

# Беспроводные линии связи и сети

Вильям Столлинс



Издательский дом "Вильямс"  
Москва • Санкт-Петербург • Киев  
2003

ББК 32.973.26-018.2.75

С81

УДК 681.3.07

Издательский дом "Вильямс"

Зав. редакцией А.В. Слепцов

Перевод с английского А.В. Высоцкого, Н.А. Голобородько, Е.Г. Грозы,  
А.В. Назаренко, К.Ю. Рулик

Под редакцией А.В. Назаренко

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

Столлингс В.

С81 Беспроводные линии связи и сети. : Пер. с англ. — М. : Издательский дом "Вильямс", 2003. — 640 с. : ил. — Парал. тит. англ.

ISBN 5-8459-0409-9 (рус.)

Беспроводная связь является весьма перспективной и быстро развивающейся областью науки и техники, исчерпывающий обзор которой представлен в данной книге. В ней дано теоретическое обоснование различных технических решений и предлагаются рекомендации по их практическому внедрению; описываются стандарты, разработанные для беспроводной передачи данных, организации беспроводных сетей и линий связи, а также других родственных областей. Автор приводит ссылки на различные дополнительные источники информации — печатные издания и Web-ресурсы. Для лучшего усвоения материала в конец каждой главы помещены задачи и вопросы для самопроверки.

Книга будет полезна специалистам в области связи, студентам, изучающим соответствующие курсы в высших учебных заведениях, а также людям, желающим самостоятельно овладеть основными концепциями организации и применения беспроводных сетей.

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 Prentice Hall, Inc.

Authorized translation from the English language edition published by Prentice Hall, Inc., Copyright © 2002

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without permission from the Publisher.

Russian language edition published by Williams Publishing House according to the Agreement with R&I Enterprises International, Copyright © 2003

ISBN 5-8459-0409-9 (рус.)  
ISBN 0-1304-0864-6 (англ.)

© Издательский дом "Вильямс", 2003  
© Prentice-Hall, Inc., 2002

# ОГЛАВЛЕНИЕ

Предисловие	17
Глава 1. Введение	21
<b>Часть 1. Предварительная информация</b>	<b>33</b>
Глава 2. Основы передачи	35
Глава 3. Сети связи	69
Глава 4. Протоколы и набор TCP/IP	99
<b>Часть 2. Технология беспроводной связи</b>	<b>129</b>
Глава 5. Антенны и распространение	131
Глава 6. Методы кодирования сигналов	167
Глава 7. Расширенный спектр	205
Глава 8. Кодирование и защита от ошибок	243
<b>Часть 3. Организация беспроводных сетей</b>	<b>293</b>
Глава 9. Спутниковая связь	295
Глава 10. Беспроводные сотовые сети	325
Глава 11. Беспроводные системы и беспроводные абонентские линии	385
Глава 12. Протоколы Mobile IP и WAP	435
<b>Часть 4. Беспроводные локальные сети</b>	<b>487</b>
Глава 13. Технология беспроводных локальных сетей	489
Глава 14. Стандарт беспроводных локальных сетей IEEE 802.11	505
Глава 15. Bluetooth	535
<b>Часть 5. Приложения</b>	<b>579</b>
Приложение А. Стандарты и организации стандартизации	581
Приложение Б. Анализ трафика	591
Приложение В. Анализ Фурье	601
Приложение Г. Протоколы управления каналом передачи данных	607
<i>Словарь терминов</i>	615
<i>Литература</i>	623
<i>Сокращения</i>	631
<i>Предметный указатель</i>	634

# СОДЕРЖАНИЕ

<b>Предисловие</b>	<b>17</b>
Задачи этой книги	17
Для кого предназначена книга	18
Internet-услуги для преподавателей и студентов	18
Благодарности	18
<b>Глава 1. Введение</b>	<b>21</b>
1.1. Беспроводные технологии достигают зрелости	22
1.2. Сотовая революция	23
1.3. Глобальная сотовая сеть	24
1.4. Широкополосные технологии	25
1.5. Проблемы беспроводной связи	25
1.6. Структура книги	26
Часть I. Предварительная информация	26
Часть II. Технология беспроводной связи	27
Часть III. Организация беспроводных сетей	28
Часть IV. Беспроводные локальные сети	29
1.7. Ресурсы Internet	29
Web-сайты для данной книги	29
Другие Web-сайты	30
Группы новостей USENET	31
<b>Часть I. Предварительная информация</b>	<b>33</b>
<b>Глава 2. Основы передачи</b>	<b>35</b>
2.1. Сигналы для передачи информации	36
Основные понятия временного представления сигнала	36
Основные понятия частотного представления сигнала	39
Связь между скоростью передачи данных и шириной полосы	41
2.2. Аналоговая и цифровая передача данных	44
Аналоговые и цифровые данные	44
Аналоговые и цифровые сигналы	45
Аналоговая и цифровая передача	48
2.3. Пропускная способность канала	49
Ширина полосы по Найквисту	50
Формула Шеннона для пропускной способности	50
2.4. Передающие среды	53
Наземная связь с использованием СВЧ	55
Спутниковая связь с использованием СВЧ	56
Широковещательное радио	58
Связь в инфракрасном диапазоне	59
2.5. Уплотнение	59

- 2.6. Рекомендуемая литература
- 2.7. Термины, вопросы и задачи
  - Основные термины
  - Вопросы
  - Задачи

Приложение 2А. Децибелы и интенсивность сигнала

### **Глава 3. Сети связи**

- 3.1. Локальные, городские и глобальные сети
  - Глобальные сети
  - Локальные сети
  - Городские сети
- 3.2. Методы коммутации
- 3.3. Коммутация каналов
- 3.4. Коммутация пакетов
  - Использование коммутации пакетов
  - Размер пакета
- 3.5. АТМ
  - Логические соединения АТМ
  - Ячейки АТМ
  - Категории услуг АТМ
- 3.6. Рекомендуемая литература и Web-сайты
- 3.7. Термины, вопросы и задачи
  - Основные термины
  - Вопросы
  - Задачи

### **Глава 4. Протоколы и набор TCP/IP**

- 4.1. Необходимость архитектуры протоколов
- 4.2. Архитектура протоколов TCP/IP
  - Уровни TCP/IP
  - Действие протоколов TCP и IP
  - Приложения TCP/IP
- 4.3. Модель OSI
  - Уровни протоколов
  - Стандартизация на основе архитектуры OSI
- 4.4. Межсетевое взаимодействие
  - Маршрутизаторы
  - Пример меж сетевого взаимодействия
- 4.5. Рекомендуемая литература
- 4.6. Термины, вопросы и задачи
  - Основные термины
  - Вопросы
  - Задачи

Приложение 4А. Протокол Internet

- IPv4
- IPv6

Приложение 4Б. Протокол управления передачей  
Управление потоком TCP

Формат сегмента TSP	124
Приложение 4В. Протокол пользовательских дейтаграмм	126
<b>Часть II. Технология беспроводной связи</b>	<b>129</b>
<b>Глава 5. Антенны и распространение</b>	<b>131</b>
5.1. Антенны	132
Диаграммы направленности	132
Типы антенн	133
Коэффициент усиления антенны	136
5.2. Режимы распространения	137
Огибание поверхности Земли	137
Отражение от верхних слоев атмосферы	141
Распространение вдоль линии прямой видимости	141
5.3. Передача сигнала в пределах линии прямой видимости	143
Затухание	144
Потери в свободном пространстве	144
Шум	147
Отношение $E_b/N_0$	149
Атмосферное поглощение	151
Многолучевое распространение	151
Преломление	152
5.4. Замирание в мобильных средах	153
Многолучевое распространение	153
Методы компенсации ошибок	157
5.5. Рекомендуемая литература	162
5.6. Термины, вопросы и задачи	162
Основные термины	162
Вопросы	162
Задачи	163
<b>Глава 6. Методы кодирования сигналов</b>	<b>167</b>
6.1. Критерии кодирования сигналов	169
6.2. Цифровые данные, аналоговые сигналы	171
Амплитудная манипуляция	173
Частотная манипуляция	173
Фазовая манипуляция	175
Производительность	179
Квадратурная амплитудная модуляция	183
6.3. Аналоговые данные, аналоговые сигналы	184
Амплитудная модуляция	184
Угловая модуляция	187
6.4. Аналоговые данные, цифровые сигналы	190
Импульсно-кодовая модуляция	191
Дельта-модуляция	194
Производительность	197
6.5. Рекомендуемая литература	199
6.6. Термины, вопросы и задачи	199
Основные термины	199

Приложение 6А. Доказательство теоремы о дискретном представлении	202
<b>Глава 7. Расширенный спектр</b>	<b>205</b>
7.1. Понятие расширенного спектра	206
7.2. Расширение спектра со скачкообразной перестройкой частоты	207
Основы использования	207
FHSS с использованием MFSK	210
Анализ производительности	212
7.3. Расширение спектра методом прямой последовательности	213
DSSS с использованием BPSK	213
Анализ производительности	216
7.4. Множественный доступ с кодовым разделением	218
Основные принципы	218
CDMA для системы DSSS	221
7.5. Создание последовательностей расширения	222
Псевдослучайные последовательности	223
Ортогональные коды	233
Множественное расширение	235
7.6. Рекомендуемая литература	235
7.7. Термины, вопросы и задачи	236
Основные термины	236
Вопросы	236
Задачи	236
<b>Глава 8. Кодирование и защита от ошибок</b>	<b>243</b>
8.1. Выявление ошибок	244
Проверка четности	245
Циклическая проверка четности с избыточностью	246
8.2. Блочные коды с коррекцией ошибок	253
Принципы блочных кодов	255
Коды Хэмминга	259
Циклические коды	262
Коды БХЧ	267
Коды Рида-Соломона	268
Чередование блоков	269
8.3. Сверточные коды	270
Декодирование	271
Турбокодирование	275
8.4. Автоматический запрос повторной передачи	277
Управление потоком данных	278
Защита от ошибок	282
8.5. Рекомендуемая литература	285
8.6. Термины, вопросы и задачи	286
Основные термины	286
Вопросы	286
Задачи	287
<b>Часть III. Организация беспроводных сетей</b>	<b>293</b>
<b>Глава 9. Спутниковая связь</b>	<b>295</b>

9.1. Параметры и конфигурации спутника	296
Спутниковые орбиты	297
Полосы частот	304
Ухудшение качества связи	305
Конфигурации спутниковой сети	309
9.2. Распределение пропускной способности — частотное разделение	310
Уплотнение с частотным разделением (FDM)	311
Множественный доступ с частотным разделением (FDMA)	312
9.3. Распределение пропускной способности — временное разделение	317
9.4. Рекомендуемая литература	323
9.5. Термины, вопросы и задачи	323
Основные термины	323
Вопросы	324
Задачи	324
<b>Глава 10. Беспроводные сотовые сети</b>	<b>325</b>
10.1. Принципы сотовой связи	326
Организация сотовой сети	326
Функционирование сотовой системы	331
Эффекты распространения радиоволн в мобильной связи	335
Переключение	336
Регулирование мощности	339
Регулирование трафика	342
10.2. Аналоговые системы первого поколения	347
Распределение спектра	347
Функционирование систем первого поколения	348
Каналы управления в системе AMPS	349
10.3. Системы TDMA второго поколения	350
Сотовые системы первого и второго поколений	350
Множественный доступ с временным разделением каналов	352
Вопросы проектирования мобильных беспроводных систем TDMA	353
Глобальная система мобильной связи	356
Архитектура сети GSM	356
Аспекты радиосвязи	359
Протокольная архитектура передачи сигналов GSM	363
10.4. Системы CDMA второго поколения	365
Множественный доступ с кодовым разделением каналов	365
Вопросы проектирования мобильных беспроводных систем CDMA	366
IS-95	368
Прямой канал системы IS-95	368
Обратный канал системы IS-95	372
10.5. Системы третьего поколения	374
Альтернативные интерфейсы	375
Вопросы проектирования систем CDMA	377
10.6. Рекомендуемая литература и Web-сайты	378
10.7. Термины, вопросы и задачи	380
Основные термины	380
Вопросы	380



## **Глава 11. Беспроводные системы и беспроводные абонентские линии**

### **11.1. Беспроводные системы**

Дуплекс с временным разделением

Формат кадра DECT

Работа схемы DECT

Адаптивная дифференциальная импульсно-кодовая модуляция

### **11.2. Беспроводные абонентские линии связи**

Роль линий WLL

Распространение сигнала в системах WLL

Ортогональное уплотнение с частотным разделением

Многоканальные многоточечные распределительные услуги (MMDS)

Локальные многоточечные распределительные услуги (LMDS)

### **11.3. Стандарт IEEE 802.16**

Архитектура IEEE 802.16

Услуги

Уровень MAC стандарта IEEE 802.16.1

Физический уровень стандарта IEEE 802.16.1

### **11.4. Рекомендуемая литература и Web-сайты**

### **11.5. Термины, вопросы и задачи**

Основные термины

Вопросы

Задачи

Приложение 11А. Линейный фильтр с предсказанием

## **Глава 12. Протоколы Mobile IP и WAP**

### **12.1. Mobile IP**

Работа Mobile IP

Обнаружение

Регистрация

Туннелирование

### **12.2. Протокол беспроводных приложений**

Обзор архитектуры

Язык разметки для беспроводных приложений

WMLScript

Среда беспроводных приложений

Беспроводный сеансовый протокол

Протокол беспроводных транзакций

Протокол безопасности беспроводного транспортного уровня

Беспроводный протокол дейтаграмм

### **12.3. Рекомендуемая литература и Web-сайты**

### **12.4. Термины, вопросы и задачи**

Основные термины

Вопросы

Задачи

Приложение 12А. Протокол управляющих сообщений Internet

Обнаружение маршрутизатора

Приложение 12Б. Аутентификация сообщений

Односторонняя хэш-функция

Ключевой хэш-код	483
Приложение 12В. Примитивы и параметры служб	484
<b>Часть IV. Беспроводные локальные сети</b>	<b>487</b>
<b>Глава 13. Технология беспроводных локальных сетей</b>	<b>489</b>
13.1. Обзор	490
Применение беспроводных локальных сетей	490
Требования к беспроводным локальным сетям	495
Технология беспроводных локальных сетей	496
13.2. Инфракрасные локальные сети	498
Плюсы и минусы	498
Методы передачи	498
13.3. Сети с расширенным спектром	501
Конфигурация	501
Вопросы передачи	501
13.4. Сети с узкополосной СВЧ-передачей	502
Лицензируемые узкополосные системы	502
Нелицензируемые узкополосные системы	502
13.5. Рекомендуемая литература и Web-сайты	503
13.6. Термины и вопросы	503
Основные термины	503
Вопросы	504
<b>Глава 14. Стандарт беспроводных локальных сетей IEEE 802.11</b>	<b>505</b>
14.1. Архитектура IEEE 802	506
Архитектура протоколов	506
Формат кадра MAC	508
Управление логическим каналом	509
14.2. Архитектура и услуги IEEE 802.11	512
Архитектура IEEE 802.11	513
Услуги IEEE 802.11	514
14.3. Управление доступом к среде в IEEE 802.11	517
Надежная доставка данных	517
Управление доступом	518
Кадр MAC	523
Вопросы безопасности	527
14.4. Физический уровень IEEE 802.11	529
Исходный физический уровень IEEE 802.11	529
IEEE 802.11a	532
IEEE 802.11b	532
14.5. Рекомендуемая литература и web-сайты	533
14.6. Термины и вопросы	533
Основные термины	533
Вопросы	534
<b>Глава 15. Bluetooth</b>	<b>535</b>
15.1. Обзор	536
Области применения Bluetooth	536
Стандарты Bluetooth	538

Архитектура протоколов	540
Модели использования	542
Пикосети и рассеянные сети	542
15.2. Радиоспецификация	545
15.3. Узкополосная спецификация	546
Перестройка частоты	546
Физические каналы	548
Пакеты	549
Исправление ошибок	554
Логические каналы	557
Управление каналом	558
Аудио	561
Безопасность	564
15.4. Спецификация администратора связей	566
15.5. Протокол управления логическим каналом и адаптации (L2CAP)	570
Каналы L2CAP	571
Пакеты L2CAP	571
Сигнальные команды	573
Качество обслуживания	575
15.6. Рекомендуемая литература и Web-сайты	577
15.7. Термины и вопросы	578
Основные термины	578
Вопросы	578

## **Часть V. Приложения** **579**

<b>Приложение А. Стандарты и организации стандартизации</b>	<b>581</b>
А.1. Важность стандартов	582
А.2. Стандарты и регулирование	583
А.3. Международный телекоммуникационный союз	583
Сектор радиосвязи	584
Сектор стандартизации телекоммуникаций	584
Схема работы	585
А.4. Стандарты Internet и общество Internet	585
Организации Internet и документы RFC	586
Процесс стандартизации	586
Документы, не проходящие путем стандарта	588
А.5. Стандарты IEEE 802	588
Модель реализации локальных сетей	588
Структура IEEE 802	589
<b>Приложение Б. Анализ трафика</b>	<b>591</b>
Б.1. Основные понятия	592
Б.2. Многосерверные модели	593
Бесконечное число источников, очистка неудачных вызовов	595
Конечное число источников, очистка неудачных вызовов	596
Задержка неудачных вызовов	597
Резюме	599
Б.3. Рекомендуемая литература	599

<b>Приложение В. Анализ Фурье</b>	<b>601</b>
В.1. Представление периодического сигнала в виде ряда Фурье	602
В.2. Разложение в ряд Фурье аperiodического сигнала	604
Спектральная плотность мощности и ширина полосы	605
В.3. Рекомендуемая литература	606
<b>Приложение Г. Протоколы управления каналом передачи данных</b>	<b>607</b>
Г.1. HDLC	608
Структура кадра	608
Функционирование	610
<b>Словарь терминов</b>	<b>615</b>
<b>Литература</b>	<b>623</b>
<b>Сокращения</b>	<b>631</b>
<b>Предметный указатель</b>	<b>634</b>

# ПРЕДИСЛОВИЕ

## Задачи этой книги

Сегодня беспроводные технологии стали наиболее динамичной областью в сфере связи и организации сетей. Быстрое распространение мобильных телефонов, различных спутниковых служб, а в последнее время еще и беспроводной Internet приводит к значительным изменениям в характере коммуникаций и сетей. В данной книге исследуются важнейшие проблемы этой области, сосредоточенные в следующих основных категориях.

- **Технология и архитектура.** Охарактеризовать методы беспроводной связи и организации сетей, а также определить их различия помогает небольшая группа таких компонентов, как диапазон частот, схема кодирования сигналов, схема исправления ошибок и архитектура сети.
- **Типы сетей.** Данная книга посвящена изучению важнейших типов беспроводных сетей, включая спутниковые, сотовые, сети стационарного радиодоступа и беспроводные локальные сети.
- **Конструктивные подходы.** В книге исследуются альтернативные решения и оцениваются их относительные преимущества.
- **Приложения.** На основе беспроводных инфраструктур, особенно протокола Mobile IP и беспроводного Web-доступа, разработано множество важных технологий и приложений.

В книгу включены обширный словарь специальных терминов, перечень часто используемых сокращений и библиография. В каждой главе содержатся задачи, рекомендации по дополнительным источникам информации и список Web-сайтов по данной теме. В каждой главе имеется также список ключевых слов и некоторое число вопросов для закрепления материала.

На протяжении всей книги особое внимание уделяется технологиям и стандартам. Книга представляет собой всестороннее руководство по изучению определенных стандартов беспроводных коммуникаций, например стандартов, публикуемых Международным телекоммуникационным союзом и группой IEEE 802, а также разработанных другими организациями. Выбор вопросов, рассматриваемых в книге, обусловлен значимостью таких стандартов в определении уже существующих продуктов и направлений будущих исследований в данной области.

## **Для кого предназначена книга**

Настоящая книга предназначена для широкой аудитории читателей, которым будет полезно изучение беспроводных методов связи и организации сетей, а также сопутствующих технологий. Она заинтересует студентов и специалистов в области обработки и передачи данных, проектировщиков и конструкторов, а также пользователей и операторов систем передачи данных и сетей. Книга создавалась как самостоятельное руководство. Для читателей, имеющих небольшую подготовку в области передачи данных или не имеющих ее вообще, в части I и приложениях изложены основные базовые темы.

## **Internet-услуги для преподавателей и студентов**

У данной книги существует Web-сайт поддержки преподавателей и студентов. На этом сайте содержатся ссылки на другие сайты родственной тематики, оригиналы диапозитивов рисунков и таблиц книги в формате PDF (Adobe Acrobat) и сведения о подписке на список рассылки данной книги. Web-страница поддержки находится по адресу: [WilliamStallings.com/Wireless1e.html](http://WilliamStallings.com/Wireless1e.html) (более подробная информация содержится в разделе 1.7). Существует также список рассылки, благодаря которому преподаватели, использующие эту книгу, могут обмениваться информацией, вопросами и предложениями между собой и с автором. Опечатки или другие ошибки, обнаруживаемые в книге, заносятся в список по адресу: [WilliamStallings.com](http://WilliamStallings.com).

## **Благодарности**

Множество людей великодушно предоставили нам свое время и опыт, рецензируя эту книгу, отчего она только выиграла. Сюда следует отнести тех, кто полностью или частично рецензировал рукопись и план книги. Это Марио Джерла (Mario Gerla) из Калифорнийского университета в Лос-Анджелесе, Джерри Плейс (Jerry Place) из университета Миссури в Канзас-Сити, Джон Метцнер (John Metzner) из университета штата Пенсильвания, Упкар Варшней (Upkar Varshney) из университета штата Джорджия, Питер Ра (Peter Rha) из университета Сан-Франциско и Артур Вербнер (Arthur Werbner).

Я благодарю также многих рецензентов, составивших подробные технические обзоры отдельных глав. Это Ларс Полсен (Lars Poulsen), Говард Эйзенхауэр (Howard Eisenhauer), Д. И. Дженнингс (D. E. Jennings), Поль Робишо (Paul Robichaux), Джон Адамс (John Adams), Джерри Хуан (Jerry Huang), Андреас Касенидес (Andreas Kasenides), Мунира Ахмед (Munira Ahmed), Хосейн Изадпана (Hossein Izadpanah), Аарон Кинг (Aaron King), Бенуа д'Юдекем (Benoit d'Udekem), Марко Казоле (Marco Casole), Кевин Петерсон (Kevin Peterson), Динеш Лал Прадхан (Dinesh Lal Pradhan) и Катал Мак-Дейд (Cathal Mc Daid).

Стивен Килби (Steven Kilby) помог в написании главы 1 и отредактировал много других глав. Я высоко ценю его вклад. Хотелось бы поблагодарить также Ричарда Ван Слика (Richard Van Slyke) из Политехнического университета Бруклина, который предоставил многие вопросы для самопроверки. Том Фронковьяк (Tom Fronckowiak) создал слайды PowerPoint для книги.

Наконец, я хотел бы поблагодарить многих людей, отвечавших за выход книги, которые сделали свою работу, как всегда, превосходно. Это сотрудники издательства Prentice Hall, в особенности мои редакторы Тони Холм (Toni Holm) и Алан Апт (Alan Apt) и руководитель производства Роз Кернан (Rose Kernan). Кроме них, доработкой и редактированием занимался Джейк Уорд (Jake Warde) из издательства Warde Publishers, руководила печатью книги Джоанна В. Померанц (Joanna V. Pomeranz), а литературное редактирование выполняла Патрисия П. Дейли (Patricia P. Daly).

# ГЛАВА 1

## ВВЕДЕНИЕ <sup>1</sup>

**1.1. Беспроводные технологии достигают зрелости**

**1.2. Сотовая революция**

**1.3. Глобальная сотовая сеть**

**1.4. Широкополосные технологии**

**1.5. Проблемы беспроводной связи**

**1.6. Структура книги**

**Часть I. Предварительная информация**

**Часть II. Технология беспроводной связи**

**Часть III. Организация беспроводных сетей**

**Часть IV. Беспроводные локальные сети**

**1.7. Ресурсы Internet**

**Web-сайты для данной книги**

**Другие Web-сайты**

**Группы новостей USENET**

---

<sup>1</sup> Эта глава, за исключением раздела 1.7, написана Стивеном Килби (Steven Kilby).



**Э**та книга представляет собой обзор беспроводных коммуникаций и с. Многие факторы, включая повышенную конкуренцию и внедрение цн вых технологий, привели к беспрецедентному развитию рынка беспровод средств. В данной главе мы рассмотрим некоторые наиболее важные фактс ставшие движущей силой этой революции в сфере коммуникаций.

Настоящая книга и Web-сайт ее поддержки содержат огромное количе материала. В этой главе представлен общий ход изложения, благодаря чему татель может охватить книгу одним взглядом.

## **1.1. БЕСПРОВОДНЫЕ ТЕХНОЛОГИИ ДОСТИГАЮТ ЗРЕЛОСТИ**

Гильельмо Маркони изобрел беспроводный телеграф в 1896 году.<sup>2</sup> В 1901 г он послал телеграфный сигнал через Атлантический океан от Корнуолла Сент-Джонса на Ньюфаундленде на расстоянии 1800 миль. Это изобретение с лало возможным связью двух сторон, посылающих друг другу буквы и цифр закодированные в аналоговом сигнале. В последующие сто лет развитие техно гий беспроводной связи привело к появлению радиовещания, телевидения, бильного телефона и спутников связи. Теперь можно послать информацию лю го типа почти в каждый уголок мира. В последнее время наибольшее внимаи к себе привлекают спутниковая связь, беспроводные сети и сотовая технологи

Запуск спутников связи начался в 1960-х годах. Первые спутники мои обслуживать лишь 240 телефонных каналов. В наше время с помощью спутн ков осуществляется приблизительно третья часть телефонного трафика и все лепередачи между странами [EVAN98]. Современные каналы спутниковой свя вносят в обрабатываемые ими сигналы задержку, обычно не превышающую ч верти секунды. Вскоре появятся более новые спутники, на более низких орб тах, имеющие еще меньшую задержку сигнала, которые предоставят такие ус ги передачи данных, как доступ в Internet.

Средства беспроводных сетей позволяют разворачивать глобальные, реги нальные и локальные сети без необходимости создания кабельной структур Институтом IEEE разработан стандарт 802.11 для беспроводных локальных с тей. Кроме него, над созданием интегральных технологий беспроводных сет работает промышленный консорциум Bluetooth.

Современным эквивалентом беспроводного телеграфа Маркони, обеспеч вающим двустороннюю двунаправленную связь, является сотовый или мобил ный телефон. В радиотелефонах первого поколения использовалась аналоговы технология. Эти устройства были громоздки и имели ограниченную зону обл живания, но на их примере ясно были видны преимущества мобильной систем связи. Теперешнее поколение беспроводных устройств построено уже на ба цифровой, а не аналоговой технологии. Цифровые сети могут передавать намн го больший телефонный трафик, и обеспечиваемые ими качество приема и безп асность выше, чем в аналоговых системах. Кроме того, цифровая технологи

---

<sup>2</sup> Собственно изобретение радиосвязи следует приписать Николе Тесле, который еп в 1893 г. произвел публичную демонстрацию своего устройства. В 1943 г. патенты Ма ркони были аннулированы в пользу Теслы [ENGE00].

позволяет вводить дополнительные услуги, такие, как идентификация вызывающего абонента. Беспроводные устройства следующего поколения также будут цифровыми, с подключением к Internet и использованием новых частотных диапазонов и более высоких скоростей передачи информации.

Значение беспроводной связи настолько велико, что ей принадлежит будущее. Очень немногим изобретениям удавалось так сократить расстояния. Стандарты, определяющие характер взаимодействия устройств беспроводной связи, быстро унифицируются и скоро позволят создать глобальную беспроводную сеть, предоставляющую все многообразие услуг.

## 2. СОТОВАЯ РЕВОЛЮЦИЯ

Одно лишь увеличение рынка мобильных телефонов наглядно свидетельствует о том, что в мире технологий произошла революция. В 1990 году количество пользователей сотовых телефонов составляло приблизительно 11 миллионов. В 2004 году оно, вероятно, возрастет до одного миллиарда [ECON99]. Этому росту обязательно будут способствовать устройства следующего поколения, с доступом в Internet. По одному из прогнозов к 2005 году число устройств беспроводного доступа в Internet превысит число устройств с доступом по проводным линиям [ECON99].

В настоящее время наиболее очевидный признак успеха беспроводных технологий — распространение телефонов. Начиная с 1996 года количество новых абонентов мобильных телефонов превышает количество новых абонентов стационарных телефонов [ECON99]. Так получилось по многим причинам. Мобильные телефоны удобны; они перемещаются вместе с людьми. Кроме того, по самой своей природе средства мобильной связи “знают” свое местоположение. Каждый мобильный телефон связывается с региональными базовыми станциями, расположение которых твердо установлено.

В успех мобильных телефонов внесли свой вклад различные технические новшества. Телефонные трубки стали меньше и легче, срок службы батарей увеличился, а цифровая технология повысила качество приема и позволила лучше использовать доступный спектр. Как и во многих других случаях, переход к цифровому оборудованию позволил снизить затраты на мобильную связь. В тех областях, где процветает конкуренция, цены начиная с 1996 года значительно снизились.

Во многих регионах распространение мобильных телефонов — единственный экономичный способ предоставить населению телефонную связь. Базовые станции можно создавать быстро и недорого по сравнению с рытьем траншей и прокладкой кабелей, особенно в сложных условиях местности.

Мобильные телефоны — лишь начало сотовой революции. Появляется все больше и больше новых типов беспроводных устройств, которые имеют доступ в Internet. К ним относятся персональные органайзеры и те же телефоны, но теперь уже с Web-доступом, функциями мгновенных сообщений, электронной почты и другими услугами, доступными в Internet. Беспроводные устройства в автомобилях позволяют загружать карты и маршруты. Скоро такие устройства смогут вызывать помощь, если произошла авария, или даже сообщать пользователю о самых низких ценах на бензин в ближайшей окрестности. Будут доступны и другие удобные функции. Например, восполнение через Internet запасов домашних холодильников.

Первым заданием для мобильных устройств была передача речи. Теперь цент переместился на передачу данных. Предполагается, что в следующие 10 лет сложится многомиллионный рынок услуг беспроводной передачи данных [AGRA99]. Большая часть этого рынка — “беспроводная” Internet. Беспроводные устройства потребуют использовать Internet иначе, чем стационарные. Беспроводные устройства имеют меньше возможностей отображения и ввода данных по сравнению с обычными стационарными устройствами, такими, как персональные компьютеры. Нормой станут не длинные сеансы просмотра Web-страниц, транзакции и обмен сообщениями. Поскольку беспроводные устройства “знают” свое расположение, можно будет адаптировать информацию к географическому местоположению пользователя. Таким образом, уже не пользователи будут искать информацию, а информация сможет сама находить пользователей.

### 1.3. ГЛОБАЛЬНАЯ СОТОВАЯ СЕТЬ

Сейчас не существует какой-то единой сотовой сети. Обычно существующие устройства поддерживают одну или две из несметного числа технологий и работают только в пределах зоны обслуживания одного оператора. Чтобы выйти за пределы этой модели, надо проделать много работы, особенно в определении и реализации стандартов.

Комплекс стандартов для беспроводных устройств следующего поколения разрабатывает Международный телекоммуникационный союз (International Telecommunication Union — ITU). В новых стандартах будут использоваться более высокие частоты, что позволит увеличить пропускную способность. Новые стандарты также помогут преодолеть несовместимость различных сетей первого и второго поколений, разработанных и развернутых за последние десять лет.

Преобладающей цифровой беспроводной сетью первого поколения в Северной Америке была система AMPS (Advanced Mobile Phone System — усовершенствованная система мобильной телефонной связи). Эта сеть предоставляла услуги передачи данных на базе перекрывающихся сетей CDPD (Cellular Digital Packet Data — сотовая система передачи пакетов цифровых данных) со скоростью 19,2 Кбит/с. В системе CDPD передача данных производилась в незанятые интервалы в обычных каналах телефонной связи.

Важнейшие беспроводные системы второго поколения — GSM (Global System for Mobile Communications — глобальная система мобильной связи), PCS (Personal Communications Service — персональная служба связи) IS-136 и PCS IS-95. В стандарте PCS IS-136 используется множественный доступ с временным разделением, а IS-95 — с кодовым разделением. В системах GSM и PCS IS-136 передача данных производится через выделенные каналы со скоростью 9,6 Кбит/с.

Сейчас Международный телекоммуникационный союз разрабатывает стандарт IMT-2000 (International Mobile Telecommunications — международные мобильные телекоммуникации). Эта группа стандартов должна определить глобальную интегрированную сеть. В стандартах используется диапазон 2 ГГц. Новые стандарты и новый диапазон позволят достичь скорости передачи данных 2 Мбит/с.

Кроме определения использования частот, схем кодирования и передачи, в стандартах также должно быть указано, как мобильные устройства будут взаимодействовать с Internet. В этой области работают несколько специализированных организаций по стандартам и промышленных консорциумов. В частности

Форум WAP (Wireless Application Protocol — протокол беспроводных приложений) разрабатывает общий протокол для обращения к Internet устройств с ограниченными возможностями отображения и ввода. Проблемная группа проектирования Internet (IETF) разрабатывает стандарт протокола Mobile IP, который представляет собой адаптацию общепринятого протокола IP к работе в мобильном окружении.

## 1.4. ШИРОКОПОЛОСНЫЕ ТЕХНОЛОГИИ

Все более важное место в Internet занимает мультимедиа. Графика, видео и звук в изобилии встречаются на Web-страницах. Та же тенденция наблюдается и в коммерческой связи. Например, сообщения электронной почты часто содержат приложения мультимедиа. Чтобы полностью включиться в эту среду, беспроводные сети должны обеспечивать такие же высокие скорости передачи данных, как и их стационарные аналоги. Этим скоростям позволяет достичь широкополосная технология.

Широкополосные беспроводные службы имеют те же преимущества, что и все беспроводные средства: удобство и сниженные затраты. За счет отказа от кабелей операторы могут развертывать беспроводные службы быстрее и дешевле, чем проводные. Кроме того, беспроводные службы мобильны и могут разворачиваться почти в любом месте.

Существует много проектов разработки стандартов широкополосной беспроводной связи для множества различных приложений. Эти стандарты охватывают все области — от беспроводных локальных сетей до небольших домашних сетей. Скорости передачи данных варьируются от 2 Мбит/с до более чем 100 Мбит/с. Уже доступны многие из этих технологий и число их значительно увеличится в следующие несколько лет.

Беспроводные локальные сети предоставляют возможность работы в сети там, где бывает трудно или слишком дорого развернуть стационарную инфраструктуру. Основные стандарты беспроводных локальных сетей — IEEE 802.11b в Америке и HiperLAN в Европе. Стандарт IEEE предусматривает скорость передачи данных до 11 Мбит/с. Европейский стандарт определяет максимальную скорость 24 Мбит/с, а в будущей редакции скорость передачи будет повышена до 54 Мбит/с.

Потенциальная проблема стандарта 802.11b — совместимость с Bluetooth. Спецификация беспроводных сетей Bluetooth определяет связь между такими устройствами, как ноутбуки, карманные компьютеры и мобильные телефоны. В Bluetooth и 802.11b используется один и тот же диапазон частот. При реализации в одном устройстве эти технологии, скорее всего, будут мешать друг другу.

В рамках проекта HomeRF разрабатываются стандарты беспроводной связи между домашними компьютерами и внешними устройствами. В настоящее время в HomeRF применяются скорости передачи до 2 Мбит/с, но новая редакция будет рассчитана на скорости порядка 10 Мбит/с.

## 1.5. ПРОБЛЕМЫ БЕСПРОВОДНОЙ СВЯЗИ

Беспроводные службы удобны и часто более дешевы в развертывании, чем стационарные, но они несовершенны. Существуют ограничения, трудности полити-

ческого и технического характера, которые в конечном счете могут помешать беспроводным технологиям полностью реализовать свой потенциал. Две основные проблемы — несовместимость стандартов и ограничения устройств.

Как было сказано выше, в Северной Америке существует два стандарта цифровых сотовых систем. В остальном мире существует по крайней мере один стандарт. Устройство, в котором используется PCS IS-136, не будет работать в районе, где развернутая система связи построена на стандарте PCS IS-54. Мы говорили также о невозможности применять Bluetooth и 802.11b в одном же устройстве. Это лишь два примера проблем, возникающих в ситуации, когда нет единого стандарта для всей отрасли. Отсутствие такого стандарта мешает технологиям достичь одного из настоящих идеалов беспроводной связи — повсеместного доступа к данным.

Ограничения, присущие устройствам, также сдерживают свободное движение данных. Маленький жидкокристаллический дисплей мобильного телефона не способен отобразить больше нескольких строк текста. Кроме того, большинство мобильных беспроводных устройств может обращаться лишь к небольшой части огромного множества Web-сайтов в Internet. В браузерах этих устройств используется специальный язык разметки для беспроводных приложений (Wireless Markup Language — WML), а не стандартный язык HTML.

Вероятнее всего, невозможно создать такое беспроводное устройство, которое могло бы удовлетворить любую потребность. Потенциал беспроводной связи можно реализовать, но не в каком-то одном изделии. Беспроводная связь будет иметь успех при интегрировании в различные устройства, предназначенные для удовлетворения разнообразных потребностей.

## **1.6. СТРУКТУРА КНИГИ**

Цель этого издания — дать всесторонний технический обзор основных принципов беспроводной связи, беспроводных сетей и беспроводных приложений. Книга разделена на четыре части. Читатель, уже знакомый с технологией передачи данных и организацией сетей, вполне может пропустить часть I или прочесть ее бегло. В части II излагаются базовые принципы, применимые ко всему материалу, содержащемуся в остальных частях; ее нужно читать следующей. Части III и IV не связаны друг с другом и могут быть прочтены в любом порядке. В части III все главы более или менее независимы, и читать их можно в любом порядке, с учетом вашей заинтересованности. Это относится и к главам 14 и 15 части V.

### **Часть I. Предварительная информация**

В этой части дан предварительный обзор содержания книги. Здесь рассматриваются основные вопросы передачи данных, а также набор протоколов TCP/IP. Содержание части I и приложений, находящихся в конце книги, позволяет сделать настоящее издание, насколько это возможно, самостоятельным руководством.

#### **Глава 2. Основы передачи**

В главе 2 дан краткий обзор вопросов передачи. Глава начинается с изложения некоторых концепций передачи данных, включая методы передачи с

налов, а также аналоговую и цифровую передачу. Затем рассматриваются пропускная способность канала, среды передачи и концепция уплотнения.

### **Глава 3. Сети связи**

В этой главе сравниваются основные технологии сетей связи, включая коммутацию каналов, коммутацию пакетов и систему АТМ.

### **Глава 4. Протоколы и набор TCP/IP**

Передача данных в сетях и распределенных приложениях базируется на низкоуровневых программах, которые не зависят от приложений и избавляют их от многих операций по обеспечению надежного обмена данными. Эти программы связи составляют архитектуру протоколов, наиболее важное воплощение которой — набор протоколов TCP/IP. В главе 4 представлена концепция архитектуры протоколов и дан краткий обзор TCP/IP. Кратко описывается также другая архитектура — эталонная модель OSI. В заключение рассматриваются концепции межсетевое взаимодействие и применение для этой цели TCP/IP.

## **Часть II. Технология беспроводной связи**

В этой части рассматриваются базовые технологии беспроводной передачи и кодирования аналоговых и цифровых данных для беспроводной передачи.

### **Глава 5. Антенны и распространение**

В главе 5 изложены фундаментальные принципы распространения радиоволн и сигналов СВЧ. Рассматриваются важнейшие аспекты работы антенн, затем изучаются режимы беспроводной передачи и в заключение исследуется важный вопрос замирания сигналов.

### **Глава 6. Методы кодирования сигналов**

Данные могут поступать как в аналоговой (непрерывной), так и в цифровой (дискретной) форме. Для передачи входные данные должны быть закодированы в виде электрического сигнала, приспособленного к характеристикам передающей среды. Как аналоговые, так и цифровые данные могут быть представлены либо аналоговыми, либо цифровыми сигналами; соответствующие ситуации при беспроводной передаче рассматриваются в главе 6.

### **Глава 7. Расширенный спектр**

Все большее распространение получает такая форма беспроводной коммуникации, как расширение спектра. Используются два основных метода: скачкообразная перестройка частоты и схема прямой последовательности. В главе 7 представлен обзор обоих методов. Здесь также рассматривается концепция множественного доступа с кодовым разделением каналов (CDMA), которое вместе с методами расширения спектра применяется для обеспечения множественного доступа.

### **Глава 8. Кодирование и защита от ошибок**

Системы беспроводной связи часто подвержены ошибкам, и практически во всех схемах беспроводной передачи имеются методы прямого исправления ошибок, состоящие в том, что к передаваемым данным добавляется избыточная ин-

формация для исправления битовых ошибок в приемнике. Эти методы подробно исследуются в главе 8. Кроме того, здесь рассматриваются способы использования избыточной информации для обнаружения ошибок, эти способы также применяются во многих беспроводных схемах. Наконец, методы обнаружения ошибок часто используются совместно с методами автоматического запроса повторной передачи (ARQ), благодаря чему передатчик может повторно передавать блоки данных, в которых приемник обнаружил ошибку.

## **Часть III. Организация беспроводных сетей**

В этой части рассматриваются основные типы беспроводных сетей. К ним относятся спутниковые сети, сотовые сети связи, беспроводные системы, сети стационарного радиодоступа. Рассматривается также применение протоколов Mobile IP и WAP для доступа к Internet и Web.

### **Глава 9. Спутниковая связь**

В этой главе излагаются основные принципы спутниковой связи. Рассматриваются геостационарные спутники (GEO), низкоорбитальные спутники (LEO) и среднеорбитальные спутники (MEO). Подробно исследуется ключевой вопрос распределения пропускной способности.

### **Глава 10. Беспроводные сотовые сети**

Глава 10 начинается с рассмотрения важнейших вопросов проектирования сотовых беспроводных сетей. Затем речь идет о традиционной системе мобильной телефонной связи, теперь называемой аналоговой системой первого поколения. После этого изучаются цифровые сотовые системы второго поколения и два применяемых в них основных подхода: множественный доступ с временным разделением (TDMA) и множественный доступ с кодовым разделением каналов (CDMA). В заключение предлагается обзор сетей третьего поколения.

### **Глава 11. Беспроводные системы и беспроводные абонентские линии**

В главе 11 рассматриваются две технологии, обеспечивающие беспроводный доступ в доме и офисе: беспроводные системы и беспроводные абонентские линии связи (WLL). Беспроводные системы произошли от простых беспроводных телефонов для одного пользователя, работающих в пределах дома, усовершенствованных для обслуживания большого числа пользователей на гораздо больших расстояниях. WLL (иногда также называемая радиошлейфом — RITL, или системой стационарного радиодоступа — FWA) — это система, в которой все соединения (или часть соединений) абонентов с общедоступной коммутируемой телефонной сетью (от абонента до коммутатора) происходят не по медным проводникам, а через радиоканал. В главе 11 излагаются проблемы проектирования WLL, а затем речь идет о стандарте IEEE 802.16.

### **Глава 12. Протоколы Mobile IP и WAP**

В главе 12 рассматриваются модификации протокола IP, позволяющие подключаться к Internet по беспроводному каналу. Затем исследуется протокол WAP (Wireless Application Protocol — протокол беспроводных приложений). Этот прото-

кол предоставляет пользователям мобильных телефонов и других беспроводных терминалов, таких, как пейджеры и карманные компьютеры (PDA), доступ к телефонной связи и информационным службам, включая Internet и Web.

## **Часть IV. Беспроводные локальные сети**

За последние годы появился целый класс локальных сетей, являющийся альтернативой обычных локальных сетей на базе витых пар, коаксиального кабеля и оптоволокна, — беспроводные локальные сети. Преимущество таких сетей состоит в том, что они позволяют избежать затрат на проводники — часто наиболее дорогостоящий компонент локальной сети — и подключать мобильные рабочие станции. В данной части исследуются базовые технологии беспроводных локальных сетей, а затем рассматриваются два стандартных подхода к построению таких сетей.

### **Глава 13. Технология беспроводных локальных сетей**

В беспроводных локальных сетях используется один из трех методов передачи: расширение спектра, узкополосная СВЧ-передача и передача в инфракрасном диапазоне. В главе 13 представлен обзор технологий беспроводных локальных сетей и их приложений.

### **Глава 14. Стандарт беспроводных локальных сетей IEEE 802.11**

Наиболее значительная группа стандартов, определяющих беспроводные локальные сети, установлена комитетом IEEE 802.11. В главе 14 подробно анализируются эти стандарты.

### **Глава 15. Bluetooth**

Bluetooth — это открытая спецификация беспроводной связи и организации сетей персональных компьютеров, мобильных телефонов и других беспроводных устройств. Она является одной из наиболее быстро развивающихся технологий и предназначена для использования в ограниченной области. Данная спецификация подробно рассматривается в главе 15.

## **1.7. РЕСУРСЫ INTERNET**

В Internet и Web доступно большое количество ресурсов, дополняющих эту книгу и предоставляющих информацию о развитии данной области.

### **Web-сайты для данной книги**

По адресу: [WilliamStallings.com/Wireless1e.html](http://WilliamStallings.com/Wireless1e.html) находится специальный Web-сайт данной книги. Он включает следующие разделы.

- **Useful Web sites** (Полезные Web-сайты). Здесь собраны ссылки на другие Web-сайты по данному предмету, включая сайты, перечисленные в данном приложении и на протяжении всей книги.
- **Errata sheet** (Список опечаток). Здесь помещается список типографских и других ошибок, замеченных в данной книге. По мере необходимости этот файл будет обновляться. Если вы обнаружите ошибки, будьте добры сооб-



щить мне о них по электронной почте. На том же Web-сайте находятся списки опечаток, замеченных в других моих книгах.

- **Figures (Рисунки).** Все рисунки этой книги в формате PDF (Adobe Acrobat).
- **Tables (Таблицы).** Все таблицы этой книги в формате PDF.
- **Slides (Слайды).** Набор слайдов PowerPoint, организованный по главам.
- **Internet mailing list (Список рассылки).** Здесь представлены сведения, необходимые для подписки на список рассылки по книге.
- **Wireless courses (Курсы по беспроводным коммуникациям).** Раздел содержит ссылки на Web-страницы курсов, которые читаются с использованием данной книги. На этих страницах преподаватели могут почерпнуть полезные идеи по планированию таких курсов.

Я также веду Web-сайт ресурсов для студентов компьютерных специальностей по адресу: [WilliamStallings.com/StudentSupport.html](http://WilliamStallings.com/StudentSupport.html). Этот сайт содержит документацию, сведения и ссылки для студентов и специалистов. Ссылки сгруппированы в четыре следующие категории.

- **Math (Математика).** Включает обзор основных разделов математики; основы анализа системы массового обслуживания; основы систем счисления и многочисленные ссылки на другие сайты по математике.
- **How-to (Практические рекомендации).** Советы и руководство по выполнению домашних заданий, составлению технических сообщений и подготовке докладов.
- **Research resources (Научные ресурсы).** Ссылки на полезные собрания документации, технических отчетов и библиографии.
- **Miscellaneous (Разное).** Разнообразные полезные документы и ссылки.

## Другие Web-сайты

Существует множество Web-сайтов, предоставляющих информацию, связанную с темами, которые рассматриваются в книге. В последующих главах в разделах “Рекомендуемая литература и Web-сайты” можно найти ссылки на определенные Web-сайты<sup>3</sup>.



Ниже перечислены представляющие интерес Web-сайты по общим вопросам беспроводных коммуникаций.

- **Telecommunications Technology Resources Page** ([www.webexpert.net/vasilios/telecom/telecom.htm](http://www.webexpert.net/vasilios/telecom/telecom.htm)). Превосходный источник ссылок по многочисленным темам.

---

<sup>3</sup> В книге указаны адреса URL Web-сайтов по состоянию на конец 2002 г. Поскольку эти адреса имеют тенденцию к частому изменению, рекомендуется обращаться за более новыми сведениями на сайт данной книги.

- **Vendors** ([guide.sbanetweb.com/](http://guide.sbanetweb.com/)). Ссылки на Web-сайты тысяч производителей программного и аппаратного обеспечения, а также список нескольких тысяч компьютерных и сетевых компаний в разделе “Phone Directory”.
- **Mobile and Wireless Computing Index** ([www.wireless.com/](http://www.wireless.com/)). Источник информации о беспроводных технологиях, изделиях, конференциях и публикациях.
- **Wireless Developer Network** ([wirelessdevnet.com/](http://wirelessdevnet.com/)). Новости, обучающие программы и обсуждение вопросов беспроводной связи.
- **Wireless Technology** ([www.palowireless.com/](http://www.palowireless.com/)). Поразительно объемный список ссылок по всем аспектам беспроводных коммуникаций, сетей и стандартов.

## Группы новостей USENET

Существует большое число групп новостей USENET, посвященных некоторым аспектам передачи данных и работы в сети. Как практически и во всех группах новостей, уровень “шума” в них значительно превышает уровень “полезного сигнала”, но все же стоит попробовать поискать интересующую вас информацию. Наиболее полезными можно считать следующие группы.

- `comp.std.wireless`. Обсуждение стандартов беспроводных глобальных и локальных сетей. Благодаря модераторам дискуссии в группе носят строго специальный характер.
- `comp.dcom.*`. Существует множество групп новостей по вопросам передачи данных, названия которых начинаются с “`comp.dcom`”.

**Часть I**

# **ПРЕДВАРИТЕЛЬНАЯ ИНФОРМАЦИЯ**

**Глава 2. Основы передачи**

**Глава 3. Сети связи**

**Глава 4. Протоколы и набор TCP/IP**

# ГЛАВА 2

## ОСНОВЫ ПЕРЕДАЧИ

### **2.1. Сигналы для передачи информации**

**Основные понятия временного представления сигнала**

**Основные понятия частотного представления сигнала**

**Связь между скоростью передачи данных и шириной полосы**

### **2.2. Аналоговая и цифровая передача данных**

**Аналоговые и цифровые данные**

**Аналоговые и цифровые сигналы**

**Аналоговая и цифровая передача**

### **2.3. Пропускная способность канала**

**Ширина полосы по Найквисту**

**Формула Шеннона для пропускной способности**

### **2.4. Передающие среды**

**Наземная связь с использованием СВЧ**

**Спутниковая связь с использованием СВЧ**

**Широковещательное радио**

**Связь в инфракрасном диапазоне**

### **2.5. Уплотнение**

### **2.6. Рекомендуемая литература**

### **2.7. Термины, вопросы и задачи**

**Приложение 2А. Децибелы и интенсивность сигнала**

**Ц**ель этой главы — сделать данную книгу исчерпывающим руководством для читателя, имеющего малую подготовку в области передачи данных или вообще не имеющего ее. Читатели, заинтересованные в освоении дополнительного материала, могут найти в конце главы ссылки на литературу.

## 2.1. СИГНАЛЫ ДЛЯ ПЕРЕДАЧИ ИНФОРМАЦИИ

В данной книге мы рассматриваем электромагнитные сигналы, используемые как средство передачи информации. Электромагнитный сигнал является функцией времени, хотя он также может быть выражен и как функция частоты; т.е. сигнал состоит из компонентов, имеющих различные частоты. Это приводит нас к мысли о том, что для понимания передачи данных рассмотрение сигнала с точки зрения *частотного представления* значительно важнее, чем с точки зрения *временного*. Ниже описываются оба упомянутых подхода.

### Основные понятия временного представления сигнала

Если рассматривать сигнал как функцию времени, то он может быть либо аналоговым, либо цифровым. **Аналоговым** называется сигнал, интенсивность которого во времени изменяется постепенно. Другими словами, в сигнале не имеется пауз или разрывов<sup>1</sup>. **Цифровым** называется сигнал, интенсивность которого в течение некоторого периода поддерживается на постоянном уровне, а затем изменяется также на постоянную величину. На рис. 2.1 приведены примеры сигналов обоих типов. Аналоговый сигнал может представлять речь, а цифровой — набор двоичных единиц и нулей.

Простейшим типом сигнала является **периодический сигнал**, в котором некоторая структура периодически повторяется во времени. На рис. 2.2 приведен пример периодического аналогового сигнала (синусоида) и периодического цифрового сигнала (прямоугольный сигнал, или меандр). Математическое определение: сигнал  $s(t)$  является периодическим тогда и только тогда, когда

$$s(t + T) = s(t) \quad -\infty < t < +\infty,$$

где постоянная  $T$  является периодом сигнала ( $T$  — наименьшая величина, удовлетворяющая этому уравнению). Если невозможно найти  $T$ , удовлетворяющее уравнению, сигнал называется **аперiodическим**.

---

<sup>1</sup> Это определение идеализировано. На самом деле переход от одного уровня напряжения к другому не может быть моментальным, а происходит в течение некоторого переходного периода. Тем не менее на практике цифровые сигналы являются хорошим приближением к идеальной модели с постоянными уровнями напряжения и мгновенными переходами.

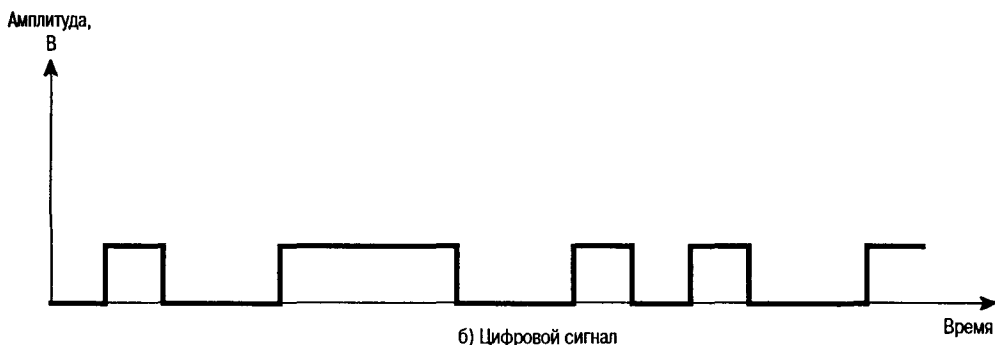
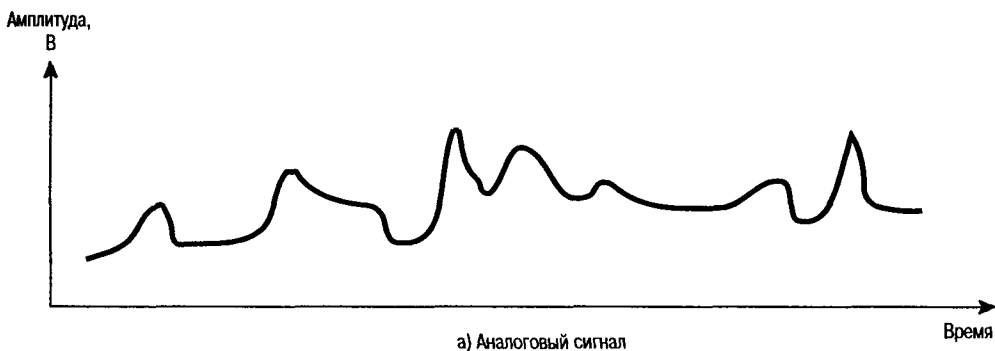


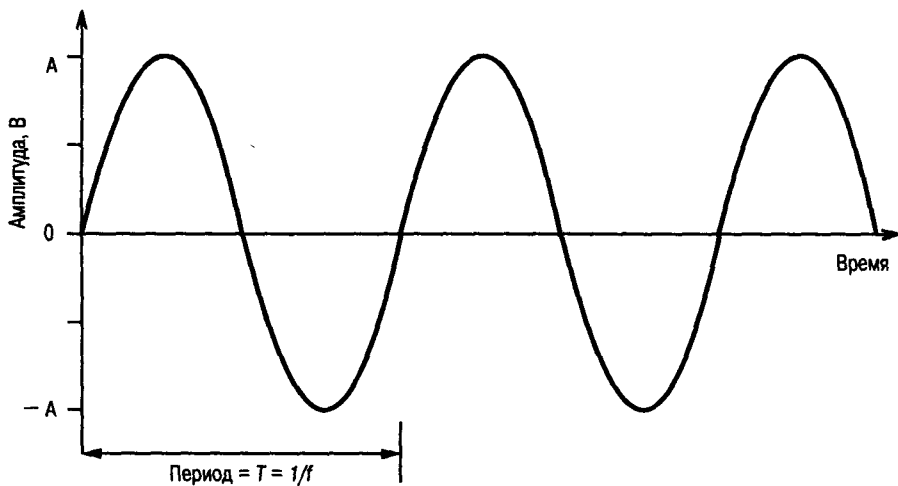
Рис. 2.1. Аналоговый и цифровой сигналы

Фундаментальным аналоговым сигналом является синусоида. В общем случае такой сигнал можно определить тремя параметрами: максимальной амплитудой  $A$ , частотой  $f$  и фазой  $\phi$ . **Максимальной амплитудой** называется максимальное значение или интенсивность сигнала во времени; измеряется максимальная амплитуда, как правило, в вольтах. **Частотой** называется темп повторения сигналов (в периодах за секунду, или герцах). Эквивалентным параметром является **период** сигнала  $T$ , представляющий собой время, за которое происходит повторение сигнала; следовательно,  $T = 1/f$ . **Фаза** является мерой относительного сдвига по времени в пределах отдельного периода сигнала (данный термин будет проиллюстрирован несколько ниже).

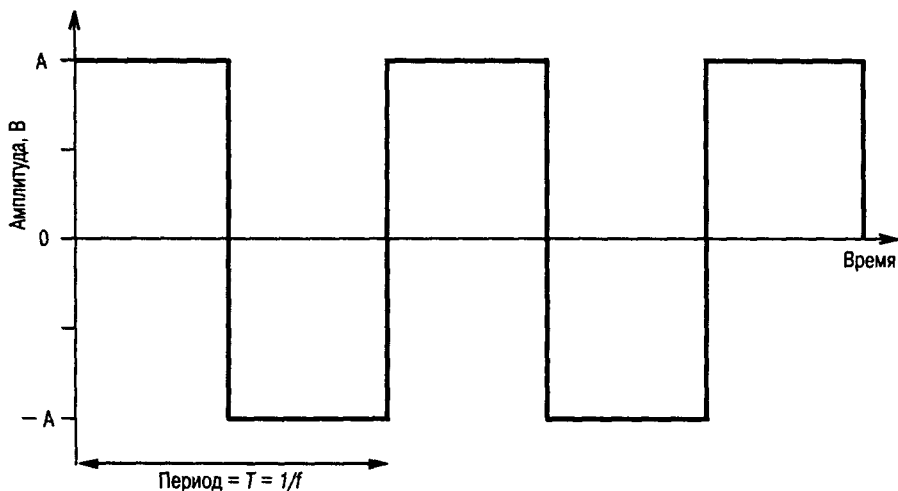
В общем случае синусоидальный сигнал можно представить в следующем виде:

$$s(t) = A \sin(2\pi f t + \phi).$$

Влияние изменения каждого из трех параметров показано на рис. 2.3. На рис. 2.3, *а* частота составляет 1 Гц; следовательно, период  $T$  равен 1 с. На рис. 2.3, *б* частота и фаза те же, но амплитуда уменьшена в два раза. На рис. 2.3, *в* частота  $f = 2$ , что эквивалентно периоду  $T = 1/2$ . Наконец, на рис. 2.3, *г* показано влияние сдвига фазы на  $\pi/4$  радиан, что составляет  $45^\circ$  ( $2\pi$  радиан =  $360^\circ = 1$  период).



а) Синусоидальный сигнал



б) Прямоугольный сигнал

Рис. 2.2. Периодические сигналы

По горизонтальной оси на рис. 2.3 отложено время; на самом же графике показана зависимость от времени величины сигнала в данной точке пространства. Подобные графики (с точностью до изменения масштаба) можно получить, если отложить на горизонтальной оси расстояние. В этом случае на графике будет изображена интенсивность сигнала в данный момент времени в зависимости от расстояния. Например, при передаче синусоидальной волны (если рассмотреть электромагнитную волну на некотором расстоянии от антенны или звук на некотором расстоянии от громкоговорителя) в отдельный момент времени интенсивность сигнала изменяется по гармоническому закону как функция расстояния от источника.

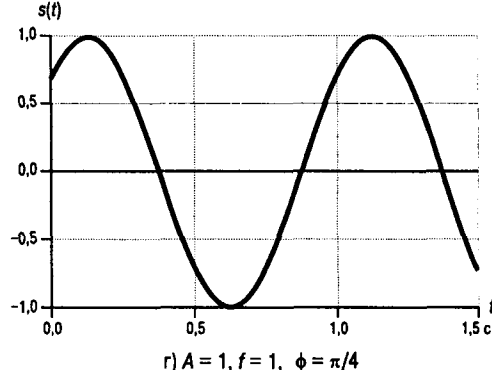
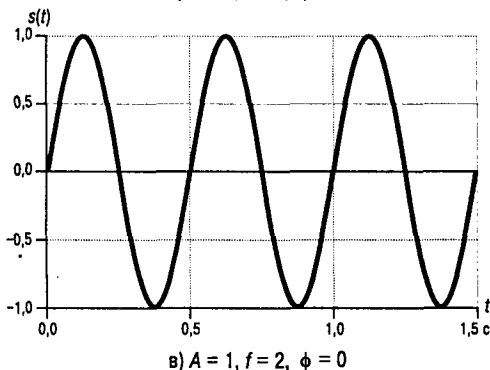
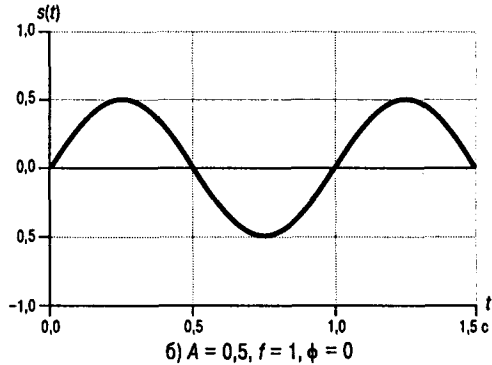
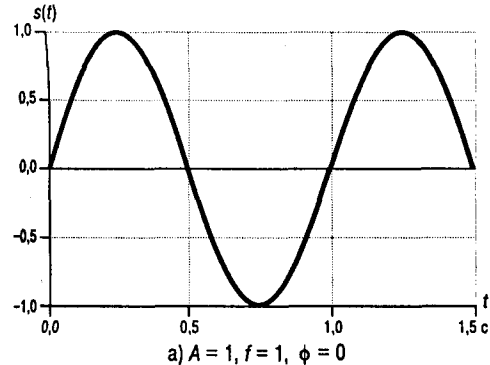


Рис. 2.3.  $s(t) = A \sin(2\pi ft + \phi)$

Существует простое соотношение между двумя синусоидальными сигналами, один из которых изменяется во времени, а другой — в пространстве. Определим длину волны сигнала  $\lambda$  как расстояние, занимаемое одним периодом или, иными словами, как расстояние между двумя точками равных фаз двух последовательных циклов. Предположим, что сигнал распространяется со скоростью  $v$ . Тогда длина волны связана с периодом следующим соотношением:  $\lambda = vT$ , что равносильно  $\lambda f = v$ . Особое значение для нашего изложения имеет случай  $v = c$ , где  $c$  — скорость света в вакууме, приблизительно равная  $3 \times 10^8$  м/с.

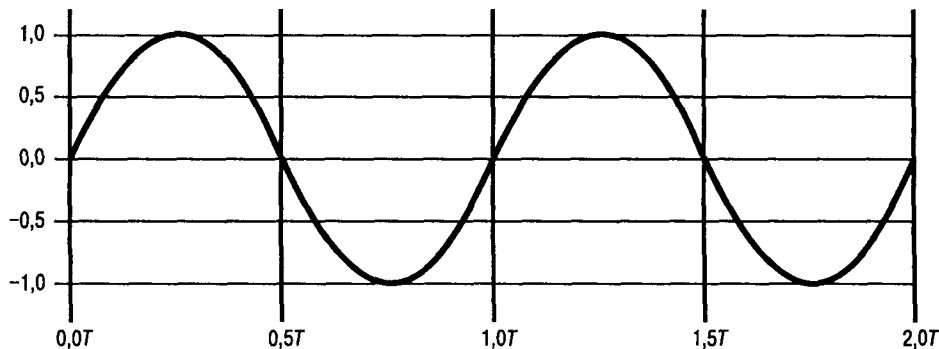
## Основные понятия частотного представления сигнала

Реальный электромагнитный сигнал составлен из многих частот. Рассмотрим, например, сигнал, показанный на рис. 2.4, в.

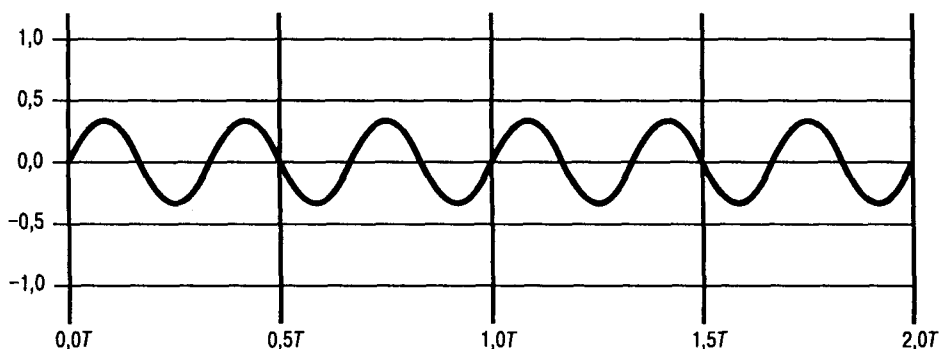
$$s(t) = (4/\pi) \times [\sin(2\pi ft) + (1/3) \sin(2\pi(3f)t)]$$

Он состоит из простых синусоидальных сигналов с частотами  $f$  и  $3f$ , показанных соответственно на рис. 2.4, а и б. Отметим два интересных момента, связанных с этим рисунком.

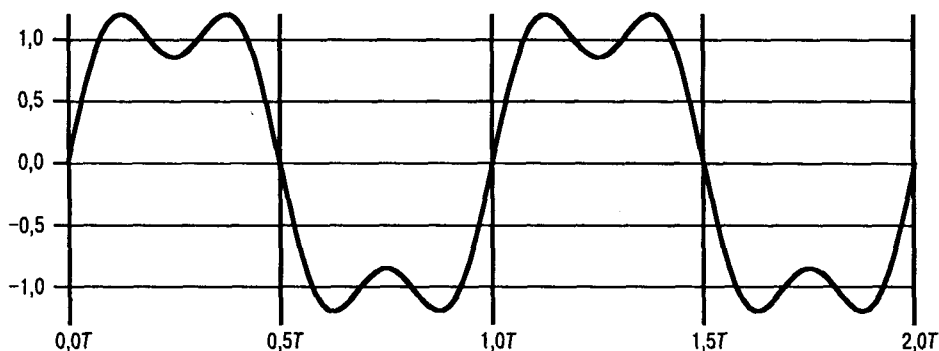




а)  $\sin(2\pi ft)$



б)  $(1/3)\sin(2\pi(3f)t)$



в)  $(4/\pi) [\sin(2\pi ft) + (1/3)\sin(2\pi(3f)t)]$

Рис. 2.4. Сложение частотных составляющих ( $T = 1/f$ )

- Вторая частота кратна первой. Если все частотные составляющие сигнала кратны одной частоте, то последняя называется **собственной частотой**.
- Период суммарного сигнала равен периоду сигнала собственной частоты. Период составляющей  $\sin(2\pi ft)$  равен  $T = 1/f$ , и, как можно увидеть на рис. 2.4, в, период сигнала  $s(t)$  также равен  $T$ .

Можно показать, применив для этой цели методику, известную как анализ Фурье, что любой сигнал складывается из синусоидальных составляющих с разными частотами. Сложив вместе достаточное количество синусоидальных сигналов с соответствующими амплитудами, частотами и фазами, можно получить электромагнитный сигнал любой формы. Аналогично, любой электромагнитный сигнал рассматривается как совокупность периодических аналоговых (синусоидальных) сигналов с разными амплитудами, частотами и фазами. Далее по ходу изложения мы увидим, как важно иметь возможность рассмотреть сигнал не как процесс изменения во времени (временное представление), а как функцию частоты (частотное представление). Для читателя, интересующегося анализом Фурье, в приложении В приводятся основы этого метода.

Спектром сигнала называется область частот, составляющих данный сигнал. Для сигнала, приведенного на рис. 2.4, *в*, спектр простирается от  $f$  до  $3f$ . Абсолютной шириной полосы сигнала называется ширина его спектра. В рассматриваемом случае (рис. 2.4, *в*) ширина полосы составляет  $3f - f = 2f$ . Многие сигналы имеют бесконечную ширину полосы, но большая часть их энергии сосредоточена в относительно узкой полосе частот, называемой **эффективной полосой**, или просто **полосой**.

## Связь между скоростью передачи данных и шириной полосы

Существует прямая связь между информационной емкостью сигнала и шириной его полосы: чем шире полоса, тем больше информации может нести сигнал. Рассмотрим очень простой пример, воспользовавшись сигналом, показанным на рис. 2.2, *б*. Предположим, что положительный импульс представляет двоичный ноль, а отрицательный — двоичную единицу. Следовательно, данный сигнал представляет двоичный поток 0101... Длительность каждого импульса равна  $1/2f$ ; следовательно, скорость передачи данных составляет  $2f$  битов в секунду (бит/с). Каковы частотные составляющие этого сигнала? Чтобы ответить на этот вопрос, вновь обратимся к рис. 2.4. При сложении синусоид с частотами  $f$  и  $3f$  мы получаем сигнал, форма которого начинает походить на форму исходного прямоугольного сигнала. Продолжим этот процесс и добавим синусоидальный сигнал с частотой  $5f$  (результат показан на рис. 2.5, *а*), а затем сигнал с частотой  $7f$  (рис. 2.5, *б*). Продолжая добавлять составляющие с нечетными частотами, кратными  $f$ , и надлежащим образом выбранными амплитудами, мы увидим, что результирующий сигнал все больше и больше приближается к прямоугольной форме.

Действительно, можно показать, что составляющие прямоугольного сигнала с амплитудами  $A$  и  $-A$  можно выразить следующим образом:

$$s(t) = A \times \frac{4}{\pi} \sum_{k=1,3,5,\dots}^{\infty} \frac{\sin(2\pi kft)}{k}.$$

Этот сигнал содержит бесконечное число частотных составляющих и, следовательно, имеет бесконечную ширину полосы. Впрочем, максимальная амплитуда  $k$ -й составляющей с частотой  $kf$  равна всего лишь  $1/k$ , поэтому большая часть энергии данного сигнала приходится на несколько первых состав-

ляющих. Что произойдет, если мы ограничим полосу только первыми тремя частотными составляющими? Ответ мы уже видели, он приведен на рис. 2.5, а. Здесь форма результирующего сигнала достаточно близка к форме исходного прямоугольного сигнала.

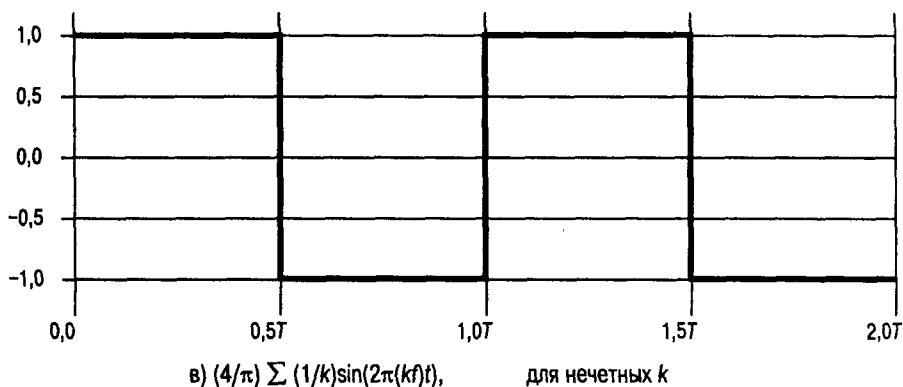
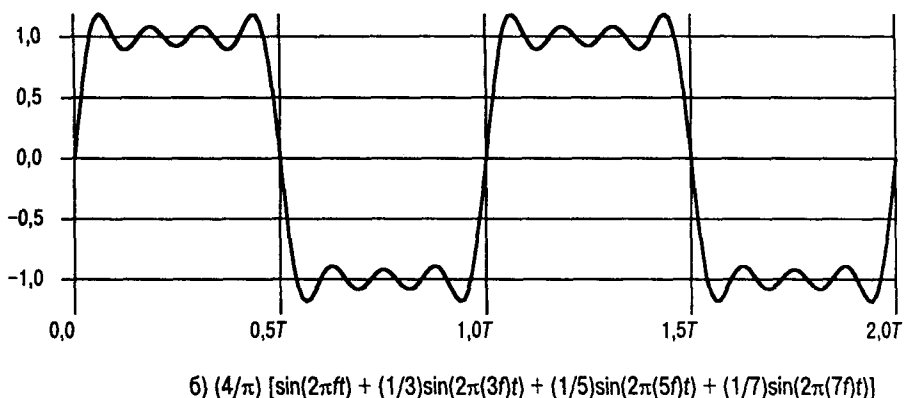
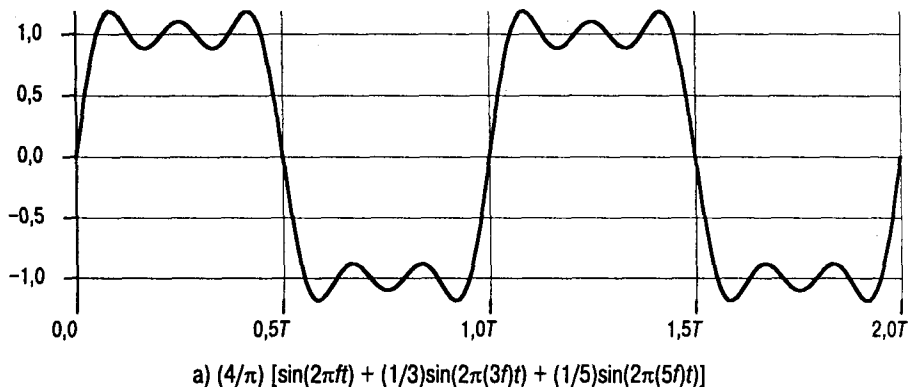


Рис. 2.5. Частотные составляющие прямоугольного сигнала ( $T = 1/f$ )

Рисунки 2.4 и 2.5 можно использовать для иллюстрации связи между скоростью передачи данных и шириной полосы. Предположим, что мы ис-

пользуем цифровую систему, способную передавать сигналы с шириной полосы 4 МГц. Попытаемся передать последовательность чередующихся нулей и единиц в виде сигнала прямоугольной формы, приведенного на рис. 2.5, а. Какой скорости передачи данных мы сможем при этом добиться? Рассмотрим три случая.

**Случай 1.** Аппроксимируем наш прямоугольный сигнал формой, показанной на рис. 2.5, а. Хотя эта форма представляет “искаженный” прямоугольный сигнал, она достаточно близка к нему, чтобы приемник мог отличить двоичный ноль от двоичной единицы. Если мы производим  $10^6$  циклов в секунду, что дает частоту  $f = 1$  МГц, то ширина полосы сигнала

$$s(t) = \frac{4}{\pi} \times \left[ \sin((2\pi \times 10^6)t) + \frac{1}{3} \sin((2\pi \times 3 \times 10^6)t) + \frac{1}{5} \sin((2\pi \times 5 \times 10^6)t) \right]$$

равна  $(5 \times 10^6) - 10^6 = 4$  МГц. Отметим, что для частоты  $f = 1$  МГц период собственной частоты равен  $T = 1/10^6 = 10^{-6} = 1$  мкс. Если мы будем рассматривать данный сигнал как поток двоичных нулей и единиц, то каждые 0,5 мкс будет передаваться один бит при скорости передачи данных  $2 \times 10^6 = 2$  Мбит/с. Следовательно, при ширине полосы 4 МГц достигается скорость передачи данных 2 Мбит/с.

**Случай 2.** Предположим теперь, что ширина полосы равна 8 МГц. Вернемся к рис. 2.5, а, подразумевая теперь частоту  $f = 2$  МГц. Используя ту же цепочку рассуждений, получим следующую ширину полосы сигнала:  $(5 \times 2 \times 10^6) - (2 \times 10^6) = 8$  МГц. Но в этом случае период  $T$  равен уже  $1/f = 0,5$  мкс. В результате один бит будет передаваться каждые 0,25 мкс при скорости передачи данных 4 Мбит/с. Следовательно, при равных других параметрах, удвоение ширины полосы приводит к удвоению возможной скорости передачи данных.

**Случай 3.** Предположим теперь, что форма сигнала, приведенная на рис. 2.4, в, достаточно для аппроксимации прямоугольного сигнала. Иначе говоря, различия между положительным и отрицательным импульсом на рис. 2.4, в достаточно, чтобы успешно использовать этот сигнал для представления последовательности нулей и единиц. Предположим, как и в случае 2, что частота  $f$  равна 2 МГц, а период  $T = 1/f = 0,5$  мкс; таким образом, один бит передается каждые 0,25 мкс, а ширина полосы сигнала составляет  $(3 \times 2 \times 10^6) - (2 \times 10^6) = 4$  МГц. Следовательно, при данной ширине полосы могут поддерживаться различные скорости передачи данных, а зависит эта скорость от способности приемника различать 0 и 1 при наличии помех и других искажений сигнала.

Подытожим наши рассуждения.

- Случай 1: ширина полосы — 4 МГц; скорость передачи данных — 2 Мбит/с.
- Случай 2: ширина полосы — 8 МГц; скорость передачи данных — 4 Мбит/с.
- Случай 3: ширина полосы — 4 МГц; скорость передачи данных — 4 Мбит/с.

Из приведенного выше изложения можем сделать следующие выводы. В общем случае любой цифровой сигнал имеет бесконечную ширину полосы. Если мы попы-

таемся передать этот сигнал через какую-то среду, передающая система наложит ограничения на ширину полосы, которую можно передать. Более того, для каждой конкретной среды справедливо следующее: чем больше передаваемая полоса, тем больше стоимость передачи. Поэтому, с одной стороны, по экономическим и практическим соображениям следует аппроксимировать цифровую информацию сигналом с ограниченной шириной полосы. С другой стороны, при ограничении ширины полосы возникают искажения, затрудняющие интерпретацию принимаемого сигнала. Чем больше ограничена полоса, тем больше искажение сигнала и тем больше потенциальная возможность возникновения ошибок при приеме.

## 2.2. АНАЛОГОВАЯ И ЦИФРОВАЯ ПЕРЕДАЧА ДАННЫХ

Термины *аналоговый* и *цифровой* приблизительно соответствуют терминам *непрерывный* и *дискретный*. В сфере передачи данных эти два термина часто используются по меньшей мере в трех контекстах: при рассмотрении данных, сигналов и передачи.

Определим данные как объекты, передающие смысл, или информацию. Сигналы — это электрическое или электромагнитное представление данных. Передача — процесс перемещения данных путем распространения сигналов по передающей среде и их обработки. Далее мы попытаемся прояснить эти абстрактные понятия, используя термины *аналоговый* и *цифровой* применительно к данным, сигналам и передаче.

### Аналоговые и цифровые данные

Понятия *аналоговые* и *цифровые* данные достаточно просты. Аналоговые данные принимают непрерывные значения из некоторого диапазона. Например, звуковые сигналы и видеосигналы представляют собой непрерывно изменяющиеся величины. Многие данные, собираемые с помощью датчиков, такие, как температура и давление, также принимают непрерывные значения. Цифровые данные, напротив, принимают только дискретные значения; примеры — текст и целые числа.

Самым привычным примером аналоговых данных является звук, который, в форме акустических волн, люди могут воспринимать непосредственно. На рис. 2.6 показан акустический спектр человеческой речи и музыки. Частотные составляющие обычной речи лежат в диапазоне приблизительно между 100 Гц и 7 кГц. Хотя значительная часть энергии речевого сигнала сосредоточена в области низких частот, исследования показали, что частоты ниже 600–700 Гц слабо влияют на улучшение разборчивости сигнала для человеческого уха. Динамический диапазон обычной речи составляет примерно 25 дБ<sup>2</sup>; т.е. мощность самого громкого крика в 300 раз больше, чем самого тихого шепота. Помимо речи, на рис. 2.6 также показаны акустический спектр и динамический диапазон музыки.

---

<sup>2</sup> О децибелах рассказано в приложении 2А.

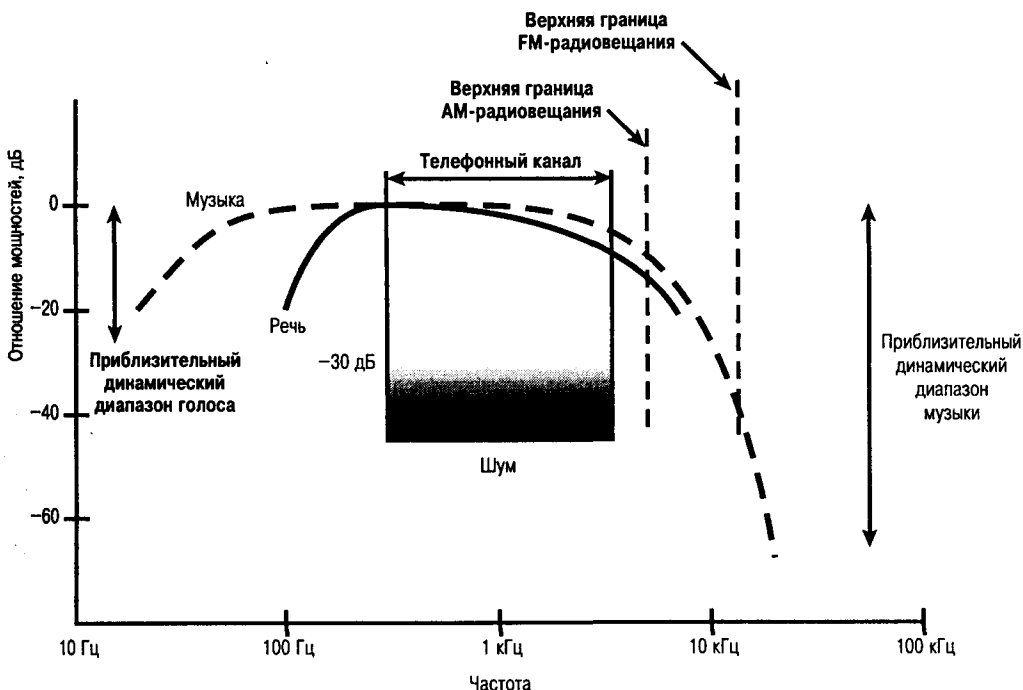


Рис. 2.6. Акустический спектр речи и музыки [CARN99]

## Аналоговые и цифровые сигналы

В системе связи информация распространяется от одной точки к другой посредством электрических сигналов. Аналоговый сигнал представляет собой непрерывно изменяющуюся электромагнитную волну, которая может распространяться через множество сред, в зависимости от частоты; в качестве примеров таких сред можно назвать проводные линии, такие, как витая пара и коаксиальный кабель, оптоволокно; этот сигнал может также распространяться через атмосферу или космическое пространство. Цифровой сигнал представляет собой последовательность импульсов напряжения, которые могут передаваться по проводной линии; при этом постоянный положительный уровень напряжения может использоваться для представления двоичного нуля, а постоянный отрицательный уровень — для представления двоичной единицы.

Основное преимущество цифровых сигналов состоит в том, что их передача в общем случае дешевле и менее восприимчива к помехам, чем передача аналоговых сигналов. Основной недостаток — цифровым сигналам затухание вредит больше, чем аналоговым. На рис. 2.7 показаны исходная последовательность импульсов напряжения, генерируемых источником, и эти же импульсы, прошедшие некоторое расстояние по передающей среде. Из-за затухания, или ослабления, мощности сигнала на высоких частотах импульсы становятся более сглаженными и низкими. Ясно, что это затухание довольно быстро может привести к потере информации, содержащейся в передаваемом сигнале.

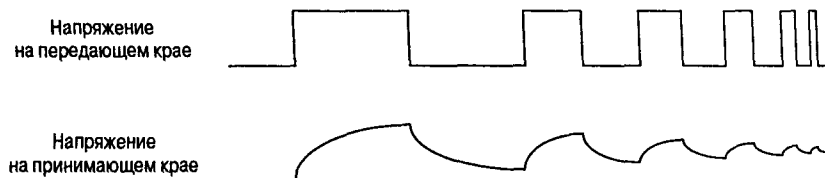


Рис. 2.7. Ослабление цифровых сигналов

И аналоговые, и цифровые данные могут быть представлены, а следовательно, и переданы, в форме как аналоговых, так и цифровых сигналов. Это показано на рис. 2.8. В общем случае аналоговые данные являются функцией времени и занимают ограниченный спектр частот. Такие данные можно непосредственно представить электромагнитным сигналом, имеющим аналогичный спектр. Наиболее наглядный пример — передача речи. Речь, представляя собой звуковые волны, содержит частотные составляющие в области 20 Гц — 20 кГц. Однако большая часть энергии речи находится в намного более узком диапазоне. Стандартный спектр речевых сигналов — 300–3400 Гц, и этого диапазона вполне хватает для разборчивой и четкой передачи речи. Именно такой диапазон обрабатывает телефонный аппарат. Все поступающие звуковые колебания в диапазоне 300–3400 Гц преобразуются в электромагнитный сигнал с подобными амплитудами и частотами. В другом аппарате выполняется обратный процесс: электромагнитная энергия преобразуется в звук.

Цифровые данные также можно представить аналоговыми сигналами, применив с этой целью модем (модулятор/демодулятор). Модем преобразует последовательность двоичных (принимающих два значения) импульсов напряжения в аналоговый сигнал, модулируя их несущей частотой. Получившийся в результате сигнал занимает определенный спектр частот с центром на несущей частоте и может распространяться в среде, подходящей для этой несущей частоты. Наиболее распространенные модемы представляют цифровые данные в спектре частот, совпадающем со спектром речи, позволяя, таким образом, передавать эти данные по обычным телефонным линиям. На другом конце линии другой модем демодулирует сигнал и восстанавливает исходные данные.

С использованием очень похожего способа аналоговые данные могут представляться и цифровыми сигналами. Устройство, выполняющее функцию передачи речевой информации, называется кодеком (кодером/декодером). По существу кодек принимает аналоговый сигнал, непосредственно представляющий речевую информацию, и аппроксимирует этот сигнал потоком битов. Приемник, со своей стороны, использует этот поток битов для воспроизведения аналоговых данных. В дальнейшем мы вернемся к этой теме.

Наконец, цифровые данные могут быть представлены непосредственно в двоичной форме, двумя уровнями напряжения. Однако для улучшения характеристик распространения двоичные данные часто перекодируются в цифровой сигнал более сложной формы, как будет показано далее.

Каждая из четырех возможных комбинаций (они представлены в табл. 2.1, а) находит широкое применение. Причины выбора какой-либо конкретной комбинации различны. Ниже мы постарались перечислить наиболее типичные из них.

- **Цифровые данные, цифровой сигнал.** В общем случае оборудование для кодирования цифровых данных цифровым сигналом дешевле и проще, чем оборудование для модулирования цифровых данных аналоговым сигналом.
- **Аналоговые данные, цифровой сигнал.** Преобразование аналоговых данных в цифровую форму позволяет использовать современное цифровое оборудование передачи и коммутации.

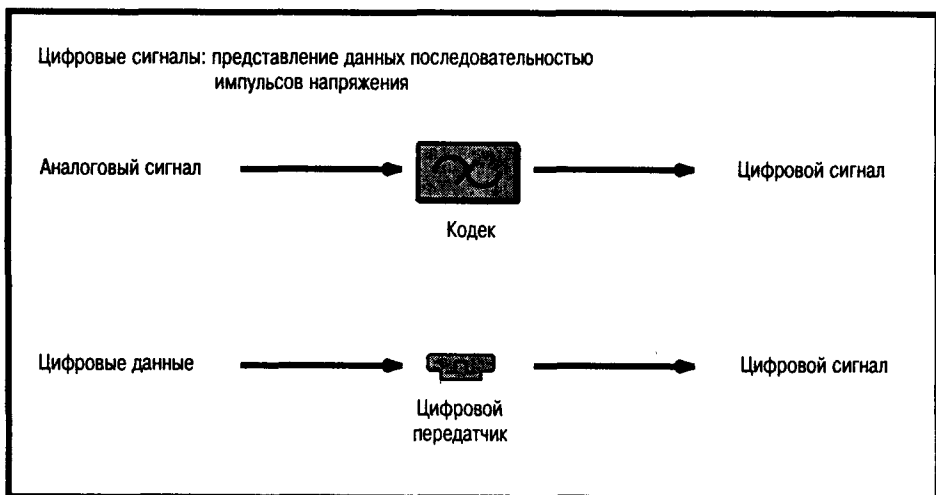
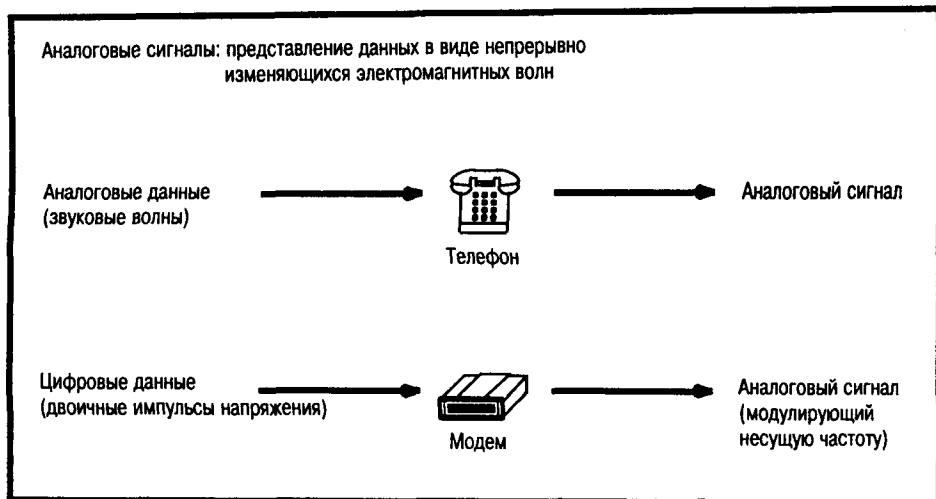


Рис. 2.8. Аналоговая и цифровая передача аналоговых и цифровых данных

- **Цифровые данные, аналоговый сигнал.** По некоторым передающим средам, таким, как оптоволокно и ненаправляемые среды, могут распространяться только аналоговые сигналы.
- **Аналоговые данные, аналоговый сигнал.** Аналоговые данные легко поддаются преобразованию в аналоговый сигнал.



Таблица 2.1. Аналоговая и цифровая передача

а) Данные и сигналы		
	Аналоговый сигнал	Цифровой сигнал
Аналоговые данные	Два варианта: 1) сигнал имеет тот же спектр, что и аналоговые данные; 2) аналоговые данные кодируются с целью получения сигнала в другой области спектра	Аналоговые данные кодируются с использованием кодека для создания потока битов
Цифровые данные	Цифровые данные кодируются с использованием модема для создания аналогового сигнала	Два варианта: 1) сигнал состоит из двух уровней напряжения, представляющих два двоичных значения; 2) цифровые данные кодируются для создания цифрового сигнала с желаемыми свойствами
б) Обращение с сигналами		
	Аналоговая передача	Цифровая передача
Аналоговый сигнал	Передается через усилители; такой же подход, если сигнал используется для представления аналоговых или цифровых данных	Предполагает использование аналогового сигнала для представления цифровых данных. Сигнал передается через ретрансляторы, на каждом из которых цифровые данные восстанавливаются из поступающего сигнала и используются для создания нового отправляемого аналогового сигнала
Цифровой сигнал	Не используется	Цифровые данные являются потоком нулей и единиц, которые могут представлять цифровые данные или быть закодированными аналоговыми данными. Сигнал передается через ретрансляторы, на каждом из которых поток нулей и единиц восстанавливается из поступающего сигнала и используется для создания нового отправляемого цифрового сигнала

## Аналоговая и цифровая передача

Как аналоговые, так и цифровые сигналы могут распространяться в соответствующих передающих средах. Но конкретный способ обращения с этими сигналами является функцией передающей системы. Существующие методы передачи данных обобщены в табл. 2.1, б. Аналоговая передача — это средство передачи аналоговых сигналов, причем передаваемые данные не конкретизируются: сигнал может представлять как аналоговые (например, речь), так и цифровые данные (например, двоичные данные, проходящие через модем). В любом случае аналоговый сигнал испытывает затухание (ослабляется), что ограничивает длину линии передачи. Чтобы сигнал можно было передавать на большие расстояния, в аналоговые передающие системы вводятся усилители, повышающие энергию сигнала. К сожалению, усилители также усиливают и шумовые составляющие сигнала. С каждым усилителем, последовательно расположенным на пути следования сигнала, этот сигнал становится все более искаженным. Для аналоговых

данных, таких, как речь, небольшое искажение не существенно, и данные остаются понятными, чего нельзя сказать о цифровых данных: в них последовательные усилители вносят ошибки.

**Цифровая передача данных**, наоборот, связана с содержанием сигнала. Мы уже говорили, что цифровой сигнал можно передать только на ограниченное расстояние, пока затухание не нарушит целостности данных. Для передачи цифровых данных на большие расстояния используются ретрансляторы, которые принимают цифровой сигнал, восстанавливают закодированную комбинацию нулей и единиц и передают новый сигнал. Таким образом происходит компенсация затухания.

Тот же метод может использоваться и при передаче аналогового сигнала, если он переносит цифровую информацию. Для этого в расположенных соответствующим образом точках передающей системы помещаются не усилители, а ретрансляторы. Такой ретранслятор восстанавливает цифровую информацию из аналогового сигнала и создает новый, чистый аналоговый сигнал, препятствуя, таким образом, накоплению помех.

## 2.3. ПРОПУСКНАЯ СПОСОБНОСТЬ КАНАЛА

Существует множество факторов, способных исказить или повредить сигнал. Наиболее распространенные из них — помехи или шумы, представляющие собой любой нежелательный сигнал, который смешивается с сигналом, предназначенным для передачи или приема, и искажает его. Помехи и другие искажения рассматриваются в главе 5. Для цифровых данных возникает вопрос: насколько эти искажения ограничивают возможную скорость передачи данных. Максимально возможная при определенных условиях скорость, при которой информация может передаваться по конкретному тракту связи, или каналу, называется **пропускной способностью канала**.

Существуют четыре понятия, которые мы попытаемся связать воедино.

- **Скорость передачи данных.** Это скорость в битах в секунду (бит/с), с которой могут передаваться данные.
- **Ширина полосы.** Это ширина полосы передаваемого сигнала, ограничиваемая передатчиком и природой передающей среды. Выражается в периодах в секунду, или герцах (Гц).
- **Шум.** Средний уровень шума в канале связи.
- **Уровень ошибок.** Частота появления ошибок. Ошибкой считается прием 1 при переданном 0 и наоборот.

Проблема, стоящая перед нами, заключается в следующем: средства связи недешевы и, в общем случае, чем шире их полоса, тем дороже они стоят. Более того, все каналы передачи, представляющие практический интерес, имеют ограниченную ширину полосы. Ограничения обусловлены физическими свойствами передающей среды или преднамеренными ограничениями ширины полосы в самом передатчике, сделанными для предотвращения интерференции с другими источниками. Естественно, нам хотелось бы максимально эффективно использовать имеющуюся полосу. Для цифровых данных это означает, что для определенной полосы желательно получить максимально возможную при существующей

щем уровне ошибок скорость передачи данных. Главным ограничением при достижении такой эффективности являются помехи.

## Ширина полосы по Найквисту

Рассмотрим вначале канал в отсутствие шумов. В таких условиях ограничения на скорость передачи данных накладывает только ширина полосы сигнала. По Найквисту это ограничение формулируется следующим образом: если скорость передачи сигнала равна  $2B$ , то с этой скоростью могут передаваться сигналы с частотами, не превышающими  $B$ . Справедливо и обратное: для сигнала с шириной полосы  $B$  наивысшая скорость передачи составляет  $2B$ . Это ограничение возникает вследствие эффекта межсимвольной интерференции, такой, как интерференция при искажении, вызванном запаздыванием.<sup>3</sup> В приложении 6А выводится одно из следствий этого факта, полезное при разработке схем кодирования цифровых объектов в аналоговые.

Отметим, что в предыдущем абзаце мы упоминали о скорости передачи сигнала. Если передаваемый сигнал является двоичным (т.е. имеет два уровня напряжения), то скорость передачи данных, которая может поддерживаться при полосе  $B$  Гц, равна  $2B$  бит/с. В качестве примера рассмотрим телефонный канал, который используется для передачи цифровых данных с помощью модема. Предположим, что ширина полосы равна 3100 Гц. В этом случае пропускная способность  $C$  канала составляет  $2B = 6200$  бит/с. Впрочем, как будет показано в главе 6, можно применять сигналы более чем с двумя уровнями напряжения; т.е. каждая сигнальная посылка может представлять несколько битов. Если, например, в качестве сигналов используются четыре возможных уровня напряжения, то каждая сигнальная посылка может представлять два бита. При многоуровневой передаче данных формула Найквиста принимает следующий вид:

$$C = 2B \log_2 M,$$

где  $M$  — количество дискретных сигналов или уровней напряжения. Следовательно, для  $M = 8$ , как в некоторых модемах, пропускная способность равна 18 600 бит/с.

Итак, при заданной ширине полосы скорость передачи данных может быть увеличена за счет увеличения количества различных элементов сигнала. Впрочем, при таком подходе на приемник ложится дополнительная нагрузка: в каждый момент приемник должен распознавать уже не два, а  $M$  возможных элементов сигнала. Кроме того, на практике величину  $M$  ограничивают помехи и прочие искажения сигнала, возникающие в линии передачи.

## Формула Шеннона для пропускной способности

Из формулы Найквиста видно, что при неизменных остальных параметрах удвоение ширины полосы удваивает скорость передачи данных. Рассмотрим теперь связь между скоростью передачи данных, шумом и уровнем ошибок. Наличие шума может привести к повреждению одного или нескольких битов. При увеличении скорости

---

<sup>3</sup> Искажение, вызванное запаздыванием, возникает, когда задержка распространения сигнала по передающей среде неодинакова для различных частотных составляющих этого сигнала.

передачи данных биты становятся “короче”, поэтому при данном шуме поражается уже большее количество битов. Следовательно, чем выше скорость передачи данных при определенном уровне шума, тем выше уровень ошибок.

Пример воздействия шума на цифровой сигнал показан на рис. 2.9. Здесь шум складывается из фоновых помех относительно умеренного уровня и случайных всплесков импульсных помех. Цифровую информацию можно восстановить из сигнала путем дискретизации полученной формы сигнала, т.е. ее измерения через определенные промежутки времени. Как легко увидеть, случайного шума достаточно для изменения 1 на 0 или 0 на 1.

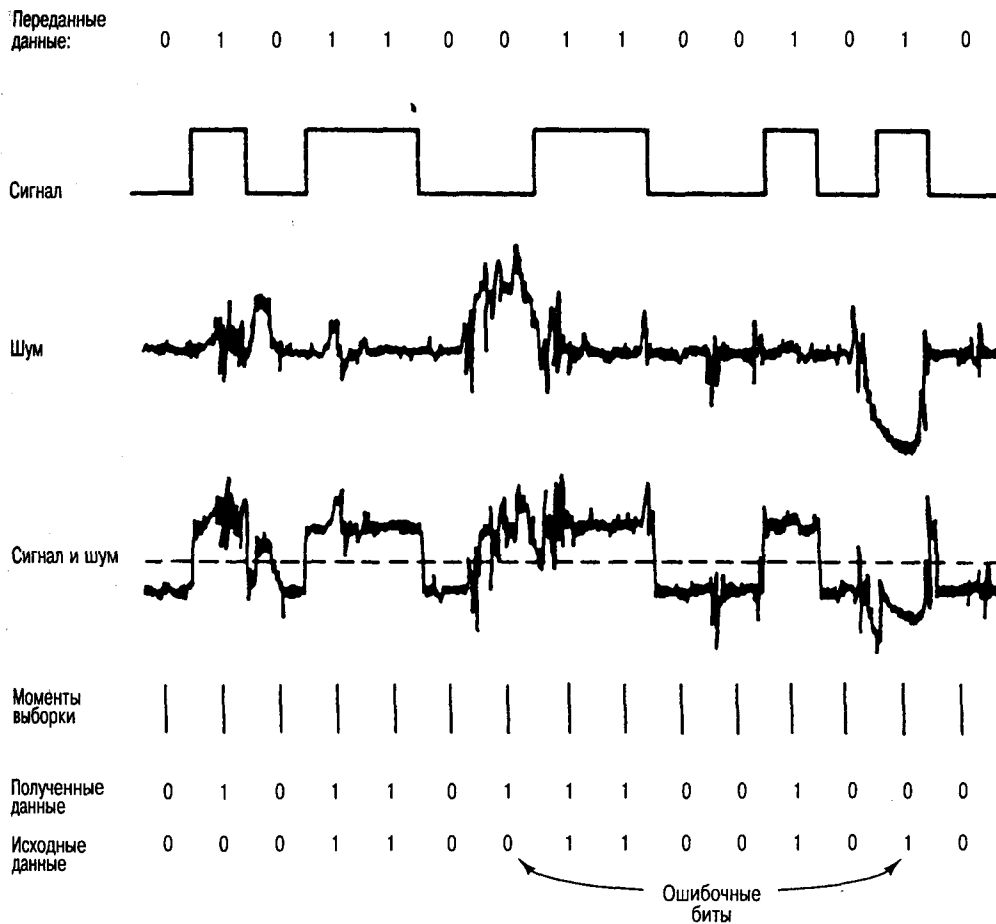


Рис. 2.9. Воздействие шума на цифровой сигнал

Все упомянутые выше понятия можно явно связать формулой, выведенной математиком Клодом Шенноном (Claude Shannon). Как мы только что показали, чем выше скорость передачи данных, тем больший ущерб может нанести нежелательный шум. При данном уровне шума следует ожидать, что сигнал большей интенсивности имеет больше шансов прибыть по назначению неповрежденным. Ключевым параметром, присутствующим в этих рассуждениях, является отно-

шение сигнал/шум (signal-to-noise ratio — SNR или  $S/N^4$ ). Оно представляет собой отношение мощности сигнала к мощности шума, присутствующего в определенный момент передачи. Как правило, данное отношение измеряется в приемнике, поскольку именно в этой точке предпринимается попытка обработать сигнал и устранить нежелательный шум. Для удобства это отношение часто представляется в децибелах:

$$\text{SNR}_{\text{дб}} = 10 \lg \frac{\text{мощность сигнала}}{\text{мощность шума}}.$$

Данная формула в децибелах выражает превышение уровня сигнала над уровнем шума. Большое значение этого отношения свидетельствует о высоком качестве сигнала и, следовательно, необходимости введения меньшего числа промежуточных ретрансляторов.

Отношение сигнал/шум довольно важно при передаче цифровых данных, поскольку оно задает верхнюю границу возможной скорости передачи. Для максимальной пропускной способности канала Шенноном был получен следующий результат:

$$C = B \log_2 (1 + \text{SNR}),$$

где  $C$  — пропускная способность канала в битах в секунду, а  $B$  — ширина полосы канала в герцах. Сама формула Шеннона дает теоретически достижимый максимум. На практике, однако, достигаются значительно меньшие скорости передачи данных. Одной из причин этого является то, что в формуле учитывается только белый шум (т.е. тепловой) и не учитываются импульсные помехи, амплитудные искажения или искажения, вызванные запаздыванием.

Пропускная способность, выведенная в предыдущей формуле, называется *безошибочной*. Шеннон доказал, что если действительная скорость передачи данных по каналу меньше безошибочной пропускной способности, то при использовании соответствующих сигнальных кодов теоретически можно добиться безошибочной передачи данных по каналу. К сожалению, теорема Шеннона не дает способа нахождения таких кодов, но она дает критерий измерения производительности реальных схем связи.

Приведем несколько поучительных наблюдений, касающихся данной формулы. Кажется, что при фиксированном уровне шума скорость передачи данных можно увеличить за счет увеличения ширины полосы или интенсивности сигнала. Стоит отметить, впрочем, что при увеличении интенсивности сигнала в системе возникают нелинейные эффекты, приводящие к увеличению интермодуляционных помех. Кроме того, поскольку шум считается белым, то более широкая полоса предполагает введение в систему большего шума. Поэтому при увеличении ширины полосы  $B$  отношение сигнал/шум уменьшается, а не увеличивается.

---

<sup>4</sup> В некоторых источниках используется обозначение SNR, а в других —  $S/N$ . Кроме того, в отдельных случаях такая величина, если она безразмерна, обозначается  $\text{SNR}$  или  $S/N$ , а если выражена в децибелах —  $\text{SNR}_{\text{дб}}$  или  $S/N_{\text{дб}}$ . Иногда обозначение SNR или  $S/N$  применяется и для величины в децибелах. В данной книге мы будем использовать SNR и  $\text{SNR}_{\text{дб}}$ .

**Пример.** Рассмотрим пример, относящийся к формулам и Найквиста, и Шеннона. Предположим, что спектр канала простирается от 3 до 4 МГц, а отношение сигнал/шум (SNR) составляет 24 дБ. Тогда

$$\begin{aligned}B &= 4 \text{ МГц} - 3 \text{ МГц} = 1 \text{ МГц}, \\ \text{SNR}_{\text{дБ}} &= 24 \text{ дБ} = 10 \lg (\text{SNR}), \\ \text{SNR} &= 251.\end{aligned}$$

Используем формулу Шеннона

$$C = 10^6 \times \log_2 (1 + 251) = 10^6 \times 8 = 8 \text{ Мбит/с.}$$

Это теоретический предел и, как уже говорилось, достичь его нереально. Предположим, впрочем, что мы его достигли. Сколько, по формуле Найквиста, для этого потребуется уровней сигнала? Получаем

$$\begin{aligned}C &= 2B \log_2 M, \\ 8 \times 10^6 &= 2 \times (10^6) \times \log_2 M, \\ 4 &= \log_2 M, \\ M &= 16.\end{aligned}$$

## 2.4. ПЕРЕДАЮЩИЕ СРЕДЫ

В системе передачи данных **передающая среда** является физическим путем между передатчиком и приемником. Передающие среды можно разделить на направляемые и ненаправляемые. В обоих случаях связь принимает форму распространения электромагнитных волн. В **направляемых средах** волны направляются по твердым проводникам, например, по медным витым парам, медным коаксиальным кабелям и оптическим волокнам. Атмосфера и открытый космос представляют собой примеры уже **ненаправляемых сред**, предоставляющих средства передачи электромагнитных сигналов, но не направляющих их; используемая в этом случае форма передачи обычно называется **беспроводной**.

Характеристики и качество передачи данных определяются как характеристиками среды, так и характеристиками сигнала. В направляемой среде более важным фактором, накладывающим ограничения на передачу, является сама среда. При определении характеристик передачи в ненаправляемой среде более важным фактором является не сама среда, а ширина полосы сигнала, генерируемого передающей антенной. Одной из ключевых характеристик сигналов, передаваемых антенной, является их направленность. В общем случае сигналы низких частот являются ненаправленными, т.е. сигнал распространяется во всех направлениях от антенны. Сигналы высоких частот, наоборот, можно сфокусировать в направленный луч.

На рис. 2.10 изображен спектр электромагнитных волн и обозначены частоты, на которых функционируют различные направляемые среды, и указаны методы ненаправленной передачи.<sup>5</sup> В остальной части этого раздела мы дадим краткий обзор ненаправляемых или беспроводных сред.

<sup>5</sup> Обратите внимание на логарифмический масштаб. Базовые сведения об этом масштабе приводятся в обзоре основных разделов математики на Web-сайте ресурсов для студентов компьютерных специальностей по адресу: [WilliamStallings.com/StudentSupport.html](http://WilliamStallings.com/StudentSupport.html).

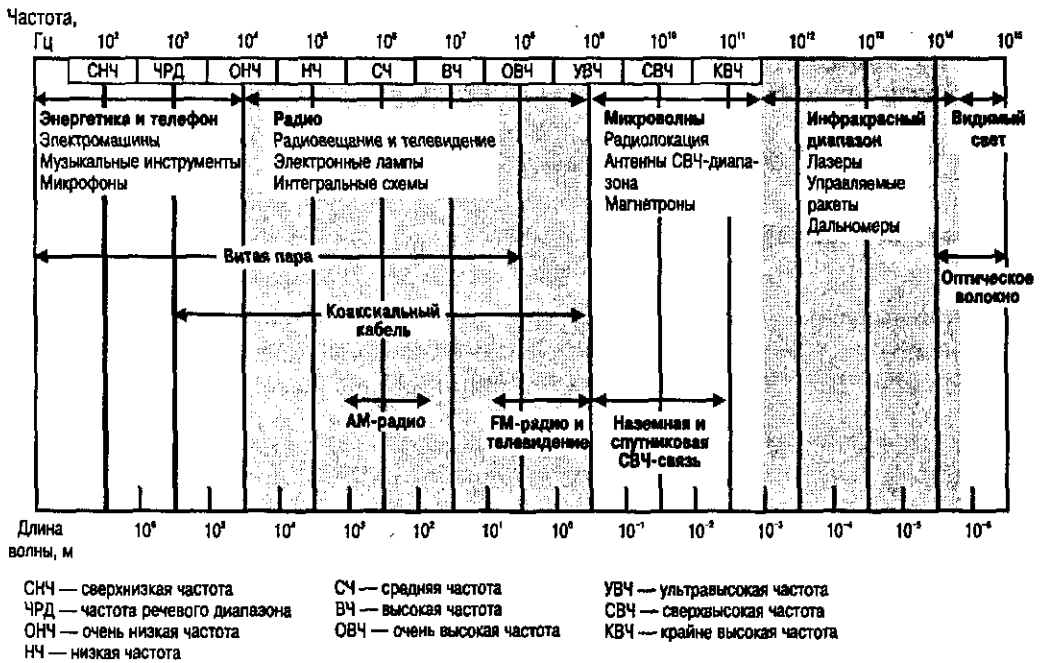


Рис. 2.10. Электромагнитный спектр телекоммуникационных технологий

В ненаправляемых средах передача и прием производятся с помощью антенны. При передаче антенна излучает электромагнитную энергию в среду (обычно воздух), а при приеме антенна улавливает электромагнитные волны из окружающей среды. Существуют, по сути, две формы беспроводной передачи: направленная и ненаправленная. При направленной передаче передающая антенна излучает сфокусированный электромагнитный луч; следовательно, передающая и приемная антенны должны быть тщательно нацелены. При ненаправленной передаче передаваемый сигнал распространяется во всех направлениях и может быть принят множеством антенн. Вообще чем выше частота сигнала, тем легче его сфокусировать в направленный луч.

Для нас представляют интерес три основных диапазона частот, в которых можно вести беспроводную передачу. Частоты в диапазоне от 2 ГГц (1 ГГц = 10<sup>9</sup> Гц) до 40 ГГц называются сверхвысокими частотами (СВЧ). На этих частотах возможно создание узконаправленных лучей, сверхвысокие частоты также вполне подходят для двухточечной передачи и используются в спутниковой связи. Для ненаправленной связи удобно использовать частоты из диапазона 30 МГц–1 ГГц. Этот диапазон частот мы будем называть диапазоном широковещательного радио.

Еще один важный для ближней связи диапазон частот — инфракрасная часть спектра. Она охватывает частоты приблизительно от  $3 \times 10^{11}$  до  $2 \times 10^{14}$  Гц. Передача в инфракрасном диапазоне полезна в приложениях ближней двухточечной и многоточечной связи в замкнутом пространстве, например комнате.

# Наземная связь с использованием СВЧ

## Физическое описание

Самым распространенным типом антенны СВЧ-диапазона является параболическая антенна. Ее диаметр обычно составляет около 3 м. Антенна жестко фиксируется и фокусирует узкий луч для передачи на приемную антенну, которая находится на расстоянии прямой видимости. Чтобы увеличить расстояние между антеннами СВЧ-диапазона, их обычно располагают на значительной высоте над землей. Такое расположение также способствует передаче сигнала над промежуточными препятствиями. Для передачи на дальние расстояния используется цепочка ретрансляционных вышек, а двухточечные СВЧ-каналы в пределах желаемого расстояния связываются в единые пучки.

## Применение

Основной отраслью применения систем наземной СВЧ-связи является служба дальней телекоммуникации, где системы этого типа представляют альтернативу использованию коаксиального кабеля или оптоволокна. При передаче на одинаковое расстояние средства СВЧ-связи требуют значительно меньшего числа усилителей или ретрансляторов, чем коаксиальный кабель, но, в то же время, эти средства позволяют передавать сигнал лишь на расстояние прямой видимости. СВЧ обычно используются как для передачи речи, так и для передачи телевизионного изображения.

Все интенсивнее СВЧ-связь начинает применяться еще в одной области, а именно — для создания коротких двухточечных каналов связи между зданиями. Такие каналы используются в локальных системах телевидения или для связи между локальными сетями. СВЧ-связь ближнего действия также может применяться в так называемых обходных приложениях: фирма может, например, установить СВЧ-канал для сообщения со средствами дальней связи в обход местной телефонной компании.

Существует еще две важных сферы применения СВЧ: сотовые системы и системы стационарного радиодоступа. О них речь пойдет в части III.

## Характеристики передачи

Передача в СВЧ-диапазоне покрывает значительную часть электромагнитного спектра. Для передачи обычно используются частоты от 2 до 40 ГГц. Чем выше используемая частота, тем шире возможная полоса и тем выше потенциальная скорость передачи данных. В табл. 2.2 показана ширина полос и скорость передачи данных нескольких типичных систем.

Таблица 2.2. Производительность типичных цифровых систем СВЧ-связи

Диапазон, ГГц	Ширина полосы, МГц	Скорость передачи данных, Мбит/с
2	7	12
6	30	90
11	40	135
18	220	274



Основным источником потерь в системах СВЧ-связи, как и в любых других передающих системах, является затухание. Для СВЧ (и радиочастот) величина потерь может быть выражена в следующем виде:

$$L = 10 \lg \left( \frac{4\pi d}{\lambda} \right)^2 \text{ дБ}, \quad (2.1)$$

где  $d$  — расстояние,  $\lambda$  — длина волны, измеряемые в одних единицах. Следовательно, потери зависят от квадрата расстояния. В отличие от применения витой пары и коаксиального кабеля потери зависят от расстояния логарифмически (в децибелах — линейно). Следовательно, при использовании систем СВЧ-связи ретрансляторы или усилители могут располагаться реже, среднее расстояние между ними обычно равно 10–100 км. Другим источником искажений является интерференция. В связи с растущей популярностью СВЧ-связи зоны передачи перекрываются, и интерференция становится постоянной опасностью. По этой причине распределение частотных диапазонов строго регулируется.

Самым распространенным диапазоном систем дальней связи, является полоса частот от 4 до 6 ГГц. Вследствие увеличения загрузки этих частот начинает использоваться диапазон 11 ГГц. Следующий диапазон, 12 ГГц, применяется как составляющая систем кабельного телевидения. Каналы СВЧ-связи используются для доставки телесигналов к местным объектам абонентского телевидения; после чего сигналы распределяются по отдельным абонентам посредством коаксиального кабеля. Самые высокие частоты, входящие в СВЧ-диапазон, применяются в коротких двухточечных каналах связи между зданиями; как правило, с этой целью используется диапазон 22 ГГц. Высокие частоты менее пригодны для передачи на большие расстояния, так как подвержены значительному затуханию, но вполне подходят для передачи на короткие расстояния. Кроме того, на более высоких частотах можно обойтись меньшими и более дешевыми антеннами.

## Спутниковая связь с использованием СВЧ

### Физическое описание

Спутник связи является, по сути, ретрансляционной станцией волн СВЧ-диапазона. Он используется для соединения двух или большего числа наземных СВЧ-приемников/передатчиков, называемых наземными станциями. Спутник принимает передачи в одном частотном диапазоне (восходящий канал), усиливает или повторяет этот сигнал и передает его на другой частоте (нисходящий канал). Каждый орбитальный спутник работает с несколькими диапазонами частот, называемыми *транспондерными каналами*, или просто *транспондерами*.

### Применение

Спутниковую связь можно считать таким же важным технологическим прорывом, как и оптоволоконную связь. Перечислим важнейшие сферы применения спутниковой связи.

- Телевещание.
- Международное телефонное сообщение.
- Частные сети.

Широковещательная спутниковая связь великолепно подходит для распространения телевизионных сигналов и с этой целью широко используется в США и во всем мире. Телевизионная сеть, в ее традиционном понимании, предусматривает централизованное распространение программ. Программы передаются на спутник, а затем транслируются на станции, которые распространяют их индивидуальным абонентам. Одна из трансляционных сетей США, PBS (Public Broadcasting Service), распространяет свои телевизионные программы практически исключительно по спутниковым каналам. Другие коммерческие компании также в значительной степени используют спутники, кроме того, в системах кабельного телевидения постоянно растет количество программ, получаемых со спутников. Новейшим применением спутниковой технологии в телевидении являются спутники непосредственного вещания (Direct Broadcast Satellite — DBS), в которых видеосигналы со спутника передаются непосредственно конечному пользователю. Снижение цен и размеров приемных антенн сделало использование спутников непосредственного вещания экономически оправданным, значительное количество каналов, использующих такие спутники, уже эксплуатируется или находится в стадии планирования.

Передача через спутник также используется для создания двухточечных магистралей между телефонными станциями в общедоступных телефонных сетях. Эта среда является оптимальной для загруженных магистралей и соперничает с наземными системами во многих протяженных международных каналах связи.

Наконец, спутники могут использоваться и в коммерческих целях. Поставщик спутника может распределить его пропускную способность на несколько каналов и сдать эти каналы в аренду отдельным коммерческим пользователям. Пользователь, имеющий антенны в определенных городах, создает частную сеть на основе спутникового канала. Такое применение спутников достаточно дорого и доступно только крупным организациям, которым требуется передавать большие объемы информации.

### **Характеристики передачи**

Для спутниковой передачи оптимален диапазон частот от 1 до 10 ГГц. На частотах меньше 1 ГГц имеются значительные помехи от естественных источников, в том числе помехи космического происхождения, солнечные, атмосферные, а также помехи со стороны различных электронных устройств. На частоте выше 10 ГГц велико поглощение сигнала атмосферой и атмосферными осадками.

Большинство современных спутников, предоставляющих услуги двухточечной связи, использует при передаче с земли на спутник (восходящий канал) полосу частот от 5,925 до 6,425 ГГц, а при передаче со спутника на землю (нисходящий канал) полосу частот от 3,7 до 4,2 ГГц. Эта комбинация носит общее название диапазона 4/6 ГГц. Отметим, что частоты, используемые в восходящем и нисходящем каналах, различны. Чтобы обеспечить продолжительную работу без интерференции, частоты приема и передачи должны быть разными. Поэтому сигналы, полученные с наземной станции на одной частоте, должны возвращаться уже на другой частоте.

Диапазон 4/6 ГГц входит в оптимальную зону от 1 до 10 ГГц, но этот диапазон уже переполнен. Вследствие наличия источников помех, действующих на остальных частотах оптимальной зоны (особенно это относится к наземным станциям, использующим СВЧ), использование этих частот невозможно. Поэтому был разработан

диапазон 12/14 ГГц (восходящий канал: 14–14,5 ГГц; нисходящий канал: 11,7–12,2 ГГц). В данном диапазоне частот требуется преодолеть проблемы затухания. Впрочем, можно использовать меньшие и более дешевые наземные приемники. Ввиду возможного заполнения и этого диапазона планируется использовать диапазон 19/29 ГГц (восходящий канал: 27,5–31,0 ГГц; нисходящий канал: 17,7–21,2 ГГц). В указанном диапазоне частот затухание увеличивается, но при этом имеется возможность использовать более широкую полосу (2500 МГц, а не 500 МГц), а также меньшие и более дешевые приемники.

Следует отметить некоторые особенности спутниковой связи. Во-первых, поскольку задействованы большие расстояния, то при передаче от одной наземной станции к другой возникает задержка распространения, составляющая около четверти секунды. Эта задержка заметна даже в обычных телефонных разговорах. Она также порождает проблемы в сферах защиты от ошибок и управления потоком данных, которые будут рассматриваться в последующих главах. Во-вторых, спутниковая связь с использованием СВЧ обеспечивает возможность широко вещания. Многие станции могут передавать на спутник, и передача со спутника может приниматься многими станциями.

## Широковещательное радио

### Физическое описание

Принципиальным различием между широковещательным радио и передачей с использованием СВЧ является то, что первое является ненаправленным, а последняя — направленной. Следовательно, широковещательное радио обходится без параболических антенн, а используемые антенны не требуются точно ориентировать.

### Применение

Термин *радио* используется для обозначения частот, входящих в диапазон от 3 кГц до 300 ГГц. При использовании нами неофициального термина *широковещательное радио* имеются в виду частоты метрового диапазона и часть диапазона ультравысоких частот: от 30 МГц до 1 ГГц. На этих частотах работают FM-радио, а также телевидение метрового и дециметрового диапазонов. Данный диапазон также используется в некоторых приложениях сетей передачи данных.

### Характеристики передачи

Для широковещательной связи эффективным является диапазон частот от 30 МГц до 1 ГГц. В отличие от распространения электромагнитных волн низкой частоты, ионосфера прозрачна для радиоволн с частотой выше 30 МГц. Поэтому передача на этих волнах ограничена расстоянием прямой видимости, и удаленные передатчики не будут интерферировать с другими из-за отражения сигнала от атмосферы. Отметим также, что в отличие от СВЧ волны широковещательного радио менее чувствительны к поглощению дождем.

Как и при СВЧ, зависимость поглощения от расстояния для радиоволн выражается формулой (2.1) и составляет  $10 \lg \left( \frac{4\pi d}{\lambda} \right)^2$ . Поскольку длина волны ра-

диапазона больше, чем в предыдущем случае, то, соответственно, затухание будет меньше.

Основным источником искажений при широковещательной связи является многолучевая интерференция. Вследствие отражения от поверхности земли, воды, а также естественных или искусственных объектов создаются разные пути прохождения сигнала между антеннами. Этот эффект часто наблюдается при появлении на экране телевизора множественных изображений, вызванных прохождением самолета.

## Связь в инфракрасном диапазоне

Для связи в инфракрасном диапазоне используются приемники/передатчики (трансиверы), модулирующие некогерентный инфракрасный свет. Трансиверы либо должны сами находиться в пределах прямой видимости, либо должно использоваться отражение от светлоокрашенной поверхности, например потолка комнаты.

Одним из важных различий связи в инфракрасном диапазоне и СВЧ-диапазоне является то, что лучи, используемые в первом случае, не проникают сквозь стены. Следовательно, интерференция и проблемы конфиденциальности, имеющиеся в СВЧ-системах, в данном случае отсутствуют. Более того, с инфракрасным диапазоном не связаны проблемы, касающиеся выделения частот, так как для использования этого диапазона лицензии не требуется.

## 2.5. УПЛОТНЕНИЕ

Как при локальной, так и при дальней связи почти всегда оказывается, что пропускная способность передающей среды больше, чем необходимо для передачи отдельного сигнала. Для повышения эффективности системы среде следовало бы передавать несколько сигналов. Такое совместное использование передающей среды называется *уплотнением*.

На рис. 2.11 показана функция уплотнения в ее простейшей форме. Мультиплексор (устройство уплотнения) имеет  $n$  входов. Далее мультиплексор соединяется одним каналом связи с демультиплексором (устройством разуплотнения). Канал связи может содержать  $n$  отдельных информационных каналов. Мультиплексор объединяет (уплотняет) данные из  $n$  входных линий и передает их по каналу связи высокой емкости. Демультиплексор принимает уплотненный поток данных, распределяет (разуплотняет) данные по информационным каналам и доставляет их на соответствующие выходные линии.

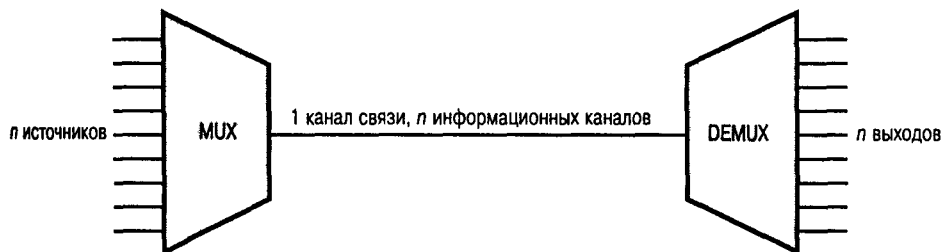


Рис. 2.11. Уплотнение

Широкое использование уплотнения в технологиях передачи данных можно объяснить наличием следующих преимуществ.

1. Более высокая скорость передачи данных, более рентабельное использование средства связи. Иначе говоря, для данной системы и на данном расстоянии стоимость передачи 1 Кбит/с снижается с увеличением скорости передачи средства связи. Соответственно, снижается стоимость аппаратуры передачи-приема с увеличением скорости передачи данных.
2. Большинству отдельных устройств передачи данных требуется поддержание относительно невысокой скорости передачи. Например, для многих терминальных и компьютерных приложений обычно более чем достаточно скорости передачи данных 64 Кбит/с.

Вышеизложенное сформулировано для устройств передачи цифровой информации, но приведенные утверждения относятся и к средствам передачи речи. Иначе говоря, чем выше пропускная способность средства связи (для телефонных каналов), тем ниже стоимость отдельного канала, а пропускная способность, необходимая отдельному телефонному каналу, незначительна.

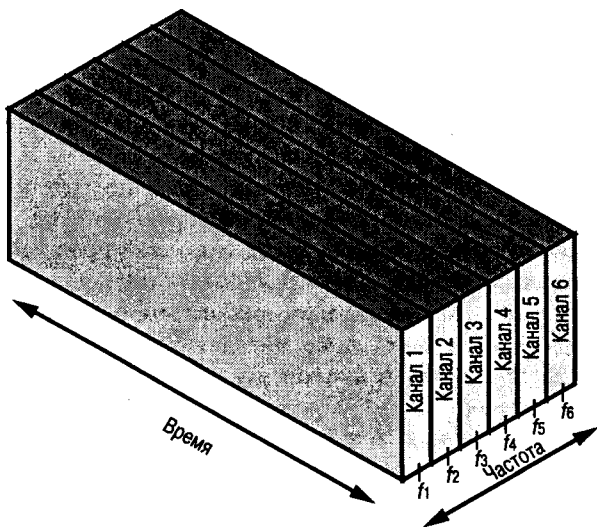
В сфере телекоммуникации широко распространены два метода уплотнения: уплотнение с частотным разделением (Frequency-Division Multiplexing — FDM) и уплотнение с временным разделением (Time-Division Multiplexing — TDM).

Уплотнение с частотным разделением, или просто частотное уплотнение, возможно в том случае, если полезная ширина полосы пропускания передающей среды превышает ширину полосы данного сигнала. Если все сигналы будут модулированы разными несущими частотами, достаточно удаленными друг от друга, так, чтобы полосы сигналов не перекрывались, можно будет передавать несколько сигналов одновременно. Простой пример такой передачи показан на рис. 2.12, а. Сигналы от шести различных источников подаются в мультиплексор, где они модулируются на шести различных несущих частотах ( $f_1, \dots, f_6$ ). Каждый модулированный сигнал требует определенной полосы, centered на несущей частоте (такая полоса называется каналом). Для предотвращения интерференции каналы разделяются защитными полосами частот — неиспользуемыми участками спектра. На рисунке эти полосы не показаны.

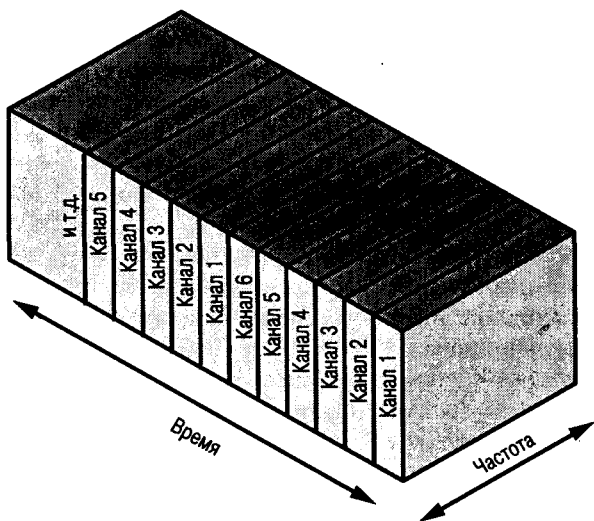
В качестве примера можно привести уплотнение речевых сигналов для телефонной связи. Мы уже указывали, что полезный спектр речи — от 300 до 3400 Гц. Таким образом, в полосе шириной 4 кГц помещается речевой сигнал и достаточная защитная полоса. Как в Северной Америке (стандарт Bell Systems), так и во всем мире (стандарт ITU-T) стандартная схема уплотнения речевых сигналов — двенадцать телефонных каналов по 4 кГц каждый, занимающих частоты от 60 до 108 кГц. Для каналов связи более высокой емкости обоими стандартами определяются дальнейшие уровни группирования каналов по 4 кГц.

Уплотнение с временным разделением возможно в том случае, если доступная скорость передачи данных (иногда ошибочно называемая шириной полосы) передающей среды превышает требуемую скорость передачи цифровых сигналов. Чередуя фрагменты отдельных сигналов во времени, можно передавать по одному каналу много цифровых сигналов. Чередование может быть побитовым, или же чередоваться могут целые блоки байтов. Например, мультиплексор, изображенный на рис. 2.12, б, имеет шесть входов, на каждый из которых подается информация со скоростью, скажем, 9,6 Кбит/с.

В этом случае для передачи данных от всех шести источников необходим канал с пропускной способностью не менее 57,6 Кбит/с. Как и при частотном уплотнении, последовательность временных интервалов, выделенная для отдельного источника, называется каналом. Цикл временных интервалов (по одному на каждый источник) называется кадром.



а) Уплотнение с частотным разделением



б) Уплотнение с временным разделением

Рис. 2.12. Частотное и временное уплотнение

Схема временного уплотнения, показанная на рис. 2.12, б, называется также синхронным временным уплотнением, поскольку временные интервалы установлены заранее и фиксированы. Асинхронное временное уплотнение, напротив, позволяет динамически изменять распределение времени между разными источниками. Далее, если явно не указано иное, под временным уплотнением мы будем подразумевать именно синхронное временное уплотнение.

Общее описание системы с синхронным временным уплотнением приведено на рис. 2.13. Требуется уплотнить в одной передающей среде набор сигналов  $[m_i(t), i = 1, n]$ . Эти сигналы — носители цифровой информации, и, в общем случае, они являются цифровыми. Данные, поступающие из каждого источника, временно заносятся в буфер. Размер каждого буфера равен, как правило, одному биту или одному символу. Буфера последовательно просматриваются, и из их содержимого формируется составной цифровой поток данных  $m_c(t)$ . Данная операция производится достаточно быстро, чтобы каждый буфер мог очиститься перед поступлением последующих данных. Следовательно, скорость передачи данных  $m_c(t)$  должна быть не меньше суммы скоростей передачи источников  $m_i(t)$ . Цифровой сигнал  $m_c(t)$  может быть либо передан в исходном виде, либо пропущен через модем и преобразован в аналоговую форму. В любом случае передача обычно является синхронной.

Формат переданных данных может быть подобен формату, представленному на рис. 2.13, б. Данные сгруппированы в кадры, каждый из которых содержит цикл временных интервалов, или слотов. В каждом кадре каждому источнику данных выделяется один или несколько слотов. Последовательность слотов различных кадров, относящихся к одному источнику, называется каналом. Размер слота равен размеру буфера передатчика и составляет, как правило, один бит или байт (символ).

При наличии асинхронных источников можно использовать метод чередования байтов. Каждый временной интервал содержит один информационный символ. Обычно стартовый и стоповый биты каждого символа удаляются перед передачей и повторно вводятся в приемнике (это делается для повышения эффективности). Технология чередования битов используется как с синхронными, так и с асинхронными источниками. В этом случае каждый временной интервал содержит один бит.

В приемнике чередующиеся данные разуплотняются и направляются в соответствующий буфер назначения. Каждому входному источнику  $m_i(t)$  соответствует собственный выходной источник, принимающий данные с той же скоростью, с которой они генерировались.

Синхронное временное уплотнение называется синхронным не потому, что используется синхронная передача, а потому, что распределение временных интервалов по источникам является заранее установленным и фиксированным. Временные интервалы, соответствующие каждому источнику, передаются вне зависимости от того, есть ли у этого источника данные для передачи. Разумеется, это происходит и при использовании частотного уплотнения. В обоих случаях пропускная способность канала связи используется не полностью, зато мы выигрываем с точки зрения простоты реализации. Впрочем, даже при фиксированном распределении слотов устройства с синхронным временным уплотнением могут работать с источниками, имеющими разные скорости передачи данных.

При этом наиболее медленному устройству ввода может выделяться один слот в цикле, а более быстрые устройства будут использовать несколько слотов.

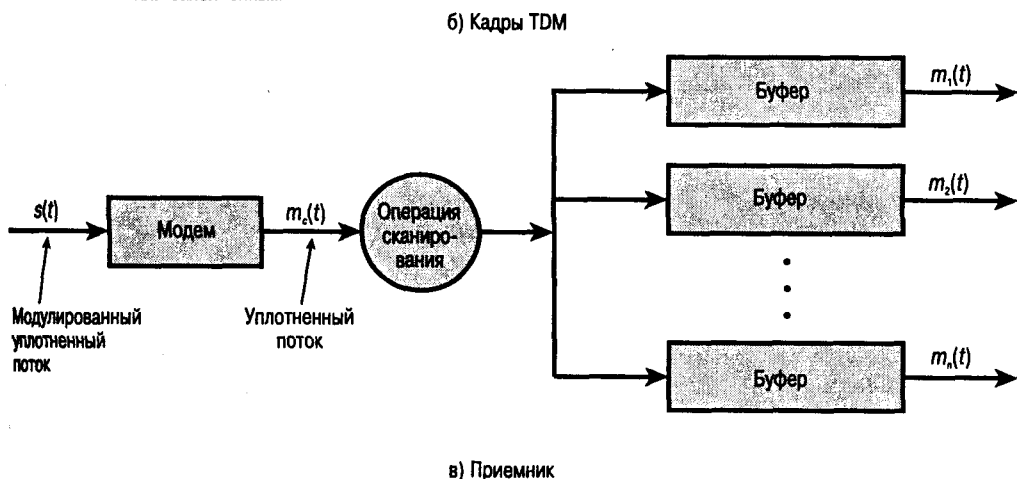
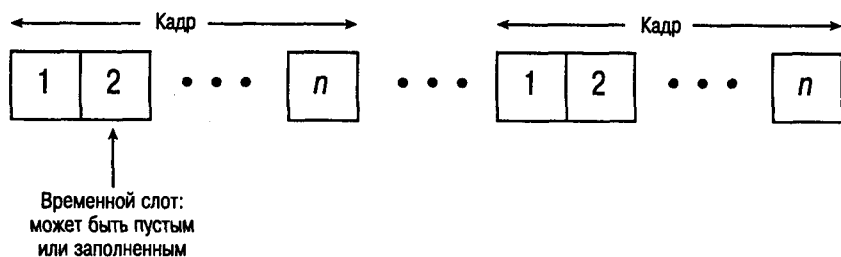
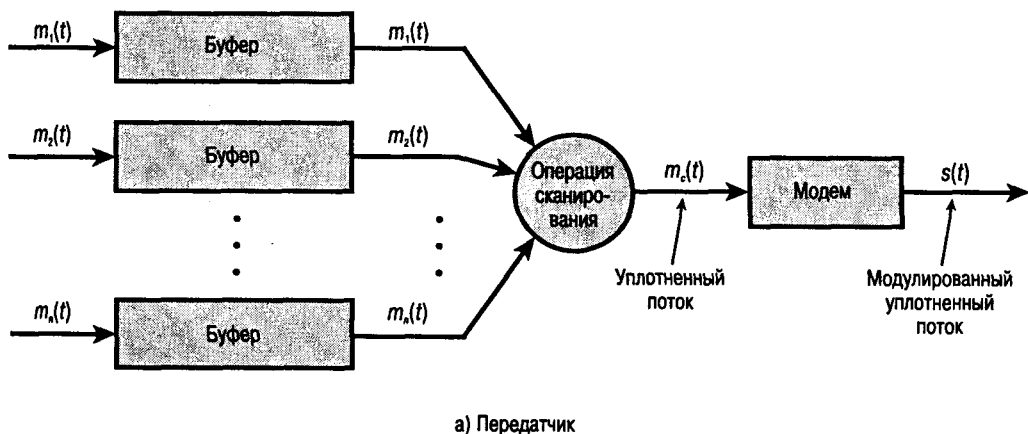


Рис. 2.13. Система с синхронным временным уплотнением

Один из примеров систем с временным уплотнением — стандартная схема передачи речевой информации при импульсно-кодовой модуляции, называемая в терминологии АТ&Т службой T1. Данные принимаются из каждого источника по одной выборке (7 бит) за раз. В службе T1 уплотняются 24 источника, поэто-



му каждый кадр включает  $8 \times 24 = 192$  бит данных и сигналов управления. Еще один заключительный бит добавляется для поддержания синхронизации. Таким образом, кадр состоит из 193 бит и содержит семибитовые выборки от каждого источника. Поскольку выборки производятся 8000 раз в секунду, требуемая скорость передачи данных составляет  $8000 \times 193 = 1,544$  Мбит/с. Как и при частотном уплотнении, для следующих уровней группирования установлены более высокие скорости передачи данных.

Возможности временного уплотнения не ограничиваются цифровыми сигналами. Чередование во времени можно применять и для аналоговых сигналов. Кроме того, в случае использования аналоговых сигналов возможно сочетание временного и частотного уплотнения. Полоса пропускания передающей системы разделяется по частоте на некоторое число каналов, в каждом из которых используется временное уплотнение.

## 2.6. РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

Все темы данной главы более подробно рассматриваются в [STAL00]. Удачное изложение этого материала также содержится в [FREE99]. Всестороннее рассмотрение вопросов как аналоговой, так и цифровой связи представлено в [COUC01].

COUC01 Couch L. *Digital and Analog Communication Systems*. — Upper Saddle River, NJ: Prentice Hall, 2001.

FREE99 Freeman R. *Fundamentals of Telecommunications*. — New York: Wiley, 1999.

STAL00 Stallings W. *Data and Computer Communications, Sixth Edition*. — Upper Saddle River, NJ: Prentice Hall, 2000 // Столлингс В. *Компьютерные системы передачи данных*, 6-е изд. — М.: Издательский дом "Вильямс", 2002.

## 2.7. ТЕРМИНЫ, ВОПРОСЫ И ЗАДАЧИ

### Основные термины

аналоговая передача	направляемые среды	уплотнение
аналоговые данные	ненаправляемые среды	фаза
аналоговый сигнал	основная частота	цифровая передача
апериодический	период	цифровой сигнал
беспроводная передача	периодический	цифровые данные
временное	пропускная способность	частота
представление	канала	частотное представление
временное уплотнение (TDM)	радио	частотное уплотнение (FDM)
децибел (дБ)	СВЧ	ширина полосы
длина волны	синхронное временное	широковещательное
инфракрасный диапазон	уплотнение	радио
максимальная	спектр	шум
амплитуда	спутниковая	
наземная СВЧ-связь	СВЧ-связь	
	средства связи	

## Вопросы

1. Укажите различия между аналоговыми и цифровыми электромагнитными сигналами.
2. Каковы три важнейшие характеристики периодического сигнала?
3. Сколько радиан составляет полная окружность (360 градусов)?
4. Каково соотношение между длиной и частотой синусоидальной волны?
5. Каково соотношение между спектром сигнала и шириной его полосы?
6. Что такое “затухание”?
7. Дайте определение пропускной способности канала.
8. Какие основные факторы влияют на пропускную способность канала?
9. Укажите различия между направляемыми и ненаправляемыми средами.
10. Каковы главные преимущества и недостатки передачи в диапазоне СВЧ?
11. Что такое “спутник непосредственного вещания”?
12. Почему спутник должен иметь разные частоты восходящего и нисходящего каналов?
13. Укажите существенные различия между широкополосным радио и СВЧ.
14. Почему уплотнение так выгодно?
15. Как избежать интерференции при частотном уплотнении?
16. Объясните, как работает синхронное временное уплотнение.

## Задачи

1. Сигнал имеет основную частоту 1000 Гц. Каков его период?
2. Выразите в наиболее простой форме следующие сигналы.
  - а.  $\sin(2\pi ft - \pi) + \sin(2\pi ft + \pi)$
  - б.  $\sin 2\pi ft + \sin(2\pi ft - \pi)$
3. Звук можно промоделировать синусоидальной функцией. Сравните относительные частоты и длины волн музыкальных нот. Скорость звука примите равной 330 м/с, а частоты звуков следующими:

Нота	C (до)	D (ре)	E (ми)	F (фа)	G ( соль)	A (ля)	B (си)	C (до)
Частота	264	297	330	352	396	440	495	582
4. Определите, какой функцией задается штриховая кривая на рис. 2.14, если сплошная кривая задается функцией  $\sin(2\pi t)$ . Иными словами, если штриховую кривую можно описать функцией  $A \sin(2\pi ft + \phi)$ , каковы значения  $A$ ,  $f$  и  $\phi$ ?

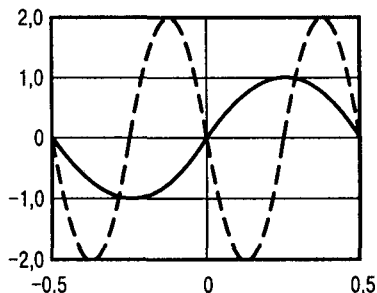


Рис. 2.14. Иллюстрация к задаче 2.4

5. Разложите сигнал  $(1 + 0,1 \cos 5t) \cos 100t$  на линейную комбинацию гармонических функций и найдите амплитуду, частоту и фазу каждой компоненты. (Подсказка: используйте тождество для  $\cos a \cos b$ .)
6. На рис. 2.5 показан эффект удаления из прямоугольного сигнала высших гармоник, когда осталось лишь несколько низших гармоник. Какой вид будет иметь сигнал в противоположном случае (т.е. если будут оставлены высшие гармоники, а удалены несколько низших гармоник)?
7. Какова пропускная способность канала телетайпа с шириной полосы 300 Гц и отношением сигнал/шум 3 дБ?
8. Цифровая система передачи сигналов должна работать со скоростью 9600 бит/с.
  - а. Какова минимальная необходимая ширина полосы канала, если сигнальная посылка кодируется четырехбитовым словом?
  - б. Рассмотрите предыдущую задачу для восьмибитовых слов.
9. Изучите работы Шеннона и Найквиста о пропускной способности каналов. Каждая определяет ограничение скорости передачи данных по каналу, но базируются они на разных подходах. Определите связь между ними.
10. Определите требуемое отношение сигнал/шум для достижения в канале пропускной способности 20 Мбит/с при ширине полосы 3 МГц.
11. Покажите, что увеличение частоты передачи или расстояния между передающей и приемной антеннами в два раза вызывает уменьшение принимаемой мощности на 6 дБ.
12. Заполните пустые клеточки следующей таблицы, содержащей приблизительные значения мощностей для различных уровней в децибелах.

Децибелы	1	2	3	4	5	6	7	8	9	10
Ослабление			0,5							0,1
Усиление			2							10

13. Усилитель имеет коэффициент усиления по напряжению 30 дБ. Какому отношению напряжений соответствует этот коэффициент?
14. Мощность сигнала на выходе усилителя составляет 20 Вт. Чему равен выход в децибел-ваттах?

## ПРИЛОЖЕНИЕ 2А. ДЕЦИБЕЛЫ И ИНТЕНСИВНОСТЬ СИГНАЛА

В каждой передающей системе очень важен такой параметр, как интенсивность сигнала. При распространении сигнала в передающей среде происходит ослабление его интенсивности, или *затухание*. Для компенсации таких потерь в различных точках среды могут вводиться усилители, сообщающие сигналу дополнительную интенсивность.

Усиление и снижение интенсивности сигнала, а также его относительные уровни обычно выражаются в децибелах. Ниже перечислены причины такого выбора.

- Интенсивность сигнала часто снижается по логарифмическому закону, поэтому ослабление проще выражать в децибелах, являющихся логарифмическими единицами.
- Суммарное усиление или ослабление сигнала в каскадном канале передачи можно вычислить с помощью простых операций сложения и вычитания.

Итак, *децибелом* (дБ) называется мера отношения между двумя уровнями сигнала:

$$G_{\text{дБ}} = 10 \lg \frac{P_{\text{вых}}}{P_{\text{вх}}},$$

где

$G_{\text{дБ}}$  — усиление в децибелах;

$P_{\text{вх}}$  — мощность на входе;

$P_{\text{вых}}$  — мощность на выходе;

$\lg$  — логарифм по основанию 10.

Связь между значениями в децибелах и степенями числа 10 показана в табл. 2.3.

**Таблица 2.3. Значения в децибелах**

Отношение мощностей	дБ	Отношение мощностей	дБ
$10^1$	10	$10^{-1}$	-10
$10^2$	20	$10^{-2}$	-20
$10^3$	30	$10^{-3}$	-30
$10^4$	40	$10^{-4}$	-40
$10^5$	50	$10^{-5}$	-50
$10^6$	60	$10^{-6}$	-60

В литературе наблюдается некоторая несогласованность в употреблении терминов “усиление” и “ослабление”. Если значение  $G_{\text{дБ}}$  положительно, оно означает на практике увеличение мощности. Например, усиление 3 дБ означает, что мощность сигнала удвоилась. Если же значение  $G_{\text{дБ}}$  отрицательно, оно означает на практике уменьшение мощности. Например, усиление -3 дБ означает, что мощность уменьшилась вдвое, т.е. произошло ослабление. Обычно при этом говорят, что имеется ослабление 3 дБ. Однако, согласно некоторым источникам, эта величина должна называться усилением -3 дБ. Более правильно будет сказать, что отрицательное усиление соответствует положительному ослаблению. Таким образом, ослабление в децибелах определяется как

$$L_{\text{дБ}} = -10 \lg \frac{P_{\text{вых}}}{P_{\text{вх}}} = 10 \lg \frac{P_{\text{вх}}}{P_{\text{вых}}}.$$

**Пример.** Если на входе линии передачи уровень мощности сигнала составляет 10 мВт, а на некотором расстоянии — 5 мВт, то ослабление сигнала можно выразить следующим образом:

$$L_{\text{дБ}} = 10 \lg(10/5) = 10 (0,3) = 3 \text{ дБ}.$$

Отметим, что в децибелах выражается относительное, а не абсолютное отличие. Ослабление сигнала с 1000 мВт до 500 мВт также является ослаблением на 3 дБ.

Децибелы используются и для измерения отношения напряжений, учитывая тот факт, что мощность пропорциональна квадрату напряжения:

$$P = \frac{V^2}{R},$$

где

$P$  — мощность, рассеиваемая на сопротивлении  $R$ ;

$V$  — напряжение на сопротивлении  $R$ .

Следовательно,

$$L_{\text{дБ}} = 10 \lg \frac{P_{\text{вх}}}{P_{\text{вых}}} = 10 \lg \frac{V_{\text{вх}}^2 / R}{V_{\text{вых}}^2 / R} = 20 \lg \frac{V_{\text{вх}}}{V_{\text{вых}}}.$$

**Пример.** Использование децибелов полезно при определении усиления или снижения мощности, происходящего на последовательности передающих элементов. Рассмотрим, например, последовательность элементов, на вход которой подается мощность 4 мВт, первый элемент является линией передачи с ослаблением 12 дБ (усилением  $-12$  дБ), второй элемент — это усилитель с усилением 35 дБ, а третий — линия передачи с ослаблением 10 дБ. Суммарное усиление равно  $(-12 + 35 - 10) = 13$  дБ. Вычислим мощность на выходе:

$$G_{\text{дБ}} = 13 = 10 \lg (P_{\text{вых}}/4 \text{ мВт}),$$

$$P_{\text{вых}} = 4 \times 10^{1.3} \text{ мВт} = 79,8 \text{ мВт}.$$

Значения в децибелах связаны с относительными амплитудами или изменениями амплитуд, но никак не с абсолютными уровнями. Было бы удобно представить абсолютный уровень мощности также в децибелах, чтобы можно было легко вычислять усиление или снижение мощности по отношению к исходному сигналу. Единица дБВт (децибел-ватт) широко используется в приложениях СВЧ-связи. В качестве эталонного уровня выбрана величина 1 Вт, и ей присвоено значение 0 дБВт. Абсолютный уровень мощности в децибел-ваттах определяется следующим образом:

$$\text{мощность, дБВт} = 10 \lg \frac{\text{мощность, Вт}}{1 \text{ Вт}}.$$

**Пример.** Мощности 1000 Вт соответствует 30 дБВт, а мощности 1 мВт соответствует  $-30$  дБВт.

Широко используется другая производная единица — дБмВт (децибел-милливатт). В этом случае за эталонный уровень мощности принимается 1 мВт, т.е. 0 дБмВт соответствует 1 мВт. Следовательно,

$$\text{мощность, дБмВт} = 10 \lg \frac{\text{мощность, мВт}}{1 \text{ мВт}}.$$

Полезно запомнить следующие соотношения:

$$+30 \text{ дБмВт} = 0 \text{ дБВт},$$

$$0 \text{ дБмВт} = -30 \text{ дБВт}.$$

# ГЛАВА 3

## СЕТИ СВЯЗИ

### 3.1. Локальные, городские и глобальные сети

Глобальные сети

Локальные сети

Городские сети

### 3.2. Методы коммутации

### 3.3. Коммутация каналов

### 3.4. Коммутация пакетов

Использование коммутации пакетов

Размер пакета

### 3.5. АТМ

Логические соединения АТМ

Ячейки АТМ

Категории услуг АТМ

### 3.6. Рекомендуемая литература и Web-сайты

### 3.7. Термины, вопросы и задачи

**В** данной главе представлен обзор различных подходов к организации сетей связи. Начинается глава с обзора типов сетей, основанных на различном географическом охвате. Затем исследуются сети с коммутацией каналов, сети с коммутацией пакетов и сети АТМ.

### 3.1. ЛОКАЛЬНЫЕ, ГОРОДСКИЕ И ГЛОБАЛЬНЫЕ СЕТИ

Локальные сети (local area network — LAN), городские сети (metropolitan area network — MAN) и глобальные сети (wide area network — WAN) представляют собой примеры сетей связи. Все данные категории, а также несколько частных случаев представлены на рис. 3.1. Для сравнения представлен типичный диапазон параметров многопроцессорного компьютера.

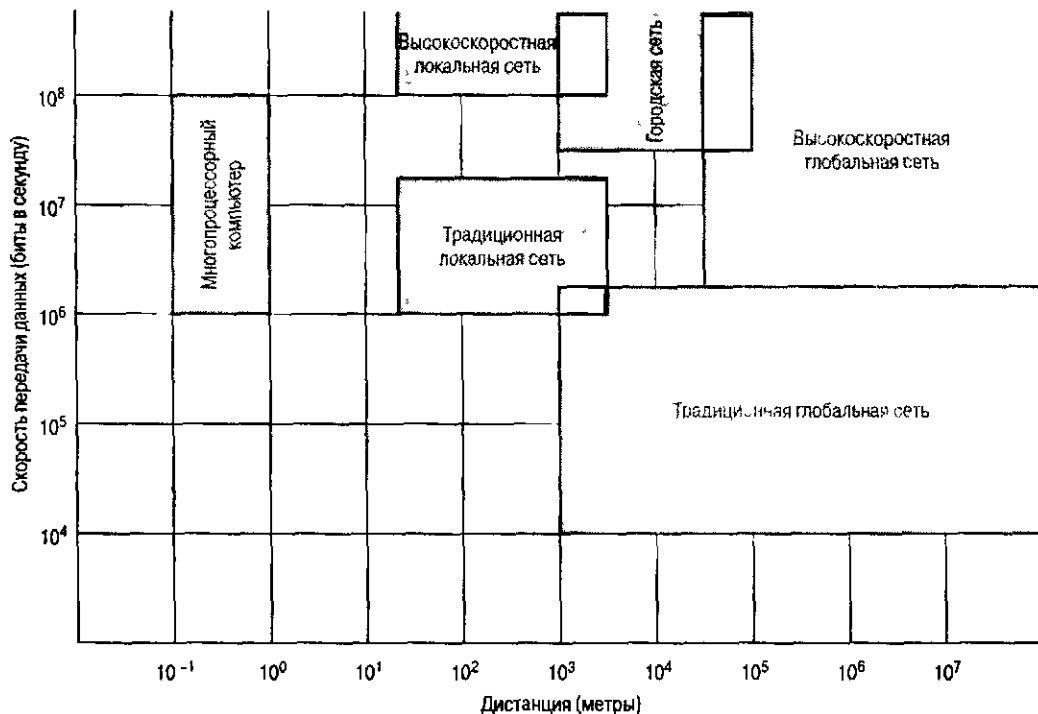


Рис. 3.1. Сравнение многопроцессорных систем, локальных, городских и глобальных сетей

#### Глобальные сети

Глобальные сети, как правило, охватывают большие географические области, требуют возможности пересечения границ общедоступных каналов и зависят, по крайней мере частично, от этих каналов. Обычно глобальная сеть состоит из некоторого числа взаимосвязанных коммутирующих узлов. Сообщение, передаваемое

мое любимым устройством, определенным путем проходит через внутренние узлы к заданному адресату.

Традиционно глобальные сети предоставляют абонентам относительно скромные возможности. Обычным является подключение к сети передачи данных либо телефонной сети посредством модема со скоростью передачи порядка 64 000 бит/с или меньше. Коммерческие абоненты могут получить более высокие скорости передачи, используя услугу T1; здесь обычной является уже скорость 1,544 Мбит/с.

Продолжающаяся разработка средств оптоволоконной связи привела к стандартизации еще более высоких скоростей передачи в глобальных сетях, в настоящее время такие услуги получили широкое распространение. Данные сети предлагают пользователям скорости передачи данных порядка десятков и сотен мегабитов в секунду, при этом используется технология, известная как АТМ (asynchronous transfer mode — асинхронный режим передачи).

## Локальные сети

Локальная сеть, как и глобальная, представляет собой сеть связи, соединяющую ряд устройств и предоставляющую средства для обмена информацией между этими устройствами. Ниже перечислены некоторые основные различия между локальными и глобальными сетями.

1. Область действия локальной сети невелика; как правило, это отдельное здание или группа зданий. Как мы увидим, различие в географических областях действия приводит к различным техническим решениям.
2. Обычно локальная сеть принадлежит той организации, которая владеет подключенными к ней устройствами. Для глобальных сетей такое бывает значительно реже — по меньшей мере значительная часть сетевого имущества не имеет владельца. Из сказанного вытекают два следствия. Во-первых, при выборе локальной сети нужно быть предельно внимательным, поскольку могут потребоваться существенные затраты (по сравнению с расходами на телефонный или выделенный канал для глобальных сетей) как на приобретение, так и на обслуживание. Во-вторых, ответственность за управление локальной сетью лежит исключительно на пользователе.
3. Внутренняя скорость передачи данных в локальных сетях, как правило, значительно выше, чем в глобальных.

Простой пример локальной сети, в котором представлены некоторые из ее характеристик, показан на рис. 3.2. Все устройства подключены к совместно используемой среде передачи. Передача от любого устройства сети может приниматься всеми остальными устройствами, присоединенными к этой сети.

Традиционные локальные сети обеспечивают передачу данных со скоростью 1–20 Мбит/с. Такой скорости, хотя и являющейся значительной, становится недостаточно при существующем темпе роста числа устройств, мультимедийных приложений и увеличении использования архитектуры клиент/сервер. Таким образом, в последнее время велась разработка высокоскоростных локальных сетей, имеющих скорость от 100 Мбит/с до 10 Гбит/с.





Рис. 3.2. Простая локальная сеть

## Городские сети

Городские сети (Metropolitan Area Network — MAN) — это промежуточное звено между локальными и глобальными сетями. Интерес к городским сетям объясняется тем, что возможностей традиционных двухточечных сетей и сетей с коммутацией бывает недостаточно для удовлетворения растущих потребностей организаций. Технология АТМ обещает удовлетворить многие запросы в отношении высоких скоростей, но кроме этого существует требование к наличию частных и общедоступных сетей, дающих большую (и дешевую) пропускную способность на значительных расстояниях. Подход с высокоскоростным совместным использованием сети (как он определен в стандартах локальных сетей) в масштабе города открывает большие возможности. Как видно на рис. 3.1, городские сети используются на больших площадях и с более высокими скоростями, чем локальные сети, хотя с точки зрения географического охвата строго разделить эти сети затруднительно.

Основным рынком для городских сетей являются заказчики с большими требованиями к пропускной способности в городских масштабах. Городская сеть предназначена для удовлетворения этих требований за меньшую цену и при большей эффективности, чем могут предложить локальные телефонные компании.

## 3.2. МЕТОДЫ КОММУТАЦИИ

При передаче данных за пределами локальной области связь, как правило, осуществляется путем переноса данных от источника к адресату через сеть промежуточных узлов коммутации; на этом принципе коммутируемой сети иногда также реализуются локальные и городские сети. Узлы коммутации не меняют информационного содержимого передачи; их назначение — просто перемещать данные от узла к узлу, пока не будет достигнут адресат. Конечные устройства, которые должны быть соединены, можно назвать *станциями*. Станции могут быть компьютерами, терминалами, телефонами или другими устройствами. Коммутирующие устройства, служащие для обеспечения связи, мы будем называть *узлами*. Узлы связаны с друг другом в некоторую топологию посредством каналов передачи. Каждая станция подключена к узлу, а совокупность узлов называется *сетью связи*.

На рис. 3.3 изображена простая сеть. Данные, входящие в сеть от станции, направляются адресату путем коммутации от узла к узлу. Например, данные от станции А, предназначенные для станции F, передаются на узел 4. Затем они могут быть направлены адресату через узлы 5 и 6 или 7 и 6. Необходимо учесть следующее.

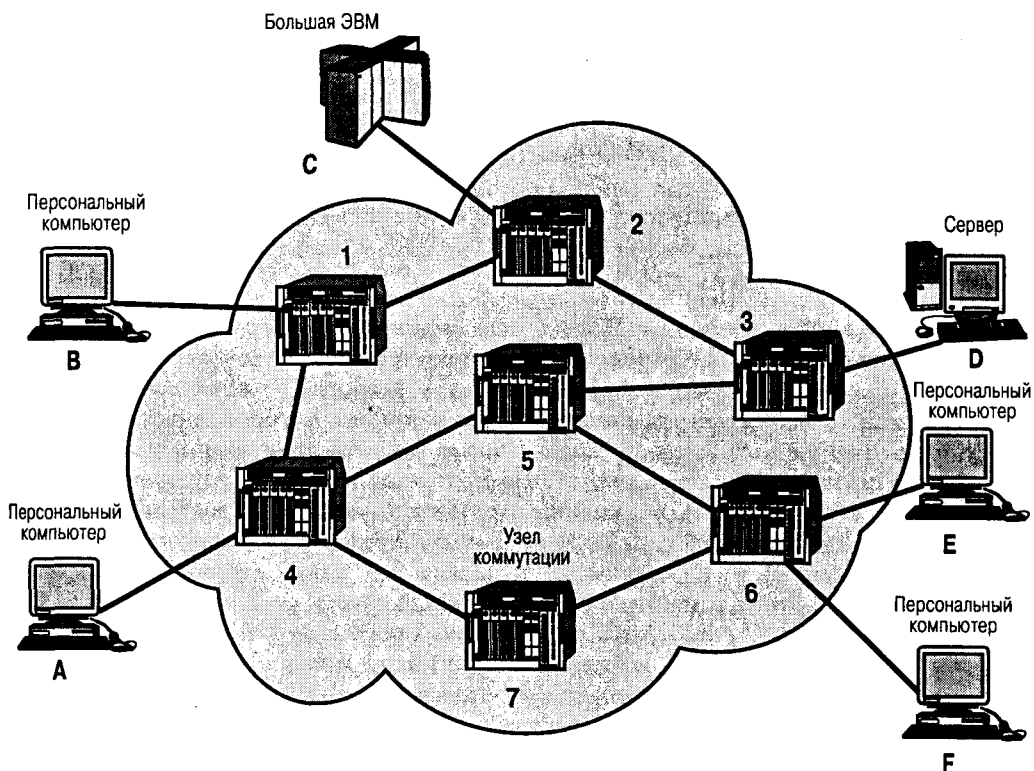


Рис. 3.3. Простая коммутируемая сеть

1. Некоторые узлы соединяются только с другими узлами (например, 5 и 7). Их единственная задача — внутренняя (по отношению к сети) коммутация данных. К другим узлам подключены одна или более станций; такие узлы, в дополнение к своей коммутирующей функции, принимают и передают данные подключенным станциям.
2. Каналы между узлами обычно уплотняются с использованием либо частотного (Frequency Division Multiplexing — FDM), либо временного (Time Division Multiplexing — TDM) уплотнения.
3. Обычно сеть не вполне связана, т.е. прямые каналы соединяют не все возможные пары узлов. В то же время желательно иметь между каждыми двумя станциями несколько возможных маршрутов через сеть. Это повышает надежность сети.

В глобальных коммутируемых сетях применяются две совершенно различные технологии: коммутация каналов и коммутация пакетов. Они отличаются способом передачи информации узлами от одного канала к другому на пути от источника к адресату. В этой главе мы рассмотрим особенности обеих технологий, после чего перейдем к важному частному случаю коммутации пакетов, известному как поэтапная передача, или АТМ.

### 3.3. КОММУТАЦИЯ КАНАЛОВ

Коммутация каналов является доминирующей технологией передачи речи и данных и останется таковой в обозримом будущем. Связь с коммутацией каналов предполагает, что имеется заранее определенный тракт связи между двумя станциями. Этот тракт — связанная последовательность каналов между узлами сети. В каждом физическом канале для данного соединения выделяется логический канал. Связь с коммутацией каналов включает три фазы, которые можно рассмотреть на примере на рис. 3.3.

- 1. Установление соединения.** Прежде чем можно будет передать какие-то сигналы, должно быть установлено сквозное соединение (от станции к станции). Например, станция А посылает узлу 4 запрос, требуя соединения со станцией Е. Как правило, канал от А до 4 является выделенным, т.е. часть соединения уже существует. Узел 4 должен найти следующий участок маршрута по направлению к узлу 6. На основе данных маршрутизации и сведений о доступности и, возможно, стоимости узел 4 выбирает канал к узлу 5, занимает в нем свободный логический канал (с использованием частотного или временного уплотнения) и посылает сообщение с требованием соединения с Е. Теперь у нас создан выделенный тракт от А через 4 до 5. Поскольку к узлу 4 может быть подключено много станций, этот узел должен быть способен создавать внутренние тракты от многих станций ко многим узлам. Остальная часть процесса выполняется подобным образом. Узел 5 выделяет канал к узлу 6 и внутри сети подключает этот канал к каналу от узла 4. Узел 6 завершает соединение, устанавливая канал к Е. В заключение определяется, занята станция Е или готова к приему соединения.
- 2. Передача данных.** Теперь можно передать по сети информацию от А к Е. Данные могут быть аналоговыми или цифровыми, в зависимости от природы сети. Поскольку развитие средств связи происходит в направлении к полностью интегрированным цифровым сетям, господствующим методом как для голоса, так и для данных становится цифровая (двоичная) передача. Тракт связи образуют: канал А-4, внутренняя коммутация в узле 4; канал 4-5, внутренняя коммутация в 5; канал 5-6, внутренняя коммутация в 6; канал 6-Е. Как правило, соединение является дуплексным.
- 3. Разрыв соединения.** После некоторого периода передачи данных соединение завершается, обычно это действие инициирует одна из двух станций. Узлам 4, 5 и 6 должны быть переданы сигналы на освобождение выделенных ресурсов.

Заметьте, что тракт соединения устанавливается до того, как начинается передача данных. Следовательно, канал между каждыми двумя узлами должен иметь резерв пропускной способности, а каждый узел должен иметь свободную коммутационную способность, чтобы обслужить требуемое соединение. Коммутаторы должны уметь самостоятельно выделять эти ресурсы и определять маршрут через сеть.

Коммутация каналов может быть довольно неэффективной. Ресурсы каналов выделяются на весь срок действия соединения, даже если данные фактически не передаются. При передаче голосовых данных степень использования может быть довольно высокой, но все же не достигает 100%. При соединении терминала и компь-

ютера канал большую часть времени может оставаться неиспользуемым. На степень использования влияет задержка начала передачи сигналов, требуемая для установления связи. Но как только соединение будет установлено, наличие сети становится практически незаметным для пользователей. Информация передается с постоянной скоростью без задержек, исключая лишь задержку при распространении по каналам. Задержка на каждом узле незначительна.

Метод коммутации каналов был создан для обслуживания голосового обмена, но теперь он используется и при обмене данными. Наиболее известный пример сети с коммутацией каналов — телефонная сеть общего пользования (рис. 3.4). Она фактически представляет собой совокупность национальных сетей, которые соединяются для обслуживания международных звонков. Хотя эта сеть первоначально была разработана и построена для аналоговых абонентов-телефонов, она обслуживает значительный поток данных через модемы и постепенно превращается в цифровую сеть. Еще один известный пример использования коммутации каналов — ведомственные телефонные сети, применяемые для соединения телефонов в здании или офисе. Коммутация каналов также применяется в частных сетях. Как правило, такую сеть создает корпорация или другая большая организация для связи между своими филиалами. Такая сеть обычно состоит из ведомственных систем в каждом филиале, соединенных выделенными линиями, предоставленными каким-либо оператором связи, например AT&T. Последний распространенный пример использования коммутации каналов — коммутатор данных. Коммутатор данных подобен ведомственной телефонной станции, но предназначен для соединения устройств обработки цифровых данных, таких, как терминалы и компьютеры.

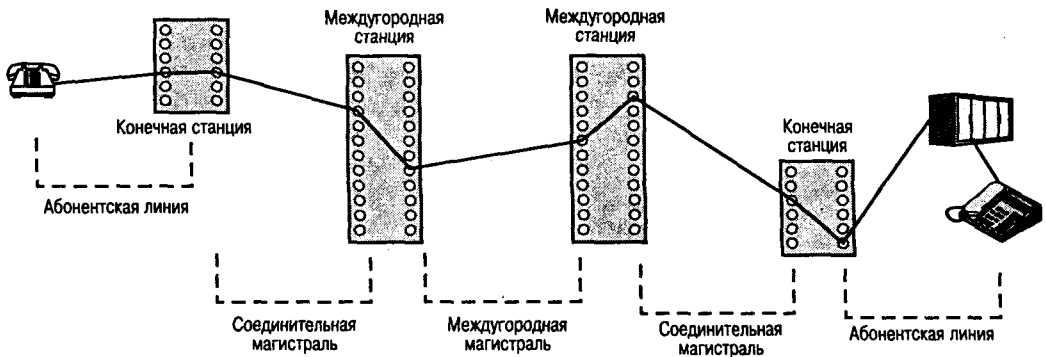


Рис. 3.4. Пример соединения через общедоступную сеть с коммутацией каналов

Общедоступную сеть связи можно описать с использованием четырех универсальных архитектурных компонентов.

- **Абоненты.** Устройства, которые подключаются к сети. До сих пор большую часть абонентских устройств в общедоступных сетях связи составляют телефоны, но доля устройств передачи данных год от года растет.
- **Абонентская линия.** Канал между абонентом и сетью. Называется также *абонентским шлейфом* или *абонентским каналом*. Почти во всех абонентских линиях используются витые пары. Длина абонентской линии

обычно составляет от нескольких километров до нескольких десятков километров.

- **Коммутаторы.** Центры коммутации в сети. Центр коммутации, который непосредственно обслуживает абонентов, называется *конечной станцией*. Как правило, конечная станция обслуживает тысячи абонентов в ограниченной области. В США имеется более 19 000 конечных станций, поэтому нереально для каждой конечной станции иметь прямую линию к каждой из других конечных станций; потребовалось бы порядка  $2 \times 10^8$  каналов. Взамен этого применяются промежуточные коммутаторы.
- **Магистралы.** Каналы между коммутаторами. Магистралы, благодаря частотному или временному уплотнению, содержат множество каналов звуковой частоты. Раньше магистралы назывались *многоканальными линиями связи*.

Абоненты соединяются непосредственно с конечной станцией, которая коммутирует телефонный обмен между абонентами и между абонентом и другими коммутаторами. Другие коммутаторы отвечают за маршрутизацию и коммутацию телефонного обмена между конечными станциями. Это различие показано на рис. 3.5. Соединение между двумя абонентами, подключенными к одной и той же конечной станции, создается таким же образом, как описано выше. Если абоненты подключены к разным конечным станциям, соединение между ними состоит из цепи соединений через одну или более промежуточных станций. На рисунке соединение между абонентами *a* и *b* создается путем их простой коммутации через конечную станцию. Установка соединения между *c* и *d* более сложна. На конечной станции абонента *c* устанавливается соединение между линией абонента *c* и одним каналом магистралы с временным уплотнением к промежуточному коммутатору. В промежуточном коммутаторе этот канал соединяется с каналом с временным уплотнением, ведущим к конечной станции абонента *d*. На этой конечной станции канал соединяется с линией абонента *d*.

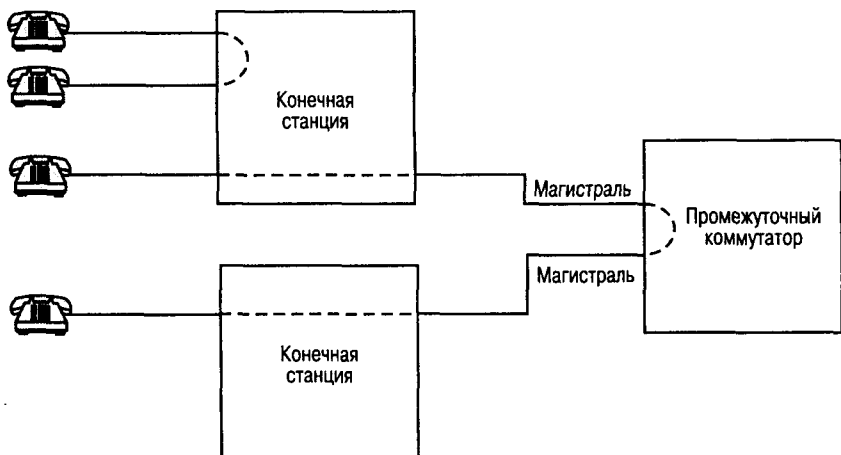


Рис. 3.5. Установка соединения

Технология коммутации каналов развивалась под воздействием тех же требований, которые предъявлялись к передаче голосовых сигналов. Одно из таких

требований — как можно меньшая задержка при передаче сигналов и, конечно, отсутствие изменений во время этой задержки. Должна была поддерживаться постоянная скорость передачи сигнала, поскольку передача и прием осуществляются при одной и той же скорости. Выполнение этих требований необходимо, чтобы происходил обычный разговор людей. Кроме того, качество принимаемого сигнала должно быть достаточно высоким, чтобы, как минимум, обеспечивать разборчивость речи.

Коммутация каналов стала широко распространенной и доминирующей потому, что она хорошо подходит для аналоговой передачи голосовых сигналов. В сегодняшнем цифровом мире ее неэффективность очевидна. Однако, несмотря на свою неэффективность, коммутация каналов остается привлекательной технологией как для локальных, так и для глобальных сетей. Одно из ее главных преимуществ — незаметность для пользователя. Когда соединение установлено, для подключенных станций оно кажется прямым; никакая дополнительная сетевая логика на этих станциях не требуется.

### 3.4. КОММУТАЦИЯ ПАКЕТОВ

Сети дальней связи с коммутацией каналов первоначально предназначались для передачи голосовых сигналов, и до сих пор большую часть информационного обмена в этих сетях составляют голосовые сообщения. Основная отличительная особенность сетей с коммутацией каналов состоит в том, что ресурсы сети выделяются для конкретного сеанса связи. Для телефонных соединений такая схема гарантирует высокую степень использования, так как большую часть времени говорит или одна, или другая сторона. Но когда сети с коммутацией каналов начали все шире применяться для передачи данных, проявились два недостатка.

- В типичном сеансе передачи данных “терминал—хост” (например, подключение персонального компьютера к серверу базы данных) большую часть времени канал не занят. Поэтому для передачи данных коммутация каналов неэффективна.
- В сети с коммутацией каналов предполагается передача данных с постоянной скоростью. Таким образом, оба общающихся устройства должны работать с одинаковой скоростью передачи данных. Поскольку серверы и рабочие станции могут быть очень разными, это условие значительно ограничивает область применения коммутации каналов.

Чтобы понять, как эти проблемы решаются в сети с пакетной коммутацией, рассмотрим механизм коммутации пакетов. Данные передаются в виде коротких пакетов. Обычно длина пакета ограничена 1000 октетами (байтами). Если отправитель должен послать более длинное сообщение, оно разбивается на ряд пакетов, как показано на рис. 3.6. Каждый пакет содержит часть данных пользователя (или все данные, если сообщение невелико) и некоторую управляющую информацию. Эта информация включает, как минимум, сведения, необходимые, чтобы провести пакет через сеть и доставить нужному получателю. На каждом узле по пути следования пакета он принимается, на короткое время сохраняется и передается на следующий узел.

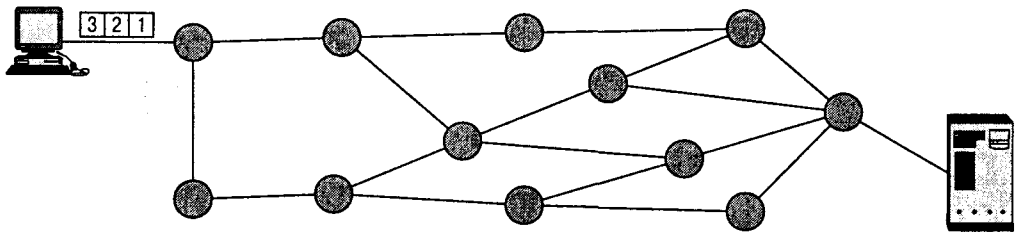


Рис. 3.6. Применение пакетов

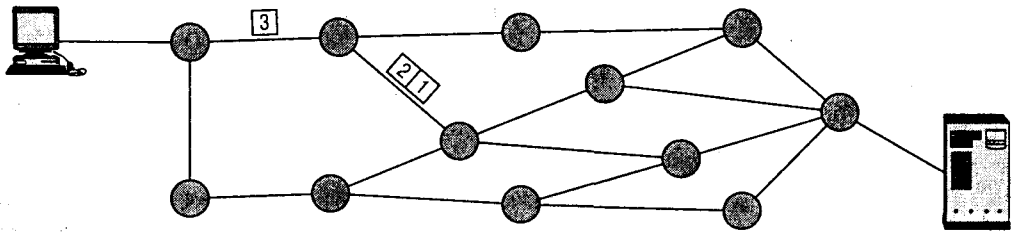
Использование данной технологии представлено на рис. 3.7. Передающий компьютер или другое устройство посылает сообщение как последовательность пакетов (а). Каждый пакет содержит управляющую информацию, которая указывает целевую станцию (компьютер, терминал, и т.д.). В первую очередь пакеты передаются узлу, к которому присоединена передающая станция. По мере того, как пакеты прибывают на этот узел, они ненадолго сохраняются, после чего определяется следующий маршрутный участок на пути к адресату, и пакеты передаются дальше. Передача начинается в тот момент, когда нужная линия становится доступной (б). Со временем все пакеты пройдут через сеть и поступят к нужному адресату.

Технология коммутации пакетов имеет определенные преимущества перед технологией коммутации каналов.

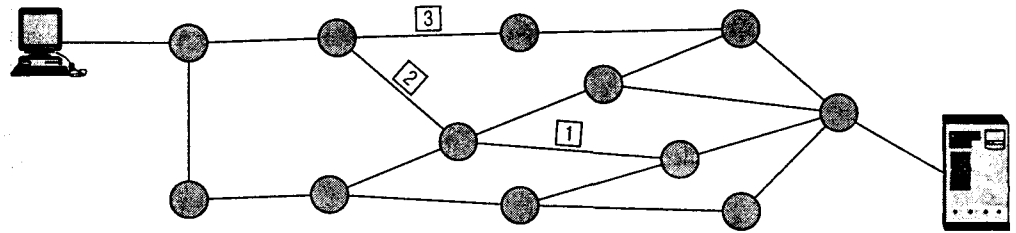
- Линия используется более эффективно, так как каждый канал между двумя узлами может поочередно использоваться несколькими пакетами. Пакеты стоят в очереди и передаются по каналу с максимально возможной скоростью. При коммутации каналов, напротив, время передачи по каналу установлено заранее по принципу уплотнения с временным разделением. Следовательно, большую часть выделенного времени канал может простаивать, так как соединение, для которого он выделен, неактивно.
- В сети с пакетной коммутацией можно изменять скорость передачи данных. Две станции с разной скоростью передачи могут обмениваться пакетами, поскольку каждая из них соединяется со своим узлом на "своей" скорости передачи данных.
- Когда в сети с коммутацией каналов увеличивается нагрузка, некоторые соединения блокируются, т.е. сеть не принимает новые запросы на соединение, пока нагрузка не уменьшится. Сеть с пакетной коммутацией, напротив, продолжает принимать пакеты, но задержка при доставке увеличивается.
- Можно применять приоритеты. Иначе говоря, если на узле ожидает очереди на передачу некоторое количество пакетов, пакеты с более высоким приоритетом будут переданы первыми. Следовательно, эти пакеты будут доставлены с меньшей задержкой, чем низкоприоритетные.



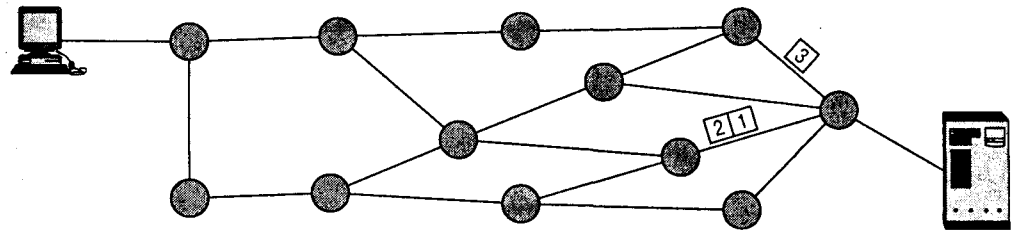
a)



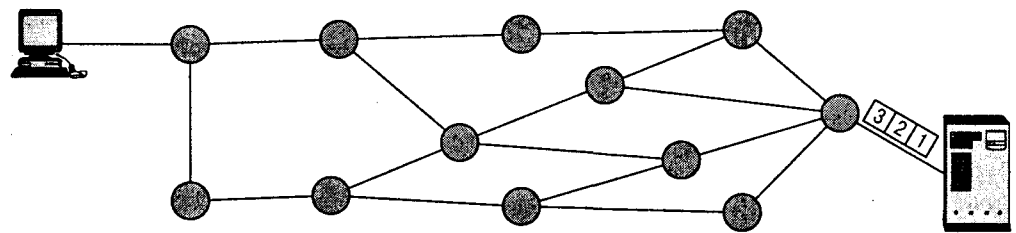
б)



в)



г)



д)

Рис. 3.7. Коммутация пакетов: использование дейтаграмм



Технология коммутации пакетов имеет также определенные недостатки по сравнению с коммутацией каналов.

- При каждом проходе пакета через узел коммутационной сети возникает задержка, отсутствующая при коммутации каналов. Эта задержка не меньше длины пакета в битах, деленной на входящую пропускную способность канала в битах в секунду; т.е. времени, уходящего на помещение пакета в внутренний буфер. Кроме того, могут возникать дополнительные задержки, связанные с обработкой и организацией очереди в узле.
- Поскольку пакеты, передаваемые данным источником данному адресату, могут иметь различный размер, передаваться по различным маршрутам и задерживаться в узлах на различное время, общее время задержки пакета может значительно варьироваться. Это явление, именуемое дрожью, может быть нежелательным для некоторых приложений (например, приложений реальной связи, в том числе передачи речи по телефону и видео реального времени).
- Для передачи пакетов через сеть к каждому пакету необходимо добавить служебную информацию, включающую адрес назначения и, часто, информацию по упорядочению пакетов. Все это снижает процент использования доступной пропускной способности. В сети с коммутацией каналов данные служебные издержки становятся ненужными сразу же после настройки канала.
- При передаче с коммутацией пакетов требуется большая обработка, чем при коммутации каналов, когда обработка на промежуточных узлах практически отсутствует.

## Использование коммутации пакетов

Если станция должна отправить через сеть с пакетной коммутацией сообщение, длина которого превышает максимальный размер пакета, она разбивает сообщение на пакеты и отправляет их по одному. Возникает вопрос, как этот поток пакетов будет обработан сетью, задача которой — переслать их по некоторому маршруту и доставить нужному получателю? В настоящее время для этого используются два метода: дейтаграмм и виртуального канала.

В случае использования метода **дейтаграмм** каждый пакет обрабатывается независимо от других, без всякой связи с пакетами, которые были переданы раньше. Рассмотрим результаты такого подхода (рис. 3.7). Каждый узел выбирает очередной узел по пути следования пакета, учитывая информацию, полученную от соседних узлов и касающуюся трафика, сбоев в линии связи и т.п. Таким образом, пакеты, следующие к одному адресату, идут по разным маршрутам и могут прибывать в порядке, отличном от исходного. В данном примере правильный порядок определяет узел на выходе сети. В некоторых сетях пакеты упорядочиваются адресатом. Кроме того, в ходе передачи по сети пакет может быть разрушен. Например, при кратковременном сбое узла коммутации пакетов все пакеты в его очередях будут утрачены. Как и ранее, обнаружить потерю пакета и выяснить, как его восстановить, должен либо целевой узел, либо узел на выходе сети. При этом методе каждый независимо обрабатываемый пакет называется дейтаграммой.

В случае использования метода **виртуального канала** маршрут прохождения пакетов устанавливается заранее. Как только маршрут установлен, пакеты следуют только по нему (рис. 3.8). Поскольку этот маршрут установлен на все время существования логического соединения, он несколько напоминает тракт в сети с коммутацией каналов и поэтому называется *виртуальным каналом*. Каждый пакет помимо данных содержит вместо адреса назначения идентификатор виртуального канала. Каждый узел на заранее установленном маршруте знает, куда направить такой пакет; принимать решение о маршрутизации не нужно. В каждый момент у каждой станции может быть несколько виртуальных каналов к любой другой станции, а также виртуальные каналы к нескольким станциям.

Таким образом, главная особенность метода виртуального канала состоит в том, что маршрут между станциями устанавливается до начала передачи данных. Следует заметить, что этот маршрут не является специально выделенным трактом, как при коммутации каналов. Как и при пакетной коммутации, пакеты записываются в буфер на каждом узле и накапливаются в очереди для вывода через канал, по которому также могут передаваться другие пакеты из других виртуальных каналов. Отличие от метода дейтаграмм заключается в том, что узел не должен принимать решения о маршрутизации для каждого пакета. Это делается только один раз для всех пакетов, использующих данный виртуальный канал.

Если две станции должны обмениваться данными в течение длительного времени, метод виртуальных каналов обнаруживает некоторые преимущества. Во-первых, сеть дает виртуальному каналу дополнительные возможности, такие, как упорядочение и защита от ошибок. Упорядочение обеспечивается тем, что все пакеты следуют по одному и тому же маршруту и поэтому доставляются получателю в исходном порядке. Защита от ошибок предусматривает не только правильный порядок получения пакетов, но и сохранность информации в них. Например, если пакет в последовательности от узла 4 к узлу 6 не прибывает на узел 6 или прибывает с ошибкой, узел 6 может потребовать повторно передать этот пакет с узла 4. Еще одно преимущество состоит в том, что по виртуальному каналу пакеты должны быстрее проходить через сеть, так как не нужно принимать решение о маршрутизации для каждого пакета на каждом узле.

Метод дейтаграмм имеет одно преимущество — отсутствие этапа установки соединения. Таким образом, если станция должна послать лишь один или несколько пакетов, дейтаграмма окажется более быстрой. Кроме того, метод дейтаграмм, как более простой, имеет большую гибкость. Например, если в одной части сети возникает перегрузка, входящие дейтаграммы могут быть направлены по другим, незагруженным, частям. В случае использования виртуального канала пакеты следуют по заранее установленному маршруту, и поэтому сети труднее приспособиться к перегрузке. Еще одно преимущество — дейтаграммы по самой своей сути более надежны. При использовании виртуальных каналов если один узел выходит из строя, все проходящие через него виртуальные каналы теряются. Если же используются дейтаграммы, последующие пакеты могут найти новый маршрут в обход этого узла.

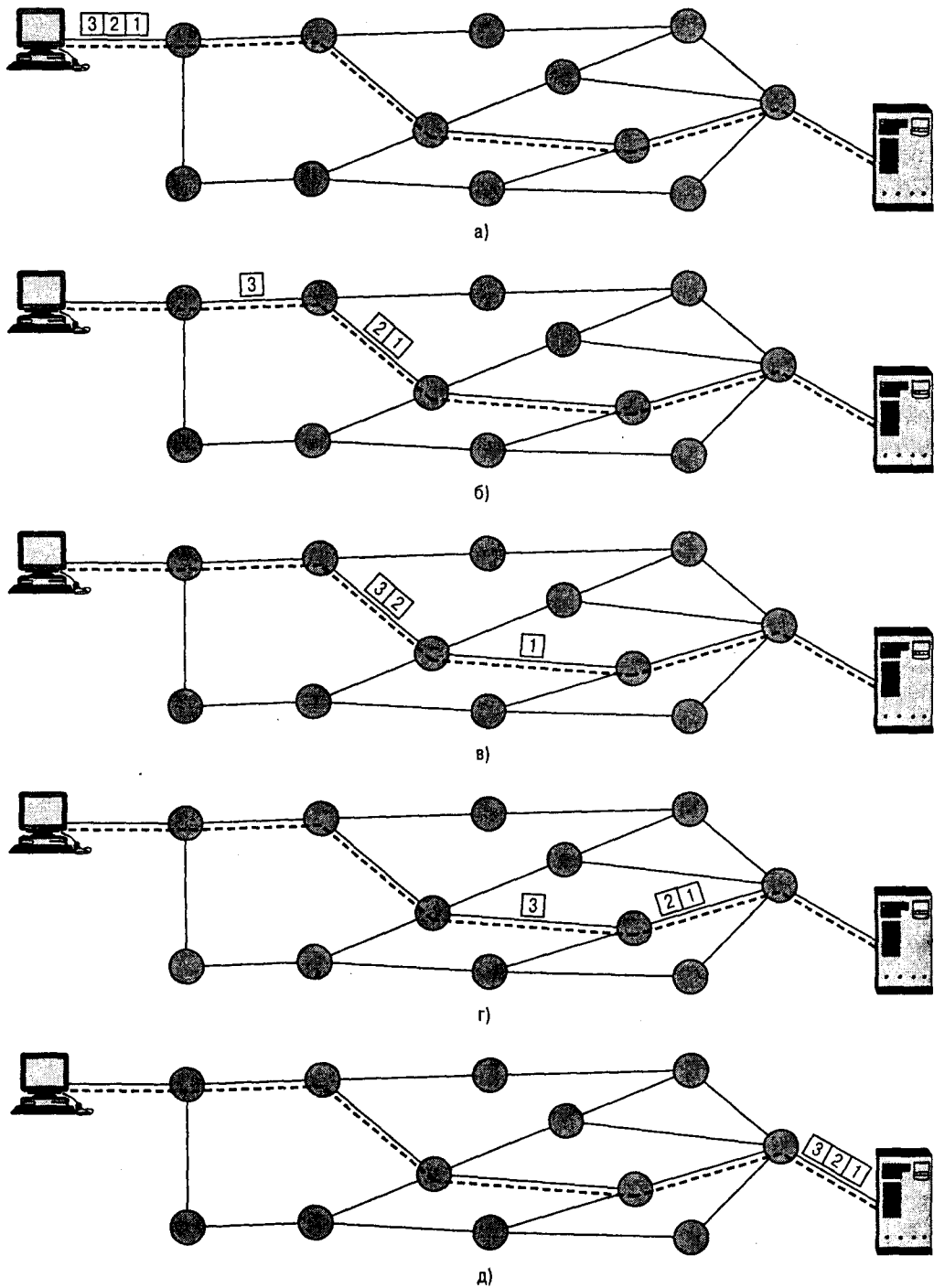


Рис. 3.8. Коммутация пакетов: использование виртуальных каналов

# Размер пакета

Между размером пакета и временем передачи имеется тесная взаимосвязь, как показано на рис. 3.9. В этом примере рассматривается виртуальный канал от станции А через узлы 4 и 1 к станции В. Передаваемое сообщение содержит 40 октетов, а каждый пакет включает 3 октета управляющей информации, которая помещается в начале каждого пакета и называется заголовком.

- а) Сообщение из 1 пакета    б) Сообщение из 2 пакетов    в) Сообщение из 5 пакетов    г) Сообщение из 10 пакетов

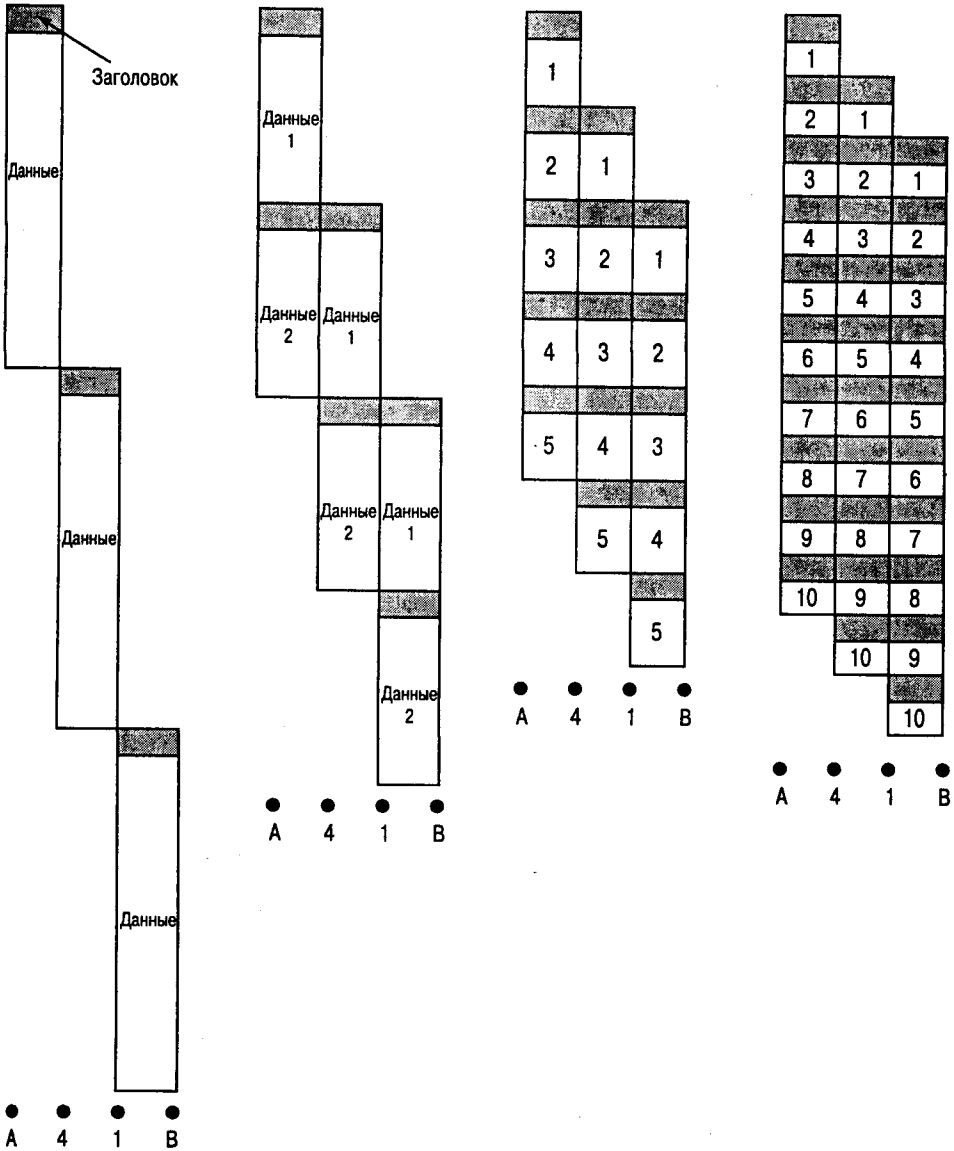


Рис. 3.9. Влияние размера пакета на время передачи

Пусть все сообщение отправляется как один пакет размером в 43 октета (3 октета заголовка и 40 октетов данных). Сначала пакет передается со станции А на узел 4 (рис. 3.9, а). Когда весь пакет будет получен, его можно будет передать с 4 на 1. Когда весь пакет получен на узле 1, он передается на станцию В. Без учета времени коммутации полное время передачи составляет 129 октет-тактов (43 октета × 3 передачи пакета).

Предположим теперь, что сообщение разбито на два пакета, каждый из которых содержит 20 октетов сообщения и, конечно, 3 октета заголовка, или управляющей информации. В этом случае узел 4 может начинать передачу первого пакета, как только он прибывает от А, не ожидая прибытия второго пакета. Из-за этого смещения передачи полное время передачи уменьшается до 92 октет-тактов. Если сообщение разбито на пять пакетов, каждый промежуточный узел может начинать передачу еще раньше и время передачи снижается еще больше, до 77 октет-тактов. Однако разбивка на все меньшие пакеты в конечном счете вызывает увеличение, а не уменьшение задержки, как показано на рис. 3.9, г. Дело в том, что каждый пакет содержит заголовки фиксированного размера, и большее количество пакетов означает большее количество этих заголовков. Кроме того, в нашем примере не показаны задержки на обработку и ожидание в очереди на каждом узле. Эти задержки также увеличиваются с ростом числа пакетов в одном сообщении. Однако в следующем разделе мы увидим, что даже при очень малом размере пакета (53 октета) можно достичь эффективной организации сети.

## 3.5. АТМ

Асинхронный режим передачи (Asynchronous Transfer Mode — АТМ), также именуемый поэлементной, или ячейечной, передачей (cell relay), в некоторых отношениях напоминает коммутацию пакетов. Как и в этой технологии, данные при передаче разбиваются на отдельные фрагменты. АТМ похож на пакетную коммутацию еще и тем, что здесь возможно создание нескольких логических соединений в одном физическом интерфейсе. В случае использования АТМ поток информации в каждом логическом соединении разделяется на пакеты фиксированного размера, называемые **ячейками**.

АТМ — это упрощенный протокол с минимальными возможностями управления потоком данных и защиты от ошибок. Благодаря этому служебные задержки на обработку ячеек АТМ и количество служебной информации, включаемой в каждую ячейку, невелики, что позволяет АТМ работать при высоких скоростях передачи данных. Кроме того, применение ячеек фиксированного размера упрощает обработку, необходимую на каждом узле АТМ, что также облегчает использование АТМ при высоких скоростях передачи.

### Логические соединения АТМ

Логические соединения в АТМ называются соединениями по виртуальному каналу (Virtual Channel Connections — VCC), подобному виртуальному каналу в сети с коммутацией пакетов. Соединение VCC — это основной объект коммутации в сети АТМ. Оно устанавливается через сеть между двумя конечными пользователями, и по нему с переменной скоростью проходит дуплексный поток ячеек

фиксированного размера. VCC также используются для связи между сетью и пользователем (передачи управляющих сигналов) и внутрисетевого обмена (управления сетью и маршрутизации).

В АТМ введен второй подуровень обработки, в котором используется понятие виртуального тракта (рис. 3.10). Соединение по виртуальному тракту (Virtual Path Connection — VPC) представляет собой пучок VCC, имеющих одни и те же конечные точки. Таким образом, все ячейки, передаваемые по всем VCC в одном VPC, коммутируются совместно.



Рис. 3.10. Структура соединений АТМ

Концепция виртуального тракта была разработана в связи с распространением высокоскоростных сетей, в которых стоимость управления сетью составляет все большую часть общей стоимости сети. Метод виртуальных каналов позволяет сдерживать рост стоимости управления за счет группирования соединений, использующих одни и те же тракты, в единый блок. В этом случае операции управления сетью применяются к небольшому количеству групп соединений, а не к массе отдельных соединений.

Ниже перечислены некоторые преимущества использования виртуальных трактов.

- **Упрощенная структура сети.** Транспортные функции сети можно разделить на применяемые к отдельному логическому соединению (виртуальному каналу) и применяемые к группе логических соединений (виртуальному тракту).
- **Повышенная производительность и надежность.** Сеть должна работать с меньшим количеством объединенных объектов.
- **Уменьшенный объем обработки и небольшая длительность установки соединения.** Большая часть работы выполняется уже при создании виртуального тракта. Благодаря резервированию пропускной способности соединения по виртуальному каналу в расчете на будущие запросы, новые VCC можно устанавливать, выполняя простые функции управления в конечных точках виртуального тракта; обработка запроса на промежуточных узлах не нужна. Таким образом, добавление новых виртуальных каналов в существующий виртуальный тракт требует минимальной дополнительной обработки.
- **Расширенные сетевые услуги.** Виртуальный тракт используется внутри сети, но видим и для конечного пользователя. Благодаря этому пользователь может задавать закрытые группы пользователей или закрытые сети пучков виртуальных каналов.

На рис. 3.11 приведена общая схема процесса установки соединения с использованием виртуальных каналов и виртуальных трактов. Процесс установки виртуального тракта и процесс установки конкретного соединения по виртуальному каналу разделены.

- Механизмы управления виртуального тракта включают определение маршрутов, резервирование пропускной способности и сохранение информации о состоянии соединения.
- Чтобы можно было создать виртуальный канал, должно существовать соединение по виртуальному тракту с пропускной способностью, достаточной для поддержки этого канала, и соответствующим качеством обслуживания. Виртуальный канал создается путем записи требуемой информации о состоянии (отображения виртуального канала на виртуальный тракт).

Терминология виртуальных трактов и виртуальных каналов, используемая в стандартах, довольно сложна; она приводится в табл. 3.1.

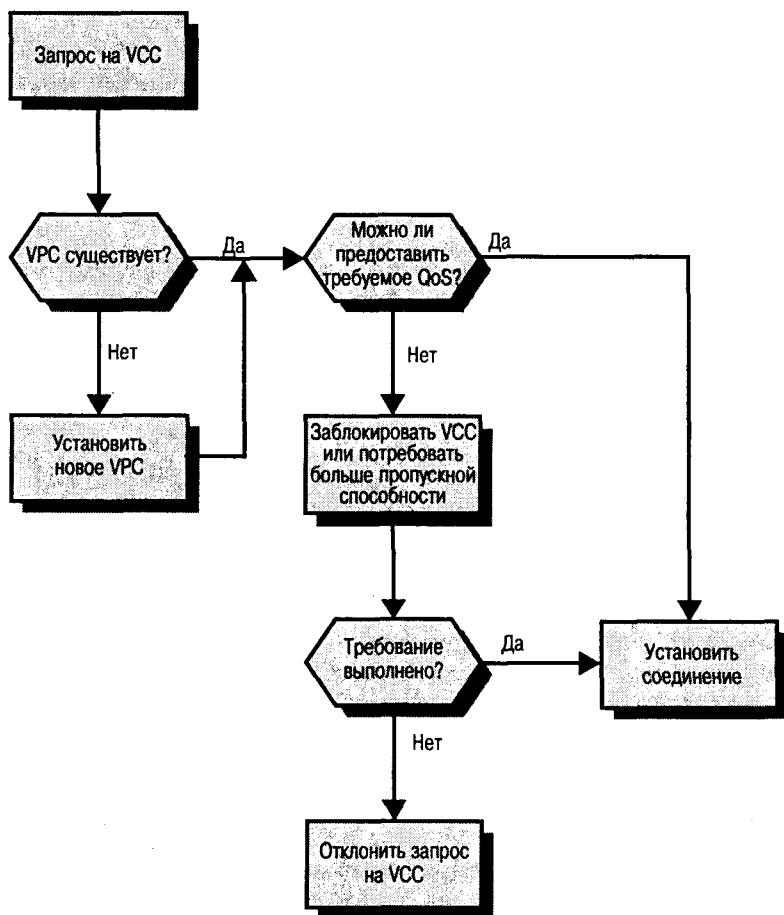


Рис. 3.11. Процесс установки соединения с использованием виртуальных трактов

**Таблица 3.1. Терминология виртуальных каналов и виртуальных трактов**

---

**Виртуальный канал (VC)**

Универсальный термин, предназначенный для описания однонаправленной транспортировки ячеек ATM, объединенных общим значением уникального идентификатора

**Связь виртуального канала**

Средство однонаправленной транспортировки ячеек ATM от точки, в которой присваивается значение VCI, до точки, в которой это значение преобразуется или сбрасывается

**Идентификатор виртуального канала (VCI)**

Уникальная числовая метка, идентифицирующая отдельную связь VC для данного VPC

**Соединение по виртуальному каналу (VCC)**

Объединение связей VC, соединяющее две точки, в которых пользователи службы ATM обращаются к уровню ATM. Соединения VCC могут служить для передачи информации между пользователями, между пользователем и сетью или внутри сети. Для ячеек, принадлежащих к одному VCC, сохраняется последовательность передачи

**Виртуальный тракт (VP)**

Универсальный термин, предназначенный для описания однонаправленной транспортировки ячеек ATM, принадлежащих к виртуальным каналам, объединенным общим значением уникального идентификатора

**Связь виртуального тракта**

Группа связей VC, определенных общим значением VPI между точкой, в которой присваивается значение VPI, и точкой, в которой это значение преобразуется или сбрасывается

**Идентификатор виртуального тракта (VPI)**

Идентифицирует отдельный канал VP

**Соединение по виртуальному тракту (VPC)**

Объединение связей VP, соединяющее точку, в которой присваиваются значения VCI, и точку, в которой эти значения преобразуются или удаляются (т.е. простирающееся на длину пучка связей VC, использующих один и тот же VPI). Соединения VPC могут служить для передачи информации между пользователями, между пользователем и сетью или внутри сети

---

### **Использование соединений по виртуальному каналу**

Конечными точками VCC могут быть конечные пользователи, объекты сети или конечный пользователь и объект сети. Во всех этих случаях в пределах VCC сохраняется исходная последовательность ячеек, иначе говоря, ячейки доставляются в том же порядке, в котором они отправлены. Рассмотрим три примера использования VCC.

- **Между конечными пользователями.** Этот вариант может применяться для сквозной передачи пользовательских данных, а также для передачи сигналов управления между конечными пользователями, как будет описано ниже. VPC между конечными пользователями предоставляет им пропускную способность в полном объеме; за организацию соединений VCC внутри VPC отвечают сами конечные пользователи, если только пропускной способности VPC достаточно для работы всех VCC.
- **Между конечным пользователем и объектом сети.** Этот вариант используется для передачи сигналов управления между пользователем и сетью, как будет показано ниже. Соединения VPC между пользователем и сетью могут



применяться для объединения трафика от конечных пользователей к коммутатору или серверу сети.

- **Между двумя объектами сети.** Данное соединение используется для функций управления трафиком и маршрутизации. Внутрисетевые соединения VPC могут применяться для определения общего маршрута обмена информацией управления сетью.

### **Характеристики виртуальных трактов и виртуальных каналов**

В Рекомендации ITU-T I.150 перечислены следующие характеристики соединений по виртуальному каналу.

- **Качество обслуживания.** Пользователю VCC предоставляется качество обслуживания, определяемое такими параметрами, как коэффициент потери ячеек (отношение числа потерянных ячеек к числу переданных) и разброс задержки ячеек.
- **Коммутируемые и полупостоянные соединения по виртуальному каналу.** Коммутируемое VCC представляет собой соединение по требованию, для установки и завершения которого необходимо передавать сигналы управления вызовами. Полупостоянное VCC — долговременное соединение, устанавливаемое операцией конфигурирования или управления сетью.
- **Сохранность последовательности ячеек.** Порядок следования ячеек в пределах VCC сохраняется.
- **Согласование параметров трафика и контроль использования.** Параметры трафика для каждого VCC можно согласовывать между пользователем и сетью. Сеть контролирует поступление ячеек в VCC, что позволяет поддерживать согласованные параметры.

Возможно согласование таких параметров трафика, как средняя скорость передачи, пиковая скорость передачи и длительность пикового периода. Чтобы справиться с перегрузкой и управлять существующими и запрашиваемыми VCC, для сети могут потребоваться разнообразные стратегии. На самом простом уровне сеть может предотвращать перегрузку, просто отклоняя новые запросы на VCC. Кроме того, ячейки могут аннулироваться в случае нарушения согласованных параметров или нарастания перегрузки. В критических ситуациях сеть может закрывать существующие соединения.

В I.150 также приводятся характеристики VPC. Первые четыре из этих характеристик совпадают с характеристиками VCC. Иными словами, качество обслуживания, коммутируемые и полупостоянные VPC, сохранность последовательности ячеек и согласование параметров трафика, контроль использования также входят в число характеристик VPC. Такое повторение характеристик необходимо по нескольким причинам. Во-первых, это дает некоторую гибкость в удовлетворении требований, которые ставятся перед сетью. Во-вторых, сеть должна учитывать все требования к VPC, чтобы согласовывать в пределах VPC установку виртуальных каналов с заданными характеристиками. Наконец, после установки VPC конечные пользователи могут согласовывать параметры создания новых VCC. Характеристики VPC устанавливают правила, которым должны следовать конечные пользователи при выборе вариантов соединений.

Кроме того, соединения VPC имеют пяту, дополнительную характеристику.

- **Ограничение идентификаторов виртуальных каналов в пределах VPC.** Один или несколько идентификаторов (номеров) виртуальных каналов могут быть недоступны для пользователя VPC и зарезервированы для использования сетью. В качестве примера можно привести номера VCC для управления сетью.

## Ячейки АТМ

В режиме асинхронной передачи применяются ячейки фиксированного размера, в которые входят заголовок, состоящий из 5 октетов, и информационное поле — из 48 октетов. Использование малых ячеек фиксированного размера дает некоторые преимущества. Во-первых, небольшие ячейки уменьшают задержку ожидания высокоприоритетных ячеек, поскольку низкоприоритетная ячейка, занявшая ресурс (например, передатчик), освободит его быстрее. Во-вторых, ячейки фиксированного размера коммутируются более эффективно, что важно для очень высоких скоростей передачи данных АТМ. При фиксированном размере ячейки проще реализовать механизм коммутации в аппаратуре.

### Формат заголовка

На рис. 3.12, а показан формат заголовка ячейки в интерфейсе между пользователем и сетью. На рис. 3.12, б показан формат заголовка, используемый внутри сети. Внешнее, по отношению к сети, поле общего управления потоком, реализующее сквозные функции, не поддерживается. Вместо этого поле идентификатора виртуального тракта расширяется с 8 до 12 бит. Это позволяет поддерживать большее число соединений VPC и включать соединения поддержки абонентов и соединения, требуемые для управления сетью.

Поле **общего управления потоком** (Generic Flow Control — GFC) существует в заголовке ячейки только в интерфейсе между сетью и пользователем, а внутри сети отсутствует. Следовательно, его можно использовать для управления потоком ячеек только в данном интерфейсе. Возможное назначение этого поля — помочь пользователю в управлении трафиком при различных уровнях качества обслуживания. В общем случае механизм GFC используется для смягчения кратковременных перегрузок в сети.

**Идентификатор виртуального тракта** (Virtual Path Identifier — VPI) образует поле маршрутизации для сети. Он включает 8 бит в интерфейсе между сетью и пользователем и 12 бит во внутрисетевом интерфейсе, разрешая поддержку увеличенного количества соединений VPC. **Идентификатор виртуального канала** (Virtual Channel Identifier — VCI) применяется для маршрутизации до конечного пользователя. Таким образом, функции VCI во многом сходны с функциями точки доступа к службе.

Поле **типа полезной нагрузки** (Payload Type — PT) указывает тип информации в информационном поле. Интерпретация битов поля PT показана в табл. 3.2. Значение 0 в первом бите обозначает информацию пользователя (т.е. информацию следующего, более высокого уровня). В этом случае второй бит показывает, имела ли место перегрузка; третий бит — бит типа служеб-

ного модуля данных (SDU)<sup>1</sup> — это 1-битовое поле, определяющее один из двух типов SDU ATM, связанных с соединением. Термин SDU относится к полезной нагрузке ячейки, занимающей 48 октетов. Значение 1 в первом бите поля типа полезной нагрузки указывает, что данная ячейка содержит информацию управления сетью или управления операциями. Это позволяет включать в пользовательские VCC ячейки управления сетью, не оказывая влияния на данные пользователя. Таким образом, поле PT может обеспечивать внутрисетевую передачу информации управления.

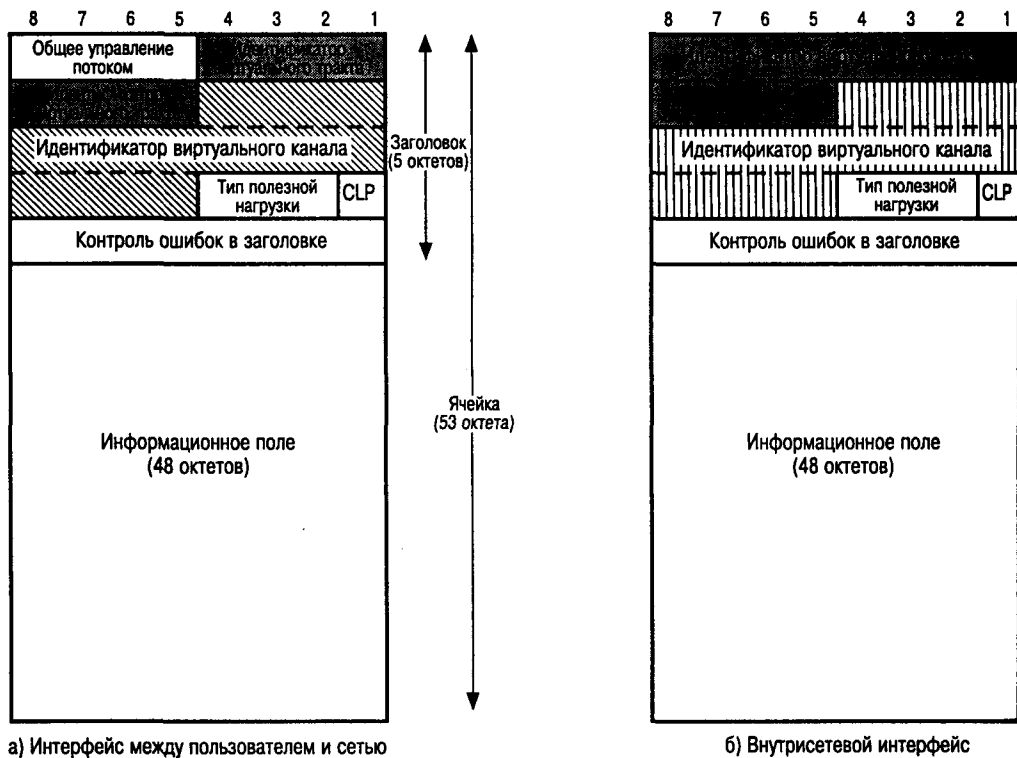


Рис. 3.12. Формат ячейки ATM

Таблица 3.2. Значения поля типа полезной нагрузки (PT)

Значение PT	Интерпретация
000	Ячейка данных пользователя, перегрузки не было, тип SDU = 0
001	Ячейка данных пользователя, перегрузки не было, тип SDU = 1
010	Ячейка данных пользователя, перегрузка была, тип SDU = 0
011	Ячейка данных пользователя, перегрузка была, тип SDU = 1

<sup>1</sup> Это термин, используемый в документах Форума ATM. В документации ITU-T данный бит называется *битом индикации между пользователями ATM* (ATM-user-to-ATM-user — AAU). Значение этих двух терминов одинаково.

значение РТ	Интерпретация
00	Ячейка функций ОАМ сегмента
01	Ячейка сквозных функций ОАМ
10	Ячейка управления ресурсами
11	Зарезервировано для будущих функций

SDU — служебный модуль данных (Service Data Unit).

ОАМ — управление операциями, администрирование, эксплуатация (Operations, Administration, and Maintenance).

Бит приоритета потери ячеек (Cell Loss Priority — CLP) определяет порядок действий сети в случае перегрузки. Значение 0 указывает, что это ячейка относительно более высокого приоритета, которая не должна отклоняться, если только есть другие возможности. Значение 1 указывает, что эту ячейку сеть может отклонять. С помощью данного поля пользователь может добавлять в сеть дополнительные ячейки (в результате чего скорость передачи превышает согласованную) с  $CLP = 1$ , которые будут доставляться адресату, если сеть не перегружена. Сеть может присваивать этому полю значение 1 для любой ячейки данных, нарушающей параметры трафика, согласованные между пользователем и сетью. Тогда коммутатор, устанавливающий это значение, признает, что данная ячейка нарушает согласованные параметры трафика, но еще может быть обработана. Затем, если в какой-то другой точке сети возникает перегрузка, эта ячейка помечается как подлежащая аннулированию до аннулирования ячеек, нарушающих согласованные пределы трафика.

Восьмибитовое поле контроля ошибок в заголовке (Header Error Control — HEC) может использоваться для коррекции 1-битовых ошибок и для выявления 2-битовых ошибок. В большинстве существующих протоколов, которые содержат поле защиты от ошибок, сам код ошибки получается гораздо короче данных, на базе которых он вычисляется. Это позволяет легко обнаруживать ошибки. В случае использования АТМ эти данные и сами по себе невелики — 32 бит при 8 битах кода. Благодаря этому код можно использовать не только для обнаружения ошибок, но и, в некоторых случаях, для их исправления. Это возможно потому, что код имеет достаточную избыточность, чтобы можно было исправлять некоторые типичные ошибки.

Функция контроля ошибок обеспечивает как восстановление после ошибок одиночных битов в заголовке, так и малую вероятность доставки ячеек с ошибочными заголовками при наличии пакетов ошибок. Для оптоволоконных систем передачи характерно сочетание ошибок одиночных битов с относительно большими пакетами ошибок. В некоторых системах передачи невозможно применить функцию исправления ошибок, которая занимает много времени.

## Категории услуг АТМ

Сеть АТМ позволяет одновременно передавать много различных типов трафика, включая звук, видеопередачу и пульсирующие потоки ТСП. Хотя каждый такой поток трафика обрабатывается как поток ячеек, состоящий из 53 октетов, проходящих по виртуальному каналу, способ обработки каждого потока данных в сети зависит от характеристик потока трафика и требований

приложения. Например, видеотрафик в реальном времени должен доставляться с минимальным разбросом задержки.

В этом разделе мы дадим обзор категорий услуг АТМ, на основании которых конечная система определяет тип требуемых услуг. Форумом АТМ определены следующие категории услуг.

- Услуги реального времени:
  - постоянная скорость передачи (CBR);
  - переменная скорость передачи в реальном времени (rt-VBR).
- Услуги модельного времени:
  - переменная скорость передачи в модельном времени (nrt-VBR);
  - доступная скорость передачи (ABR);
  - неопределенная скорость передачи (UBR).

### Услуги реального времени

Наиболее важное различие между приложениями касается допустимых величин и разброса задержки (последний параметр также называется дрожанием). Приложения реального времени, как правило, предлагают пользователю поток информации, который должен воспроизводить поток на передающей стороне. Например, поток аудио- или видеоинформации должен поступать к пользователю непрерывно и равномерно. Неравномерность или слишком большая степень потерь приводят к значительному снижению качества. На приложения, основанные на взаимодействии между людьми, накладываются жесткие требования относительно задержки. Как правило, задержка более нескольких сотен миллисекунд уже становится заметной и раздражающей. Следовательно, потребность в сетях АТМ, которые могли бы коммутировать и доставлять данные в реальном времени, довольно высока.

Услуги CBR (constant bit rate — постоянная скорость передачи) определяются наиболее просто. Они используются приложениями, в которых требуется фиксированная скорость передачи данных, доступная в течение всего существования соединения и имеющая относительно жесткое ограничение по максимальной величине задержки передачи. Услуги CBR обычно применяются для передачи аудио- и видеоинформации без сжатия. К приложениям CBR относятся следующие:

- проведение видеоконференций;
- интерактивный обмен аудиоинформацией (например, телефонная связь);
- распространение аудио- и видеоинформации (например, телевидение, дистанционное обследование, платное телевидение);
- доставка аудио- и видеоинформации (например, видео- или аудиобиблиотеки).

Категория **rt-VBR** (real-time variable bit rate — переменная скорость передачи в реальном времени) предназначена для приложений с высокими требованиями к скорости, т.е. тех, где величина и разброс задержки жестко ограничены. Основное различие между областями применения **rt-VBR** и **CBR** состоит в том, что приложения **rt-VBR** работают со скоростью передачи, которая меняется со временем. Иными словами, источники **rt-VBR** работают неравномерно. На

пример, при сжатии видеосигнала стандартными методами получается последовательность кадров разного размера. Поскольку для видеотрансляции в реальном времени требуется постоянный темп передачи кадров, фактическая скорость передачи данных будет меняться.

Услуги rt-VBR дают сети большую гибкость, чем CBR. В этом режиме возможно статистическое уплотнение в одном и том же объеме выделенной пропускной способности многих соединений, каждому из которых, тем не менее, предоставляются требующиеся для него услуги.

### Услуги модельного времени

Услуги модельного времени предназначены для приложений с неравномерным трафиком и отсутствием жестких требований к величине и разбросу задержки. Соответственно, сеть обладает большей гибкостью в обработке таких потоков трафика и большими возможностями применения статистического уплотнения, что повышает эффективность сети.

Для некоторых приложений, не работающих в реальном времени, можно заранее описать ожидаемый поток трафика, благодаря чему сеть может обеспечить значительно более высокое качество обслуживания (QoS) в отношении потерь и задержки. Для таких приложений применимы услуги nrt-VBR (non-real-time variable bit rate — переменная скорость передачи в модельном времени). В этом режиме конечная система указывает пиковую скорость передачи ячеек, долговременную или среднюю скорость и меру возможной неравномерности или сгруппированности ячеек. Имея эти сведения, сеть может так распределить ресурсы, чтобы обеспечить относительно низкую задержку и минимальную степень потери ячеек.

Услуги nrt-VBR можно применять для передачи данных с жесткими требованиями ко времени отклика. В качестве примеров можно привести заказ авиабилетов, банковские транзакции и контроль технологических процессов.

В каждый момент времени определенная часть пропускной способности сети ATM используется для передачи в режиме CBR и двух разновидностей режима VBR. Может существовать также запас незадействованной пропускной способности, что объясняется следующими причинами: 1) не все ресурсы были выделены для трафика CBR и VBR; 2) неравномерный характер трафика VBR приводит к тому, что в некоторые моменты используется меньшая пропускная способность, чем была выделена. Всю эту неиспользуемую пропускную способность можно задействовать для услуг UBR (unspecified bit rate — неопределенная скорость передачи). Эти услуги подходят для приложений, допускающих переменный размер задержки и потерю некоторых ячеек. Такими признаками обычно обладает трафик на базе TCP. В режиме UBR ячейки передаются в порядке поступления, при этом используется пропускная способность, не занятая другими услугами; величина задержки и степень потери ячеек могут меняться. Источник UBR не получает никаких гарантий сохранения параметров передачи и никаких сведений о возможной перегрузке; такой режим называется обслуживанием “без обязательств” (best-effort service). К примерам приложений UBR относятся следующие:

передача текста, цифровых данных, изображений, обмен сообщениями, распространение и доставка информации;

- работа удаленных терминалов (например, дистанционное управление компьютером).

Приложения неравномерного действия, работающие по надежному сквозному протоколу, такому, как TCP, могут обнаруживать перегрузку в сети, замечая увеличение времени двусторонней задержки и увеличение доли аннулированных пакетов. Однако TCP не содержит механизма, позволяющего справедливо распределять ресурсы между отдельными подключениями. Кроме того, TCP не обладает способностью сдерживать перегрузку на возможно более низком уровне, используя явную информацию от перегруженных узлов сети.

Чтобы улучшить обслуживание неравномерно работающих источников, которые иначе использовали бы UBR, были определены услуги ABR (available bit rate — доступная скорость передачи). Приложение, использующее ABR, указывает возможную пиковую скорость передачи ячеек (peak cell rate — PCR) и требуемую минимальную скорость передачи ячеек (minimum cell rate — MCR). Сеть распределяет ресурсы таким образом, чтобы все приложения ABR получили пропускную способность по крайней мере на уровне MCR. Затем вся остальная пропускная способность равномерно распределяется между всеми источниками ABR. Справедливость распределения пропускной способности гарантируется явной обратной связью от источников ABR. Весь объем пропускной способности, не используемый источниками ABR, остается доступным для трафика UBR.

В качестве примера приложения, использующего ABR, можно привести соединение локальных сетей между собой. В этом случае конечными системами, подключенными к сети ATM, являются маршрутизаторы.

На рис. 3.13 показана схема распределения ресурсов в сети в стационарном состоянии (без добавления или удаления виртуальных каналов).

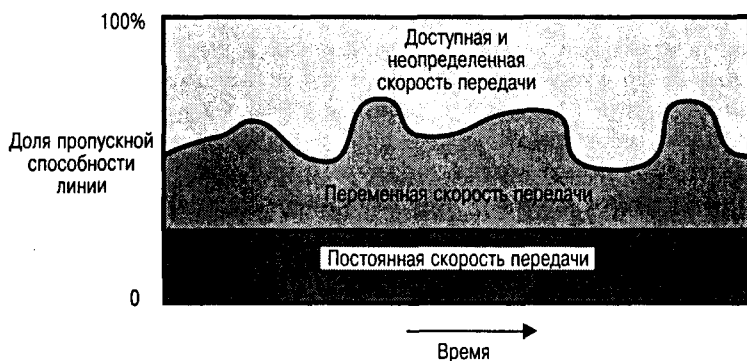


Рис. 3.13. Различные услуги ATM

## 3.6. РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА И WEB-САЙТЫ

Как и подобает, технологии с большим стажем — коммутации каналов — посвящено много литературы. В качестве книг, содержащих исчерпывающую информацию по данному предмету, можно порекомендовать [BELL00] и [FREE96]. Коммутации

пакетов уделено много внимания в работах [SPOH97], [BERT92] и [SPRA91]. Обширный обзор технологии ATM представлен в [MCDY99] и [BLAC99]. Все вопросы, поднимавшиеся в данной главе, подробно рассмотрены в книге [STAL00].

BELL00 Bellamy J. *Digital Telephony*. — New York: Wiley, 2000 // Беллами Дж. *Цифровая телефония*. — М.: Радиосвязь, 1986.

BERT92 Bertsekas D., Gallager R. *Data networks*. — Englewood Cliffs, NJ: Prentice Hall, 1992. // Бертсекас Д., Галлагер Р. *Сети передачи данных*. — М.: Мир, 1998.

BLAC99 Black U. *ATM. Volume I: Foundation for Broadband Networks*. — Upper Saddle River, NJ: Prentice Hall, 1992.

FREE96 Freeman R. *Telecommunication System Engineering*. — New York: Wiley, 1996.

MCDY99 McDysan D., Spohn D. *ATM: Theory and Application*. — New York: McGraw-Hill, 1999.

SPOH97 Spohn D. *Data Network Design*. — New York: McGraw-Hill, 1997.

SPRA91 Spragins J., Hammond J., Pawlikowski K. *Telecommunications Protocols and Design*. — Reading, MA: Addison-Wesley, 1991.

STAL00 Stallings W. *Data and Computer Communications*, 6th ed. — Upper Saddle River, NJ: Prentice Hall, 2000 // Столлинс В. *Компьютерные системы передачи данных*, 6-е изд.: Пер. с англ. — М.: Издательский дом “Вильямс”, 2002.



### Рекомендуемые Web-сайты

- ATM Forum Web Site ([www.atmforum.com/](http://www.atmforum.com/)). Содержит технические спецификации, “белые книги” и электронные копии журнала *53 Bytes* — издания Форума.
- Cell Relay Retreat ([cell-relay.indiana.edu/cell-relay](http://cell-relay.indiana.edu/cell-relay)). Содержит архивы списка рассылки по передаче ячеек, ссылки на множество документов и Web-сайтов по ATM.

## ТЕРМИНЫ, ВОПРОСЫ И ЗАДАЧИ

### Основные термины

Монетский шлейф	глобальная сеть	коммутация пакетов
режим асинхронной передачи (ATM)	городская сеть	локальная сеть
виртуальный канал	дейтаграмма	пакет
виртуальный тракт	заголовков	поячеечная передача
	коммутация каналов	ячейка

### Вопросы

1. Сформулируйте различия между локальными и глобальными сетями.
2. Почему для каждой пары станций полезно иметь более одного связующего маршрута передачи данных?
3. На какое приложение рассчитаны сети с коммутацией каналов?



4. Сформулируйте различия между статической и альтернативной маршрутизацией в сети с коммутацией каналов.
5. Что такое “полупостоянное соединение”?
6. Какие скорости передачи данных предлагает ISDN?
7. Объясните различие между работой схем с использованием дейтаграмм и виртуальных каналов.
8. Назовите ограничения, возникающие при использовании для передачи данных сети с коммутацией каналов.
9. Чем отличается виртуальный канал от виртуального тракта?

## Задачи

1. Заданы следующие параметры коммутируемой сети:

$N$  — число переходов между двумя заданными конечными системами;

$L$  — длина сообщения в битах;

$B$  — скорость передачи данных в битах в секунду во всех каналах;

$P$  — фиксированный размер пакета в битах;

$H$  — количество служебных битов (размер заголовка) в пакете;

$S$  — время установления соединения (коммутация каналов или виртуальный канал) в секундах;

$D$  — задержка распространения на один переход в секундах.

- а. При  $N = 4$ ,  $L = 3200$ ,  $B = 9600$ ,  $P = 1024$ ,  $H = 16$ ,  $S = 0,2$ ,  $D = 0,001$  вычислите суммарную задержку для коммутации каналов, пакетной коммутации с виртуальными каналами и пакетной коммутации с дейтаграммами. Подтверждение не производится. Задержка обработки на узлах не учитывается.
- б. Выведите общие формулы для трех методик, указанных в предыдущем пункте, показывающие, при каких условиях задержки будут равны.
2. Какое значение  $P$  в функции от  $N$ ,  $L$  и  $H$  приводит к минимуму суммарной задержки в сети с дейтаграммами? Предположим, что  $L$  намного больше  $P$ , а  $D$  равно нулю.
3. Рассмотрим простую телефонную сеть, состоящую из двух конечных станций и одного промежуточного коммутатора с дуплексной магистралью на 1 МГц между каждой конечной станцией и промежуточным коммутатором. В среднем каждый телефон делает четыре звонка за 8-часовой рабочий день со средней продолжительностью соединения 6 мин. Десять процентов соединений являются междугородными. Каково максимальное количество телефонов, которое может обслужить конечная станция?
4. Найдите ошибку в следующем рассуждении. Для пакетной коммутации требуется, чтобы к каждому пакету были добавлены биты адреса и управления. В результате появляются значительные служебные издержки. При коммутации каналов маршрут устанавливается незаметно и служебные биты не нужны.
  - а. Следовательно, при коммутации каналов нет дополнительных издержек.
  - б. Так как коммутация каналов не вызывает дополнительных издержек, канал должен использоваться более эффективно, чем при пакетной коммутации.
5. Если допустить, что сбои на любой станции или узле сети невозможны, может ли пакет быть доставлен не туда, куда следует?

6. Рассмотрим сеть с пакетной коммутацией из  $N$  узлов, связанных в следующие топологии.

- а. Звезда: один центральный узел без подключенной станции; все остальные узлы подключены к центральному узлу.
- б. Кольцо: каждый узел соединяется с двумя другими узлами, образуя замкнутый круг.
- в. Полносвязная сеть: каждый узел непосредственно соединен со всеми другими узлами.

Для каждого варианта укажите среднее число переходов между станциями.

7. Одно из основных решений при разработке АТМ касалось использования ячеек фиксированной или переменной длины. Рассмотрим это решение с точки зрения эффективности. Определим коэффициент эффективности передачи как

$$N = \frac{\text{число информационных октетов}}{\text{число информационных октетов} + \text{число служебных октетов}}.$$

- а. Рассмотрим использование пакетов фиксированной длины. В этом случае служебные издержки определяются только служебными октетами. Пусть

$L$  — размер поля данных ячейки в октетах;

$H$  — размер заголовка ячейки в октетах;

$X$  — число информационных октетов, передаваемых в одном сообщении.

Получите выражение для  $N$ . *Подсказка:* в выражении должен использоваться оператор  $\lceil \cdot \rceil$ , где  $\lceil Y \rceil$  — наименьшее целое число, которое больше или равно  $Y$ .

- б. Если ячейки имеют переменную длину, то служебную информацию составляет заголовок вместе с флагами для разграничения ячеек или дополнительным полем длины в заголовке. Пусть  $H_v$  — дополнительное количество служебных октетов, необходимое для использования ячеек переменной длины. Получите выражение для  $N$  в зависимости от  $X$ ,  $H$  и  $H_v$ .
- в. Пусть  $L = 48$ ,  $H = 5$  и  $H_v = 1$ . Постройте график зависимости  $N$  от размера сообщения для ячеек фиксированной и переменной длины. Прокомментируйте результаты.

8. Еще один ключевой вопрос, возникавший при разработке АТМ, — размер поля данных для ячеек фиксированного размера. Рассмотрим этот вопрос с точки зрения эффективности и времени задержки.

- а. Предположим, что имеет место длительная передача, т.е. все ячейки заполнены. Получите выражение для коэффициента эффективности  $N$  в зависимости от  $H$  и  $L$ .
- б. Задержка пакетирования — это задержка, вносимая в поток передачи вследствие необходимости записывать данные в буфер, пока не будет заполнен весь пакет. Получите выражение для этой задержки в зависимости от  $L$  и скорости передачи источника  $R$ .
- в. Обычные скорости передачи данных для кодирования речи — 32 Кбит/с и 64 Кбит/с. Постройте график зависимости задержки пакетирования от  $L$  для этих двух скоростей; отложите ось  $Y$  с левой стороны с максимальным значением 2 мс. На этой же диаграмме постройте график зависимости от  $L$  коэффициента эффективности передачи; отложите ось  $Y$  с правой стороны с максимальным значением 100%. Прокомментируйте результаты.

9. Рассмотрим передачу сжатого видеоизображения по сети ATM. Пусть стандартные ячейки ATM передаются через 5 узлов. Скорость передачи данных 43 Мбит/с.

- а. Чему равно время передачи одной ячейки через один коммутатор?
- б. Каждый коммутатор может передать ячейки из другого трафика, причем все такие ячейки считаются имеющими низший приоритет. Если коммутатор занят передачей ячейки, наша ячейка должна дождаться конца передачи. Если коммутатор свободен, наша ячейка передается немедленно. Чем равно максимальное время между прибытием видеоячейки на первый коммутатор и выходом ее из пятого (последнего) коммутатора? Временем распространения и временем коммутации можно пренебречь; время передачи время ожидания освобождения коммутатора следует учесть.
- в. Пусть известно, что каждый коммутатор 60% времени используется низкоприоритетным трафиком. Иными словами, в любой момент подхода ячейки к коммутатору он занят с вероятностью 0,6. Пусть время завершения передачи среднем равно половине времени передачи ячейки. Чему равно среднее время между прибытием видеоячейки на первый коммутатор и выходом ее из пятого коммутатора?
- г. Более интересным параметром является не задержка а дрожь, представляющая собой разброс задержек. Используя данные, содержащиеся в предыдущих пунктах, определите, соответственно, максимальную и среднюю дрожь задержки.

Во всех случаях предполагается, что различные случайные события независимы друг от друга; например, игнорируется типичная неравномерность подобной передачи.

# ГЛАВА 4

## ПРОТОКОЛЫ И НАБОР TCP/IP

**4.1. Необходимость архитектуры протоколов**

**4.2. Архитектура протоколов TCP/IP**

Уровни TCP/IP

Действие протоколов TCP и IP

Приложения TCP/IP

**4.3. Модель OSI**

Уровни протоколов

Стандартизация на основе архитектуры OSI

**4.4. Межсетевое взаимодействие**

Маршрутизаторы

Пример межсетевого взаимодействия

**4.5. Рекомендуемая литература**

**4.6. Термины, вопросы и задачи**

**Приложение 4А. Протокол Internet**

IPv4

IPv6

**Приложение 4Б. Протокол управления передачей**

Управление потоком TCP

Формат сегмента TCP

**Приложение 4В. Протокол пользовательских  
дейтаграмм**

**Д**анная глава начинается с представления концепции многоуровневой архитектуры протоколов. Затем рассматривается наиболее распространенная архитектура такого типа — набор протоколов TCP/IP. TCP/IP — это концепция Internet, представляющая основу для разработки широкого диапазона стандартов связи компьютеров. В настоящее время практически все поставщики компьютеров предлагают поддержку данной архитектуры. Другой заслуживающей упоминания архитектурой является эталонная модель OSI (Open System Interconnection — взаимодействие открытых систем). OSI — это стандартизованная архитектура, которая часто используется для описания функций связи, но реализуется в настоящее время редко.

После обсуждения архитектур протоколов исследуется важная концепция межсетевого взаимодействия (internetworking). Организации обязательно потребуются более одной сети связи, и будут нужны некоторые средства для обеспечения взаимодействия этих сетей — так что снова возникнут вопросы, связанные с архитектурой протоколов.

## **4.1. НЕОБХОДИМОСТЬ АРХИТЕКТУРЫ ПРОТОКОЛОВ**

При обмене данными между компьютерами, терминалами и/или другими устройствами обработки данных круг проблем может быть довольно широким. Рассмотрим, например, передачу файла с одного компьютера на другой. Для этого требуется наличие тракта передачи данных между этими компьютерами, прямого или проходящего через сеть связи. Но это еще не все. Обычно требуется выполнить следующие задачи.

1. Система источника должна либо активизировать прямой тракт передачи данных, либо идентифицировать для сети связи систему назначения.
2. Система источника должна удостовериться, что система назначения готова принять данные.
3. Приложение, выполняющее в системе источника передачу файла, должно удостовериться, что программа управления файлами в системе назначения готова получить и сохранить файл для указанного пользователя.
4. Если форматы файлов, используемые в двух системах, несовместимы, то одна или другая система должна выполнить преобразование формата.

Очевидно, что две компьютерные системы должны достаточно тесно взаимодействовать. Вместо того чтобы соответствующую логику реализовать как один модуль, задача взаимодействия разбивается на подзадачи, каждая из которых решается отдельно. В протокольной архитектуре модули располагаются в виде вертикального стека. Каждый уровень стека выполняет некоторую группу родственных функций, требуемых для связи систем. В реализации более примитивных функций он полагается на нижестоящий уровень и не интересуется подробностями этой реализации. Кроме того, каждый уровень предлагает услуги вышестоящему уровню. В идеале уровни должны быть определены так, чтобы изменения на одном из них не затрагивали другие уровни.

Разумеется, в связи участвуют две стороны, поэтому в обеих системах должны иметься одинаковые наборы многоуровневых функций. В таком контексте связь осуществляется между соответствующими, или *одноранговыми* (peer), уровнями сообщающихся систем. Передача происходит посредством форматированных блоков информации, управление которыми производится через принимаемые соглашения, называемые *протоколами*. Основными элементами протоколов являются следующие.

- **Синтаксис.** Включает такие понятия, как формат блоков данных.
- **Семантика.** Включает управляющую информацию, необходимую для координации и обработки ошибок.
- **Синхронизация.** Включает согласование скорости и упорядочение блоков информации.

## 2. АРХИТЕКТУРА ПРОТОКОЛОВ TCP/IP

Протокол TCP/IP является результатом исследований и разработок, проведенных в экспериментальной сети с коммутацией пакетов ARPANET, которые финансировались Управлением перспективных исследовательских программ США (Defense Advanced Research Projects Agency — DARPA). Обычно говорят о наборе протоколов TCP/IP. Этот набор состоит из большого числа протоколов, имеющих статус стандартов (IAB) Internet.

### Уровни TCP/IP

В общем случае в связи участвуют три агента: приложения, компьютеры и сети. Примеры приложений — передача файлов и электронная почта. Приложения, рассматриваемые в данной книге, являются распределенными, включающими обмен данными между двумя компьютерными системами. Эти (и другие) приложения выполняются на компьютерах, которые часто поддерживают множественные одновременно выполняемые приложения. Компьютеры связываются в сети, и обмениваемые данные передаются от одного компьютера другому через сеть. Таким образом, передача данных от одного приложения другому включает (1) передачу данных компьютеру, на котором “проживает” приложение, и (2) доставку (в пределах компьютера) данных целевому приложению.

В соответствии с рассмотренными концепциями задачи протокола TCP/IP можно разделить на пять относительно независимых уровней.

- Физический уровень.
- Уровень доступа к сети.
- Межсетевой уровень.
- Транспортный, или межузловой, уровень.
- Уровень приложений.

**Физический уровень** включает физический интерфейс между устройством передачи данных (например, рабочей станцией или компьютером) и передающей средой или сетью. Задача этого уровня — установление характеристик среды передачи, природы сигналов, скорости передачи и прочих подобных параметров.

**Уровень доступа к сети** связан с обменом данными между конечной системой и сетью, к которой она подключена. Передающий компьютер должен предоставить сети адрес компьютера назначения, чтобы сеть могла проложить маршрут передачи данных к конечному пункту. Передающий компьютер может затребовать определенные услуги, предоставляемые сетью, такие, например, как установление приоритета. Какое именно программное обеспечение будет использоваться на этом уровне, зависит от типа используемой сети; для сетей с коммутацией каналов, коммутацией пакетов (например, сетей с протоколом X.25), локальных сетей (например, сети Ethernet) и других типов сетей разработаны различные стандарты. Таким образом, функции, связанные с доступом к сети, имеет смысл выделить в отдельный уровень. При подобном подходе выбор остальных программ связи, обеспечивающих функции над уровнем доступа к сети, не будет зависеть от типа используемой сети. Программное обеспечение высших уровней должно работать одинаково корректно в любой сети, к которой присоединен компьютер.

Уровень доступа к сети обеспечивает связь с сетью и определяет маршрут продвижения данных по сети для двух конечных систем, подключенных к одной сети. Если устройства подключены к различным сетям, требуется механизм, позволяющий данным переходить из одной сети в другую. Это функция **межсетевого уровня**. Задачу межсетевого обмена на этом уровне выполняет протокол IP (Internet Protocol). Он реализуется не только в конечных системах, но и на маршрутизаторах. Маршрутизатор — это узел обработки, соединяющий две сети; его основная функция — передача данных из одной сети в другую по маршруту от источника к адресату.

Независимо от природы приложений, выполняющих обмен данными, существует обычное требование надежности такого обмена. Иными словами, мы хотели бы быть уверены в том, что все данные достигли приложения-адресата и что они поступили в том же порядке, в каком были отправлены. Механизмы, обеспечивающие надежную доставку данных, относятся к **транспортному, или межузловому, уровню**. Для этой цели обычно используется протокол TSP (Transmission Control Protocol — протокол управления передачей).

Наконец, **уровень приложений** содержит логику, необходимую для поддержки различных пользовательских приложений. Для каждого типа приложения (например, передачи файлов) требуется предназначенный именно для него отдельный модуль.

## Действие протоколов TSP и IP

На рис. 4.1 показано, как протоколы TSP/IP сконфигурированы для осуществления связи. Общее средство связи может состоять из множества сетей, каждая из них часто называется *подсетью*. Для соединения компьютера с подсетью используется некий протокол доступа к сети, например логика Ethernet. Он позволяет хосту посылать данные через сеть другому хосту или, если хост принадлежит другой сети, маршрутизатору. Для этого во всех конечных системах и маршрутизаторах реализуется протокол IP. Он действует как ретранслятор, перемещая блок данных от одного хоста к другому через один или несколько маршрутизаторов. Протокол TSP, наоборот, реализуется только в конечных системах и отслеживает блок данных с целью обеспечения надежной доставки данных соответствующему приложению.

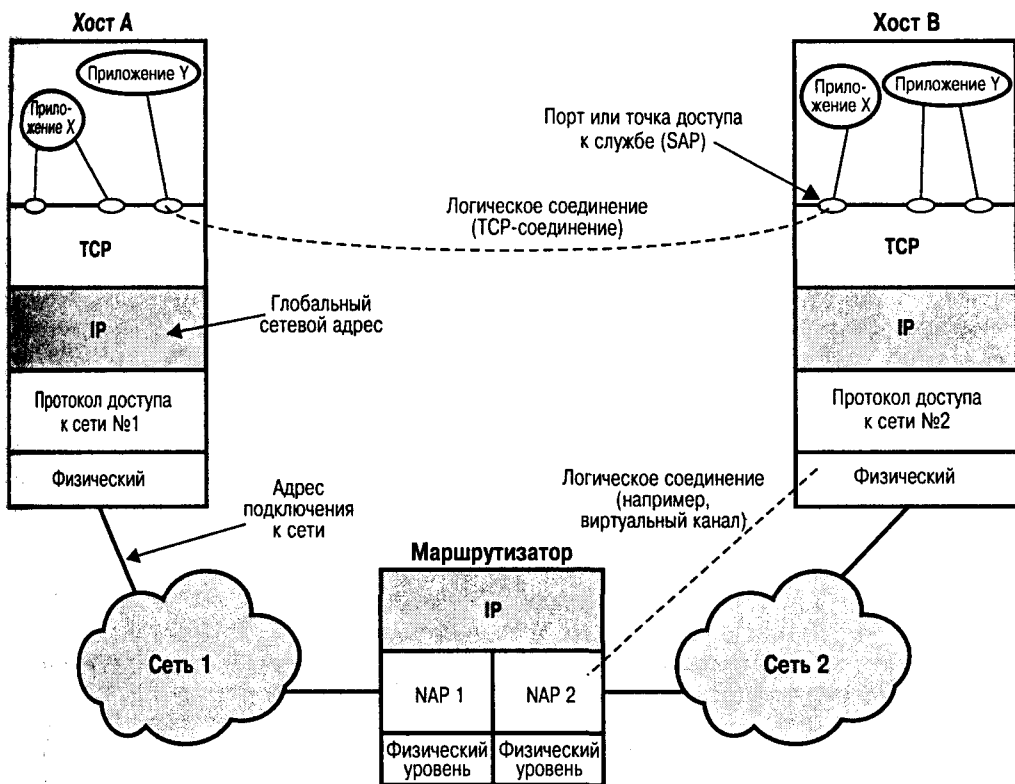


Рис. 4.1. Концепции TCP/IP

Для организации успешного сообщения каждый объект системы должен иметь уникальный адрес. Фактически требуется два уровня адресации. Каждый хост подсети должен иметь уникальный глобальный подсетевой адрес для доставки данных нужному хосту. Каждый процесс хоста должен также иметь адрес, уникальный для этого хоста, что позволяет протоколу сквозной передачи (TCP) доставлять данные соответствующему процессу. Адреса, используемые в последнем случае, называются *портами*.

Проследим простую операцию. Предположим, что процесс, связанный с портом 1 хоста А, желает передать сообщение другому процессу, связанному с портом 3 хоста В. Первый процесс спускает сообщение протоколу TCP, указывая адресовать его хосту В, порт 3. Протокол TCP передает это сообщение протоколу IP с указанием о пересылке его хосту В. Отметим, что протоколу IP не требуется указывать, для какого порта предназначено сообщение, нужно лишь указать, что данные предназначены для хоста В. Далее, протокол IP спускает сообщение на уровень доступа к сети (например, логике Ethernet) с указанием о пересылке его маршрутизатору J (первому маршрутизатору на пути к хосту В).

Для управления этой операцией должны передаваться не только пользовательские данные, но и управляющая информация, как это показано на рис. 4.2. Допустим, что передающий процесс генерирует блок данных и передает его протоколу TCP, который, стремясь улучшить управление, может разбить этот блок на меньшие части. К каждому такому фрагменту протокол TCP присоединяет



управляющую информацию (называемую заголовком TCP), формируя при этом *сегмент TCP*. Присоединенную управляющую информацию должен будет использовать одноранговый объект протокола TCP хоста В. Ниже приводятся примеры информации, которая может содержаться в таком заголовке.

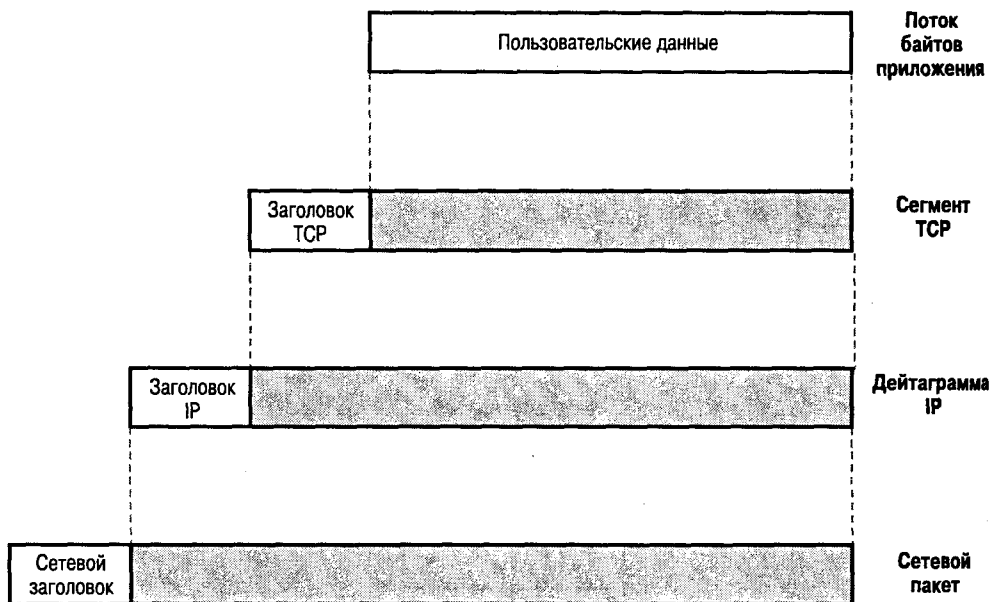


Рис. 4.2. Модули данных протокола в архитектуре TCP/IP

- **Порт назначения.** Когда объект TCP хоста В получает сегмент, он должен знать, кому доставить эти данные.
- **Порядковый номер.** Протокол TCP присваивает передаваемым определенному порту сегментам порядковые номера, чтобы при нарушении порядка объект TCP хоста В мог его восстановить.
- **Контрольная сумма.** Передающий протокол TCP включает в сегмент код представляющий собой функцию содержимого остальной части сегмента. Принимающий протокол TCP выполняет такие же вычисления и сравнивает результат с поступившим кодом. Несоответствие результатов показывает что в процессе передачи произошла ошибка.

Далее протокол TCP передает каждый сегмент протоколу IP с указанием передаче его хосту В. Эти сегменты должны передаваться через одну или несколько подсетей и ретранслироваться через один или несколько промежуточных маршрутизаторов. Описанная операция также требует использования управляющей информации, поэтому протокол IP присоединяет к данным заголовок с управляющей информацией, формируя таким образом *дейтаграмму IP*. В качестве примера информации, содержащейся в заголовке IP, можно назвать адрес хоста назначения (в нашем примере — хоста В).

На последнем этапе каждая дейтаграмма IP предоставляется уровню доступа к сети с целью передачи ее через первую подсеть, находящуюся на ее пути адресату. Уровень доступа к сети присоединяет к данным собственный заголовок

вок, создавая таким образом пакет, или кадр. Этот пакет передается через подсеть к маршрутизатору J. Информация, содержащаяся в заголовке пакета, необходима для передачи данных через подсеть. Ниже приводится информация, которая может содержаться в этом заголовке.

- **Адрес подсети назначения.** Подсеть должна знать, к какому из присоединенных устройств доставить пакет.
- **Запросы на функции.** Протокол доступа к сети может потребовать, чтобы сеть предоставила некоторые функции, например установление приоритета.

Маршрутизатор J отсекает заголовок пакета и изучает заголовок IP. На основании информации об адресе назначения, полученной из заголовка IP, модуль IP маршрутизатора направляет дейтаграмму IP через подсеть 2 к хосту В. Для этого дейтаграмма вновь инкапсулируется, т.е. к ней добавляется заголовок доступа к сети.

Когда хост В получает данные, происходит обратный процесс. На каждом уровне удаляется соответствующий заголовок, а оставшийся сегмент передается на вышестоящий уровень, и так продолжается до тех пор, пока исходные пользовательские данные не будут доставлены процессу-адресату.

Кстати, общее название блока данных, обмениваемого на любом протокольном уровне, — **модуль данных протокола** (protocol data unit — PDU). Таким образом сегмент TCP — это PDU уровня TCP (PDU TCP).

## Приложения TCP/IP

Для работы над протоколом TCP стандартизовано несколько приложений. Наиболее распространенные из них перечислены ниже.

**Простой протокол передачи почты (Simple Mail Transfer Protocol — SMTP)** предоставляет основные возможности электронной почты. Он обеспечивает механизм передачи сообщений между отдельными хостами. В число возможностей протокола SMTP входят списки рассылки, уведомление о получении и пересылка. Протокол SMTP не указывает способа создания сообщений, для этого требуются некоторые локальные средства редактирования или встроенные возможности электронной почты. После создания сообщения протокол SMTP принимает его и использует протокол TCP для отправки сообщения модулю SMTP другого хоста. После этого целевой модуль SMTP использует локальный пакет электронной почты для запоминания входящего сообщения в пользовательском почтовом ящике.

**Протокол передачи файлов (File Transfer Protocol — FTP)** используется для отправления по команде пользователя файлов из одной системы в другую. Эта услуга может оказываться как текстовым, так и двоичным файлам, а протокол предоставляет функции управления доступом пользователя к этим файлам. Когда пользователь запрашивает передачу файла, протокол FTP устанавливает соединение TCP с целевой системой для обмена управляющими сообщениями. Это позволяет передать идентификатор пользователя и его пароль, а также позволяет пользователю указать файл и требуемые действия. После получения согласия на передачу файла устанавливается второе соединение TCP для выполнения передачи данных. Файл передается через установленное соединение, при этом в него не включаются никакие “довески” в виде заголовков или управляющей информации на уровне приложений. После завершения передачи устанавливается контрольное соединение для передачи сообщений о завершении передачи и готовности к приему новых команд передачи файлов.

Протокол TELNET предоставляет возможность удаленного входа в систему, что позволяет пользователю через терминал или персональный компьютер войти в систему удаленного компьютера и действовать, как будто существует непосредственная связь с этим компьютером. Изначально этот протокол разрабатывался для работы с простыми терминалами с режимом прокрутки. Протокол TELNET, по существу, реализован в виде двух модулей. Один из них — пользовательский TELNET (User TELNET) — взаимодействует с модулем ввода-вывода терминала с целью сообщения с локальным терминалом. Этот модуль преобразует символы действительных терминалов в стандартные сетевые символы и наоборот. Другой — серверный TELNET (Server TELNET) — взаимодействует с приложением, действуя как псевдотерминал, так что удаленный терминал кажется приложению локальным. Передача данных между пользовательским и серверным протоколами TELNET возложена на TCP-соединение.

### 4.3. МОДЕЛЬ OSI

Модель взаимодействия открытых систем (Open Systems Interconnection — OSI) разработана Международной организацией по стандартизации (International Organization for Standardization — ISO)<sup>1</sup> в качестве модели для архитектуры компьютерной связи и основы для разработки стандартов протоколов.

#### Уровни протоколов

Модель OSI включает следующие семь уровней.

- Уровень приложений.
- Уровень представления.
- Сеансовый уровень.
- Транспортный уровень.
- Сетевой уровень.
- Канальный уровень.
- Физический уровень.

На рис. 4.3 иллюстрируется модель OSI; дается краткое описание функций, выполняемых на каждом уровне. Назначение модели OSI состоит в том, что для выполнения этих функций на каждом уровне должны быть разработаны протоколы.

Разработчики модели OSI предполагали, что сама модель и протоколы, разработанные в ее рамках, будут доминировать в среде передачи данных, вытеснив со временем частные реализации протоколов и конкурирующие обобщенные модели, такие, как набор протоколов TCP/IP. Этого не произошло. Хотя в рамках модели OSI и были разработаны многие полезные протоколы, в целом семиуровневая модель не нашла широкого применения. Вместо нее стала доминировать как раз архитектура TCP/IP. Это произошло по нескольким причинам. Пожалуй, важнейшей из них является та, что ключевые протоколы TCP/IP совершенствовались и тестировались

---

<sup>1</sup> ISO — это не аббревиатура (иначе писалось бы IOS), а слово, происходящее от греческого *isos*, означающего *равный*.

то время, когда сходные протоколы OSI находились лишь на этапе разработки. Когда встал вопрос о необходимости взаимодействия сетей, готовым к использованию был только набор TCP/IP. Другая причина успеха TCP/IP заключается в том, что модель OSI является излишне сложной, все задачи ее семи уровней реализуются при меньшем числе уровней TCP/IP.

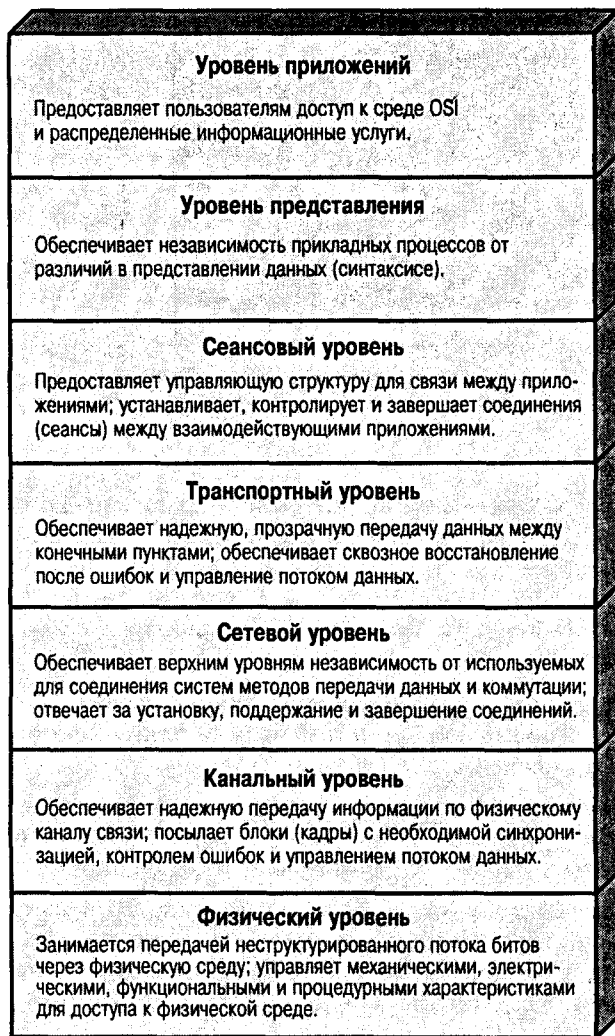


Рис. 4.3. Уровни модели OSI

На рис. 4.4 показаны уровни архитектур TCP/IP и OSI с приблизительным соотношением функций, выполняемых в рамках обеих моделей.

OSI	TCP/IP
Уровень приложений	Уровень приложений
Уровень представления	
Семантический уровень	
Транспортный уровень	Транспортный (межузловой) уровень
Сетевой уровень	Межсетевой уровень
Канальный уровень	Уровень доступа к сети
Физический уровень	Физический уровень

Рис. 4.4. Сравнение архитектур протоколов TCP/IP и OSI

## Стандартизация на основе архитектуры OSI

Основной целью разработки модели OSI было желание предоставить основу для стандартизации<sup>2</sup>. Поскольку модель только в общих терминах определяет функции, выполнение которых требуется на данном уровне, в рамках этой модели для каждого уровня можно разработать один или несколько стандартов протоколов. Отметим две особенности модели OSI, делающие ее удачной основой для выработки стандартов.

- Поскольку функции каждого уровня четко определены, стандарты могут разрабатываться одновременно и независимо для каждого уровня, что ускоряет процесс стандартизации.
- Благодаря четкому определению границ между уровнями изменения стандартов одного уровня не затрагивают программное обеспечение, существующее на других уровнях, что облегчает введение новых стандартов.

Рис. 4.5 иллюстрирует использование модели OSI в качестве такой основы для стандартизации. Общие функции связи разделены на семь различных уровней. Т.е. общие функции разбиваются на несколько модулей, что максимально упрощает сопряжение между модулями. Помимо этого, используются структурные принципы сокрытия информации: нижние уровни имеют дело с большим объемом деталей, в то время как верхние уровни от этих деталей не зависят. Еще одной особенностью структуры является то, что каждый уровень предоставляет услуги для вышестоящего уровня и протокол для однорангового уровня других систем.

На рис. 4.6 более детально показана природа стандартизации, требуемой на каждом уровне. При этом ключевыми являются такие три элемента.

<sup>2</sup> Понятия, вводимые в разделе, применимы и для архитектуры TCP/IP.

Общая информация

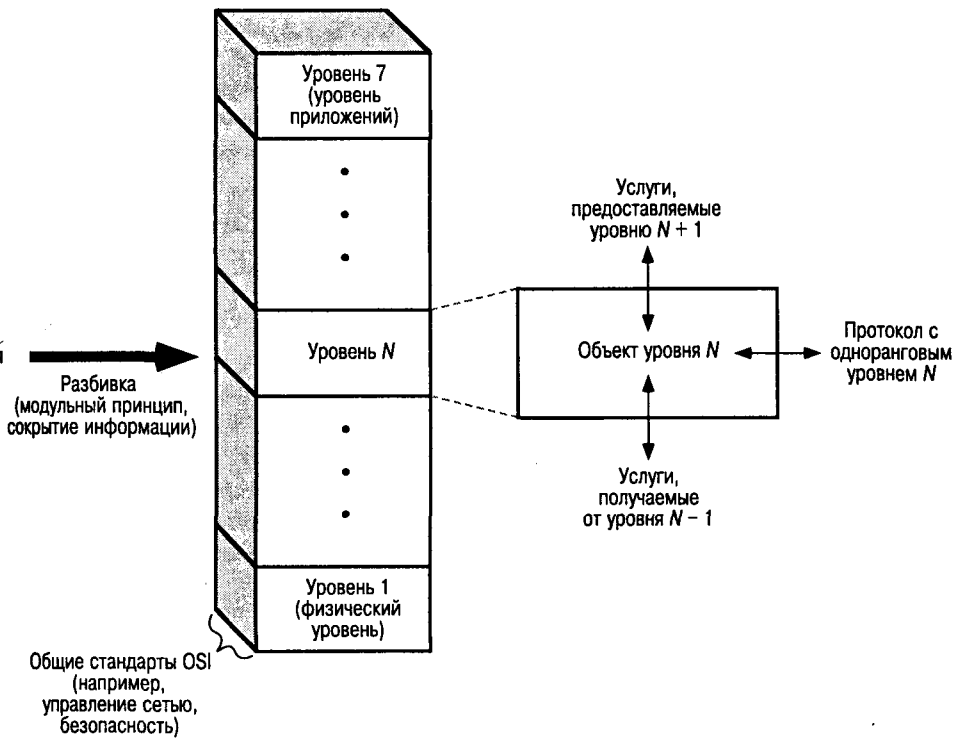


Рис. 4.5. Архитектура OSI как основа для стандартизации

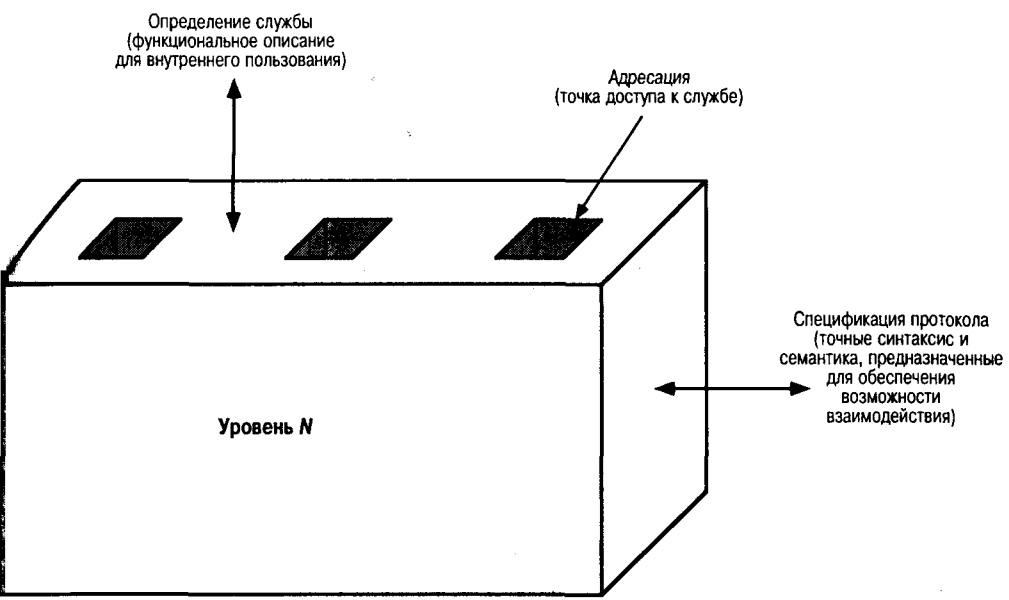


Рис. 4.6. Стандарты, характерные для уровня

- **Спецификация протокола.** Два объекта протокола, принадлежащих к одноранговым уровням различных систем, взаимодействуют посредством протокола. Поскольку в процессе задействованы две различные открытые системы, то протокол должен быть точно определен. Это касается формата обмениваемых PDU, семантики всех полей и допустимой последовательности PDU.
- **Определение услуг.** Стандарты требуются не только для протоколов, оперирующих на данном уровне, но и для услуг, которые каждый уровень предоставляет вышестоящему уровню. Определение услуг, как правило, равносильно функциональному описанию, где указано, какие услуги предоставляются, но не указано, каким образом.
- **Адресация.** Каждый уровень предоставляет услуги объектам вышестоящего уровня. Обращение к этим объектам происходит через точку доступа к службе (SAP). Таким образом, сетевая точка доступа к службе (NSAP) указывает на транспортный объект, являющийся пользователем службы сетевого уровня.

Необходимость предоставления для открытых систем точной спецификации протокола очевидна. Следующие два пункта требуют некоторых пояснений. Перечислим причины использования в качестве определения службы функционального описания. Во-первых, взаимодействие между двумя соседними уровнями ограничивается одной открытой системой и не затрагивает других. Следовательно, хотя одноранговые уровни различных систем предоставляют одинаковые услуги для вышестоящих уровней, способы предоставления услуг могут отличаться в различных системах, при этом не утрачивается возможность взаимодействия систем. Во-вторых, смежные уровни часто реализуются на одном процессоре. В этом случае было бы неплохо позволить программисту использовать аппаратное обеспечение и операционную систему для создания максимально эффективного интерфейса.

Что касается адресации, то использование на каждом уровне механизма адресации, реализованного в виде SAP, позволяет каждому уровню уплотнять многочисленных пользователей вышестоящего уровня. Уплотнение может и не происходить на каждом уровне, но модель это позволяет.

## 4.4. МЕЖСЕТЕВОЕ ВЗАИМОДЕЙСТВИЕ

В большинстве случаев локальная или глобальная сети не являются обособленным объектом. Чтобы удовлетворить все потребности организации, могут понадобиться несколько локальных сетей различных типов. Для решения вопросов производительности или безопасности возможно использование нескольких сетей одного типа. Кроме того, организация может иметь несколько географически раздельных локальных сетей и настраивать их взаимодействие через глобальную сеть с целью централизации управления и распределенного обмена информацией.

В табл. 4.1 перечислены некоторые часто используемые термины в области взаимодействия сетей, или межсетевого обмена. С точки зрения пользователя, связанная совокупность сетей может выглядеть как одна большая сеть. Однако если все составляющие сети сохраняют свою идентичность, а для осуществления связи сквозь границы сетей необходимы специальные механизмы, такая конфигурация часто называется **объединенной сетью**, или **internet** (с малой буквы). Важнейшим примером internet является сеть **Internet** (с большой буквы). По-

сколькy Internet развилась из исследовательской сети с коммутацией пакетов, она послужила основой для разработки технологии межсетевого взаимодействия и моделью для частных объединенных сетей в пределах организаций. Позже такие сети получили название **intranet**.

**Таблица 4.1. Термины межсетевого обмена**

**Сеть связи**

Средство обеспечения передачи данных между устройствами, подключенными к сети **internet**, или объединенная сеть

Совокупность сетей связи, объединенных мостами и/или маршрутизаторами

**intranet**

Корпоративная объединенная сеть, предоставляющая основные возможности Internet, в частности World Wide Web. Сеть intranet используется в пределах организации для внутренних целей и может быть изолированной, отдельной объединенной сетью или иметь связи с Internet

**Конечная система (ES)**

Устройство, подключенное к одной из сетей объединенной сети и используемое для поддержки приложений или услуг для конечных пользователей

**Промежуточная система (IS)**

Устройство, применяемое для соединения двух сетей и предоставляющее связь между конечными системами, подключенными к разным сетям

**Мост**

Промежуточная система, применяемая для соединения двух локальных сетей, в которых используются схожие сетевые протоколы. Мост играет роль фильтра адресов, отбирая из одной локальной сети пакеты, предназначенные для получателей в другой локальной сети, и передавая их. Мост не изменяет содержимого пакетов и ничего не добавляет к ним. Мост работает на уровне 2 модели OSI

**Маршрутизатор**

Промежуточная система, применяемая для соединения двух локальных сетей со схожими или различными сетевыми протоколами. Маршрутизатор использует протокол межсетевого обмена, имеющийся на каждом маршрутизаторе и в каждой конечной системе сети. Маршрутизатор работает на уровне 3 модели OSI

Каждая составляющая подсеть в объединенной сети поддерживает связь между устройствами, подключенными к данной сети; эти устройства называются **конечными системами** (End System — ES). Кроме того, сети связаны устройствами, которые в документах ISO называются **промежуточными системами** (Intermediate System — IS). Промежуточные системы создают тракт связи и выполняют необходимые функции ретрансляции и маршрутизации, в результате чего становится возможным обмен данными между устройствами, подключенными к различным подсетям в составе internet.

Особый интерес представляют два типа IS: мосты и маршрутизаторы. Различия между ними определяются типом протоколов, используемых для логики межсетевого взаимодействия. Так, мост работает на уровне 2 семиуровневой архитектуры взаимодействия открытых систем (OSI) и выполняет функцию ретранслятора кадров между сходными сетями. **Маршрутизатор** работает на уровне 3 архитектуры OSI и передает пакеты между потенциально различными сетями. И в том, и в другом случае предполагается, что используются одни и те же протоколы верхнего уровня.



О функциях маршрутизаторов речь шла ранее в этой главе в контексте протокола IP. В то же время поскольку маршрутизаторы играют важную роль в общей схеме межсетевого взаимодействия, на них стоит остановиться подробнее.

## Маршрутизаторы

Для организации межсетевого взаимодействия отличающихся сетей используются маршрутизаторы, связывающие различные подсети. К средствам межсетевого взаимодействия предъявляются следующие требования.

1. Предоставление канала связи между сетями.
2. Обеспечение маршрутизации и обмена данными между процессами в конечных системах различных сетей.
3. Выполнение указанных функций таким способом, который не требовал бы изменения архитектуры любой из составляющих подсетей.

Из третьего требования вытекает, что маршрутизатор должен приспосабливаться к множеству различий между сетями, включая следующие.

- **Различные схемы адресации.** В сетях могут использоваться различные схемы присвоения адресов устройствам. Например, в сети IEEE 802 каждому устройству выделяется 48-битовый двоичный адрес; в сети ATM обычно применяется 15-значный десятичный адрес (каждая цифра которого кодируется четырьмя битами, в результате чего получается 60-битовый адрес). Таким образом, необходима некоторая глобальная адресация, а также служба каталогов.
- **Различные максимальные размеры пакетов.** Возможно, пакеты из одной сети придется разбивать на меньшие части для другой сети. Этот процесс называется *сегментацией*, или *фрагментацией*. Например, Ethernet разрешает использование пакетов до 1500 байт; максимальный размер пакета в сетях X.25 составляет 1000 байт. Пакет, передаваемый системой Ethernet и полученный маршрутизатором для пересылки через сеть X.25, возможно, придется разбить на более мелкие части.
- **Интерфейсы.** Аппаратные и программные интерфейсы различных сетей отличаются. Концепция маршрутизации не должна зависеть от этих различий.
- **Надежность.** Различные сетевые службы могут предлагать как надежные сквозные виртуальные каналы, так и ненадежные услуги. Работа маршрутизатора не должна зависеть от предположений о надежности сети.

Перечисленные требования наилучшим образом удовлетворяет протокол межсетевого взаимодействия, такой как IP, реализованный во всех конечных системах и маршрутизаторах.

## Пример межсетевого взаимодействия

На рис. 4.7 представлена конфигурация, которую мы используем для иллюстрации взаимодействия протоколов при организации межсетевого взаимодействия. В данном случае нас будут интересовать сервер, присоединенный к глобальной

Сервер ATM, рабочая станция, присоединенная к локальной сети IEEE 802, и маршрутизатор, соединяющий названные сети<sup>3</sup>. Маршрутизатор обеспечивает соединение между сервером и рабочей станцией, которое позволяет данным конечным системам игнорировать детали связующих сетей.

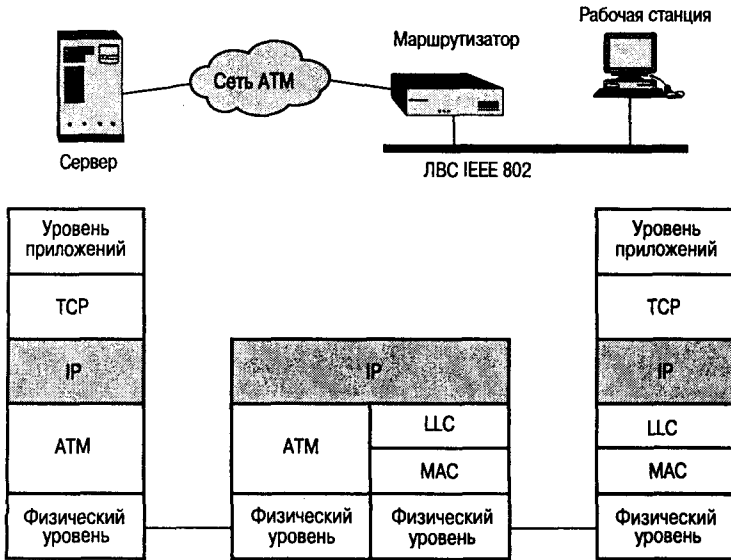


Рис. 4.7. Пример конфигурации для иллюстрации использования TCP/IP

На рис. 4.8–4.10 отображены типичные этапы передачи блока данных, такого, как файл или Web-страница, от сервера через internet к приложению на рабочей станции. В приведенном примере сообщение проходит всего через один маршрутизатор. Перед тем как данные можно будет передавать, приложение и транспортный уровень сервера устанавливают (совместно с соответствующим уровнем на рабочей станции) приемлемые базовые правила сеанса связи. В число вопросов, подлежащих согласованию, входят: используемый знаковый код, метод выявления ошибок и т.п. С этой целью на каждом уровне используется соответствующий протокол, после чего происходит передача сообщения.

<sup>3</sup> Протокольная архитектура IEEE 802 описана в разделе 14.1. На данный момент достаточно сказать, что архитектура состоит из физического уровня, уровня управления доступом к среде (MAC), на котором решаются вопросы адресации и защиты от ошибок, и уровня управления логическим каналом (LLC), на котором рассматриваются логические соединения и определяется пользователь LLC.

### 1. Подготовка данных

Протокол уровня приложений подготавливает для передачи блок данных. Например, сообщение электронной почты (SMTP), файл (FTP) или блок пользовательского ввода (TELNET)

### 2. Использование общего синтаксиса

Если требуется, данные конвертируются в форму, приемлемую для адресата. Это может включать иной код символов, использование шифрования и/или сжатия

### 3. Сегментация данных

TCP может разбить блок данных на несколько сегментов, которые будут упорядочиваться приемником. Каждый сегмент TCP включает заголовок с порядковым номером и контрольную последовательность кадра для выявления ошибок

### 4. Дублирующиеся сегменты

С каждого сегмента TCP снимается копия на случай потери или повреждения сегмента при передаче. Копия удаляется после получения от другого объекта TCP подтверждения об успешном приеме сегмента

### 5. Фрагментация сегментов

IP может разбить сегмент TCP на несколько дейтаграмм, чтобы согласовать размер пакетов данных с требованиями промежуточной сети. Каждая дейтаграмма содержит заголовок с адресом назначения, контрольной последовательностью кадра и другую управляющую информацию

### 6. Кадрование

К каждой дейтаграмме IP добавляет заголовок ATM, в результате чего образуется ячейка ATM. Заголовок включает идентификатор соединения и поле защиты от ошибок в заголовке

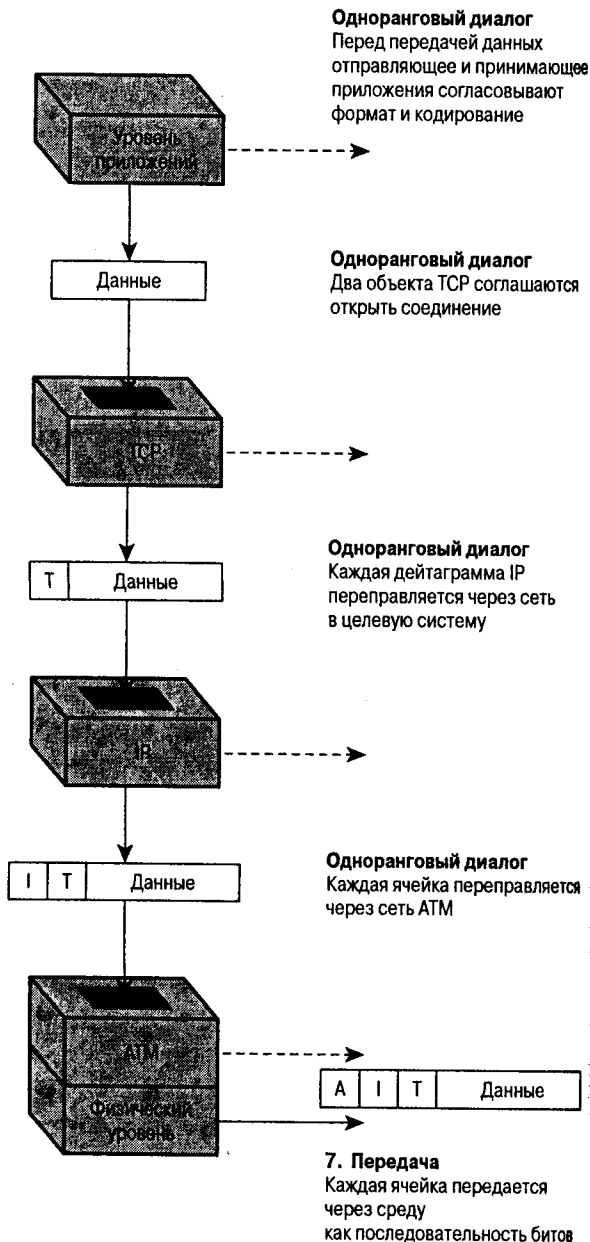


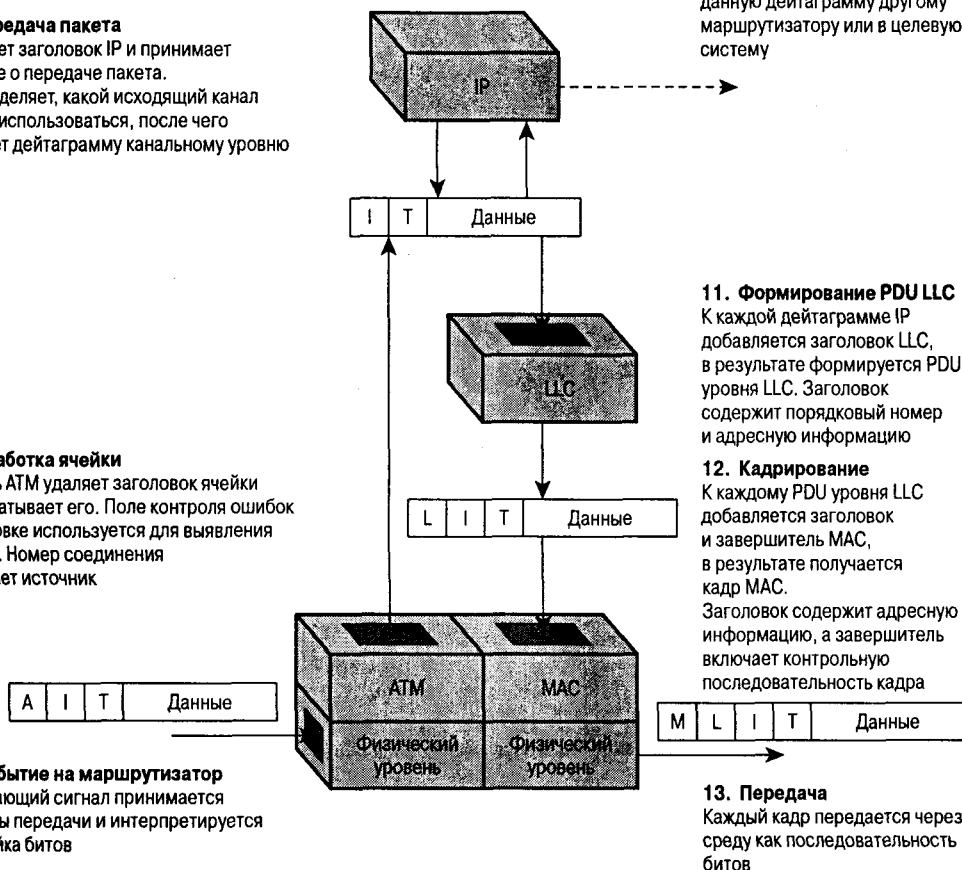
Рис. 4.8. Работа TCP/IP: действия на стороне отправителя

**10. Передача пакета**  
 IP изучает заголовок IP и принимает решение о передаче пакета. Он определяет, какой исходящий канал должен использоваться, после чего передает дейтаграмму канальному уровню

**9. Обработка ячейки**  
 Уровень ATM удаляет заголовок ячейки и обрабатывает его. Поле контроля ошибок в заголовке используется для выявления ошибок. Номер соединения указывает источник

**8. Прибытие на маршрутизатор**  
 Поступающий сигнал принимается из среды передачи и интерпретируется как ячейка битов

**Одноранговый диалог**  
 Маршрутизатор передает данную дейтаграмму другому маршрутизатору или в целевую систему



**11. Формирование PDU LLC**  
 К каждой дейтаграмме IP добавляется заголовок LLC, в результате формируется PDU уровня LLC. Заголовок содержит порядковый номер и адресную информацию

**12. Кадрирование**  
 К каждому PDU уровня LLC добавляется заголовок и завершитель MAC, в результате получается кадр MAC. Заголовок содержит адресную информацию, а завершитель включает контрольную последовательность кадра

**13. Передача**  
 Каждый кадр передается через среду как последовательность битов

Рис. 4.9. Работа TCP/IP: действия на маршрутизаторе

## 20. Доставка данных

Приложение выполняет необходимые преобразования, в том числе разуплотнение и дешифровку, и направляет данные требуемому файлу или другому адресату

## 19. Сборка пользовательских данных

Если TCP разбил пользовательские данные на несколько сегментов, данные собираются в блок и передаются приложению

## 18. Обработка сегмента TCP

TCP удаляет заголовок, проверяет контрольную последовательность кадра и передает подтверждение при корректности контрольной последовательности. Кроме того, осуществляется управление потоком

## 17. Обработка дейтаграммы IP

IP удаляет заголовок. Обработывается контрольная последовательность кадра и другая управляющая информация

## 16. Обработка PDU уровня LLC

Уровень LLC удаляет заголовок и обрабатывает его. Порядковый номер используется для управления потоком и защиты от ошибок

## 15. Обработка кадра

Уровень MAC удаляет заголовок и завершитель и обрабатывает их. Контрольная последовательность кадра используется для выявления ошибок

## 14. Прибытие по назначению

Входной сигнал принимается из среды передачи и интерпретируется как кадр, состоящий из битов

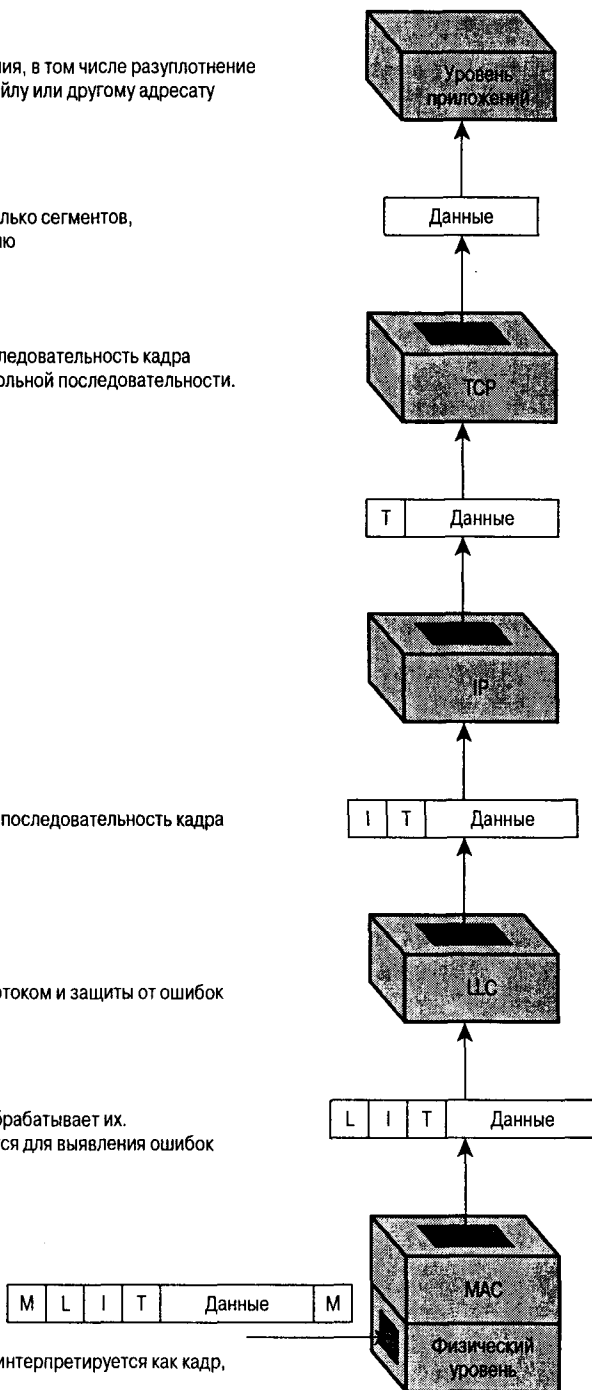


Рис. 4.10. Работа TCP/IP: действия в приемнике

## 4.5. РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

Весьма подробно все темы, упомянутые в данной главе, рассмотрены в книге [STAL00]. Полезным справочником по TCP/IP является [MURH98], где технически точно и основательно рассмотрены протоколы набора TCP/IP.

MURH98 Murhammer M. et al. *TCP/IP: Tutorial and Technical Overview*. — Upper Saddle River, NJ: Prentice Hall, 1998.

STAL00 Stallings W. *Data and Computer Communications, 6th ed.* — Upper Saddle River, NJ: Prentice Hall, 2000 // Столлингс В. *Компьютерные системы передачи данных*, 6-е изд. : Пер. с англ. — М. : Издательский дом “Вильямс”, 2002.

## 4.6. ТЕРМИНЫ, ВОПРОСЫ И ЗАДАЧИ

### Основные термины

internet	межсетевое	протокол управления
Intranet	взаимодействие	передачей (TCP)
IPv6	модуль данных	сетевой уровень
архитектура протоколов	протокола (PDU)	точка доступа к службе
заголовок	порт	(SAP)
контрольная	протокол	транспортный уровень
последовательность	протокол Internet (IP)	уровень приложений
кадра (FCS)	протокол	физический уровень
контрольная сумма	пользовательских	
маршрутизатор	дейтаграмм (UDP)	

### Вопросы

1. Какую основную функцию выполняет уровень доступа к сети?
2. Какие задачи выполняет транспортный уровень?
3. Что такое “протокол”?
4. Что такое “модуль данных протокола” (PDU)?
5. Что такое архитектура протоколов?
6. Что собой представляет TCP/IP?
7. Какие преимущества многоуровневого подхода можно увидеть на примере архитектуры TCP/IP?
8. Что такое “маршрутизатор”?

### Задачи

1. Используя многоуровневую модель, приведенную на рис. 4.11, опишите процессы заказа и доставки пиццы с указанием взаимодействий, происходящих на каждом уровне.
2. а. Премьер-министры Франции и Китая должны провести переговоры по телефону, но ни один из них не знает языка другого. Мало того, ни у одного нет переводчика, владеющего двумя языками. Впрочем, в штате обоих имеются переводчики, владеющие английским языком. Для изображения возникшей ситуации

нарисуйте диаграмму, подобную той, что приведена на рис. 4.11, и опишите взаимодействия, происходящие на каждом ее уровне.

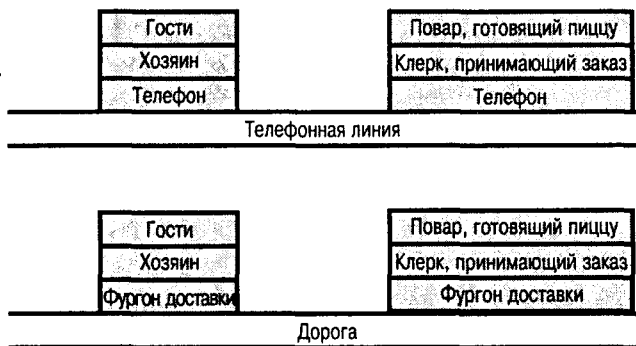


Рис. 4.11. Архитектура для задачи 4.11

6. Предположим, что переводчик китайского премьер-министра может работать только с японским языком, а переводчик французского — только с немецким, при этом в Германии имеется переводчик с японского языка на немецкий и наоборот. Нарисуйте новую диаграмму, на которой была бы отображена схема переговоров, и опишите предполагаемый телефонный разговор.
3. Укажите основные недостатки использования многоуровневого подхода по отношению к протоколам.
4. Две армии “синих”, расположенные на противоположных холмах, готовятся атаковать расположенную в низине армию “красных”. Армия “красных” способна одержать победу над любой из армий “синих” в отдельности, но если они атакуют одновременно, “красные” обречены на поражение. Сообщение между армиями “синих” происходит посредством ненадежной системы связи (пехотинца). Командующий одной из армий “синих” желает атаковать в полдень. Перед ним стоит следующая проблема: если он пошлет сообщение с приказом об атаке, он не может быть уверен, что оно дойдет по назначению. Подтверждение получения приказа также может не дойти. Существует ли протокол, использование которого позволит армиям “синих” избежать поражения?
5. Широковещательная сеть — это сеть, в которой передача одной присоединенной станции принимается всеми станциями, присоединенными к среде. Примеры: локальные сети с шинной топологией, такие, как Ethernet, и беспроводные радиосети. Обоснуйте необходимость либо ненужность сетевого уровня (уровня 3 архитектуры OSI) в широковещательной сети.
6. Среди принципов, использованных ISO при определении уровней OSI, были следующие.
  - Число уровней должно быть достаточно малым, чтобы избежать громоздкости структуры и реализации, но достаточно большим, чтобы функции различных уровней отличались задействованными процессами или технологиями.
  - Границы уровней выбираются по принципу минимизации числа и величины взаимодействий сквозь границы.

Основываясь на этих принципах, спроектируйте архитектуру из восьми уровней и приведите аргументы в ее защиту. Сделайте то же самое для шестиуровневой архитектуры.

7. В ситуации, изображенной на рис 4.2, один PDU  $N$ -го уровня инкапсулируется в PDU уровня  $(N - 1)$ . Существует возможность разбивки одного PDU  $N$ -го уровня на несколько PDU уровня  $(N - 1)$  (процесс сегментации) или объединения нескольких PDU  $N$ -го уровня в один PDU уровня  $(N - 1)$  (процесс объединения в блоки).
  - а. Требуется ли при сегментации наличие в каждом сегменте уровня  $(N - 1)$  копии заголовка  $N$ -го уровня?
  - б. Требуется ли при объединении в блоки, чтобы каждый PDU  $N$ -го уровня имел собственный заголовок, или эти данные можно объединить в единый PDU  $N$ -го уровня с единым заголовком  $N$ -го уровня?
8. Зачем нужен протокол UDP? Почему пользовательская программа не может прямо обращаться к IP?

## ПРИЛОЖЕНИЕ 4А. ПРОТОКОЛ INTERNET

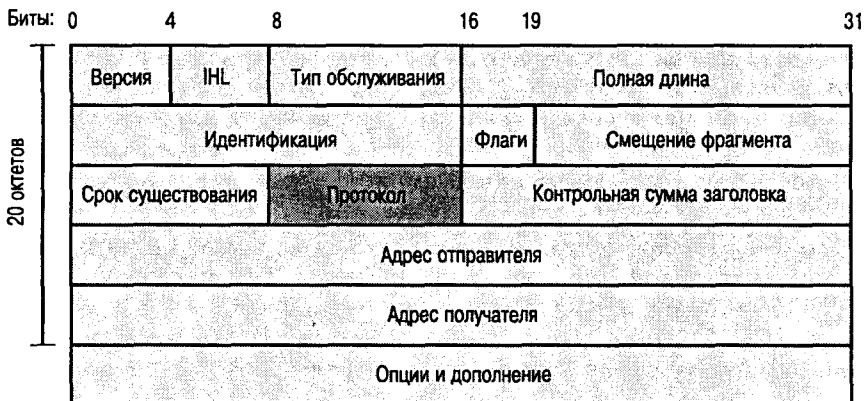
Пожалуй, важнейшим протоколом набора TCP/IP является протокол Internet (IP). В предыдущие десятилетия использовалась версия IPv4. Относительно недавно была стандартизована новая версия, IPv6, пока не получившая широкого распространения.

### IPv4

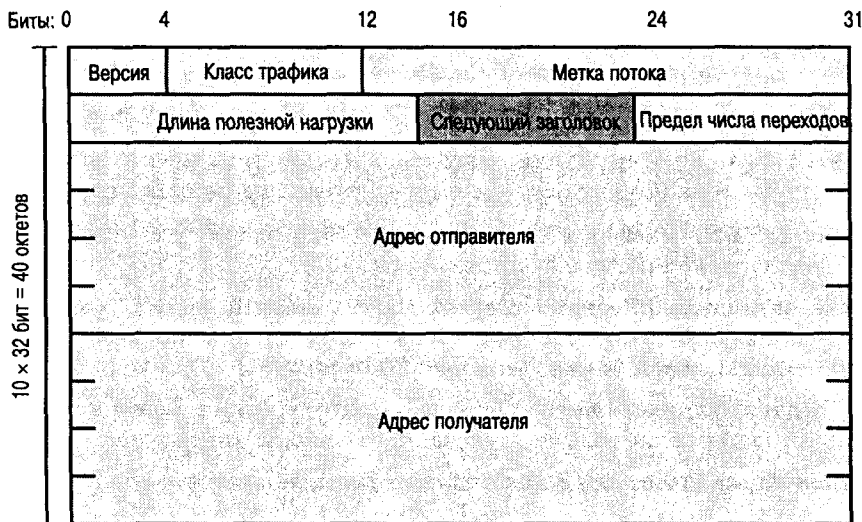
На рис. 4.12, а изображен формат заголовка IP, минимальный размер которого — 20 октетов, или 160 бит. Этот формат содержит следующие поля.

- **Версия (4 бит).** Указывает номер версии, благодаря чему возможно дальнейшее развитие протокола. Значение поля — 4.
- **Длина заголовка internet (Internet Header Length — IHL) (4 бит).** Длина заголовка в 32-разрядных словах. Минимальное значение — 5, что соответствует минимальной длине заголовка 20 октетов.  
**Тип обслуживания (8 бит).** Управление IP-модулями конечной системы и маршрутизаторов посредством указанного приоритета пакета.
- **Полная длина (16 бит).** Полная длина пакета в октетах.
- **Идентификация (16 бит).** Порядковый номер, который вместе с адресом отправителя, адресом получателя и протоколом пользователя уникально идентифицирует пакет. Таким образом, этот номер должен быть уникальным для каждого сочетания адресов отправителя и получателя и протокола пользователя в течение времени, пока пакет остается в объединенной сети.
- **Флаги (3 бит).** В настоящее время определены только два бита. Бит продолжения (More) используется при фрагментации и повторной сборке. Бит “не фрагментировать” (Don't Fragment), когда он установлен, запрещает фрагментацию. Этот бит может быть полезен, если известно, что получатель не может собирать фрагменты. Однако если этот бит установлен, дейтаграмма будет отклоняться, если ее размер превышает максимально допустимый размер пакета в какой-то из подсетей на пути следования. Поэтому при установке бита может понадобиться использование маршрутизации от источника, чтобы обойти подсети с малым максимальным размером пакета.





а) Заголовок IPv4



б) Заголовок IPv6

Рис. 4.12. Заголовки IP

- **Смещение фрагмента (13 бит).** Указывает позицию этого фрагмента в исходном пакете в единицах по 64 бит. Подразумевается, что длина поля данных во всех фрагментах, кроме последнего, должна быть кратна 64 бит.
- **Срок существования (Time to Live — TTL) (8 бит).** Определяет, какое время (в секундах) пакет может оставаться в internet. Каждый маршрутизатор, обрабатывающий пакет, должен уменьшать TTL по крайней мере на единицу, так что TTL несколько напоминает счетчик переходов.

- **Протокол (8 бит).** Указывает протокол вышестоящего уровня, который должен принять поле данных в системе получателя. Следовательно, данное поле идентифицирует тип очередного заголовка в пакете после заголовка IP.
- **Контрольная сумма заголовка (16 бит).** Код обнаружения ошибок, относящийся только к заголовку. Так как значения некоторых полей заголовка (например, срок существования, поля фрагментации) могут изменяться в процессе передачи, эта сумма проверяется и вычисляется заново в каждом маршрутизаторе. Контрольная сумма представляет собой результат 16-разрядного сложения в дополнительном коде всех 16-разрядных слов в заголовке. При вычислении поле контрольной суммы обнуляется.
- **Адрес отправителя (32 бит).** Разнообразные варианты назначения битов позволяют по-разному обозначать сеть и конечную систему, подключенную к указанной сети (7 и 24 бит, 14 и 16 бит, 21 и 8 бит).
- **Адрес получателя (32 бит).** Такие же характеристики, как и в адресе отправителя.
- **Опции (переменная длина).** Закодированные опции, требуемые отправителем, могут включать метку безопасности, маршрутизацию от источника, предопределенную маршрутизацию и временные метки.
- **Дополнение (переменная длина).** Используется для дополнения заголовка пакета до длины, кратной 32 бит.

Поля адреса отправителя и получателя в заголовке IP содержат 32-разрядный глобальный адрес internet, обычно состоящий из идентификатора сети и идентификатора хоста. Адрес допускает различные схемы расположения битов, обозначающих сеть и хост, как показано на рис. 4.13. Такое кодирование обеспечивает гибкость при назначении адресов хостам и позволяет объединять в internet сети разных размеров. Так, для указанных ниже условий лучше всего подходят сети трех классов.

- **Класс А.** Небольшое количество сетей со множеством хостов в каждой.
- **Класс В.** Среднее количество сетей со средним количеством хостов в каждой.
- **Класс С.** Много сетей с небольшим количеством хостов в каждой.

В конкретном окружении может быть целесообразным использование всех адресов, относящихся к одному классу. Например, в корпоративной сети, состоящей из множества локальных сетей отделов, может быть желательным, чтобы все адреса принадлежали к классу С. Однако формат адресов позволяет использовать в одной объединенной сети все три класса адресов одновременно, что и сделано во всемирной Internet. Сочетание разных классов приемлемо для объединенной сети, состоящей из нескольких больших сетей, множества малых сетей и какого-то количества сетей среднего размера.

Адреса IP обычно записываются в так называемой *десятичной нотации с разделительными точками*, в которой каждый октет в 32-разрядном адресе представляется десятичным числом. Например, IP-адрес 11000000 11100100 00010001 00111001 записывается как 192.228.17.57.

Следует заметить, что все адреса сетей класса А начинаются с двоичного 0. Адреса сетей с первым октетом 0 (двоичное 00000000) и 127 (двоичное 01111111) зарезервированы; следовательно, остается 126 потенциально возможных номеров для се-

тей класса А, соответствующих первому в адресе десятичному числу в диапазоне от 1 до 126. Адреса сетей класса В начинаются с двоичного 10; следовательно, первое десятичное число в адресе класса В будет находиться в пределах от 128 до 191 (двоичные 10000000–10111111). Второй октет также входит в адрес сети класса В, поэтому общее количество адресов в классе В составляет  $2^{14} = 16\,384$ . Для адресов класса С первое десятичное число находится в пределах от 192 до 223 (11000000–11011111). Общее количество адресов класса С составляет  $2^{21} = 2\,097\,152$ .

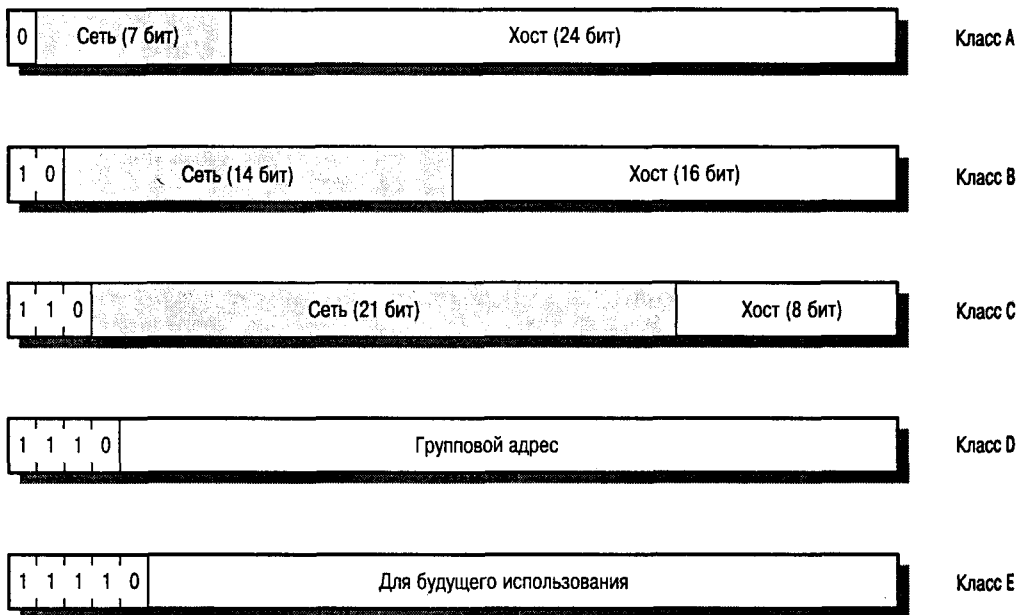


Рис. 4.13. Форматы адресов IP

## IPv6

В 1995 году Проблемная группа проектирования Internet (Internet Engineering Task Force — IETF), занимающаяся разработкой стандартов протоколов для Internet, выпустила спецификацию для протокола IP следующего поколения, известного как IPng (RFC 1752). Данная спецификация послужила основой разработки стандарта, известного как IPv6 (RFC 2460). По сравнению с существующим (IPv4) протоколом IP протокол IPv6 имеет улучшенные функциональные возможности, связанные с адаптацией к более высоким скоростям современных сетей и смешиванию информационных потоков, включая графику и видео. При этом основным мотивом разработки нового протокола была все же потребность в большем количестве адресов. В протоколе IPv4 для идентификации источника и адресата используется 32-битовый адрес. При впечатляющем росте Internet и частных сетей, присоединенных к Internet, этой длины адреса становится недостаточно для предоставления адресов всем нуждающимся системам. Как показано на рис. 4.12, б, заголовок протокола IPv6 включает 128-битовые поля адреса источника и адресата. Ожидается, что в конечном счете все установленное оборудование, которое использует протоколы TCP/IP, перейдет на протокол IPv6, но этот процесс может продолжаться годы, если не десятилетия.

## ПРИЛОЖЕНИЕ 4Б. ПРОТОКОЛ УПРАВЛЕНИЯ ПЕРЕДАЧЕЙ

Большинство приложений, использующих набор TCP/IP, при надежной доставке данных полагаются на протокол TCP; TCP, в свою очередь, полагается на IP в вопросах адресации и маршрутизации.

Ниже описан один из ключевых механизмов TCP — управление потоком. Затем характеризуются функциональные возможности TCP, для чего разбираются элементы заголовка TCP.

### Управление потоком TCP

Как и большинство протоколов, обеспечивающих управление потоком, TCP использует разновидность механизма раздвижного окна. Эта разновидность отличается от той, что применяется в других протоколах, таких, как LLC, HDLC и X.25, тем, что здесь разделяются подтверждение принятых модулей данных и предоставление права передачи следующих модулей.

Механизм управления потоком, используемый TCP, называется схемой кредитов (*credit allocation scheme*). В этой схеме считается, что каждый передаваемый октет данных имеет порядковый номер. Кроме данных, в заголовке каждого переданного сегмента имеются три поля, связанных с управлением потоком: порядковый номер (*SN*), номер подтверждения (*AN*) и окно (*W*). Когда транспортный объект отправляет сегмент, он помещает в поле данных сегмента порядковый номер первого октета. Принимающий транспортный объект подтверждает получение сегмента с помощью обратного сегмента, в котором ( $AN = i, W = j$ ). Это означает следующее.

- Все октеты до порядкового номера  $SN = i - 1$  подтверждены; следующий ожидаемый октет имеет порядковый номер  $AN = i$ .
- Разрешается отправить дополнительное окно из  $W = j$  октетов данных; т.е. октеты, соответствующие порядковым номерам от  $i$  до  $i + j - 1$ .

Этот механизм представлен на рис. 4.14. Для простоты восприятия показан поток данных только в одном направлении, и предполагается, что в каждом сегменте отправляется 200 октетов данных. Вначале, в процессе установки соединения, порядковые номера у отправителя и получателя синхронизируются, и объект А получает начальный кредит на передачу 1400 октетов, начиная с номера 1001. Передав 600 октетов в трех сегментах, А сокращает свое окно до размера 800 октетов (номера с 1601 до 2400). Приняв эти сегменты, объект В подтверждает получение всех октетов до 1600 и выдает кредит на 1000 октетов. Это означает, что А может передать октеты с номерами от 1601 до 2600 (5 сегментов). Однако к моменту, когда сообщение от В достигает объекта А, тот уже передал два сегмента, содержащих октеты 1601–2000 (в рамках начального кредита). Таким образом, оставшийся кредит А в этот момент составляет лишь 600 октетов (3 сегмента). В процессе обмена объект А передвигает задний край своего окна при каждой передаче, но передний край он передвигает, только получая кредит.

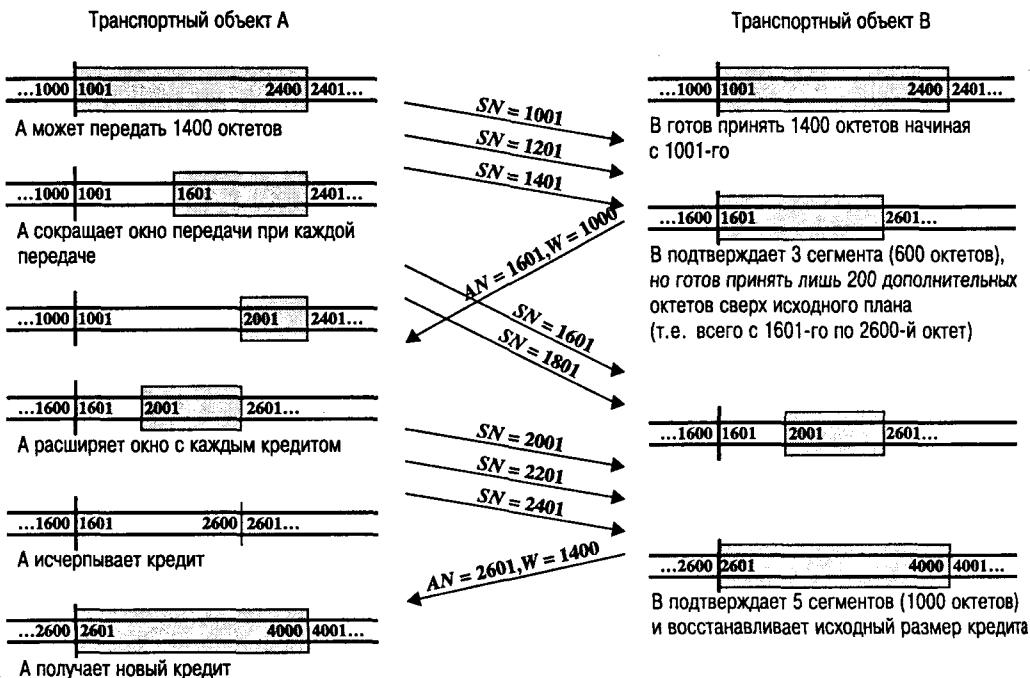


Рис. 4.14. Пример механизма кредитов

Механизм выделения кредитов очень гибок. Предположим, например, что последним объектом В передан сообщение с  $(AN = i, W = j)$  и что последний октет данных, принятый В, имел номер  $i - 1$ . Тогда:

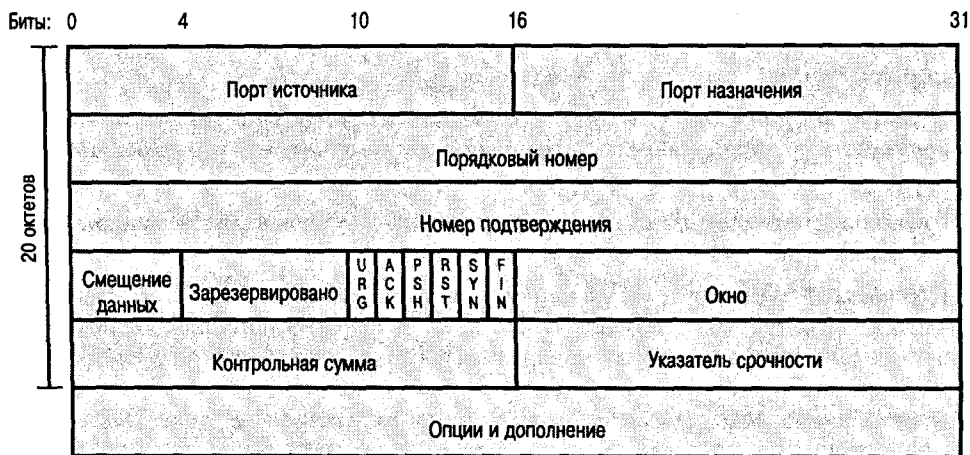
- чтобы увеличить кредит до размера  $k$ , где  $k > j$  без получения новых данных, В передает  $(AN = i, W = k)$ ;
- чтобы подтвердить прибывший сегмент, содержащий  $m$  октетов данных ( $m < j$ ) без выделения дополнительного кредита, В передает  $(AN = i + m, W = j - m)$ .

## Формат сегмента TCP

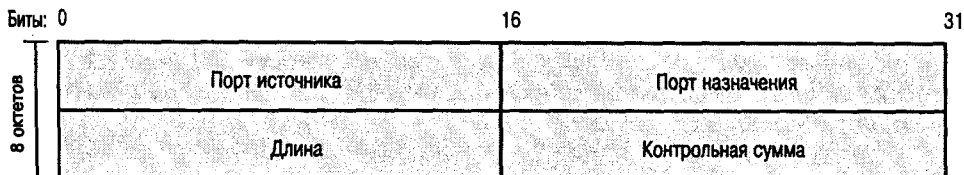
В TCP применяется модуль данных протокола только одного типа, называемый *сегментом TCP*. Формат заголовка показан на рис. 4.15, а. Поскольку все механизмы протокола должны обслуживать один и тот же заголовок, он сделан довольно большим, с минимальной длиной 20 октетов. Заголовок включает следующие поля.

- **Порт источника (16 бит).** Пользователь TCP-отправитель.
- **Порт назначения (16 бит).** Пользователь TCP-получатель.
- **Порядковый номер (32 бит).** Порядковый номер первого октета данных в этом сегменте, за исключением случая, когда установлен флаг SYN. Если SYN установлен, это начальный порядковый номер (initial sequence number — ISN), и первый октет данных имеет номер ISN + 1.

- **Номер подтверждения (32 бит).** Каскадное подтверждение. Содержит порядковый номер следующего октета данных, который предполагает получить объект ТСР.



а) Заголовок TCP



б) Заголовок UDP

Рис. 4.15. Заголовки TCP и UDP

- **Смещение данных (4 бит).** Количество 32-разрядных слов в заголовке.
- **Зарезервировано (6 бит).** Зарезервировано для будущего использования.
- **Флаги (6 бит).**
  - **URG.** Обратить внимание на поле указателя срочности.
  - **ACK.** Обратить внимание на поле подтверждения.
  - **PSH.** Функция продвижения.
  - **RST.** Сброс соединения.
  - **SYN.** Синхронизация порядковых номеров.
  - **FIN.** От отправителя больше не будет данных.

- **Окно (16 бит).** Выделение кредита (в октетах) для управления потоком данных. Содержит число октетов данных, начиная с номера подтверждения, разрешенное для приема получателем.
- **Контрольная сумма (16 бит).** Поразрядное дополнение до единицы суммы по модулю  $2^{16} - 1$  всех 16-разрядных слов в сегменте вместе с псевдозаголовком (см. ниже).
- **Указатель срочности (16 бит).** Указывает на последний октет в последовательности срочных данных. Это позволяет получателю узнать, какое количество срочных данных поступило.
- **Опции (переменная длина).** Например, опция, указывающая максимальный размер сегмента для приема.

*Порядковый номер и номер подтверждения* относятся к отдельным октетам, а не к целым сегментам. Например, если сегмент содержит порядковый номер 1001 и включает 600 октетов данных, порядковый номер относится к первому октету в поле данных; следующий сегмент в логической последовательности будет иметь порядковый номер 1601. Таким образом, по своей логике протокол TCP ориентирован на потоки: он принимает поток октетов от пользователя, группирует их в сегменты, как считает нужным, и нумерует каждый октет в потоке.

Поле *контрольной суммы* вычисляется на основе всего сегмента и псевдозаголовка, добавляемого перед заголовком на время вычисления (как при передаче, так и при приеме). Псевдозаголовок включает следующие поля из заголовка IP: адреса internet и протоколы отправителя и получателя, поле длины сегмента. Добавление псевдозаголовка — это защитная мера TCP от неправильной работы IP. Иначе говоря, если IP доставит сегмент не тому хосту, которому следует, и даже если сегмент не содержит ошибочных битов, принимающий объект TCP все равно обнаружит ошибку доставки.

## ПРИЛОЖЕНИЕ 4В. ПРОТОКОЛ ПОЛЬЗОВАТЕЛЬСКИХ ДЕЙТАГРАММ

Кроме TCP, существует еще один широко используемый протокол транспортного уровня, входящий в набор TCP/IP, — протокол пользовательских дейтаграмм (User Datagram Protocol — UDP), определенный в RFC 768. Протокол UDP предоставляет услуги без установки соединения для процедур уровня приложений. Таким образом, UDP — это в своей основе ненадежная служба; доставка и защита от дубликатов не гарантируются. Однако этот протокол имеет небольшие служебные издержки и во многих случаях может быть вполне приемлемым.

Преимущества ориентированного на соединение подхода понятны. Он предоставляет такие возможности, достижимые прежде всего в соединениях, как управление потоком данных, защита от ошибок и упорядоченная доставка. Однако в некоторых условиях служба без установки соединения оказывается более пригодной. На нижних уровнях (internet, сеть) служба без установки соединения более устойчива. Примером этого может служить использование дейтаграмм при коммутации пакетов (см. главу 3). Если каждый пакет интерпретируется независимо и независимо переправляется через сеть, влияние перегрузки или потерь в

сети меньше, чем при использовании predetermined маршрута, или соединения, по которому следуют все пакеты.

Кроме того, служба без установки соединения — это своего рода “наименьший общий знаменатель” услуг, которые должны предоставляться на более высоких уровнях. Далее, даже на транспортном уровне и выше можно найти аргументы в пользу использования службы без установки соединения. Бывают ситуации, в которых служебные издержки на установку и поддержание соединения оказываются неоправданными или даже неприемлемыми. Некоторые примеры приведены ниже.

- **Внутренний сбор данных.** Включает периодическое активное или пассивное извлечение выборки из источников данных, таких, как датчики, и подачу автоматических сообщений самопроверки от средств обеспечения безопасности или компонентов сети. При контроле в реальном времени потеря однократного модуля данных не имеет серьезных последствий, так как вскоре прибывает очередное сообщение.
- **Внешнее распространение данных.** Включает широкоэвещательную передачу, объявление о новых узлах или об изменении адресов служб и передачу сигналов реального времени.
- **Отклики на запросы.** Приложения, в которых общий сервер предоставляет транзакционные услуги многим распределенным пользователям транспортных служб и для которых типична последовательность “запрос-отклик”. Использование службы регулируется на уровне приложения, и соединения на низком уровне часто оказываются ненужными и громоздкими.
- **Приложения реального времени.** Например, передача голоса и телеметрии, включая требования многократного резервирования и/или передачи в реальном времени. Эти приложения не обязательно должны иметь ориентированные на соединение функции, такие, как повторная передача.

Таким образом, на транспортном уровне имеются “ниши” как для ориентированной на соединение, так и для не требующей установки соединения службы.

Протокол UDP работает над протоколом IP. Поскольку он не требует установки соединения, у UDP очень немного обязанностей. По сути, он всего лишь дополняет IP возможностью адресации портов. Это хорошо видно на примере заголовка UDP, показанного на рис. 4.15, б. Заголовок содержит поля порта источника и порта назначения. Поле длины содержит длину всего сегмента UDP, включая заголовки и данные. Контрольная сумма вычисляется по такому же алгоритму, какой используется в TCP и IP. В UDP контрольная сумма относится ко всему сегменту UDP вместе с псевдозаголовком, добавляемым перед заголовком на время вычисления. Это тот же заголовок, который применяется в TCP. Если обнаруживается ошибка, сегмент отклоняется, и дальнейшие действия не предпринимаются.

Наличие поля контрольной суммы в UDP не обязательно. Если оно не используется, ему присваивается нулевое значение. Однако следует заметить, что контрольная сумма IP относится только к заголовку IP, а не к полю данных, которое в этом случае состоит из заголовка UDP и данных пользователя. Таким образом, если UDP не проверяет контрольной суммы, то данные пользователя никак не проверяются.



**Часть II**

# **ТЕХНОЛОГИЯ БЕСПРОВОДНОЙ СВЯЗИ**

**Глава 5. Антенны и распространение**

**Глава 6. Методы кодирования сигналов**

**Глава 7. Расширенный спектр**

**Глава 8. Кодирование и защита от ошибок**

# ГЛАВА 5

## АНТЕННЫ И РАСПРОСТРАНЕНИЕ

### 5.1. Антенны

Диаграммы направленности

Типы антенн

Коэффициент усиления антенны

### 5.2. Режимы распространения

Огибание поверхности Земли

Отражение от верхних слоев атмосферы

Распространение вдоль линии прямой видимости

### 5.3. Передача сигнала в пределах линии прямой видимости

Затухание

Потери в свободном пространстве

Шум

Отношение  $E_b/N_0$

Атмосферное поглощение

Многолучевое распространение

Преломление

### 5.4. Замирание в мобильных средах

Многолучевое распространение

Методы компенсации ошибок

### 5.5. Рекомендуемая литература

### 5.6. Термины, вопросы и задачи

**Д**анная глава дает возможность получить некоторые фундаментальные знания о процессе радиопередачи. Вначале рассматриваются антенны, после чего внимание будет уделено распространению сигнала.

## 5.1. АНТЕННЫ

Антенну можно определить как проводник (или систему проводников), используемый для излучения или улавливания электромагнитной энергии из пространства. Для передачи сигнала радиочастотные электрические импульсы передатчика с помощью антенны преобразуются в электромагнитную энергию, которая излучается в окружающее пространство (атмосфера, вода, космос). При получении сигнала энергия электромагнитных волн, поступающих на антенну, преобразуется в радиочастотные электрические импульсы, после чего подается на приемник.

Как правило, при двусторонней связи одна и та же антенна может быть использована как для приема, так и для передачи сигнала. Такой подход возможен, потому что любая антенна с равной эффективностью поставляет энергию из окружающей среды к принимающим терминалам и от передающих терминалов в окружающую среду (предполагается, что в обоих случаях используется одна и та же частота). Иными словами, характеристики антенны одинаковы для процессов получения и передачи электромагнитной энергии.

### Диаграммы направленности

Антенны излучают энергию во всех направлениях. Однако в большинстве случаев эффективность передачи сигнала для различных направлений неодинакова. Наиболее распространенным способом определения эффективности антенны является диаграмма направленности, которая представляет собой зависимость излучающих свойств антенны от пространственных координат. Один из наиболее простых типов диаграммы направленности соответствует идеальному случаю так называемой изотропной антенны. Под **изотропной антенной** понимают точку в пространстве, которая излучает энергию одинаково во всех направлениях. Диаграмма направленности для изотропной антенны представляет собой сферу, центр которой совпадает с положением антенны. Как правило, диаграммы направленности антенн представляются как двухмерное поперечное сечение трехмерной диаграммы. Подобная диаграмма для изотропной антенны представлена на рис. 5.1, *а*. Расстояние от антенны до любой точки диаграммы направленности прямо пропорционально энергии, которая была излучена антенной в данном направлении. На рис. 5.1, *б* представлен еще один идеализированный случай — направленная антенна с одним выделенным направлением излучения (вдоль горизонтальной оси).

Размер диаграммы направленности может быть произвольным. Важно лишь, чтобы в каждом направлении были соблюдены пропорции. Чтобы на основе относительного расстояния определить приведенную мощность в заданном направлении, от точки размещения антенны до пересечения с диаграммой направленности проводят прямую линию под соответствующим углом наклона. На рис. 5.1 для двух идеализированных антенн сравниваются два угла передачи сигнала (А и В). Изотропной антенне соответствует ненаправленная круговая диаграмма; векторы А и В равны по величине. Для вибратора Герца вектор В длиннее вектора А. Т.е. в направлении В излучаемая мощность больше, чем в

направлении А. Относительная длина векторов прямо пропорциональна мощности сигнала, излученной в соответствующем направлении.

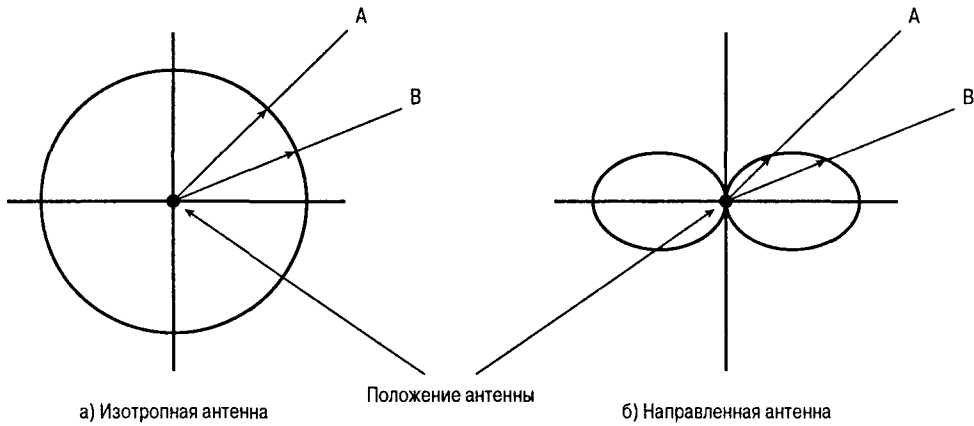


Рис. 5.1. Диаграммы направленности идеализированных антенн

Диаграммы направленности являются удобным средством определения такой меры направленности антенны, как **ширина луча**. Ширина луча (также называемая шириной луча по уровню половинной мощности) — это угол, в пределах которого излучаемая мощность составляет не меньше половины мощности, которая излучается в преимущественном направлении.

Антенна, используемая для приема сигналов, характеризуется **диаграммой приема**. Наибольшие секторы такой диаграммы определяют направление, оптимальное для получения сигналов.

## Типы антенн

### Диполи

Двумя наиболее простыми примерами антенн являются полуволновой диполь (также называемый вибратором Герца, см. рис. 5.2, а) и четвертьволновая вертикальная антенна (антенна Маркони, см. рис. 5.2, б). Полуволновой диполь состоит из двух прямых коллинеарных проводников равной длины, разделенных небольшой щелью, на которую подается сигнал. Максимальная эффективность передачи сигнала достигается при длине антенны, равной половине длины волны. Вертикальная четвертьволновая антенна наиболее часто используется в автомобильных или портативных радиолах.

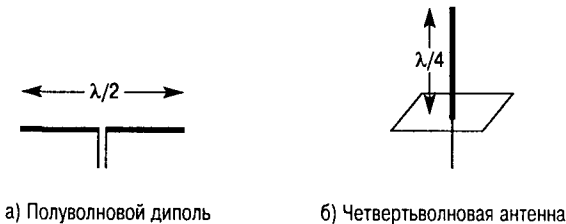


Рис. 5.2. Простейшие антенны

Полуволновой диполь характеризуется ненаправленной диаграммой излучения в одном измерении трехмерного пространства. В двух других плоскостях диаграмма по форме напоминает цифру “8” (см. рис. 5.3, *a*). Для излучения направленных сигналов могут использоваться более сложные типы антенн. Распространенный пример диаграммы направленности такой антенны представлен на рис. 5.3, *b*. В данном случае максимум мощности излучается в направлении оси  $x$ .

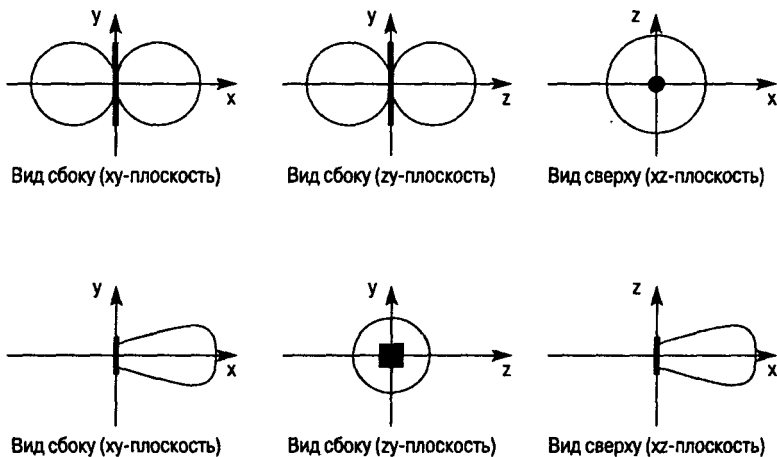
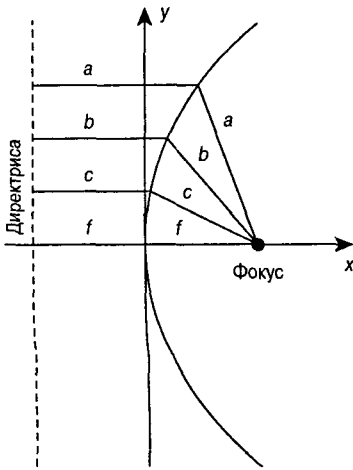


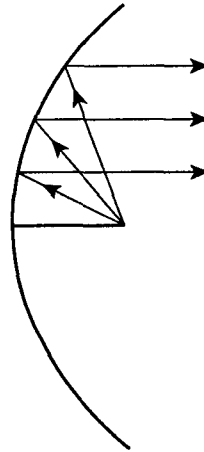
Рис. 5.3. Диаграммы направленности в трехмерном пространстве [SCH100]

## Параболическая отражающая антенна

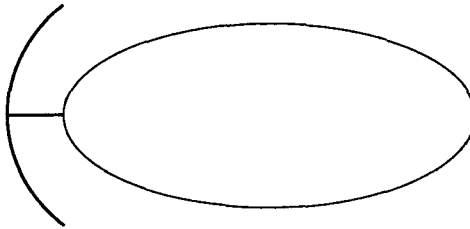
Параболические отражающие антенны используются широко, в частности, для спутниковой связи, а также в наземных системах СВЧ-связи. В школьном курсе геометрии парабола определяется как геометрическое место точек, равноудаленных от некоторой фиксированной прямой линии и фиксированной точки, которая не принадлежит этой линии. Упомянутая фиксированная точка называется *фокусом*; фиксированную линию называют директрисой (рис. 5.4, *a*). В результате обращения параболы вокруг собственной оси получается поверхность, именуемая *параболоидом*. Поперечное сечение параболоида, проведенное параллельно его оси, является параболой, тогда как при сечении, которое перпендикулярно данной оси, образуется окружность. Подобные поверхности используются при создании прожекторов, оптических телескопов и радиотелескопов, а также СВЧ-антенн. Столь широкое их применение объясняется следующим свойством: если источник электромагнитной энергии (или звука) поместить в фокус параболоида и если поверхность параболоида является отражающей, тогда волны будут отражаться параллельно оси параболоида. На рис. 5.4, *b* представлен вид подобного эффекта в поперечном сечении. Теоретически при образовании луча, параллельного оси параболоида, дисперсия отсутствует. Однако на практике дисперсия будет иметь место, поскольку объем источника энергии больше одной точки. Верно также утверждение, обратное приведенному ранее: если принимаемые волны параллельны оси отражающего параболоида, результирующий сигнал будет сконцентрирован в фокусе.



а) Парабола



б) Отражающие свойства параболической антенны (поперечное сечение)



в) Диаграмма излучения параболической антенны (поперечное сечение)

Рис. 5.4. Параболическая отражающая антенна

На рис. 5.4, в представлен типичный пример диаграммы направленности параболической отражающей антенны. В табл. 5.1 приводятся параметры ширины луча для антенн различных размеров при частоте сигнала 12 ГГц. Следует отметить, что чем больше диаметр антенны, тем более сфокусированным будет луч.

Таблица 5.1. Ширина луча для параболических отражающих антенн различного диаметра при частоте сигнала  $f = 12$  ГГц [FREE97]

Диаметр антенны, м	Ширина луча, градусы
0,5	3,5
0,75	2,33
1,0	1,75
1,5	1,166
2,0	0,875
2,5	0,7
5,0	0,35

## Коэффициент усиления антенны

Коэффициент усиления является мерой направленности антенны. Данный параметр определяется как отношение мощности сигнала, излученного в определенном направлении, к мощности сигнала, излучаемого идеальной ненаправленной (изотропной) антенной в любом направлении. Если, например, коэффициент усиления антенны равен 3 дБ, это означает, что ее сигнал сильнее сигнала изотропной антенны в данном направлении на 3 дБ (в 2 раза). Увеличение мощности сигнала в одном направлении возможно лишь за счет остальных направлений распространения. Другими словами, увеличение мощности сигнала в одном направлении влечет за собой уменьшение мощности в других направлениях. Необходимо отметить, что коэффициент усиления характеризует направленность сигнала, а не увеличение выходной мощности по отношению к входной (как может показаться из названия), поэтому данный параметр часто еще называют коэффициентом направленного действия.

Эффективная площадь антенны связана с предыдущим параметром и также зависит от размеров и формы антенны. Отношение между коэффициентом направленного действия антенны и ее эффективной площадью можно записать в следующем виде:

$$G = \frac{4\pi A_e}{\lambda^2} = \frac{4\pi f^2 A_e}{c^2}. \quad (5.1)$$

Здесь

$G$  — коэффициент направленного действия антенны;

$A_e$  — эффективная площадь;

$f$  — несущая частота;

$c$  — скорость света ( $\approx 3 \times 10^8$  м/с);

$\lambda$  — длина волны несущей.

В табл. 5.2 приводятся значения коэффициента усиления антенны и эффективной площади для некоторых распространенных форм антенн.

**Таблица 5.2. Коэффициенты усиления и эффективная площадь некоторых типов антенн [COUC01]**

Тип антенны	Эффективная площадь $A_e$ , м <sup>2</sup>	Коэффициент усиления мощности (относительно изотропной антенны)
Изотропная	$\lambda^2/4\pi$	1
Бесконечно малый диполь или контур	$1,5 \lambda^2/4\pi$	1,5
Полуволновой диполь	$1,64 \lambda^2/4\pi$	1,64
Рупорная антенна, площадь раструба $A_c$	$0,81 A_c$	$10 A_c/\lambda^2$
Параболическая, площадь лицевой поверхности $A_c$	$0,56 A_c$	$7 A_c/\lambda^2$
Турникетная (два перпендикулярно пересекающихся диполя)	$1,15 \lambda^2/4\pi$	1,15

**Пример.** Найти эффективную площадь и коэффициент усиления параболической отражающей антенны диаметром 2 м при частоте сигнала 12 ГГц. Площадь поверхности  $A = \pi r^2 = \pi$ ; эффективная площадь  $A_e = 0,56\pi$ . Длина волны  $\lambda = c/f = (3 \times 10^8)/(12 \times 10^9) = 0,025$  м.

Имеем

$$G = (7A)/\lambda^2 = (7 \times \pi)/(0,025)^2 = 35,186,$$

$$G_{\text{дБ}} = 45,46 \text{ дБ.}$$

## 5.2. РЕЖИМЫ РАСПРОСТРАНЕНИЯ

При распространении сигнал, излученный антенной, может огибать поверхность Земли, отражаться от верхних слоев атмосферы либо распространяться вдоль линии прямой видимости. Табл. 5.3 дает представление о том, какой тип распространения характерен для определенных диапазонов частот. В данной книге в основном рассматривается распространение волн вдоль линии прямой видимости. В настоящем разделе будет приведен краткий обзор других типов распространения.

### Огибание поверхности Земли

При огибании поверхности Земли (см. рис. 5.5, а) путь распространения сигнала в той или иной степени повторяет контур планеты. Передача может производиться на значительные расстояния, намного превышающие пределы прямой видимости. Данный эффект имеет место для частот до 2 МГц. На способность сигналов, принадлежащих данной полосе частот, повторять кривизну земной поверхности влияют несколько факторов. Одним из них является то, что электромагнитная волна индуцирует электрический ток на поверхности Земли; это приводит к замедлению волнового фронта около земной поверхности. В результате волновой фронт отклоняется в сторону земной поверхности и, таким образом, повторяет ее кривизну. Другим фактором является дифракция. Данное явление связано с поведением электромагнитных волн при наличии препятствий.



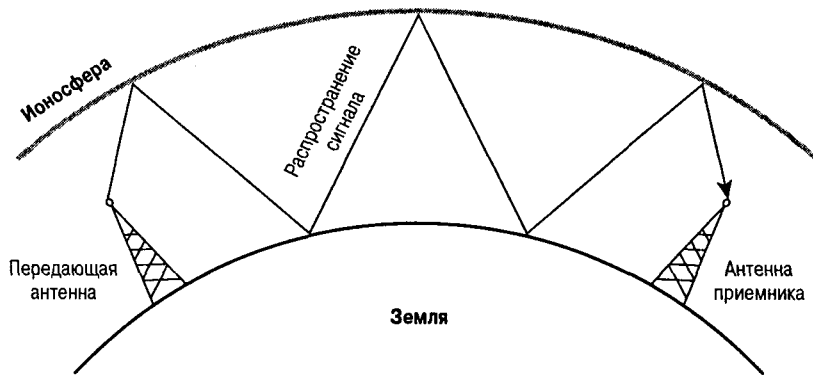
Таблица 5.3. Диапазоны частот

Полоса	Частотный диапазон	Длина волны в вакууме	Особенности распространения	Применение
Крайне низкая частота	30-300 Гц	10 000-1 000 км	Огибает поверхность Земли	Частоты линий электропередач; используются в некоторых домашних системах управления
Частота речевого диапазона	300-3 000 Гц	1 000-100 км	Огибает поверхность Земли	Используются в аналоговых линиях телефонной связи
Очень низкая частота (ОНЧ)	3-30 кГц	100-10 км	Огибает поверхность Земли; постоянный низкий уровень затухания; высокий уровень атмосферных шумов	Навигация дальнего действия; связь между подводными лодками
Низкая частота	30-300 кГц	10-1 км	Огибает поверхность Земли; несколько менее надежна по сравнению с ОНЧ; поглощается в дневное время	Навигация дальнего действия; Морские радиомаяки
Средняя частота	300-3000 кГц	1 000-100 м	Огибает поверхность Земли; отражается от верхних слоев атмосферы в ночное время; значительное затухание днем и ночью; присутствуют атмосферные шумы	Морская радиосвязь; определение направления; радиовещание в АМ-диапазоне
Высокая частота	3-30 МГц	100-10 м	Отражается от верхних слоев атмосферы; качество передачи зависит от времени года и времени суток, а также от частоты	Связь между радиодлюбителями; международное радиовещание; военная связь; связь между морским транспортом и авиатранспортом на больших расстояниях
Очень высокая частота (ОВЧ)	30-300 МГц	10-1 м	Распространяется вдоль линии прямой видимости; рассеивается вследствие температурной инверсии; присутствие космических шумов	Телевидение метрового диапазона; радиовещание в FM-диапазоне; дуплексная радиосвязь; связь с авиатранспортом в АМ-диапазоне; вспомогательная навигация для авиатранспорта
Ультравысокая частота (УВЧ)	300-3 000 МГц	100-10 см	Распространяется вдоль линии прямой видимости; присутствие космических шумов	телевидение УВЧ-диапазона; содовая телефонная связь; радарные установки; радиорелейные линии; персональные системы связи

Полоса	Частотный диапазон	Длина волны в вакууме	Особенности распространения	Применение
Сверхвысокая частота (СВЧ)	3-30 ГГц	10-1 см	Распространяется вдоль линии прямой видимости; затухание во время дождя при частоте свыше 10 ГГц; атмосферное затухание при взаимодействии с кислородом и водным паром	Спутниковая связь; радарные установки; наземные радиорелейные линии; местные линии радиосвязи
Чрезвычайно высокая частота	30-300 ГГц	10-1 мм	Распространяется вдоль линии прямой видимости; атмосферное затухание при взаимодействии с кислородом и водным паром	Экспериментальный диапазон; местные линии радиосвязи
Инфракрасное излучение	300 ГГц-400 ТГц	1 мм-330 нм	Распространяется вдоль линии прямой видимости	Локальные вычислительные сети инфракрасного диапазона; потребительские электронные устройства
Видимый свет	400-900 ТГц	770-380 нм	Распространяется вдоль линии прямой видимости	Оптическая связь



а) Распространение околосемных волн (частота до 2 МГц)



б) Отражение от верхних слоев атмосферы (частота 2-30 МГц)



в) Распространение сигнала вдоль линии визирования (частота свыше 30 МГц)

*Рис. 5.5. Типы распространения радиоволн*

Рассеяние электромагнитных волн указанного диапазона в атмосфере происходит таким образом, что в верхние атмосферные слои эти волны не попадают.

Наиболее распространенным примером связи с использованием волн, огибающих поверхность Земли, является радиовещание в АМ-диапазоне.

## Отражение от верхних слоев атмосферы

Радиоволны, отражающиеся от верхних слоев атмосферы, используются для любительской, персональной и служебной радиосвязи, а также для международного радиовещания (именно такие волны используют радиокomпании Голос Америки и Би-Би-Си). При рассматриваемом типе распространения сигнал наземной антенны отражается от ионизированных слоев верхней части атмосферы, именуемых ионосферой, по направлению к Земле. Хотя может показаться, что отражение радиоволн от ионосферы аналогично отражению от твердой поверхности, на самом деле данный эффект вызван преломлением волн. Суть эффекта преломления будет объяснена позднее.

Как показано на рис. 5.5, б, отражение сигнала от верхних слоев атмосферы (и впоследствии от Земли) может происходить многократно. Если радиоволна распространяется подобным образом, сигнал может получаться на расстоянии тысяч километров от передатчика.

## Распространение вдоль линии прямой видимости

Если частота радиосигнала превышает 30 МГц, то огибание им земной поверхности и отражение от верхних слоев атмосферы становятся невозможными. В этом случае связь должна осуществляться в пределах прямой видимости (рис. 5.5, в). При связи через спутник сигнал с частотой свыше 30 МГц не будет отражаться ионосферой. Такой сигнал может передаваться от наземной станции к спутнику и обратно при условии, что спутник не находится за пределами горизонта. При наземной связи передающая и принимающая антенны должны находиться в пределах *эффективной* линии прямой видимости. Использование термина "эффективный" связано с тем, что волны сверхвысокой частоты искривляются и преломляются атмосферой. Степень и направление искривления зависят от различных факторов. Однако, как правило, искривления сверхвысокочастотных волн повторяют кривизну поверхности Земли. Поэтому такие волны распространяются на расстояние, превышающее оптическую линию прямой видимости.

## Преломление

Перед тем как продолжить изложение материала, приведем обещанное описание явления преломления волн. Преломление происходит из-за того, что скорость распространения электромагнитной волны является функцией плотности той среды, через которую проходит волна. В вакууме электромагнитные волны (например, свет или радиосигналы) распространяются со скоростью, приблизительно равной  $3 \times 10^8$  м/с. Данную постоянную величину "с" принято называть скоростью света, хотя строго имеется в виду скорость света в вакууме. В воздухе, воде, стекле или другой прозрачной или частично прозрачной среде электромагнитные волны распространяются со скоростью ниже с.

Скорость распространения электромагнитной волны меняется при переходе между веществами разной плотности. В результате происходит разовое преломление направления распространения волны на границе раздела двух сред. Пример подобного преломления приводится на рис. 5.6. При переходе из менее плотной среды в более плотную волна изменяет направление распространения в сторону более плотного вещества. Данное явление можно наблюдать при частичном погружении стержня в воду. Наблюдаемый результат будет аналогичным

тому, что показан на рис. 5.6, т.е. стержень будет казаться изогнутым и более коротким, чем на самом деле.

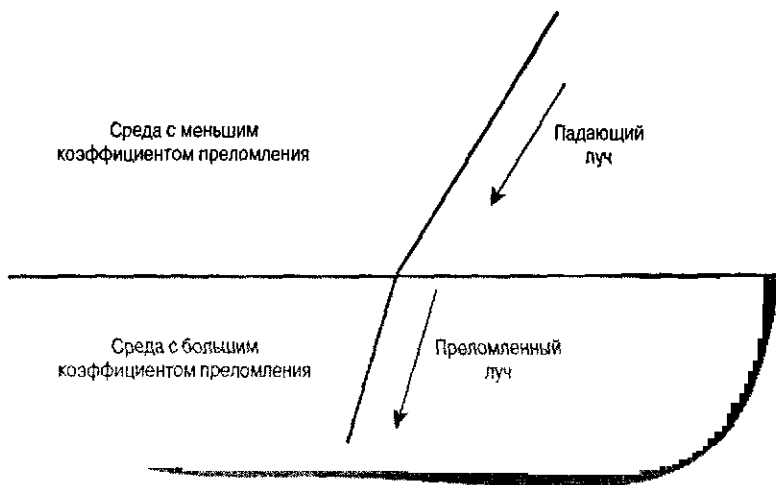


Рис. 5.6. Преломление электромагнитной волны [POOL98]

Коэффициент преломления при переходе из одной среды в другую определяется как отношение синуса угла падения к синусу угла преломленного луча. Этот коэффициент также можно выразить как отношения скоростей распространения волны в различных средах. Абсолютный коэффициент преломления в определенной среде вычисляется по отношению к вакууму. Сигналам с различной длиной волны соответствуют разные коэффициенты преломления. Таким образом, эффект преломления зависит от длины волны сигнала.

На рис. 5.6 представлено резкое, одномоментное изменение направления распространения сигнала при переходе из одной среды в другую. Если сигнал проходит через вещество, коэффициент преломления которого изменяется постепенно, отклонение от изначального направления распространения будет также происходить постепенно. При нормальных условиях коэффициент преломления атмосферы уменьшается с увеличением высоты. Таким образом, по мере приближения к поверхности Земли скорость распространения радиосигнала уменьшается. В результате происходит незначительное искривление направления распространения радиоволн, поступающих в сторону земной поверхности.

### Оптическая линия прямой видимости и линия радиовидимости

При отсутствии помех или препятствий распространению (оптическую) линию прямой видимости можно представить в следующем виде:

$$d = 3,57\sqrt{h}.$$

Здесь  $d$  — расстояние между антенной и видимым горизонтом в километрах,  $h$  — высота антенны в метрах. Эффективная линия прямой видимости (другими словами, линия радиовидимости) по отношению к видимому горизонту определяется следующим выражением (см. рис. 5.7):

$$d = 3,57\sqrt{Kh}.$$



Рис. 5.7. Видимый горизонт для оптических волн и радиоволн

Здесь  $K$  — поправочный коэффициент для учета преломления. Хорошим эмпирическим результатом является  $K = 4/3$ . Таким образом, максимальное расстояние между двумя антеннами для распространения сигнала в пределах линии прямой видимости составляет  $3,57(\sqrt{Kh_1} + \sqrt{Kh_2})$ , где  $h_1$  и  $h_2$  — высоты соответствующих антенн.

**Пример.** Для радиосвязи (предполагается распространение радиоволн в пределах прямой видимости) используются две антенны. Высота одной из них составляет 100 м, другая находится на уровне земли. Максимальное расстояние между двумя антеннами будет следующим:

$$d = 3,57\sqrt{Kh} = 3,57\sqrt{133} = 41 \text{ км.}$$

Предположим, что высота принимающей антенны составляет 10 м. Какой должна быть высота передающей антенны, чтобы сохранить прежнее значение максимально допустимого расстояния?

$$41 = 3,57(\sqrt{Kh_1} + \sqrt{13,3});$$

$$\sqrt{Kh_1} = \frac{41}{3,57} - \sqrt{13,3} = 7,84;$$

$$h_1 = 7,84^2/1,33 = 46,2 \text{ м.}$$

Таким образом, высота передающей антенны может быть уменьшена более чем на 50 м. Данный пример демонстрирует выигрыш в необходимой высоте передающей антенны при поднятии над уровнем земли принимающей антенны.

## 5.3. ПЕРЕДАЧА СИГНАЛА В ПРЕДЕЛАХ ЛИНИИ ПРЯМОЙ ВИДИМОСТИ

Для любой системы связи справедливо утверждение, что принимаемый сигнал отличается от переданного сигнала. Данный эффект является следствием различных искажений в процессе передачи. При передаче аналогового сигнала искажения приводят к его случайному изменению, что проявляется в ухудшении качества связи. Если же передаются цифровые данные, искажения приводят к появлению двоичных ошибок — двоичная единица может преобразоваться в нуль и наоборот. В этом разделе будут рассмотрены различные типы искажений, а также их влияние на пропускную способность каналов связи. В данной книге рассматривается передача радиосигналов в пределах линии прямой видимости. В этом контексте наиболее важными являются следующие типы искажений:

- затухание и амплитудное искажение сигнала;
- потери в свободном пространстве;
- шум;
- атмосферное поглощение;
- искажение вследствие многолучевого распространения;
- преломление.

## Затухание

При передаче сигнала в любой среде его интенсивность уменьшается с расстоянием. В изотропной среде такое ослабление, или затухание, в общем случае логарифмически зависит от расстояния. Как правило, для изотропной среды затухание можно выразить как постоянную потерю интенсивности (в децибелах) на единицу длины. Для неизотропной среды затухание может выражаться более сложной функцией расстояния и состава атмосферы. При рассмотрении затухания для инженера-связиста важны три фактора.

1. Полученный сигнал должен обладать мощностью, достаточной для его обнаружения и интерпретации приемником.
2. Чтобы при получении отсутствовали ошибки, мощность сигнала должна поддерживаться на уровне, в достаточной мере превышающем шум.
3. При повышении частоты сигнала затухание возрастает, что приводит к искажению.

Первые два фактора связаны с затуханием интенсивности сигнала и использованием усилителей или ретрансляторов. Для двухточечного канала связи мощность сигнала передатчика должна быть достаточной для четкого приема. В то же время интенсивность сигнала не должна быть слишком большой, так как в этом случае контуры передатчика или приемника могут оказаться перегруженными, что также приведет к искажению сигнала. Если расстояние между приемником и передатчиком превышает определенную постоянную, свыше которой затухание становится неприемлемо высоким, для усиления сигнала в заданных точках пространства располагаются ретрансляторы или усилители. Задача усиления сигнала значительно усложняется, если существует множество приемников, особенно если расстояние между ними и передающей станцией непостоянно.

Третий фактор списка известен как амплитудное искажение. Вследствие того, что затухание является функцией частоты, полученный сигнал искажается по сравнению с переданным, что снижает четкость приема. Относительная мощность частотных компонентов полученного и переданного сигналов неодинакова. Для устранения этой проблемы используются методы выравнивания искажения в определенной полосе частот. Одним из возможных подходов может быть использование устройств, усиливающих высокие частоты в большей мере, чем низкие.

## Потери в свободном пространстве

Для любого типа беспроводной связи передаваемый сигнал рассеивается по мере его распространения в пространстве. Следовательно, мощность сигнала, принимаемого антенной с постоянной эффективной площадью, будет умень-

шаться по мере удаления от передающей антенны. Для спутниковой связи упомянутый эффект является основной причиной снижения интенсивности сигнала. Даже если предположить, что все прочие причины затухания и ослабления отсутствуют, переданный сигнал будет затухать по мере распространения в пространстве. Причина этого — распространение сигнала по все большей площади. Данный тип затухания называют потерями в свободном пространстве и вычисляют через отношение мощности излученного сигнала  $P_i$  к мощности полученного сигнала  $P_r$ . Для вычисления того же значения в децибелах следует взять десятичный логарифм от указанного отношения, после чего умножить полученный результат на 10. Для идеальной изотропной антенны потери в свободном пространстве составляют

$$\frac{P_i}{P_r} = \frac{(4\pi d)^2}{\lambda^2} = \frac{(4\pi f d)^2}{c^2}.$$

Здесь

- $P_i$  — мощность сигнала передающей антенны;
- $P_r$  — мощность сигнала, поступающего на антенну приемника;
- $\lambda$  — длина волны несущей;
- $d$  — расстояние, пройденное сигналом между двумя антеннами;
- $c$  — скорость света ( $3 \times 10^8$  м/с).

Следует отметить, что параметры  $d$  и  $\lambda$  выражаются в одних и тех же единицах (например, в метрах).

Приведенное выражение можно записать в следующем виде:

$$\begin{aligned} L_{\text{вб}} &= 10 \lg \frac{P_i}{P_r} = 20 \lg \left( \frac{4\pi d}{\lambda} \right) = -20 \lg(\lambda) + 20 \lg(d) + 21,98 \text{ дБ} = \\ &= 20 \lg \left( \frac{4\pi f d}{c} \right) = 20 \lg(f) + 20 \lg(d) - 147,56 \text{ дБ}. \end{aligned} \quad (5.2)$$

На рис. 5.8 приводится зависимость потерь сигнала в свободном пространстве от пройденного расстояния<sup>1</sup>.

Для других типов антенн следует учитывать коэффициент усиления. В результате уравнение для потерь мощности сигнала в свободном пространстве принимает следующий вид:

$$\frac{P_i}{P_r} = \frac{(4\pi)^2 (d)^2}{G_r G_i \lambda^2} = \frac{(\lambda d)^2}{A_r A_i} = \frac{(cd)^2}{f^2 A_r A_i}.$$

Здесь

- $G_r$  — коэффициент усиления передающей антенны;
- $G_i$  — коэффициент усиления антенны приемника;
- $A_r$  — эффективная площадь передающей антенны;
- $A_i$  — эффективная площадь антенны приемника.

<sup>1</sup> Как отмечалось в приложении 2А, в литературе существует некоторое несоответствие в применении терминов “усиление” (gain) и “потери” (loss). При записи уравнения (5.2) использовались те же договоренности, что приводились для уравнения (2.2).



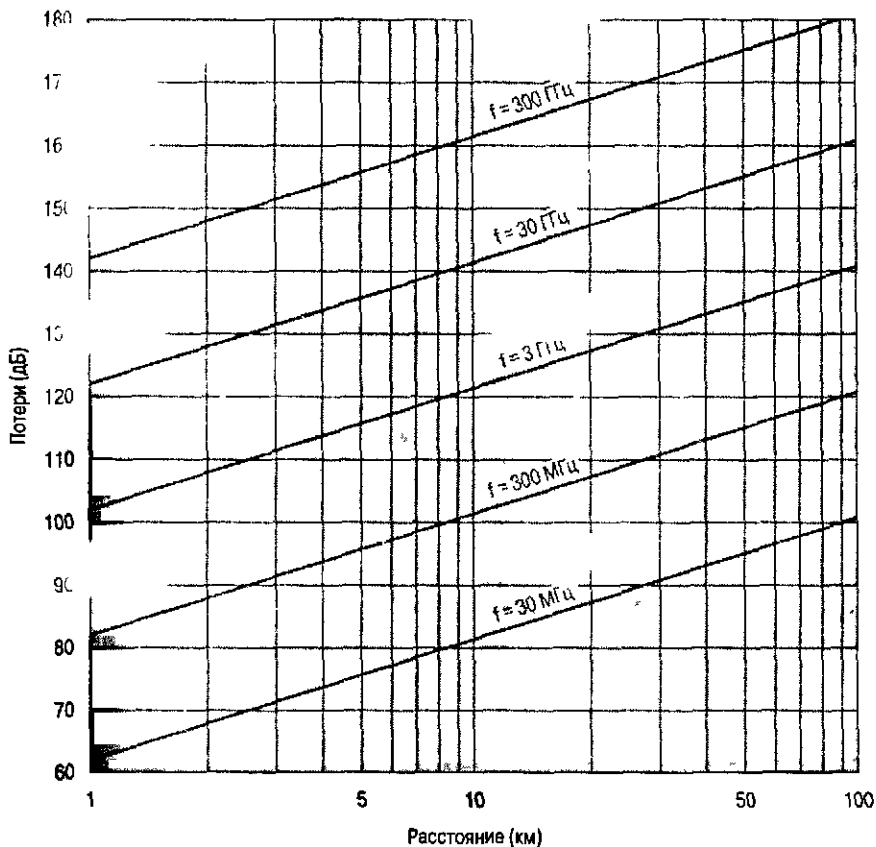


Рис. 5.8. Потери мощности сигнала в свободном пространстве

Переход от второй дроби к третьей выполнен с учетом связи коэффициента усиления антенны и ее эффективной площади, которая была определена в уравнении (5.1). Выражение для потерь в свободном пространстве можно переписать в следующем виде:

$$\begin{aligned}
 L_{\text{дБ}} &= 20 \lg(\lambda) + 20 \lg(d) - 10 \lg(A_e A_r) = \\
 &= -20 \lg(f) + 20 \lg(d) - 10 \lg(A_e A_r) + 169,54 \text{ дБ}.
 \end{aligned}
 \tag{5.3}$$

Следовательно, если размеры антенн и их разнесение в пространстве остаются неизменными, увеличение длины несущей волны (снижение несущей частоты  $f$ ) приводит к увеличению потерь в свободном пространстве. Сравним уравнения (5.2) и (5.3). Из уравнения (5.2) следует, что по мере возрастания частоты растут и потери в свободном пространстве, а потому при высоких частотах радиосигнала потери становятся значительным препятствием для связи. Однако проанализировав уравнение (5.3), можно сказать, что потери легко компенсировать, увеличивая коэффициент усиления антенны. Действительно, при работе на высоких частотах усиление увеличивается, тогда как остальные факторы, влияющие на качество связи, остаются неизменными. Из уравнения (5.2) следует, что при фиксированном расстоянии между приемником и передатчиком уве-

личение частоты приводит к возрастанию потерь в свободном пространстве на величину  $20 \lg f$ . Однако если учесть коэффициент усиления антенны, а также считать ее эффективную площадь неизменной, потери мощности сигнала в свободном пространстве составят  $-20 \lg f$ . Следовательно, при использовании более высоких частот потери мощности сигнала снижаются.

**Пример.** Найдем потери мощности сигнала в свободном пространстве для изотропной антенны. Частота несущей равна 4 ГГц, передача производится с наземной станции на синхронизированный спутник по кратчайшей траектории (35 863 км). При частоте несущей 4 ГГц длина волны равна  $(3 \times 10^8)/(4 \times 10^9) = 0,075$  м. Тогда

$$L_{\text{дб}} = -20 \lg (0,075) + 20 \lg (35,853 \times 10^6) + 21,98 = 195,6 \text{ дБ.}$$

Учтем коэффициент усиления антенн спутника и наземной станции. В большинстве случаев этот параметр составляет 44 и 48 дБ, соответственно. Вычисляем потери в свободном пространстве.

$$L_{\text{дб}} = 195,6 - 44 - 48 = 103,6 \text{ дБ.}$$

Будем считать, что мощность передающей антенны наземной станции равна 250 Вт. Какова мощность сигнала, полученного антенной спутника? Мощность антенны в 250 Вт соответствует 24 дБВт переданного сигнала. Следовательно, мощность, полученная антенной приемника, равна  $24 - 103,6 = -79,6$  дБВт.

## Шум

Для любой передачи данных справедливо утверждение, что полученный сигнал состоит из переданного сигнала, модифицированного различными искажениями, которые вносятся самой системой передачи, а также из дополнительных нежелательных сигналов, взаимодействующих с исходной волной во время ее распространения от точки передачи к точке приема. Эти нежелательные сигналы принято называть шумом. Шум является основным фактором, ограничивающим производительность систем связи.

Шумы можно разделить на четыре категории:

- тепловой шум;
- интермодуляционные шумы;
- перекрестные помехи;
- импульсные помехи.

**Тепловой шум** является результатом теплового движения электронов. Данный тип помех оказывает влияние на все электрические приборы, а также на среду передачи электромагнитных сигналов. Тепловой шум является функцией температуры и равномерно распределен по спектру частот, поэтому данный тип шума называют также белым шумом. Тепловой шум устранить нельзя, поэтому именно он определяет верхний предел производительности систем связи. Тепловой шум оказывает значительное влияние на спутниковые системы связи, поскольку сигнал, получаемый наземной станцией от спутника, достаточно слаб.

Тепловой шум, присутствующий в полосе шириной 1 Гц, для любого устройства или проводника составляет

$$N_0 = kT \text{ (Вт/Гц).}$$

Здесь<sup>2</sup>

$N_0$  — плотность мощности шумов в ваттах на 1 Гц полосы;

$k$  — постоянная Больцмана,  $k = 1,3803 \times 10^{-23}$  Дж/К;

$T$  — температура в Кельвинах (абсолютная температура).

**Пример.** Комнатная температура ( $T$ ) обычно принимается равной  $17^\circ\text{C}$ , или  $290$  К. При данной температуре плотность мощности белого шума составляет

$$N_0 = (1,3803 \times 10^{-23}) \times 290 = 4 \times 10^{-21} \text{ Вт/Гц} = -204 \text{ дБВт/Гц.}$$

Определение единицы измерения децибел-ватт (дБВт) дается в приложении 2А.

Считается, что шум не зависит от частоты. Следовательно, тепловой шум, присутствующий в полосе диапазона  $B$  Гц, можно выразить следующим образом:

$$N = kTB.$$

Запишем данное выражение, используя децибел-ватты:

$$\begin{aligned} N &= 10 \lg k + 10 \lg T + 10 \lg B = \\ &= -228,6 \text{ дБВт} + 10 \lg T + 10 \lg B. \end{aligned}$$

**Пример.** Для приемника с шириной полосы  $10$  МГц и эффективной температурой шума  $294$  К уровень теплового шума на выходе составит:

$$\begin{aligned} N &= -228,6 \text{ дБВт} + 10 \lg(294) + 10 \lg 10^7 = \\ &= -228,6 + 24,7 + 70 = \\ &= -133,9. \end{aligned}$$

Если сигналы разной частоты передаются в одной среде, может иметь место интермодуляционный шум. Интермодуляционным шумом являются помехи, возникающие на частотах, которые представляют собой сумму, разность или произведение частот двух исходных сигналов. Например, смешивание двух сигналов, передаваемых на частотах  $f_1$  и  $f_2$ , соответственно, может привести к передаче энергии на частоте  $f_1 + f_2$ . При этом данный паразитный сигнал может интерферировать с сигналом связи, передаваемым на частоте  $f_1 + f_2$ .

Интермодуляционный шум возникает вследствие нелинейности приемника, передатчика или же промежуточной системы передачи. Как правило, все указанные компоненты ведут себя как линейные системы, т.е. их выходная мощность равна входной мощности, умноженной на некоторую константу. Для нелинейных систем выходная мощность является более сложной функцией входной мощности. Нелинейность может быть вызвана неисправностью одной из деталей, использованием сигнала чрезмерной мощности или же просто природой используемого усилителя. Для указанных случаев помехи возникают на частотах, являющихся суммой или разностью частот исходных сигналов.

---

<sup>2</sup> Джоуль (Дж) является единицей измерения электрической, механической или тепловой энергии в системе СИ. Единица мощности системы СИ, ватт (Вт), соответствует одному джоулю в секунду. Кельвин (К) — единица термодинамической температуры в системе СИ. Температуре  $T$  Кельвин соответствует  $T - 273,15$  градусов Цельсия.

С **перекрестными помехами** сталкивался каждый, кто во время использования телефона параллельно слышал разговор посторонних людей. Данный тип помех возникает вследствие нежелательного объединения трактов передачи сигналов. Такое объединение может быть вызвано сцеплением близко расположенных витых пар или (значительно реже) линий коаксиального кабеля, по которым передаются множественные сигналы. Перекрестные помехи могут возникать во время приема посторонних сигналов антеннами СВЧ-диапазона. Несмотря на то что для указанного типа связи используют высокоточные направленные антенны, потерь мощности сигнала во время распространения избежать все же невозможно. Как правило, мощность перекрестных помех равна по порядку (или ниже) мощности теплового шума. Впрочем, в нелицензируемом диапазоне ISM (Industrial, Scientific and Medical Radio Frequency Band — частоты для промышленного, научного и медицинского применения) перекрестные помехи, как правило, доминируют.

Все указанные выше типы помех являются предсказуемыми (в разумных пределах) и характеризуются относительно постоянным уровнем мощности. Таким образом, вполне возможно спроектировать систему передачи сигнала, которая была бы устойчивой к указанным помехам. Однако кроме вышеперечисленных типов помех существуют так называемые **импульсные помехи**, которые по своей природе являются прерывистыми и состоят из нерегулярных импульсов или кратковременных шумовых пакетов с относительно высокой амплитудой. Причин возникновения импульсных помех может быть множество, в том числе внешние электромагнитные воздействия (например, молнии) или дефекты (поломки) самой системы связи.

Как правило, отрицательное влияние импульсных помех на процесс аналоговой передачи данных незначительно. Например, при передаче голосового сигнала может появляться потрескивание или щелчки, не сильно влияющие на разборчивость передаваемой информации. В то же время при передаче цифровых данных импульсные помехи — это основной источник ошибок. К примеру, посторонний импульс длительностью 0,01 с никак не повлияет на голосовой сигнал, однако при передаче данных со скоростью 56 Кбит/с появление такого импульса будет означать потерю 560 бит информации.

## Отношение $E_b/N_0$

В главе 2 приводилось определение отношения сигнал/шум (SNR). Для определения скорости передачи цифровых данных и уровня ошибок обычно используется другой параметр, связанный с данным отношением. Именно этот параметр — отношение энергии сигнала на 1 бит к плотности мощности шумов на 1 герц ( $E_b/N_0$ ) — используется в качестве меры производительности цифровых систем связи. Рассмотрим цифровой или аналоговый сигнал, содержащий двоичные цифровые данные, передаваемые с определенной скоростью —  $R$  бит/с. Напомним, что  $1 \text{ Вт} = 1 \text{ Дж/с}$ , и вычислим удельную энергию одного бита сигнала:  $E_b = ST_b$  (где  $S$  — мощность сигнала;  $T_b$  — время передачи одного бита). Скорость передачи данных  $R$  можно выразить в виде  $R = 1/T_b$ . Следовательно,

$$\frac{E_b}{N_0} = \frac{S/R}{N_0} = \frac{S}{kTR}$$

Отношение, выраженное в децибелах, принимает следующий вид:

$$\begin{aligned} \left( \frac{E_b}{N_0} \right)_{\text{дБ}} &= S_{\text{дБВт}} - 10 \lg R - 10 \lg k - 10 \lg T = \\ &= S_{\text{дБВт}} - 10 \lg R + 228,6 \text{ дБВт} - 10 \lg T. \end{aligned}$$

Отношение  $E_b/N_0$  имеет большое практическое значение, поскольку скорость появления ошибочных битов является (убывающей) функцией данного отношения. При известном значении  $E_b/N_0$ , требуемом для получения желаемого уровня ошибок, можно выбирать все прочие параметры в приведенном уравнении. Необходимо отметить, что для сохранения требуемого значения  $E_b/N_0$  при повышении скорости передачи данных  $R$  потребуется увеличивать мощность передаваемого сигнала по отношению к шуму.

Попробуем интуитивно понять полученный результат, вернувшись к примеру, представленному на рис. 2.9. Применим подход, использованный при рассмотрении передачи аналогового сигнала, к передаче цифровых данных. Довольно часто уровень мощности шума достаточен для изменения значения одного из битов данных. Если же увеличить скорость передачи данных вдвое, биты будут “упакованы” в два раза плотнее, и тот же посторонний сигнал приведет к потере двух битов информации. Следовательно, при неизменной мощности сигнала и шума увеличение скорости передачи данных влечет за собой возрастание уровня возникновения ошибок.

Преимуществом использования  $E_b/N_0$  в качестве характеристики систем связи по сравнению параметром SNR является зависимость последнего от ширины полосы.

**Пример.** Рассмотрим метод кодирования сигнала, для которого необходимо, чтобы отношение  $E_b/N_0$  равнялось 8,4 дБ при частоте возникновения ошибок  $10^{-4}$  (ошибочным является один бит из каждых 10 000). Если эффективная температура теплового шума равна 290 К (комнатная температура), а скорость передачи данных — 2 400 бит/с, какой должна быть мощность сигнала, чтобы преодолеть тепловой шум?

$$\begin{aligned} 8,4 &= S_{\text{дБВт}} - 10 \lg 2400 + 228,6 \text{ дБВт} - 10 \lg 290 = \\ &= S_{\text{дБВт}} - (10)(3,38) + 228,6 - (10)(2,46) \\ S &= -161,8 \text{ дБВт}. \end{aligned}$$

Параметры  $E_b/N_0$  и SNR можно связать между собой. Имеем

$$\frac{E_b}{N_0} = \frac{S}{N_0 R}.$$

Параметр  $N_0$  является плотностью мощности шума (Вт/Гц). Следовательно, шум в сигнале с шириной полосы  $B_T$  можно записать как  $N = N_0 B_T$ , так что получаем следующее:

$$\frac{E_b}{N_0} = \frac{S B_T}{N R}. \quad (5.4)$$

Еще одним важным уравнением является связь  $E_b/N_0$  со спектральной эффективностью. В главе 2 приводилось полученное Шенноном выражение для максимальной пропускной способности канала (бит/с).

$$C = B \log_2 (1 + S/N).$$

Здесь  $C$  — пропускная способность канала (бит/с);  $B$  — ширина полосы канала (Гц). Уравнение можно переписать в следующей форме:

$$\frac{S}{N} = 2^{C/B} - 1.$$

Используя уравнение (5.4) и заменяя  $B_T$  на  $B$  и  $R$  на  $C$ , соответственно, получим следующее:

$$\frac{E_b}{N_0} = \frac{B}{C} (2^{C/B} - 1).$$

Данная полезная формула представляет связь достижимой спектральной эффективности  $C/B$  и  $E_b/N_0$ .

**Пример.** Найти минимальное значение  $E_b/N_0$ , необходимое для получения спектральной эффективности 6 бит/с/Гц. Имеем  $E_b/N_0 = (1/6)(2^6 - 1) = 10,5 = 10,21$  дБ.

## Атмосферное поглощение

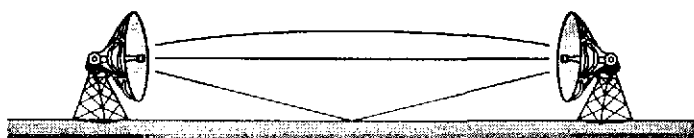
Причиной дополнительных потерь мощности сигнала между передающей и принимающей антеннами является атмосферное поглощение, при этом основной вклад в ослабление сигнала вносят водные пары и кислород. Максимальные потери мощности сигнала наблюдаются вблизи точки 22 ГГц, что вызвано взаимодействием с водяными парами. Для частот ниже 15 ГГц потери мощности немного меньше. Пик потерь мощности вследствие взаимодействия электромагнитных волн с кислородом наблюдается вблизи точки 60 ГГц; действие данного фактора ослабляется при частотах ниже 30 ГГц. Дождь и туман (капли воды, находящиеся во взвешенном состоянии в воздухе) приводят к рассеиванию радиоволн и, в конечном счете, к ослаблению сигнала. Указанные факторы могут быть основной причиной потерь мощности сигнала. Следовательно, в областях, для которых характерно значительное выпадение осадков, необходимо либо сокращать расстояние между приемником и передатчиком, либо использовать для связи более низкие частоты.

## Многолучевое распространение

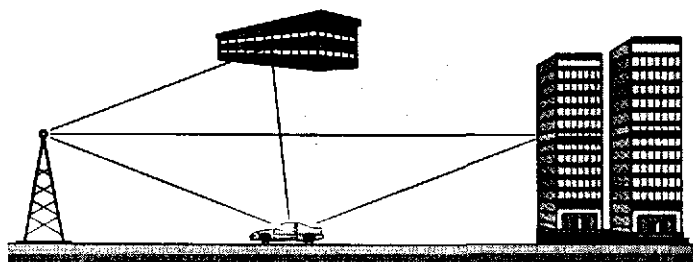
Если расположение антенн в системе связи является относительно произвольным, может случиться так, что тракт связи совпадет с линией прямой видимости между передатчиком и приемником (при отсутствии препятствий, приводящих к интерференции). Как правило, именно такое расположение выбирается для спутниковых систем или двухточечных сверхвысокочастотных каналов связи. Противоположным примером может служить мобильная телефонная связь — в данном случае препятствия на пути распространения сигнала встречаются довольно часто. Препятствия могут отражать сигнал, что приводит к появлению нескольких его копий, поступающих с различными интервалами задержки. В некоторых крайних случаях исходный сигнал может вообще отсутствовать. В зависимости от разницы между траекториями распространения прямой волны и ее отраженных копий мощность первой может

быть меньше или больше суммы мощностей отраженных сигналов. Для систем связи стационарных антенн с удачно выбранным расположением, а также связи спутник и неподвижной наземной станции может осуществляться усиление исходного сигнала и отсеечение побочных компонентов, вызванных многолучевым распространением. Единственным исключением является прохождение сигнала над поверхностью воды, отражающая поверхность которой постоянно колеблется из-за ветра. Для систем мобильной телефонной связи, а также для антенн с неудачным расположением эффекты многолучевого распространения могут быть основным фактором, влияющим на качество сигнала.

На рис. 5.9 схематически изображены основные виды многолучевого распространения, типичные для неподвижных сверхвысокочастотных устройств, а также для наземных и мобильных систем связи. В первом из упомянутых случаев кроме распространения вдоль линии прямой видимости сигнал может преломляться атмосферой. Могут также присутствовать сигнальные компоненты, вызванные отражением от земли. При мобильной связи отражающую поверхность представляют здания и другие объекты местности.



а) Сверхвысокочастотная линия визирования



б) Мобильная радиосвязь

Рис. 5.9. Примеры многолучевого распространения

## Преломление

При распространении в атмосфере радиоволны преломляются (или изгибаются). Преломление является результатом изменения скорости волны по мере изменения высоты или же происходит вследствие других изменений состояния атмосферы. При нормальных условиях скорость сигнала увеличивается с высотой, что приводит к его "изгибанию" по направлению к земле. Однако в некоторых случаях погодные условия могут приводить к таким изменениям скорости распространения в зависимости от высоты, которые значительно отличаются от обычных флуктуаций. В итоге только часть волны, передаваемой вдоль линии прямой видимости, достигнет антенны приемника. Возможна также ситуация, при которой сигнал не будет получен вообще.

## 5.4. ЗАМИРАНИЕ В МОБИЛЬНЫХ СРЕДАХ

Пожалуй, наиболее сложной проблемой, с которой может столкнуться инженер-связист, является замирание сигнала в системах мобильной связи. Термином *замирание* (fading) обозначается изменение мощности полученного сигнала во времени, вызванное изменением тракта связи или среды распространения. Для неподвижных устройств связи замирание является следствием изменений в атмосфере, например наличием или отсутствием дождя. В мобильной связи, где две антенны движутся относительно друг друга, это приводит к тому, что относительное расположение различных препятствий меняется со временем, вызывая сложные эффекты при передаче сигнала.

### Многолучевое распространение

На рис. 5.10 изображены три механизма распространения электромагнитных волн, которые оказывают влияние на многолучевое распространение сигнала. **Отражение** имеет место, когда электромагнитная волна сталкивается с препятствием, размеры которого значительно превышают длину волны. Рассмотрим в качестве примера сигнал, отраженный от земли вблизи мобильного устройства, работающего как приемник. Отражение от поверхности земли приводит к фазовому сдвигу электромагнитной волны на  $180^\circ$ . В результате может наблюдаться взаимная интерференция прямой и отраженной волны, что приводит к значительным потерям мощности полученного сигнала<sup>3</sup>. Кроме того, антенна мобильного устройства меньше большинства искусственных структур в рассматриваемой области, так что происходит многолучевая интерференция, которая может оказывать как положительное, так и отрицательное влияние на принимаемый сигнал.

**Дифракция** происходит на краю объектов, непроницаемых для электромагнитного излучения, размеры которых значительно больше длины волны. При поступлении на край такого объекта радиоволны начинают распространяться в разные стороны. При этом точку, в которой происходит дифракция, можно рассматривать как источник излучения. В результате сигнал может быть получен приемником, который находится вне линии прямой видимости передатчика.

Если препятствие имеет размеры меньше или порядка длины волны, происходит **рассеяние**. Исходный сигнал разделяется на несколько более слабых сигналов. Для сверхвысоких частот, которые обычно используются в сотовой связи, рассеяние может вызываться множеством объектов, например фонарными столбами или дорожными знаками. Такие эффекты рассеяния предсказать сложно.

Три указанных эффекта распространения по-разному влияют на производительность системы в зависимости от особенностей местности и перемещения мобильного устройства внутри ячейки. Если мобильное устройство и ретранслятор находятся в пределах прямой видимости, влияние дифракции и рассеяния невелико, хотя влияние отражения может быть значительным. Если же линия прямой видимости отсутствует (например, на улицах города) сигнал принимается в основном благодаря дифракции и рассеянию.

---

<sup>3</sup> С другой стороны, отраженный сигнал волны имеет большую траекторию распространения, что дает разность фаз вследствие временной задержки отраженного сигнала относительно прямого. Если временная задержка равна половине периода волны, оба сигнала снова оказываются в фазе.



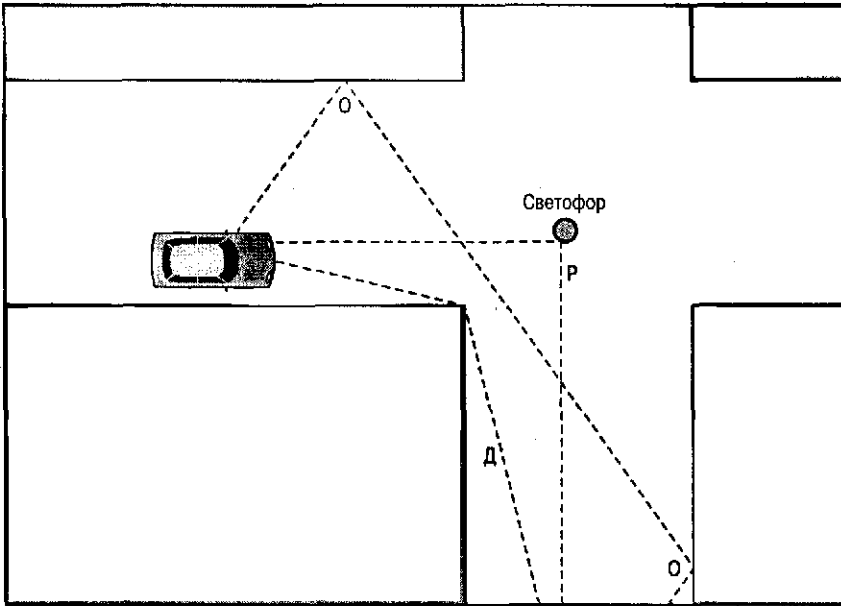
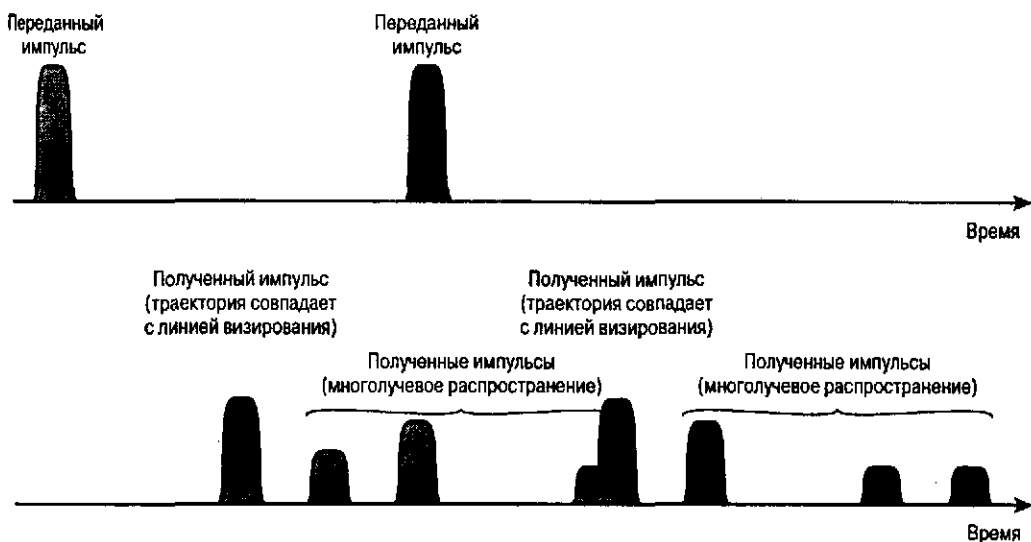


Рис. 5.10. Важные типы механизмов распространения: отражение (O), рассеяние (P), дифракция (D) [ANDE95]

### Эффекты многолучевого распространения

Как отмечалось, одним из нежелательных эффектов многолучевого распространения является разность фаз различных копий сигнала, поступающих на антенну приемника. В результате сложения уровень мощности сигнала по отношению к шуму может снизиться, что усложняет процесс распознавания сигнала приемником.

Другим явлением, важным при передаче цифровых данных, является межсимвольная интерференция (intersymbol interference — ISI). Рассмотрим передачу узконаправленного импульсного сигнала данной частоты по каналу между неподвижной антенной и мобильным устройством. На рис. 5.11 показано, что может доставить канал, если импульс передается в два разных момента времени. На верхней части рисунка представлены два импульса в момент передачи. В нижней части изображены сигналы, поступающие на приемник. В обоих случаях первый принятый импульс — это желаемый сигнал, пришедший по линии прямой видимости. Мощность этого импульса может меняться в зависимости от степени поглощения атмосферой. По мере удаления мобильного устройства от неподвижной антенны атмосферное поглощение прямого сигнала будет возрастать. Однако кроме прямого импульса могут получаться вторичные сигналы, образованные вследствие отражения, дифракции или рассеяния. Предположим, что рассматриваемый импульс содержит один или более закодированных битов информации. В этом случае одна или больше запоздалых копий импульса могут приниматься одновременно с основным импульсом следующего бита. По отношению к следующему импульсу такие запаздывающие импульсы являются шумом, затрудняющим восстановление данных.



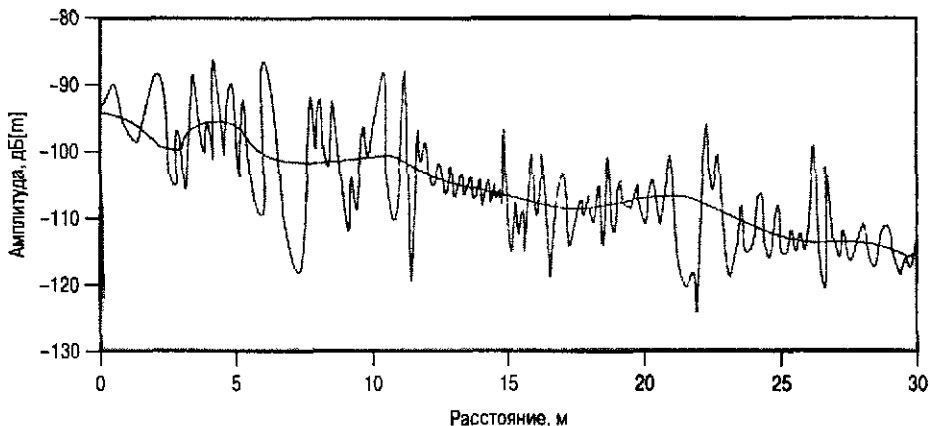
*Рис. 5.11. Два импульса при нестационарном многолучевом распространении*

При перемещении антенны мобильного устройства изменяется относительное положение различных препятствий. Соответственно, изменяются число, величина и время поступления вторичных импульсов. Это затрудняет создание методов обработки сигналов, которые бы позволяли отфильтровать эффекты многолучевого распространения и точно восстановить искомым сигнал.

### Типы замирания

Замирание сигнала в мобильных средах может быть быстрым и медленным. В примере, приведенном на рис. 5.10, по мере передвижения мобильного устройства по городской улице мощность сигнала достаточно быстро меняется при изменении положения на расстояние порядка половины длины волны. При использовании типичной для мобильных систем связи частоты 900 МГц длина волны сигнала составляет 0,33 м. Временная зависимость, приведенная на рис. 5.12, является примером пространственного изменения амплитуды полученного сигнала с частотой 900 МГц в городской среде. Следует отметить, что колебания амплитуды могут составлять от 20 до 30 дБ при незначительном изменении положения. Данное явление быстрого изменения амплитуды называют быстрым замиранием. Этот эффект оказывает влияние не только на мобильные телефоны, находящиеся в движущихся автомобилях, но также на телефоны, используемые во время ходьбы по улице.

Как правило, пользователи мобильной связи перемещаются на расстояния, значительно превосходящие длину волны сигнала. При этом меняется окружающая обстановка — человек может проходить мимо зданий разной высоты, площадей, перекрестков и т.д. При перемещении на большие расстояния происходят изменения уровня средней полученной энергии относительно уровня быстро меняющихся флуктуаций. Эти изменения показаны на рис. 5.12 с помощью плавно убывающей кривой и называются медленным замиранием.



*Рис. 5.12. Типичный пример медленного и быстрого замирания сигнала мобильного устройства в городских условиях*

Эффекты замирания можно также разделить на селективные и амплитудные. Амплитудным (неселективным) замиранием называют пропорциональные и одновременные флуктуации всех частотных компонентов полученного сигнала. Селективное замирание оказывает различное воздействие на спектральные компоненты радиосигнала. Термин *селективное замирание* обычно применяется по отношению ко всей полосе канала связи. Если затухание происходит в некоторой части полосы сигнала, говорят об селективном замирании; при амплитудном замирании рассматриваемая полоса уже полосы спектра, подверженного замиранию, и полностью находится в ней.

### Канал с замиранием

При проектировании систем связи одной из важных задач инженера является оценка влияния шума и замирания, происходящего вследствие многолучевого распространения, на каналы мобильной связи. Простейшей моделью с точки зрения анализа является канал с аддитивным белым гауссовым шумом (Additive White Gaussian Noise — AWGN). В таком канале качество передаваемого сигнала ухудшается в результате присутствия теплового шума, связанного с принимающими и передающими контурами (включая любые промежуточные ретрансляторы или усилители), а также под влиянием физических свойств самого канала. В некоторых случаях такая модель является достаточно точной (космическая связь; некоторые виды передачи посредством проводников, таких, как коаксиальный кабель). В то же время при рассмотрении наземных систем связи, в частности мобильных устройств, модель AWGN будет не лучшим выбором для инженера.

Релеевское замирание происходит тогда, когда среди нескольких прямых траекторий распространения сигналов нет ни одной доминирующей (например, траектории вдоль линии прямой видимости). Такой случай представляет собой наихудший возможный сценарий. К счастью, релеевское замирание можно рассмотреть аналитически, что позволяет получить представление о характеристиках системы в неблагоприятных условиях (например, в плотно застроенных городских районах).

Райсовское замирание характерно для ситуации, когда кроме нескольких не прямых траекторий распространения сигнала существует также траектория распространения по линии прямой видимости. Часто такая модель может применяться для устройств, находящихся в помещении, а модель Релея больше подходит для описания открытых сред. Кроме того, модель райсовского замирания можно использовать при рассмотрении “более открытого” пространства или небольших ячеек мобильной связи. В качестве характеристики канала связи применяется коэффициент  $K$ , определенный следующим образом:

$$K = \frac{\text{мощность доминирующего сигнала}}{\text{мощность отраженных сигналов}}.$$

Если  $K = 0$  (т.е. числитель равен нулю), канал считают релеевским. Канал AWGN характеризуется значением  $K = \infty$  (знаменатель равен 0). На рис. 5.13 (на основе [FREE98a] и [SKLA01]) иллюстрируется производительность системы связи при наличии шума. В данном случае уровень битовых ошибок приводится как функция  $E_b/N_0$ , и по мере роста данного параметра уровень ошибок падает. Из графика следует, что при достаточно высокой мощности сигнала относительно шума канал AWGN имеет довольно хорошую производительность. При больших значениях  $K$ , приблизительно соответствующих открытому пространству в сельской местности или микроячейке связи, райсовский канал также характеризуется хорошей производительностью. Такой уровень производительности достаточен для передачи оцифрованных голосовых сигналов, хотя при передаче данных требуются дополнительные затраты для повышения качества связи. Производительность релеевского канала является относительно неудовлетворительной, что видно из графика для амплитудного замирания и медленного замирания, и в таких случаях необходимо использовать схемы исправления ошибок. И последнее, в некоторых средах возможен так называемый наихудший релеевский сценарий. Примерами могут служить быстрое замирание в городской среде, а также замирание сигнала, который принадлежит полосе селективного замирания. В подобных случаях, каким бы ни было отношение  $E_b/N_0$ , приемлемого качества связи можно достичь только при использовании методов компенсации ошибок, которые рассматриваются ниже.

## Методы компенсации ошибок

Методы компенсации ошибок и искажений, вносимых замиранием вследствие многолучевого распространения, делятся на три основные категории: прямое исправление ошибок, адаптивное выравнивание и разнесение. Как правило, для исправления ошибок в мобильной радиосвязи используются одновременно все три метода.

### Прямое исправление ошибок

Этот метод применяется в цифровых системах связи с радиопередачей цифровых данных, оцифрованного звука или видео. Слово “прямое” означает, что приемник исправляет битовые ошибки, основываясь только на информации входящего цифрового сигнала. Такой метод противоположен обратному исправлению ошибок, когда приемник определяет наличие ошибок, после чего к передатчику отсылается запрос на повторную передачу ошибочно принятых данных. Такой подход к исправлению ошибок является непрактичным для множества

систем радиосвязи. К примеру, при спутниковой связи процесс повторной передачи нежелателен ввиду большой длительности передачи. Для мобильной связи уровень ошибок, как правило, настолько высок, что повторно переданный блок данных, скорее всего, также будет содержать ошибки. В подобных системах нужно использовать прямое исправление ошибок. Опишем суть этого метода.

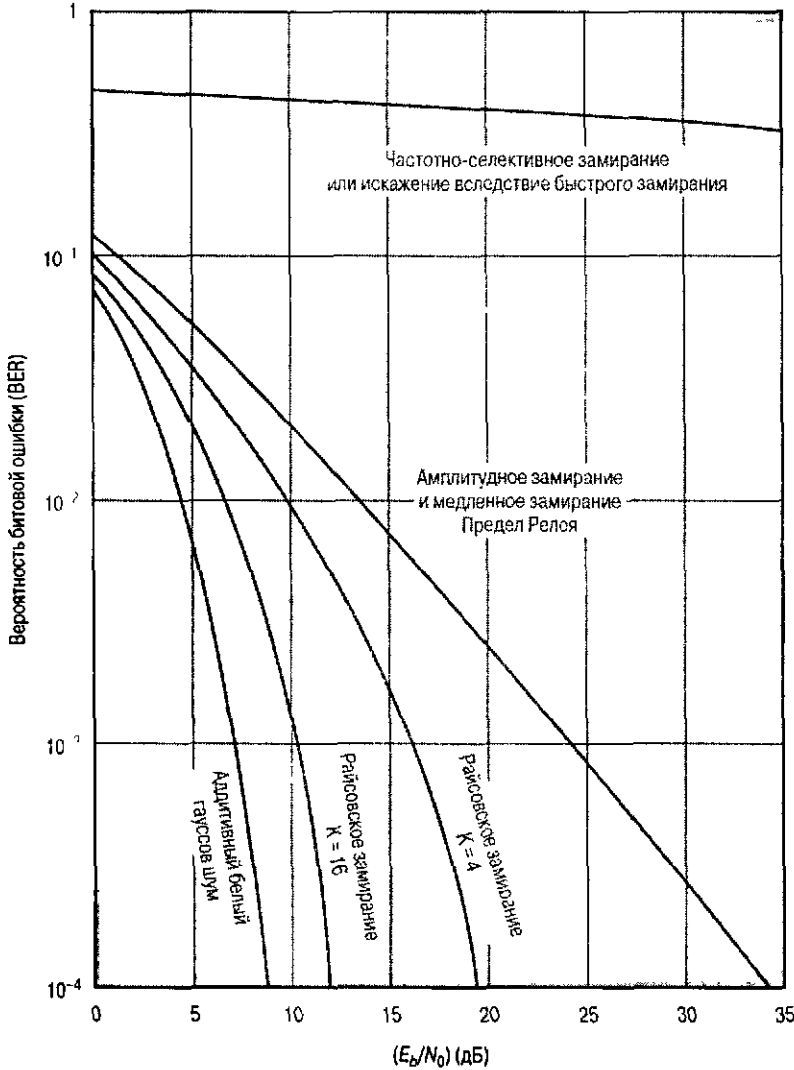


Рис. 5.13. Теоретическая зависимость уровня ошибок от  $E_b/N_0$  для различных условий замирания

1. В каждый из блоков данных передатчик включает определенное число дополнительных избыточных битов. Эти биты вычисляются как функция битов данных и составляют код коррекции ошибок.

2. Для каждого полученного блока битов (биты данных и код коррекции ошибок) приемник, используя биты данных, вычисляет новый код коррекции ошибок. Если рассчитанный код совпадает с полученным, считается, что в данном блоке ошибок нет.
3. Несовпадение полученного и рассчитанного кодов означает наличие ошибки в одном или нескольких битах. Если число ошибочных битов меньше порогового значения (которое зависит от используемого алгоритма и длины кода), приемник может определить их местонахождение и исправить ошибки.

Как правило, отношение полного числа отосланных битов к числу битов данных для систем мобильной радиосвязи составляет порядка 2–3. Может показаться, что описанный подход требует слишком больших служебных издержек, поскольку производительность системы снижается в 2–3 раза, однако мобильные среды настолько сложны, что подобная избыточность просто необходима.

Применение метода прямого исправления ошибок подробно рассматривается в главе 8.

### **Адаптивное выравнивание**

Адаптивное выравнивание может использоваться как при передаче аналоговых сигналов (например, аналоговый звук или видео), так и в цифровых системах связи (цифровые данные, оцифрованный звук или видео) и является средством борьбы с межсимвольной интерференцией. В процессе выравнивания используется метод концентрации рассеянной энергии символа в исходный интервал времени. Тема выравнивания довольно широка; соответствующие методы включают использование так называемых сосредоточенных аналоговых схем, а также сложных алгоритмов цифровой обработки сигнала. В данной главе методы цифровой обработки сигналов лишь упоминаются.

На рис. 5.14 представлен распространенный метод обработки сигнала с использованием линейного эквалайзера. В данном примере из каждого символа входного сигнала производится пять выборок, которые разделены равными промежутками времени  $\tau$ . Каждая выборка независимо взвешивается коэффициентом  $C_i$ , после чего взвешенные выборки суммируются и дают выходной сигнал. Контур называется адаптивным, поскольку значения коэффициентов  $C_i$  корректируются динамически. Обычно коэффициенты задаются с помощью заранее известной *настроечной последовательности* битов. Настроечная последовательность передается приемнику, который сравнивает полученные данные с их ожидаемыми значениями. На основе сравнительного анализа рассчитываются соответствующие значения коэффициентов  $C_i$ . Для корректного учета изменений в окружающих условиях время от времени передается новая настроечная последовательность.

Для релейских каналов (или каналов с худшим качеством связи) может требоваться включение новой настроечной последовательности в каждый передаваемый блок данных. Повторимся: хотя такой подход и требует чрезмерных затрат, эти затраты связаны с высоким уровнем ошибок в мобильных системах радиосвязи.

### **Методы разнесения**

Разнесение основывается на том факте, что в каждом отдельно взятом канале процессы замирания происходят независимо. С ошибками можно бороться, создав множество логических каналов между приемником и передатчиком и пе-

редав часть сигнала через каждый из этих каналов. Хотя с помощью такого метода и не удастся полностью предотвратить возникновение ошибок, можно снизить частоту их появления благодаря разделению передаваемого сигнала. Затем для работы с сигналом, имеющим сниженный уровень ошибок, могут применяться другие методы (выравнивание, прямое исправление ошибок).

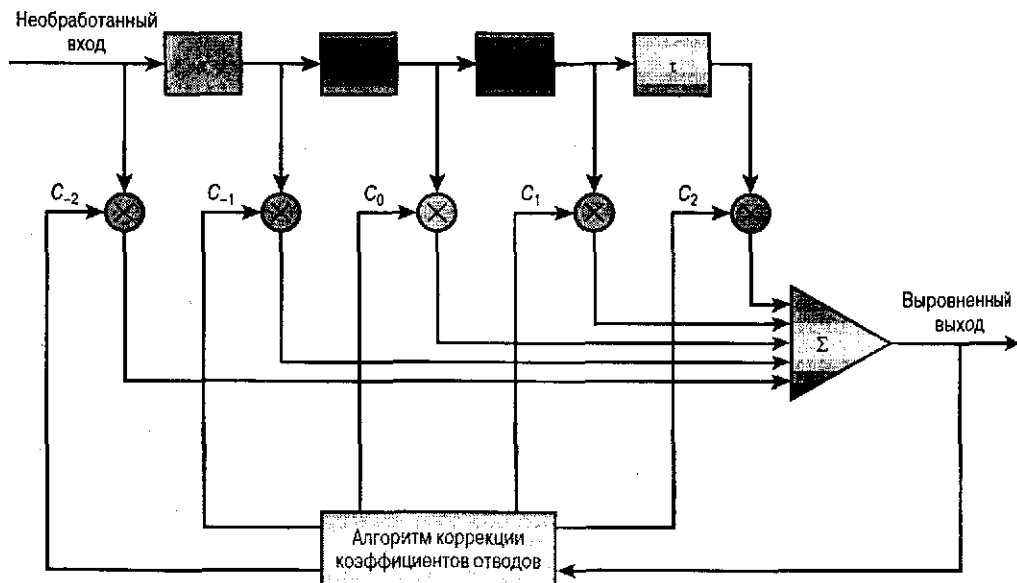


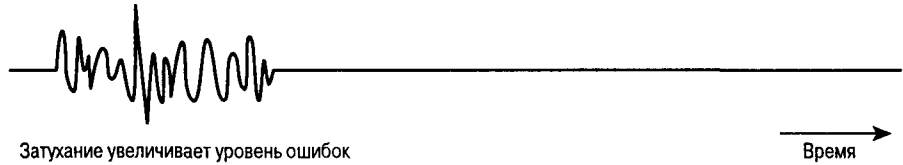
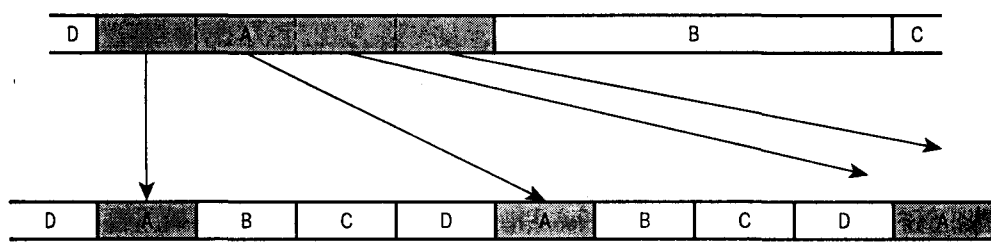
Рис. 5.14. Линейный эквалайзер [PROA94]

В некоторых схемах разнесения используются различные физические тракты передачи, и такие схемы называются **пространственным разнесением**. К примеру, для приема сообщения могут использоваться множественные антенны, расположенные на небольшом расстоянии друг от друга, при этом сигналы комбинируются так, чтобы приемник мог восстановить наиболее вероятный исходный сигнал. Другим примером может быть использование множества совместно размещенных направленных антенн, каждая из которых имеет собственный угол приема. При этом принимаемое сообщение также восстанавливается из суммарного сигнала.

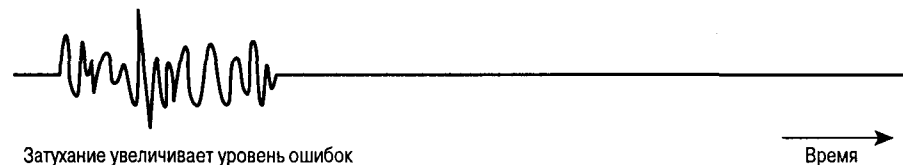
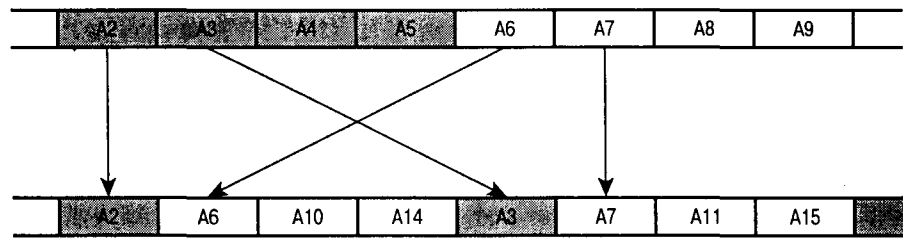
В общем случае под **разнесением** понимают методы разнесения по частоте или во времени. При **частотном разнесении** сигнал распределяется по широкому диапазону частот или передается посредством нескольких несущих. Важнейшим примером такой схемы является расширенный спектр, рассмотренный в главе 7.

Целью **временного разнесения** является распределение данных по разным временным промежуткам, в результате чего шум затрагивает меньшее число битов данных. Разнесение во времени может быть достаточно эффективным в области медленного замирания. Если мобильное устройство перемещается с небольшой скоростью, оно может сравнительно долго оставаться в области с высоким уровнем замирания. Даже если средний уровень сигнала в данной местности значительно выше интерференции, будет наблюдаться длительное возмущение. При этом требуемого эффекта не даст даже применение мощнейших кодов коррекции ошибок. Если цифровые данные передаются с использованием временного уплотнения (Time Division Multiplexing — TDM), когда множество пользователей получают доступ к одному физическо-

му каналу в разные временные интервалы (см. рис. 2.13, б), для разнесения во времени может использоваться чередование блоков данных. Эта идея иллюстрируется на рис. 5.15, а (основан на иллюстрации из [JONE93]). Отметим, что число битов, на которые оказывает влияние шум, остается неизменным, но эти биты разбросаны по большему числу логических каналов. Если в каждом канале используется схема прямого исправления ошибок, то уменьшение числа ошибок по сравнению с передачей по одному каналу существенно увеличивает шансы на успешное исправление всех ошибок. Разнесение во времени может применяться и в том случае, когда схема TDM не используется: поток передаваемых битов рассматривается как последовательность блоков, которые впоследствии перемешиваются. В примере, приведенном на рис. 5.15, б, блоки перемешиваются группами по четыре. Как и в предыдущем случае, число ошибочных битов не уменьшается, однако код исправления ошибок применяется к битовым блокам, которые разнесены во времени. Отметим, что более высокой степени разнесения можно достичь, одновременно применяя чередование TDM и перемешивание блоков.



а) Поток TDM



б) Перемешивание без использования TDM

Рис. 5.15. Разделение пакетов ошибок с помощью чередования блоков данных



Недостатком методов разнесения во времени является задержка. Чем больше степень чередования и перемешивания, тем больше времени потребуется приемнику для восстановления исходной последовательности битов.

## 5.5. РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

Все темы, затронутые в данной главе, подробно освещаются в [FREE97]. Исчерпывающее описание типов антенн и распространения приводится в [BERT00]. Исключительно понятно антенны рассмотрены в [THUR00].

**BERT00** Bertoni H. *Radio Propagation for Modern Wireless Systems*. — Upper Saddle River, NJ: Prentice Hall, 2000.

**FREE97** Freeman R. *Radio System Design for Telecommunications*. — New York: Wiley, 1997.

**THUR00** Thurwachter C. *Data and Telecommunications: Systems and Applications*. — Upper Saddle River, NJ: Prentice Hall, 2000.

## 5.6. ТЕРМИНЫ, ВОПРОСЫ И ЗАДАЧИ

### Основные термины

адаптивное	изотропная антенна	отражение от верхних
выравнивание	импульсные помехи	слоев атмосферы
амплитудное	интермодуляционный	параболическая
замирание	шум	отражающая
антенна	коэффициент	антенна
атмосферное	усиления антенны	перекрестные помехи
поглощение	линия прямой	потери в свободном
быстрое замирание	видимости	пространстве
вибратор Герца	линия	преломление
диаграмма	радиовидимости	прямое исправление
направленности	медленное замирание	ошибок
диаграмма приема	многолучевое	разнесение
диполь	распространение	рассеяние
дифракция	огибание поверхности	тепловой шум
замирание	Земли	ширина луча
затухание	оптическая линия	шум
селективное	прямой видимости	
замирание	отражение	

### Вопросы

1. Какие две функции выполняются антеннами?
2. Что такое “изотропная антенна”?
3. Какую информацию можно получить из диаграммы направленности?
4. Каким преимуществом обладает параболическая отражающая антенна?
5. Какие параметры определяют коэффициент усиления антенны?

6. Какова основная причина потерь мощности сигнала в системах спутниковой связи?
7. Назовите четыре основных типа помех и кратко опишите их.
8. Что такое “преломление”?
9. Что такое “замирание”?
10. В чем разница между дифракцией и рассеянием?
11. В чем разница между быстрым и медленным замиранием?
12. В чем разница между селективным и амплитудным замиранием?
13. Назовите три метода разнесения и кратко опишите их.

## Задачи

1. При радиосвязи в свободном пространстве мощность сигнала уменьшается прямо пропорционально квадрату расстояния от передатчика, тогда как при передаче через кабель постоянной является величина затухания (в децибелах) на километр. В приведенной ниже таблице указаны значения потерь мощности сигнала для передачи сигнала в свободном пространстве и с использованием кабеля. Заполните пустые клетки таблицы.

Расстояние, км	Радиосвязь, дБ	Кабель, дБ
1	-6	-3
2		
4		
8		
16		

2. Найдите оптимальную длину волны и частоту для полуволнового диполя длиной 10 м.
3. Известно, что максимальная глубина океана, на которой можно обнаружить электромагнитные сигналы, сгенерированные в атмосфере, увеличивается с ростом длины волны этих сигналов. Основываясь на данном факте, военные предложили использовать радиоволны большой длины (с частотой порядка 30 Гц) для связи с подводными лодками. Найдите длину антенны, если известно, что она должна быть приблизительно равна половине длины волны.
4. Максимум мощности звуковых колебаний человеческого голоса наблюдается на частоте порядка 300 Гц. Размер антенны для передачи таких частот непомерно велик, поэтому для радиопередачи голосовой сигнал модулируется на несущей с большей частотой, передающая антенна для которой имеет значительно меньший размер.
  - а. Найдите длину антенны для передачи радиоволн с частотой 300 Гц (длина антенны должна равняться половине длины волны).
  - б. Альтернативным способом передачи голоса может быть модуляция (см. главу 6). Несущая волна модулируется голосовым сигналом таким образом, что диапазон сигнала представляет собой узкую полосу, центр которой совпадает с частотой несущей. Предположим, что передающая антенна имеет длину 1 м. Какой должна быть частота несущей?
5. Довольно часто приходится слышать истории о людях, которые способны улавливать радиосигналы коронками зубов. Предположим, что коронка длиной 2,5 мм (0,0025 м) выступает в роли радиоантенны. Зная, что длина антенны равна половине длины принимаемой волны, определите частоту принимаемого сигнала.

6. В разделе 5.1 утверждалось, что если поместить источник электромагнитного излучения в фокус параболоида, поверхность которого обладает отражающими свойствами, то волны будут отражаться вдоль линий, параллельных оси параболоида. Рассмотрим в качестве примера параболу  $y^2 = 2px$ , график которой изображен на рис. 5.16. Пусть  $P(x_1, y_1)$  — точка, принадлежащая параболе;  $PF$  — линия, соединяющая фокус и точку  $P$ . Через точку  $P$  проведем прямую  $L$ , параллельную оси абсцисс, а также касательную параболы  $M$ . Угол между  $L$  и  $M$  равен  $\beta$ ; угол между  $PF$  и  $M$  —  $\alpha$ ;  $\alpha$  соответствует углу, под которым на параболе в точке  $P$  падает луч, вышедший из точки  $F$ . Поскольку угол падения равен углу отражения, луч, отраженный в точке  $P$ , должен образовать угол  $\alpha$  с прямой  $M$ . Следовательно, доказав, что  $\alpha = \beta$ , можно автоматически доказать, что лучи, исходящие из точки  $F$ , после отражения параболой будут распространяться параллельно оси абсцисс.

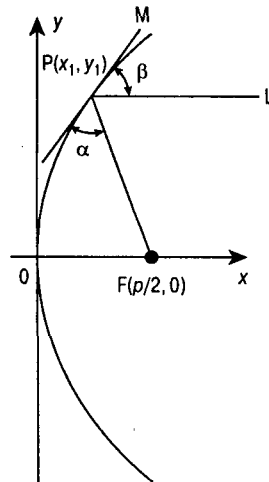


Рис. 5.16. Параболическое отражение

- а. Докажите, что  $\operatorname{tg} \beta = p/y_1$ . Подсказка: из тригонометрии известно, что коэффициент наклона линии равен тангенсу угла, который данная линия образует с положительным направлением оси абсцисс. Известно также, что коэффициент наклона прямой, являющейся касательной некоторой кривой в заданной точке, равен производной функции этой кривой в точке касания.
- б. Докажите, что  $\operatorname{tg} \alpha = p/y_1$ , из чего следует, что  $\alpha = \beta$ . Подсказка: формула тангенса разности углов  $\alpha_1$  и  $\alpha_2$ :  $\operatorname{tg}(\alpha_1 - \alpha_2) = (\operatorname{tg} \alpha_2 - \operatorname{tg} \alpha_1)/(1 + \operatorname{tg} \alpha_2 \times \operatorname{tg} \alpha_1)$ .
7. Для каждого типа антенн, перечисленных в табл. 5.2, найдите эффективную площадь и коэффициент усиления. Длина волны составляет 30 см. Найдите те же параметры для длины волны 3 мм. Площадь рупорной и параболической антенны считайте равной  $\pi$ .
8. Как правило, расстояние удобнее записывать в километрах, а частоту — в мегагерцах. Запишите уравнение (5.2), используя данные размерности.
9. Рассмотрим две антенны, каждая из которых является полуволновым диполем с коэффициентом направленного действия 3 дБ. Если переданная мощность равна 1 Вт, а расстояние между антеннами — 10 км, какой будет полученная мощность?

ность? Будем считать, что взаимное расположение антенн позволяет точно оценить коэффициент направленного действия. Используется частота 100 МГц.

10. Рассмотрим передатчик с выходной мощностью 50 Вт.

- а. Выразите выходную мощность в децибел-ваттах и децибел-милливаттах.
- б. Выходной сигнал передатчика передается с помощью антенны единичного усиления с несущей 900 МГц. Какой будет мощность принятого сигнала в децибел-милливаттах на расстоянии 100 м в вакууме?
- в. Решите задачу (б) для расстояния 10 км.
- г. Решите задачу (в), если коэффициент усиления антенны приемника равен 2.

11. Сверхвысокочастотный передатчик имеет выходную мощность 0,1 Вт при частоте 2 ГГц. Пусть данный передатчик используется в некоторой системе связи. Передающие и принимающие антенны — параболы диаметром 1,2 м.

- а. Найдите коэффициент усиления (в децибелах) каждой антенны.
- б. Учитывая коэффициент усиления антенны, найдите эффективную излученную мощность переданного сигнала.
- в. Расстояние между передающей и принимающей антеннами составляет 24 км в вакууме. Найдите номинальную мощность сигнала (в децибел-милливаттах) на выходе антенны приемника.

12. В разделе 5.2 утверждалось, что при отсутствии препятствий на пути распространения сигнала, длина линии оптической видимости выражается как

$d = 3,57 \sqrt{h}$ , где  $d$  — расстояние между антенной и линией горизонта в километрах;  $h$  — высота антенны в метрах. Выведите указанное уравнение, считая радиус Земли равным 6370 км. *Подсказка:* пусть антенна расположена перпендикулярно к поверхности Земли, а прямая, соединяющая вершину антенны с линией горизонта, является касательной к поверхности Земли в точке горизонта. Для большей наглядности начертите рисунок, изображающий антенну, линию прямой видимости и земную окружность.

13. Определите высоту телевизионной передающей антенны, которая могла бы обеспечивать зрителям прием сигнала на расстоянии 80 км.

14. Найдите уровень теплового шума в канале с шириной полосы 10 кГц. Рабочая температура составляет 50°C, мощность передаваемого сигнала — 1 000 Вт. Сравните уровень теплового шума с рабочей мощностью канала.

15. Меандр длительностью  $T = 1$  мс, изображенный на рис. 2.5, в, обрабатывается фильтром низких частот, который позволяет пропускать без поглощения волны частотой до 8 кГц.

- а. Найдите мощность выходного сигнала.
- б. Учитывая наличие на входе фильтра напряжения теплового шума  $N_0 = 0,1$  мкВт/Гц, найдите отношение выходного сигнала к шумам в децибелах.

16. Уровень мощности сигнала, полученного некоторой цифровой системой связи, равен -151 дБВт, а эффективная шумовая температура приемника — 1500 К. Найдите отношение  $E_b/N_0$  для канала связи со скоростью передачи 2 400 бит/с.

17. Луч видимого света попадает из атмосферы в воду под углом 30° к поверхности воды. Под каким углом этот луч будет распространяться в воде? При нормальных атмосферных условиях коэффициент преломления воздуха на поверхности Земли составляет 1,0003. Коэффициент преломления воды принято считать равным 4/3.

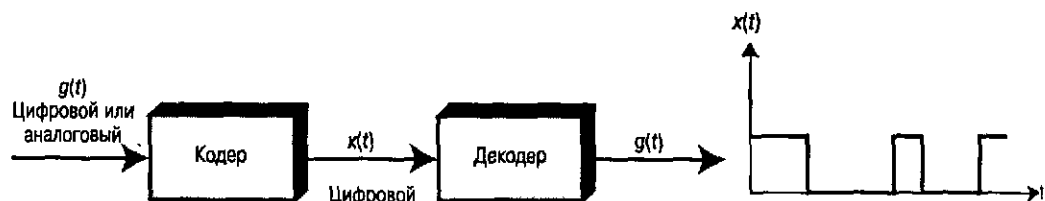
# ГЛАВА **6**

## **МЕТОДЫ КОДИРОВАНИЯ СИГНАЛОВ**

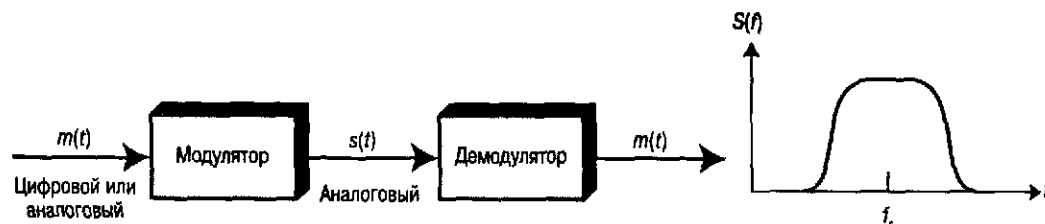
- 6.1. Критерии кодирования сигналов**
- 6.2. Цифровые данные, аналоговые сигналы**
  - Амплитудная манипуляция**
  - Частотная манипуляция**
  - Фазовая манипуляция**
  - Производительность**
  - Квадратурная амплитудная модуляция**
- 6.3. Аналоговые данные, аналоговые сигналы**
  - Амплитудная модуляция**
  - Угловая модуляция**
- 6.4. Аналоговые данные, цифровые сигналы**
  - Импульсно-кодовая модуляция**
  - Дельта-модуляция**
  - Производительность**
- 6.5. Рекомендуемая литература**
- 6.6. Термины, вопросы и задачи**
- Приложение 6А. Доказательство теоремы о дискретном представлении**

В главе 2 рассказано о различии аналоговых и цифровых данных и аналоговых и цифровых сигналов, а на рис. 2.8 показано, что данные любого вида могут кодироваться сигналом любого типа.

На рис. 6.1 акцент делается уже на задействованных процессах. При цифровой передаче сигналов исходные данные  $g(t)$ , которые могут быть как аналоговыми, так и цифровыми, кодируются цифровым сигналом  $x(t)$ . Конкретная форма этого сигнала зависит от метода кодировки и выбирается с целью оптимизации использования передающей среды. Например, может быть выбрана кодировка, максимально экономно использующая полосу или сводящая к минимуму вероятность возникновения ошибок.



а) Кодирование в цифровой сигнал



б) Модуляция в аналоговый сигнал

Рис. 6.1. Методы кодирования и модуляции

Основу аналоговой передачи сигналов составляет непрерывный сигнал с постоянной частотой, называемый *несущим сигналом*. Частота несущего сигнала выбирается совместимой с используемой передающей средой. Передавать данные с помощью несущего сигнала позволяет модуляция — процесс кодирования исходных данных несущим сигналом с частотой  $f_c$ . Все методы модуляции включают операции с одним (или более) из трех фундаментальных параметров частотного представления сигнала, которыми являются амплитуда, частота и фаза.

Поступающий сигнал  $m(t)$  может быть аналоговым или цифровым и называется *модулируемым* или *немодулированным сигналом*. Результат модулирования несущего сигнала называется *модулированным сигналом*  $s(t)$ . Как показано на рис. 6.1, б, сигнал  $s(t)$  является сигналом с ограниченной полосой (полосовым сигналом). Положение полосы в спектре связано с частотой  $f_c$ , и обычно полоса центрируется на данной частоте. Отметим, что в этом случае используемая кодировка также выбирается с целью оптимизации каких-либо характеристик передачи.

Все четыре возможные комбинации (или кодировки), изображенные на рис. 6.1, были рассмотрены в главе 2: цифровые данные, цифровой сигнал; цифровые данные, аналоговый сигнал; аналоговые данные, аналоговый сигнал; аналоговые данные, цифровой сигнал. Три последних метода являются важными в контексте беспроводной связи и широко используются.

- **Цифровые данные, аналоговый сигнал.** При беспроводной передаче цифровые данные и цифровые сигналы требуется преобразовать в аналоговые сигналы.
- **Аналоговые данные, аналоговый сигнал.** Обычно узкополосный аналоговый сигнал, такой, как речь или видео, для передачи требуется модулировать на высокочастотной несущей.
- **Аналоговые данные, цифровой сигнал.** Для улучшения качества передачи и использования преимуществ схем TDM речь обычно оцифровывается перед передачей по направляемой или ненаправляемой среде. При беспроводной передаче полученные цифровые данные требуется дополнительно модулировать на аналоговой несущей.

Три названные технологии будут рассмотрены в данной главе. Вначале будет изучен критерий, который можно использовать при оценке различных подходов в каждой категории. В главе 7 рассмотрен метод расширения спектра, который можно отнести сразу к нескольким категориям.

## 6.1. КРИТЕРИИ КОДИРОВАНИЯ СИГНАЛОВ

Дадим вначале определения некоторых терминов. Напомним, что цифровой сигнал — это последовательность дискретных прерывистых импульсов напряжения. Каждый импульс представляет собой сигнальную посылку. Для передачи двоичных данных каждый бит данных кодируется сигнальной посылкой. В простейшем случае существует взаимно-однозначное соответствие между битами и сигнальными посылками (рис. 2.9). Здесь двоичный 0 представляется более низким уровнем напряжения, а двоичная 1 — более высоким. Подобным образом цифровой поток данных можно закодировать в аналоговый сигнал как последовательность сигнальных посылок, причем каждая сигнальная посылка — это импульс постоянной частоты, фазы и амплитуды. Может существовать взаимно однозначное соответствие между элементами данных (битами) и аналоговыми сигнальными посылками. Ниже будет показано, что как для аналоговых, так и для цифровых сигналов может существовать соответствие типа “один-множество” и “множество-одна” между элементами данных и сигнальными посылками.

Продолжим определять термины. Значение термина “скорость передачи данных” очевидно; эта скорость измеряется в битах в секунду. *Длительностью* или *длиной бита* называется время, затрачиваемое передатчиком на излучение этого бита; при скорости передачи данных  $R$  длина бита равна  $1/R$ . *Скоростью модуляции* называется скорость изменения уровня сигнала. Как будет показано позднее, эта величина зависит от природы цифрового кодирования. Скорость модуляции измеряется в бодах, 1 бод соответствует одной сигнальной посылке в одну секунду. В табл. 6.1 приведены основные термины, а несколько позже приводится пример, помогающий лучше их понять.

**Таблица 6.1. Основные термины области передачи данных**

Термин	Единицы	Определение
Элемент данных	Бит	Отдельный двоичный нуль или единица
Скорость передачи данных	Биты в секунду (бит/с)	Скорость, с которой передаются элементы данных
Сигнальная посылка	Цифровая: импульс напряжения с постоянной амплитудой. Аналоговая: импульс постоянной частоты, фазы и амплитуды	Часть сигнала, занимающая кратчайший интервал сигнального кода
Скорость передачи сигнала или скорость модуляции	Сигнальная посылка в секунду (боды)	Скорость, с которой передаются элементы данных

Задачи, входящие в интерпретацию цифровых сигналов в приемнике, можно обобщить, снова обратившись к рис. 2.9. Во-первых, приемник должен знать распределение сигналов во времени. Иначе говоря, приемник должен, с определенной точностью, знать, где бит начинается и где — заканчивается. Во-вторых, приемник должен определить уровень каждого двоичного разряда: высокий он (1) или низкий (0). На рис. 2.9 для выполнения этих задач в середине интервала каждого двоичного разряда производится выборка, после чего полученная величина сравнивается с пороговым значением. Как можно видеть на том же рисунке, помехи и другие искажения могут приводить к ошибочным результатам.

Какие же факторы определяют, насколько успешной будет интерпретация сигнала, поступившего в приемник? Как было показано в главе 2, существенными являются три фактора: отношение сигнал/шум (или, лучше, отношение  $E_b/N_0$ ), скорость передачи данных и ширина полосы. При фиксированных остальных параметрах справедливы следующие утверждения.

- Увеличение скорости передачи данных приводит к увеличению скорости появления ошибочных битов, или частоты битовых ошибок (BER — bit error rate)<sup>1</sup>.
- Увеличение отношения сигнал/шум уменьшает скорость появления ошибочных битов.
- Увеличение ширины полосы позволяет увеличить скорость передачи данных.

Существует еще один фактор, позволяющий увеличить производительность — выбор схемы кодирования. Схема кодирования представляет собой простое отображение информационных битов в сигнальные посылки. Был испробован целый ряд подходов. Перед тем как перейти к их изучению, рассмотрим следующие способы оценки или сравнения различных схем.

<sup>1</sup> Данный параметр является наиболее общей характеристикой возникновения ошибок в канале передачи данных и определяется как вероятность приема ошибочного бита. Этот параметр также называется коэффициентом ошибок битов (bit error ratio). Этот термин понятнее, поскольку термин “скорость” обычно используется для обозначения изменяющегося со временем количества чего-либо. К сожалению, в большинстве книг и стандартов буква R в аббревиатуре BER расшифровывается как rate — скорость.



- **Спектр сигнала.** Некоторые особенности спектра сигнала имеют важное значение. Отсутствие высокочастотных компонентов означает, что для передачи требуется более узкая полоса. Кроме того, желательнее также, чтобы отсутствовала постоянная составляющая. Если в сигнале она присутствует, то должно существовать прямое физическое соединение передающих устройств. Если же постоянная составляющая отсутствует, то можно использовать соединение по переменному току через трансформатор; это дает превосходную электрическую изоляцию и снижает интерференцию. И, наконец, уровни искажения сигнала и интерференции зависят от спектральных свойств передаваемого сигнала. На практике передаточная функция канала всегда хуже на краях полосы пропускания. Следовательно, при хорошей структуре сигнала основная передаваемая мощность должна быть сосредоточена в середине полосы передачи. Тогда принимаемый сигнал будет менее искаженным. Этого можно добиться с помощью кодировок, позволяющих формировать спектр передаваемого сигнала.
- **Синхронизация.** Приемник должен уметь определять начало и конец каждого двоичного разряда. Это совсем не легко. Один из довольно дорогостоящих подходов состоит в выделении специального канала для синхронизации приемника и передатчика. Альтернативный вариант — создание некоторого механизма синхронизации на базе переданного сигнала. Организовать такой механизм позволяет соответствующая кодировка.
- **Интерференция сигналов и помехоустойчивость.** Отдельные кодировки позволяют добиваться лучшей производительности при наличии помех, чем другие. Выражается это, как правило, через скорость появления ошибок.
- **Стоимость и сложность.** Хотя цифровые логические схемы продолжают падать в цене, этот фактор игнорировать не стоит. В частности, чем выше скорость передачи сигналов, позволяющая получить определенную скорость передачи данных, тем выше цена. Мы увидим, что некоторые кодировки в действительности требуют большую скорость передачи сигналов, чем имеющаяся скорость передачи данных.

Обратимся теперь к обсуждению различных методов кодирования.

## 6.2. ЦИФРОВЫЕ ДАННЫЕ, АНАЛОГОВЫЕ СИГНАЛЫ

Начнем с рассмотрения передачи цифровых данных с помощью аналоговых сигналов. Примером является передача цифровых данных через телефонную сеть общего пользования. Изначально телефонная сеть была разработана для получения, коммутации и передачи аналоговых сигналов речевого диапазона частот (примерно 300–3400 Гц). В настоящее время сеть неприменима для оперирования цифровыми сигналами, поступающими от абонентов (хотя такое положение дел уже начинает меняться), поэтому цифровые устройства присоединяются к сети посредством модема (модулятора-демодулятора), преобразующего цифровые данные в аналоговые сигналы и наоборот.

В телефонной сети используются модемы, генерирующие сигнал на частотах речевого диапазона. Те же основополагающие технологии используются и в модемах,

генерирующих сигналы на высших частотах (например, в СВЧ-диапазоне). В этом разделе описаны данные технологии и предлагается краткое обсуждение технических характеристик альтернативных подходов.

Мы уже говорили о том, что в процессе модулирования задействованы одна или несколько характеристик несущего сигнала: амплитуда, частота и фаза. Соответственно, существуют три основные технологии кодирования или модуляции, выполняющие преобразование цифровых данных в аналоговый сигнал (см. рис. 6.2): амплитудная манипуляция (amplitude-shift keying — ASK), частотная манипуляция (frequency-shift keying — FSK) и фазовая манипуляция (phase-shift keying — PSK). Отметим, что во всех перечисленных случаях результирующий сигнал центрирован на несущей частоте.

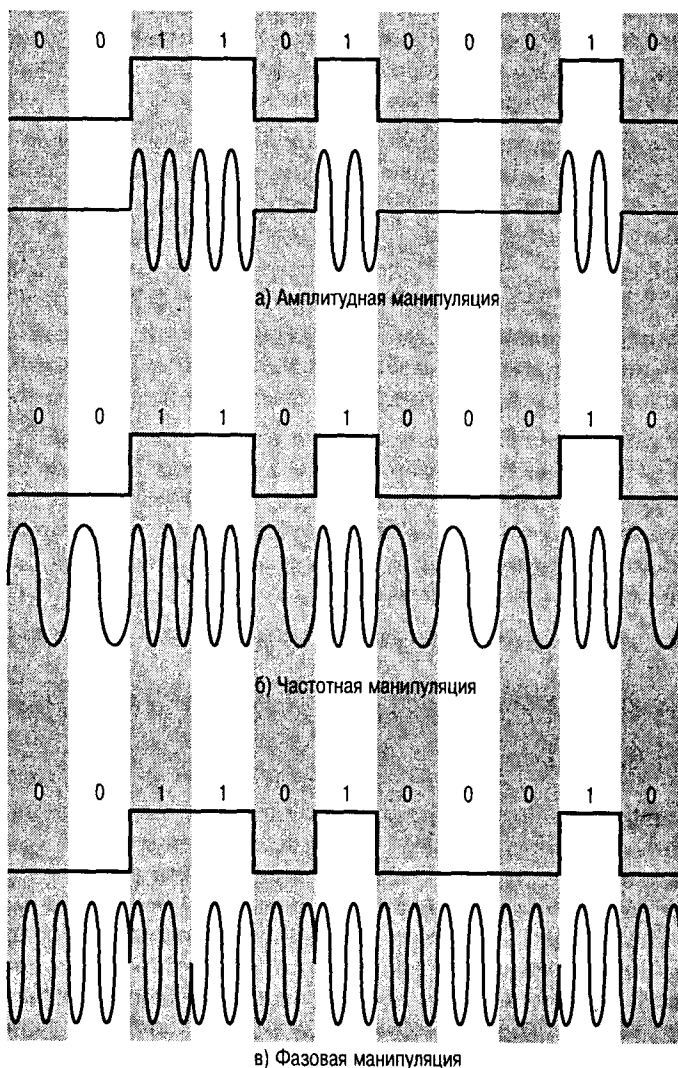


Рис. 6.2. Модуляция цифровых данных аналоговыми сигналами

## Амплитудная манипуляция

При амплитудной манипуляции два двоичных значения представляются сигналами несущей частоты с двумя различными амплитудами. Одна из амплитуд, как правило, выбирается равной нулю; т.е. одно двоичное число представляется наличием несущей частоты при постоянной амплитуде, а другое — ее отсутствием (рис. 6.2, а). Результирующий сигнал равен

$$\text{при амплитудной манипуляции } s(t) = \begin{cases} A \cos(2\pi f_c t) & \text{— двоичная 1} \\ 0 & \text{— двоичный 0} \end{cases} \quad (6.1)$$

Здесь  $A \cos(2\pi f_c t)$  — несущий сигнал. Метод амплитудной манипуляции чувствителен к внезапным скачкам напряжения и неэффективен. В телефонных линиях он обычно используется только при скоростях до 1200 бит/с.

Метод амплитудной манипуляции используется для передачи цифровых данных по оптоволокну. Для передатчиков на светодиодах справедливо предыдущее равенство. Иными словами, одна сигнальная посылка представляется световым импульсом, тогда как другая — отсутствием света. В лазерных же передатчиках имеется, как правило, фиксированный ток смещения, вызывающий излучение света с более низким уровнем. В результате одна из сигнальных посылок представляется этим уровнем, тогда как световая волна большей амплитуды представляет другую сигнальную посылку.

## Частотная манипуляция

Наиболее распространенной формой частотной манипуляции является бинарная (BFSK), в которой два двоичных числа представляются сигналами двух различных частот, расположенных около несущей. Результирующий сигнал равен

$$\text{при бинарной частотной манипуляции } s(t) = \begin{cases} A \cos(2\pi f_1 t) & \text{— двоичная 1} \\ A \cos(2\pi f_2 t) & \text{— двоичный 0} \end{cases} \quad (6.2)$$

где  $f_1$  и  $f_2$  — частоты, смещенные от несущей частоты  $f_c$  на величины, равные по модулю, но противоположные по знаку.

На рис. 6.3 приведен пример использования частотной манипуляции для дуплексной работы в телефонной линии. Данный рисунок представляет собой спецификацию модемов серии Bell System 108. Напомним, что по телефонным линиям передаются частоты приблизительно от 300 до 3400 Гц и что *дуплексная передача* означает одновременную передачу сигналов в обоих направлениях. Для организации дуплексной передачи полоса разделяется на две части. В одном направлении (передача или прием) числа 0 и 1 представляются частотами, центрированными на частоте 1170 Гц, но смещенными на 100 Гц в обе стороны. Чередувание этих частот должно дать сигнал, спектр которого лежит в левой затененной части рис. 6.3. Подобным образом при приеме или передаче в противоположном направлении модем использует частоты, смещенные на 100 Гц от частоты 2125 Гц. Диапазон полученного в результате сигнала обозначен правой затененной областью рис. 6.3. Отметим, что сигналы несколько перекрываются, поэтому между ними существует незначительная интерференция.

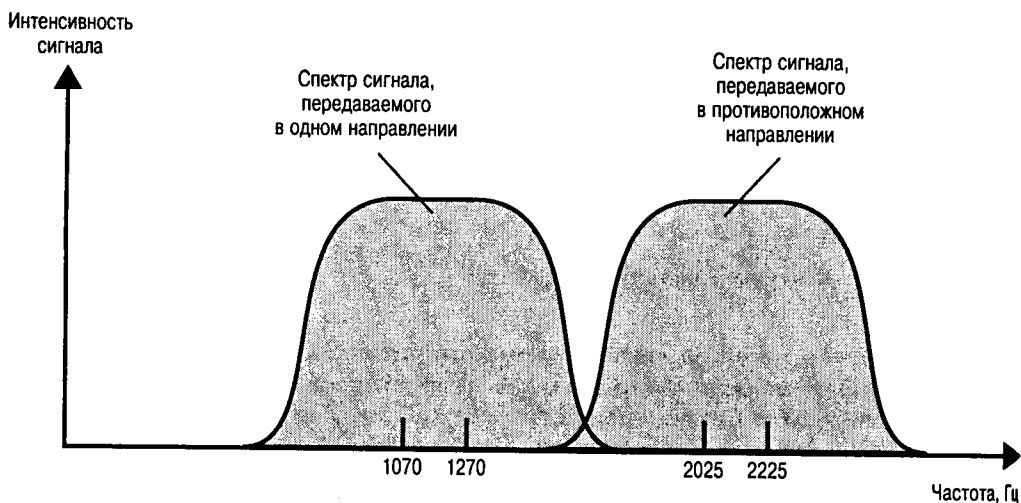


Рис. 6.3. Дуплексная передача по телефонной линии с использованием частотной манипуляции

Бинарная частотная манипуляция менее восприимчива к ошибкам, чем амплитудная манипуляция. Впрочем, в телефонных линиях она также используется при скоростях до 1 200 бит/с. Помимо этого, схема ASK применяется в высокочастотных (от 3 до 30 МГц) радиопередачах, а в локальных сетях, использующих коаксиальный кабель, она может применяться даже на более высоких частотах.

Более эффективной, но и более подверженной ошибкам, является схема многочастотной манипуляции (MFSK), в которой используются более двух частот. В этом случае каждая сигнальная посылка представляет более одного бита. Переданный сигнал MFSK (для одного периода передачи сигнальной посылки) можно определить следующим образом:

$$\text{при многочастотной манипуляции } s_i(t) = A \cos 2\pi f_i t, \quad 1 \leq i \leq M. \quad (6.3)$$

Здесь

$$f_i = f_c + (2i - 1 - M)f_d;$$

$f_c$  — несущая частота;

$f_d$  — разностная частота;

$M$  — число различных сигнальных посылок =  $2^L$ ;

$L$  — число битов на одну сигнальную посылку.

Для согласования скорости передачи данных с входным потоком битов каждая выходная сигнальная посылка передается в течение  $T_s = LT$  секунд, где  $T$  — время передачи бита (скорость передачи данных =  $1/T$ ). Таким образом, одна сигнальная посылка (представляющая собой тон постоянной частоты) кодирует  $L$  бит. Общая требуемая полоса —  $2Mf_d$ . Можно показать, что минимальный разнос частоты составляет  $2f_d = 1/T_s$ . Таким образом, модулятор требует ширины полосы  $W_d = 2Mf_d = M/T_s$ .

**Пример.** При  $f_c = 250$  кГц,  $f_d = 25$  кГц и  $M = 8$  ( $L = 3$  бит) имеем следующее распределение частот для каждой возможной 3-битовой комбинации:

- $f_1 = 75$  кГц    000
- $f_2 = 125$  кГц   001
- $f_3 = 175$  кГц   010
- $f_4 = 225$  кГц   011
- $f_5 = 275$  кГц   100
- $f_6 = 325$  кГц   101
- $f_7 = 375$  кГц   110
- $f_8 = 425$  кГц   111

Схема может поддерживать скорость передачи данных  $2f_d = 1/T_s = 50$  Кбит/с

На рис. 6.4 представлен пример схемы MFSK с  $M = 4$ . Входной поток битов кодируется по два бита, после чего передается одна из четырех возможных 2-битовых комбинаций.

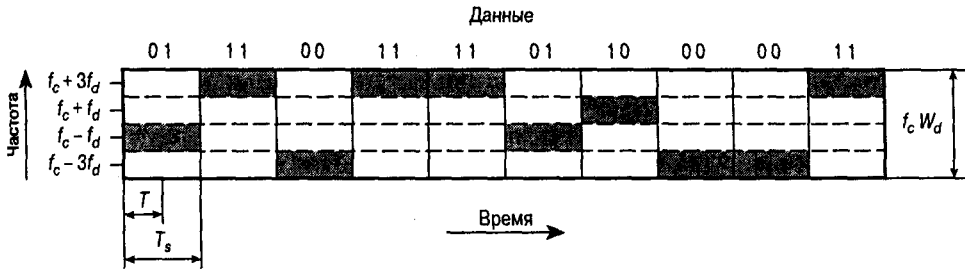


Рис. 6.4. Использование частоты схемой MFSK ( $M = 4$ )

## Фазовая манипуляция

При фазовой манипуляции для представления данных выполняется смещение несущего сигнала.

### Двухуровневая фазовая манипуляция

Простейшая схема, в которой для представления двух двоичных цифр используются две фазы, называется бинарной фазовой манипуляцией (рис. 6.2, в). Получающийся сигнал имеет следующий вид (для одного периода передачи бита):

при бинарной фазовой манипуляции

$$s(t) = \begin{cases} A \cos(2\pi f_c t) \\ A \cos(2\pi f_c t + \pi) \end{cases} = \begin{cases} A \cos(2\pi f_c t) & \text{— двоичная 1} \\ -A \cos(2\pi f_c t) & \text{— двоичный 0} \end{cases} \quad (6.4)$$

Поскольку сдвиг фазы на  $180^\circ$  ( $\pi$ ) эквивалентен умножению синусоиды на  $-1$ , может использоваться правая часть выражения (6.4). Это позволяет использовать удобную формулировку. Если имеется поток битов и  $d(t)$  определяется как

дискретная функция, значение которой равно +1 при передаче 1 и -1 при передаче 0, то переданный сигнал можно определить следующим образом:

$$\text{при бинарной фазовой манипуляции } s_A(t) = A d(t) \cos(2\pi f_c t). \quad (6.5)$$

Альтернативной формой двухуровневой PSK является дифференциальная PSK (DPSK), пример которой приведен на рис. 6.5. В данной системе двоичный 0 представляется сигнальным пакетом, фаза которого совпадает с фазой предыдущего посланного пакета, а двоичная 1 представляется сигнальным пакетом с фазой, противоположной фазе предыдущего пакета. Такая схема называется *дифференциальной*, поскольку сдвиг фаз выполняется относительно предыдущего переданного бита, а не относительно какого-то эталонного сигнала. При дифференциальном кодировании передаваемая информация представляется не сигнальными посылками, а изменениями между последовательными сигнальными посылками. Схема DPSK делает излишним строгое согласование фазы местного гетеродина приемника и передатчика. До тех пор пока предыдущая полученная фаза точна, точен и фазовый эталон.

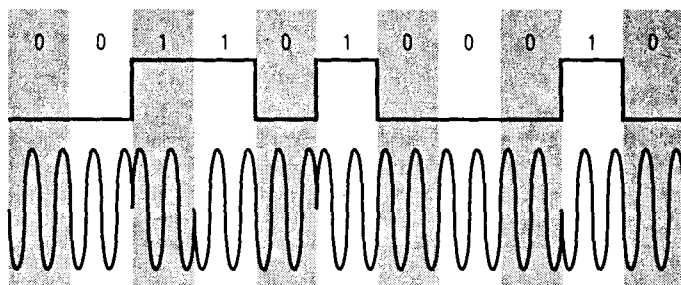


Рис. 6.5. Дифференциальная фазовая манипуляция (DPSK)

### Четырехуровневая PSK

Если каждой сигнальной посылкой представить более одного бита, то это позволит эффективнее использовать полосу сигнала. Например, в распространенной кодировке, известной как квадратурная фазовая манипуляция (quadrature phase-shift keying — QPSK), вместо сдвига фазы на  $180^\circ$ , как в кодировке PSK, используются сдвиги фаз, кратные  $\pi/2$  ( $90^\circ$ ).

$$\text{При квадратурной фазовой манипуляции } s(t) = \begin{cases} A \cos(2\pi f_c t + \frac{\pi}{4}) & \text{—11} \\ A \cos(2\pi f_c t + \frac{3\pi}{4}) & \text{—10} \\ A \cos(2\pi f_c t + \frac{5\pi}{4}) & \text{—00} \\ A \cos(2\pi f_c t + \frac{7\pi}{4}) & \text{—01} \end{cases} \quad (6.6)$$

Таким образом, каждая сигнальная посылка представляет не один бит, а два.

На рис. 6.6 в общих чертах представлена схема модуляции QPSK. Вход — поток двоичных цифр со скоростью  $R = 1/t_B$ , где  $t_B$  — ширина бита. Поток конвертируется в два отдельных потока битов со скоростью  $R/2$  каждый. Полученные два потока называются синфазным (I) и квадратурным (Q). На диаграмме верхний поток модулируется на несущей  $f_c$  путем умножения потока битов на несущую. Для удобства двоичная единица отображается в  $1/\sqrt{2}$ , а ноль — в  $-1/\sqrt{2}$ . Таким образом, двоичная единица представляется несущей с измененным масштабом, а двоичный ноль — отрицательной версией несущей с измененным масштабом; амплитуда в обоих случаях постоянна. Для модуляции нижнего потока используется та же несущая, смещенная на  $90^\circ$ . После этого два полученных сигнала складываются и передаются. Результирующий сигнал можно записать следующим образом:

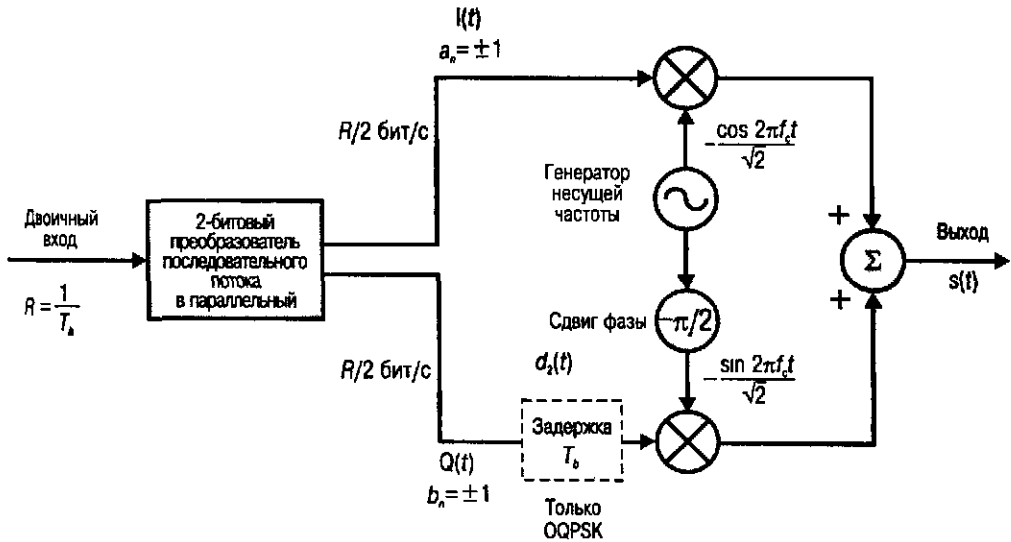


Рис. 6.6 Модуляторы QPSK и OQPSK

при квадратурной фазовой манипуляции  $s(t) = \frac{1}{\sqrt{2}} I(t) \cos 2\pi f_c t - \frac{1}{\sqrt{2}} Q(t) \sin 2\pi f_c t$ .

На рис. 6.7 приведен пример кодирования QPSK. Оба модулированных потока являются сигналами BPSK со скоростью передачи, равной половине скорости передачи исходного потока. Таким образом, скорость передачи символов в полученных сигналах равна половине скорости передачи битов на входе. Отметим, что при переходе от одного символа к другому возможно изменение фазы на  $180^\circ$  ( $\pi$ ).

На рис. 6.6 также показана разновидность схемы QPSK, именуемая QPSK со сдвигом, или ортогональная QPSK (OQPSK). Отличие заключается в том, что в квадратурный поток вводится задержка, равная времени передачи одного бита, что дает следующий сигнал:

$$s(t) = \frac{1}{\sqrt{2}} I(t) \cos 2\pi f_c t - \frac{1}{\sqrt{2}} Q(t - T_b) \sin 2\pi f_c t.$$

Двоичное число	1	2	3	4	5	6	7	8	9	10
Двоичное значение	1	-1	1	1	-1	-1	-1	1	1	1
	I	Q	I	Q	I	Q	I	Q	I	Q

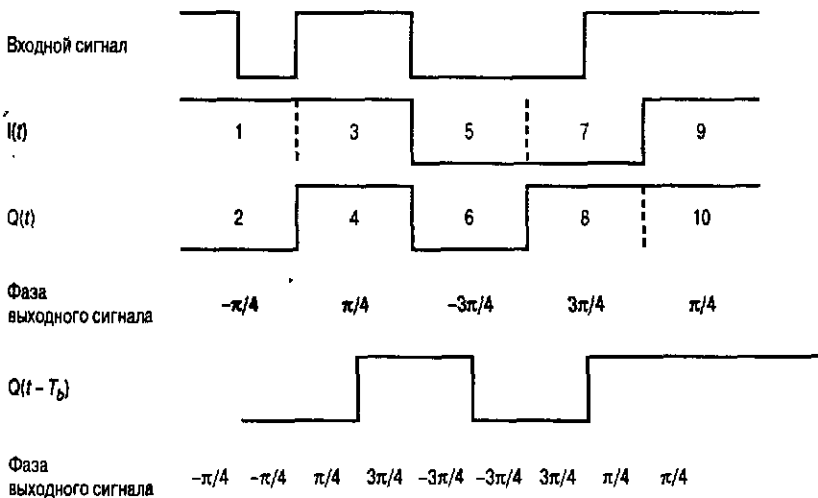


Рис. 6.7. Примеры сигналов QPSK и OQPSK

Поскольку OQPSK и QPSK отличаются только задержкой в квадратурном потоке, спектральные характеристики и вероятности ошибок обеих схем совпадают. Из рис. 6.7 можно видеть, что в каждый момент времени изменить знак может только один из двух битов в паре сигналов; следовательно, суммарное изменение фазы никогда не превысит  $90^\circ$  ( $\pi/2$ ). Эта особенность может оказаться выгодной, поскольку физические ограничения модуляторов фазы не позволяют легко выполнять значительные изменения фазы при больших скоростях передачи. Кроме того, схема OQPSK обеспечивает лучшую достоверность передачи в канале (в который входят и приемник с передатчиком), имеющем существенно нелинейные компоненты. Нелинейность приводит к расширению полосы передачи, что может вызвать интерференцию сигналов данного канала с сигналами соседних каналов. Поскольку контролировать расширение полосы намного легче при незначительных изменениях фазы, схема OQPSK имеет преимущество перед схемой QPSK.

### Многоуровневая фазовая манипуляция

Описанную схему можно расширить: передавать, например, по три бита в каждый момент времени, используя для этого восемь различных углов сдвига фаз. Более того, при каждом угле можно использовать несколько амплитуд. В стандартном модеме со скоростью 9 600 бит/с, например, использованы 12 углов сдвига фаз, причем при четырех из них используются сигналы двух различных амплитуд.

На последнем примере отлично видна разница между скоростью передачи данных  $R$  (биты в секунду) и скоростью модуляции (боды) сигнала. Предположим, что на вход этой схемы подается цифровой сигнал, кодированный следующим образом: каждый бит представляется уровнем постоянного напряжения, двоичный ноль — одним, двоичная единица — другим. Скорость передачи данных равна  $R = 1/t_b$ . В то же время в каждой сигнальной посылке кодированного сигнала содержится  $L = 4$  бит, для чего использованы  $M = 16$  различных комбинаций амплитуды и фазы. Сле-



довательно, скорость модуляции равна  $R/4$ , поскольку при каждой замене сигнальной посылки передается четыре бита. Значит, скорость передачи сигнала в линии равна 2 400 бод, а скорость передачи данных равна 9 600 бит/с. Именно таким способом, используя сложные схемы модуляции, можно достичь высоких скоростей передачи данных по телефонным линиям. Подытожим сказанное.

$$D = \frac{R}{L} = \frac{R}{\log_2 M}, \quad (6.7)$$

где

$D$  — скорость модуляции, бод;

$R$  — скорость передачи данных, бит/с;

$M$  — число различных сигнальных посылок,  $= 2^L$ ;

$L$  — число битов в сигнальной посылке.

## Производительность

При рассмотрении производительности различных методов модулирования цифровых данных с помощью аналоговых сигналов первым параметром, представляющим интерес, является ширина полосы модулированного сигнала. Зависит данный параметр от многих факторов, в том числе от используемого определения ширины полосы и методов фильтрации, применяемых для создания полосового сигнала. Приведем некоторые результаты, взятые из работы [COUC01].

Ширина полосы пропускания  $B_T$  для схемы ASK равна

$$\text{ASK} \quad B_T = (1 + r)R, \quad (6.8)$$

где  $R$  — скорость передачи битов, а  $r$  связано с методом фильтрации сигнала для создания полосы пропускания; как правило,  $0 < r < 1$ . Следовательно, ширина полосы непосредственно связана со скоростью передачи битов. Приведенная формула также справедлива и для схемы PSK.

При использовании схемы FSK ширину полосы можно выразить следующим образом:

$$\text{FSK} \quad B_T = 2\Delta F + (1 + r)R, \quad (6.9)$$

где  $\Delta F = f_2 - f_c = f_c - f_1$  — смещение модулируемой частоты относительно несущей. При использовании очень высоких частот доминирует член с  $\Delta F$ . Например, в одном из стандартов передачи FSK-кодированных сигналов по локальной сети, созданной на основе коаксиального кабеля, указаны  $\Delta F = 1,25$  МГц,  $f_c = 5$  МГц и  $R = 1$  Мбит/с; в этом случае член  $2\Delta F = 2,5$  МГц доминирует. Если же рассмотреть упоминавшийся в предыдущем параграфе модем Bell 108, то при  $\Delta F = 100$  Гц,  $f_c = 1\,170$  Гц (в одном направлении), а  $R = 300$  бит/с доминирующим будет член  $(1 + r)R$ .

При использовании многофазной передачи сигналов можно добиться значительно более эффективного использования полосы:

$$\text{MPSK} \quad B_T = \left(\frac{1+r}{L}\right)R = \left(\frac{1+r}{\log_2 M}\right)R, \quad (6.10)$$

где  $L$  — число битов, закодированных в одной сигнальной посылке, а  $M$  — количество различных сигнальных посылок.

Для многоуровневой FSK (MFSK) имеем следующее:

$$\text{MFSK } B_r = \left( \frac{(1+r)M}{\log_2 M} \right) R. \quad (6.11)$$

В табл. 6.2 показано отношение скорости передачи данных  $R$  к ширине полосы пропускания различных схем. Отметим, что данное отношение также называется эффективностью использования полосы и, как можно догадаться, является мерой эффективности, с которой полосу можно использовать для передачи данных. Итак, теперь мы можем утверждать, что преимущества многоуровневых методов передачи сигналов стали очевидными.

**Таблица 6.2. Отношение скорости передачи данных к ширине полосы пропускания различных схем кодирования цифровых данных аналоговыми сигналами**

	$r = 0$	$r = 0,5$	$r = 1$
Амплитудная манипуляция	1,0	0,67	0,5
Бинарная частотная манипуляция			
Широкополосная ( $\Delta F \gg R$ )	-0	-0	-0
Узкополосная ( $\Delta F \approx f_c$ )	1,0	0,67	0,5
Фазовая манипуляция	1,0	0,67	0,5
Многофазная передача сигналов			
$L = 4, b = 2$	2,00	1,33	1,00
$L = 8, b = 3$	3,00	2,00	1,50
$L = 16, b = 4$	4,00	2,67	2,00
$L = 32, b = 5$	5,00	3,33	2,50

Разумеется, все предыдущее обсуждение относится к спектру сигнала на входе линии связи. Ничего не говорилось о производительности в присутствии помех. Некоторые результаты, полученные на основании разумных допущений относительно передающих систем [COUC01], обобщаются на рис. 6.8. В данном случае скорость появления ошибочных битов изображена как функция отношения  $E_b/N_0$ , определенного в главе 5. Очевидно, что по мере увеличения этого отношения скорость появления ошибочных битов падает. Более того, методы DPSK и BPSK дают выигрыш около 3 дБ по сравнению с методами ASK и BFSK.

Подобная информация для различного числа уровней  $M$  схем MFSK и MSPK представлена на рис. 6.9. Следует отметить важное отличие. Для MFSK вероятность ошибки при данном значении  $E_b/N_0$  уменьшается по мере увеличения  $M$ , а для схемы MSPK справедливо обратное. С другой стороны, сравнивая формулы (6.10) и (6.11), мы видим, что эффективность использования полосы для схемы MFSK уменьшается при увеличении  $M$ , тогда как для MSPK верно обратное.

**Пример.** Чему равна эффективность использования полосы при использовании частотной, амплитудной, фазовой и квадратурной фазовой манипуляции при скорости появления ошибок  $10^{-7}$  в канале с отношением сигнал/шум, равным 12 дБ?

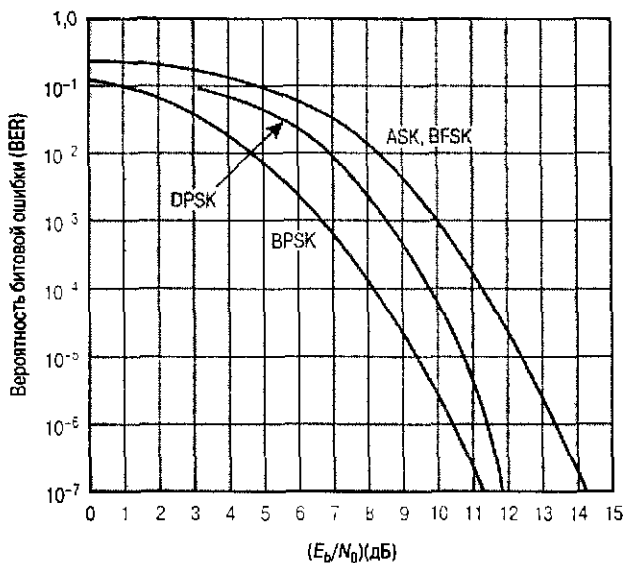


Рис. 6.8. Теоретическая частота битовых ошибок для различных схем кодирования

Используя формулу (5.4), имеем

$$\frac{E_b}{N_0} = 12 \text{ дБ} - \left( \frac{R}{B_T} \right)_{\text{дБ}}.$$

Используя рис. 6.8, получаем для частотной и амплитудной манипуляции

$$\frac{E_b}{N_0} = 14,2 \text{ дБ},$$

$$\left( \frac{R}{B_T} \right)_{\text{дБ}} = -2,2 \text{ дБ},$$

$$\frac{R}{B_T} = 0,6.$$

Используя рис. 6.8, получаем для фазовой манипуляции

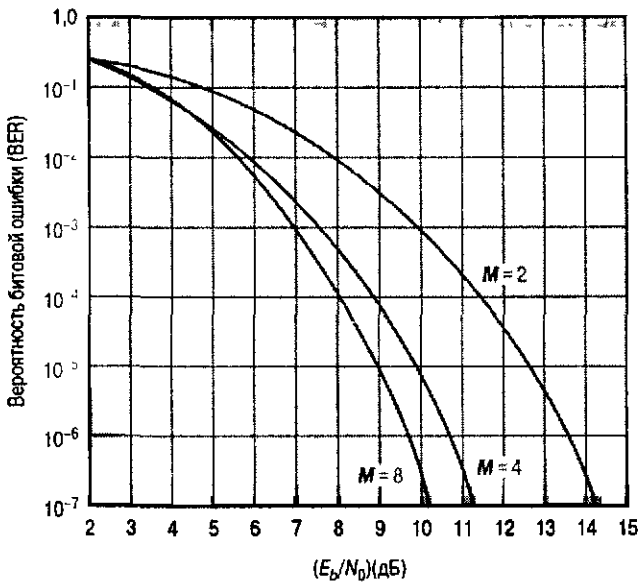
$$\frac{E_b}{N_0} = 11,2 \text{ дБ},$$

$$\left( \frac{R}{B_T} \right)_{\text{дБ}} = 0,8 \text{ дБ},$$

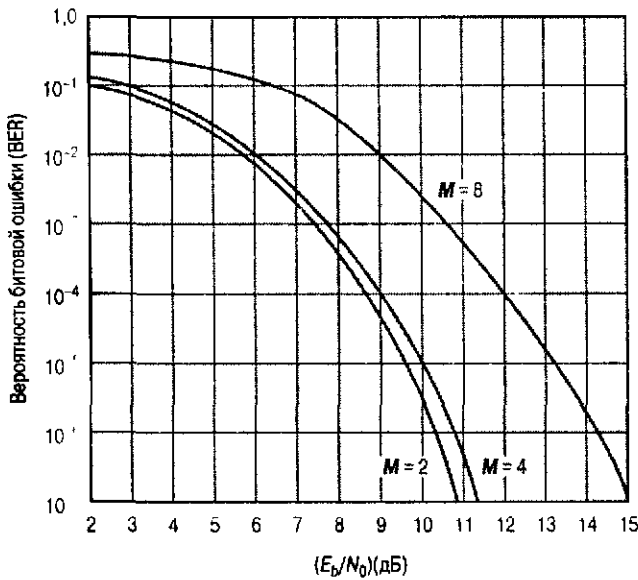
$$\frac{R}{B_T} = 1,2.$$

Чтобы получить ответ для квадратурной фазовой манипуляции, требуется учесть, что скорость двоичной передачи равна  $D = R/2$ . Следовательно,

$$\frac{R}{B_T} = 2,4.$$



а) Многоуровневая FSK (MFSK)



б) Многоуровневая PSK (MPSK)

Рис. 6.9. Теоретическая частота битовых ошибок для многоуровневых FSK и PSK

Как показано в предыдущем примере, эффективности использования полосы при амплитудной и частотной манипуляциях равны, эффективность использования полосы при фазовой манипуляции выше, а значительного улучшения можно добиться при использовании многоуровневой передачи сигналов.

## Квадратурная амплитудная модуляция

Квадратурная амплитудная модуляция (quadrature amplitude modulation — QAM) является популярным методом аналоговой передачи сигналов, используемым в некоторых беспроводных стандартах. Данная схема модуляции совмещает в себе амплитудную и фазовую модуляции. В методе QAM использованы преимущества одновременной передачи двух различных сигналов на одной несущей частоте, но при этом задействованы две копии несущей частоты, сдвинутые относительно друг друга на  $90^\circ$ . При квадратурной амплитудной модуляции обе несущие являются амплитудно-модулированными. Итак, два независимых сигнала одновременно передаются через одну среду. В приемнике эти сигналы демодулируются, а результаты объединяются с целью восстановления исходного двоичного сигнала.

В общих чертах модуляционная схема QAM показана на рис. 6.10. Со скоростью  $R$  бит/с на вход поступает поток двоичных цифр. Этот поток разбивается на два потока (биты попеременно распределяются по двум отдельным потокам), передаваемых со скоростью  $R/2$  бит/с каждый. Обратимся к рисунку: верхний поток модулируется на несущей частоте  $f_c$  с использованием схемы ASK, для чего двоичный поток умножается на несущую. Таким образом, двоичный ноль представляется отсутствием несущей волны, а двоичная единица — наличием несущей волны постоянной амплитуды. Для модулирования нижнего потока та же несущая волна смещается на  $90^\circ$ , после чего вновь используется схема ASK. Затем два модулированных сигнала складываются и передаются вместе. Суммарный переданный сигнал можно записать следующим образом:

$$\text{QAM } s(t) = d_1(t)\cos 2\pi f_c t + d_2(t)\sin 2\pi f_c t.$$

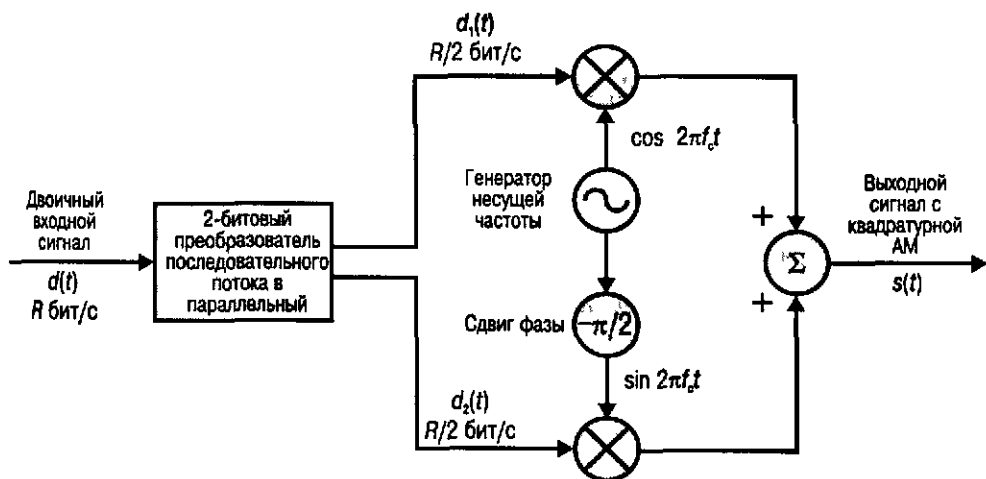


Рис. 6.10. Модулятор QAM

При использовании двухуровневой амплитудной манипуляции каждый из двух потоков может находиться в одном из двух состояний, а объединенный поток — в одном из  $2 \times 2 = 4$  состояний. При использовании четырехуровневой манипуляции (т.е. четырех различных уровней амплитуды) объединенный поток будет находиться в одном из  $4 \times 4 = 16$  состояний. Уже ре-

лизованы системы, имеющие 64 или даже 256 состояний. Чем больше число состояний, тем выше скорость передачи данных, возможная при определенной ширине полосы. Разумеется, как указывалось ранее, чем больше число состояний, тем выше потенциальная частота возникновения ошибок вследствие помех или поглощения.

## 6.3. АНАЛОГОВЫЕ ДАННЫЕ, АНАЛОГОВЫЕ СИГНАЛЫ

Модулирование можно определить как процесс объединения входного сигнала  $m(t)$  и несущего сигнала с частотой  $f_c$  с целью создания сигнала, полоса которого (как правило) центрирована на частоте  $f_c$ . Для цифровых сигналов причина модулирования должна быть очевидной: при наличии только средств аналоговой передачи модулирование необходимо для преобразования цифровых данных в аналоговую форму. Причина использования модулирования при наличии аналоговых данных не настолько очевидна. Все же речевые сигналы передаются по телефонным линиям с сохранением исходного спектра (такой процесс называется *немодулированной передачей*). Среди причин аналогового модулирования аналоговых сигналов можно выделить следующие.

- Эффективная передача может требовать более высоких частот. При ненаправленной передаче передавать немодулированные сигналы практически невозможно; в этом случае потребовались бы километровые антенны.
- Модуляция допускает частотное уплотнение, важную технологию, о которой речь шла в главе 2.

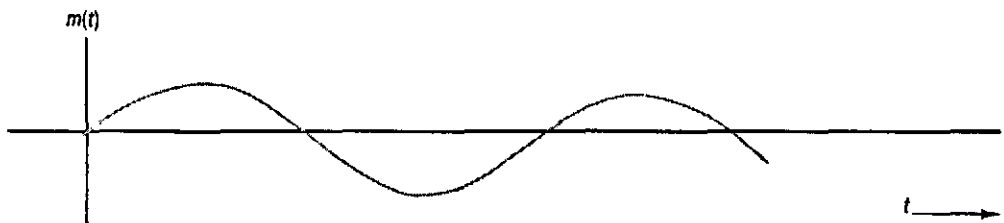
В данном разделе мы рассмотрим основные технологии модуляции с использованием аналоговых данных: амплитудную модуляцию (АМ), частотную модуляцию (FM) и фазовую модуляцию (PM). Как и ранее, в процессе модуляции используются три основные характеристики сигнала.

### Амплитудная модуляция

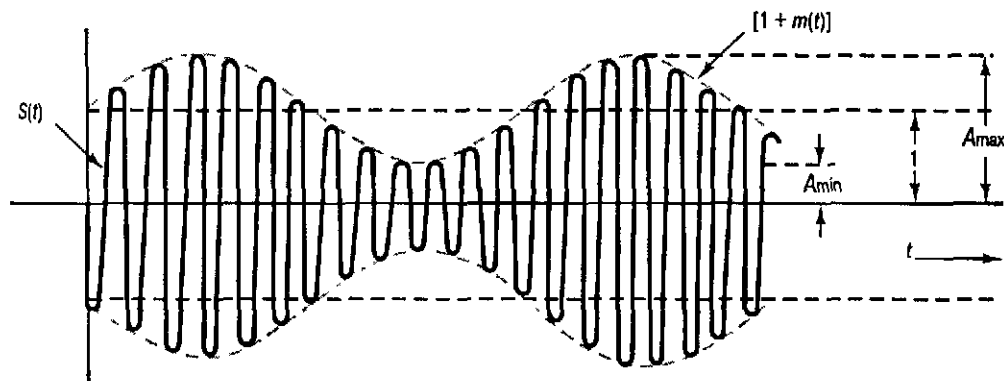
Амплитудная модуляция (АМ) представляет собой простейшую форму модуляции, она изображена на рис. 6.11. Математически данный процесс можно описать следующим образом:

$$\text{для амплитудной модуляции } s(t) = [1 + n_a x(t)] \cos 2\pi f_c t, \quad (6.12)$$

где  $\cos 2\pi f_c t$  — несущий, а  $x(t)$  — входной сигнал (несущий информацию), оба приведенные к единичной амплитуде. Параметр  $n_a$  называется *коэффициентом модуляции* и представляет собой отношение амплитуд входного сигнала и несущего. Согласно использованной ранее форме записи входной сигнал равен  $m(t) = n_a x(t)$ . Единица в предыдущей формуле представляет собой постоянную составляющую, предупреждающую, как будет показано далее, потерю информации. Данная схема также известна как *двухполосная связь с передачей несущей* (double sideband transmitted carrier — DSBTC).



а) Модулирующая синусоида



б) Результирующий амплитудно-модулированный сигнал

Рис. 6.11. Амплитудная модуляция

**Пример.** Получим выражение для сигнала  $s(t)$ , если его амплитуда модулируется сигналом  $x(t) = \cos 2\pi f_m t$ .

Имеем

$$s(t) = [1 + n_a \cos 2\pi f_m t] \cos 2\pi f_c t.$$

Используя тригонометрические формулы, получаем

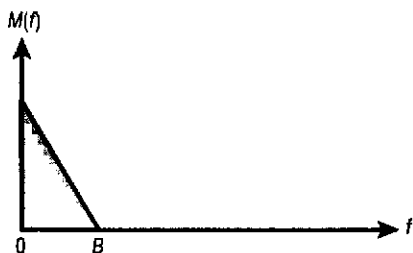
$$s(t) = \cos 2\pi f_c t + \frac{n_a}{2} \cos 2\pi(f_c - f_m)t + \frac{n_a}{2} \cos 2\pi(f_c + f_m)t.$$

Результирующий сигнал состоит из гармоники исходной несущей частоты и двух гармоник, смещенных относительно несущей на  $f_m$  Гц.

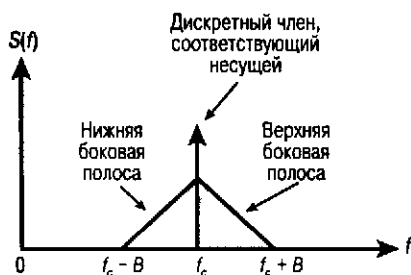
Из предыдущей формулы и рис. 6.11 видно, что при амплитудной модуляции происходит умножение входного сигнала на несущий. Огибающая результирующего сигнала равна  $[1 + n_a x(t)]$ , и поскольку  $n_a < 1$ , то огибающая точно воспроизводит исходный сигнал. Если  $n_a > 1$ , то огибающая пересечет временную ось и информация будет потеряна.

Полезно посмотреть на спектр амплитудно-модулированного сигнала (рис. 6.12). Спектр сигнала состоит из спектра исходного несущего сигнала и спектра входного сигнала, сдвинутого на  $f_c$ . Часть спектра  $|f| > |f_c|$  называется *верхней боковой полосой*, а часть спектра  $|f| < |f_c|$  — *нижней боковой полосой*. И верхняя, и нижняя боковые полосы точно копируют исходный спектр  $M(f)$ , причем в нижней боковой по-

лосе частоты следуют в обратном порядке. Рассмотрим в качестве примера речевой сигнал с полосой от 300 до 3 000 Гц, модулированный на несущей частоте 60 кГц. Результирующий сигнал состоит из верхней полосы, лежащей в диапазоне от 60,3 до 63 кГц, нижней — в диапазоне от 57 до 59,7 кГц, и несущей частоты 60 кГц.



а) Спектр модулирующего сигнала



б) Спектр амплитудно-модулированного сигнала с несущей  $f_c$

Рис. 6.12. Спектр амплитудно-модулированного сигнала

Отметим важное соотношение:

$$P_t = P_c \left( 1 + \frac{n_a^2}{2} \right),$$

где  $P_t$  — общая переданная мощность сигнала  $s(t)$ , а  $P_c$  — мощность, переданная на несущей частоте. Хотелось бы, чтобы  $n_a$  было максимально большим, чтобы большая часть мощности сигнала использовалась для передачи информации. Впрочем,  $n_a$  все же должно быть меньше 1.

Должно быть очевидным, что, вследствие наличия в каждой боковой полосе полного спектра сигнала  $m(t)$ , сигнал  $s(t)$  содержит лишние компоненты. Этот факт используется при *однополосной связи* (single sideband — SSB): одна из боковых полос и несущая частота отделяются, а передается только вторая боковая полоса. Перечислим основные преимущества такого подхода.

- Требуется только половина ширины полосы; т.е.  $B_T = B$ , где  $B$  — ширина полосы исходного сигнала. В методе же передачи несущей с двойной боковой полосой  $B_T = 2B$ .



- Для передачи сигнала требуется меньшая мощность, поскольку она уже не расходуется на передачу несущей и одной из боковых полос. Еще одним вариантом является использование метода передачи двойной боковой полосы с подавлением несущей (double sideband suppressed carrier — DSBSC), при котором несущая частота отфильтровывается, а передаются обе боковые полосы. Этот метод несколько экономит мощность, но использует такую же полосу, как и метод передачи несущей с двойной боковой полосой.

Недостатком метода подавления несущей является то, что несущая может использоваться для синхронизации. Предположим, например, что исходный аналоговый сигнал содержит цифровые данные, закодированные сигналом с амплитудной манипуляцией. Чтобы приемник правильно расшифровал данные, он должен знать начальную точку каждого периода прохождения бита. Постоянная несущая предоставляет механизм синхронизации, с помощью которого рассчитывается время поступления бита. Альтернативой является метод подавления боковой полосы (VSB — vestigial sideband), при котором используется одна боковая полоса и несущая сниженной мощности.

## Угловая модуляция

Частотная и фазовая модуляции являются частными случаями угловой модуляции. Модулированный сигнал выражается следующим образом:

$$\text{при угловой модуляции } s(t) = A_c \cos[2\pi f_c t + \phi(t)]. \quad (6.13)$$

При фазовой модуляции фаза пропорциональна модулирующему сигналу:

$$\text{при фазовой модуляции } \phi(t) = n_p m(t), \quad (6.14)$$

где  $n_p$  — коэффициент фазовой модуляции.

При частотной модуляции модулирующему сигналу пропорциональна производная фазы:

$$\text{при частотной модуляции } \phi'(t) = n_f m(t), \quad (6.15)$$

где  $n_f$  — коэффициент частотной модуляции.

Чтобы более подробно объяснить вышесказанное с математической точки зрения, можно рассмотреть следующий пример. Фаза сигнала  $s(t)$  в любой момент времени равна  $2\pi f_c t + \phi(t)$ . Мгновенное отклонение фазы от несущего сигнала равно  $\phi(t)$ . При фазовой модуляции это мгновенное отклонение фазы пропорционально  $m(t)$ . Поскольку частоту можно определить как скорость изменения фазы сигнала, то мгновенная частота сигнала  $s(t)$  равна

$$2\pi f_i(t) = \frac{d}{dt} [2\pi f_c t + \phi(t)],$$

$$f_i(t) = f_c + \frac{1}{2\pi} \phi'(t),$$

т.е. мгновенное отклонение частоты от несущей равно  $\phi'(t)$ , что при частотной модуляции пропорционально  $m(t)$ .

На рис. 6.13 показаны амплитудная, частотная и фазовая модуляции синусоидальной волны. Формы сигналов с частотной и фазовой модуляцией весьма схожи. В действительности их невозможно различить без знания модулирующей функции.

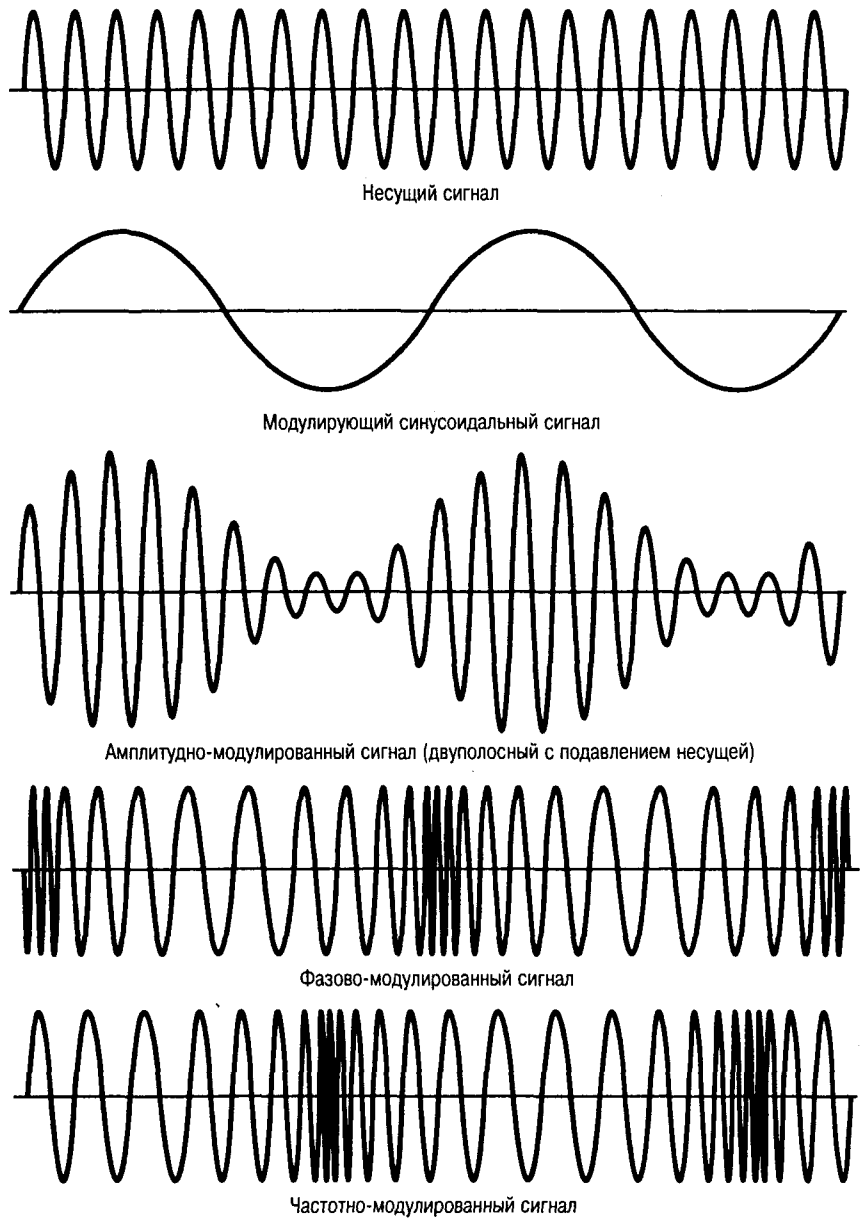


Рис. 6.13. Амплитудная, фазовая и частотная модуляции синусоидальной несущей синусоидальным сигналом

Приведем несколько наблюдений, связанных с процессом частотной модуляции. Легко видеть, что максимальное отклонение  $\Delta F$  следующее:

$$\Delta F = \frac{1}{2\pi} n_f A_m \text{ Гц},$$

где  $A_m$  — максимальное значение  $m(t)$ . Следовательно, увеличение  $m(t)$  приведет к увеличению  $\Delta F$ , которое, как можно предположить, увеличит полосу передачи  $B_T$ . Впрочем, как видно из рис. 6.13, это не увеличивает средний уровень мощности частотно-модулированного сигнала, равный  $A_c^2/2$ . При амплитудной же модуляции имеем абсолютно противоположную картину: уровень модуляции влияет на мощность амплитудно-модулированного сигнала, но не влияет на ширину полосы.

**Пример.** Определим выражение для  $s(t)$ , если его фаза модулируется сигналом  $\phi(t) = n_p \cos 2\pi f_m t$ . Предположим, что  $A_c = 1$ . Непосредственно из формул раздела получаем

$$s(t) = \cos[2\pi f_c t + n_p \cos 2\pi f_m t].$$

Мгновенное отклонение фазы от несущего сигнала равно  $n_p \cos 2\pi f_m t$ . Угол сдвига фаз сигнала изменяется по синусоиде от его немодулированного значения, причем максимальное отклонение равно  $n_p$ .

Последнее выражение можно записать, используя тригонометрические тождества Бесселя:

$$s(t) = \sum_{n=-\infty}^{\infty} J_n(n_p) \cos\left(2\pi f_c t + 2\pi n f_m t + \frac{n\pi}{2}\right),$$

где  $J_n(n_p)$  — функция Бесселя первого рода  $n$ -го порядка. Используя свойство

$$J_{-n}(x) = (-1)^n J_n(x),$$

последнее равенство можно записать в следующем виде:

$$s(t) = J_0(n_p) \cos 2\pi f_c t + \sum_{n=1}^{\infty} J_n(n_p) \left[ \cos\left(2\pi(f_c + n f_m)t + \frac{n\pi}{2}\right) + \cos\left(2\pi(f_c - n f_m)t + \frac{(n+2)\pi}{2}\right) \right].$$

Результирующий сигнал содержит гармонику исходной несущей частоты, а также набор боковых полос, смещенных от  $f_c$  на все возможные частоты, кратные  $f_m$ . При  $n_p \ll 1$  члены высших порядков быстро уменьшаются.

**Пример.** Определим выражение для  $s(t)$ , если его частота модулируется по закону  $\phi'(t) = -n_f \sin 2\pi f_m t$ . Вид  $\phi'(t)$  выбран для удобства. Имеем

$$\phi(t) = - \int n_f \sin 2\pi f_m t \, dt = \frac{n_f}{2\pi f_m} \cos 2\pi f_m t.$$

Следовательно,

$$s(t) = \cos\left[2\pi f_c t + \frac{n_f}{2\pi f_m} \cos 2\pi f_m t\right],$$

$$s(t) = \cos \left[ 2\pi f_c t + \frac{\Delta F}{f_m} \cos 2\pi f_m t \right].$$

Мгновенный сдвиг частоты от несущей равен  $-n_f \sin 2\pi f_m t$ . Частота сигнала изменяется по синусоиде от немодулированного значения, причем максимальное значение сдвига фаз равно  $n_f$  радиан в секунду.

Формула для частотно-модулированного сигнала идентична формуле для фазово-модулированного сигнала с заменой  $\Delta F/f_m$  на  $n_p$ . Следовательно, разложение по функциям Бесселя также аналогично.

Так же, как и в случае с амплитудной модуляцией, частотная и фазовая модуляции порождают сигнал с полосой, центрированной на частоте  $f_c$ . Впрочем, теперь мы можем видеть, что размеры этих полос весьма отличаются. Амплитудная модуляция представляет собой линейный процесс, поэтому она порождает частоты, равные сумме и разности несущего сигнала и гармоник модулируемого сигнала. Следовательно, для амплитудной модуляции

$$B_T = 2B.$$

В то же время угловая модуляция включает член вида  $\cos(\phi(t))$ , который нелинеен и порождает большой диапазон частот. По сути, для модулируемой синусоиды с частотой  $f_m$  сигнал  $s(t)$  будет содержать гармоники с частотами  $f_c + f_m$ ,  $f_c + 2f_m$  и т.д. В более общем случае для передачи частотно- или фазово-модулированного сигнала требуется бесконечная полоса. На практике хорошим эмпирическим правилом, известным как правило Карсона (Carson) [COUC01], является следующее:

$$B_T = 2(\beta + 1)B,$$

где

$$\beta = \begin{cases} n_p A_m & \text{для фазовой модуляции} \\ \frac{\Delta F}{B} = \frac{n_f A_m}{2\pi B} & \text{для частотной модуляции} \end{cases}$$

Формулу для частотной модуляции мы можем переписать в таком виде:

$$B_T = 2\Delta F + 2B. \tag{6.16}$$

Следовательно, и частотная, и фазовая модуляции требуют большей полосы, чем амплитудная.

## 6.4. АНАЛОГОВЫЕ ДАННЫЕ, ЦИФРОВЫЕ СИГНАЛЫ

В этом разделе мы рассмотрим процесс трансформации аналоговых данных в цифровые сигналы. Строго говоря, уместнее будет назвать этот процесс *преобразованием аналоговых данных в цифровые данные*, или *оцифровыванием данных*. После преобразования аналоговых данных в цифровые могут произойти изменения. Наиболее частыми являются три случая.

1. Цифровые данные могут передаваться с использованием кодировки NRZ-L<sup>2</sup>. В этом случае мы фактически имеем прямой переход аналоговых данных в цифровой сигнал.
2. Цифровые данные могут кодироваться как цифровой сигнал с использованием кодировки, отличной от NRZ-L. В этом случае требуются дополнительные шаги.
3. Цифровые данные могут преобразовываться в аналоговый сигнал с использованием одного из методов модуляции, описанных в разделе 6.2.

Последняя процедура, кажущаяся несколько странной, проиллюстрирована на рис. 6.14, где представлен процесс оцифровки речевой информации с последующим ее преобразованием в аналоговый сигнал с амплитудной манипуляцией. Такая последовательность действий делает возможной цифровую передачу в смысле, определенном в главе 2. Оцифрованную речевую информацию можно рассматривать как цифровые данные, даже несмотря на то что требования передачи (например, использование СВЧ) предписывают необходимость применения аналоговых сигналов.



Рис. 6.14. Оцифровка аналоговых данных

Устройство, с целью передачи выполняющее преобразование аналоговых данных в цифровую форму с последующим восстановлением исходных аналоговых данных из цифровых, называется кодеком (кодер-декодером). В этом разделе мы рассмотрим два основных метода, используемых в кодеках, импульсно-кодировую модуляцию и дельта-модуляцию. Заканчивается данный раздел обсуждением сравнительных характеристик.

## Импульсно-кодировая модуляция

Импульсно-кодировая модуляция (pulse code modulation — PCM) опирается на теорию о дискретном представлении, которая формулируется следующим образом.

Если из сигнала  $f(t)$  через постоянные промежутки времени производятся выборки, причем скорость выборки больше, чем удвоенная максимальная частота сигнала, то такие выборки содержат всю информацию об исходном

<sup>2</sup> Метод без возврата к нулевому уровню (Nonreturn to Zero-Level — NRZ-L) — это наиболее распространенный и наиболее легкий способ передачи цифровых сигналов путем использования разных уровней напряжения для двух различных двоичных чисел: Постоянное положительное напряжение представляет двоичный 0, а постоянное отрицательное напряжение — двоичную 1. Основное назначение данной кодировки — генерирование или интерпретация цифровых данных посредством терминалов или других устройств. Если для передачи используется какой-то другой код, он обычно создается системой передачи сигнала NRZ-L.

сигнале. Восстановить функцию  $f(t)$  из выборок можно с помощью фильтров нижних частот.

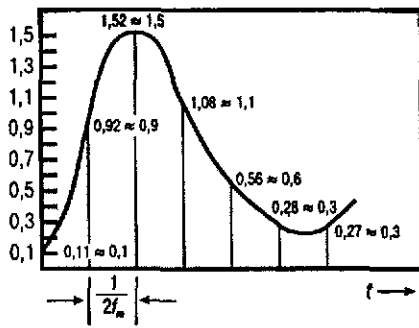
Для читателей, заинтересовавшихся этим вопросом, в приложении 6А приводится доказательство данной теоремы. Если речевая информация ограничена частотами, лежащими ниже 4000 Гц, чего вполне достаточно для ее разборчивости, то 8 000 выборок в секунду будет достаточно для полного изображения речевого сигнала. Отметим, впрочем, что эти выборки являются аналоговыми и называются выборками с *амплитудно-импульсной модуляцией* (pulse amplitude modulation — PAM). Чтобы преобразовать эти выборки в цифровые, каждой из них должен быть присвоен двоичный код. На рис. 6.15 показано, как каждая выборка аппроксимируется путем ее “квантования” на один из 16 различных уровней. После этого каждая выборка может представляться четырьмя битами. Но поскольку оцифрованные значения являются только аппроксимациями, то точно восстановить исходный сигнал невозможно. Улучшить качество восстанавливаемого речевого сигнала можно с помощью восьмибитовой выборки, допускающей 256 уровней квантования. В этом случае качество сигнала сравнимо с качеством, получаемым при аналоговой передаче. Отметим, что такая выборка предполагает скорость передачи данных 8 000 выборок в секунду  $\times$  8 бит в выборке = 64 Кбит/с.

Итак, при импульсно-кодовой модуляции из непрерывного сигнала с непрерывной амплитудой (аналогового сигнала) вначале создается цифровой сигнал. Этот цифровой сигнал состоит из  $n$ -битовых блоков, причем каждый  $n$ -бит содержит амплитуду модулированного импульса. После получения сигнала процесс повторяется в обратной последовательности с целью восстановления аналогового сигнала. Отметим, впрочем, что в процессе нарушаются условия теоремы о дискретном представлении. Вследствие оцифровывания модулированного импульса исходный сигнал является теперь только аппроксимацией и не может быть восстановлен точно. Этот эффект известен как ошибка квантования или квантовый шум. Отношение сигнал/шум для квантового шума даже в работе [GIBS93] и составляет

$$\text{SNR}_{\text{дБ}} = 20 \lg 2^n + 1,76 \text{ дБ} = 6,02n + 1,76 \text{ дБ}.$$

Следовательно, каждый дополнительный бит, используемый для оцифровывания, увеличивает отношение сигнал/шум приблизительно на 6 дБ или в 4 раза.

Улучшить метод импульсно-кодовой модуляции можно за счет схемы, известной как *нелинейное кодирование*, что, по сути, означает неравномерное размещение уровней квантования. Использование равномерного размещения невыгодно тем, что средняя абсолютная ошибка для каждой выборки одинакова, вне зависимости от уровня сигнала. Следовательно, значения с меньшей амплитудой искажаются сравнительно сильнее. Если же для сигналов с низкими амплитудами использовать большее число уровней квантования, а для сигналов с большими амплитудами — меньшее число уровней, то можно достичь заметного снижения уровня искажений сигнала, как показано, например, на рис. 6.16.



а)

Число	Двоичный эквивалент	Сигнал РСМ
0	0000	—
1	0001	—
2	0010	—
3	0011	—
4	0100	—
5	0101	—
6	0110	—
7	0111	—
8	1000	—
9	1001	—
10	1010	—
11	1011	—
12	1100	—
13	1101	—
14	1110	—
15	1111	—

б)

Рис. 6.15. Импульсно-кодовая модуляция

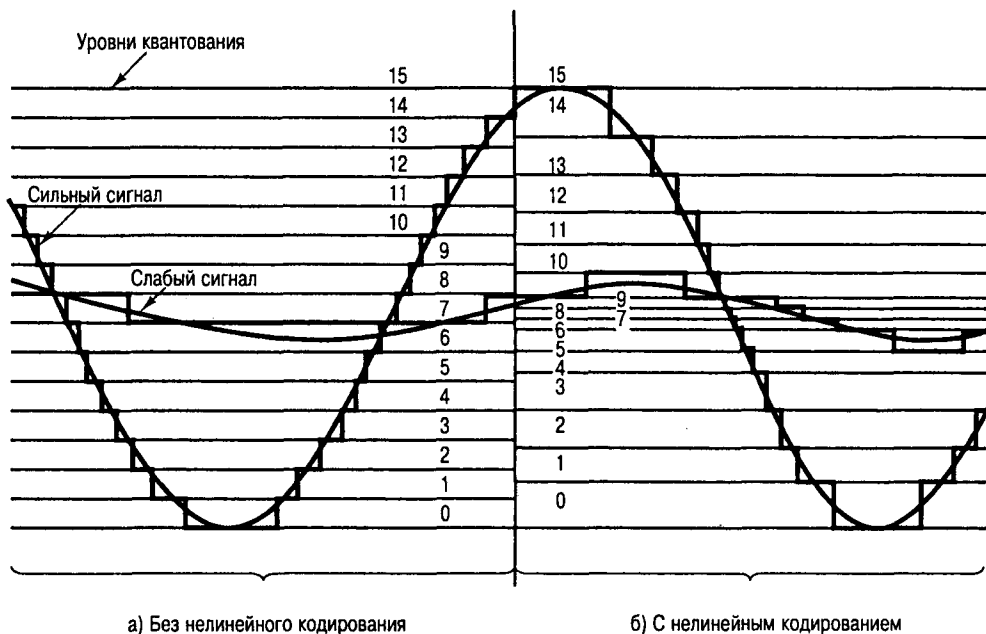


Рис. 6.16. Результат использования нелинейного кодирования

Того же результата можно добиться, используя и квантование с постоянным шагом, но компандируя (усиливая-сжимая) входящий аналоговый сигнал. Компандирование — это процесс, при котором сильные сигналы на входе усиливаются меньше, чем слабые, вследствие чего происходит сжатие диапазона интенсивности сигнала. На выходе же выполняется обратная операция. Пример функций компандирования приведен на рис. 6.17. Отметим, что на входе компандера эффект заключается в сжатии выборки, так что большие значения уменьшаются относительно меньших. Таким образом, при фиксированном числе уровней квантования сигналам низкого уровня выделяется больше уровней. На выходе компандер расширяет выборки, и сжатые величины восстанавливают исходное значение.

Нелинейное кодирование может значительно улучшить отношение сигнал/шум импульсно-кодовой модуляции. Для речевых сигналов, например, можно получить улучшение от 24 до 30 дБ.

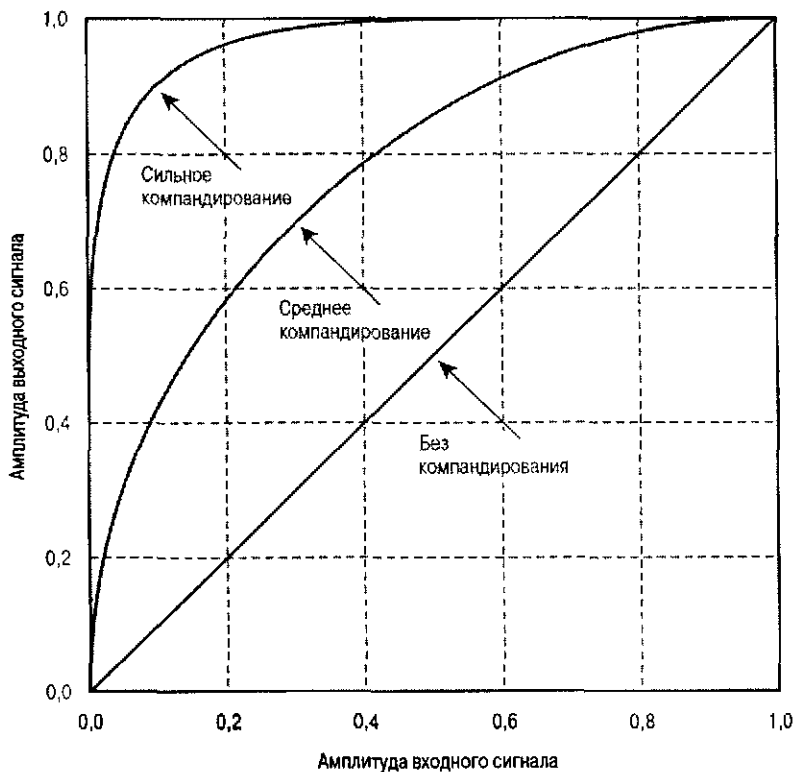
## Дельта-модуляция

Для улучшения производительности импульсно-кодовой модуляции или для упрощения этого вида кодировки могут использоваться различные методики. Одной из наиболее популярных альтернатив является дельта-модуляция (delta modulation — DM).

При использовании дельта-модуляции поступающий аналоговый сигнал аппроксимируется ступенчатой функцией, значение которой в каждом интервале выборки  $T$ , увеличивается или уменьшается на один уровень квантования  $\Delta$ . Пример такой функции приведен на рис. 6.13, где ступенчатая функция наклад-



дывается на исходный аналоговый сигнал. Важной характеристикой данной ступенчатой функции является бинарное поведение: в каждый момент взятия выборки значение функции уменьшается или увеличивается на постоянную величину. Следовательно, выход процесса дельта-модуляции каждой выборки можно представить одной двоичной цифрой. По сути, двоичный поток генерируется путем аппроксимации производной аналогового сигнала, а не его амплитуды: 1 генерируется, если ступенчатая функция возрастает на следующем интервале, и 0 — в противоположном случае.



*Рис. 6.17. Пример функций компрессирования*

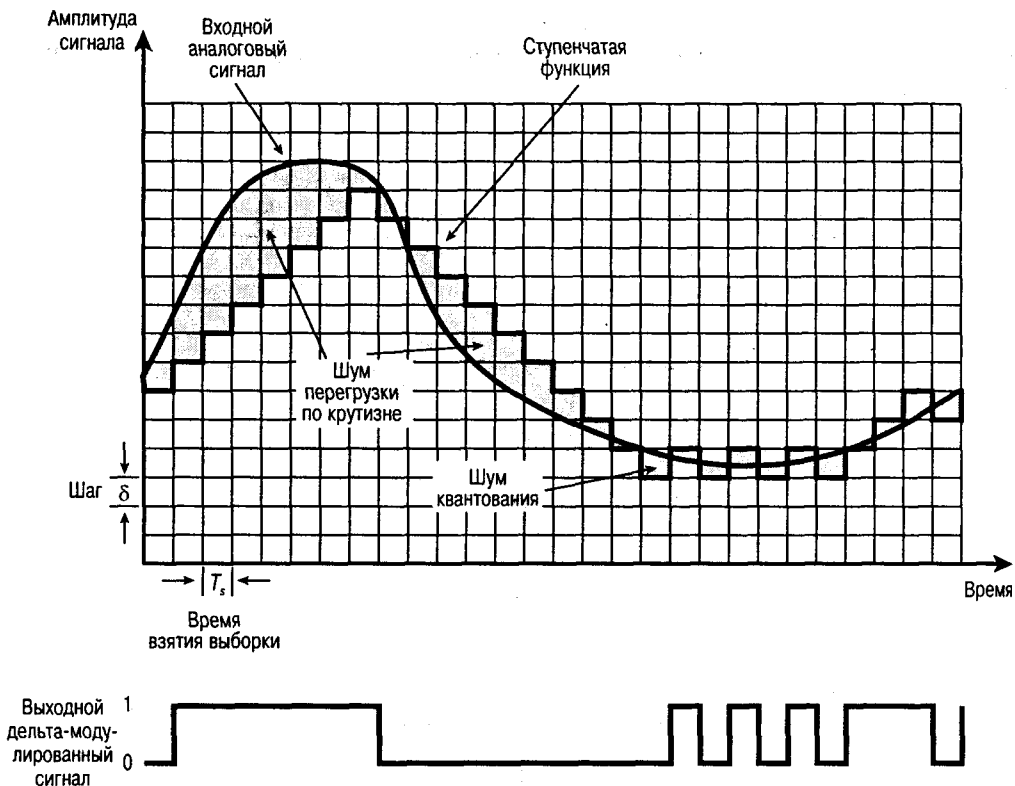
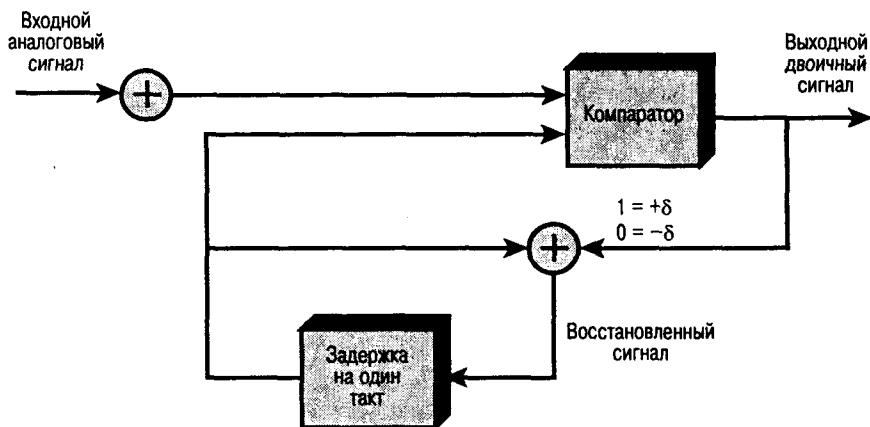


Рис. 6.18. Пример дельта-модуляции

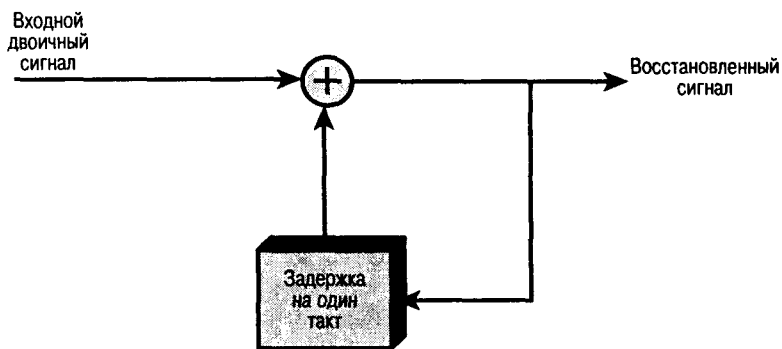
Направление перехода (вверх или вниз), происходящего в каждый момент взятия выборки, определяется так, чтобы ступенчатая функция максимально соответствовала исходному аналоговому сигналу. Логика процесса изображена на рис. 6.19 и представляет, по сути, механизм обратной связи. При передаче происходит следующее: в каждый момент взятия выборки входной аналоговый сигнал сравнивается с последним значением аппроксимирующей ступенчатой функции. Если величина оцифровываемого сигнала превышает значение ступенчатой функции, то генерируется 1; в противном случае генерируется 0. Таким образом, изменение ступенчатой функции всегда происходит по направлению к входному сигналу. Выходом же процесса дельта-модуляции является двоичная последовательность, которую приемник может использовать для восстановления ступенчатой функции. После этого полученная функция сглаживается посредством некоторого интеграционного процесса или при пропускании ее через фильтр нижних частот, что позволяет создать аналоговую аппроксимацию входного аналогового сигнала.

В схеме дельта-модуляции имеется два важных параметра: размер шага  $\delta$ , выделенного каждому двоичному символу, и частота выборки. Как показано на рис. 6.18, выбор  $\delta$  является компромиссом между двумя типами ошибок или шумов. При очень медленном изменении аналогового сигнала будет существовать квантовый шум, возрастающий по мере увеличения  $\delta$ . Однако если аналого-

вая форма волны изменяется настолько быстро, что ступенчатая функция не может за ней поспеть, то возникает шум перегрузки по крутизне, возрастающий по мере уменьшения  $\delta$ .



а) Передача



б) Прием

Рис. 6.19. Дельта-модуляция

Очевидно, что точность схемы можно улучшить за счет увеличения частоты выборки. Впрочем, в результате увеличивается и скорость передачи данных в выходном сигнале.

Основным преимуществом дельта-модуляции по сравнению с импульсно-кодовой модуляцией является простота ее реализации. В то же время при равных скоростях передачи данных отношение сигнал/шум выше при импульсно-кодовой модуляции.

## Производительность

Хорошее воспроизведение речи можно получить при использовании импульсно-кодовой модуляции с 128 уровнями квантования, или семибитовой кодировки ( $2^7 = 128$ ). Речевой сигнал занимает полосу, не превышающую 4 кГц. Следова-

тельно, согласно теореме о дискретном представлении, выборки можно делать с частотой 8 000 выборок в секунду. Значит, скорость передачи цифровых данных, кодированных с использованием импульсно-кодовой модуляции, составляет  $8\,000 \times 7 = 56$  Кбит/с.

Рассмотрим в связи с этим требования к полосе. Аналоговый речевой сигнал занимает полосу шириной 4 кГц. Использование импульсно-кодовой модуляции позволяет преобразовать этот аналоговый сигнал в цифровой со скоростью передачи 56 Кбит/с. Но согласно критерию Найквиста (см. главу 2) данный цифровой сигнал требует полосы порядка 28 кГц. При рассмотрении же сигналов с широкими полосами различия оказываются еще больше. Например, в схеме импульсно-кодовой модуляции, широко применяющейся в цветном телевидении, использована десятибитовая кодировка, в которой для передачи сигнала с полосой 6,6 МГц используется скорость до 92 Мбит/с. Несмотря на это, использование цифровых технологий для передачи аналоговых данных становится все более и более популярным. Можно выделить следующие основные причины такой популярности.

- Поскольку ретрансляторы заменены усилителями, то дополнительные помехи не возникают.
- Как будет показано далее, при передаче цифровых сигналов используется уплотнение с разделением времени, в то время как при передаче аналоговых сигналов — уплотнение с разделением частоты. При временном разделении комбинационных помех не возникает, тогда как мы видели, что при частотном разделении эти шумы представляют проблему.
- Обращение к цифровой передаче сигналов позволяет использовать более эффективные технологии цифровой коммутации.

К тому же были разработаны более эффективные кодировки. При передаче речи разумной целью является достижение скорости порядка 4 Кбит/с. Поскольку при передаче видео большинство элементов изображения не изменяются от кадра к кадру, методы сжатия видеоизображения путем удаления избыточных данных позволяют снизить требования к передаче до 15 Мбит/с, а для медленно изменяющихся кадров, таких, например, которые встречаются в видеоконференциях, — до 64 Кбит/с или даже меньше.

Отметим, в заключение, что во многих случаях использование телекоммуникационной системы приводит и к аналоговой обработке цифровых данных, и к цифровой обработке аналоговых данных. Подавляющее большинство локальных конечных устройств в телекоммуникационной сети являются аналоговыми, а сеть сама по себе использует смесь аналоговых и цифровых технологий. Следовательно, цифровые данные на пользовательском терминале могут преобразовываться в аналоговые посредством модема, после этого оцифровываться с помощью кодека, и, возможно, после доставки данных по назначению описанный процесс повторится в обратном порядке.

Итак, телекоммуникационные средства обрабатывают аналоговые сигналы представляемые в виде речи или цифровых данных. Характеристики сигналов в этих случаях достаточно различны. Если речевые сигналы находятся ближе к нижней части полосы (см. рис. 3.9), то цифровые сигналы в аналоговой кодировке более равномерно распределены по полосе и, следовательно, содержат больше высокочастотных составляющих. Исследования показали, что вследствие наличия

этих высоких частот для оцифровывания аналоговых сигналов, представляющих цифровые данные, технологии, основанные на импульсно-кодовой модуляции, предпочтительнее технологий, основанных на использовании дельта-модуляции.

## 6.5. РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

Существует много хороших справочников по аналоговым схемам модуляции цифровых данных. Рекомендуемым выбором являются работы [COUC01], [XION00] и [GLOV98]; кроме того, в этих работах удачно описаны цифровые и аналоговые схемы модуляции для аналоговых данных.

Исключительно ясное описание методов преобразования цифровых и аналоговых данных в аналоговые сигналы и аналоговых данных в цифровые сигналы дано в [PEAR92]. Другая всесторонняя трактовка тем данной главы представлена в работе [SKLA01].

Представляющая интерес интерпретация понятий побитовой скорости передачи, бодов и ширины полосы дается в книге [FREE98]. Рекомендуемым вводным курсом, в котором раскрываются понятия, связанные с эффективностью использования полосы и схемами кодирования, является работа [SKLA93].

- COUC01 Couch L. *Digital and Analog Communication Systems*. — Upper Saddle River, NJ: Prentice Hall, 2001.
- FREE98 Freeman R. Bits, Symbols, Baud, and Bandwidth. — *IEEE Communications Magazine*, April 1998.
- GLOV98 Glover I., Grant P. *Digital Communications*. — Upper Saddle River, NJ: Prentice Hall, 1998.
- PEAR92 Pearson J. *Basic Communication Theory*. — Englewood Cliffs, NJ: Prentice Hall, 1992.
- SKLA93 Sklar B. Defining, Designing, and Evaluating Digital Communication Systems. — *IEEE Communications Magazine*, November 1993.
- SKLA01 Sklar B. *Digital Communications: Fundamentals and Applications*. — Upper Saddle River, NJ; Prentice Hall, 2001. // Склар, Б. *Цифровая связь: теоретические основы и практическое применение* — М.: Издательский дом "Вильямс", 2003.

## 6. ТЕРМИНЫ, ВОПРОСЫ И ЗАДАЧИ

### Основные термины

амплитудная манипуляция (ASK)	квадратурная амплитудная модуляция (QAM)	фазовая модуляция (PM)
амплитудная модуляция (AM)	квадратурная фазовая манипуляция (QPSK)	частота битовых ошибок (BER)
дельта-модуляция (DM)	модуляция несущая частота	частотная манипуляция (FSK)
дифференциальная фазовая манипуляция (DPSK)	угловая модуляция	частотная модуляция (FM)
импульсно-кодовая модуляция (PCM)	фазовая манипуляция (PSK)	

## Вопросы

1. Что такое “дифференциальное кодирование”?
2. Какую функцию выполняет модем?
3. Назовите три основных преимущества цифровой передачи перед аналоговой.
4. Как при амплитудной манипуляции представляются двоичные значения, и какое ограничение имеет данный подход?
5. Что такое “NRZ-L”? Назовите основной недостаток данной схемы кодирования.

## Задачи

1. На рис. 6.20 показан демодулятор сигнала с квадратурной амплитудной модуляцией, соответствующий модулятору, показанному на рис. 6.10. Покажите, что такая схема восстанавливает два сигнала,  $d_1(t)$  и  $d_2(t)$ , которые затем могут объединяться с целью восстановления исходного входа.

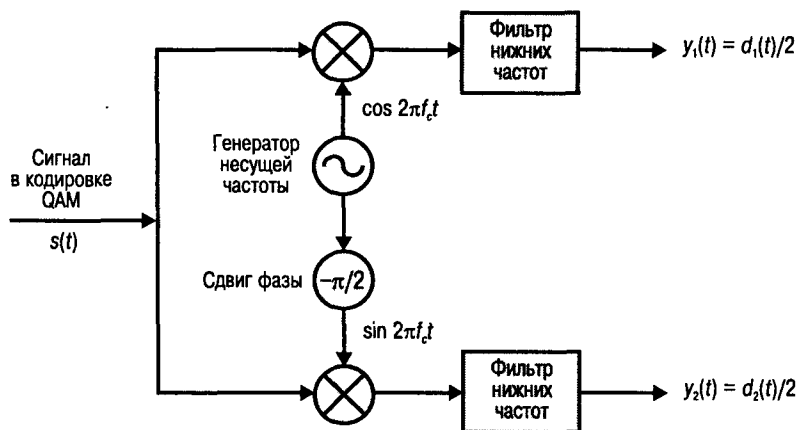


Рис. 6.20. Демодулятор сигнала с квадратурной амплитудной модуляцией

2. Синусоидальная волна должна использоваться в двух различных схемах передачи сигналов: а) PSK; б) QPSK. Длительность сигнальной посылки равна  $10^{-5}$  с. Определите отношение  $E_b/N_0$  (в дБ) для каждого случая, если уровень принятого сигнала

$$s(t) = 0,005 \sin(2\pi \cdot 10^6 t + \theta) \text{ В,}$$

а измеренная в приемнике мощность шума равна  $2,5 \times 10^{-8}$  Вт.

3. Определите отношение сигнал/шум, необходимое для получения эффективности использования полосы 1,0 при амплитудной, частотной, фазовой и квадратурной фазовой манипуляции. Частоту появления ошибок считайте равной  $10^{-6}$ .
4. Сигнал в кодировке NRZ-L пропускается через фильтр с  $r = 0,5$ , после чего модулируется на несущей частоте. Скорость передачи данных равна 2 400 бит/с. Оцените ширину полосы для амплитудной и частотной манипуляций. При рассмотрении частотной манипуляции считайте, что используются частоты 50 кГц и 55 кГц.
5. Предположим, что телефонная линия откорректирована и позволяет производить полосовую передачу данных в диапазоне частот от 600 до 3000 Гц. Ширина доступной полосы равна 2400 Гц. Оцените полосу, требуемую для передачи со скоростью

2400 бит/с при  $r=1$  QPSK-кодированного сигнала и для восьмиуровневой передачи — сигналов со скоростью 4 800 бит/с. Достаточно ли этой полосы?

6. Почему при кодировании аналоговых сигналов, представляющих цифровые данные, кодировка PCM предпочтительнее дельта-модуляции?
7. Являются ли функции модема и кодека взаимно-обратными (т.е. может ли инвертированный модем использоваться в качестве кодека и наоборот)?
8. Сигнал квантуется с использованием 10-битовой импульсно-кодовой модуляции. Найдите отношение сигнал/квантовый шум.
9. Рассмотрим аудиосигнал, спектр которого принадлежит диапазону от 300 до 3 000 Гц. Предположим, что для генерации сигнала в кодировке PCM делается 7 000 выборок в секунду.
  - а. Для отношения сигнал/шум 30 дБ определите необходимое число равномерно размещенных квантовых уровней.
  - б. Определите требуемую скорость передачи данных.
10. Определите шаг  $\delta$ , необходимый для предотвращения шума перегрузки кривизны, как функцию частоты гармоник сигнала, имеющей наивысшую частоту. Амплитуды всех гармоник сигнала считайте равными  $A$ .
11. В устройство импульсно-кодового модулирования подается сигнал с полным напряжением 10 В. Для генерации из сигнала 8-битового закодированного сигнала используется квантование с равномерным размещением уровней. Максимальное приведенное квантованное напряжение равно  $1 - 2^{-8}$ . Определите: а) приведенный шаг, б) действительный шаг в вольтах, в) действительный максимальный квантовый уровень в вольтах, г) приведенное разрешение, д) действительное разрешение, е) разрешение в процентах.

12. Из аналогового сигнала, показанного на рис. 6.21, требуется получить дельта-модулированный сигнал. Период взятия выборки и размер шага определяются с помощью сетки рисунка. Для этого периода также показаны первый сигнал, вышедший из модулятора, и начало ступенчатой функции. Продолжите ступенчатую функцию и полностью получите сигнал, вышедший из модулятора. Укажите области с искажением, вызванным шумом перегрузки кривизны.

13. Рассмотрим сигнал с угловой модуляцией

$$s(t) = 10 \cos[(10^8)\pi t + 5 \sin 2\pi(10^3)t].$$

Определите максимальное отклонение фазы и максимальное отклонение частоты.

14. Рассмотрим сигнал с угловой модуляцией

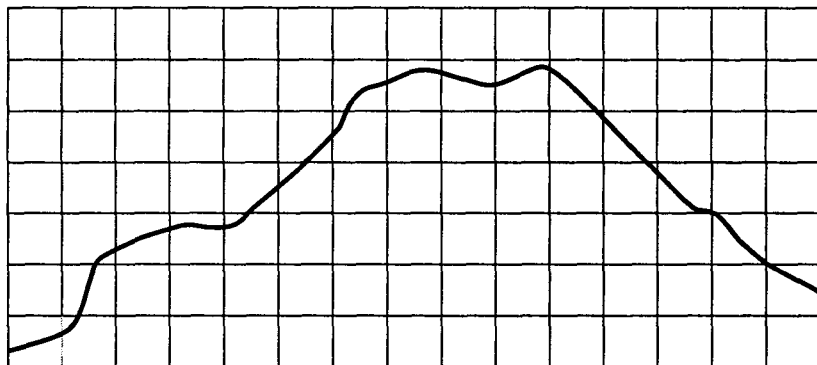
$$s(t) = 10 \cos[2\pi(10^6)t + 0,1 \sin (10^3)\pi t].$$

- а. Выразите  $s(t)$  как фазово-модулированный сигнал с  $n_p = 10$ .
- б. Выразите  $s(t)$  как частотно-модулированный сигнал с  $n_f = 10\pi$ .

15. Пусть  $m_1(t)$  и  $m_2(t)$  — информационные сигналы, а  $s_1(t)$  и  $s_2(t)$  — соответствующие модулированные сигналы с несущей  $f_c$ .

- а. Покажите, что при простой амплитудной модуляции из сигнала  $m_1(t) + m_2(t)$  получается модулированный сигнал, равный линейной комбинации сигналов  $s_1(t)$  и  $s_2(t)$ . По этой причине амплитудная модуляция иногда называется линейной модуляцией.
- б. Покажите, что при простой фазовой модуляции из сигнала  $m_1(t) + m_2(t)$  получается модулированный сигнал, не равный линейной комбинации сигнала

лов  $s_1(t)$  и  $s_2(t)$ . По этой причине фазовая модуляция иногда называется *нелинейной модуляцией*.



Дельта-модулированный сигнал



Рис. 6.21. Пример дельта-модуляции

## ПРИЛОЖЕНИЕ 6А. ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ О ДИСКРЕТНОМ ПРЕДСТАВЛЕНИИ

Теорему о дискретном представлении можно сформулировать следующим образом.

Если

- $x(t)$  — сигнал с ограниченной полосой  $f_h$ ,
- $p(t)$  — выборочный сигнал, состоящий из импульсов, взятых через интервал  $T_s = 1/f_s$ , где  $f_s$  — частота дискретизации,
- $x_s(t) = x(t)p(t)$  — дискретный сигнал,

то сигнал  $x(t)$  можно восстановить непосредственно из сигнала  $x_s(t)$  тогда и только тогда, когда  $f_s \geq 2f_h$ .

### Доказательство

Поскольку сигнал  $p(t)$  состоит из равномерной последовательности импульсов, то он является периодическим сигналом и может быть представлен в виде ряда Фурье:

$$p(t) = \sum_{n=-\infty}^{\infty} P_n e^{2\pi i n f_s t}.$$



Получаем

$$x_s(t) = x(t)p(t) = \sum_{n=-\infty}^{\infty} P_n x(t) e^{2\pi n f_s t}.$$

Рассмотрим теперь Фурье-образ функции  $x_s(t)$ :

$$X_s(f) = \int_{-\infty}^{\infty} x_s(t) e^{-2\pi i f t} dt.$$

Подставляя выражение для  $x_s(t)$ , имеем

$$X_s(f) = \int_{-\infty}^{\infty} \sum_{n=-\infty}^{\infty} P_n x(t) e^{2\pi n f_s t} e^{-2\pi i f t} dt.$$

После упрощения получаем

$$X_s(f) = \sum_{n=-\infty}^{\infty} P_n \int_{-\infty}^{\infty} x(t) e^{-2\pi i (f - n f_s) t} dt.$$

Исходя из определения преобразования Фурье, можем записать следующее:

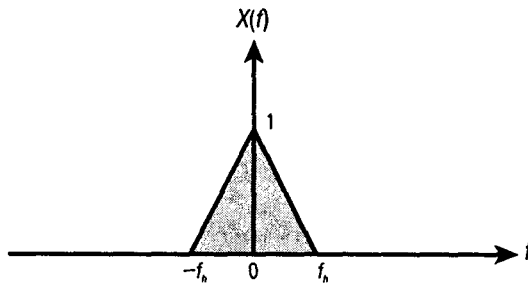
$$X(f - n f_s) = \int_{-\infty}^{\infty} x(t) e^{-2\pi i (f - n f_s) t} dt,$$

где  $X(f)$  — Фурье-образ функции  $x(t)$ . Подставив это в предыдущую формулу, получим

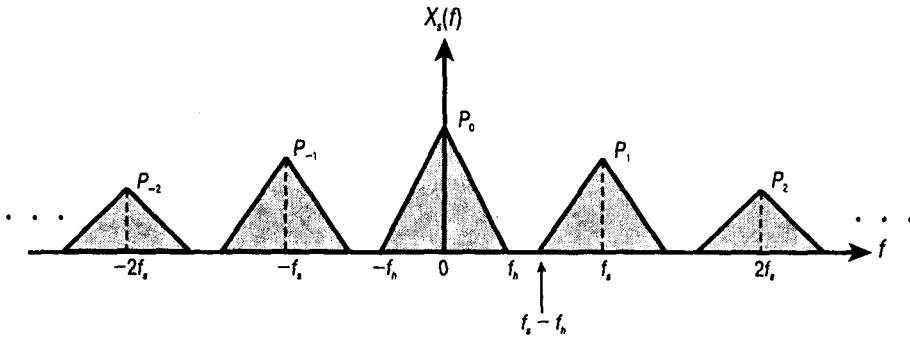
$$X_s(f) = \sum_{n=-\infty}^{\infty} P_n X(f - n f_s).$$

Последняя формула имеет интересную интерпретацию, представленную на рис. 6.22, где мы предполагали, не нарушая общности, что полоса сигнала  $x(t)$  лежит в диапазоне от 0 до  $f_h$ . Спектр сигнала  $x_s(t)$  состоит из спектра сигнала  $x(t)$  и спектра, который получается из спектра сигнала  $x(t)$  при переносе каждой гармоники на несущую частоту. Каждый из перенесенных спектров умножается на соответствующий коэффициент ряда Фурье для  $p(t)$ . Если  $f_s \geq 2f_h$ , то эти перенесенные спектры не перекрываются, а спектр сигнала  $x(t)$  умножается на коэффициент  $P_0$ , входящий в  $X_s(f)$ . Восстановление спектра исходного сигнала  $x(t)$  произойдет при передаче сигнала  $X_s(t)$  через полосовой фильтр  $f \geq f_s$ . В форме уравнения это выглядит следующим образом:

$$X_s(f) = P_0 X(f), \quad \frac{-f_s}{2} \leq f \leq \frac{f_s}{2}.$$



а) Спектр сигнала  $x(t)$



б) Спектр сигнала  $x_s(t)$

Рис. 6.22. Иллюстрация к теореме о дискретном представлении

# ГЛАВА 7

## РАСШИРЕННЫЙ СПЕКТР

- 7.1. Понятие расширенного спектра
- 7.2. Расширение спектра со скачкообразной перестройкой частоты
  - Основы использования
  - FHSS с использованием MFSK
  - Анализ производительности
- 7.3. Расширение спектра методом прямой последовательности
  - DSSS с использованием BPSK
  - Анализ производительности
- 7.4. Множественный доступ с кодовым разделением
  - Основные принципы
  - CDMA для системы DSSS
- 7.5. Создание последовательностей расширения
  - Псевдослучайные последовательности
  - Ортогональные коды
  - Множественное расширение
- 7.6. Рекомендуемая литература
- 7.7. Термины, вопросы и задачи

**В** технологиях радиосвязи крайне важную роль играет расширенный спектр. Данный метод не подпадает ни под одну из категорий, определенных в предыдущей главе, поскольку может быть использован для передачи как цифровых, так и аналоговых данных с помощью аналогового сигнала.

Изначально метод расширенного спектра создавался для разведывательных и военных целей. Основная идея метода состоит в том, чтобы распределить информационный сигнал по широкой полосе радиодиапазона, что в итоге позволит значительно усложнить подавление или перехват сигнала. Первая разработанная схема расширенного спектра известна как метод перестройки частоты<sup>1</sup>. Более современной схемой расширенного спектра является метод прямой последовательности. Оба метода используются в различных стандартах и продуктах беспроводной связи.

Ниже, после краткого обзора, названные методы расширенного спектра рассматриваются подробно. Кроме того, в данной главе будет исследован метод множественного доступа на основе расширения спектра.

## 7.1. ПОНЯТИЕ РАСШИРЕННОГО СПЕКТРА

На рис. 7.1 приведены ключевые элементы системы расширенного спектра. Входной сигнал поступает на каналный кодер, который генерирует аналоговый сигнал со сравнительно узкой полосой, центрированной на определенной частоте. Далее сигнал модулируется с помощью последовательности чисел, именуемой кодом расширения, или расширяющей последовательностью. Обычно, хотя и не всегда, код расширения создается генератором случайных чисел. В результате модуляции полоса передаваемого сигнала значительно расширяется (другими словами, расширяется спектр сигнала). После приема сигнал демодулируется с использованием того же кода расширения. Последний шаг — сигнал подается на каналный декодер для восстановления данных.

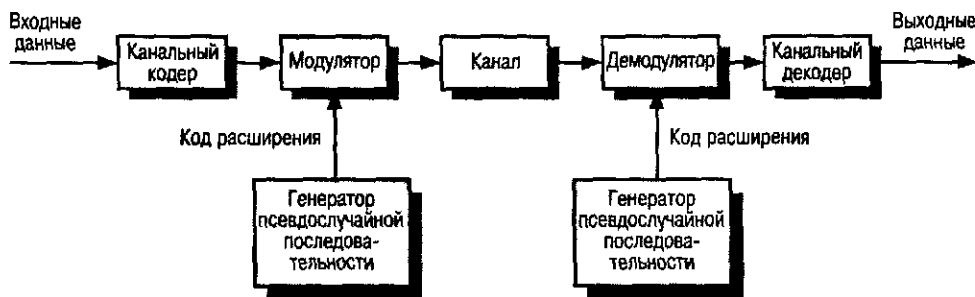


Рис. 7.1. Общая схема цифровой системы связи с использованием расширенного спектра

Избыток спектра дает возможность получить следующие преимущества.

<sup>1</sup> Как бы невероятно это не звучало, но расширение спектра методом перестройки частоты придумала голливудская кинозвезда Хеди Ламарр (Hedy Lamarr) в 1940 году в возрасте 26 лет. В 1942 году Ламарр запатентовала свое изобретение (патент США 2 292 387 от 11 августа 1942 г.) совместно с партнером, который начал принимать участие в работе несколько позже. Девушка не получила никакой прибыли от патента, считая открытый ею метод связи своим вкладом в участие США во второй мировой войне.

- Невосприимчивость сигнала к различным типам шумов, а также к искажениям, вызванным многолучевым распространением. Впервые расширенный спектр был использован в военных целях благодаря устойчивости расширенного сигнала к попыткам подавления.
- Расширенный спектр позволяет скрывать и шифровать сигналы. Восстановить зашифрованные данные сможет только пользователь, которому известен код расширения.
- Несколько пользователей могут одновременно использовать одну полосу частот при крайне малой взаимной интерференции. Данное свойство используется в технологии мобильной связи, известной как уплотнение с кодовым разделением (code division multiplexing — CDM), или множественный доступ с кодовым разделением (code division multiple access — CDMA).

## 7.2. РАСШИРЕНИЕ СПЕКТРА СО СКАЧКООБРАЗНОЙ ПЕРЕСТРОЙКОЙ ЧАСТОТЫ

При использовании расширенного спектра со скачкообразной перестройкой частоты (frequency-hopping spread spectrum — FHSS) передача сигнала производится с помощью определенных наборов частот, имеющих свойства случайных последовательностей. Перестройка частоты сигнала происходит через определенные интервалы времени. Изменение рабочих частот приемника синхронизируется с передатчиком, что позволяет получить сигнал. В то же время при попытке подслушать сеанс связи будут слышны лишь неразборчивые звуки, а создание намеренных помех на одной из частот приведет к уничтожению только нескольких битов сигнала.

### Основы использования

На рис. 7.2 приведен пример передачи сигнала со скачкообразной перестройкой частоты (frequency hopping — FH). Для передачи FH-сигнала резервируется определенное количество каналов. Как правило, используют  $2^k$  несущих частот, которые составляют  $2^k$  каналов. Расстояние между несущими частотами (а следовательно, ширина каждого канала) обычно равно ширине полосы входного сигнала. При передаче каждый канал используется в течение фиксированного интервала времени; в стандарте IEEE 802.11, например, этот интервал равен 300 мс. В течение такого интервала производится передача некоторого количества закодированных определенным образом битов (возможно, частей битов, см. далее). Последовательность использования каналов задается кодом расширения, и поскольку приемник и передатчик используют один и тот же код, переходы между каналами выполняются синхронно.

На рис. 7.3 приводится типичная блок-схема системы связи со скачкообразной перестройкой частоты. При передаче двоичные данные подаются на модулятор, работающий с использованием определенного метода цифро-аналогового кодирования, например частотной манипуляции (frequency-shift keying — FSK) или же двоичной фазовой манипуляции (binary phase shift keying — BPSK).

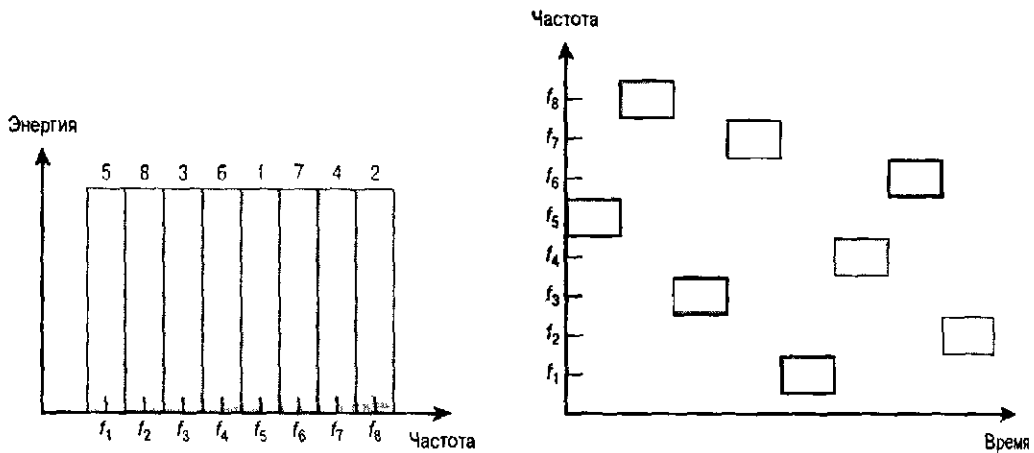


Рис. 7.2. Пример скачкообразной перестройки частоты

Полученный сигнал центрирован на некоторой базовой частоте. Изображенный далее генератор псевдослучайных чисел применяется для получения индексов таблицы используемых частот. Именно псевдослучайная последовательность является упоминавшимся ранее кодом расширения. Каждые  $k$  бит на выходе генератора определяют одну из  $2^k$  несущих частот. Для каждого из последующих интервалов времени (которому соответствует  $k$  бит псевдослучайного кода) выбирается новая несущая частота. Эта частота модулируется сигналом исходного модулятора. Форма полученного сигнала не изменится, однако он будет центрирован на выбранной частоте. Приемник демодулирует полученный сигнал расширенного спектра с помощью той же последовательности частот (основанной на псевдослучайном коде), что использовалась при модуляции. Для получения выходных данных полученный сигнал еще раз демодулируется.

В схеме на рис. 7.3 показано перемножение двух сигналов. Приведем пример подобного умножения, используя BFSK в качестве схемы модуляции данных. Сигнал FSK на входе системы FHSS можно определить следующим образом (сравните с уравнением (6.2)):

$$s_d(t) = A \cos(2\pi(f_0 + 0,5(b_i + 1)\Delta f)t) \quad \text{при } iT < t < (i+1)T. \quad (7.1)$$

Здесь

$A$  — амплитуда сигнала;

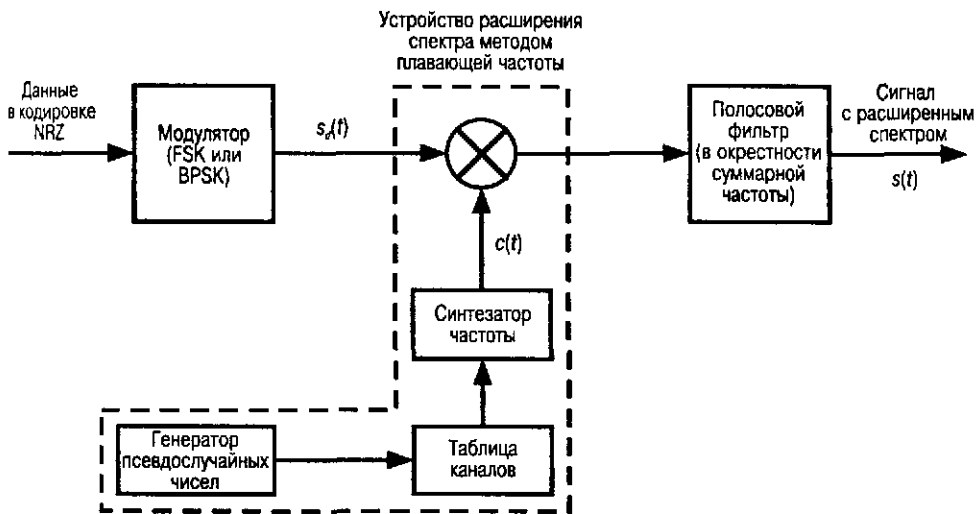
$f_0$  — базовая частота;

$b_i$  — значение  $i$ -го бита данных (+1 соответствует двоичной 1; -1 — двоичному 0);

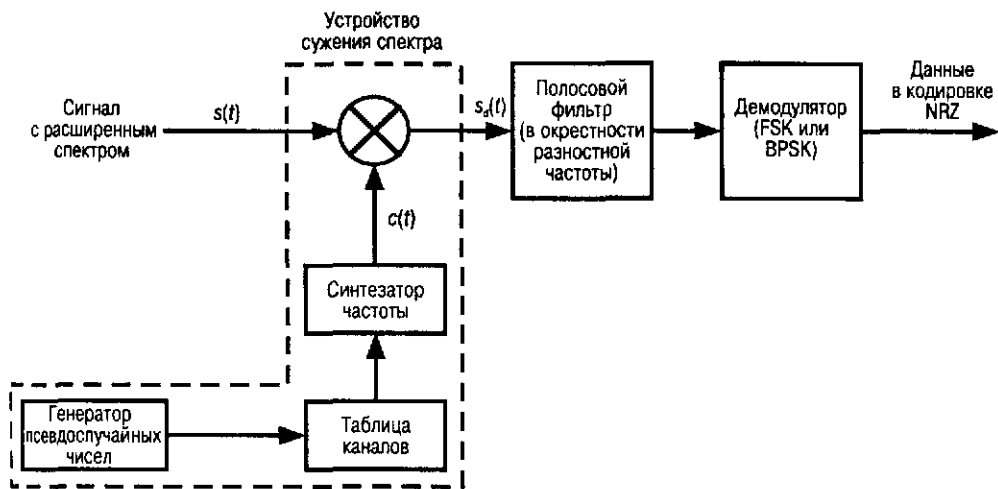
$\Delta f$  — разнесение частот;

$T$  — длительность передачи одного бита; скорость передачи данных равна  $1/T$ .

Таким образом, в течение интервала передачи  $i$ -го бита частота сигнала будет равна  $f_0$ , если бит имеет значение -1, и  $f_0 + \Delta f$  при значении бита +1.



а) Передатчик



б) Приемник

Рис. 73. Система связи расширенного спектра со скачкообразной перестройкой частоты

Синтезатор частот генерирует последовательность тонов, частота которых изменяется через равные промежутки времени. Переходы в наборе  $2^k$  частот определяются  $k$  битами псевдослучайного кода. Будем считать время передачи на одной частоте равным времени передачи одного бита, будем также пренебрегать фазовой разницей между информационным сигналом  $s_n(t)$  и сигналом расширения  $c(t)$ . Результирующий сигнал во время  $i$ -го интервала передачи (передачи  $i$ -го бита) можно выразить в следующем виде:

$$p(t) = s_d(t)c(t) = A \cos(2\pi(f_0 + 0,5(b_i + 1)\Delta f)t) \cos(2\pi f_i t).$$

Здесь  $f_i$  — частота сигнала, сгенерированного синтезатором частот во время  $i$ -го интервала передачи. Используя тригонометрическое равенство<sup>2</sup>  $\cos(x)\cos(y) = (1/2)(\cos(x + y) + \cos(x - y))$ , получим следующее:

$$p(t) = 0,5 A [\cos(2\pi(f_0 + 0,5(b_i + 1)\Delta f + f_i)t) + \cos(2\pi(f_0 + 0,5(b_i + 1)\Delta f - f_i)t)].$$

Полосовой фильтр (см. рис. 7.3) позволяет отсеять разность частот, пропуская только их сумму. Полученный сигнал FHSS может быть записан в таком виде:

$$s(t) = 0,5 A \cos(2\pi(f_0 + 0,5(b_i + 1)\Delta f + f_i)t). \quad (7.2)$$

Итак, частота информационного сигнала в течение  $i$ -го интервала передачи равна  $f_0 + f_i$ , если значение бита  $-1$ , и  $f_0 + f_i + \Delta f$  при значении  $+1$ .

Приемник получит сигнал, описываемый только что определенной функцией  $s(t)$ . Умножение полученного сигнала на копию сигнала расширения дает в результате следующее:

$$p(t) = s(t)c(t) = 0,5A \cos(2\pi(f_0 + 0,5(b_i + 1)\Delta f + f_i)t) \cos(2\pi f_i t).$$

Снова используя тригонометрическое равенство, получим такой результат:

$$p(t) = s(t)c(t) = 0,25 A [\cos(2\pi(f_0 + 0,5(b_i + 1)\Delta f + f_i + f_i)t) + \cos(2\pi(f_0 + 0,5(b_i + 1)\Delta f)t)].$$

Для отсеивания разности частот используется полосовой фильтр (рис. 7.3), в результате получается сигнал, который описывается функцией  $s_d(t)$ , определенной в формуле (7.1).

$$0,25A \cos(2\pi(f_0 + 0,5(b_i + 1)\Delta f)t).$$

## FHSS с использованием MFSK

С технологией FHSS часто используется многочастотная манипуляция (multiple frequency shift keying — MFSK). Как упоминалось в главе 6, схема MFSK предусматривает использование  $M = 2^L$  различных частот для кодирования входного цифрового сигнала по  $L$  бит за такт. Передаваемый сигнал описывается следующей функцией (уравнение (6.3)):

$$s_i(t) = a \cos 2\pi f_i t, \quad 1 \leq i \leq M.$$

Здесь

$$f_i = f_c + (2i - 1 - M)f_d;$$

$f_c$  — частота несущей;

$f_d$  — разностная частота;

$M$  — количество различных сигнальных посылок,  $M = 2^L$ ;

$L$  — число битов на сигнальную посылку.

В схеме FHSS частота сигнала MFSK меняется с периодичностью  $T_c$  секунд. Перестройка частоты производится путем модулирования сигнала MFSK несущей FHSS. В результате сигнал MFSK передается по соответствующему каналу FHSS. При скорости передачи данных  $R$  время передачи одного бита составляет

<sup>2</sup> Математическую памятку, содержащую краткий обзор тригонометрических равенств, можно найти по адресу: [WilliamStallings.com/StudentSupport.html](http://WilliamStallings.com/StudentSupport.html).



$T = 1/R$  секунд. Время, необходимое для передачи сигнальной посылки, равно  $T_s = LT$ . Если  $T_c$  больше или равно  $T_s$ , модуляцию с расширением спектра принято называть расширением спектра с медленной скачкообразной перестройкой частоты, в противном случае говорят о быстрой перестройке частоты<sup>3</sup>. Подытожим: расширение спектра с медленной перестройкой частоты —  $T_c \geq T_s$ ; расширение спектра с быстрой перестройкой частоты —  $T_c < T_s$ .

На рис. 7.4 приведен пример медленной схемы FHSS с использованием сигнала MFSK, изображенного на рис. 6.4. Здесь  $M = 4$ , т.е. для кодирования каждых 2 бит входных данных за такт используется четыре разных частоты. Каждая сигнальная посылка является дискретным частотным тоном. Полная ширина полосы MFSK составляет  $W_d = Mf_d$ . В данном примере использовалась схема FHSS с  $k = 2$ . Следовательно, всего использовано  $4 = 2^k$  различных каналов, ширина полосы каждого из которых равна  $W_d$ . Полная ширина полосы FHSS равна  $W_s = 2^k W_d$ . Для выбора одного из четырех каналов используется по 2 бит псевдослучайной последовательности. Выбранный канал используется на протяжении времени передачи двух сигнальных посылок, или 4 бит ( $T_c = 2T_s = 4T$ ).

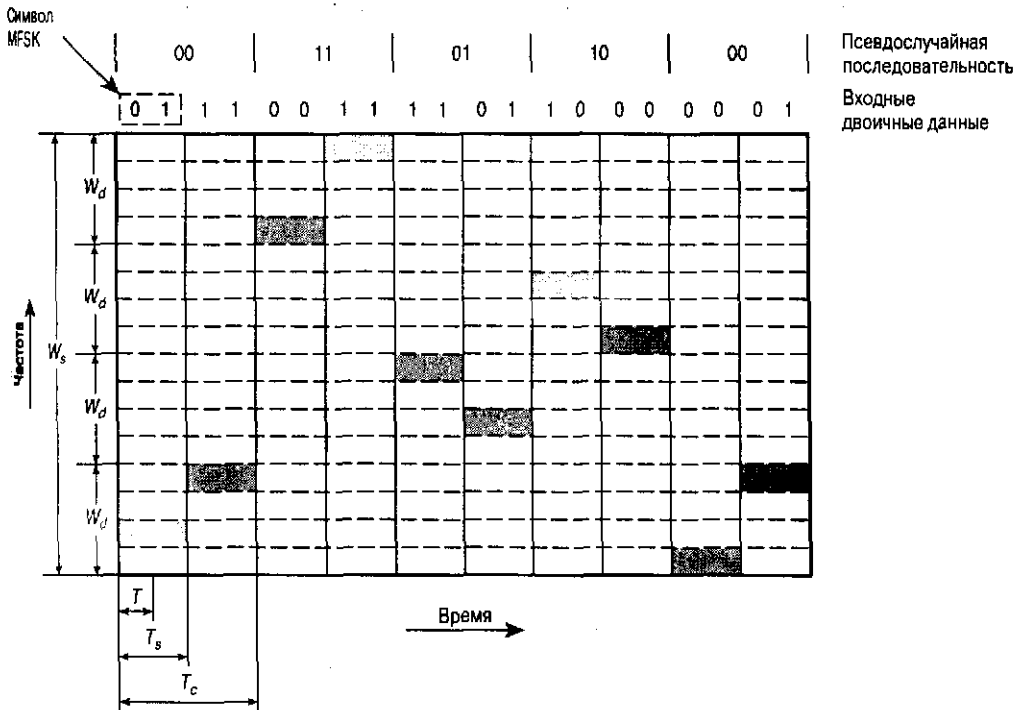


Рис. 7.4. Расширение спектра с медленной перестройкой частоты (модуляция MFSK,  $M = 4$ ,  $k = 2$ )

<sup>3</sup> Некоторые авторы используют иную классификацию (например, [PICK82]): несколько скачков при передаче одного бита называется быстрой перестройкой частоты; передача нескольких битов на одной частоте — медленной перестройкой; передача одного бита на одной частоте не относится ни к одной из категорий. В данной книге использовано более общее определение, в котором фигурируют сигнальные посылки, а не биты данных.

На рис. 7.5 приведен пример быстрой схемы FHSS с применением того же сигнала MFSK ( $M = 4, k = 2$ ). В данном примере каждая сигнальная посылка соответствует двум частотным тонам. Снова  $W_d = Mf_d$  и  $W_s = 2^k W_d$ .  $T_s$  и  $T_c$  связаны как  $T_s = 2T_c = 2T$ . Вообще при наличии шума или намеренных помех быстрая схема дает лучшее качество связи, чем медленная. Например, если для передачи каждой сигнальной посылки используется три или более частот (элементарных сигналов), приемник принимает решение о том, какая сигнальная посылка была послана, основываясь на большинстве корректно принятых элементарных сигналов.

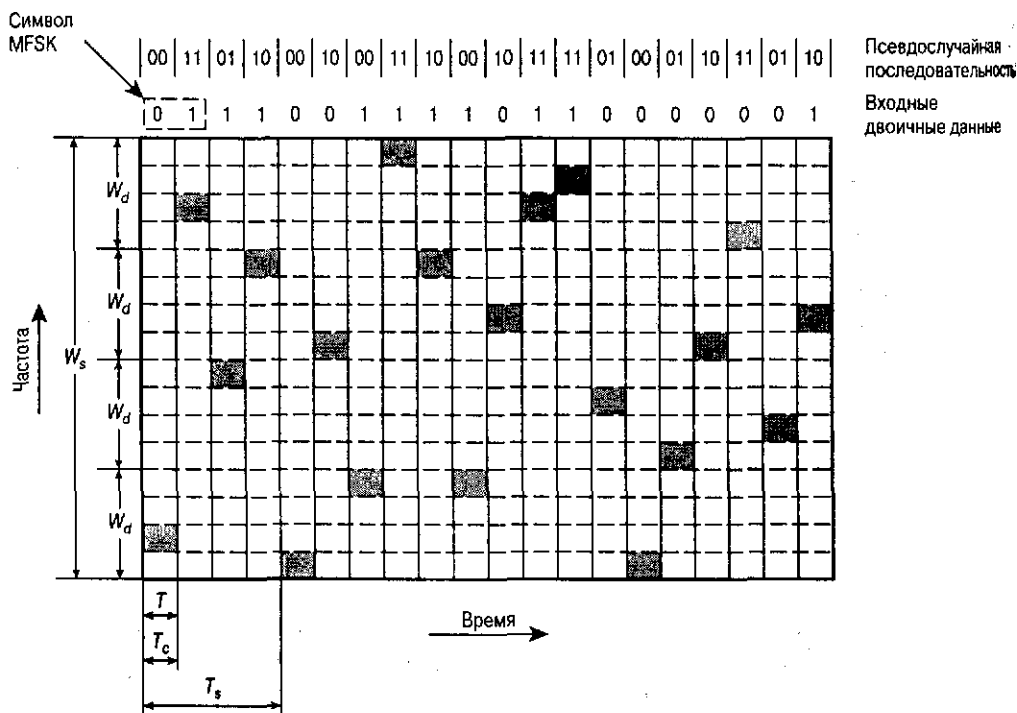


Рис. 7.5. Расширение спектра с быстрой перестройкой частоты (модуляция MFSK,  $M = 4, k = 2$ )

## Анализ производительности

Как правило, FHSS предусматривает использование большого числа частот, следовательно,  $W_s$  намного больше  $W_d$ . Одним из преимуществ такого подхода является большая устойчивость системы с большим значением  $k$  к воздействию намеренных помех. Предположим, например, что имеется передатчик MFSK с шириной полосы  $W_d$  и источник намеренных помех с полосой такой же ширины и фиксированной мощностью  $S_j$  на несущей частоте. Тогда отношение энергии сигнала на бит к плотности энергии шума на герц записывается в следующем виде:

$$\frac{E_b}{N_j} = \frac{E_b W_d}{S_j}$$

При использовании скачкообразной перестройки частоты генератор намеренных помех вынужден будет создавать шум на всех  $2^k$  частотах. Поскольку мощность генератора помех фиксирована, мощность шума на каждой из частот будет равна  $S/2^k$ . Значит, отношение мощности сигнала к мощности шума возрастет на величину, именуемую коэффициентом расширения спектра.

$$G_p = 2^k = \frac{W_s}{W_d}. \quad (7.3)$$

### 7.3. РАСШИРЕНИЕ СПЕКТРА МЕТОДОМ ПРЯМОЙ ПОСЛЕДОВАТЕЛЬНОСТИ

При расширении спектра методом прямой последовательности (direct sequence spread spectrum — DSSS) после обработки исходного сигнала кодом расширения каждому исходному биту ставится в соответствие несколько битов передаваемого сигнала. Степень расширения спектра прямо пропорциональна количеству битов кода. Другими словами, 10-битовый код расширяет полосу частот сигнала в 10 раз больше, чем 1-битовый код.

Один из методов применения DSSS — комбинирование цифрового информационного потока и битовой последовательности кода расширения с использованием исключающего ИЛИ. Операция исключающего ИЛИ выполняется согласно следующим правилам:

$$0 \oplus 0 = 0 \quad 0 \oplus 1 = 1 \quad 1 \oplus 0 = 1 \quad 1 \oplus 1 = 0$$

Пример такого комбинирования приводится на рис. 7.6. Отметим, что бит данных, равный единице, инвертирует биты кода; если же бит данных равен нулю, биты кода расширения передаются без изменений. Комбинация двух последовательностей битов имеет такую же скорость передачи, как и последовательность кода расширения. Следовательно, полоса комбинированной последовательности больше полосы последовательности данных. В данном примере скорость передачи последовательности битов кода в четыре раза превышает аналогичный параметр для битов данных.

#### DSSS с использованием BPSK

Рассмотрим использование схемы DSSS на практике, предполагая применение модуляции BPSK. Для обозначения двоичных данных удобнее будет использовать не нуль и единицу, а “+1” и “-1” соответственно. Как было показано в уравнении (6.5), сигнал BPSK можно описать следующей формулой:

$$s_d(t) = A d(t) \cos(2\pi f_c t). \quad (7.4)$$

Здесь

$A$  — амплитуда сигнала;

$f_c$  — несущая частота;

$d(t)$  — дискретная функция, принимающая значение +1, если соответствующий бит потока данных равен 1, и -1, когда бит данных равен 0.

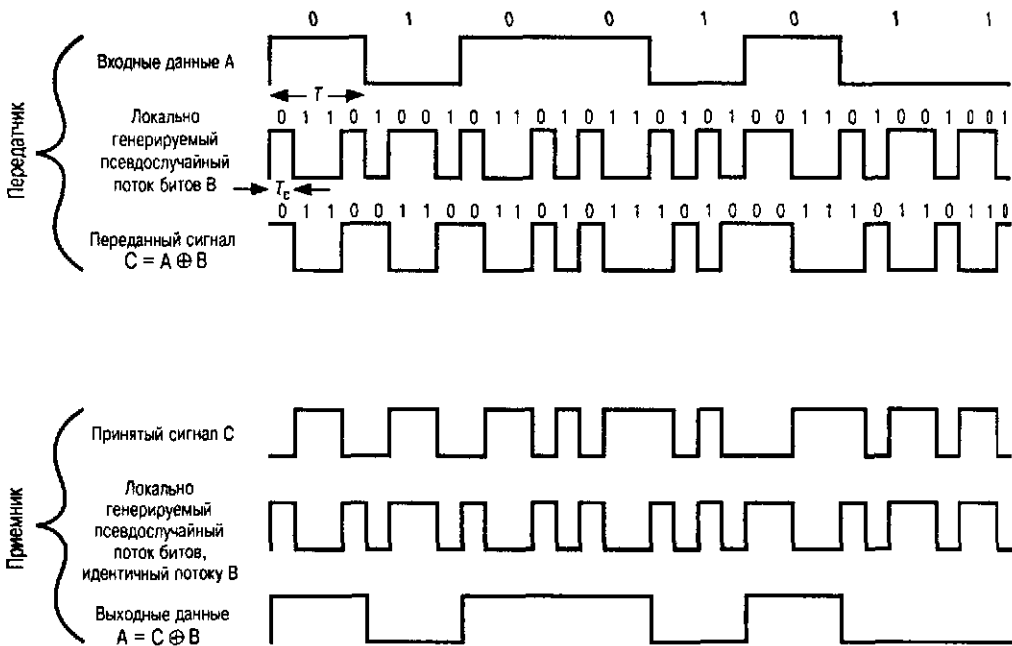


Рис. 7.6. Пример использования расширения спектра методом прямой последовательности

Чтобы получить сигнал DSSS, необходимо умножить  $s_A(t)$  на функцию  $c(t)$ , которая соответствует псевдослучайной последовательности и принимает значения  $-1$  и  $+1$ :

$$s(t) = A d(t)c(t) \cos(2\pi f_c t). \quad (7.5)$$

При поступлении сигнала на приемник он еще раз умножается на  $c(t)$ . Поскольку  $c(t) \times c(t) = 1$ , в результате умножения будет восстановлен исходный сигнал:

$$s(t)c(t) = A d(t)c(t)c(t) \cos(2\pi f_c t) = s_A(t).$$

Формулу (7.5) можно интерпретировать двояко, откуда следуют две реализации описанного метода. Первая интерпретация — умножение  $c(t)$  на  $d(t)$  с последующим применением модуляции BPSK (именно такой подход рассматривался выше). Можно также использовать альтернативный подход — модуляцию по схеме BPSK потока данных  $d(t)$  с последующим умножением полученной функции  $s_A(t)$  на  $c(t)$ .

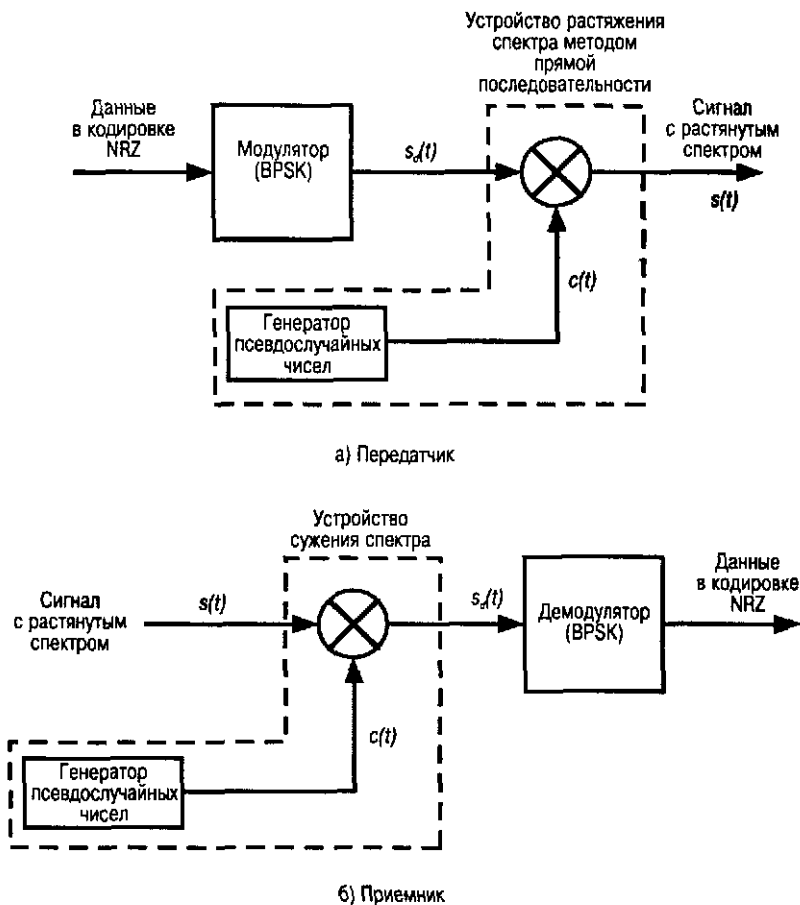


Рис 7.7. Система расширения спектра методом прямой последовательности

Реализация второй трактовки приведена на рис. 7.7. Пример использования этого подхода изображен на рис. 7.8.

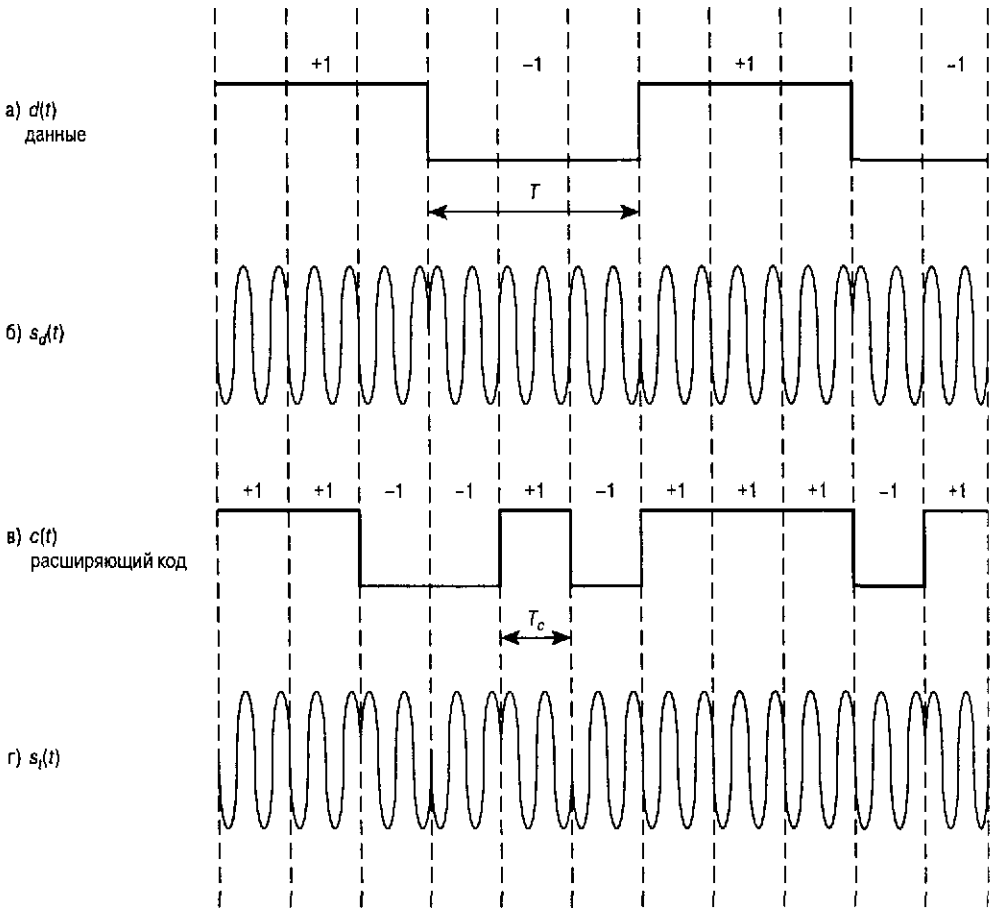


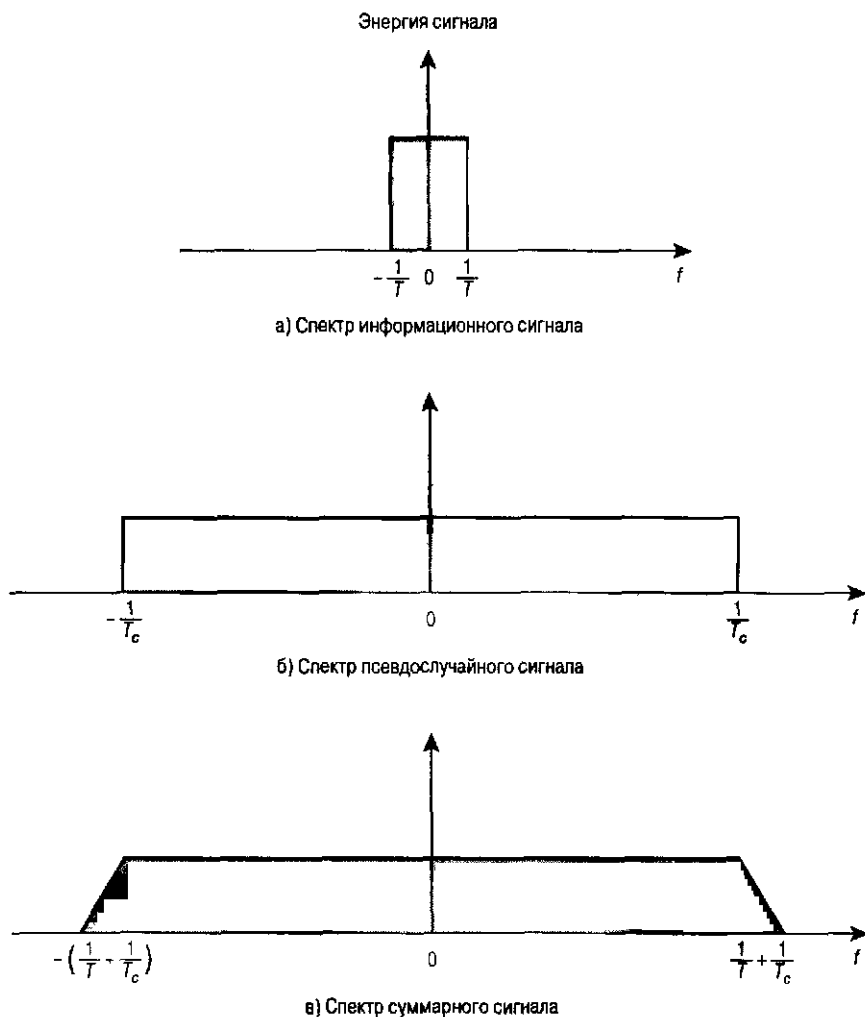
Рис. 7.8. Пример системы расширения спектра методом прямой последовательности (модуляция BPSK)

## Анализ производительности

Расширение спектра при использовании схемы DSSS определить довольно просто (рис. 7.9). В нашем примере ширина полосы одного бита информационного сигнала равна  $T$ , что соответствует скорости передачи данных  $1/T$ . Следовательно, в зависимости от кодировки ширина спектра сигнала будет составлять порядка  $2/T$ . Подобным образом, спектр псевдослучайного сигнала равен  $2/T_c$ . Получающийся расширенный спектр изображен на рис. 7.9, в. Степень расширения прямо зависит от скорости передачи псевдослучайной последовательности.

Как и для схемы FSSS, представление об эффективности DSSS можно получить, проанализировав устойчивость системы связи к подавлению. Предположим, что намеренная помеха ставится на центральной частоте системы DSSS. Сигнал помех имеет следующий вид:

$$s_j(t) = \sqrt{2S_j} \cos(2\pi f_j t).$$



*Рис. 7.9. Приблизительный спектр сигнала DSSS*

Полученный сигнал можно представить так:

$$s_r(t) = s(t) + s_f(t) + n(t).$$

Здесь

- $s(t)$  — переданный сигнал;
- $s_f(t)$  — сигнал намеренных помех;
- $n(t)$  — аддитивный белый шум;
- $S_f$  — мощность сигнала помех.

Устройство сужения спектра в приемнике умножает  $s_r(t)$  на  $c(t)$ . Компонент сигнала, соответствующий намеренным помехам, можно записать в следующем виде:

$$y_j(t) = \sqrt{2S_f} c(t) \cos(2\pi f_c t) . .$$

Таким образом, имеем простое применение модуляции BPSK к несущему тону. Следовательно, мощность несущей  $S_j$  распределена в полосе, ширина которой приблизительно равна  $2/T_c$ . В то же время демодулятор BPSK (рис. 7.7), следующий за устройством сужения спектра, включает полосовой фильтр с шириной полосы  $2/T$ , который согласован с данными BPSK. Значит, большая часть мощности помех отфильтровывается. Хотя строго следует учитывать влияние множества факторов, мощность намеренных помех, которые не были отсеяны полосовым фильтром, можно записать приблизительно:

$$S_j F = S_j (2/T) / (2/T_c) = S_j (T_c/T).$$

Таким образом, использование расширенного спектра снизило мощность намеренных помех в  $(T_c/T)$  раз. Величина, обратная данному коэффициенту, выражает выигрыш в отношении сигнал/шум:

$$G_p = \frac{T}{T_c} = \frac{R_c}{R} = \frac{W_s}{W_d}. \quad (7.6)$$

Здесь

$R_c$  — скорость передачи данных кода расширения;

$R$  — скорость передачи данных;

$W_d$  — ширина полосы сигнала;

$W_s$  — ширина полосы сигнала расширенного спектра.

Результат подобен полученному ранее для схемы FHSS (уравнение (7.3)).

## 7.4. МНОЖЕСТВЕННЫЙ ДОСТУП С КОДОВЫМ РАЗДЕЛЕНИЕМ

### Основные принципы

Множественный доступ с кодовым разделением (code division multiple access — CDMA) является одним из методов уплотнения с использованием расширенного спектра. Рассмотрение принципов работы CDMA начнем с информационного сигнала со скоростью передачи  $D$ . В соответствии со схемой, индивидуальной для каждого пользователя и именуемой пользовательским кодом, каждый бит данных разбивается на  $k$  элементарных сигналов. Канал, используемый для передачи данных, характеризуется скоростью передачи  $kD$  элементарных сигналов/с. Рассмотрим в качестве примера вариант  $k = 6^4$ . Для простоты код будем рассматривать как последовательность значений “1” и “-1”. На рис. 7.10 изображены коды трех пользователей (A, B, C), каждый из которых использует в процессе связи приемник базовой станции R. Код пользователя A имеет вид  $c_A = \langle 1, -1, -1, 1, -1, 1 \rangle$ . Для пользователей B и C коды равны, соответственно,  $c_B = \langle 1, -1, -1, 1, 1, 1 \rangle$  и  $c_C = \langle 1, 1, -1, 1, 1, -1 \rangle$ .

<sup>4</sup> Пример предоставлен профессором Политехнического университета Бруклина (Polytechnic University of Brookline) Ричардом Ван Сликом (Rychard Van Slyke).



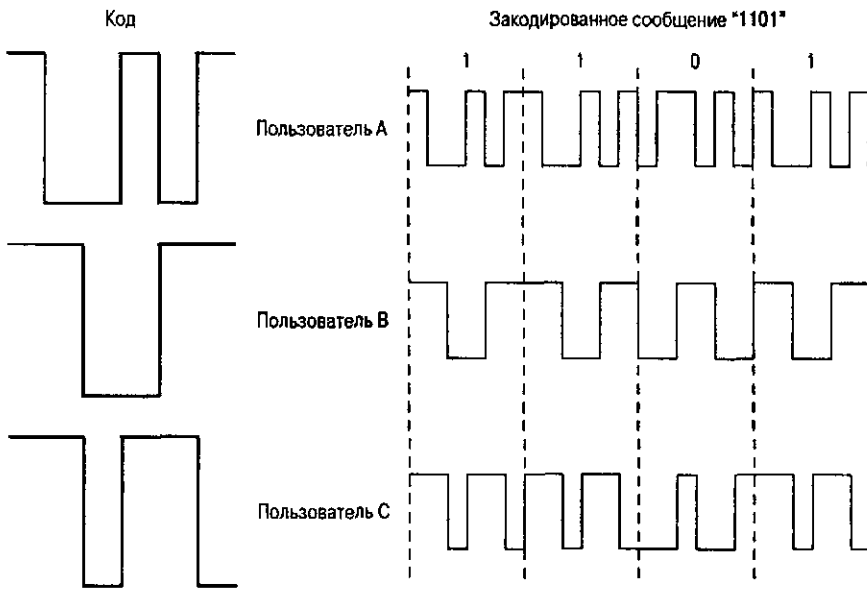


Рис. 7.10. Пример CDMA

Рассмотрим процесс связи пользователя А с базовой станцией. Предполагается, что базовой станции известен код А. Для простоты будем считать, что процесс связи является синхронизированным, т.е. что базовой станции известны моменты времени, в которые следует ожидать код. Для передачи одного бита информации А передает свой код в виде последовательности элементарных сигналов  $\langle 1, -1, -1, 1, -1, 1 \rangle$ . Если передаваемый бит является нулем, пользователь передает код-дополнение ("1" заменяется "-1" и наоборот)  $\langle -1, 1, 1, -1, 1, -1 \rangle$ . Приемник базовой станции декодирует последовательности элементарных сигналов. В нашем примере для декодирования полученной последовательности элементарных сигналов  $d = \langle d_1, d_2, d_3, d_4, d_5, d_6 \rangle$  приемник R использует следующую функцию (предполагается, что необходимо установить связь с пользователем  $u$ , код которого  $\langle c_1, c_2, c_3, c_4, c_5, c_6 \rangle$  известен приемнику):

$$S_u(d) = d_1 \times c_1 + d_2 \times c_2 + d_3 \times c_3 + d_4 \times c_4 + d_5 \times c_5 + d_6 \times c_6.$$

Индекс  $u$  в  $S_u(d)$  означает, что нас интересует пользователь  $u$ . Применим эту формулу для пользователя А (рис. 7.10). Если А пересылает бит со значением 1, тогда последовательность  $d$  равна  $\langle 1, -1, -1, 1, -1, 1 \rangle$ . Подставив соответствующие значения, получим выражение для  $S_A(d)$ :

$$S_A(1, -1, -1, 1, -1, 1) = 1 \times 1 + (-1) \times (-1) + (-1) \times (-1) + 1 \times 1 + (-1) \times (-1) + 1 \times 1 = 6$$

Если пользователь А пересылает бит со значением 0 ( $d = \langle -1, 1, 1, -1, 1, -1 \rangle$ ), выражение изменится следующим образом:

$$S_A(-1, 1, 1, -1, 1, -1) = (-1) \times 1 + 1 \times (-1) + 1 \times (-1) + (-1) \times 1 + 1 \times (-1) + (-1) \times 1 = -6.$$

Обратите внимание, что независимо от структуры передаваемой последовательности  $d$ , состоящей из "1" и "-1", всегда выполняется неравенство

$-6 \leq S_A(d) \leq 6$ . Максимальные значения 6 и -6 соответствуют прямому и дополнительному коду пользователя А. Таким образом, если значение  $S_A$  равно 6, это означает, что от пользователя А получен бит, равный 1. При  $S_A$ , равном -6, считается, что пользователь А передает бит со значением 0. Все прочие значения  $S_A$  соответствуют либо сеансам передачи других пользователей, либо ошибке. Итак, в чем смысл приведенного примера? Ответ на вопрос очевиден, если рассмотреть декодирование сигнала пользователя В с помощью кода  $S_A$  (т.е. декодирование производится посредством неправильного кода). Если пользователь В пересылает бит 1, последовательность  $d$  равна  $\langle 1, 1, -1, -1, 1, 1 \rangle$ . Тогда

$$S_A(1, 1, -1, -1, 1, 1) = (-1) \times 1 + (-1) \times (-1) + 1 \times (-1) + 1 \times 1 + (-1) \times (-1) + (-1) \times 1 = 0.$$

Таким образом, нежелательный сигнал пользователя В не виден вообще. Легко проверить, что если В будет пересылать бит со значением 0, декодер снова даст результат  $S_A = 0$ . Если пользователи А и В одновременно передают сигналы  $s_A$  и  $s_B$ , соответственно, а декодер приемника является линейным, то имеет место равенство  $S_A(s_A + s_B) = S_A(s_B) + S_A(s_A) = S_A(s_A)$ . Равенство справедливо, поскольку при использовании кода А декодер попросту игнорирует сигналы пользователя В. Коды пользователей А и В, для которых выполняется условие  $S_A(c_B) = S_B(c_A) = 0$ , называются *ортгоналными*. Такие коды достаточно удобно использовать, однако на практике достичь полной ортогональности сложно. В большинстве случаев при  $X \neq Y$  абсолютное значение  $S_X(c_Y)$  мало, но не равно нулю. Таким образом, легко различить варианты  $X \neq Y$  и  $X = Y$ . В нашем примере  $S_A(c_C) = S_C(c_A) = 0$ , а  $S_B(c_C) = S_C(c_B) = 2$ . В последнем случае сигнал пользователя С не будет полностью отфильтрован декодером и внесет незначительный вклад в декодированный сигнал. С помощью декодера  $S_B$  приемник может выделить сигнал, переданный и, даже если в рассматриваемой ячейке одновременно передают собственные сигналы еще несколько пользователей.

Результаты приведенного выше обсуждения обобщаются в табл. 7.1.

Таблица 7.1. Пример CDMA

**а) Коды пользователей**

Пользователь А	1	-1	-1	1	-1	1
Пользователь В	1	1	-1	-1	1	1
Пользователь С	1	1	-1	1	1	-1

**б) Передача сигнала пользователем А**

Передача (бит данных = 1)	1	-1	-1	1	-1	1	
Кодовое слово приемника	1	-1	-1	1	-1	1	
Умножение	1	1	1	1	1	1	= 6
Передача (бит данных = 0)	-1	1	1	-1	1	-1	
Кодовое слово приемника	1	-1	-1	1	-1	1	
Умножение	-1	-1	-1	-1	-1	-1	= -6

## в) Передача сигнала пользователем В; приемник пытается восстановить передачу пользователя А

Передача (бит данных = 1)	1	1	-1	-1	1	1	
Кодовое слово приемника	1	-1	-1	1	-1	1	
Умножение	1	-1	1	-1	-1	1	= 0

## г) Передача сигнала пользователем С; приемник пытается восстановить передачу пользователя В

Передача (бит данных = 1)	1	1	-1	1	1	-1	
Кодовое слово приемника	1	1	-1	-1	1	1	
Умножение	1	1	1	-1	1	-1	= 2

## д) Передача сигнала пользователями В и С; приемник пытается восстановить передачу пользователя В

В (бит данных = 1)	1	1	-1	-1	1	1	
С (бит данных = 1)	1	1	-1	1	1	-1	
Суммарный сигнал	2	2	-2	0	2	0	
Кодовое слово приемника	1	1	-1	-1	1	1	
Умножение	2	2	2	0	2	0	= 8

На практике приемник CDMA способен либо полностью отфильтровать побочные сигналы пользователей, либо вклад таких сигналов выражается в шуме незначительной мощности. В то же время система связи может выйти из строя, если одновременно с пользователем, сигнал которого пытается получить приемник, доступ к тому же каналу пытаются получить множество других пользователей или если один либо несколько мешающих сигналов имеют слишком большую мощность (например, из-за нахождения пользователя близко к приемнику), что называется “проблемой расположения” (near/far problem).

## CDMA для системы DSSS

Рассмотрим CDMA как систему DSSS-BPSK. На рис. 7.11 приводится пример такой системы с  $n$  пользователями, которые для передачи применяют индивидуальные, взаимно ортогональные псевдослучайные последовательности (сравните с рис. 7.7). В процессе передачи сигнала определенным пользователем последовательность данных  $d_i(t)$  модулируется согласно схеме BPSK. Полученный сигнал с шириной полосы  $W$ , умножается на код расширения пользователя  $c_i(t)$ . Все сигналы (а также шум) поступают на антенну приемника. Предположим, что приемник пытается восстановить данные, полученные от пользователя 1. Полученный сигнал умножается на код расширения пользователя 1, после чего демодулируется. В результате ширина полосы этой части суммарного сигнала, соответствующей пользователю 1, уменьшается пропорционально скорости передачи данных и становится равной полосе исходного сигнала до расширения спектра. Поскольку остальная часть суммарного сигнала ортогональна коду рас-

ширину пользователя 1, она будет по-прежнему иметь ширину полосы  $W_s$ . Таким образом, энергия нежелательных сигналов распределена по широкой полосе спектра, а искомый сигнал сконцентрирован в узкой полосе. Если в контуре демодулятора присутствует полосовой фильтр, искомый сигнал можно успешно восстановить.

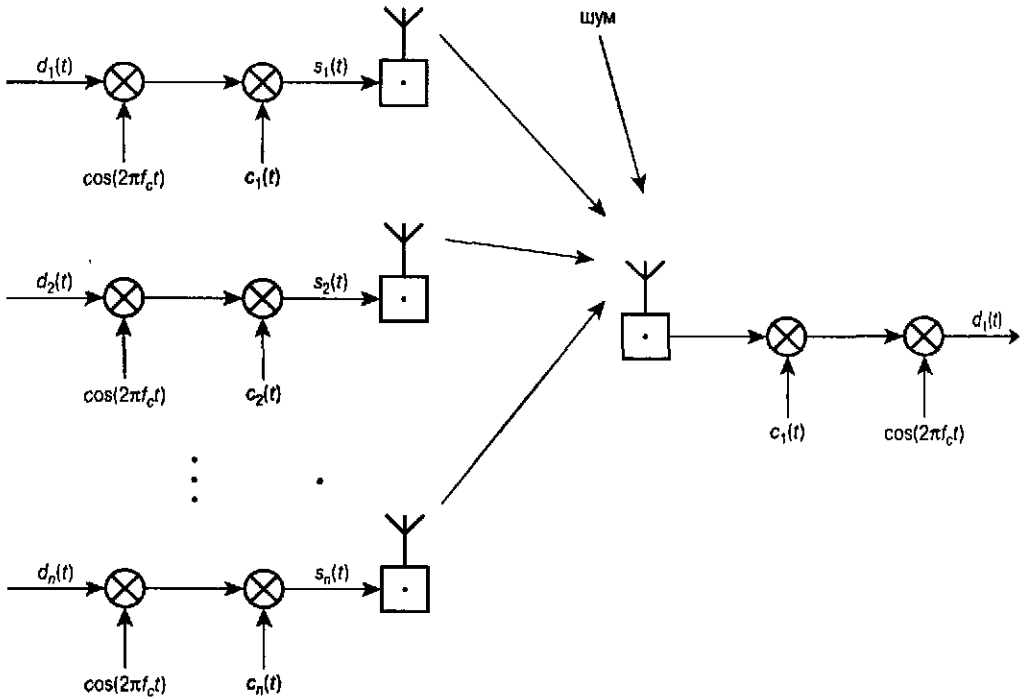


Рис. 7.11. CDMA в среде DSSS

## 7.5. СОЗДАНИЕ ПОСЛЕДОВАТЕЛЬНОСТЕЙ РАСШИРЕНИЯ

Как уже упоминалось, последовательность расширения спектра  $c(t)$  представляет собой последовательность двоичных чисел, которая известна как приемнику, так и передатчику. Расширение спектра состоит в умножении (применении исключающего ИЛИ) входных данных на код расширения. При этом скорость передачи данных для кода расширения больше скорости передачи входной информации. Сужение после получения сигнала производится посредством кода расширения, точно синхронизированного с принятым сигналом.

Скорость передачи данных сигнала равна скорости передачи последовательности расширения, следовательно, увеличивается скорость передачи данных сигнала, а это требует увеличения ширины полосы. Кроме того, возрастает избыточность системы. Коды расширения выбираются так, чтобы результирующий сигнал был шумоподобным; следовательно, код расширения должен содержать приблизительно равное число нулей и единиц и мало или нуль повторяющихся

последовательностей. Если коды расширения используются в среде CDMA, существует дополнительное требование — отсутствие корреляции. При получении множества сигналов, каждый из которых характеризуется индивидуальным кодом расширения, необходимо, чтобы приемник мог расшифровать отдельный сигнал, используя соответствующий код. Сигналы с расширенным спектром не должны коррелировать между собой, чтобы все сигналы, за исключением искомого, были подобны шуму, что позволит избежать интерференции при сужении определенного сигнала. Вследствие введения высокой избыточности при расширении сужение позволяет справиться с интерференцией с посторонними сигналами, которые присутствуют в той же полосе.

Существует две категории последовательностей расширения спектра: ортогональные коды и псевдослучайные последовательности. Псевдослучайные последовательности обычно применяются в системах FHSS и DSSS, не использующих схему CDMA. В системах DSSS с CDMA используются как псевдослучайные последовательности, так и ортогональные коды. Оба названных подхода рассмотрены ниже.

## **Псевдослучайные последовательности**

В идеальном случае последовательность расширения спектра представляет собой случайный ряд двоичных единиц и нулей. В то же время, поскольку приемник и передатчик должны иметь копии кода, генерирование последовательности расширения спектра должно быть предсказуемым. При этом последовательность должна обладать свойствами случайного ряда чисел. Данные требования могут быть выполнены с помощью генератора псевдослучайных чисел, создающего повторяемую периодическую последовательность, обладающую свойствами случайного ряда.

Псевдослучайные последовательности генерирует алгоритм, в котором используется некоторое исходное (или начальное) число. Алгоритм является детерминированным, поэтому генерируемые числа не являются статистически случайными. В то же время при достаточно хорошем алгоритме получающаяся последовательность чисел успешно пройдет многие разумные тесты на случайность, подобные числа часто называют псевдослучайными. Важным является то, что предсказать последовательность без знания алгоритма и начального числа невозможно. Следовательно, успешно декодировать сигнал сможет только приемник, совместно с передатчиком владеющий нужной информацией.

Псевдослучайные последовательности нашли широкое применение в компьютерных технологиях и системах связи. На сегодняшний день методы использования таких последовательностей достаточно хорошо развиты. Сначала мы рассмотрим необходимые свойства псевдослучайной последовательности, после чего перейдем к анализу метода генерации последовательностей, который чаще всего используется в системах связи расширенного спектра.

### **Свойства псевдослучайных последовательностей**

Два наиболее важных свойства псевдослучайных последовательностей — непредсказуемость и случайность.

При генерации последовательности псевдослучайных чисел традиционной задачей является обеспечение соответствия этой последовательности четко определенным критериям случайности. Два приведенных ниже критерия используются для подтверждения случайного характера числовых последовательностей.

- **Равномерное распределение.** Распределение чисел последовательности должно быть равномерным; т.е. все числа последовательности должны появляться с приблизительно одинаковой частотой. Для последовательности двоичных чисел данное определение необходимо расширить, поскольку членами последовательности являются только два числа (0 и 1). В общем случае желательными являются такие свойства.
  - **Свойство баланса.** В длинной последовательности количество двоичных единиц должно приближаться к  $1/2$  от общего числа.
  - **Серийность.** Серией (run) называют последовательность, полностью состоящую из двоичных нулей или единиц. Появление другой цифры означает начало новой серии. Длина приблизительно половины всех серий должна равняться 1; четверть всех серий должны иметь длину 2; одна восьмая — длину 3 и т.д.
- **Независимость.** Ни одно из чисел последовательности не может быть получено с помощью других чисел, также принадлежащих последовательности.

Хотя критерии соответствия числовой последовательности какому-либо виду статистического распределения (например, равномерному) определены довольно четко, критериев независимости членов последовательности не существует. Впрочем, с помощью определенных тестов можно показать, что члены последовательности не являются независимыми. В общем случае такие тесты проводятся до тех пор, пока не появится достаточная уверенность в том, что элементы последовательности независимы.

Для реализации практических задач, таких, как использование расширенного спектра, существует еще одно требование.

- **Свойство корреляции.** Если определенный промежуток последовательности периодически сравнивается с циклическим сдвигом этого же промежутка, то разность числа совпадений и числа несовпадений не должна превышать 1.

### Использование линейного регистра сдвига с обратной связью

Генератор псевдослучайных чисел, который используется в системе расширенного спектра, обычно состоит из контура исключяющего ИЛИ и линейного регистра сдвига с обратной связью (linear feedback shift register — LFSR). Фактически регистр LFSR является цепью 1-битовых элементов памяти. Каждый элемент имеет линию выхода, которая предоставляет информацию о текущем хранящемся значении, а также линию ввода. Через дискретные промежутки времени (такты) текущее значение элемента памяти заменяется значением из линии ввода. Весь регистр LFSR синхронизирован, поэтому на каждом такте происходит 1-битовый сдвиг всего регистра.

Реализация контура выглядит следующим образом.

1. LFSR содержит  $n$  бит.
2. Присутствуют контуры исключяющего ИЛИ: от 1 до  $(n - 1)$ .
3. Наличие или отсутствие контура соответствует наличию или отсутствию соответствующего члена порождающего многочлена (объясняется ниже)  $P(X)$ , исключая член  $X^n$ .

Для описания псевдослучайного регистра LFSR используются два эквивалентных метода. Генератор можно рассматривать как реализацию суммы (по схеме исключающего ИЛИ) членов:

$$B_n = A_0 B_0 \oplus A_1 B_1 \oplus A_2 B_2 \oplus \dots \oplus A_{n-1} B_{n-1}. \quad (7.7)$$

На рис. 7.12 графически изображено уравнение (7.7). На практике при создании регистра контуры умножения не используются; вместо этого для  $A_i = 0$  удаляется соответствующий контур исключающего ИЛИ. На рис. 7.13, а приводится пример 4-битового регистра LFSR, реализующего следующее уравнение:

$$B_3 = B_0 \oplus B_1. \quad (7.8)$$

Использование регистра сдвига имеет несколько преимуществ. Последовательности, созданные LFSR, могут выглядеть почти случайными в течение достаточно большого промежутка времени, что помогает сделать сигнал подобным шуму. Регистры LFSR легко реализовать на аппаратном уровне, они работают с достаточно высокими скоростями; это важно, поскольку скорость расширения больше скорости передачи данных.

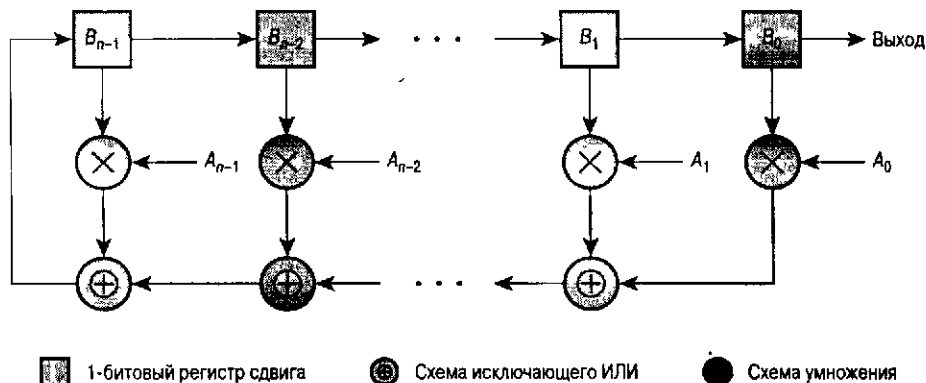
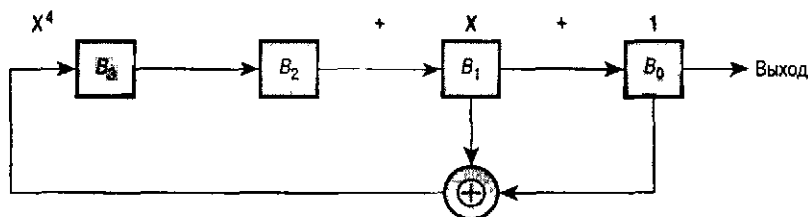


Рис. 7.12. Генератор двоичных последовательностей с использованием линейного регистра сдвига с обратной связью

Можно доказать, что выходной сигнал LFSR является периодическим с максимальным периодом  $N = 2^n - 1$ . Последовательность, полностью состоящая из нулей, может возникнуть только если начальное содержимое всех разрядов регистра LFSR равно нулю или все коэффициенты в уравнении (7.7) равны нулю (отсутствует обратная связь). Всегда можно найти такую конфигурацию обратной связи, что период будет равен  $N$ ; сгенерированные в таком случае последовательности называют *последовательностями максимальной длины* (maximal-length sequences), или *m-последовательностями*. Такие *m-последовательности* применяются для синхронизации с помощью приемника и в схемах множественного доступа (например, CDMA), что будет объяснено позднее.

На рис. 7.13, б показана генерация *m-последовательности* контуром LFSR, который изображен на рис. 7.13, а. Регистр LFSR реализует уравнение (7.8) с начальным состоянием 1000 ( $B_3 = 1; B_2 = 0; B_1 = 0; B_0 = 0$ ). Рис. 7.13, б представляет собой таблицу, в которой показано содержимое LFSR в различные моменты времени. Каждая строка таблицы соответствует текущим значениям, которые хранятся в четырех

элементах регистра. Кроме того, указан выход контура исключающего ИЛИ. Последний столбец содержит значение выходного бита,  $B_0$ . Отметим, что выходная последовательность повторяется с периодом 15 бит. Период последовательности (равный длине  $m$ -последовательности) можно вычислить следующим образом:  $15 = 2^4 - 1$ . Как видно из табл. 7.2, генерируемые значения  $m$ -последовательности не зависят от исходного состояния регистра LFSR (за исключением состояния 0000). При различных исходных состояниях регистра создание  $m$ -последовательности начинается с различных точек, но сама она не меняется.



а) Реализация регистра сдвига

Состояние	$B_3$	$B_2$	$B_1$	$B_0$	$B_0 \oplus B_1$	Выход
Исходное = 0	1	0	0	0	0	0
1	0	1	0	0	0	0
2	0	0	1	0	1	0
3	1	0	0	1	1	1
4	1	1	0	0	0	0
5	0	1	1	0	1	0
6	1	0	1	1	0	1
7	0	1	0	1	1	1
8	1	0	1	0	1	0
9	1	1	0	1	1	1
10	1	1	1	0	1	0
11	1	1	1	1	0	1
12	0	1	1	1	0	1
13	0	0	1	1	0	1
14	0	0	0	1	1	1
15 = 0	1	0	0	0	0	0

б) Пример для исходного состояния 1000

Рис. 7.13. Контур с регистрами сдвига для генерации псевдослучайной последовательности

Таблица 7.2. 15 псевдослучайных последовательностей, сгенерированных схемой, изображенной на рис. 7.13, а

Исходное состояние	Выходная последовательность
1000	000100110101111
0100	001001101011110
0010	010011010111100
1001	100110101111000
1100	001101011110001



Исходное состояние	Выходная последовательность
0110	011010111100010
1011	110101111000100
0101	101011110001001
1010	010111100010011
1101	101111000100110
1110	011110001001101
1111	111100010011010
0111	111000100110101
0011	110001001101011
0001	100010011010111

Изменяя параметр  $A$ , в уравнении (7.7), для любого данного размера регистра LFSR можно сгенерировать определенное число уникальных  $m$ -последовательностей. В табл. 7.3 приводится длина последовательности и число уникальных  $m$ -последовательностей, которые можно сгенерировать при различных размерах регистра LFSR.

Таблица 7.3. Последовательности максимальной длины, сгенерированные регистром сдвига

Число каскадов регистра сдвига, $N$	Длина последовательности, $L = 2^N - 1$	Число $m$ -последовательностей	Пример порождающего многочлена	Параметр полинома $B_n$ (см. рис. 7.12)
2	3	1	$X^2 + X + 1$	$B_1 \oplus B_0$
3	7	2	$X^3 + X + 1$	$B_1 \oplus B_0$
4	15	2	$X^4 + X + 1$	$B_1 \oplus B_0$
5	31	6	$X^5 + X^2 + 1$	$B_2 \oplus B_0$
6	63	6	$X^6 + X + 1$	$B_1 \oplus B_0$
7	127	18	$X^7 + X + 1$	$B_1 \oplus B_0$
8	255	16	$X^8 + X^6 + X^5 + X + 1$	$B_6 \oplus B_1 \oplus B_1 \oplus B_0$
9	511	48	$X^9 + X^4 + 1$	$B_4 \oplus B_0$
10	1 023	60	$X^{10} + X^3 + 1$	$B_3 \oplus B_0$
11	2 047	176	$X^{11} + X^2 + 1$	$B_2 \oplus B_0$
12	4 095	144	$X^{12} + X^7 + X^4 + X^3 + 1$	$B_7 \oplus B_4 \oplus B_3 \oplus B_0$
13	8 191	630	$X^{13} + X^4 + X^3 + X + 1$	$B_4 \oplus B_3 \oplus B_1 \oplus B_0$
14	16 383	756	$X^{14} + X^{12} + X^{11} + X + 1$	$B_{12} \oplus B_{11} \oplus B_1 \oplus B_0$
15	32 767	1800	$X^{15} + X + 1$	$B_1 \oplus B_0$
20	1 048 575	2400	$X^{20} + X^3 + 1$	$B_3 \oplus B_0$

Контур LFSR можно эквивалентно определить через порождающий многочлен. Порождающий многочлен (или генератор)  $P(X)$ , соответствующий уравнению (7.7), можно записать в следующем виде:

$$P(X) = A_0 + A_1X + A_2X^2 + \dots + A_{n-1}X^{n-1}. \quad (7.9)$$

Важным свойством порождающего многочлена является то, что величина, обратная ему, — это последовательность, сгенерированная соответствующим регистром LFSR. Например, для 3-битового LFSR с полиномом  $P(X) = 1 + X + X^3$  сгенерированная последовательность находится как  $1/(1 + X + X^3)$ . Процедура деления представлена на рис. 7.14, а результат представлен ниже.

$$1 + X + X^2 + (0 \times X^3) + X^4 + (0 \times X^5) + (0 \times X^6).$$

Далее последовательность повторяется; это означает, что выход регистра имеет такой вид:

$$1110100.$$

Поскольку период полученной последовательности равен  $7 = 2^3 - 1$ , это —  $m$ -последовательность. Отметим, что деление производилось не совсем привычным способом. Причина такого отличия: в операциях по модулю 2 (или при использовании исключающего ИЛИ) вычитание и сложение дают один и тот же результат.

Список порождающих многочленов, которые генерируют  $m$ -последовательности для регистров LFSR разного размера, приводится в табл. 7.3.

$$\begin{array}{r}
 \begin{array}{r}
 1 \\
 - 1 + X + \quad X^3 \\
 \hline
 X + \quad X^3 \\
 - X + X^2 + \quad X^4 \\
 \hline
 X^2 + X^3 + X^4 \\
 - X^2 + X^3 + \quad X^5 \\
 \hline
 X^4 + X^5 \\
 - X^4 + X^5 + \quad X^7 \\
 \hline
 X^7 \\
 - X^7 + X^8 + \quad X^{10} \\
 \hline
 X^8 + \quad X^{10} \\
 - X^8 + X^9 + \quad X^{11} \\
 \hline
 \end{array}
 \quad \left| \begin{array}{l}
 1 + X + X^3 \\
 \hline
 1 + X + X^2 + \quad X^4 + \quad X^7 + X^8 + \dots
 \end{array} \right.
 \end{array}$$

Рис. 7.14.  $1/(1 + X + X^3)$

### Свойства $m$ -последовательностей

Несколько свойств  $m$ -последовательностей делают их предпочтительными для использования в системах расширенного спектра.

**Свойство 1.** Любая  $m$ -последовательность содержит  $2^{n-1}$  единиц и  $2^{n-1} - 1$  нулей.

**Свойство 2.** Если перемещать окно длиной  $n$  вдоль выходной последовательности, то за  $N$  сдвигов (где  $N = 2^{n-1}$ ) каждая последовательность, состоящая из  $n$  элементов (исключая последовательность из одних нулей), появится ровно один раз.

**Свойство 3.** Любая  $m$ -последовательность содержит: одну серию единиц длиной  $n$ ; одну серию нулей длиной  $n - 1$ ; по одной серии нулей и единиц длиной  $n - 2$ ; по две серии нулей и единиц длиной  $n - 3$ ; и, в общем случае, по  $2^{n-3}$  серий нулей и единиц длиной 1.

Для множества систем связи последовательность нулей и единиц заменяется последовательностью  $\pm 1$ , при этом двоичный нуль представляется  $-1$ , а двоичная единица —  $1$ <sup>5</sup>. По определению периодическую автокорреляционную функцию полученной последовательности можно записать в таком виде:

$$R(\tau) = \frac{1}{N} \sum_{k=1}^N B_k B_{k-\tau}. \quad (7.10)$$

Тогда можно добавить следующее свойство.

**Свойство 4.** Периодическая автокорреляционная функция  $m$ -последовательности, состоящей из  $\pm 1$ , имеет такой вид:

$$R(\tau) = \begin{cases} 1 & \tau = 0, n, 2N, \dots \\ -\frac{1}{N} & \text{в противном случае} \end{cases}$$

На рис. 7.15 изображен общий вид автокорреляционной функции для  $m$ -последовательностей, а также автокорреляционная функция для  $m$ -последовательностей, сгенерированных 4-битовым регистром LFSR.

По сути, корреляцию можно определить как меру схожести двух разных наборов данных. Корреляция определяется в диапазоне от  $-1$  до  $1$  со следующей интерпретацией значений.

Корреляционное значение	Интерпретация
1	Две последовательности полностью совпадают
0	Две последовательности абсолютно не связаны
-1	Последовательности являются зеркальными образами друг друга

Другие значения соответствуют частичной корреляции. Автокорреляционная функция, определение которой приводится в формуле (7.10), — это корреляция последовательности со всеми возможными фазовыми сдвигами этой же последовательности. Корреляционное значение истинно случайных данных близко к 0 для всех автокорреляционных функций с любым фазовым сдвигом, отличным от нуля. Для  $m$ -последовательностей данное утверждение выполняется. Можно показать, что автокорреляционная функция  $m$ -последовательности (единственный острый пик) существенно облегчает синхронизацию приемника.

<sup>5</sup> Некоторые авторы используют обратное обозначение — двоичному нулю ставится в соответствие “+1”, а двоичной единице — “-1”. Существенной разницы в представлениях нет, достаточно лишь непротиворечиво использовать однажды принятую договоренность.

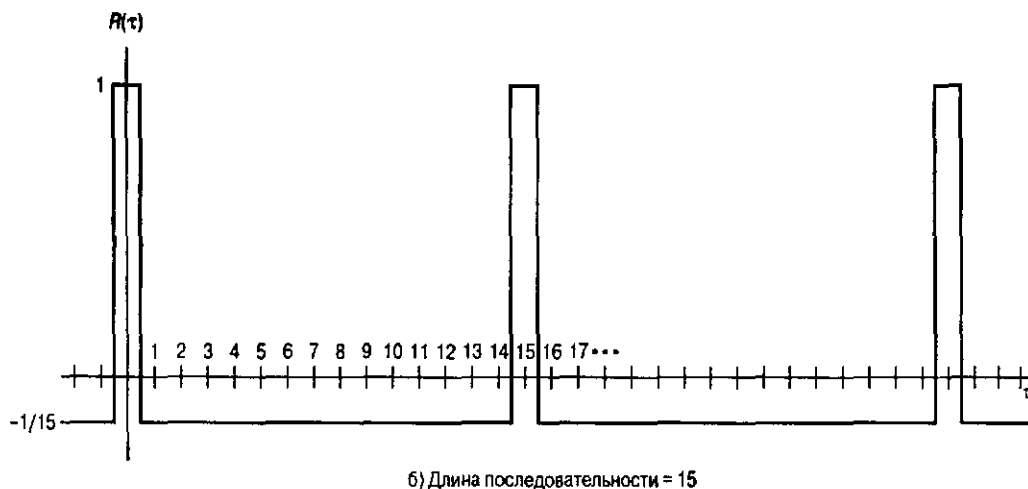
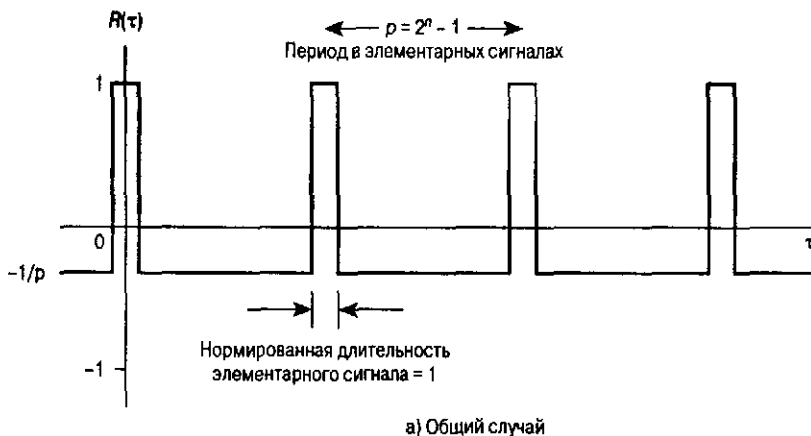


Рис. 7.15. Автокорреляционная функция псевдослучайных последовательностей

При рассмотрении расширенного спектра другой важной функцией является функция взаимной корреляции. В данном случае сравниваются два сигнала, полученные из разных источников, а не одна последовательность с разными фазовыми сдвигами. Функция взаимной корреляции двух источников  $A$  и  $B$  определяется так:

$$R_{A,B}(\tau) = \frac{1}{N} \sum_{k=1}^N A_k B_{k-\tau} \quad (7.11)$$

В общем случае значение этой функции при сравнении некоторого сигнала с истинно случайной последовательностью мало, что дает следующие преимущества.

1. Взаимная корреляция  $m$ -последовательности и шума незначительна, что можно использовать при отсеивании шумов приемником.
2. Взаимная корреляция двух  $m$ -последовательностей незначительна, что используется в системах CDMA, поскольку приемник получает возможность

выборочного отсеивания сигналов расширенного спектра, созданных с использованием разных  $m$ -последовательностей.

### Последовательности Голда

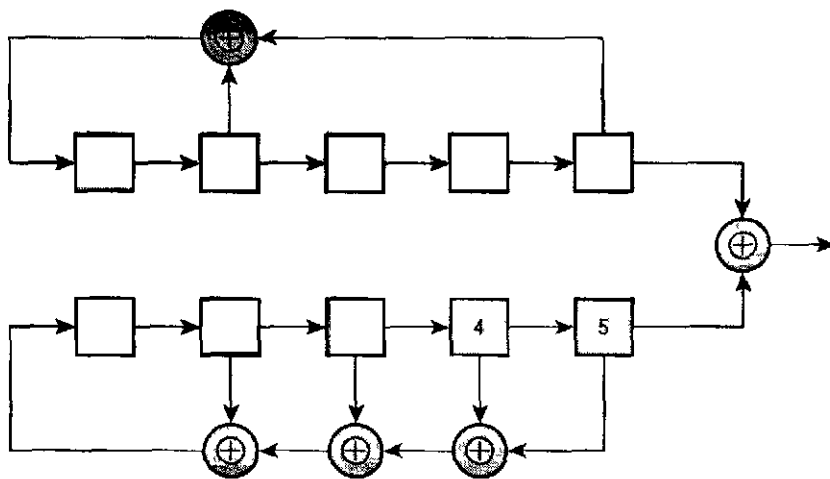
Описанные выше  $m$ -последовательности достаточно легко генерируются и широко применяются в системах FHSS и DSSS, не использующих CDMA. Впрочем, для систем DSSS с CDMA  $m$ -последовательности не являются оптимальными. Для связи по схеме CDMA требуется создать семейство последовательностей расширения (по одной для каждого пользователя), коды которых должны иметь определенную взаимную корреляцию. Как правило, с  $m$ -последовательностями этого сделать не удастся, подходящими оказываются последовательности Голда (Gold sequences). Привлекательность последовательностей Голда объясняется тем, что для генерации множества уникальных кодов требуется единственный простой контур.

Последовательность Голда создается путем применения операции исключения ИЛИ к двум синхронизированным  $m$ -последовательностям, соответствующий пример приводится на рис. 7.16, а. В общем случае длина результирующей последовательности не будет максимальной. Кроме того, желаемые последовательности Голда можно сгенерировать только с помощью предпочтительных пар  $m$ -последовательностей. Такие пары получаются из таблиц пар или генерируются соответствующим алгоритмом. В работе [DIXO94] приводится список предпочтительных пар и алгоритма их генерации. В данном разделе будут приведены только общие математические принципы, на основе которых создается код Голда.

Пусть имеется  $m$ -последовательность, которая представляется двоичным вектором  $a$  длины  $N$ , и путем выборки каждого  $q$ -го символа  $a$  создается новая последовательность  $a'$ . Количество выбранных символов последовательности  $a'$  должно быть равно  $N$ , так что для набора недостающих элементов используются последовательные копии  $a$ . Полученную последовательность  $a'$  называют децимацией  $a$  и обозначают  $a' = a[q]$ . Период последовательности  $a'$  не обязательно равен  $N$ , поэтому  $a'$  не обязательно будет  $m$ -последовательностью. Можно доказать, что  $a'$  будет  $m$ -последовательностью с периодом  $N$  тогда и только тогда, когда  $\text{НОД}(n, q) = 1$  (НОД — наибольший общий делитель). Другими словами,  $n$  и  $q$  не должны иметь общих множителей за исключением 1. Для создания последовательности Голда необходимо найти предпочтительную пару  $m$ -последовательностей  $a$  и  $a' = a[q]$ , которая удовлетворяет следующим условиям:

1.  $n/4 \neq 0$ ; т.е. все  $n$ , кроме 0, 4, 8, 12, ... ;
2.  $q$  — нечетное; кроме того,  $q = (2^k + 1)$  или  $q = (2^{2k} - 2^k + 1)$  для некоторого  $k$ ;
3.  $\text{НОД}(n, k) = \begin{cases} 1 & \text{при нечетном } n \\ 2 & \text{при } n = 2 \text{ по модулю } 4 \end{cases}$ .

Для регистров сдвига длины  $n$  взаимная корреляция последовательностей Голда, полученных с помощью предпочтительных пар, ограничена величиной  $R$ :  $|R| \leq 2^{(n+1)/2} + 1$  для нечетных  $n$ ; и  $|R| \leq 2^{(n+2)/2} + 1$  для четных  $n$ .



а) Реализация регистра сдвига

Последовательность 1: 1111100011011101010000100101100  
 Последовательность 2: 1111100100110000101101010001110  
 Применение исключающего ИЛИ после сдвига 0: 00000001111011011111011101000010  
 Применение исключающего ИЛИ после сдвига 1: 0000101010111100001010000110001  
 ⋮  
 Применение исключающего ИЛИ после сдвига 30: 1000010001000101000110001101011

б) Полученные последовательности Голда

Рис. 7.16. Пример создания набора последовательностей Голда [DINA98]

При наличии предпочтительной пары набор кодов Голда составляется следующим образом:  $\{a, a', a \oplus a, a \oplus Da', a \oplus D^2a', \dots, a \oplus D^{N-1}a'\}$ , где  $D$  — элемент задержки; т.е.  $D$  представляет 1-битовый сдвиг  $a'$  относительно  $a$ . Для создания кодов Голда начальное состояние двух регистров сдвига устанавливается равным векторам, состоящим только из единиц. К полученным последовательностям применяется операция исключающего ИЛИ, в результате чего получается одна последовательность Голда. Таким образом получают три последовательности. Остальные последовательности создаются так: вторая из первых двух последовательностей смещается на один бит, после чего повторно применяется операция исключающего ИЛИ. Процесс продолжается, пока не будут перебраны все возможные сдвиги, причем после применения исключающего ИЛИ получается очередная последовательность набора. Можно показать, что для предпочтительной пары 5-битовых регистров сдвига любой сдвиг начального состояния от 0 до 30 бит даст новую последовательность Голда (сдвиг на 31 бит соответствует нулевому сдвигу). Таким образом, созданный набор последовательностей Голда включает в себя две исходные  $m$ -последовательности, а также все сгенерированные последовательности — всего 33 (рис. 7.16, б).

В общем случае период любого кода последовательности Голда, сгенерированной двумя  $n$ -битовыми регистрами сдвига, равен  $N = 2^n - 1$ , т.е. равен периоду  $m$ -

последовательностей. В семействе кодов Голда насчитывается  $(N + 2)$  кода. В качестве подтверждения эффективности кодов Голда можно сказать, что при  $n = 13$  ( $N = 2^{13} - 1 = 8191$ ) существует 630  $m$ -последовательностей (табл. 7.3) и имеются пары таких последовательностей, корреляционное значение которых  $R = 703$ , тогда как последовательность Голда гарантирует выбор пар с  $R \leq (2^{(n+1)/2} + 1) = 129$ .

### Последовательности Касами

Еще одним важным семейством псевдослучайных последовательностей являются последовательности Касами (Kasami sequences), наиболее вероятной сферой применения которых являются беспроводные схемы третьего поколения. Последовательности Касами определяются процедурой, сходной с той, что использовалась для кодов Голда. Различают большие и малые наборы последовательностей Касами.

Для четных  $n$  создается **малый набор** последовательностей Касами, состоящий из  $M = 2^{n/2}$  разных последовательностей, период каждой равен  $N = 2^n - 1$ . Первым шагом в создании набора является децимация с шагом  $q = 2^{n/2} + 1$   $m$ -последовательности  $a$  с периодом  $N$ . Можно показать, что полученная последовательность  $a'$  имеет период  $2^{n/2} - 1$ . Следующий шаг — дублирование одного периода  $a'$   $q$  раз, после чего получается последовательность длиной  $(2^{n/2} - 1)(2^{n/2} + 1) = N$ . Например, для  $n = 10$  период  $a$  составляет  $2^{10} - 1 = 1023$ , а период  $a'$  будет равен  $2^5 - 1 = 31$ . За 1023 бита последовательности  $a'$  последовательность с периодом 31 бит повторится 33 раза. Последний шаг в создании набора Касами — применение операции исключающего ИЛИ к  $N$  битам последовательности  $a$  и  $N$  битам  $a'$ , а также  $2^{n/2} - 1$  циклических сдвигов на 1 бит последовательности  $a'$ .

Можно показать, что максимальная взаимная корреляция полученного набора последовательностей равна  $2^{n/2} - 1$ . Это меньше, чем для последовательностей Голда, поэтому последовательности Касами эффективнее.

**Большой набор** последовательностей Касами также состоит из определенных рядов, каждый из которых имеет период  $N = 2^n - 1$  для четных  $n$  и содержит последовательности Голда и малые наборы последовательностей Касами. Набор генерируется путем децимации  $m$ -последовательности  $a$  с периодом  $N$  при двух различных значениях  $q$ . При  $q = 2^{n/2} + 1$  создается последовательность  $a'$ , при  $q = 2^{(n+2)/2} + 1$  — последовательность  $a''$ . Набор формируется после применения операции исключающего ИЛИ к  $a$ ,  $a'$ ,  $a''$  и различным сдвигам  $a'$  и  $a''$ . Можно показать, что максимальная взаимная корреляция для большого набора последовательностей Касами равна  $2^{(n+2)/2}$ .

### Ортогональные коды

В отличие от псевдослучайных последовательностей, ортогональные коды — это наборы последовательностей, для которых все попарные значения функции взаимной корреляции равны нулю. Ортогональный набор последовательностей имеет следующее свойство:

$$\sum_{k=0}^{M-1} \phi_i(k\tau)\phi_j(k\tau) = 0, \quad i \neq j.$$

Здесь  $M$  — длина каждой последовательности набора;  $\phi_i$  и  $\phi_j$  —  $i$ -й и  $j$ -й члены набора, соответственно;  $t$  — время передачи одного бита.

Ортогональные коды фиксированной и переменной длины применяются в системах CDMA. В такой системе каждому мобильному пользователю выделяется одна из последовательностей набора в качестве кода расширения, при этом взаимная корреляция между всеми пользователями равна нулю.

### Коды Уолша

Код Уолша (Walsh code) — наиболее распространенный ортогональный код, используемый в системах CDMA. Набор кодов Уолша длиной  $n$  состоит из  $n$  строк матрицы Уолша  $n \times n$ . Матрица определяется следующим рекуррентным выражением:

$$W_1 = (0) \quad W_{2n} = \begin{pmatrix} W_n & W_n \\ W_n & -W_n \end{pmatrix}.$$

Здесь  $n$  — размерность матрицы; верхняя черта соответствует применению логического НЕ к битам матрицы. Для матриц Уолша справедливо следующее: каждая строка ортогональна любой другой строке, а также любой строке, к которой применена операция логического НЕ.

На рис. 7.17 представлены матрицы Уолша размерностью 2, 4 и 8. Напомним, что для вычисления взаимной корреляции 1 заменяется +1, а 0 — -1.

$$W_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad W_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} \quad W_8 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

Рис. 7.17. Матрицы Уолша

Ортогональные коды расширения, такие, как последовательности Уолша, могут использоваться, только если все пользователи одного канала CDMA синхронизированы с точностью до малой доли элементарного сигнала. Из-за того что взаимная корреляция различных сдвигов последовательностей Уолша не равна нулю, при отсутствии точной синхронизации требуются псевдослучайные последовательности.

### Ортогональные коды переменной длины

Системы CDMA третьего поколения проектируются таким образом, чтобы поддерживать связь с пользователями, имеющими разную скорость передачи данных. Следовательно, эффективная поддержка возможна, если коды расширения используются с разными скоростями, но при этом не нарушается их ортогональность. Предположим, что минимальная скорость передачи данных равна  $R_{\min}$ , а



каждая последующая скорость в 2 раза больше предыдущей. Если для передачи со скоростью  $R_{\text{ном}}$  используется последовательность расширения длины  $N$ , то для передачи каждого бита данных тратится ( $N = 2^n$ ) бит последовательности расширения (передается последовательность для информационного бита 0; затем передается дополнительный код для бита 1); таким образом, скорость передачи сигнала равна  $\sqrt{R_{\text{ном}}}$ . Для скорости передачи данных  $2R_{\text{ном}}$  и последовательности длиной  $N/2 = 2^{n-1}$  получаем ту же скорость передачи сигнала  $\sqrt{R_{\text{ном}}}$ . Вообще для получения выходной скорости сигнала  $2^k R_{\text{ном}}$  необходимо использовать код длиной  $2^{n-k}$ .

Наборы ортогональных последовательностей переменной длины легко получаются из матриц Уолша разных размерностей. Более подробно данный вопрос рассмотрен в [DINA98].

## Множественное расширение

При наличии достаточной ширины полосы чрезвычайно эффективным является метод множественного расширения (multiple spreading). Обычно это делается так: с помощью ортогонального кода расширяется скорость передачи данных, что позволяет достичь взаимной ортогональности всех пользователей одной ячейки; затем спектр полученного сигнала расширяется псевдослучайной последовательностью, что дает крайне низкую взаимную корреляцию пользователей разных ячеек. При таком двойном расширении ортогональные коды называются кодами распределения по каналам (channalization codes), а псевдослучайные коды — кодами скремблирования (scrambling codes). Пример использования метода множественного расширения приводится в главе 10 при рассмотрении стандарта IS-95.

## 7.6. РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

Исчерпывающий анализ расширенного спектра приводится в [PETE95] и [DIXO94]. В [TANT98] собраны многие важные публикации по данной тематике, в том числе [PICK82], являющаяся отличным введением в область расширенного спектра.

Хороший обзор псевдослучайных последовательностей в системах расширенного спектра приведен в [MACW76], а обзор CDMA — в [PRAS98].

DIXO94 Dixon R. *Spread Spectrum Systems with Commercial Applications*. — New York: Wiley, 1994.

MACW76 Macwilliams F., Sloane N. Pseudo-Random Sequences and Arrays. — In: *Proceedings of the IEEE*, December 1976 (имеется также в [TANT98]).

PETE95 Peterson R., Ziemer R., Borth D. *Introduction to Spread Spectrum Communications*. — Englewood Cliffs, NJ: Prentice Hall, 1995.

PICK82 Pickholtz R., Schilling D., Milstein L. Theory of Spread Spectrum Communications — A Tutorial. — *IEEE Transactions on Communications*, May 1982 (имеется также в [TANT98]).

PRAS98 Prasad R., Ojanpera T. An Overview of CDMA Evolution: Toward Wideband CDMA. — *IEEE Communications Surveys*, Fourth Quarter 1998. (<http://www.comsoc.org/>.)

TANT98 Tantaratana S., Ahmed K., eds. *Wireless Applications of Spread Spectrum Systems: Selected Readings*. — Piscataway, NJ: IEEE Press, 1998.

## 7.7. ТЕРМИНЫ, ВОПРОСЫ И ЗАДАЧИ

### Основные термины

<i>m</i> -последовательность	последовательность	расширение спектра с
автокорреляция	расширения	быстрой
взаимная корреляция	псевдослучайная	скачкообразной
код расширения	последовательность	перестройкой частоты
корреляция	расширение спектра	(быстрая схема FHSS)
линейный регистр	методом прямой	расширение спектра с
сдвига с обратной	последовательности	медленной
связью (LFSR)	(DSSS)	скачкообразной
множественный доступ с	расширение спектра	перестройкой частоты
кодovým разделением	методом	(медленная схема
(CDMA)	скачкообразной	FHSS)
ортогональность	перестройки частоты	расширенный спектр
	(FHSS)	элементарный сигнал

### Вопросы

1. Какая связь существует между шириной полосы сигнала до и после кодирования с использованием расширенного спектра?
2. Перечислите три основных преимущества расширенного спектра.
3. Что такое “расширение спектра со скачкообразной перестройкой частоты”?
4. В чем различие между медленной и быстрой схемами FHSS?
5. Что такое “расширение спектра методом прямой последовательности”?
6. Какая связь существует между скоростью передачи данных до и после кодирования сигнала с использованием DSSS?
7. Что такое “множественный доступ с кодovým разделением каналов”?
8. Объясните различие между автокорреляцией и взаимной корреляцией.

### Задачи

1. Пусть поток данных со скоростью 56 Кбит/с необходимо передать с помощью сигнала расширенного спектра.
  - а. Найдите требуемую ширину полосы канала при следующих значениях SNR: 0,1; 0,01 и 0,001.
  - б. Для обычной системы связи (без использования расширенного спектра) разумная эффективность использования полосы — 1 бит/с/Гц. Другими словами, для передачи потока данных со скоростью 56 Кбит/с используется полоса 56 кГц. Найдите максимальное SNR, при котором передача будет производиться с приемлемым уровнем ошибок. Сравните результат с результатом, полученным при использовании расширенного спектра.

*Подсказка:* см. анализ пропускной способности канала связи в разделе 2.3.

2. Полоса системы FHSS  $W_s = 400$  МГц; ширина полосы отдельного канала равна 100 Гц. Чему равно минимальное число битов псевдослучайного кода, необходимое для перестройки частоты?

3. Система FHSS с модуляцией MFSK ( $M=4$ ) использует 1 000 различных частот. Найдите коэффициент расширения спектра.

4. В приведенной ниже таблице иллюстрируется работа системы FHSS в течение одного полного периода псевдослучайной последовательности.

Время	0	1	2	3	4	5	6	7	8	9	10	11
Входные данные	0	1	1	1	1	1	1	0	0	0	1	0
Частота	$f_1$		$f_3$		$f_{27}$		$f_{26}$		$f_8$		$f_{10}$	
Псевдослучайная последовательность	001				110				011			

Время	12	13	14	15	16	17	18	19
Входные данные	0	1	1	1	1	0	1	0
Частота	$f_1$		$f_3$		$f_2$		$f_2$	
Псевдослучайная последовательность	001				001			

- Найдите период псевдослучайной последовательности.
- Система использует один из методов частотной манипуляции (FSK). Назовите этот метод.
- Сколько битов содержится в одном символе?
- Найдите число используемых частот FSK.
- Найдите длину псевдослучайной последовательности в течение интервала передачи на одной частоте.
- Какой тип перестройки частоты используется системой — быстрый или медленный?
- Найдите полное число возможных изменений частоты.
- Найдите дисперсию восстановленного сигнала.

5. В приведенной ниже таблице характеризуется работа системы FHSS. Используется та же псевдослучайная последовательность, что и в предыдущей задаче.

Время	0	1	2	3	4	5	6	7	8	9	10	11
Входные данные	0	1	1	1	1	1	1	0	0	0	1	0
Частота	$f_1$	$f_{21}$	$f_{11}$	$f_3$	$f_3$	$f_3$	$f_{22}$	$f_{10}$	$f_0$	$f_0$	$f_1$	$f_{22}$
Псевдослучайная последовательность	001	110	011	001	001	001	110	011	001	001	001	110

Время	12	13	14	15	16	17	18	19
Входные данные	0	1	1	1	1	0	1	0
Частота	$f_9$	$f_1$	$f_3$	$f_3$	$f_{22}$	$f_{11}$	$f_3$	$f_3$
Псевдослучайная последовательность	01	00	00	00	11	01	00	00
	1	1	1	1	0	1	1	1

- Найдите период псевдослучайной последовательности.
- Система использует один из методов частотной манипуляции (FSK). Назовите этот метод.
- Сколько битов содержится в одном символе?
- Найдите число используемых частот FSK.
- Найдите длину псевдослучайной последовательности в течение интервала передачи на одной частоте.

- е. Какой тип перестройки частоты используется системой — быстрый или медленный?
- ж. Найдите полное число возможных изменений частоты.
- з. Найдите дисперсию восстановленного сигнала.
5. Рассмотрим схему MFSK со следующими параметрами:  $f_c = 250$  кГц;  $f_d = 25$  кГц;  $M = 8$  ( $L = 3$  бит).
- а. Найдите распределение частот для каждой из 8 возможных комбинаций трех битов данных.
- б. С данной схемой MFSK требуется использовать технологию FHSS с  $k = 2$ ; т.е. система может перестраиваться на 4 различных несущих. Используйте результаты пункта а и найдите распределение  $4 \times 8 = 32$  частот.
6. На рис. 7.18 (на основе иллюстрации из [BELL00]) приводится упрощенная схема кодирования и декодирования CDMA. Имеется семь логических каналов, в каждом из которых применяется схема DSSS с 7-битовым кодом расширения. Все источники данных считаются синхронизированными. Если все семь источников одновременно передадут один бит информации в виде 7-битовой последовательности, приемник получит комбинированный сигнал. При совпадении двух положительных или отрицательных чисел сигнал будет усиливаться; два противоположных числа будут взаимно уничтожаться. Для декодирования данных определенного канала приемник умножает полученный комбинированный сигнал на соответствующий код расширения, суммирует полученные результаты, после чего отрицательному числу ставится в соответствие двоичный 0, а положительному — двоичная 1.
- а. Найдите коды расширения всех семи каналов.
- б. Найдите выходной сигнал приемника для канала 1; найдите значение бита в канале.
- в. Решите пункт б для канала 2.

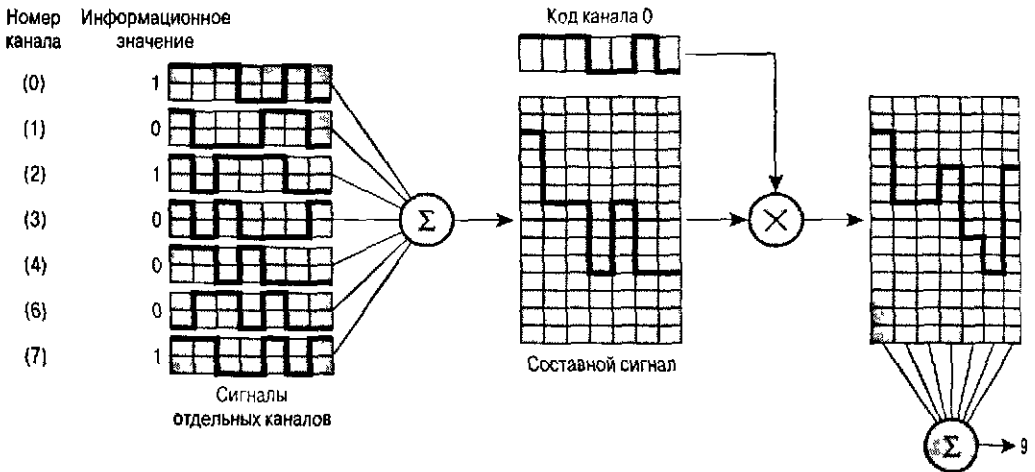


Рис. 7.18. Пример кодирования и декодирования для 7-канальной системы CDMA

8. Для системы, описанной в задаче 7, найдите взаимную корреляцию канала 0 и остальных шести каналов, используя найденные значения кодов расширения.

9. Наиболее распространенным современным методом генерации псевдослучайных чисел является метод линейного сравнения. В алгоритме фигурирует 4 параметра:

$m$  — модуль;  $m > 0$ ;

$a$  — множитель;  $0 \leq a < m$ ;

$c$  — инкремент;  $0 \leq c < m$ ;

$X_0$  — начальное значение  $0 \leq X_0 < m$ .

Последовательность псевдослучайных чисел  $\{X_n\}$  можно получить с помощью следующего итеративного уравнения:

$$X_{n+1} = (aX_n + c) \text{ по модулю } m.$$

Если параметры  $m$ ,  $a$ ,  $c$  и  $X_0$  являются целыми числами, результатом будет последовательность целых чисел, каждое из которых принадлежит промежутку  $0 \leq X_n < m$ . Необходимой характеристикой генератора псевдослучайных чисел является то, что созданная последовательность должна выглядеть случайной. Несмотря на то что на самом деле последовательность не является случайной (поскольку была создана неким алгоритмом), существуют определенные статистические критерии проверки последовательности на случайность. Другая желательная характеристика: функция должна быть порождающей с полным периодом. Т.е. перед повторением функция должна генерировать все числа в промежутке от 0 до  $m$ .

При использовании алгоритма линейного сравнения параметры, с помощью которых получается последовательность с полным периодом, не обязательно будут наилучшими для создания случайных чисел. Рассмотрим в качестве примера два генератора.

$$X_{n+1} = (6X_n) \text{ по модулю } 13,$$

$$X_{n+1} = (7X_n) \text{ по модулю } 13.$$

Запишите две полученные последовательности и убедитесь, что обе они имеют полный период. Какая из последовательностей, по вашему мнению, является более случайной?

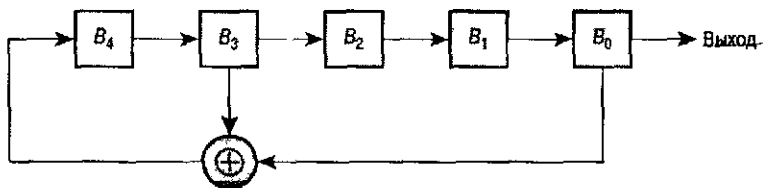
10. Длительные последовательности разных случайных чисел могут генерироваться, только если параметр  $m$  имеет максимальное значение. Принято, что  $m$  приблизительно равно максимальному положительному числу, которое способен отобразить компьютер. В общем случае значение  $m$  выбирают равным (или приблизительно равным)  $2^{31}$ . Многие эксперты рекомендуют использовать значение  $2^{31} - 1$ . Может возникнуть вопрос: почему не использовать значение  $2^{31}$ , ведь это число можно отобразить без использования дополнительных битов, а операции по модулю будут выполняться легче. Однако в общем случае лучше использовать операции по модулю  $2^k - 1$ , а не по модулю  $2^k$ . Почему?
11. При любом использовании псевдослучайных чисел (для шифрования, моделирования или статистических расчетов) довольно небезопасно слепо полагаться на генератор случайных чисел, который, возможно, обнаружится в системной библиотеке компьютера. В [PARK88] доказывалось, что многие современные учебники и программные пакеты основаны на использовании некорректных алгоритмов создания псевдослучайных чисел. Данное упражнение позволит протестировать вашу систему. Процедура проверки основывается на теореме, приписываемой Эрнесто Сезаро (Ernesto Cesaro, доказательство теоремы приводится в [KNUT98]). Теорема формулируется следующим образом: вероятность того, что наибольший общий делитель

двух случайно выбранных чисел будет равен 1, равна  $\frac{6}{\pi^2}$ . Используйте теорему для

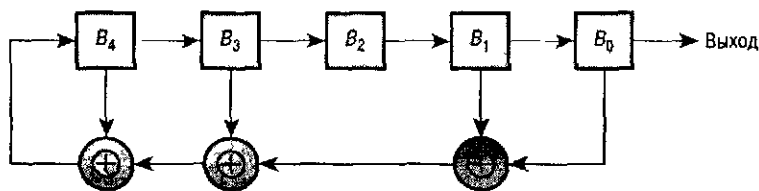
создания программы статистического определения значения  $\pi$ . Основная программа должна вызывать три подпрограммы: генератор случайных чисел из системной библиотеки; подпрограмму вычисления наибольшего общего делителя с помощью алгоритма Евклида; подпрограмму вычисления квадратных корней. Если последние две подпрограммы не включены в вашу системную библиотеку, напишите их самостоятельно. Цикл основной программы должен повторяться многократно для получения вышеупомянутого значения. Если полученный вами результат приблизительно равен 3,14 — поздравляем! Если нет, то с большой вероятностью результат окажется близким к 2,7. Объясните причину появления именно такого некорректного результата вычислений.

12. Задача является демонстрацией того, как для создания  $m$ -последовательностей могут использоваться различные контуры LFSR.

- Предположим, что начальное состояние контура LFSR, изображенного на рис. 7.19, а, — 1 0 0 0. Изобразите процесс создания  $m$ -последовательности в форме таблицы, подобной той, что приведена на рис. 7.13, б.
- Рассмотрим контур, представленный на рис. 7.19, б. Повторите пункт а для того же начального состояния контура. Докажите, что данный контур генерирует  $m$ -последовательность, которая не совпадает с последовательностью, создание которой описано в предыдущем примере.



а)



б)

Рис. 7.19. 5 каскадные контуры LFSR различной конфигурации

- Докажите, что коды в матрице Уолша  $8 \times 8$  взаимно ортогональны. Используйте факт, что произведение любых двух кодов матрицы дает 0.
- Рассмотрим систему CDMA. Пользователи  $A$  и  $B$  применяют коды Уолша  $(-1 \ 1 \ -1 \ 1 \ -1 \ 1)$  и  $(-1 \ -1 \ 1 \ 1 \ -1 \ -1 \ 1 \ 1)$  соответственно.
  - Найдите выход приемника, если пользователь  $A$  передает информационный бит 1, а пользователь  $B$  не передает данных.
  - Найдите выход приемника, если пользователь  $A$  передает бит 0, а пользователь  $B$  не передает данных.

- в. Найдите выход приемника, если пользователи  $A$  и  $B$  одновременно передают биты 1. Считайте, что мощность сигналов, полученных от  $A$  и  $B$ , одинакова.
- г. Найдите выход приемника, если пользователь  $A$  передает бит 0, а пользователь  $B$  передает бит 1. Считайте, что мощность сигналов, полученных от  $A$  и  $B$ , одинакова.
- д. Найдите выход приемника, если пользователь  $A$  передает бит 1, а пользователь  $B$  передает бит 0. Считайте, что мощность сигналов, полученных от  $A$  и  $B$ , одинакова.
- е. Найдите выход приемника, если пользователи  $A$  и  $B$  передают биты 0. Считайте, что мощность сигналов, полученных от  $A$  и  $B$ , одинакова.
- ж. Найдите выход приемника, если пользователи  $A$  и  $B$  одновременно передают биты 1. Считайте, что мощность сигнала, полученного от  $B$ , в два раза превышает мощность сигнала  $A$ . Разницу мощности сигналов можно представить следующим образом: компоненты сигнала  $A$  записываются как  $(+1, -1)$ , а компоненты сигнала  $B$  — как  $(+2, -2)$ .
- з. Найдите выход приемника, если пользователь  $A$  передает бит 0, а пользователь  $B$  передает бит 1. Считайте, что мощность сигнала, полученного от  $B$ , в два раза превышает мощность сигнала  $A$ .

# ГЛАВА 8

## КОДИРОВАНИЕ И ЗАЩИТА ОТ ОШИБОК

- 8.1. Выявление ошибок
  - Проверка четности
  - Циклическая проверка четности с избыточностью
- 8.2. Блочные коды с коррекцией ошибок
  - Принципы блочных кодов
  - Коды Хэмминга
  - Циклические коды
  - Коды БХЧ
  - Коды Рида-Соломона
  - Чередование блоков
- 8.3. Сверточные коды
  - Декодирование
  - Турбокодирование
- 8.4. Автоматический запрос повторной передачи
  - Управление потоком данных
  - Защита от ошибок
- 8.5. Рекомендуемая литература
- 8.6. Термины, вопросы и задачи



**В** предыдущих главах уже говорилось об искажении сигнала в процессе передачи, а также о зависимости частоты появления ошибок от скорости передачи данных и отношения сигнал/шум. Ошибки будут всегда, независимо от конструкции системы передачи, так что в переданном кадре один или несколько битов обязательно изменят свое значение.

Существуют три наиболее распространенных орудия борьбы с ошибками в процессе передачи данных:

- коды обнаружения ошибок;
- коды с коррекцией ошибок, называемые также схемами прямого исправления ошибок (forward error correction — FEC);
- протоколы с автоматическим запросом повторной передачи (automatic repeat request — ARQ).

Код обнаружения ошибок позволяет довольно легко установить наличие ошибки. Как правило, подобные коды используются совместно с определенными протоколами канального или транспортного уровня (см. рис. 4.4), имеющими схему ARQ. В схеме ARQ приемник попросту отклоняет блок данных, в котором была обнаружена ошибка, после чего передатчик передает этот блок повторно. Коды прямого исправления ошибок позволяют не только обнаружить ошибки, но и исправить их, не прибегая к повторной передаче. Схемы FEC часто используются в беспроводной передаче, где повторная передача крайне неэффективна, а уровень ошибок довольно высок.

Данная глава посвящена изучению трех названных выше методов борьбы с ошибками.

## 8.1. ВЫЯВЛЕНИЕ ОШИБОК

Далее будем считать, что данные передаются как одна или несколько непрерывных последовательностей битов, которые называют кадрами. Определим вероятности, связанные с возникновением ошибок в переданных кадрах.

$P_b$  — вероятность появления единичного ошибочного бита; именуется также частотой появления ошибочных битов (bit error rate — BER).

$P_1$  — вероятность безошибочного приема кадра.

$P_2$  — вероятность того, что используемый алгоритм выявления ошибок не позволяет обнаружить ошибку в кадре.

$P_3$  — вероятность того, что используемый алгоритм выявления ошибок позволяет обнаружить все ошибки в кадре.

Рассмотрим для начала пример, когда при передаче данных схемы выявления ошибок не используются. В этом случае вероятность обнаружения всех ошибок ( $P_3$ ) равна нулю. Чтобы найти значения остальных вероятностей, предположим, что каждый бит может быть ошибочным с равной вероятностью, т.е.  $P_b$  — постоянная независимая величина для каждого бита. Тогда можно записать:

$$P_1 = (1 - P_b)^f,$$

$$P_2 = 1 - P_1,$$

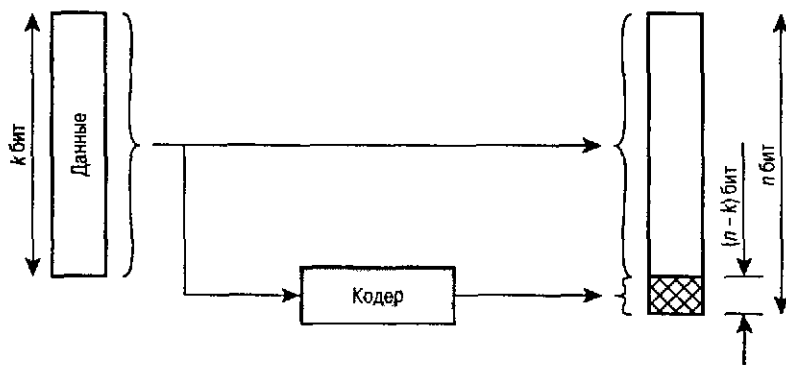
где  $F$  — число битов в кадре. Иными словами, вероятность получения кадра без ошибок уменьшается с ростом вероятности битовой ошибки. Кроме того, вероятность отсутствия ошибок в полученном кадре уменьшается с увеличением длины кадра; чем длиннее кадр, тем больше в нем битов и тем больше вероятность ошибочности одного из них.

Чтобы проиллюстрировать приведенные соотношения, рассмотрим простой пример. Для цифровых сетей с интеграцией услуг (Integrated Services Digital Network — ISDN) стандартной задачей является поддержание частоты появления ошибочных битов в канале 64 Кбит/с ниже  $10^{-6}$ , по крайней мере для 90% интервалов длительностью 1 мин. Теперь предположим, что требования пользователя являются сравнительно скромными — в течение суток в среднем должно появляться не более одного кадра с необнаруженной битовой ошибкой. Длину кадров примем равной 1000 бит. В течение суток может быть передано  $5,529 \times 10^6$  кадров. Вычислим максимально допустимую частоту появления ошибочных кадров.  $P_2 = 1/(5,529 \times 10^6) = 0,18 \times 10^{-6}$ . Для принятого нами значения  $P_b = 10^{-6}$   $P_1 = (0,999999)^{1000} = 0,999$ . Следовательно,  $P_2 = 10^{-3}$ , что на три порядка превышает требования.

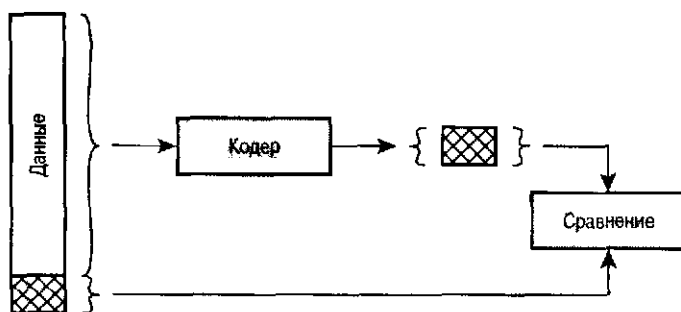
Полученный результат свидетельствует о необходимости применения схем обнаружения ошибок. Работа всех методов обнаружения ошибок основывается на следующем принципе: к информационному кадру передатчиком добавляется последовательность битов, которые составляют код обнаружения ошибок (рис. 8.1). Этот код вычисляется как функция переданных битов. Обычно для информационного блока из  $k$  бит алгоритм обнаружения ошибок дает код, имеющий  $n-k$  бит, причем  $(n-k) < k$ . Код обнаружения ошибок (иногда называемый контрольными битами) присоединяется к блоку данных, в результате чего получается последовательность из  $n$  бит, которая и передается. Приемник разделяет полученную последовательность на  $k$  бит данных и  $(n-k)$  бит кода обнаружения ошибок. Основываясь на битах данных, приемник вычисляет код, после чего сверяет результат с принятым кодом обнаружения ошибок. Если два кода не совпадают, имеется ошибка. Следовательно, параметр  $P_3$  — это вероятность того, что в кадре присутствует ошибка и она обнаружена с помощью используемой схемы. Параметр  $P_2$  называют остаточным уровнем ошибок.  $P_2$  — это вероятность того, что ошибка не будет обнаружена, несмотря на использование схемы выявления ошибок.

## Проверка четности

Наиболее простой метод обнаружения ошибок — добавление бита четности в конец каждого блока данных. Типичный пример: передача знаков, во время которой бит четности добавляется к каждому 7-битовому знаку. Значение этого контрольного бита выбирается так, чтобы общее число единиц в знаке было или нечетным (отрицательная четность), или четным (положительная четность). Например, если передатчик передает 1110001, используя схему с отрицательной четностью, в конце последовательности будет добавлена двоичная единица. В итоге передана будет последовательность 11100011. После получения знака приемник проверяет количество двоичных единиц, и если оно нечетное, считается, что ошибок нет. Если же в процессе передачи один бит (или любое нечетное их количество) изменил значение — будет обнаружено наличие ошибки. Необходимо отметить, что если свое значение изменят два бита (или любое четное их количество), ошибка обнаружена не будет. Как правило, положительная четность используется при синхронной передаче данных, а отрицательная — при асинхронной.



а) Отправитель



б) Приемник

Рис. 8.1. Процесс обнаружения ошибок

Проверка четности не является надежным средством обнаружения ошибок, поскольку шумовые импульсы обычно довольно продолжительны (в особенности при высоких скоростях передачи данных), чтобы изменить более одного бита.

### Циклическая проверка четности с избыточностью

Циклическая проверка четности с избыточностью (cyclic redundancy check — CRC) — это один из наиболее широко используемых и надежных методов обнаружения ошибок. Принцип работы данного метода сводится к следующему: для блока из  $k$  бит (сообщения) передатчик генерирует так называемую контрольную последовательность кадра (frame check sequence — FCS) из  $(n - k)$  бит. При этом результирующая последовательность (состоящая из данных и FCS) должна делиться без остатка на заданную константу. После передачи приемник делит полученную последовательность на эту константу и, если деление не дало остатка, считает, что ошибки в процессе передачи отсутствовали<sup>1</sup>.

<sup>1</sup> Данная процедура несколько отличается от той, что приведена на рис. 8.1. Как будет показано далее, приемник может использовать для деления  $k$  полученных информационных битов, а результат сравнивать с  $(n - k)$  контрольными битами.

Для лучшего понимания этой процедуры мы опишем ее тремя способами: через арифметические действия по модулю 2, действия с полиномами и цифровую логику.

### Арифметика по модулю 2

В вычислениях по модулю 2 используется двоичное сложение без переноса в старший разряд (операция исключающего ИЛИ). Двоичное вычитание без переноса также интерпретируется как операция исключающего ИЛИ. Проиллюстрируем сказанное:

$$\begin{array}{r}
 1111 \quad 1111 \quad 11001 \\
 + 1010 \quad - 0101 \quad \times 11 \\
 \hline
 0101 \quad 1010 \quad 11001 \\
 \phantom{0101 \quad 1010} \quad \underline{110010} \\
 \phantom{0101 \quad 1010} \quad 101011
 \end{array}$$

Определим следующие параметры:

$T$  —  $n$ -битовый кадр, который необходимо передать;

$D$  —  $k$ -битовый блок данных (сообщение); первые  $k$  бит кадра  $T$ ;

$F$  —  $(n - k)$ -битовая контрольная последовательность кадра; последние  $(n - k)$  бит кадра  $T$ ;

$P$  —  $(n - k + 1)$ -битовый предопределенный делитель.

Нам нужно, чтобы  $T$  нацело делилось на  $P$ . Очевидно, что

$$T = 2^{n-k}D + F.$$

Т.е. умножая  $D$  на  $2^{n-k}$ , мы фактически сдвигаем его влево на  $(n - k)$  бит, и заполняем оставшееся место нулями, а полученную последовательность дополняем нулями. После добавления  $F$  получаем конкатенацию  $D$  и  $F$ , т.е.  $T$ . Необходимо, чтобы  $P$  было делителем  $T$ . Рассмотрим деление  $2^{n-k}D$  на  $P$ :

$$\frac{2^{n-k}D}{P} = Q + \frac{R}{P}. \quad (8.1)$$

Имеем частное и некоторый остаток. Поскольку деление производилось по модулю 2, остаток должен быть по крайней мере на один бит короче делителя. Используем полученный остаток в качестве контрольной последовательности кадра. Тогда

$$T = 2^{n-k}D + R. \quad (8.2)$$

Удовлетворяет ли  $R$  условию отсутствия остатка при делении  $T$  на  $P$ ? Для того чтобы проверить это, запишем:

$$\frac{T}{P} = \frac{2^{n-k}D + R}{P} = \frac{2^{n-k}D}{P} + \frac{R}{P}.$$

Подставив уравнение (8.1), получим:

$$\frac{T}{P} = Q + \frac{R}{P} + \frac{R}{P}.$$

Поскольку сумма по модулю 2 равных двоичных чисел дает в результате 0, можем записать:

$$\frac{T}{P} = Q + \frac{R+R}{P} = Q.$$

Поскольку в последнем выражении остатка нет, мы доказали, что  $T$  нацело делится на  $P$ . Итак, мы получили простой способ создания контрольной последовательности кадра:  $2^{n-k}D$  делится на  $P$ , после чего  $(n-k)$ -битовый остаток используется как FCS. После получения сигнала приемник делит  $T$  на  $P$ . Если остаток равен нулю, ошибки отсутствуют.

Рассмотрим простой пример.

**Пример.**

1. Дано:

сообщение  $D$  — 1010001101 (10 бит);

последовательность  $P$  — 110101 (6 бит);

5-битовая контрольная последовательность кадра  $R$  — необходимо вычислить.

Итак,  $n = 15, k = 10, (n - k) = 5$ .

2. Сообщение умножается на  $2^5$ , что дает в результате 101000110100000.

3. Полученная последовательность делится на  $P$ :

$$\begin{array}{r}
 2^{n-k}D \longrightarrow 101000110100000 \\
 \underline{110101} \\
 111011 \\
 \underline{110101} \\
 111010 \\
 \underline{110101} \\
 111110 \\
 \underline{110101} \\
 101100 \\
 \underline{110101} \\
 110010 \\
 \underline{110101} \\
 01110 \longleftarrow R
 \end{array}
 \qquad
 \begin{array}{r}
 110101 \longleftarrow P \\
 \hline
 1101010110 \longleftarrow Q
 \end{array}$$

4. Полученный остаток складывается с  $2^5D$ , результат ( $T = 101000110101110$ ) передается.

5. При отсутствии ошибок принятая последовательность  $T$  не отличается от переданной. Приемник выполняет деление на  $P$ :

$$\begin{array}{r}
 T \rightarrow \begin{array}{r} 101000110101110 \\ \underline{110101} \\ 111011 \\ \underline{110101} \\ 111010 \\ \underline{110101} \\ 111110 \\ \underline{110101} \\ 101111 \\ \underline{110101} \\ 110101 \\ \underline{110101} \\ 0 \end{array} \\
 \end{array}
 \qquad
 \begin{array}{r}
 \begin{array}{|l} 110101 \\ \hline 1101010110 \end{array} \leftarrow P \\
 \leftarrow Q
 \end{array}$$

Поскольку деление не дало остатка, считается, что ошибки отсутствуют.

Последовательность  $P$  на один бит длиннее контрольной последовательности кадра, а ее точный вид зависит от ожидаемого типа ошибок. Как минимум старший и младший биты  $P$  должны быть единицами.

Существует простой способ определения наличия одной или нескольких ошибок. Ошибка — это изменение значения бита, что эквивалентно применению операции исключающего ИЛИ к данному биту и двоичной единице (добавлению к биту единицы по модулю 2):  $0 + 1 = 1$ ;  $1 + 1 = 0$ . Таким образом, ошибки в  $n$ -битовом кадре можно представить как  $n$ -битовое поле с единицами в каждом ошибочном разряде. Полученный таким образом кадр  $T_r$  можно записать в следующем виде:

$$T_r = T \oplus E,$$

где

$T$  — переданный кадр;

$E$  — карта ошибок с единицами в местах появления ошибок;

$T_r$  — полученный кадр.

При наличии ошибки ( $E \neq 0$ ) приемник не сможет ее обнаружить тогда и только тогда, когда  $T_r$  делится на  $P$  (или, эквивалентно,  $E$  делится на  $P$ ). Интуитивно можно предположить, что вероятность такого события мала.

### Полиномы

Процесс циклической проверки четности с избыточностью можно также рассмотреть, представив все значения в форме полиномов фиктивной переменной  $X$  с двоичными коэффициентами, которые соответствуют битам двоичного числа. Как и выше, арифметические операции выполняются по модулю 2. Процесс проверки можно описать следующим образом:

$$\frac{X^{n-k} D(X)}{P(X)} = Q(X) + \frac{R(X)}{P(X)},$$

$$T(X) = X^{n-k} D(X) + R(X).$$

Сравните приведенные уравнения с (8.1) и (8.2).

**Пример.** Используя предыдущий пример, получаем: последовательности  $D = 1010001101$  соответствует полином  $D(X) = X^9 + X^7 + X^3 + X^2 + 1$ , последовательности  $P = 110101$  — полином  $P(X) = X^5 + X^4 + X^2 + 1$ , последовательности  $R = 01110$  —  $R(X) = X^3 + X^2 + X$ . Полиномиальное деление, соответствующее приведенному ранее двоичному делению, показано на рис. 8.2.

$$\begin{array}{r}
 2^3 D(X) \longrightarrow \begin{array}{r}
 X^{14} + X^{12} + X^8 + X^7 + X^5 \\
 \underline{X^{14} + X^{13} + X^{11} + X^9} \\
 X^{13} + X^{12} + X^{11} + X^9 + X^8 \\
 \underline{X^{13} + X^{12} + X^{10} + X^8} \\
 X^{11} + X^{10} + X^9 + X^7 \\
 \underline{X^{11} + X^{10} + X^8 + X^6} \\
 X^9 + X^8 + X^7 + X^6 + X^5 \\
 \underline{X^9 + X^8 + X^6 + X^4} \\
 X^7 + X^5 + X^4 \\
 \underline{X^7 + X^6 + X^4 + X^2} \\
 X^6 + X^5 + X^2 \\
 \underline{X^6 + X^5 + X^3 + X} \\
 X^3 + X^2 + X \longleftarrow R(X)
 \end{array} \\
 \left. \begin{array}{l}
 X^5 + X^4 + X^2 + 1 \longleftarrow P(X) \\
 X^9 + X^8 + X^6 + X^4 + X^2 + X \longleftarrow Q(X)
 \end{array} \right\}
 \end{array}$$

Рис. 8.2. Пример полиномиального деления

Ошибка не обнаружится только в том случае, если соответствующий полином  $E(X)$  делится на  $P(X)$ . Можно показать [РЕТЕ61, РАМА88], что при надлежащем выборе полинома  $P(X)$  выявляются такие ошибки:

- все 1-битовые ошибки, если  $P(X)$  имеет более одного ненулевого члена;
- все 2-битовые ошибки, если  $P(X)$  имеет делитель из трех членов;
- любое нечетное количество ошибок, если в разложении  $P(X)$  по множителям присутствует  $(X + 1)$ ;
- любой пакет ошибок<sup>2</sup>, длина которого не превышает  $n - k$ , или, эквивалентно, не превышает длину контрольной последовательности кадра;
- часть пакета ошибок длиной  $n - k + 1$ ; часть равна  $1 - 2^{-(n-k-1)}$ ;
- часть пакета ошибок длиной более  $n - k + 1$ ; часть равна  $1 - 2^{-(n-k)}$ .

Кроме того, можно показать, что если все последовательности ошибок считать равновероятными, то для пакета ошибок длиной  $r + 1$  вероятность появления необнаруженной ошибки (т.е. вероятность того, что  $E(X)$  делится на  $P(X)$ ) равна  $1/2^{r-1}$ . Для пакета ошибок большей длины вероятность появления необнаруженной ошибки равна  $1/2^r$ , где  $r$  — длина контрольной последовательности кадра.

Наиболее широко используются четыре полинома  $P(X)$ :

<sup>2</sup> Пакет ошибок (burst error) длиной  $B$  — это непрерывная последовательность из  $B$  ошибочных битов.

CRC-12	$X^{12} + X^{11} + X^3 + X^2 + X + 1$
CRC-16	$X^{16} + X^{15} + X^2 + 1$
CRC-CCITT	$X^{16} + X^{12} + X^5 + 1$
CRC-32	$X^{32} + X^{26} + X^{23} + X^{22} + X^{16} + X^{12} + X^{11} + X^{10} + X^8 + X^7 + X^5 + X^4 + X^2 + X + 1$

Система CRC-12 используется для передачи потоков 6-битовых знаков; здесь генерируются 12-битовые FCS. Системы CRC-16 и CRC-CCITT применяются для передачи 8-битовых знаков в США и Европе, соответственно; здесь создаются контрольные последовательности длиной 16 бит. В большинстве случаев этих полиномов достаточно, хотя иногда (например, в стандартах двухточечной синхронной передачи) как вариант указан полином CRC-32.

### Цифровая логика

Процесс циклической проверки четности с избыточностью можно представить (и, как правило, он так и реализуется) как схему деления, состоящую из элемента исключающего ИЛИ и регистра сдвига. Регистр сдвига представляет собой строку 1-битовых ячеек памяти. Каждая ячейка имеет выходную шину, показывающую текущее хранимое значение, и входную шину. В дискретные моменты времени, которые называются тактами, значения ячеек памяти замещаются значениями, указанным во входной шине. Замена происходит синхронно во всем регистре, так что в результате значения ячеек регистра сдвигаются на один бит.

Реализация схемы выглядит следующим образом.

1. Регистр, содержащий  $(n - k)$  бит (по размеру контрольной последовательности кадра).
2. До  $(n - k)$  элементов исключающего ИЛИ.
3. Наличие или отсутствие логического элемента соответствует наличию или отсутствию члена в полиноме-делителе  $P(X)$ , исключая члены 1 и  $X^{n-k}$ .

**Пример.** Понять архитектуру схемы деления можно с помощью примера, приведенного на рис. 8.3. В примере, как и ранее, используются такие величины:

$$\text{сообщение } D = 1010001101; D(X) = X^9 + X^7 + X^3 + X^2 + 1;$$

$$\text{делитель } P = 110101; P(X) = X^5 + X^4 + X^2 + 1.$$

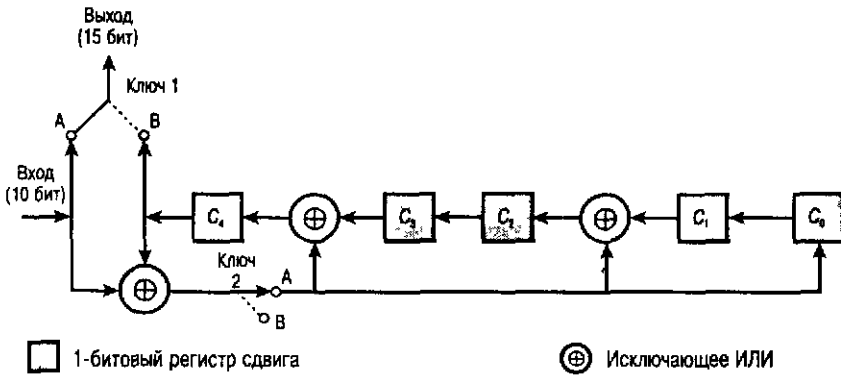
На рис. 8.3, а представлена реализация регистра сдвига. Процесс начинается с очистки регистра (все ячейки обнуляются). После этого передаваемое сообщение (делимое) побитово вводится в регистр, начиная со старшего бита. На рис. 8.3, б приведена таблица, которая иллюстрирует пошаговую работу схемы по мере введения отдельных битов. Строки таблицы содержат значения пяти ячеек регистра сдвига в соответствующие моменты времени. Кроме того, в строках таблицы приводятся значения на выходе трех схем исключающего ИЛИ. Последнее число в каждой строке — значение следующего входного бита, который станет доступен для работы на следующем этапе.

Отметим, что операция исключающего ИЛИ влияет на значения ячеек  $C_4$ ,  $C_2$  и  $C_0$  при следующем сдвиге, что идентично рассмотренному ранее процессу двоичного деления. Процесс выполняется для всех битов передаваемого сообщения. Для обеспечения корректности выходного сигнала используются два ключа. При вводе битов данных оба ключа находятся в положении А. В результате за первых 10 ша-



гов входные биты подаются в регистр сдвига и также используются в качестве выходных битов. По окончании обработки последнего бита данных регистр сдвига содержит остаток деления (FCS), выделенный серым цветом. При вводе последнего бита данных в регистр оба ключа устанавливаются в положение В. В этом случае: 1) все логические элементы больше не изменяют значения битов; 2) за следующие 5 шагов на выход подаются 5 бит CRC.

В приемнике используется аналогичная логика. Принятые биты последовательности  $M$  вводятся в регистр сдвига по мере поступления. Если ошибки отсутствуют, то после обработки  $M$  регистр сдвига будет содержать последовательность  $R$ . После этого начинает поступать переданная последовательность  $R$ . В результате на выход регистра будут подаваться двоичные нули и по завершении приема все ячейки будут иметь значение 0.



а) Реализация регистра сдвига

Исходное состояние	$C_4$	$C_3$	$C_2$	$C_1$	$C_0$	$C_4 \oplus C_3 \oplus I$	$C_4 \oplus C_1 \oplus I$	$C_4 \oplus I$	$I = \text{ввод}$
Этап 1	1	0	1	0	1	1	1	1	0
Этап 2	1	1	1	1	1	1	1	0	1
Этап 3	1	1	1	1	0	0	0	1	0
Этап 4	0	1	0	0	1	1	0	0	0
Этап 5	1	0	0	1	0	1	0	1	0
Этап 6	1	0	0	0	1	0	0	0	1
Этап 7	0	0	0	1	0	1	0	1	1
Этап 8	1	0	0	0	1	1	1	1	0
Этап 9	1	0	1	1	1	0	1	0	1
Этап 10	0	0	0	0	0				

б) Пример для последовательности битов 1010001101

Рис. 8.3. Схема с регистром сдвига для деления на полином  $X^5 + X^4 + X^2 + 1$

На рис. 8.4 представлена общая архитектура реализации регистра сдвига CRC для полинома  $P(X) = \sum_{i=0}^{n-k} A_i X^i$ , где  $A_0 = A_{n-k} = 1$ , все остальные  $A_i$  равны 0 или 1<sup>3</sup>.

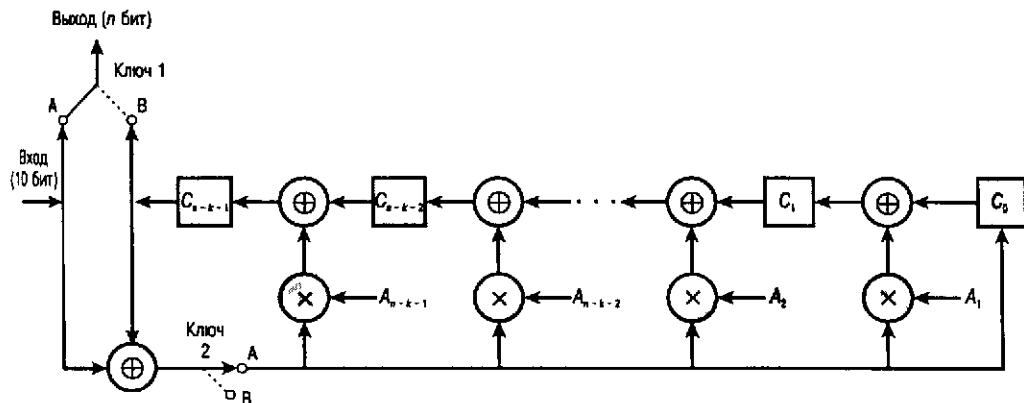


Рис. 8.4. Общая архитектура CRC для реализации полинома-делителя  $1 + A_1X + A_2X^2 + \dots + A_{n-1}X^{n-k-1} + X^{n-k}$

## 8.2. БЛОЧНЫЕ КОДЫ С КОРРЕКЦИЕЙ ОШИБОК

Методы обнаружения ошибок широко применяются на практике в протоколах управления каналами передачи данных, таких, как HDLC, а также в транспортных протоколах, таких, как TCP. В то же время под использованием кодов обнаружения ошибок подразумевается повторная передача блоков данных согласно процедуре ARQ (этот метод подробно рассматривается в разделе 8.4). Для беспроводных приложений такой подход неприемлем по двум причинам.

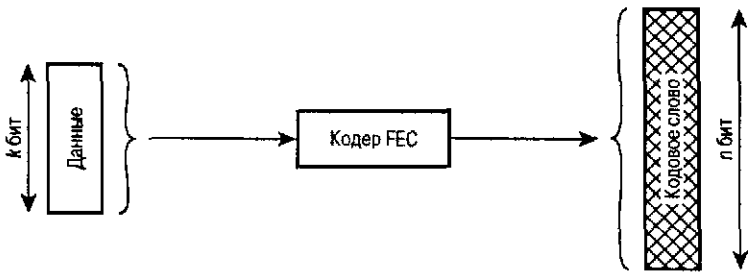
1. Уровень ошибок в беспроводном канале может быть довольно высок; в результате потребуется значительное число повторных передач.
2. В некоторых случаях, особенно в спутниковой связи, задержка распространения сигнала довольно велика по сравнению с временем передачи одного кадра. Как показывается в разделе 8.4, при повторной передаче, как правило, передается кадр, содержащий ошибку, а также все последующие кадры. При большом расстоянии между приемником и передатчиком ошибка в одном кадре приводит к необходимости повторной передачи множества кадров.

Вместо повторной передачи было бы лучше, если бы приемник мог исправлять ошибки в полученном сигнале, используя информацию, содержащуюся в самом сигнале. На рис. 8.5 представлена схема реализации этой идеи. С помощью кодера FEC

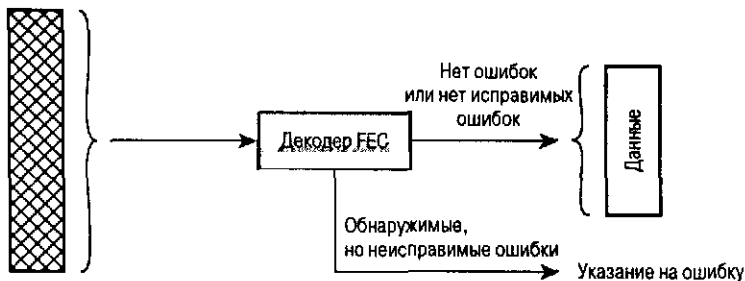
<sup>3</sup> Часто сдвиг в регистре CRC производится слева направо, т.е. в направлении, противоположном используемому при аналогичном двоичном делении. Поскольку старший бит двоичных чисел обычно записывается слева, более уместным кажется все же использование регистра со сдвигом справа налево.

(forward error correction — прямое исправление ошибок) передатчик преобразует каждый  $k$ -битовый блок данных в  $n$ -битовый блок ( $n > k$ ), именуемый кодовым словом, который затем передается (в беспроводной связи для передачи используется созданный модулятором аналоговый сигнал). При распространении сигнал может подвергаться воздействию шума, что может привести к появлению ошибочных битов. Приемник демодулирует полученный сигнал, преобразовывая его в строку битов, подобную переданной, но, возможно, с ошибками. Полученный блок данных обрабатывается декодером FEC, в результате возможны такие ситуации.

1. При отсутствии ошибочных битов вход декодера FEC идентичен исходному кодовому слову, так что на выход декодера поступает исходный блок данных.
2. Декодер может обнаружить и исправить определенные последовательности ошибок. Поэтому даже если вход декодера отличается от исходного кодового слова, из него можно получить исходные данные.
3. Некоторые последовательности ошибок могут быть обнаружены декодером, но не могут быть исправлены. В этом случае декодер сообщает о наличии неисправимой ошибки.
4. Наличие некоторых (обычно довольно редких) последовательностей ошибок не может быть обнаружено декодером. В результате декодер преобразовывает входной  $n$ -битовый блок в  $k$ -битовую последовательность, которая отличается от переданной, но которую кодер считает правильной.



а) Отправитель



б) Приемник

Рис. 8.5. Прямое исправление ошибок

Каким образом декодер исправляет ошибки? По сути, исправление ошибок производится с помощью добавления избыточных данных к передаваемому со-

общению. Избыточность позволяет приемнику восстановить исходное сообщение даже при наличии определенного уровня ошибок. В данном разделе будет рассмотрен широко используемый класс кодов с коррекцией ошибок, известный как блочные коды. В начале будут рассмотрены общие принципы, после чего мы перейдем к анализу конкретных кодов.

Отметим, что довольно часто коды с коррекцией ошибок используются по схеме, изображенной на рис. 8.1 для кодов обнаружения ошибок. Т.е. в алгоритме FEC к входному  $k$ -битовому блоку данных добавляется  $(n-k)$  контрольных битов; в результате размер передаваемого блока составляет  $n$  бит; все биты исходного  $k$ -битового блока содержатся в полученном  $n$ -битовом блоке. Для некоторых схем прямого исправления ошибок (например, сверточных кодов, описанных в разделе 8.3) входная  $k$ -битовая последовательность так преобразовывается в  $n$ -битовое кодовое слово, что исходные  $k$  бит не фигурируют явно в кодовом слове.

## Принципы блочных кодов

Для начала определим термины, которыми мы будем оперировать. Расстоянием Хэмминга (Hamming distance)  $d(v_1, v_2)$  между двумя  $n$ -битовыми двоичными последовательностями  $v_1$  и  $v_2$  называют число несовпадающих разрядов  $v_1$  и  $v_2$ . Например, если

$$v_1 = 011011, v_2 = 110001,$$

то

$$d(v_1, v_2) = 3.$$

Рассмотрим теперь метод блочного кодирования с целью коррекции ошибок. Пусть требуется передать определенное количество  $k$ -битовых блоков данных. Вместо передачи каждого блока как последовательности  $k$  бит, преобразуем каждую  $k$ -битовую последовательность в уникальное  $n$ -битовое кодовое слово.

**Пример.** Для  $k = 2$  и  $n = 5$  имеем следующее присваивание:

Блок данных	Кодовое слово
00	00000
01	00111
10	11001
11	11110

Предположим, что кодовое слово было получено в виде последовательности битов 00100. Поскольку эта последовательность не соответствует ни одному из кодовых слов, приемник обнаружил ошибку. Как можно ее исправить? Точно узнать, какой из блоков данных был передан, невозможно, поскольку шум мог изменить 1, 2, 3, 4 или даже все 5 переданных битов. Отметим, впрочем, что для преобразования приемлемого кодового слова 00000 в полученную последовательность достаточно изменения одного бита. Соответственно, для преобразования 00111 в 00100 нужно изменить два бита; 11110 в 00100 — три бита; 11001 в 00100 — четыре бита. Таким образом, можно сделать вывод, что с наибольшей вероятностью был передан блок 00000, т.е. искомый блок данных — 00. Подобное рассуждение и есть логика коррекции ошибки. Используя понятие “расстояние Хэмминга”, его можно представить в следующем виде:

$d(00000, 00100) = 1$ ;  $d(00111, 00100) = 2$ ;  $d(11001, 00100) = 4$ ;  $d(11110, 00100) = 3$ .

Таким образом, искомое правило коррекции ошибок можно сформулировать так: если получено кодовое слово с ошибкой, его истинное значение принимается равным кодовому слову, которое находится на минимальном расстоянии Хэмминга от искаженного слова. Это правило применимо только тогда, когда для каждого ошибочного кодового слова имеется только одно реальное кодовое слово, находящееся от него на минимальном расстоянии Хэмминга.

Для нашего примера приведенное условие не выполняется. Полное количество возможных последовательностей равно  $2^5 = 32$ , из которых 4 используются как кодовые слова, а 28 являются ошибочными. Для ошибочных кодовых слов можно составить таблицу.

Ошибочное кодовое слово	Минимальное расстояние	Существующее кодовое слово	Ошибочное кодовое слово	Минимальное расстояние	Существующее кодовое слово
00001	1	00000	10000	1	00000
00010	1	00000	10001	1	11001
00011	1	00111	10010	2	00000 или 11110
00100	1	00000	10011	2	00111 или 11001
00101	1	00111	10100	2	00000 или 11110
00110	1	00111	10101	2	00111 или 11001
01000	1	00000	10110	1	11110
01001	1	11001	10111	1	00111
01010	2	00000 или 11110	11000	1	11001
01011	2	00111 или 11001	11010	1	11110
01100	2	00000 или 11110	11011	1	11001
01101	2	00111 или 11001	11100	1	11110
01110	1	11110	11101	1	11001
01111	1	00111	11111	1	11110

В 8 случаях ошибочная последовательность находится на расстоянии 2 от двух различных существующих кодовых слов. Любая такая последовательность может возникнуть в результате двух битовых ошибок, и при этом приемник не может однозначно выбрать правильное кодовое слово. Таким образом, ошибка обнаруживается, но не исправляется. В то же время для всех 1-битовых ошибок полученная последовательность находится на расстоянии 1 только от одного правильного кодового слова, что позволяет принять правильное решение. Следовательно, данный код можно использовать для исправления всех 1-битовых ошибок, но в то же время он не позволяет исправлять ошибки в двух битах. Для иллюстрации сказанного рассмотрим расстояния между корректными кодовыми словами:

$$d(00000, 00111) = 3; \quad d(00000, 11001) = 3; \quad d(00000, 11110) = 4;$$

$$d(00111, 11001) = 4; \quad d(00111, 11110) = 3; \quad d(11001, 11110) = 3.$$

Минимальное расстояние между двумя правильными кодовыми словами равно 3. Следовательно, при появлении 1-битовой ошибки расстояние между исходной и

ошибочной последовательностью составит 1; а расстояние между ошибочной последовательностью и всеми прочими правильными кодовыми словами будет не менее 2. Таким образом, рассмотренный код позволяет всегда исправлять 1-битовые ошибки. Кроме того, будут обнаружены все 2-битовые ошибки.

Приведенные ранее примеры иллюстрируют важные свойства блочных кодов с коррекцией ошибок. Блочный код  $(n, k)$  преобразует  $k$  бит данных в  $n$ -битовые кодовые слова. В большинстве случаев каждое корректное кодовое слово — это исходные  $k$  бит данных плюс  $(n - k)$  контрольных битов. Таким образом, структура блочного кода эквивалентна структуре функции вида  $v_c = f(v_d)$ , где  $v_d$  — вектор, состоящий из  $k$  бит данных;  $v_c$  — вектор, состоящий из  $n$  бит кодового слова.

В блочном коде  $(n, k)$  имеется  $2^k$  приемлемых кодовых слов из  $2^n$  возможных. Отношение числа избыточных битов к числу битов данных  $(n - k)/k$  принято называть избыточностью кода (code redundancy); отношение количества битов данных к полному числу битов  $(k/n)$  называют степенью кодирования (code rate). Степень кодирования — это мера того, какая дополнительная полоса потребуется после кодирования, если скорость передачи данных мы хотим сохранить такой же, какая была до кодирования. Например, при степени кодирования  $1/2$  для сохранения скорости передачи системе потребуется полоса, в два раза большая, чем та, которая необходима для передачи некодированного сигнала. Для рассмотренного выше примера степень кодирования равна  $2/5$ , следовательно, для сохранения скорости передачи потребуется увеличить ширину полосы в  $2,5$  раза. Т.е. если скорость передачи сигнала на входе кодера равна  $1$  Мбит/с, то для сохранения прежних параметров скорость выходного сигнала должна быть равна  $2,5$  Мбит/с.

Для кода, состоящего из кодовых слов  $w_1, w_2, \dots, w_s$ , где  $s = 2^k$ , минимальное расстояние кода  $d_{\min}$  определяется следующим образом:

$$d_{\min} = \min_{i \neq j} [d(w_i, w_j)].$$

Можно показать, что если для кода выполняется неравенство  $d_{\min} \geq 2t + 1$  ( $t$  — некоторое положительное целое число), то с помощью данного кода можно исправить все символы, содержащие до  $t$  ошибочных битов, включительно. Если  $d_{\min} \geq 2t$ , то можно исправить все символы, содержащие до  $(t - 1)$  ошибочных битов. Кроме того, будут обнаружены все символы с  $t$  ошибочными битами, исправить которые, в общем случае, нельзя. Верно и обратное утверждение — каждый код, позволяющий исправлять до  $t$  ошибочных битов, должен удовлетворять условию  $d_{\min} \geq 2t + 1$ . Для каждого кода, позволяющего исправлять до  $(t - 1)$  ошибочных битов и обнаруживать все символы с  $t$  ошибками, должно выполняться условие  $d_{\min} \geq 2t$ .

Связь между  $d_{\min}$  и  $t$  можно записать другим способом, выразив максимальное количество битовых ошибок в кодовом слове, гарантированно исправляемых кодом, в таком виде:

$$t = \left\lfloor \frac{d_{\min} - 1}{2} \right\rfloor,$$

где  $\lfloor x \rfloor$  — наибольшее целое число, не превышающее  $x$  (например,  $\lfloor 6,3 \rfloor = 6$ ). Более того, если нас интересует только обнаружение ошибок, но не их исправление, количество обнаруживаемых ошибок  $t$  удовлетворяет следующему равенству:

$$t = d_{\min} - 1.$$

Последнее выражение можно понять интуитивно, если вспомнить, что  $d_{\min}$  битовых ошибок может изменить одно корректное кодовое слово на другое. При любом меньшем количестве ошибок этого произойти не может.

Выбирая блочный код, нужно учитывать следующие соображения.

1. При данных  $n$  и  $k$  предпочтительным является максимально возможное  $d_{\min}$ .
2. Код должен быть сравнительно простым для кодирования и декодирования, требуя минимальной памяти и времени обработки.
3. Для уменьшения ширины требуемой полосы число избыточных битов  $(n - k)$  должно быть небольшим.
4. Для снижения уровня ошибок число избыточных битов  $(n - k)$  должно быть большим.

Очевидно, что два последних требования противоречат друг другу, так что нужно выбирать некоторый компромиссный вариант.

Перед тем как перейти к рассмотрению примеров кодов, обратимся к рис. 8.6, который основан на иллюстрации из [ЛЕВО98]. Графики подобного типа довольно часто приводятся в литературе для демонстрации эффективности разных схем кодирования. Напомним (см. главу 6), что кодирование может использоваться для снижения необходимого значения  $E_b/N_0$ , что позволяет достичь заданного уровня битовых ошибок.<sup>4</sup> Методы кодирования (модуляции), рассмотренные в главе 6, основываются на определении сигнальных посылок, представляющих биты. Кодирование, рассматриваемое в данной главе влияет на отношение  $E_b/N_0$ . На рис. 8.6 кривая справа соответствует системе модуляции без кодирования; в затененной области параметры системы могут быть улучшены. В этой области при заданном  $E_b/N_0$  достигается меньшее значение BER. Верно и обратное утверждение — при заданном BER требуется меньшее значение  $E_b/N_0$ . Кривая, расположенная левее, представляет собой типичный результат обработки сигнала кодом со степенью кодирования  $1/2$  (в этом случае число битов данных равно числу контрольных битов). Для частоты возникновения ошибок  $10^{-5}$  использование кодирования позволяет снизить  $E_b/N_0$  на 2,77 дБ. Это улучшение называется эффективностью кодирования. Эффективность кодирования — это снижение необходимого уровня  $E_b/N_0$  для системы с кодированием по сравнению с системой без кодирования (подразумевается одна и та же модуляция) для достижения заданной частоты ошибок.

Необходимо отметить, что частота появления ошибок при коде со степенью  $1/2$  — это частота появления неисправленных ошибок, а  $E_b$  — энергия на один бит данных. Поскольку степень кодирования равна  $1/2$ , каждому биту данных соответствуют два канальных бита, следовательно удельная энергия кодированного бита в два раза меньше удельной энергии бита данных, т.е. разница  $E_b/N_0$  составляет 3 дБ. Если рассмотреть удельную энергию кодированного бита такой системы, уровень битовых ошибок в канале составит  $2,4 \times 10^{-2}$ , или 0,024.

В завершение необходимо отметить, что если значение  $E_b/N_0$  — ниже определенного порогового, методы кодирования отрицательно сказываются на произ-

---

<sup>4</sup>  $E_b/N_0$  — отношение энергии сигнала на один бит к плотности мощности шума на один герц; определение и подробный анализ данного параметра приведены в главе 5.

водительности системы. В примере, приведенном на рис. 8.6, это пороговое значение равно приблизительно 5,4 дБ. При более низком значении  $E_b/N_0$  использование дополнительных контрольных битов снижает удельную энергию битов данных, что ведет к увеличению числа ошибок. Если значение  $E_b/N_0$  выше порогового, способность кода исправлять ошибки (эффективность кодирования) позволяет компенсировать снижение  $E_b$  и улучшить работу системы.

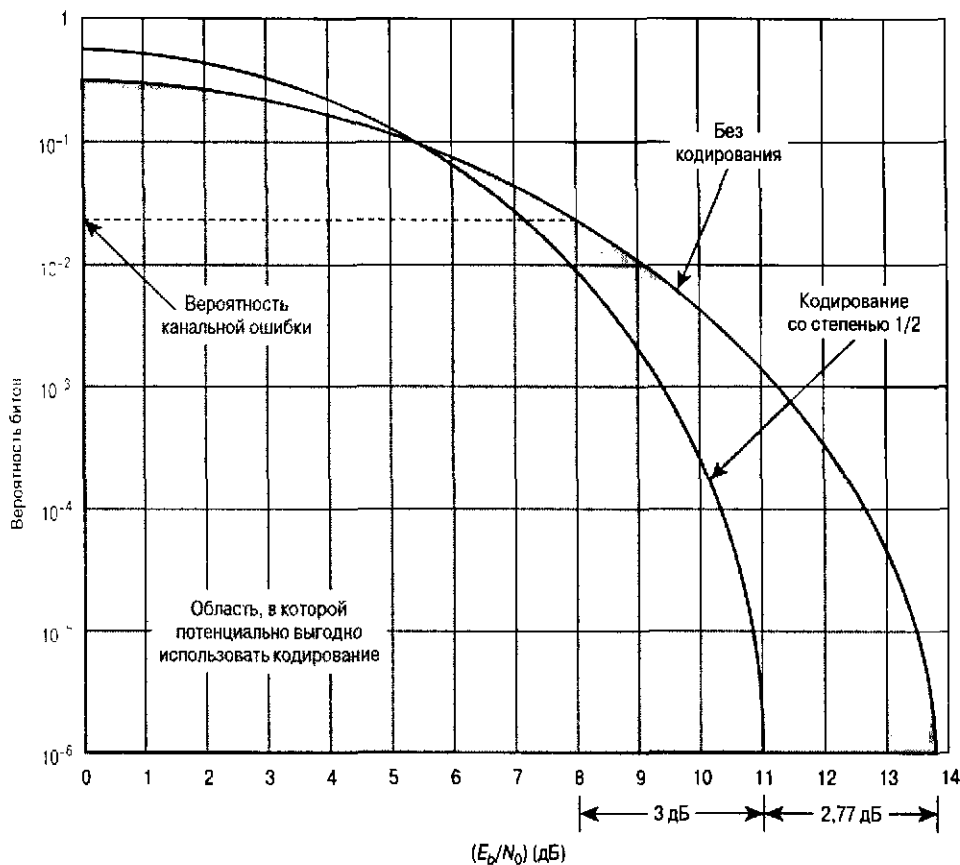


Рис. 8.6. Улучшение работы системы с помощью кодирования

Перейдем к рассмотрению конкретных примеров кодов с коррекцией ошибок.

### Коды Хэмминга

Коды Хэмминга — это семейство блочных кодов с коррекцией ошибок ( $n, k$ ), которые характеризуются следующими параметрами

Длина блока	$n = 2^m - 1$
Количество битов данных	$k = 2^m - m - 1$
Количество контрольных битов	$n - k = m$
Минимальное расстояние	$d_{\min} = 3$



Здесь  $m \geq 3$ . Коды Хэмминга просты в использовании и легко поддаются анализу, однако редко используются на практике. Рассмотрим этот тип кодов, чтобы проиллюстрировать на их примере некоторые фундаментальные принципы работы блочных кодов.

Коды Хэмминга созданы для исправления 1-битовых ошибок. Для начала определим необходимую длину кода. Коды Хэмминга применяются так же, как методы определения ошибок (см. рис. 8.1) — в процессе кодирования сохраняются  $k$  бит данных и добавляются  $(n - k)$  контрольных битов. При декодировании используются две последовательности из  $(n - k)$  бит, одна из которых является кодовым словом входящего сигнала, а другая рассчитывается на основе полученных битов данных. Последовательности побитово сравниваются с помощью логического исключающего ИЛИ. Результат сравнения называют синдромом. Биту синдрома присваивается значение 0, если биты двух последовательностей совпадают, и 1 — в противном случае.

Синдром — это  $(n - k)$ -битовое слово в диапазоне от 0 до  $2^{(n - k)} - 1$ . Значение 0 свидетельствует об отсутствии ошибки. Если же ошибка присутствует, ее местоположение определяется из синдрома. Поскольку ошибочным может быть любой из  $k$  бит данных или  $(n - k)$  проверочных битов, должно выполняться следующее соотношение:

$$2^{(n - k)} - 1 \geq k + (n - k) = n.$$

Приведенное уравнение определяет количество битов, необходимое для исправления 1-битовой ошибки в слове, содержащем  $k$  бит данных. В табл. 8.1 приводится число контрольных битов, необходимое для последовательностей данных разной длины.

Таблица 8.1. Требования к кодам Хэмминга

Количество битов данных	Исправление 1-битовых ошибок		Исправление 1-битовых ошибок; обнаружение 2-битовых ошибок	
	Количество контрольных битов	Увеличение блока в %	Количество контрольных битов	Увеличение блока в %
8	4	50	5	62,5
16	5	31,25	6	37,5
32	6	18,75	7	21,875
64	7	10,94	8	12,5
128	8	6,25	9	7,03
256	9	3,52	10	3,91

Для удобства генерируемый синдром должен обладать следующими свойствами.

- Если синдром состоит только из нулей — ошибки не обнаружены.
- Если один и только один бит синдрома равен 1 — ошибка присутствует в одном из контрольных битов; в этом случае исправлять ошибку не нужно.
- Если синдром содержит более одного бита со значением 1, он является указателем на положение ошибки в слове для исправления которой указанный бит инвертируется.

Для получения заданных характеристик контрольные биты и биты данных следующим образом комбинируются в  $n$ -битовый блок. Начиная с младшего (крайнего правого) разряда через промежутки, которые соответствуют степеням 2 (т.е. на позиции 1, 2, 4, ...,  $2^{(n-k)}$ ), вводятся контрольные биты Хэмминга. Оставшиеся позиции заполняются битами данных. Для расчета контрольных битов каждый бит данных со значением 1 представляется двоичным значением, соответствующим его положению в последовательности. Таким образом, если 9-й бит равен 1, то соответствующее ему значение — 1001. Затем ко всем битам последовательности применяют операцию исключающего ИЛИ, в результате получают биты кода Хэмминга. В приемнике операция исключающего ИЛИ применяется ко всем элементам последовательности со значением 1. В этом случае операция применяется и к контрольным битам, и к битам данных. Поскольку положение контрольных битов определяется степенями 2, операцию исключающего ИЛИ можно применить ко всем информационным элементам последовательности, имеющим значение 1, и коду Хэмминга (состоящему из контрольных битов). Ошибки отсутствуют, если результатом операции будет 0. Если результат отличен от нуля — он представляет собой синдром, значение которого соответствует положению ошибочного бита.

**Пример.** Код Хэмминга (8, 4) характеризуется распределением, представленным в табл. 8.2. Имеется 8-битовый блок данных 00111001. Значение 1 имеют четыре бита (в таблице затенены), к позициям этих битов применяется операция исключающего ИЛИ, в результате которой получаем код Хэмминга 0111 — четыре контрольных бита. Следовательно, передается блок 001101001111. Предположим, что третий бит данных (позиция 6) ошибочен и изменил значение с 0 на 1. Тогда приемником будет получен следующий блок: 001101101111. Полученный код Хэмминга по-прежнему равен 0111. Приемник применяет операцию исключающего ИЛИ к коду Хэмминга и всем значениям позиций ненулевых битов данных. В результате будет получена последовательность 0110. Ненулевой результат свидетельствует о наличии ошибки и указывает, что ошибочный бит находится на позиции 6.

Таблица 8.2. Расположение битов данных и контрольных битов

а) Переданный блок												
Разряд	12	11	10	9	8	7	6	5	4	3	2	1
Номер разряда	1100	1011	1010	1001	1000	0111	0110	0101	0100	0011	0010	0001
Бит данных	D8	D7	D6	D5		D4	D3	D2		D1		
Контрольный бит					C8				C4		C2	C1
Переданный блок	0	0	1	1	0	1	0	0	1	1	1	1
Коды			1010	1001		0111				0011		

б) Расчет контрольных битов перед передачей сигнала	
Разряд	Код
10	1010
9	1001

Разряд												Код	
7												0111	
3												0011	
исключающее ИЛИ = C8 C4 C2 C1												0111	
в) Полученный блок													
Разряд	12	11	10	9	8	7	6	5	4	3	2	1	
Номер разряда	1100	1011	1010	1001	1000	0111	0110	0101	0100	0011	0010	0001	
Бит данных	D8	D7	D6	D5	D4		D3	D2	D1				
Контрольный бит					C8			C4		C2			C1
Переданный блок	0	0	1	1	0	1	1	0	1	1	1	1	
Коды				1010		1001		0111		0110		0011	
г) Расчет контрольных битов после приема сигнала													
Разряд												Код	
												0111	
10												1010	
9												1001	
7												0111	
6												0110	
3												0011	
исключающее ИЛИ = синдром												0111	

Описанный код известен как *код исправления однобитовых ошибок* (single-error-correcting — SEC). Одной из разновидностей этого кода является код исправления 1-битовых и обнаружения 2-битовых ошибок (single-error-correcting, double-error-detecting — SEC-DED). Как показано в табл. 8.1, для таких кодов требуется на один бит больше, чем для кодов SEC. Этот дополнительный бит является битом четности для всего кодового блока.

## Циклические коды

Большинство используемых на практике блочных кодов с коррекцией ошибок относятся к категории циклических. Для циклических кодов справедливо следующее: если  $n$ -битовая последовательность  $c = (c_0, c_1, \dots, c_{n-1})$  является кодовым словом, то последовательность  $(c_{n-1}, c_0, c_1, \dots, c_{n-2})$ , полученная с помощью циклического сдвига  $c$  на одну позицию вправо, также используется в качестве кодового слова. Данный класс кодов можно легко кодировать и декодировать с использованием линейных регистров сдвига с обратной связью (linear feedback shift register — LFSR). Примерами циклических кодов являются коды Боуза-Чоудхури-Хоквенгема (БХЧ) и коды Рида-Соломона.

Реализация циклического кодера как регистра LFSR подобна реализации, приведенной на рис. 8.4 для кодов обнаружения ошибок. Основное отличие состоит в том, что вход кода CRC имеет произвольную длину и в результате полу-

чается контрольный код CRC фиксированной длины, тогда как циклический код с коррекцией ошибок генерирует контрольный код  $(n - k)$  бит) на основе входной последовательности фиксированной длины  $(k)$  бит).

На рис. 8.7 представлена реализация в виде LFSR декодера циклического блочного кода. Сравните данный рисунок с логикой кодера, приведенной на рис. 8.4. Отметим, что в кодере  $k$  бит данных используются как вход для получения в регистре сдвига кода длиной  $(n - k)$  бит. Для декодера входом являются полученные  $n$  бит, которые содержат  $k$  бит данных, за которыми следуют  $(n - k)$  контрольных битов. При отсутствии ошибок после первых  $k$  тактов регистр сдвига содержит последовательность контрольных битов, идентичную переданной. После оставшихся  $(n - k)$  тактов регистр сдвига будет содержать код-синдром.

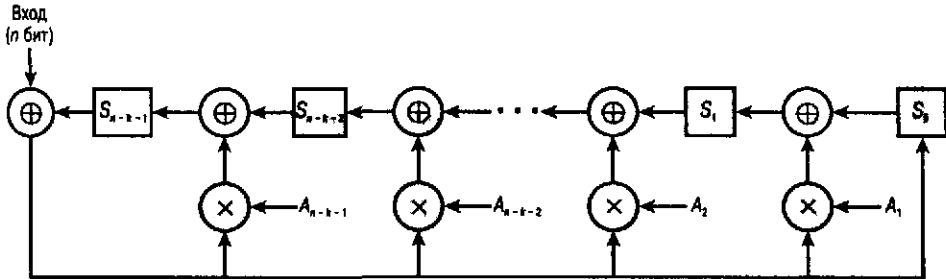


Рис. 8.7. Генератор блочного синдрома для делителя  $1 + A_1X + A_2X^2 + \dots + A_{n-1}X^{n-k-1} + X^{n-k}$

Для декодирования циклического кода используется следующая процедура.

1. С помощью полученных битов вычисляется код-синдром. Вычисления проводятся аналогично тому, как кодер обрабатывает биты данных для получения контрольного кода.
2. Если все биты синдрома равны нулю — ошибки отсутствуют.
3. Если синдром отличен от нуля, производится дополнительная обработка для исправления ошибки.

Значение синдромов можно понять, изучив блочный код с использованием полиномов. Как и при проверке четности с избыточностью, конкретный циклический код можно представить полиномиальным делителем, именуемым порождающим многочленом (или генератором). Для кода  $(n, k)$  порождающий многочлен можно записать в таком виде:

$$P(X) = 1 + \sum_{i=1}^{n-k-1} A_i X^i + X^{n-k},$$

где каждый из коэффициентов  $A_i$  может принимать значения 0 или 1, что соответствует двоичному разряду в делителе. Например, для  $P = 11001$  многочлен  $P(X)$  имеет вид  $X^4 + X^3 + 1$ . Аналогично последовательность битов данных представляется полиномом  $D(X)$ , а контрольный код — полиномом  $C(X)$ . Напомним (см. обсуждение циклической проверки четности с избыточностью), что контрольный код определяется следующим образом.

$$\frac{X^{n-k}D(X)}{P(X)} = Q(X) + \frac{C(X)}{P(X)}$$

Т.е. блок данных  $D(X)$  сдвигается влево на  $(n-k)$  бит и делится на  $P(X)$ . В результате получим частное  $Q(X)$  и остаток  $C(X)$  длиной  $(n-k)$  бит. Передаваемый блок — это конкатенация  $D(X)$  и  $C(X)$ :

$$T(X) = X^{n-k}D(X) + C(X). \quad (8.3)$$

При отсутствии ошибок передачи  $T(X)$  должно делиться на  $P(X)$  без остатка, что легко продемонстрировать:

$$\frac{T(X)}{P(X)} = \frac{X^{n-k}D(X)}{P(X)} + \frac{C(X)}{P(X)} = \left( Q(X) + \frac{C(X)}{P(X)} \right) + \frac{C(X)}{P(X)} = Q(X). \quad (8.4)$$

Последнее выражение справедливо в соответствии с правилами арифметики по модулю 2 ( $a + a = 0$ ). Следовательно, если ошибки отсутствуют,  $T(X)$  делится на  $P(X)$  без остатка.

Если один или более битов являются ошибочными, полученный блок  $Z(X)$  будет иметь такой вид:

$$Z(X) = T(X) + E(X),$$

где  $E(X)$  —  $n$ -битовый полином ошибок, содержащий 1 в каждом двоичном разряде, где в  $Z(X)$  имеется ошибка. Если  $Z(X)$  передается через регистр LFSR, который показан на рис. 8.7, фактически производится деление  $Z(X)/P(X)$ , что в результате дает синдром  $S(X)$  длиной  $(n-k)$  бит:

$$\frac{Z(X)}{P(X)} = B(X) + \frac{S(X)}{P(X)}, \quad (8.5)$$

где  $B(X)$  — частное, а  $S(X)$  — остаток деления. Следовательно,  $S(X)$  является функцией  $Z(X)$ . Каким же образом полученный результат можно использовать для исправления ошибок? Для ответа на этот вопрос, запишем уравнение (8.5) в развернутой форме:

$$\frac{Z(X)}{P(X)} = B(X) + \frac{S(X)}{P(X)}, \quad (8.6)$$

$$\frac{T(X) + E(X)}{P(X)} = B(X) + \frac{S(X)}{P(X)},$$

$$Q(X) + \frac{E(X)}{P(X)} = B(X) + \frac{S(X)}{P(X)},$$

$$\frac{E(X)}{P(X)} = [Q(X) + B(X)] + \frac{S(X)}{P(X)}.$$

Видим, что деление  $E(X)/P(X)$  дает тот же остаток, что и  $Z(X)/P(X)$ . Следовательно, значение синдрома  $S(X)$  зависит только от ошибочных битов и не зависит от

начальной последовательности битов (переданного значения  $T(X)$ ). Если ошибочные биты  $E(X)$  можно получить из синдрома  $S(X)$ , то ошибки в  $Z(X)$  можно исправить посредством простого сложения:

$$Z(X) + E(X) = T(X) + E(X) + E(X) = T(X).$$

Поскольку  $S(X)$  зависит только от  $E(X)$ , возможности блочного циклического кода определить очень легко. Синдром состоит из  $(n - k)$  бит, следовательно, он может принимать  $2^{n-k}$  возможных значений. Нулевой синдром указывает на отсутствие ошибок. Следовательно, всего можно исправить  $(2^{n-k} - 1)$  различных ошибочных комбинаций. Чтобы с помощью кода  $(n, k)$  можно было исправить все возможные однобитовые ошибки, должно выполняться неравенство  $n \leq (2^{n-k} - 1)$ . Исправление всех 1- и 2-битовых ошибок требует выполнения следующего неравенства:

$$\left( n + \frac{n(n-1)}{2} \right) \leq (2^{n-k} - 1).$$

Способ получения  $E(X)$  из  $S(X)$  может зависеть от используемого кода. Наиболее простой подход — построить таблицу, которая ставила бы в соответствие значениям  $E(X)$  значения  $S(X)$ . После этого потребуется простой способ выполнения поиска в такой таблице.

**Пример<sup>5</sup>.** Рассмотрим код  $(7, 4)$  с порождающим многочленом  $P(X) = X^3 + X^2 + 1$ . Имеем  $7 = 2^3 - 1$ , следовательно с помощью этого кода можно исправить все 1-битовые ошибки. В табл. 8.3, а приводятся все используемые кодовые слова. Отметим, что  $d_{\text{мин}} = 3$ , значит, данный код — это код с коррекцией 1-битовых ошибок. Например, для блока данных 1010 имеем  $D(X) = X^3 + X$  и  $X^{n-k}D(X) = X^6 + X^4$ . Деление выполняется согласно уравнению (8.4):

$$\begin{array}{r}
 2^3 D(X) \longrightarrow \begin{array}{r} X^6 + X^4 \\ \underline{X^6 + X^5 + X^3} \\ X^5 + X^4 + X^3 \\ \underline{X^5 + X^4 + X^2} \\ X^3 + X^2 \\ \underline{X^3 + X^2 + 1} \\ 1 \longleftarrow S(X) \end{array} \qquad \begin{array}{l} \left| \begin{array}{l} X^3 + X^2 + 1 \longleftarrow P(X) \\ X^3 + X^2 + 1 \longleftarrow Q(X) \end{array} \right. \end{array}
 \end{array}$$

Далее используя уравнение (8.3), получаем  $T(X) = X^6 + X^4 + 1$ , что соответствует кодовому слову 1010001.

Для коррекции ошибок требуется таблица синдромов, показанная в табл. 8.3, б. Например, для последовательности ошибок 1000000  $E(X)$  будет равно  $X^6$ . Используя последнюю строку уравнения (8.6), рассчитываем:

<sup>5</sup> Данный пример взят из [LATH98].

$$\begin{array}{r}
 Z(X) \longrightarrow X^6 \\
 \hline
 X^6 + X^5 + \quad X^3 \\
 \hline
 X^5 + \quad X^3 \\
 \hline
 X^5 + X^4 + \quad X^2 \\
 \hline
 X^4 + X^3 + X^2 \\
 \hline
 X^4 + X^3 + \quad X \\
 \hline
 X^2 + X \longleftarrow S(X)
 \end{array}
 \quad
 \begin{array}{r}
 X^3 + X^2 + 1 \longleftarrow P(X) \\
 \hline
 X^3 + X^2 + X \longleftarrow Q(X) + B(X)
 \end{array}$$

Таким образом,  $S = 110$ . Остальные ячейки табл. 8.3, б вычисляются аналогично. Теперь предположим, что был получен блок 1101101, или  $Z(X) = X^6 + X^5 + X^3 + X^2 + 1$ . Используя уравнение (8.5), запишем:

$$\begin{array}{r}
 Z(X) \longrightarrow X^6 + X^5 + \quad X^3 + X^2 + 1 \\
 \hline
 X^6 + X^5 + \quad X^3 \\
 \hline
 X^2 + 1 \longleftarrow S(X)
 \end{array}
 \quad
 \begin{array}{r}
 X^3 + X^2 + 1 \longleftarrow P(X) \\
 \hline
 X^3 \longleftarrow B(X)
 \end{array}$$

Следовательно,  $S = 101$ . Используя табл. 8.3, б, получаем  $E = 0001000$ . Тогда

$$T = 1101101 \oplus 0001000 = 1100101.$$

Итак, согласно табл. 8.3, а, был передан блок данных 1100.

**Таблица 8.3. Циклический код с коррекцией 1-битовых ошибок (7,4)**

**а) Таблица используемых кодовых слов**

Блок данных	Кодовое слово
0000	0000000
0001	0001101
0010	0010111
0011	0010101
0100	0100011
0101	0101110
0110	0110100
0111	0111001
1000	1000110
1001	1001011
1010	1010001
1011	1011100
1100	1100101
1101	1101000
1110	1110010
1111	1111111

б) Таблица синдромов, соответствующих 1-битовым ошибкам

Ошибочная комбинация	Синдром
0000001	001
0000010	010
0000100	100
0001000	101
0010000	111
0100000	011
1000000	110

### Коды БХЧ

Коды БХЧ являются одним из наиболее мощных циклических блочных кодов и получили широкое применение в беспроводных приложениях. Для любой пары положительных целых чисел  $m$  и  $t$  существуют двоичные коды БХЧ  $(n, k)$  со следующими параметрами.

Длина блока	$n = 2^m - 1$
Количество контрольных битов	$n - k \leq mt$
Минимальное расстояние	$d_{\min} \geq 2t + 1$

С помощью такого кода можно исправить все слова, содержащие  $t$  (или менее) ошибок. Порождающий многочлен кода БХЧ можно создать из множителей полинома  $(X^{2^m-1} + 1)$ . Использование кодов БХЧ предоставляет некоторую свободу выбора параметров (длина блока, степень кодирования). В табл. 8.4 перечислены параметры БХЧ для кодов длиной до  $2^8 - 1$ . В табл. 8.5 приводятся некоторые порождающие многочлены БХЧ.

Таблица 8.4. Параметры кодов БХЧ

$n$	$k$	$t$	$n$	$k$	$t$	$n$	$k$	$t$	$n$	$k$	$t$	$n$	$k$	$t$
7	4	1	63	30	6	127	64	10	255	207	6	255	99	23
15	11	1		24	7		57	11		199	7		91	25
	7	2		18	10		50	13		191	8		87	26
	5	3	16	11	43	14	187	9	79	27				
31	26	1	10	13	36	15	179	10	71	29				
	21	2	7	15	29	21	171	11	63	30				
	16	3	127	120	1	22	23	163	12	55	31			
11	5	113		2	15	27	155	13	47	42				
63	6	7	106	3	8	31	147	14	45	43				
	57	1	99	4	255	247	1	139	15	37	45			
	51	2	92	5		239	2	131	18	29	47			
	45	3	85	6		231	3	123	19	21	55			
	39	4	78	7		223	4	115	21	13	59			
	36	5	71	9		215	5	107	22	9	63			



Таблица 8.5. Порождающие многочлены кодов БХЧ

$n$	$k$	$t$	$P(X)$
7	4	1	$X^3 + X + 1$
15	11	1	$X^4 + X + 1$
15	7	2	$X^8 + X^7 + X^6 + X^4 + 1$
15	5	3	$X^{10} + X^8 + X^4 + X^4 + X^2 + X + 1$
31	26	1	$X^4 + X^2 + 1$
31	21	2	$X^{10} + X^9 + X^8 + X^6 + X^5 + X^3 + 1$

Для декодирования сигналов БХЧ было разработано большое количество методов, требующих меньше памяти, чем прямой поиск в таблице. Одна из наиболее простых схем была предложена в работе [BERL80]. Ее основная идея состоит в вычислении полинома локализации ошибки с последующим нахождением его корней. Сложность данного алгоритма увеличивается как квадрат количества ошибок, которые нужно исправить.

## Коды Рида-Соломона

Коды Рида-Соломона (Reed-Solomon codes — RS codes) — это широко используемый подкласс недвоичных кодов БХЧ. При использовании кодов Рида-Соломона данные обрабатываются порциями по  $m$  бит, именуемыми символами. Код  $(n, k)$  характеризуется следующими параметрами.

Длина символа	$m$ бит
Длина блока	$n = (2^m - 1)$ символов = $m(2^m - 1)$ бит
Длина блока данных	$k$ символов
Размер контрольного кода	$n - k = 2t$ символов = $m(2t)$ бит
Минимальное расстояние	$d_{\min} = (2t + 1)$ символов

Таким образом, алгоритм кодирования расширяет блок  $k$  символов до размера  $n$ , добавляя  $(n - k)$  избыточных контрольных символов. Как правило,  $m$  является степенью 2; широко используется значение  $m = 8$ .

**Пример.** Пусть  $t = 1$ ,  $m = 2$ . Обозначая символы как 0, 1, 2, 3, их двоичные эквиваленты можно записать как 0 = 00; 1 = 01; 2 = 10; 3 = 11. Код имеет следующие параметры:

$$n = 2^2 - 1 = 3 \text{ символа} = 6 \text{ бит,}$$

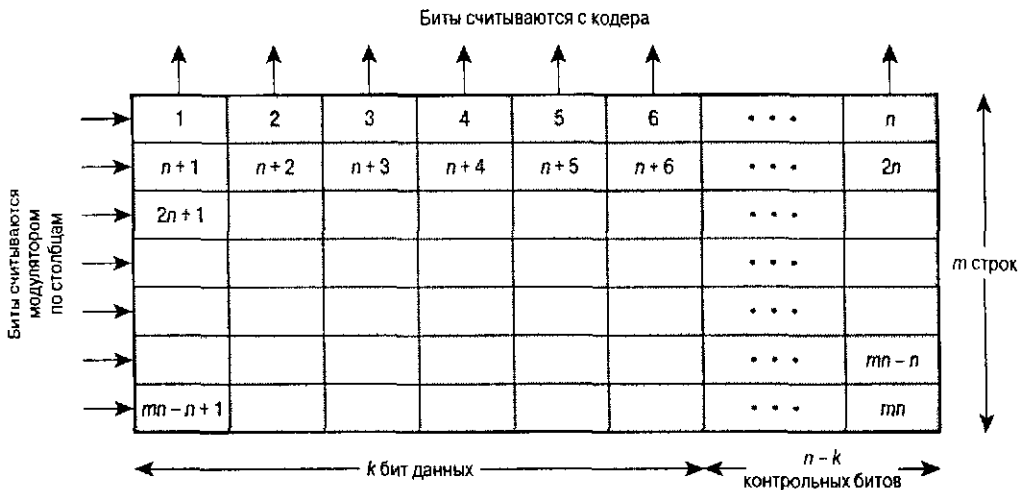
$$n - k = 2 \text{ символа} = 4 \text{ бит.}$$

С помощью данного кода можно исправить любой пакет ошибок, который искажает 2-битовый символ.

Коды Рида-Соломона удобны для исправления пакетов ошибок. Данный тип кодов характеризуется высокоэффективным использованием избыточности, длина блоков и размеры символов могут легко приспособливаться под сообщения разных размеров. Кроме того, для таких кодов существуют эффективные методы кодирования.

## Чередование блоков

В беспроводных системах с блочными кодами широко используется метод чередования блоков; пример такого чередования приводится на рис. 8.8. Преимущество чередования состоит в том, что приемник распределяет пакет ошибок, искаживший некоторую последовательность битов, по большому числу блоков, благодаря чему становится возможным исправление ошибок. Чередование выполняется с помощью чтения и записи данных в различном порядке. На рис. 8.8 представлен простой и широко применяемый метод чередования. Здесь данные, которые необходимо передать, заносятся в прямоугольную матрицу, каждая строка которой состоит из  $n$  бит, что равно размеру блока, а считываются эти данные по столбцам. В результате  $k$  бит данных и соответствующие им  $(n - k)$  контрольных битов, которые вместе составляют блок из  $n$  бит, рассеиваются и перемешиваются с битами других блоков. После приема исходный порядок битов восстанавливается. Если во время передачи пакет помех воздействует на некоторую последовательность битов, то все эти биты оказываются разнесенными по различным блокам. Следовательно, от любой контрольной последовательности требуется возможность исправления лишь небольшой части от общего количества инвертированных битов. В частности, любой пакет ошибок длиной  $l = mb$  разбивается на  $m$  пакетов длиной  $b$ . После некоторого размышления легко опять справедливость такого утверждения. Предположим, что используется код  $(n, k)$ , с помощью которого можно исправить все комбинации из  $t$  (или менее) битов;  $t = \lfloor (n - k)/2 \rfloor$ . Если степень чередования равна  $m$ , то в результате будет получен код  $(mn, mk)$ , с помощью которого можно исправить ошибки, затрагивающие до  $mt$  бит.



*Примечание* Числа в ячейках матрицы указывают порядок считывания битов  
После чередования биты подаются на выход в следующем порядке: 1,  $n + 1$ ,  $2n + 1$  и т.д.

Рис. 8.8. Чередование блоков

## 8.3. СВЕРТОЧНЫЕ КОДЫ

Блочные коды являются одной из двух категорий кодов с коррекцией ошибок, широко используемых при беспроводной передаче. Вторая категория — это сверточные коды. Блочный код  $(n, k)$  обрабатывает данные блоками по  $k$  бит, генерируя на выходе блок из  $n$  бит ( $n > k$ ) для каждого  $k$ -битового блока на входе. Если прием и передача данных происходят относительно непрерывным потоком, то блочный код (в частности, с большим значением  $n$ ) может быть не так удобен, как код, который генерирует избыточные биты непрерывно. В последнем случае обнаружение и исправление ошибок выполняется непрерывно, и именно в этом состоит преимущество сверточных кодов.

Сверточный код задается тремя параметрами:  $n$ ,  $k$  и  $K$ . Код  $(n, k, K)$  обрабатывает входящие данные порциями по  $k$  бит и генерирует выходную последовательность, состоящую из  $n$  бит для каждых  $k$  бит входа. До этого момента принципы работы сверточных и блочных кодов не отличаются. Для сверточных кодов  $n$  и  $k$ , как правило, являются очень малыми числами. Разница между двумя типами кодов состоит в том, что сверточные коды используют память, которая характеризуется *длиной кодового ограничения*  $K$ . По сути, текущая  $n$ -битовая выходная последовательность кода  $(n, k, K)$  зависит не только от значений текущего входного блока, состоящего из  $k$  бит, но также и от предыдущих  $(K - 1)$   $k$ -битовых блоков. Следовательно, текущая выходная  $n$ -битовая последовательность является функцией последних  $(K \times k)$  входных битов.

Принципы работы сверточных кодов удобно рассмотреть на конкретном примере, представленном на рис. 8.9. Здесь приведены два альтернативных представления кода. Рис. 8.9, *а* — это регистр сдвига, который наиболее удобен для описания и реализации процесса кодирования. Рис. 8.9, *б* — это эквивалентное представление, удобное для изучения процесса декодирования.

Для кода  $(n, k, K)$  регистр сдвига содержит последние  $(K \times k)$  входных битов; в исходном состоянии все ячейки регистра содержат нули<sup>6</sup>. Кодер генерирует  $n$  выходных битов, после чего наиболее “старые”  $k$  бит регистра стираются и вводится новая  $k$ -битовая последовательность. Хотя выходные  $n$  бит зависят от  $(K \times k)$  входных битов, степень кодирования равна отношению  $k$  входных битов к  $n$  выходных битов. Следовательно, как и для блочного кода, степень кодирования равна  $k/n$ . Наиболее широко используемые двоичные кодеры имеют  $k = 1$ ; соответственно, длина регистра такого кодера равна  $K$ . В рассматриваемом примере (рис. 8.9, *а*) используется код  $(2, 1, 3)$ . Здесь кодер преобразует входной бит  $u_n$  в два выходных бита —  $v_{n1}$  и  $v_{n2}$ , используя три последних полученных бита. Первый сгенерированный бит поступает из верхнего логического контура ( $v_{n1} = u_n \oplus u_{n-1} \oplus u_{n-2}$ ) а второй — из нижнего ( $v_{n2} = u_n \oplus u_{n-2}$ ).

---

<sup>6</sup> Иногда в литературе указано, что используется на одну ячейку регистра сдвига меньше, а входные биты подаются как в ячейку памяти, так и в контуры исключяющего ИЛИ; обе схемы эквивалентны.

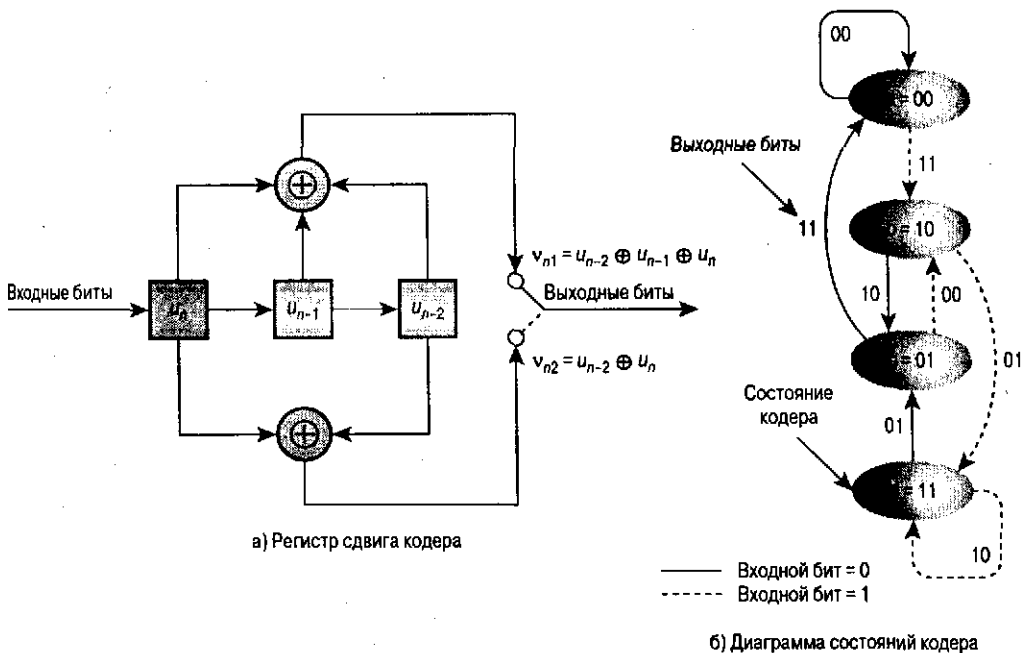


Рис. 8.9. Сверточный кодер (2, 1, 3)

Для любого данного  $k$ -битового входа существует  $2^{k(K-1)}$  различных функций, отображающих  $k$  входных битов в  $n$  выходных. Решение о том, какая из этих функций будет использована, зависит от истории последних  $(K-1)$   $k$ -битовых входных блоков. Следовательно, сверточный код можно представить как конечный автомат. Автомат имеет  $2^{k(K-1)}$  различных состояний, переход между которыми определяется последними  $k$  входными битами и дает  $n$  выходных битов. Начальное состояние автомата — все нули. Для примера, приведенного на рис. 8.9, б существует 4 состояния, по одному для каждой возможной комбинации последних двух битов. Следующий входной бит инициирует переход и дает два выходных бита. Например, если последние два бита — 10 ( $u_{n-1} = 1$ ;  $u_{n-2} = 0$ ), а следующий бит равен 1 ( $u_n = 1$ ), то текущим состоянием будет  $b(10)$ , а следующим состоянием —  $d(11)$ . Выход имеет такой вид:

$$v_{n1} = u_{n-2} \oplus u_{n-1} \oplus u_n = 0 \oplus 1 \oplus 1 = 0,$$

$$v_{n2} = 0 \oplus 1 = 1.$$

## Декодирование

Чтобы описать процесс декодирования, можно расширить диаграмму состояний, показав в кодере хронологическую последовательность битов. Если диаграмма состояний расположена вертикально, как на рис. 8.9, б, то расширенная диаграмма, называемая решетчатой, строится путем воспроизведения состояний и представления переходов между состояниями в горизонтальном

направлении слева направо, в соответствии с течением времени или посредством ввода данных (рис. 8.10). Если длина кодового ограничения  $K$  велика, то решетчатая диаграмма будет слишком громоздкой. В таком случае для отображения переходов можно использовать  $2^{K-2}$  упрощенных решетчатых фрагментов. На рис. 8.11 приводится подобное отображение для кода (2, 1, 7). Здесь показано каждое состояние кодера и даны определения всех ветвей диаграммы.

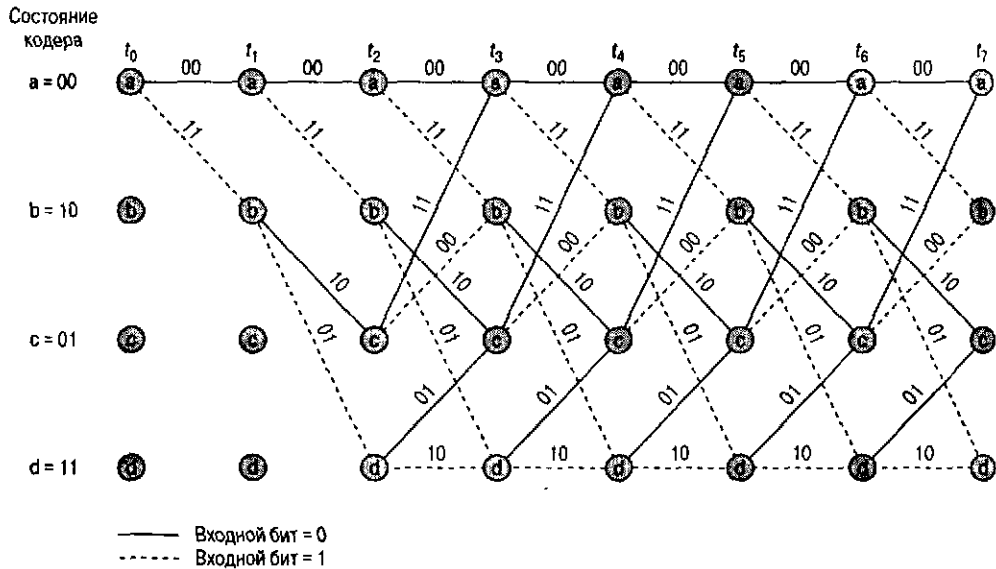
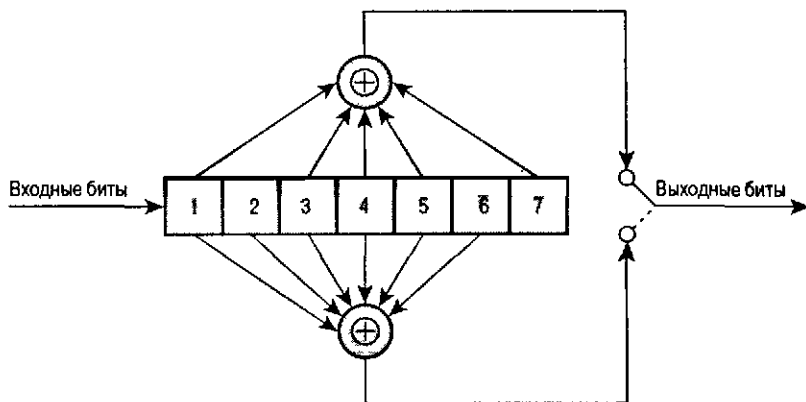


Рис. 8.10. Решетчатая диаграмма для кодера, изображенного на рис. 8.9

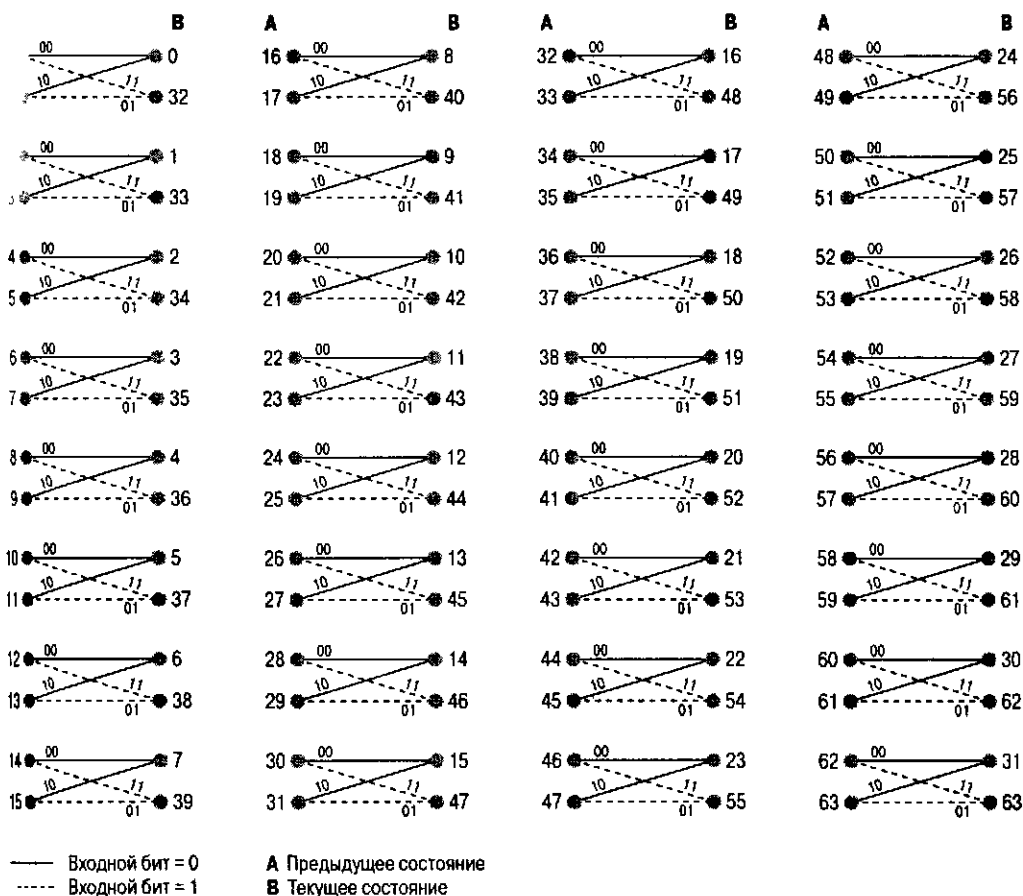
Любая корректная выходная последовательность определяется маршрутом в решетчатой диаграмме. Для рассматриваемого примера маршрут a-b-c-b-d-c-a генерируется входной последовательностью 1011000 и дает на выходе последовательность 11 10 00 01 01 11 00. При некорректном переходе, например a-c, декодер пытается исправить ошибку, определив, какому входу с наибольшей вероятностью соответствует некорректный выход.

Для сверточных кодов разработано большое количество алгоритмов исправления ошибок. Пожалуй, важнейшим из них является код Витерби (Viterbi code). По сути, метод Витерби — это сравнение полученной последовательности со всеми возможными переданными последовательностями. Алгоритм подбирает маршрут через решетчатую диаграмму, кодированная версия которого отличается от полученной наименьшим числом элементов. Как только выбран правильный маршрут, декодер может восстановить входные биты данных из выходных битов кода.

В зависимости от метрики, используемой для определения отличия полученной и истинной последовательностей, существует несколько разновидностей алгоритма Витерби. Для представления идеи алгоритма используем широко распространенную метрику — расстояние Хэмминга.



а) Диаграмма регистра сдвига



б) Диаграммы переходов

Рис. 8.11. Решетчатые диаграммы для кодера (2, 1, 7) [JAIN90]

Представим полученную закодированную последовательность в виде слова  $w = w_0w_1w_2, \dots$  и попытаемся найти наиболее вероятный маршрут через решетчатую диаграмму. Для каждого момента времени  $i$  и для каждого состояния перечислим *активный маршрут* (или маршруты) через диаграмму к заданному состоянию. Под активным маршрутом подразумевают корректный маршрут через решетчатую диаграмму, для которого расстояние Хэмминга до полученного слова минимально до момента времени  $i$ . В момент времени  $i$  пометим каждое состояние расстоянием от его активного маршрута до полученного слова. Используется следующее соотношение:

$$\text{расстояние до полученного слова} = \text{расстояние на предыдущем шаге} + \text{расстояние для последнего перехода.} \quad (8.7)$$

Алгоритм выполняется в  $(b + 1)$  шаг, где  $b$  — предопределенный размер окна. Для кода  $(n, k, K)$  декодирование первого выходного  $n$ -битового блока  $(w_0w_1w_2\dots w_{n-1})$  проходит следующим образом.

- **Шаг 0.** Начальное состояние решетки в момент времени 0 помечается нулем, поскольку на этот момент маршруты не отличаются.
- **Шаг  $i + 1$ .** Для каждого состояния  $S$  в момент времени  $(i + 1)$  с помощью уравнения (8.7) находятся все ведущие к нему активные маршруты. Состояние  $S$  помечается длиной данного маршрута (маршрутов).
- **Шаг  $b$ .** Алгоритм прекращается в момент времени  $b$ . Если в этот момент времени все активные маршруты начинаются с одного отрезка, и метка этого отрезка —  $x_0x_1x_2\dots x_{n-1}$ , то первый кодовый блок  $w_0w_1w_2\dots w_{n-1}$  исправляется на  $x_0x_1x_2\dots x_{n-1}$ . Если же существуют два активных начальных отрезка, ошибка является неисправимой.

После принятия и исправления, при необходимости, первого кодового блока окно декодирования сдвигается на  $n$  бит вправо, после чего выполняется декодирование следующего блока.

**Пример<sup>7</sup>.** На рис. 8.12 приводится пример применения алгоритма Витерби к последовательности 10010100101100, ... с размером окна декодирования  $b = 7$  (подразумевается использование кодера, определенного на рис. 8.9 и 8.10). Линии на рисунке соответствуют корректным маршрутам решетчатой диаграммы. Жирные линии отображают текущие активные маршруты. На шаге 1 принимается последовательность  $w_0w_1 = 10$ . Корректными являются две последовательности: 00 и 11. Для обеих последовательностей расстояние до полученной последовательности равно 1. Определяются два активных маршрута, каждый с меткой состояния "1". На следующем шаге имеем  $w_1w_2 = 01$ . Используя уравнение (8.7), находим разницу между входом и четырьмя возможными корректными состояниями (сверху вниз): 2, 2, 3 и 1. На данном этапе все возможные корректные маршруты считаются активными. Некоторые корректные маршруты на шаге 3 перестают быть активными. Это происходит из-за того, что такие маршруты отбрасываются, если в текущем состоянии существует другой корректный маршрут с меньшим расстоянием. Например, последовательность состояний a-a-a-a

<sup>7</sup> За основу взят пример из [ADAM91].

характеризуется тремя несовпадениями; последовательность состояний  $a-b-c-a$  — четырьмя. После шага 7 все активные маршруты проходят через начальный отрезок  $a-b$ , выход для которого равен 11. Алгоритм исправляет  $w_0w_1$  на 11 и переходит к следующему блоку ( $w_1w_2$ ). Отметим, что если бы размер окна был равен 5, ошибка была бы неисправимой.

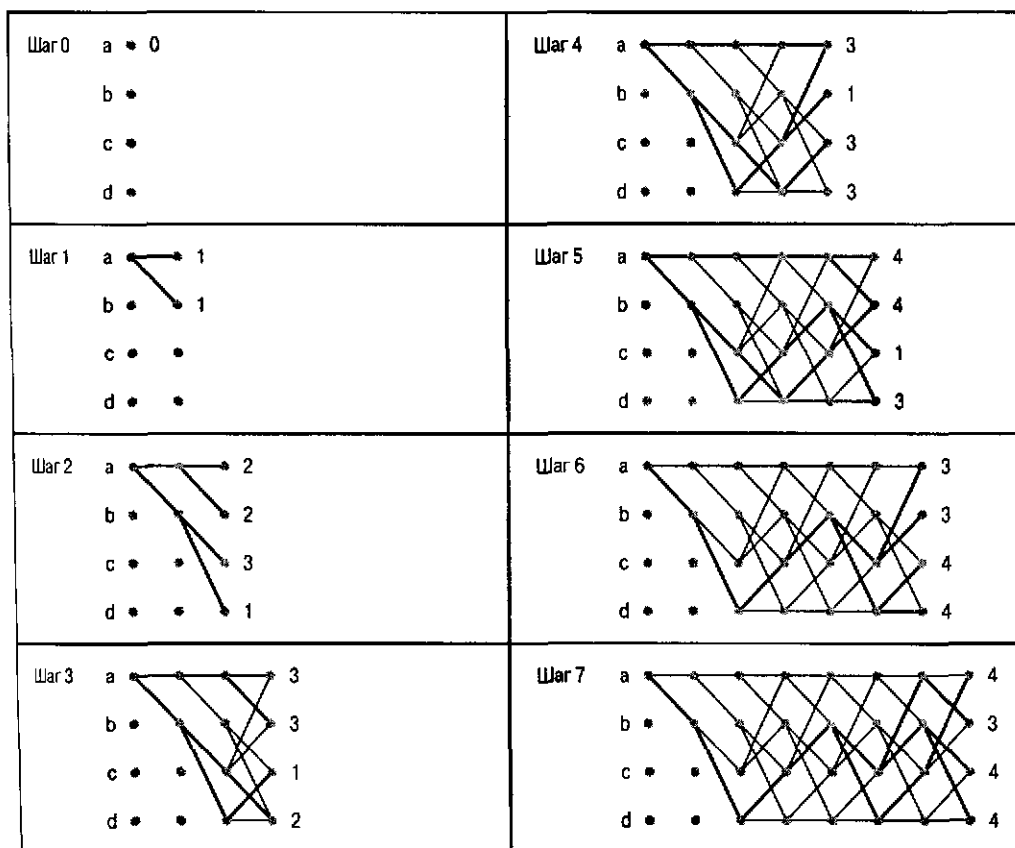


Рис. 8.12. Алгоритм Витерби для  $w = 10010100101100 \dots$  с размером окна декодирования  $b = 7$

Сверточные коды дают возможность достичь хорошей производительности в каналах с шумом, где искаженной является большая доля переданных битов. Следовательно, данные коды все шире используются в беспроводных приложениях.

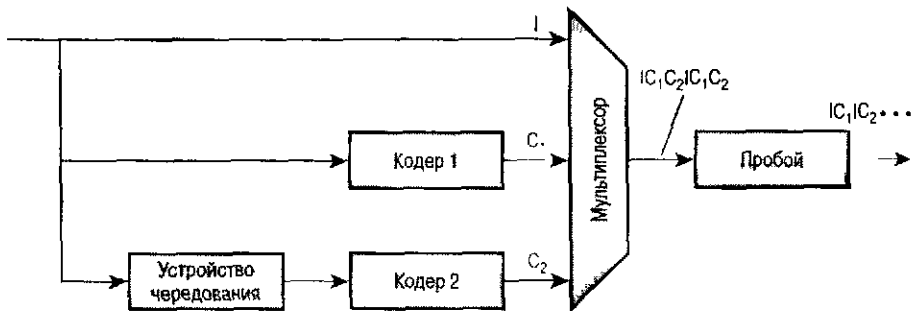
## Турбокодирование

По мере увеличения скорости передачи данных в беспроводных приложениях исправление ошибок становится все более важной задачей при проектировании систем. В последнее время в беспроводных системах третьего поколения получил широкое распространение новый класс кодов, известных как турбокоды. Использование турбокодов позволяет достичь очень близкой к пределу Шеннона вероятности ошибок, и эти коды можно эффективно реализовать для высокоскоростной связи. На данный момент спроектировано довольно много кодеров и де-

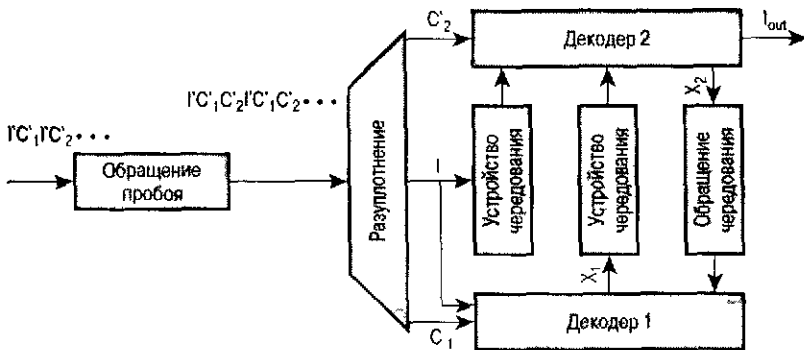


кодеров, использующих турбокоды, основанные преимущественно на сверточном кодировании. В данном разделе представлен общий обзор таких устройств.

На рис. 8.13, а изображен турбокодер. В приведенной схеме кодер дублируется дважды. Первый кодер получает поток входных битов, для каждого из которых генерирует единственный контрольный бит  $C_1$ . Вход второго кодера — это подвергнутый чередованию исходный входной поток, дающий последовательность контрольных битов  $C_2$ . Затем исходный входной бит и два контрольных бита уплотняются; в результате получается последовательность  $I_1C_{11}C_{21}I_2C_{12}C_{22}$ , ... т.е. за первым входным битом следует контрольный бит кодера 1, за ним идет контрольный бит кодера 2 и т.д. Степень кодирования результирующей последовательности равна  $1/3$ . Если использовать только половину контрольных битов, чередуя выходы двух кодеров, можно достичь степени кодирования  $1/2$  (данный процесс называют *пробоем*). В системах связи третьего поколения используются обе степени кодирования:  $1/2$  и  $1/3$ .



а) Кодер



б) Декодер

Рис. 8.13. Турбокодирование и декодирование [СОТТ00]

Отметим, что каждый кодер генерирует для каждого входного бита один контрольный бит, при этом входной бит сохраняется. В рассмотренном методе сверточного кодирования (см. например, рис. 8.9, а) входные биты не сохраняются и используются множественные выходные биты ( $n$  выходных контрольных битов для  $k$  входных битов). При турбокодировании используется разновидность

сверточного кода, известная как рекурсивный систематический сверточный код (recursive systematic convolutional — RSC). Для типичного RSC-кодера  $(2, 1, K)$  один из двух рассчитанных контрольных битов по обратной связи подается в регистр сдвига, а на выход идет другой контрольный бит (рис. 8.14). Выход кодера состоит из чередующихся входных и контрольных битов. RSC-кодер имеет такую же решетчатую структуру, как и обычный сверточный кодер; они также обладают сходными статистическими свойствами.

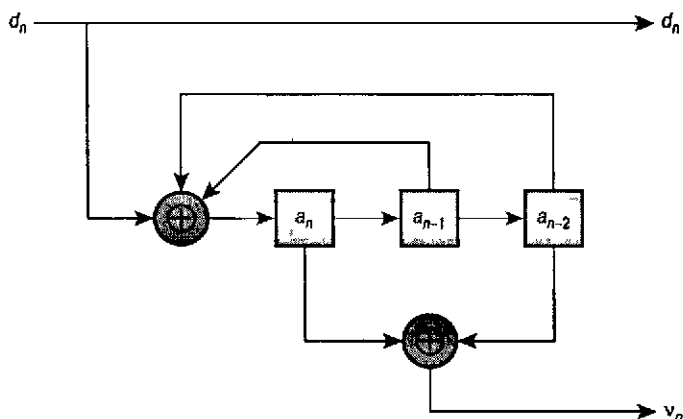


Рис. 8.14. RSC-кодер

На рис. 8.15 показано, как турбокодер можно реализовать с использованием двух RSC-кодеров; для пробы используется ключ, в результате степень кодирования равна  $1/3$  (без использования ключа степень кодирования равна  $S$ ).

На рис. 8.13, б представлена общая диаграмма турбодекодера. Недостающие контрольные биты вычисляются при необходимости. Декодер 1 начинает работу первым, используя значения  $I'$  и  $P'_1$  для создания битов коррекции  $X_1$ . Биты  $I'$  и  $X_1$ , а также значения  $P'_2$  подаются на декодер 2. Для правильного упорядочения битов требуется повторно произвести операцию чередования. Декодер 2 использует все входные данные для создания битов коррекции  $X_2$ , которые после восстановления исходного порядка битов вводятся в декодер 1 для второй итерации алгоритма декодирования. По завершении достаточного числа итераций из  $I'$  и  $X_2$  генерируется выходной бит.

## 8.4. АВТОМАТИЧЕСКИЙ ЗАПРОС ПОВТОРНОЙ ПЕРЕДАЧИ

Автоматический запрос повторной передачи (automatic repeat request — ARQ) — это механизм, применяемый в протоколах управления каналами передачи данных и транспортными протоколами, который основывается на использовании определенного кода обнаружения ошибок, например, циклической проверки четности с избыточностью, которая была описана в разделе 8.1. Механизм защиты от ошибок ARQ тесно связан с механизмом управления потоком данных, который также входит в упомянутые протоколы. Ниже рассматривается управле-

ние потоком данных, после чего анализируется схема ARQ. При обсуждении будем называть блок данных, переданный одним объектом протокола другому, модулем данных протокола (protocol data unit — PDU). Определение этого термина приводится в главе 4.

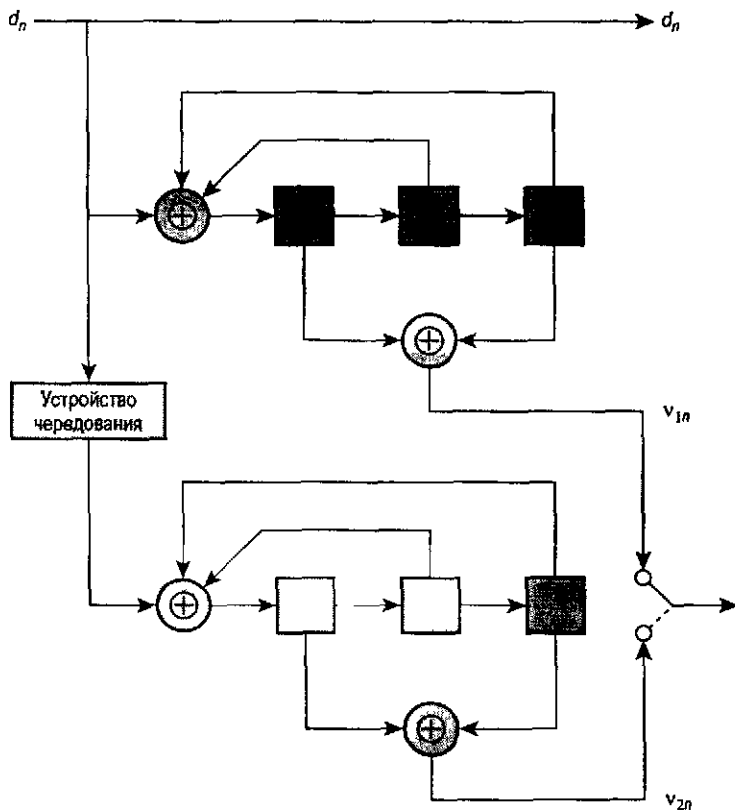


Рис. 8.15. Параллельное соединение двух RSC-кодеров [SKLA97b]

## Управление потоком данных

Управление потоком данных — это метод, гарантирующий, что передающий объект не перегрузит своими данными принимающий объект. Как правило, получающий объект отводит для данных буфер некоторого максимального объема. После принятия данных приемник должен определенным образом обработать полученную информацию (например, изучить заголовок и отделить его от PDU) перед тем, как передать ее на более высокий уровень. Отсутствие управления потоком данных может вызвать перегрузку буфера приемника.

Рассмотрим для начала механизмы управления потоком при отсутствии ошибок. Модель, которую мы будем использовать, приводится на рис. 8.16, а и представляет собой вертикальную диаграмму зависимости состояния передачи от времени. Преимущество такого представления в том, что оно позволяет показать временную структуру и взаимосвязь отправленных и принятых данных. Каждая стрелка соответствует отдельному PDU, который перемещается по каналу связи

между двумя станциями. Данные пересылаются в виде последовательности PDU, причем каждый PDU содержит информационную часть и некоторую управляющую информацию. Пока будем считать, что все переданные PDU успешно приняты — ни один из них не потерян и не получен с ошибками. Более того, будем считать, что PDU получены в той последовательности, в которой были отправлены. Следует учесть к тому же, что в процессе передачи каждый PDU задерживается на некоторое переменное время.

Как правило, перед отправкой блока или потока данных блок разбивается на меньшие блоки и информация передается с помощью нескольких PDU. Это вызвано рядом причин.

- Размер буфера приемника может быть ограничен.
- Увеличение времени передачи повышает вероятность появления ошибок, что повлечет за собой повторную передачу всего PDU. Если размер PDU мал, ошибки обнаруживаются быстрее, кроме того, повторной передачи требуют блоки меньшего размера.
- В среде совместного использования (например, ЛВС) желательно, чтобы одна станция не занимала среду на длительный период времени, приводя к откладыванию передач других станций.

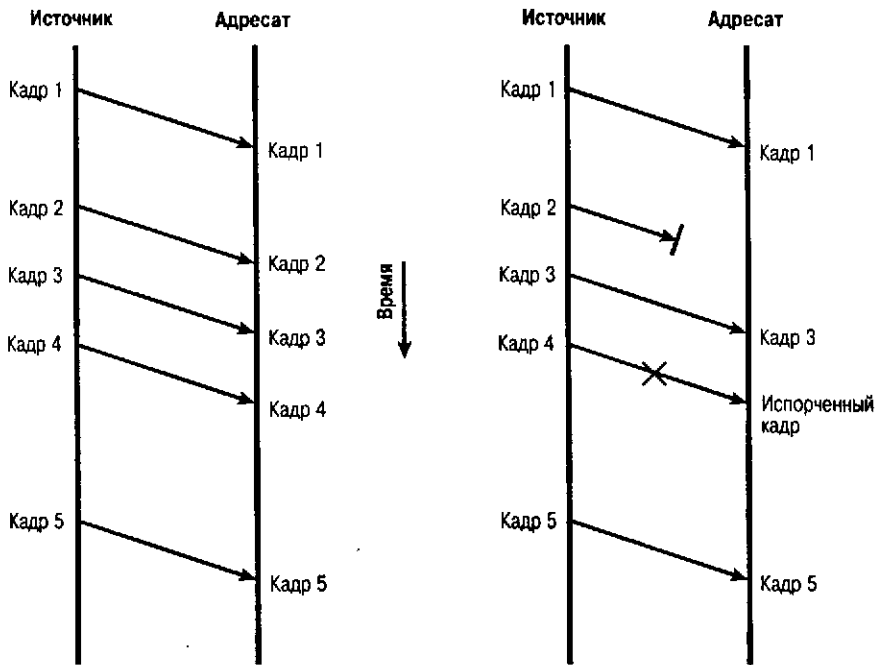


Рис. 8.16. Модель передачи PDU

Как правило, протоколы с механизмом управления потоком позволяют проводить одновременную передачу множества PDU. Рассмотрим работу двух стан-

ций, А и В, подключенных к полнодуплексному каналу. Станция В выделяет буфер для  $W$  PDU. Итак, В способна принять PDU в количестве  $W$  штук, а станция А может переслать  $W$  PDU без ожидания подтверждения. Для отслеживания подтвержденных PDU каждому из них присваивается порядковый номер. Станция В подтверждает получение PDU, отправляя сообщение, которое включает в себя порядковый номер следующего ожидаемого PDU. Данное сообщение также свидетельствует, что станция В готова принять следующие  $W$  PDU, начиная с того, номер которого указан. Данная схема может использоваться для подтверждения получения нескольких PDU. Например, станция В может получить PDU 2, 3 и 4 но задержать отправку подтверждения до прихода PDU 4. Передавая подтверждение, которое содержит порядковый номер 5, станция тем самым одновременно сообщает о получении PDU 2, 3 и 4. Станция А поддерживает список порядковых номеров, разрешенных к пересылке; станция В создает список порядковых номеров PDU, которые она готова получить. Каждый из этих списков можно считать окном модулей PDU, а саму описанную операцию принято называть управлением потоком методом раздвижных окон (shding-window flow control).

Пришло время сделать несколько важных замечаний. Поскольку порядковый номер занимает место в PDU, очевидно, что его размер ограничен. Например, для 3-битового поля номера могут быть от 0 до 7. Т.е. PDU нумеруются по модулю 8 (после номера 7 следует 0). В общем случае для  $k$ -битового поля порядковые номера могут принадлежать интервалу от 0 до  $2^k - 1$ , а PDU нумеруются по модулю  $2^k$ .

На рис. 8.17 представлен процесс управления потоком с помощью раздвижных окон. Предполагается, что используется 3-битовый порядковый номер (PDU нумеруются от 0 до 7, после чего номера повторно используются для следующих PDU). Затененным прямоугольником обозначены те PDU, которые могут быть отправлены. В данном примере передатчик может отправить пять PDU, начиная с номера 0. После пересылки каждого PDU затененное окно уменьшается, после получения каждого подтверждения окно увеличивается. PDU, которые находятся между вертикальной линией и затененной областью, уже отосланы, но подтверждения о их получении еще нет. Как будет показано далее, передатчик должен сохранять в буфере резервные копии этих PDU на случай, если потребуются повторная передача.

Размер окна не обязательно должен быть максимальным для заданной длины порядковых номеров. Например, для 3-битового порядкового номера размер окна для станций, использующих протокол управления потоком методом раздвижных окон, может выбираться равным 4.

**Пример.** Иллюстрация к этому примеру приведена на рис. 8.18. Предполагается наличие 3-битового поля порядковых номеров и максимальный размер окна в 7 PDU. В начальном состоянии окна А и В разрешают А передать семь PDU, начиная с порядкового номера 0 (F0). После передачи трех PDU (F0, F1, F2) без получения подтверждения А уменьшает размер своего окна до 4 и сохраняет копии трех переданных PDU. Теперь А может передать еще четыре PDU, начиная с порядкового номера 3. После этого В передает кадр RR (receive ready — “готов к приему”) 3, что значит “Получены все PDU до номера 2 включительно, станция готова принять PDU номер 3; фактически она готова принять семь PDU, начиная с порядкового номера 3” После передачи этого

подтверждения А снова может передать семь PDU, на этот раз начиная с порядкового номера 3. А может также удалить хранящиеся в буфере PDU, получение которых было подтверждено. Станция А переходит к передаче PDU с номерами 3, 4, 5 и 6. Станция В возвращает сообщения RR 4, подтверждая получение F3 и разрешая передачу PDU, начиная с F4 до F2. Пока передается это подтверждение, А успевает переслать F4, F5 и F6, и теперь А может отправить четыре PDU, начиная с F7.

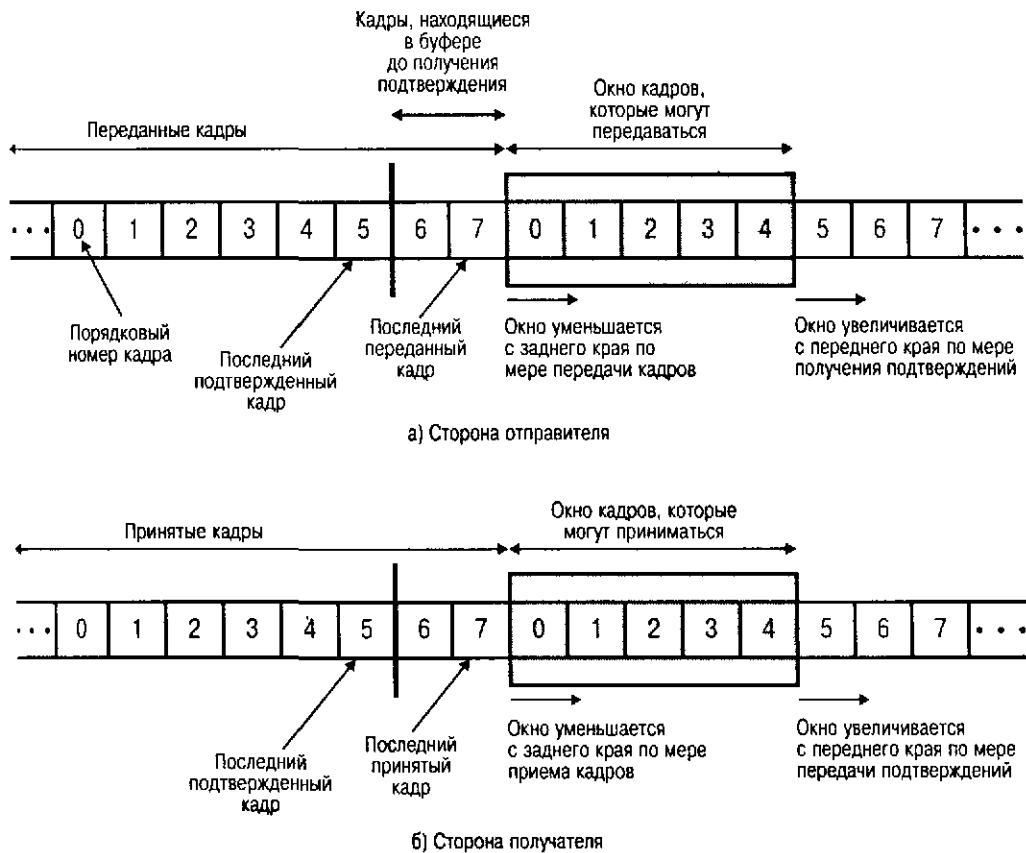


Рис. 8.17. Схема раздвижного окна

Описанный выше механизм представляет некоторую форму управления потоком: приемник должен быть способен принять семь PDU после подтверждения о принятии одного. Помимо этого, большая часть протоколов позволяет станции на другой стороне обрывать поток кадров, для чего посылается сообщение RNR (Receive Not Ready — “не готов к приему”), подтверждающее прием предшествующих кадров, но запрещающее передачу последующих. Иначе говоря, сообщение RNR 5 означает: “Приняты все кадры до четвертого включительно, но больше станция не может принять ни одного”. Для повторного открытия окна станция через некоторое время должна посылать обычное подтверждение приема.

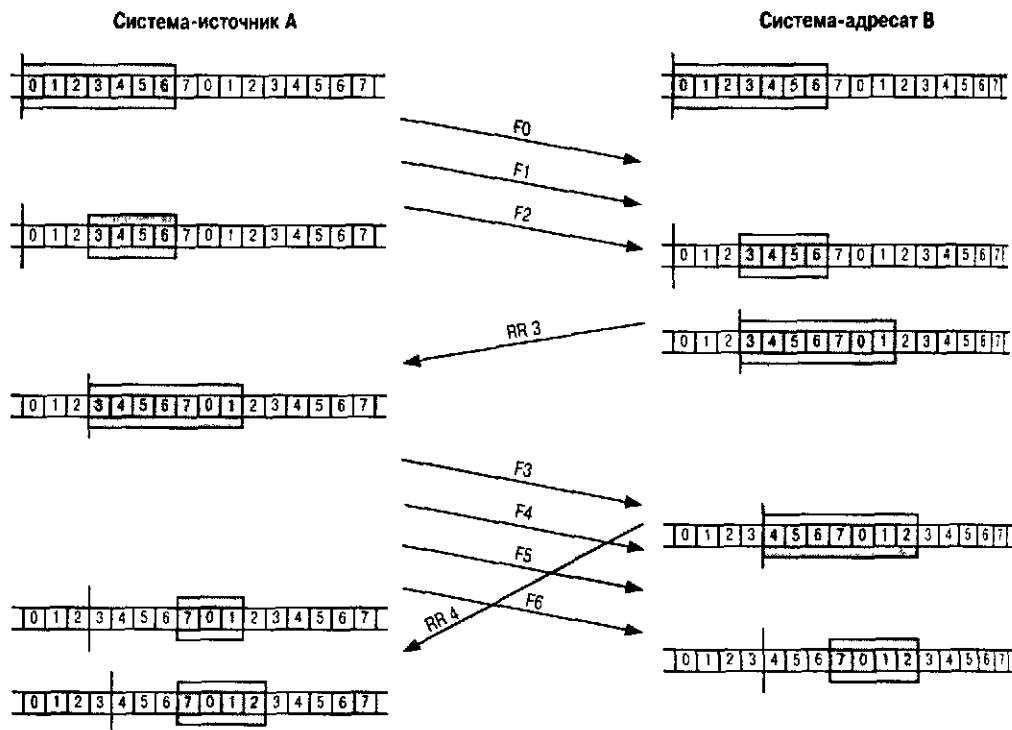


Рис. 8.18. Пример протокола раздвижных окон

Итак, до сих пор мы рассматривали передачу только в одном направлении. Если между двумя станциями происходит обмен данными, то каждая из них должна удерживать два окна, одно для передачи и второе для приема, кроме того, каждая сторона должна посылать другой данные и подтверждения приема. Для обеспечения эффективной поддержки этого требования обычно предусматривается функция, известная как *каскадное подтверждение* (piggybacking). Каждый *информационный PDU* включает поле, содержащее порядковый номер этого кадра, и поле, содержащее порядковый номер, используемый для подтверждения. Таким образом, если станции нужно передать данные и подтверждение приема, она передает их в одном кадре, экономя пропускную способность канала. Понятно, что если станции требуется передать только подтверждение приема, без данных, то она передает отдельный *PDU подтверждения*. Если же станции пора передавать только данные, то вместе с ними она должна повторно передать последнее отправленное подтверждение приема. Это делается потому, что кадр данных содержит поле для номера подтверждения, и в данное поле требуется занести какое-то значение. Когда станция получает такое дублирующее подтверждение, она его просто игнорирует.

## Защита от ошибок

Защитой от ошибок называются механизмы, позволяющие выявлять и исправлять ошибки, появляющиеся в процессе передачи *PDU*. Модель, которую будем использовать мы и которая соответствует типичному случаю, приведена на

рис. 8.16, б. Как и раньше, предполагаем, что данные передаются в виде последовательности PDU; PDU прибывают в том же порядке, в котором были посланы; каждый переданный PDU перед приемом задерживается на некоторое случайное и потенциально переменное время. Помимо этого, мы будем допускать возможность возникновения ошибок двух типов.

- **Потеря PDU.** PDU не доходит до адресата. Помеха может повредить PDU до такой степени, что приемник не определит, что PDU был передан.
- **Поврежденный PDU.** Получен PDU, поддающийся распознаванию, но имеющий несколько ошибочных битов (измененных в процессе передачи).

Самые распространенные методы защиты от ошибок включают некоторые или все следующие составляющие.

- **Выявление ошибок.** Приемник обнаруживает ошибки и отклоняет ошибочные PDU.
- **Положительное подтверждение приема.** Адресат возвращает положительное подтверждение для успешно принятых, безошибочных PDU.
- **Повторная передача после истечения времени ожидания.** Источник повторно передает PDU, подтверждение приема которого не было получено по прошествии некоторого предопределенного времени.
- **Отрицательное подтверждение приема и повторная передача.** Адресат возвращает отрицательное подтверждение приема для PDU, в которых были замечены ошибки. После этого источник повторно передает такие PDU.

Совокупность этих механизмов называется **автоматическим запросом повторной передачи** (Automatic Repeat Request — ARQ); результатом использования запроса ARQ является превращение ненадежного канала передачи данных в надежный. В настоящее время наиболее распространенным является **возвратный запрос ARQ** (go-back-N ARQ). Возвратный запрос ARQ основывается на управлении потоком методом раздвижных окон, описание которого приводится выше.

При использовании этого метода станция может передавать набор PDU, последовательно пронумерованных по модулю какой-то максимальной величины. Число находящихся в обращении неподтвержденных PDU определяется размером окна, для этого поток регулируется методом раздвижных окон. При отсутствии ошибок получатель в обычном порядке подтвердит получение входящих PDU, отправив сообщение RR (Receive Ready — “готов к приему”, или каскадное подтверждение). Если же станция-получатель обнаружит в PDU ошибку, то она пошлет отрицательное подтверждение REJ (reject — отклонение). В этом случае станция-получатель отклоняет данный PDU и все последующие входящие PDU, пока ошибочный бит не будет получен правильно. Следовательно, когда станция-отправитель получает сообщение REJ, она должна повторно передать ошибочный PDU плюс все последующие PDU, которые были переданы на данный момент.

Рассмотрим следующий случай: станция А передает PDU станции В. После каждой передачи станция А настраивает таймер подтверждения на только что переданный PDU. Предположим, что ранее станция В успешно приняла  $(i-1)$ -й PDU, а станция А только что передала  $i$ -й PDU. В методе возвратного запроса ARQ учтены приведенные ниже возможности. Рассмотрим три случая.



1. **Поврежденный PDU.** Если получен неверный PDU (т.е. станция В обнаружила ошибку), он отклоняется, и станция В не предпринимает никаких действий по отношению к этому PDU. Возможны два варианта.
  - а. В течение некоторого времени станция А посылает  $(i + 1)$ -й PDU. Станция В получает этот PDU несвоевременно и посылает станции А сообщение REJ  $i$ . В ответ на это станция А должна передать PDU  $i$  и все последующие PDU.
  - б. В течение некоторого времени станция А не присылает дополнительных PDU. Станция В не получает ничего и, соответственно, не посылает ни сообщения RR, ни сообщения REJ. По истечении времени, на которое установлен таймер станции А, станция передает PDU RR, содержащий Р-бит, значение которого установлено равным 1. Станция В интерпретирует PDU RR с таким Р-битом как команду, для подтверждения приема которой требуется отправить сообщение RR, указывающее следующий PDU, ожидаемый станцией В, т.е. в нашем случае PDU  $i$ . После того как станция А получает указанное сообщение, она повторно передает PDU  $i$ .
2. **Поврежденное сообщение RR.** Здесь также имеются два варианта.
  - а. Станция В получает  $i$ -й PDU и посылает сообщение RR  $(i + 1)$ , которое теряется в пути. Вследствие кумулятивности сообщений (т.е. наличие сообщения RR 6 означает, что были получены все PDU вплоть до пятого) может случиться так, что станция А получит следующее сообщение RR, подтверждающее получение следующего PDU, причем произойдет это до истечения предопределенного времени ожидания подтверждения получения PDU  $i$ .
  - б. Если же время ожидания подтверждения получения PDU  $i$  истекло, то станция А передает ту же команду RR, что и в варианте б случая 1. При этом станция А запускает другой таймер, называемый Р-битовым таймером. Если станция В не ответит на команду RR или если ответ будет поврежден, то время ожидания Р-битового таймера истечет. В этот момент станция А снова попытается передать команду RR и перезапустить Р-битовый таймер. Данная процедура будет повторяться до тех пор, пока не будет получено подтверждение приема или число повторов не превысит некоторого заданного максимального значения. В последнем случае будет инициализирована процедура перезапуска.
3. **Поврежденное сообщение REJ.** При потере сообщения REJ происходит то же, что и в варианте б случая 1.

Пример потока PDU с использованием возвратного запроса ARQ показан на рис. 8.19, а. Вследствие наличия в канале связи задержки распространения ко времени прибытия на станцию-источник подтверждения (положительного или отрицательного) эта станция успевает передать два дополнительных PDU, помимо того, подтверждение которого ожидается. Следовательно, когда прибывает PDU 6, станция В немедленно посылает сообщение REJ 5. Когда станция А получает это сообщение, ей приходится повторно передавать не только PDU 5, но и PDU 6 и 7. Отметим, что передатчик должен удерживать копии всех неподтвержденных PDU.

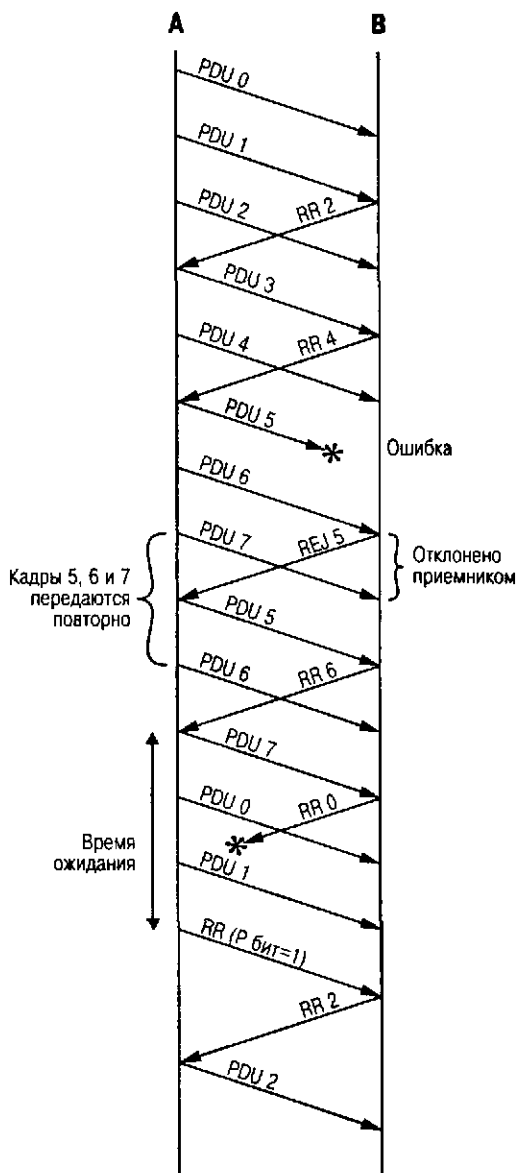


Рис. 8.19. Возвратный запрос ARQ

## 8.5. РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

Классическое описание кодов обнаружения ошибок и циклической проверки целостности с избыточностью (CRC) приводится в [ПЕТЕ61]. Хорошим введением в CRC является [РАМА88].

[АДАМ91] содержит исчерпывающий анализ кодов с коррекцией ошибок. В [КЛА01] имеется четкий, хорошо написанный раздел по данной теме. [БЕРЛ87]

и [BHAR83] представляют собой полезные обзорные статьи. В [ASH90] в доступной форме приводится математический анализ кодов с коррекцией ошибок. Подробный анализ турбокодов содержится в [SKLA97] и [BERR96]. Более детальное рассмотрение темы можно найти в [VUCE00].

- ADAM91 Adamek J. *Foundations of Coding*. — New York: Wiley, 1991.  
 ASH90 Ash R. *Information Theory*. — New York: Dover, 1990.  
 BERL87 Berlekamp E., Peile R., Pope S. The Application of Error Control to Communications. — *IEEE Communications Magazine*, April 1987.  
 BERR96 Berrou C., Glavieux A. Near Optimum Error Correcting Codes and Decoding Turbo Codes. — *IEEE Transactions on Communications*, October 1996.  
 BHAR83 Bhargava V. Forward Error Correction Schemes for Digital Communications. — *IEEE Communications Magazine*, January 1983.  
 PETE61 Peterson W., Brown D. Cyclic Codes for Error Detection. — In: *Proceedings of the IEEE*, January 1961.  
 RAMA88 Ramabadran T., Gaitonde S. A Tutorial on CRC Computations. — *IEEE Micro*, August 1988.  
 SKLA97 Sklar B. A Primer on Turbo Code Concepts. — *IEEE Communications Magazine*, December 1997.  
 SKLA01 Sklar B. *Digital Communications: Fundamentals and Applications*. — Upper Saddle River, NJ: Prentice Hall, 2001. // Склар Б. *Цифровая связь: теоретические основы и практическое применение*. — М.: Издательский дом "Вильямс", 2003.  
 VUCE00 Vucetic B., Yuan J. *Turbo Codes: Principles and Applications*. — Boston: Kluwer Academic Publishers, 2000.

## 8.6. ТЕРМИНЫ, ВОПРОСЫ И ЗАДАЧИ

### Основные термины

автоматический запрос повторной передачи (ARQ)	контрольная последовательность битов	решетчатая диаграмма сверточный код турбокод
бит четности	контрольная последовательность кадра	управление потоком данных
блочный код с коррекцией ошибок	модуль данных протокола (PDU)	управление потоком методом раздвижных окон
возвратный запрос ARQ зацита от ошибок	обнаружение ошибок	циклическая проверка четности с избыточностью
исправление ошибок	проверка четности	циклический код
код с коррекцией ошибок	прямое исправление ошибок	чередование блоков
код Хэмминга	расстояние Хэмминга	
словое слово		

### Вопросы

1. Дайте определение бита четности.
2. Что такое CRC?
3. Почему считается, что с помощью CRC можно обнаружить больше ошибок, чем при использовании проверки четности?

4. Перечислите три разных способа описания алгоритма CRC.
5. Можно ли создать код с коррекцией ошибок, который бы исправлял некоторые (но не все) 2-битовые ошибки? Ответ аргументируйте.
6. Что представляют параметры  $n$  и  $k$  в блочном коде с коррекцией ошибок ( $n, k$ )?
7. Что представляют параметры  $n, k$  и  $K$  в сверточном коде ( $n, k, K$ )?
8. Что представляет собой решетчатая диаграмма сверточного кода?
9. Назовите два ключевых элемента защиты от ошибок.
10. Объясните принцип работы возвратного запроса ARQ.

## Задачи

1. Почему при вычислении контрольной последовательности кадра используется не обычная арифметика, а арифметика по модулю 2?
2. Рассмотрим кадр, состоящий из двух 4-битовых знаков. Предположим, что вероятность битовой ошибки независима для каждого из битов и равна  $10^{-3}$ .
  - а. Найдите вероятность того, что полученный кадр содержит по крайней мере одну ошибку.
  - б. Прибавьте к каждому знаку бит проверки четности. Какой теперь будет вероятность появления ошибки?
3. Используя полином CRC-МККТТ, найдите 16-битовый код CRC для сообщения, первый бит которого равен 1, а все последующие — 0.
  - а. Используя длинное деление.
  - б. Используя механизм регистра сдвига, приведенный на рис. 8.3.
4. Объясните, почему при отсутствии ошибок регистр сдвига CRC будет содержать только нули. Проиллюстрируйте на примере.
5. Найдите CRC для  $P = 110011$  и  $M = 11100011$ .
6. Для создания 4-битовой контрольной последовательности 11-битового сообщения генерируется CRC. Порождающий многочлен равен  $X^4 + X^3 + 1$ .
  - а. Изобразите схему регистра сдвига для выполнения поставленной задачи (см. рис. 8.4).
  - б. Закодируйте последовательность битов данных 10011011100 (крайний левый бит является младшим), используя порождающий многочлен. Найдите кодовое слово.
  - в. Предположив, что бит 7 (считая от младшего бита) кодового слова является ошибочным, покажите, что алгоритм обнаружит данную ошибку.
7. В стандартах связи, таких, как HDLC, широко используется модифицированная процедура CRC, которая определяется следующим образом:

$$\frac{X^{16}M(X) + X^kL(X)}{P(X)} = Q + \frac{R(X)}{P(X)},$$

контрольная последовательность кадра =  $L(X) + R(X)$ ,

где

$$L(X) = X^{15} + X^{14} + X^{13} + \dots + X + 1,$$

$k$  — количество проверяемых битов (поля адресации, управления и информации).

- а. Опишите результат выполнения данной процедуры.
  - б. Перечислите потенциальные преимущества процедуры.
  - в. Продемонстрируйте реализацию регистра сдвига для  $P(X) = X^{16} + X^{12} + X^5 + 1$ .
8. На рис. 8.20 представлен контур полиномиального деления, дающий тот же результат, что и контур, приведенный на рис. 8.4.

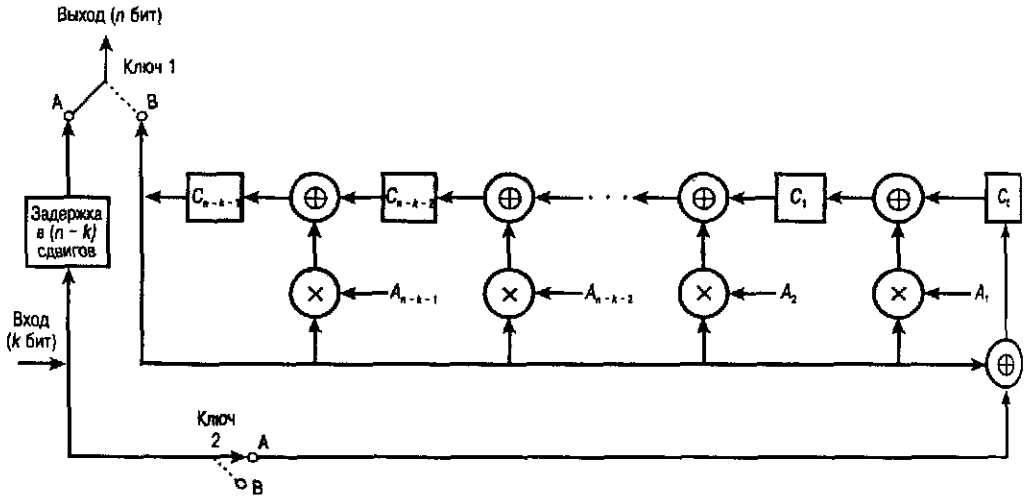


Рис. 8.20. Альтернативная архитектура CRC для делителя  $1 + A_1X + A_2X^2 + \dots + A_{n-k-1}X^{n-k-1} + X^{n-k}$

- а. Используя предложенную структуру, изобразите регистр LFSR для деления на  $X^6 + X^5 + X^2 + 1$ , эквивалентный тому, что приведен на рис. 8.3, а.
  - б. Покажите последовательность шагов, позволяющих получить результирующую CRC, используя метод, представленный на рис. 8.3. В конце работы регистр должен содержать начальное значение 01110. Подсказка: помните, что между входом и выходом существует задержка  $n - k$ , поэтому для получения окончательного результата нужно  $(n - k)$  дополнительных сдвигов.
  - в. Преимущество структуры, представленной на рис. 8.20, перед контуром, изображенным на рис. 8.4, состоит в том, что в первом случае явно видна реализация процесса длинного деления. Почему?
  - г. Какой недостаток присущ контуру, изображенному на рис. 8.20?
9. Найти попарные расстояния Хэмминга для следующих кодовых слов.
- а. 00000, 10101, 01010.
  - б. 000000, 010101, 101010, 110110.
10. В разделе 8.2 были рассмотрены блочные коды с коррекцией ошибок, при использовании которых решения принимались на основе минимального расстояния. Т.е. если код состоит из  $s$  равновероятных кодовых слов длиной  $n$ , для каждой полученной последовательности  $\mathbf{v}$  приемник выбирает кодовое слово  $\mathbf{w}$ , для которого расстояние  $d(\mathbf{w}, \mathbf{v})$  будет минимальным. Необходимо доказать, что эта схема является "идеальной" в том смысле, что приемник всегда выбирает кодовое слово, для которого максимальна вероятность выбора  $\mathbf{w}$  при данном  $\mathbf{v}$ ,  $p(\mathbf{w} | \mathbf{v})$ . Поскольку все кодовые слова считаются равновероятными, то для кодового слова с максимальным значением  $p(\mathbf{w} | \mathbf{v})$  будет максимальной и вероятность  $p(\mathbf{v} | \mathbf{w})$ .

- а. Поскольку  $w$  будет получено в виде  $v$ , в процессе передачи должно произойти точно  $d(w, v)$  ошибок, причем ошибочные биты должны появиться в местах несовпадения  $w$  и  $v$ . Пусть  $\beta$  — вероятность того, что заданный бит будет передан ошибочно, а  $n$  — длина кодового слова. Выразите  $p(v | w)$  как функцию  $\beta$ ,  $d(w, v)$  и  $n$ . *Подсказка:* количество ошибочных битов равно  $d(w, v)$ , а количество правильно переданных битов —  $n - d(w, v)$ .
- б. Сравните  $p(v | w_1)$  и  $p(v | w_2)$  двух различных кодовых слов  $w_1$  и  $w_2$ , вычислив отношение  $p(v | w_1)/p(v | w_2)$ .
- в. Приняв  $0 < \beta < 0,5$ , покажите, что  $p(v | w_1) > p(v | w_2)$  тогда и только тогда, когда  $d(v | w_1) > d(v | w_2)$ . Из этого следует, в частности, что кодовое слово  $w$ , которому соответствует максимальное значение  $p(v | w)$ , является словом, расстояние от которого до  $v$  минимально.
11. В разделе 8.2 утверждается, что если для заданного целого положительного числа  $t$  код удовлетворяет условию  $d_{\min} \geq 2t + 1$ , то использование такого кода позволяет исправить все кадры, имеющие до  $t$  ошибочных битов включительно. Докажите это утверждение. *Подсказка:* начните со следующего наблюдения — чтобы кодовое слово  $w$  было декодировано, как другое кодовое слово  $w'$ , полученная последовательность по крайней мере должна быть так же близка к  $w'$ , как к  $w$ .
12. Для кода Хэмминга, представленного в табл. 8.2, запишите формулы вычисления контрольных битов как функции битов данных.
13. Для кода Хэмминга, представленного в табл. 8.2, покажите, что произойдет, если ошибочным будет не бит данных, а контрольный бит.
14. Рассмотрим хранящееся в памяти системы 8-битовое слово 11000010. Используя алгоритм Хэмминга, определите, какие контрольные биты должны храниться в памяти вместе с данными. Покажите, каким образом был получен ответ.
15. Для 8-битового слова 00111001 в памяти должны храниться контрольные биты 0111. Предположим, что после считывания слова вычисленные значения контрольных битов составляют 1101. Какая последовательность данных была считана из памяти?
16. Сколько контрольных битов необходимо использовать для кода с коррекцией ошибок Хэмминга, чтобы обнаружить 1-битовые ошибки в 1024-битовом слове?
17. Выполните деление  $f(X) = X^6 + 1$  на  $g(X) = X^4 + X^3 + X + 1$ . Проверьте полученный результат, умножив его на  $g(X)$ . У вас должно получиться  $f(X)$ .
18. Рассмотрите пример, приведенный в табл. 8.3.
- а. Изобразите регистр LFSR.
- б. Используя схему, подобную той, которая представлена на рис. 8.3, б, покажите, что контрольные биты блока данных 1010 имеют значение 001.
19. Используя структуру чередования  $n = 4$ ,  $m = 6$  (рис. 8.8), докажите, что имеют место следующие свойства.
- а. Любой пакет из  $m$  последовательных ошибок даст в результате отдельные (разделенные промежутками по меньшей мере  $n$  бит) ошибки на выходе устройства восстановления исходного порядка битов.
- б. Любой пакет из  $bm$  последовательных ошибок ( $b > 1$ ) даст в результате пакеты на выходе устройства восстановления исходного порядка битов, состоящие не более чем из  $\lceil b \rceil$  ошибок. Каждый пакет на выходе отделен от других пакетов по меньшей мере  $\lfloor b \rfloor$  битами. Запись  $\lceil b \rceil$  обозначает наименьшее целое число, не меньше  $b$ ;  $\lfloor b \rfloor$  — наибольшее целое, не больше  $b$ .

- в. Периодическая последовательность 1-битовых ошибок, разделенная промежутками в  $m$  бит, приведет к появлению одного пакета ошибок длиной  $m$  бит на выходе устройства восстановления исходного порядка битов.
- г. Если пренебречь временем задержки распространения в канале, задержка между входом и выходом устройства чередования составит  $2(n(m-1)+1)$ . Для начала процесса передачи достаточно заполнения  $(n(m-1)+1)$  ячеек, такого же числа ячеек достаточно для начала обращения чередования в приемнике.
20. Рассмотрим сверточный кодер с параметрами  $(v_{n1} = u_n \oplus u_{n-2})$  и  $(v_{n2} = u_{n-1} \oplus u_{n-2})$ .
- Изобразите для данного кодера реализацию регистра сдвига, подобную той, что приведена на рис. 8.9, а.
  - Изобразите для данного кодера диаграмму состояний, подобную той, что приведена на рис. 8.9, б.
  - Изобразите для данного кодера решетчатую диаграмму, подобную диаграмме, приведенной на рис. 8.9, в.
21. Предположим, что для кодера, описанного в задаче 20, начальные значения ячеек регистра сдвига равны 0, а после передачи последнего бита информации передаются еще два бита, равные 0.
- Для чего используются дополнительные два бита?
  - Найдите закодированную последовательность, соответствующую последовательности данных 1101011. Считайте, что первым в кодер вводится крайний левый бит.
22. Простейшая форма управления потоком, известная как управление потоком с остановками (stop-and-wait flow control), реализуется следующим образом. Источник передает кадр. После получения данного кадра приемник передает подтверждение, указывающее на желание принять следующий кадр. Источник должен ожидать прибытия подтверждения и лишь затем начинать следующую передачу. Таким образом, адресат может остановить поток данных, просто не отправив подтверждения получения очередного кадра. Рассмотрим полудуплексный двухточечный канал связи с использованием схемы с остановками, когда передается последовательность сообщений, причем каждое сообщение разбивается на несколько кадров. Ошибки и служебные издержки не учитываются.
- Как увеличение размера сообщений (т.е. уменьшение их общего числа) скажется на использовании канала связи? Остальные параметры считайте фиксированными.
  - Как увеличение числа кадров при фиксированном размере сообщения скажется на использовании канала связи?
  - Как увеличение размера кадров скажется на использовании канала связи?
23. Рассмотрите рис. 8.21. На узле А создаются кадры и посылаются узлу С через узел В. Определите минимальную скорость передачи между узлами В и С, не вызывающую переполнения буферов узла В. Используйте следующую информацию.

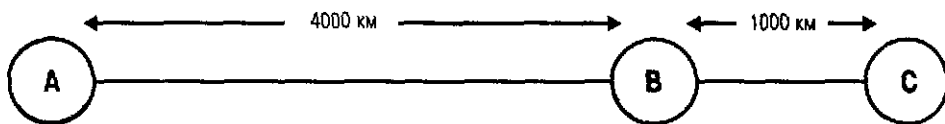


Рис. 8.21. Иллюстрация к задаче 8.23

- Скорость передачи данных между узлами А и В равна 100 Кбит/с.
- Задержка распространения для обеих линий составляет 5 мкс/км.
- Линии между узлами являются дуплексными.
- Размер всех информационных кадров равен 1000 бит; размером кадров подтверждения приема можно пренебречь.
- Между узлами А и В используется протокол раздвижных окон с размером окна 3.
- Между узлами В и С используется процедура с остановками.
- Ошибки отсутствуют.

*Подсказка:* чтобы не вызвать переполнения буфера узла В, среднее число кадров, поступающих на узел и покидающих его, должно оставаться неизменным в течение длительного интервала времени.

24. Скорость передачи данных в канале равна  $R$  бит/с, а задержка распространения составляет  $t$  секунд на километр. Расстояние между передающим и принимающим узлами равно  $L$  км. Между узлами идет обмен кадрами фиксированного размера  $B$  бит. Выведите формулу минимального размера поля порядкового номера кадра в зависимости от  $R$ ,  $t$ ,  $B$  и  $L$  (для этого рассмотрите максимальное использование канала). Размер кадров АСК считайте пренебрежимо малым, а процесс обработки на узлах — мгновенным.
25. Два соседних узла (А и В) используют протокол раздвижных окон с 3-битовым порядковым номером. Используется возвратный механизм ARQ с размером окна 4. Предполагая, что передачу ведет станция А, а прием осуществляет станция В, покажите позиции окна в следующие моменты времени.
- а. До передачи кадров станцией А.
  - б. После передачи станцией А кадров 0, 1, 2 и подтверждения станцией В кадров 0, 1, причем команды АСК уже получены станцией А.
  - в. После передачи станцией А кадров 3, 4, 5 и подтверждения станцией В кадра 4, причем команда АСК уже получена станцией А.
26. Две станции сообщаются посредством спутникового канала связи со скоростью передачи 1 Мбит/с и задержкой распространения 270 мс. Спутник служит исключительно для ретрансляции данных, получаемых с одной станции на другую, с пренебрежимо малой коммутационной задержкой. Чему равна максимальная пропускная способность кадров HDLC размером 1024 бит с 3-битовым порядковым номером, т.е. какова пропускная способность информационных битов, передаваемых в кадрах HDLC?



**Часть III**

# **ОРГАНИЗАЦИЯ БЕСПРОВОДНЫХ СЕТЕЙ**

**Глава 9. Спутниковая связь**

**Глава 10. Беспроводные сотовые сети**

**Глава 11. Беспроводные системы и беспроводные  
абонентские линии**

**Глава 12. Протоколы Mobile IP и WAP**

# ГЛАВА 9

## СПУТНИКОВАЯ СВЯЗЬ

### 9.1. Параметры и конфигурации спутника

Спутниковые орбиты

Полосы частот

Ухудшение качества связи

Конфигурации спутниковой сети

### 9.2. Распределение пропускной способности — частотное разделение

Уплотнение с частотным разделением (FDM)

Множественный доступ с частотным разделением  
(FDMA)

### 9.3. Распределение пропускной способности — временное разделение

### 9.4. Рекомендуемая литература

### 9.5. Термины, вопросы и задачи

**З**начение спутниковой связи для развития телекоммуникаций и средств передачи данных трудно переоценить. Появление спутниковой связи здесь сыграло не менее важную роль, чем внедрение оптоволокна. В первом разделе этой главы приводится обзор ключевых понятий и параметров беспроводной связи, связанных с используемыми здесь спутниковыми антеннами. Оставшаяся часть главы посвящена механизмам распределения пропускной способности.

## 9.1. ПАРАМЕТРЫ И КОНФИГУРАЦИИ СПУТНИКА

Сердцем системы спутниковой связи является антенна спутника, находящегося на стационарной орбите. С помощью одного или нескольких таких спутников, используемых как космические ретрансляторы, осуществляется связь между двумя или несколькими станциями, принадлежащими одной системе спутниковой связи и расположенными на Земле или близ Земли. Системы антенн, расположенных на Земле или близ Земли, называют **наземными станциями**. Канал передачи данных с наземной станции на спутник называется **восходящим (uplink)**, а канал передачи данных в обратном направлении — **нисходящим (downlink)**. Электронное оборудование спутника, которое принимает сигналы восходящего канала и преобразует их в сигналы нисходящего канала, называется **транспондером**.

Спутники связи можно классифицировать по таким признакам.

- **Зона обслуживания:** может быть глобальной, региональной или национальной. Чем больше зона обслуживания, тем больше спутников будет задействовано в организации сети.
- **Тип услуг:** существуют спутники стационарной службы связи (fixed service satellite — FSS), радиовещательной службы (broadcast service satellite — BSS) и мобильной службы (mobile service satellite — MSS). В этой главе будут рассматриваться преимущественно спутники FSS и BSS.
- **Характер использования:** коммерческие, военные, любительские или экспериментальные.

При проектировании станций беспроводной связи следует учитывать те многочисленные особенности, которые отличают станции, базирующиеся на спутниках, от наземных станций.

- Зона обслуживания спутниковой системы намного превышает зону обслуживания наземной системы. Для одной антенны спутника, находящегося на геостационарной орбите, доступно около одной четвертой поверхности Земли.
- Такие ресурсы космического аппарата, как мощность и выделенная ширина полосы, весьма ограничены. Поэтому на стадии проектирования нужно выбирать оптимальное соотношение между параметрами наземной станции и спутника.
- Условия, в которых находятся общающиеся спутники, не так меняются со временем, как условия связи спутника с наземной станцией или же как условия связи двух наземных беспроводных антенн. Поэтому каналы связи спутник-спутник можно рассчитать с довольно высокой степенью точности.

- Если передатчик и приемник находятся в зоне обслуживания одного спутника, то затраты на передачу данных не зависят от расстояния между ними.
- Легко внедряются широковещательные, многоадресные и двухточечные приложения.
- Пользователю доступен очень широкий диапазон частот или высокая скорость передачи данных.
- В целом качество данных, передаваемых с помощью спутника, поддерживается исключительно высоким, несмотря на кратковременные отключения или ухудшение качества связи.
- Для спутников, находящихся на геостационарной орбите, задержка распространения сигнала с земли на спутник и обратно равна примерно одной четвертой секунды.
- Передающие наземные станции во многих случаях могут принимать и собственные сигналы.

## Спутниковые орбиты

Существует несколько классификаций орбит спутников.

1. Орбита может быть круговой, с центром окружности, расположенном в центре Земли, или эллиптической, для которой центр Земли находится в одном из фокусов эллипса.
2. Спутник может вращаться вокруг Земли в разных плоскостях. **Экваториальные орбиты** расположены в плоскости земного экватора. **Полярные орбиты** проходят над обоими полюсами Земли. Все остальные орбиты называются **наклонными**.
3. По высоте над уровнем моря орбиты классифицируются следующим образом: геостационарные околоземные орбиты (geostationary earth orbit — GEO), средние околоземные орбиты (medium earth orbit — MEO) и низкие околоземные орбиты (low earth orbit — LEO). Подробнее орбиты всех типов рассмотрены позднее.

### Расстояние

На рис. 9.1 приведена геометрическая схема, в соответствии с которой рассчитывается зона обслуживания спутника. Ключевым параметром данной схемы является угол **возвышения** наземной станции  $\theta$ , который является углом между горизонтальной линией (т.е. линией, касательной к поверхности Земли в точке расположения антенны) и направлением основного луча антенны, нацеленного непосредственно на спутник. Максимальная величина зоны обслуживания спутника получается при нулевом угле возвышения. Тогда зона обслуживания во всех направлениях ограничивается только оптическим горизонтом спутника. Однако существуют по крайней мере три проблемы, которые не позволяют строить антенны наземных станций с нулевым углом возвышения и определяют **минимальный угол возвышения** [INGL97].

1. Нельзя игнорировать здания, деревья и другие наземные объекты, которые могут находиться на пути луча, идущего от антенны. Подобные помехи приводят к ослаблению сигнала, так как происходит его частичное поглощение или искажение вследствие многократного отражения луча.

2. Чем меньше значение угла возвышения, тем больше атмосферное поглощение, так как при малых углах возвышения сигнал проходит в атмосфере большее расстояние.
3. На качество приема также неблагоприятно влияет электрический шум, обусловленный высокой температурой около поверхности Земли.

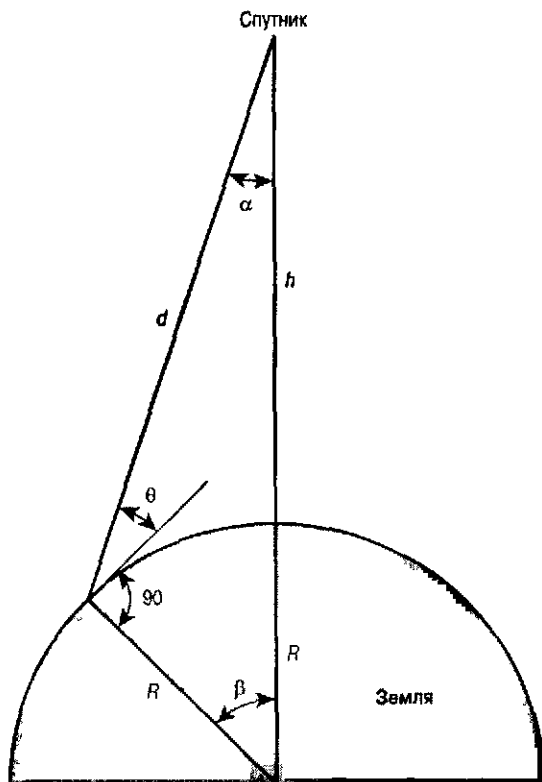


Рис. 9.1. Охват и угол возвышения

При проектировании нисходящего канала принято использовать минимальный угол возвышения, который, в зависимости от частоты сигнала, может составлять  $5\text{--}20^\circ$ . Для восходящего канала Федеральная комиссия по средствам связи США предписывает использовать угол возвышения не менее  $5^\circ$ .

Угол охвата  $\beta$  — это мера части земной поверхности, которая видна со спутника с учетом минимального угла возвышения. Угол  $\beta$  задает окружность на поверхности Земли, центр которой находится в точке, расположенной непосредственно под спутником. Справедливо следующее соотношение<sup>1</sup>:

<sup>1</sup> В первом выражении используется теорема синусов, которая гласит, что стороны любого треугольника пропорциональны синусам противоположных углов. Во втором выражении используется тот факт, что сумма углов треугольника равна  $\pi$ . В третьем — тригонометрическое тождество  $\sin(x) = \cos(x - \pi/2)$ .

$$\frac{R}{R+h} = \frac{\sin(\alpha)}{\sin(\theta + \frac{\pi}{2})} = \frac{\sin(\frac{\pi}{2} - \beta - \theta)}{\sin(\theta + \frac{\pi}{2})} = \frac{\cos(\beta + \theta)}{\cos(\theta)}.$$

Здесь

$R$  — радиус Земли, который равен 6370 км;

$h$  — высота орбиты (т.е. кратчайшее расстояние от спутника до Земли);

$\beta$  — угол охвата;

$\theta$  — минимальный угол возвышения.

Расстояние от спутника до ближайшей к нему точки зоны обслуживания рассчитывается по следующей формуле:

$$\frac{d}{R+h} = \frac{\sin(\beta)}{\sin(\theta + \frac{\pi}{2})} = \frac{\sin(\beta)}{\cos(\theta)}, \quad (9.1)$$

$$d = \frac{(R+h)\sin(\beta)}{\cos(\theta)} = \frac{R\sin(\beta)}{\sin(\alpha)}$$

Следовательно, задержка кругового обращения сигнала принадлежит такому диапазону:

$$\frac{2h}{c} \leq t \leq \frac{2(R+h)\sin(\beta)}{c(\cos\theta)},$$

где  $c$  — скорость света, которая равна примерно  $3 \times 10^8$  м/с.

Размер зоны обслуживания спутника обычно выражается через диаметр покрываемой области, равный  $2\beta R$ , где  $\beta$  выражается в радианах.

На рис. 9.2 показаны период обращения, зона обслуживания (выраженная через радиус охваченной области земной поверхности), а также максимальная задержка кругового обращения сигнала.

## Геостационарные спутники

В настоящее время самыми распространенными среди спутников связи являются геостационарные (ГЕО). Любопытно, что первым, кто выдвинул идею использования таких спутников, был писатель-фантаст Артур Кларк, пославший в 1945 году статью на эту тему в журнал *Wireless World*. Идея заключалась в следующем: если спутник находится на круговой орбите на высоте 35 863 км над поверхностью Земли и движется в плоскости земного экватора, то угловая частота вращения такого спутника будет совпадать с угловой частотой вращения Земли и спутник все время будет находиться над одной и той же точкой на экваторе<sup>2</sup>. Такая орбита изображена на рис. 9.3 с соблюдением масштаба. На орбите показано

<sup>2</sup> Вместо термина *геостационарный* часто используется термин *геосинхронный*. Дотошный читатель все же почувствует разницу, так как геосинхронной можно назвать любую круговую орбиту на высоте 35 863 км, а геостационарной можно назвать только ту из геосинхронных орбит, которая имеет нулевое склонение, что позволяет спутнику зависать над определенным участком экватора.

множество символов, соответствующих спутникам, что также отражает реальную картину. На геостационарной орбите действительно находится довольно много спутников, некоторые из них расположены вплотную друг к другу.

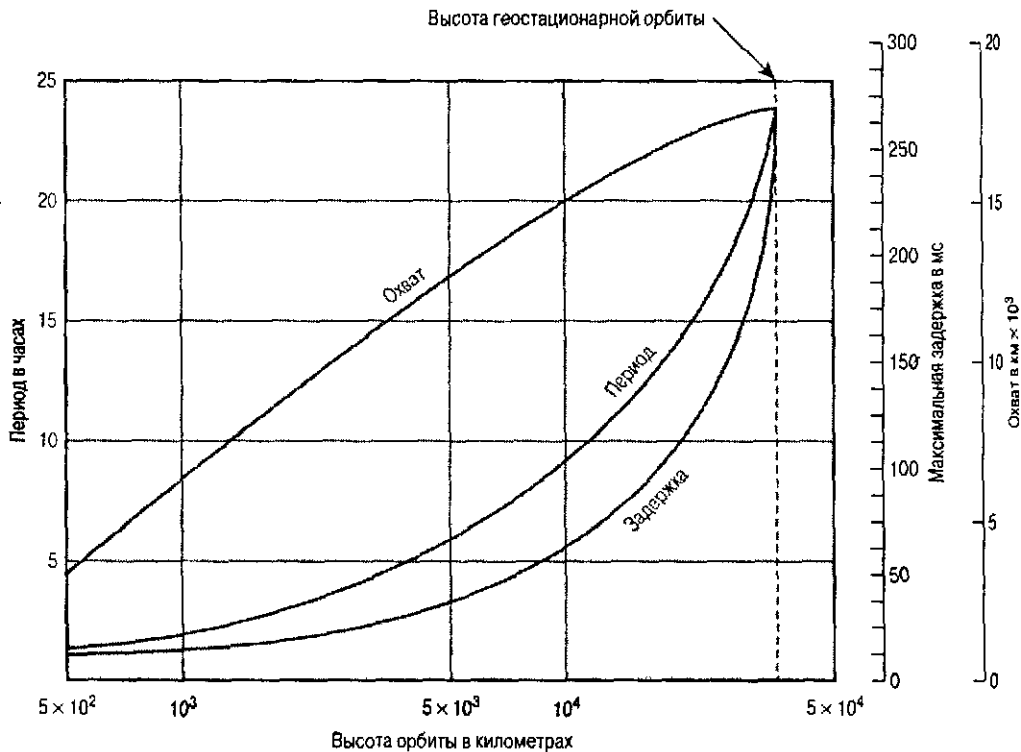


Рис. 9.2. Параметры спутника как функция высоты орбиты

Геостационарные орбиты имеют ряд преимуществ, которые выгодно отличают их от орбит других типов.

- Так как спутник не движется относительно Земли, то не возникает проблем с изменением частоты сигнала из-за относительного движения спутника и наземных антенн (обусловленного эффектом Доплера).
- Упрощается процедура отслеживания спутника с наземных станций.
- Спутник, находящийся на высоте 35 863 км над Землей, может связаться примерно с четвертой частью земной поверхности. Для того чтобы покрыть все населенные зоны Земли, исключая участки близ северного и южного полюсов, понадобится вывести на геостационарную орбиту всего три спутника, расположив их на расстоянии  $120^\circ$  друг от друга.

С другой стороны, есть и недостатки.

- После прохождения расстояния свыше 35 000 км сигнал может стать довольно слабым.
- Полярные области и приполярные участки северного и южного полушарий практически недоступны для геостационарных спутников.

- Несмотря на то что скорость света равна 300 000 км/с, задержка прохождения сигнала из точки на экваторе, расположенной под спутником, на спутник и обратно довольно существенна.

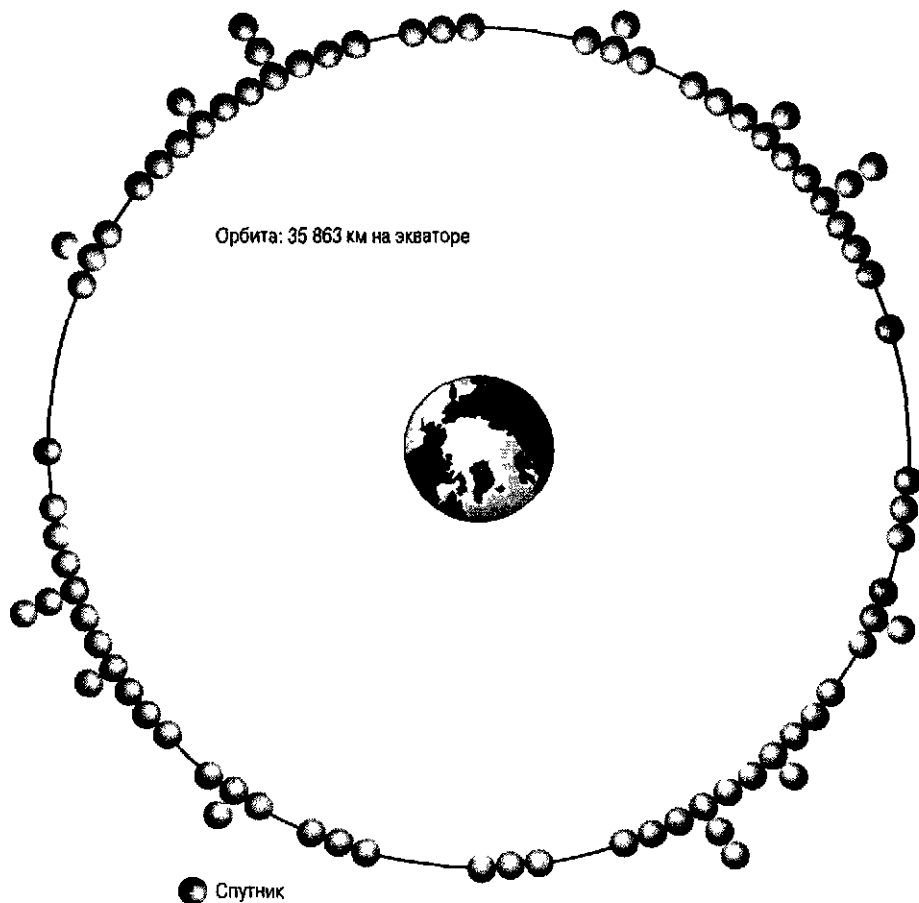


Рис. 9.3. Геостационарная орбита (GEO)

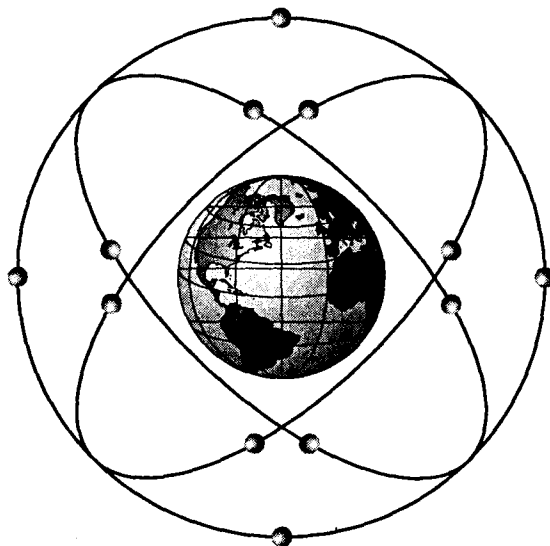
Задержка связи между двумя наземными абонентами, расположенными непосредственно под спутником, в действительности составляет  $(2 \times 35\,863) / 300\,000 = 0,24$  с. Для других абонентов, не находящихся непосредственно под спутником, задержка связи еще больше. Если использовать спутниковую связь для телефонных переговоров, то суммарный интервал между завершением фразы одного абонента и ответом другого удваивается и получается равным примерно половине секунды. Такую задержку трудно не заметить. Еще одним свойством геостационарных спутников является то, что выделенные для них частоты используются над очень большим участком поверхности Земли. Для многоточечных приложений, таких, как передача телевизионных программ, это даже к лучшему, однако для двухточечной связи спектр частот геостационарного спутника будет использоваться очень неэкономно. Можно использовать специальные антенны с направленным или



управляемым лучом, которые позволяют ограничить область, покрываемую сигналом спутника, и таким образом контролировать размеры области, в пределах которой принимается сигнал спутника. Для решения некоторых из перечисленных проблем были разработаны другие типы орбит. Для персональных устройств связи третьего поколения важную роль играют *спутники на низких околоземных орбитах (LEOS)* и *спутники на средних околоземных орбитах (MEOS)*.



а) Околоземная орбита: часто — полярная орбита на высоте 500–1500 км



б) Средняя околоземная орбита: наклонена к экватору, высота 5000–12 000 км

Рис. 9.4. Орбиты низких и средних околоземных спутников

## Спутники LEO

Спутники LEO (рис. 9.4, а) имеют следующие характеристики.

- Круговые или эллиптические орбиты на высоте до 2000 км. Все предлагаемые и реальные системы располагаются на высоте от 500 до 1500 км.
- Период орбиты равен 1,5–2 ч.
- Диаметр зоны обслуживания равен примерно 8000 км.
- Задержка кругового распространения сигнала составляет не более 20 мс.
- Максимальное время, в течение которого спутник виден с фиксированной точки на поверхности Земли (в пределах радиогоризонта), достигает 20 мин.
- Ввиду высокой скорости относительного движения спутника и стационарного расположения наземной станции оборудование системы связи должно быть способно учитывать большие доплеровские сдвиги, которым подвергается частота сигнала.
- Для спутников LEO велико сопротивление атмосферы, поэтому орбита спутника постепенно деформируется.

Чтобы ввести в действие такую систему связи, нужно довольно много орбитальных плоскостей; кроме того, на каждой орбите должны находиться по несколько спутников. Тогда при сообщении двух наземных станций сигнал, как правило, будет передаваться с одного спутника на другой.

Спутники LEO имеют ряд преимуществ по сравнению с геостационарными спутниками. Кроме уже упомянутого сокращения задержки распространения сигнала, принимаемый сигнал, отправленный со спутника LEO, гораздо сильнее сигнала со спутника GEO при той же энергии передачи. Зону обслуживания спутника LEO можно локализовать с гораздо большей степенью точности, так что можно эффективнее распорядиться спектром частот, выделенным для спутника LEO. Именно поэтому данная технология сейчас предложена для связи с мобильными и персональными терминалами, для функционирования которых нужны более сильные сигналы. С другой стороны, чтобы 24 часа обеспечивать широкую зону обслуживания, нужно много спутников LEO.

Было выдвинуто несколько коммерческих предложений по использованию кластеров спутников LEO для предоставления услуг связи. Эти предложения можно разделить на две категории.

- **Малые кластеры LEO.** Такие кластеры, предназначенные для работы при частоте связи ниже 1 ГГц, используют полосу частот шириной не более 5 МГц и обеспечивают скорость передачи данных до 10 Кбит/с. Эти системы предназначены для поиска, слежения и низкоскоростного обмена сообщениями. Примером такой спутниковой системы является система Orbcomm, которая стала первой из малых действующих систем LEO: ее первые два спутника были запущены в апреле 1995 года. Система Orbcomm разрабатывалась для поиска и пакетной передачи данных и предназначалась для работы с небольшими пакетами данных длиной 6–250 байт. Она используется для отслеживания передвижения прицепов, железнодорожных составов, тяжелого оборудования и других удаленных и мобильных активов. Такую систему можно также использовать для наблюдения за удаленными приборами, резервуарами-хранилищами нефти и газа, скважинами и трубопроводами или для поддержки связи с рабо-

чими, в какой бы точке земного шара они не находились. Для передачи данных на спутники этой системы используются частоты от 148,00 до 150,05 МГц, а для передачи сигналов со спутников — частоты от 137,00 до 138,00 МГц. В распоряжении системы Orbcomm находятся более 30 спутников на низких околоземных орбитах. Она поддерживает абонентскую скорость передачи данных на спутник, равную 2,4 Кбит/с, и со спутника — 4,8 Кбит/с.

- **Большие кластеры LEO.** Такие системы работают на частотах более 1 ГГц и поддерживают скорость передачи данных до нескольких мегабайтов в секунду. Эти системы стремятся предоставлять те же услуги, что и малые кластеры LEO, а также дополнительные услуги по передаче голоса и по определению местоположения. Примером большой системы LEO является система Globalstar. Ее спутники довольно примитивны. В отличие от некоторых малых систем LEO, на спутниках этой системы не установлено оборудования для обработки данных или связи между спутниками. Обработка данных большей частью проводится наземными станциями системы. В системе используется технология CDMA, подобная стандарту CDMA сотовой связи. Нисходящая связь с мобильными пользователями осуществляется в диапазоне S-полосы (около 2 ГГц). Система Globalstar тесно интегрирована с традиционными звуковыми службами-носителями. Все звонки должны обрабатываться на наземных станциях. Группа спутников системы Globalstar состоит из 48 действующих спутников и 8 запасных. Эти спутники находятся на орбитах на высоте 1413 км.

### Спутники MEO

Спутники MEO (рис. 9.4, б) имеют следующие характеристики.

- Круговая орбита, расположенная на высоте 5000–12 000 км.
- Период орбиты около 6 ч.
- Диаметр зоны обслуживания колеблется от 10 000 до 15 000 км.
- Задержка кругового распространения сигнала составляет менее 50 мс.
- Максимальное время, в течение которого спутник виден из фиксированной точки на земной поверхности (находящейся в пределах радиогоризонта), составляет несколько часов.

В системах MEO не требуется так много переключений между спутниками, как в системах LEO. Значения таких параметров спутника MEO, как задержка распространения сигнала со спутника MEO на Землю и его требуемая мощность, выше, чем у спутников LEO, однако существенно меньше, чем у геостационарных спутников. Для новой программы ICO, разработанной в январе 1995 года, была предложена система MEO. Запуски спутников MEO начались в 2000 году. На орбиты высотой 10 400 км планировалось поместить двенадцать спутников (из них два запасных). Спутники должны быть равномерно распределены между двумя плоскостями и иметь угол наклона к экватору в 45°. Системе MEO предлагается использовать для предоставления таких услуг, как цифровая передача речи, данных, факсимильных сообщений, широковещательных уведомлений, и для обмена сообщениями.

### Полосы частот

В табл. 9.1 перечислены полосы частот, доступные для спутниковой связи. Обратите внимание, что чем выше частота полосы, тем больше ее доступная ширина.

Впрочем, полосы более высокой частоты сильнее подвержены искажениям передачи. Службам мобильной спутниковой связи (MSS) выделены частоты в полосах S и L. Если сравнить эти полосы с полосами более высоких частот, то здесь наблюдается большее преломление луча и его проникновение через физические преграды, такие, как листва и неметаллические структуры. Эти характеристики являются крайне важными для услуг мобильной связи. Кроме того, полосы L и S широко используются для других наземных приложений, поэтому в настоящее время существует конкуренция между поставщиками услуг в СВЧ-диапазоне за возможность использования полос L и S.

**Таблица 9.1. Полосы частот для спутниковой связи**

Полоса	Диапазон частот, ГГц	Суммарная ширина полосы, ГГц	Распространенные приложения
L	1–2	1	Мобильная спутниковая связь (MSS)
S	2–4	2	Службы MSS, NASA, исследование дальнего космоса
C	4–8	4	Спутники стационарной службы связи (FSS)
X	8–12,5	4,5	Военные службы FSS, исследования Земли и метеорологические спутники
Ku	12,5–18	5,5	Службы FSS, радиовещательные спутниковые службы (BSS)
K	18–26,5	8,5	Службы FSS и BSS
Ka	26,5–40	13,5	Службы FSS

При предоставлении службе полосы частот диапазоны для восходящего и нисходящего каналов определяются отдельно, причем частота восходящего канала всегда выше. Как известно, для сигнала с более высокой частотой рассеяние или потери в свободном пространстве больше, чем для его низкочастотного дополнения. Впрочем, наземные станции располагают достаточной мощностью, позволяющей компенсировать низкую производительность на высоких частотах.

## Ухудшение качества связи

Производительность спутникового канала связи зависит от трех факторов:

- расстояния между антенной наземной станции и антенной спутника;
- в нисходящем канале — от расстояния между антенной наземной станции и “точкой прицела” спутника;
- атмосферного поглощения.

Рассмотрим данные факторы подробнее.

### Расстояние

Напомним формулу (5.2), выражение для потерь в свободном пространстве:

$$L_{\text{дБ}} = 10 \lg \frac{P_t}{P_r} = 10 \lg \left( \frac{4\pi d}{\lambda} \right) = -20 \lg(\lambda) + 20 \lg(d) + 21,98 \text{ дБ}. \quad (9.2)$$

Здесь

$P_t$  — мощность сигнала на передающей антенне;

$P_r$  — мощность сигнала на принимающей антенне;

$\lambda$  — длина волны несущей;

$d$  — расстояние распространения между антеннами.

Величины  $d$  и  $\lambda$  измеряются в одних единицах (например, в метрах).

Чем выше частота сигнала (т.е. чем меньше его длина волны), тем большими будут потери. Для геостационарного спутника потери в свободном пространстве на экваторе составляют:

$$L_{дБ} = -20 \lg(\lambda) + 20 \lg(35,863 \times 10^6) + 21,98 \text{ дБ} = -20 \lg(\lambda) + 173,07.$$

Потери для точек земной поверхности, удаленных от экватора, но все еще видимых со спутника, будут несколько больше. Максимальное расстояние (от спутника до горизонта) для геостационарного спутника равно 42 711 км. Потери в свободном пространстве при прохождении сигналом такого расстояния равны:

$$L_{дБ} = -20 \lg(\lambda) + 174,59.$$

Затухание сигнала как функции частоты и высоты орбиты показано на рис. 9.5.

### След спутника

Для работы на СВЧ, которые используются в спутниковой связи, применяются узконаправленные антенны. Таким образом, сигнал со спутника не распространяется во все стороны, а нацеливается на определенную точку Земли. Выбор точки прицела производится с учетом местоположения и размеров области, которая должна покрываться. На центральную точку этой области придется наиболее мощный сигнал, интенсивность сигнала будет спадать по мере удаления от центральной точки в любом направлении. Этот эффект обычно демонстрируется на примере модели, называемой следом спутника (satellite footprint), представленной на рис. 9.6. Показано, какая эффективная часть излученной энергии антенны попадает в каждую точку на территории США. Результат зависит от мощности сигнала, поступающего в принимающую антенну, а также от направленности передающей антенны. В рассматриваемом примере мощность сигнала, принимаемого в Арканзасе, составляет +36 дБВт, а в Массачусетсе — +32 дБВт. Чтобы получить реальную мощность сигнала, принимаемого в каждой точке покрываемой области, из эффективной мощности, указанной на рисунке, следует вычесть потери в свободном пространстве.

### Атмосферное поглощение

Главной причиной затухания сигнала в атмосфере является наличие кислорода, от которого, конечно же, никуда не деться, а также влаги. Поглощение, обусловленное наличием воды, происходит в туман и в дождь. Еще одним фактором, влияющим на затухание сигнала, является значение угла возвышения спутника, который зависит от положения наземной станции (угол  $\theta$  на рис. 9.1). Чем меньше угол возвышения, тем больший путь в атмосфере придется пройти сигналу. Наконец, величина атмосферного затухания зависит от частоты сигнала. В общем случае чем выше частота сигнала,

тем сильнее он затухает. На рис. 9.7 показана характерная величина затухания как функция угла возвышения для частот полосы С. Конечно, затухание, обусловленное туманом и дождем, возникает только в случае присутствия этих явлений в атмосфере.

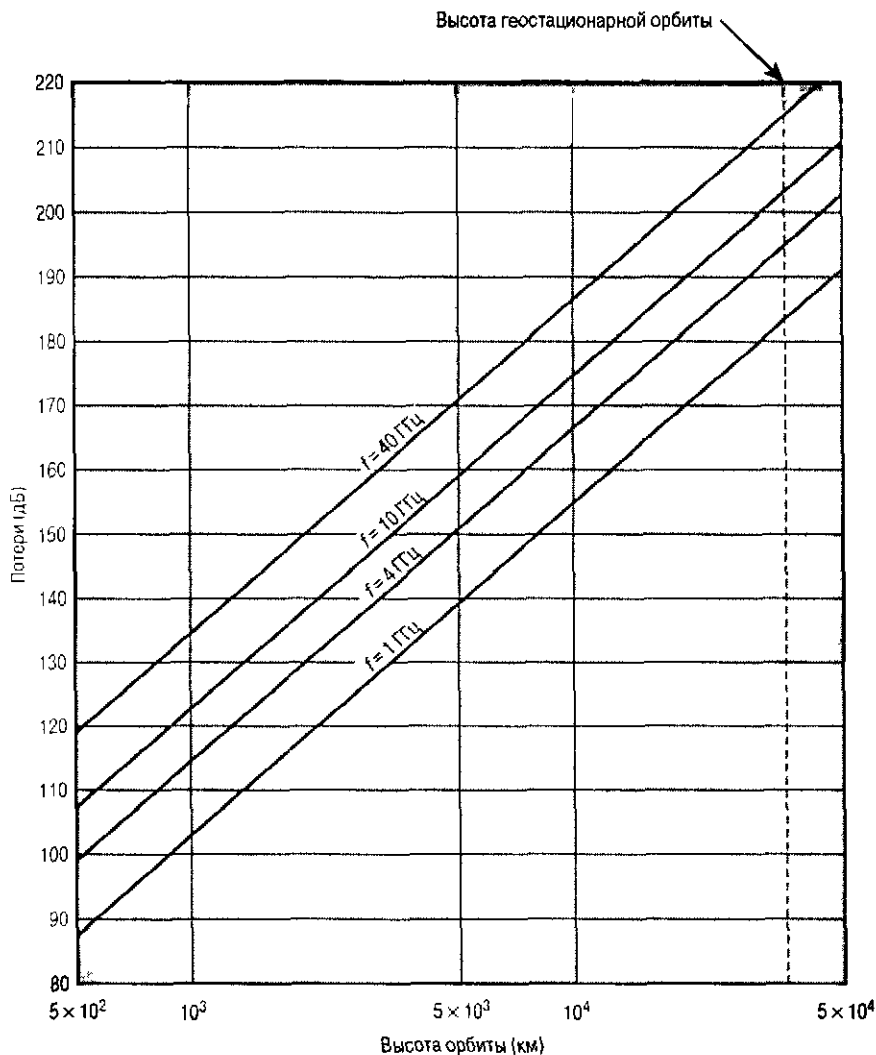


Рис. 9.5. Минимальные потери в свободном пространстве как функция высоты орбиты

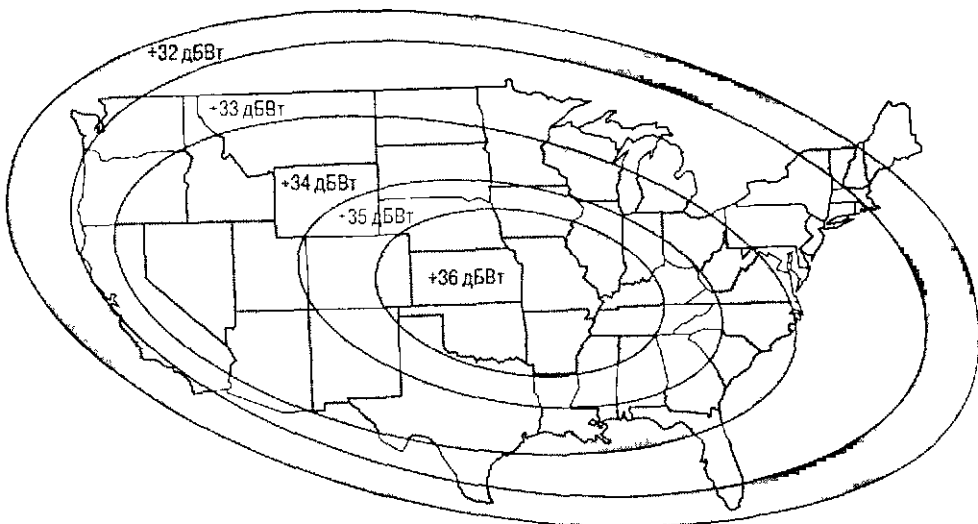


Рис. 9.6. Типичный след спутника

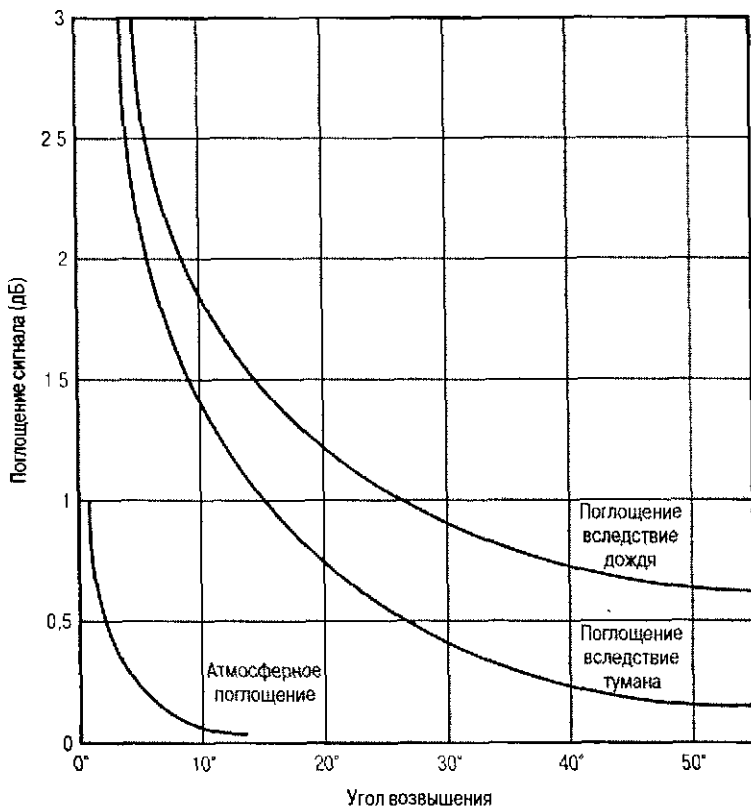


Рис. 9.7. Затухание сигнала (полосы С), обусловленное атмосферным поглощением

## Конфигурации спутниковой сети

На рис. 9.8 схематично изображены две наиболее распространенные конфигурации спутниковых систем связи. В первой конфигурации спутник используется для обеспечения двухточечной связи между двумя удаленными наземными антеннами. Во второй конфигурации спутник обеспечивает сообщения между одним наземным передатчиком и несколькими наземными приемниками.

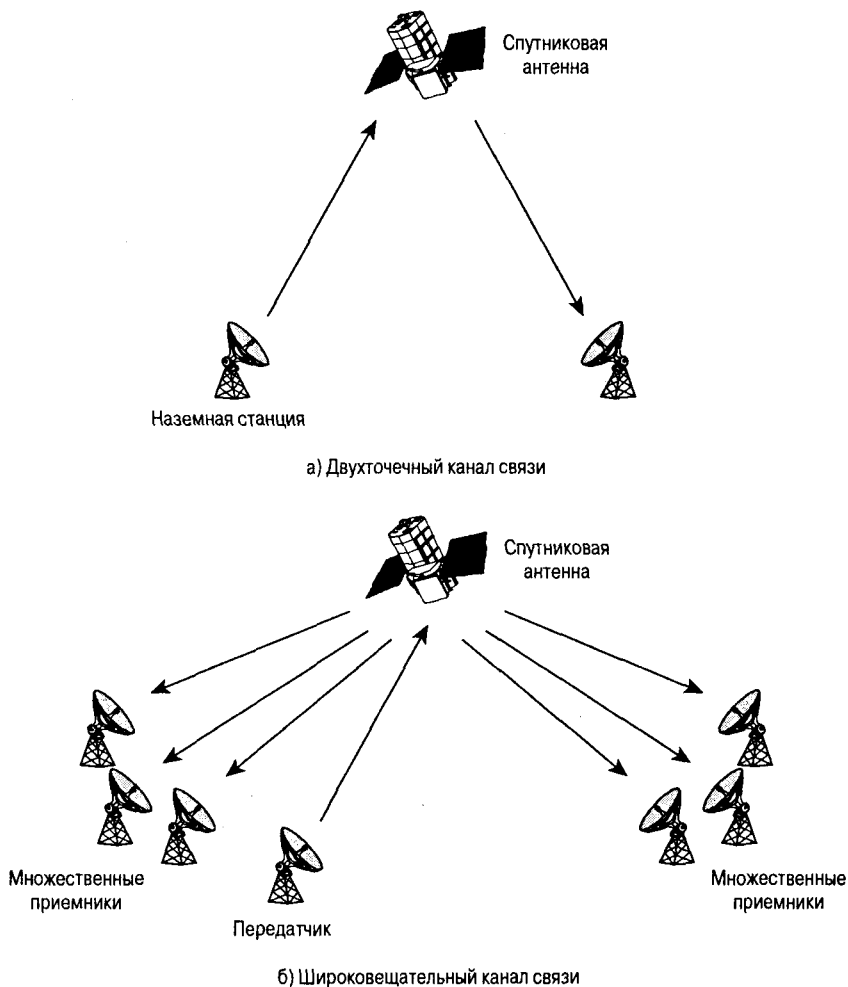


Рис. 9.8. Конфигурации систем спутниковой связи

Существует также разновидность второй конфигурации, в которой осуществляется двухсторонняя связь между комплексом наземных станций, состоящим из одного центрального концентратора и множества удаленных абонентских станций. Такой тип конфигурации, показанный на рис. 9.9, используется в системах VSAT (Very Small Aperture Terminal — терминал со сверхмалой апертурой луча). Недорогими антеннами VSAT оснащены многие абонентские станции. Установив определенный порядок, эти станции совместно используют пропускную



способность спутника для передачи данных станции-концентратору. Концентратор может обмениваться сообщениями с любым абонентом и ретранслировать сообщения, идущие от одного абонента другому.

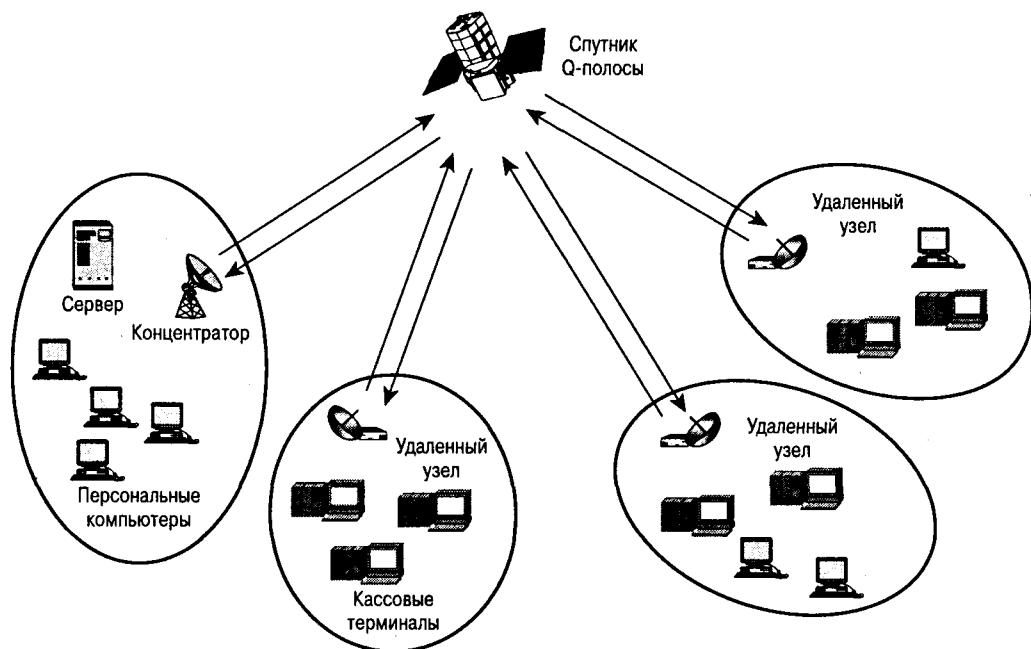


Рис. 9.9. Конфигурация типичной системы VSAT

## 9.2. РАСПРЕДЕЛЕНИЕ ПРОПУСКНОЙ СПОСОБНОСТИ — ЧАСТОТНОЕ РАЗДЕЛЕНИЕ

Как правило, один геостационарный спутник работает в довольно широкой полосе частот (например, 500 МГц), которая делится на несколько каналов, имеющих меньшую ширину (например, по 40 МГц). В каждом из таких каналов требуется распределить пропускную способность. Иногда, например при телевещании или передаче единичного потока цифровых данных со скоростью 50 Мбит/с, весь канал выделяется для одного пользователя или приложения. В то же время если отбросить такие крайние случаи, то можно сказать, что экономное использование спутника невозможно без разделения канала между несколькими пользователями. Поэтому задача распределения, в основном, сводится к уплотнению каналов, о котором рассказывалось в главе 2. В некоторых случаях распределение проводится под централизованным управлением, осуществляемым обычно со спутника. Пропускная способность может также распределяться динамически посредством команд, передаваемых наземными станциями. Ниже на примерах рассмотрены оба случая.

Все стратегии распределения относятся к одной из трех категорий.

- Множественный доступ с частотным разделением (Frequency Division Multiple Access — FDMA).

- Множественный доступ с временным разделением (Time Division Multiple Access — TDMA).
- Множественный доступ с кодовым разделением (Code-Division Multiple Access — CDMA).

В этом и следующих разделах мы изучим схемы FDMA и TDMA. Схема CDMA уже была рассмотрена в главе 7.

## Уплотнение с частотным разделением (FDM)

Как уже упоминалось, полная пропускная способность спутника связи делится на несколько каналов. Это верхний уровень уплотнения, далее пропускная способность распределяется внутри каждого канала. На рис. 9.10 приведен пример схемы уплотнения с частотным разделением (FDM), типичной для геостационарных спутников связи. Эта схема используется в спутниках Galaxy корпорации PanAmSat<sup>3</sup>. Используемые спутником частоты принадлежат полосе C, ширина полосы спутника равна 500 МГц и разбита на 24 канала по 40 МГц. Втиснуть 24 канала в полосу шириной 500 МГц удается посредством многократного использования частоты: каждая из выделяемых частот используется двумя несущими, поляризации которых ортогональны. В каждый канал шириной 40 МГц входит защитная полоса, составляющая 4 МГц, так что реальная ширина каждого канала равна 36 МГц. При использовании двухточечной конфигурации (рис. 9.8, а) каждый канал можно применять для различных альтернативных целей. Например:

- 1200 каналов речевого диапазона (VF);
- один поток данных со скоростью 50 Мбит/с;
- 16 каналов со скоростью 1,544 Мбит/с в каждом;
- 400 каналов со скоростью 64 Кбит/с в каждом;
- 600 каналов со скоростью 40 Кбит/с в каждом;
- один аналоговый видеосигнал;
- от шести до девяти цифровых видеосигналов.

Ширина полосы для аналогового видеосигнала может показаться слишком большой. Рассчитаем ее, используя правило Карсона (уравнение (6.16)), рассмотренное в главе 6. Ширина полосы видеосигнала, сопровождаемого аудиосигналом, приблизительно равна 6,8 МГц. Затем суммарный сигнал посредством частотной модуляции (FM) переносится на несущую 6 ГГц. Максимальное отклонение частоты сигнала при использовании этой процедуры равно  $\Delta F = 12,5$  МГц. Определим ширину полосы передаваемого сигнала:

$$B_T = 2\Delta F + 2B = 2(12,5 + 6,8) = 38,6 \text{ МГц,}$$

это приемлемо для транспондера на 36 МГц.

В цифровом видео использование сжатия может привести к тому, что скорость передачи данных в одном канале составит 3–5 Мбит/с, в зависимости от того, насколько видеоматериал насыщен движением.

<sup>3</sup> PanAmSat является крупнейшим спутниковым оператором. Эта частная компания предоставляет услуги спутниковой связи по всему миру.

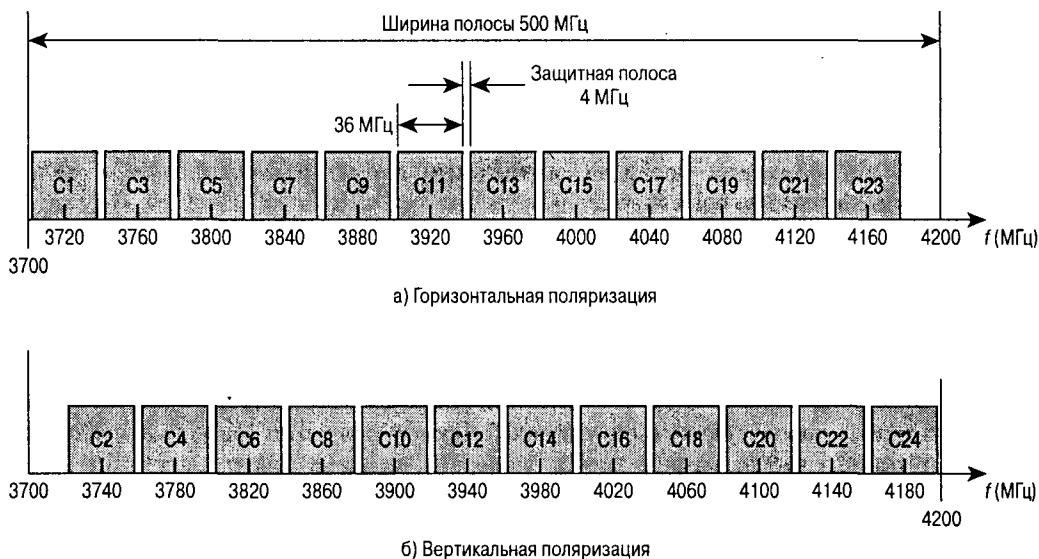


Рис. 9.10. Типичный план распределения частот спутникового транспондера для нисходящих каналов связи (чтобы составить план для восходящих каналов связи, к приведенным на рисунке значениям следует прибавить 2225 МГц)

## Множественный доступ с частотным разделением (FDMA)

В предыдущем разделе предполагалось, что спутник используется в качестве промежуточного устройства, обеспечивающего, по сути, двухточечную связь между двумя наземными станциями. Поскольку зона обслуживания спутника весьма велика, он может выполнять гораздо больше функций. Например, для серии спутников INTELSAT<sup>4</sup> канал шириной 36 МГц можно разделить, используя уплотнение с частотным разделением, на несколько меньших каналов, в каждом из которых можно использовать частотную модуляцию. Каждый из этих меньших каналов, в свою очередь, несет несколько сигналов речевого диапазона (VF), для чего используется уплотнение с частотным разделением. Возможность получения несколькими наземными станциями доступа к одному и тому же каналу называется схемой FDMA (frequency division multiple access — множественный доступ с частотным разделением).

Количество подканалов, на которые можно разбить спутниковый канал с помощью технологии FDMA, ограничено тремя факторами:

<sup>4</sup> INTELSAT (International Telecommunications Satellite Organization — Международная организация спутниковых телекоммуникаций) — это консорциум национальных поставщиков спутниковой связи, основанный согласно международному договору. INTELSAT занимается приобретением космических аппаратов, учитывая глобальный спрос на телефонные и телевизионные услуги. Действует консорциум как оптовый продавец, предоставляя участки космического пространства конечным пользователям, работающим на собственных наземных станциях. Представителем Соединенных Штатов в этом консорциуме является COMSAT. В настоящее время INTELSAT находится в процессе приватизации.

- тепловой шум;
- комбинационные помехи;
- перекрестные помехи.

Эти термины были определены и рассмотрены в главе 5. Воздействие первых двух факторов прямо противоположно. Передаваемый сигнал очень малой интенсивности будет искажаться фоновым шумом. При очень большой интенсивности сигнала нелинейные эффекты, имеющие место в усилителях спутников, приведут к сильным комбинационным помехам. Перекрестные помехи происходят при попытках увеличить пропускную способность путем многократного использования частот. Перекрестные помехи ограничивают применение этой практики, однако не сводят ее на нет. Полосу частот можно многократно использовать в том случае, если имеются антенны, которые могут излучать два поляризованных сигнала одинаковой частоты с ортогональными поляризациями. Как и ранее, если интенсивность сигнала слишком высока, то становится значительной и интерференция.

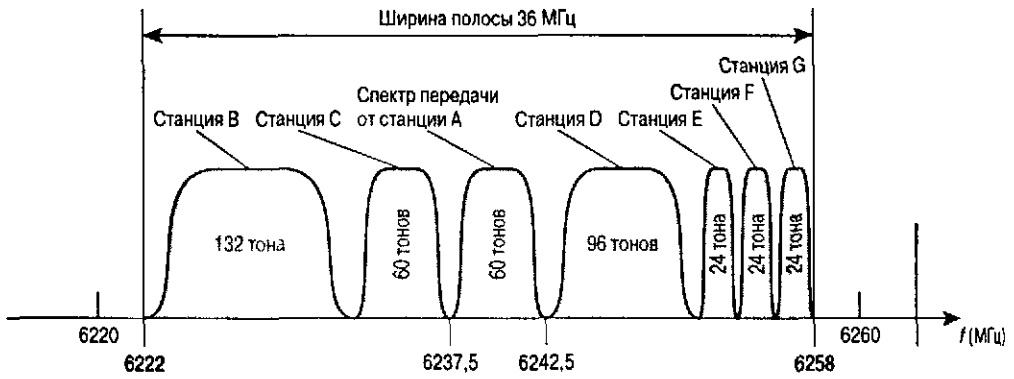
Возможны две формы FDMA.

- **Множественный доступ с фиксированным распределением (fixed-assignment multiple access — FAMA).** Распределение пропускной способности спутникового канала между множеством станций производится заранее. В результате значительная часть пропускной способности не используется, поскольку спрос на частоты может меняться в процессе связи.
- **Множественный доступ с распределением по запросу (demand-assignment multiple access — DAMA).** Распределение пропускной способности среди множества станций меняется с изменением спроса на частоты. Это позволяет оптимально распределить пропускную способность спутника.

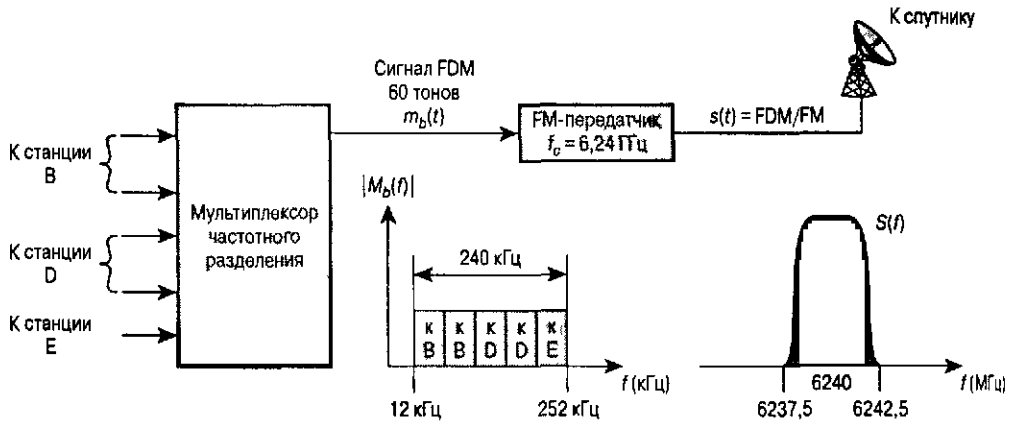
## FAMA-FDMA

На рис. 9.11 приведен частный пример схемы FAMA-FDMA, в которой семь наземных станций совместно используют восходящий канал связи с пропускной способностью 36 МГц. Нарисовать подобную диаграмму для нисходящего канала связи предлагаем читателю. Станция А выделяется полоса частот шириной 5 МГц, от 6237,5 до 6242,5 МГц, в которой можно разместить 60 голосовых каналов, используя уплотнение с разделением частот в сочетании с частотной модуляцией (FDM-FM). Т.е. FDM используется для разделения полосы на 60 каналов, а частотная модуляция — для модулирования каналов на несущей 6240 МГц. Как указано на рисунке, трафик от станции А к другим станциям выглядит следующим образом: 24 канала к станции В, 24 канала к станции D и 12 каналов к станции Е. Оставшийся спектр канала, общая ширина которого составляет 36 МГц, распределяется между другими наземными станциями в соответствии с их потребностью в трафике. В связи с этим примером следует отметить несколько поучительных моментов.

- На схеме показаны элементы как технологии FAMA, так и технологии FDMA. Термин FAMA отражает тот факт, что между станциями наперед заданы логические связи. Поэтому на рис. 9.11 показано, что станция А имеет три прямые двухточечные связи — по одной со станциями В (24 канала), D (24 канала) и Е (12 каналов). Термин FDMA отражает тот факт, что множество станций имеют доступ к спутниковой связи, используя при этом разные полосы частот.



а) Распределение частот в восходящем канале транспондера



б) Передающее оборудование наземной станции А

Рис. 9.11. Схема фиксированного распределения частот FDMA для спутниковой связи [COUC01]

- Несмотря на то что наземная станция может передавать на спутник только одну несущую (например, станция А передает на частоте 6,24 ГГц в полосе шириной 5 МГц), она должна быть способна к приему по крайней мере по одной несущей от каждой удаленной станции, с которой она хочет поддерживать связь (к примеру, станция А должна принимать три несущие, которые являются частью передачи станций В, D и Е).
- Спутник не выполняет функций коммутатора, хотя он принимает от различных источников отдельные участки канала шириной 36 МГц. На спутнике происходит только прием сигналов в пределах этого спектра, их преобразование в полосу, центрированную на частоте 4 ГГц, и повторная передача.
- Используемая полоса частот имеет довольно значительную ширину. Например, станция А должна передавать 60 голосовых каналов, которые занимают всего 240 кГц (т.е. один канал занимает 4 кГц), а выделена для этой станции полоса шириной 5 МГц. Такая широкая полоса требуется из-за использования частотной (а не амплитудной) модуляции, которая позволяет

сохранить сигнал при передаче его на большие расстояния, а также минимизировать требования к мощности спутника.

## DAMA-FDMA

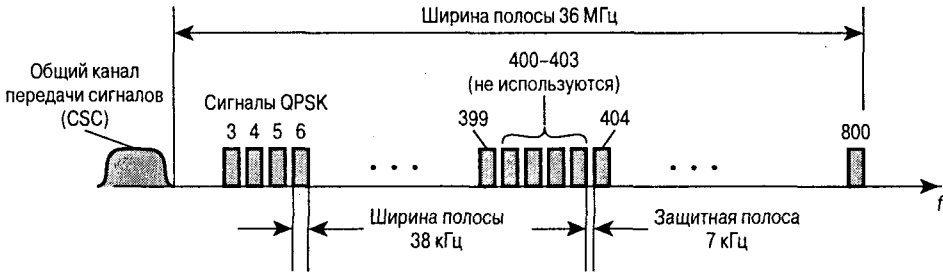
Только что описанная схема FAMA-FDMA оказывается не очень эффективной. В полосе С используемая ширина каждого канала обычно равна 36 МГц. В одной схеме FDMA, используемой корпорацией INTELSAT, эта полоса поделена на 7 блоков по 5 МГц, каждый из которых несет группу из 60 голосовых каналов. Таким образом, их общее количество составляет 420. В рассмотренном примере (рис. 9.11) также присутствуют 420 голосовых каналов. Если полосу частот разделить на 14 подканалов шириной по 2,5 МГц, то в каждом подканале можно будет разместить 48 голосовых каналов, т.е. всего 336 каналов. Эффективнее всего оказывается избегать создания групп и просто делить полосу в 36 МГц на отдельные голосовые каналы. Эта схема известна как “один канал — одна несущая” (single channel per carrier — SCPC).

Схема SCPC в настоящее время реализована в полосе С. Один канал шириной 36 МГц разделяется на 800 аналоговых каналов по 45 кГц, каждый из которых выделяется для симплексного канала связи с использованием частотной модуляции. Существует также цифровой вариант схемы SCPC, в котором используется модуляция QPSK, дающая в той же полосе 45 кГц скорость передачи 54 Кбит/с; этого достаточно для передачи оцифрованной речи. С помощью технологии FAMA пары (полнодуплексных) каналов распределяются по парам наземных станций. Как правило, каналы каждой наземной станции приходится дополнительно уплотнять, поскольку она обслуживает некоторое количество пользовательских станций. Если каждая наземная станция работает с большим числом пользовательских станций, то даже при использовании технологии FAMA достигается высокая степень связности. Тогда, как и в обычной схеме FDMA, спутник принимает частоты в пределах канала шириной всего 36 МГц, транслирует их в полосу, центрированную на частоте 4 ГГц, и передает канал всем станциям.

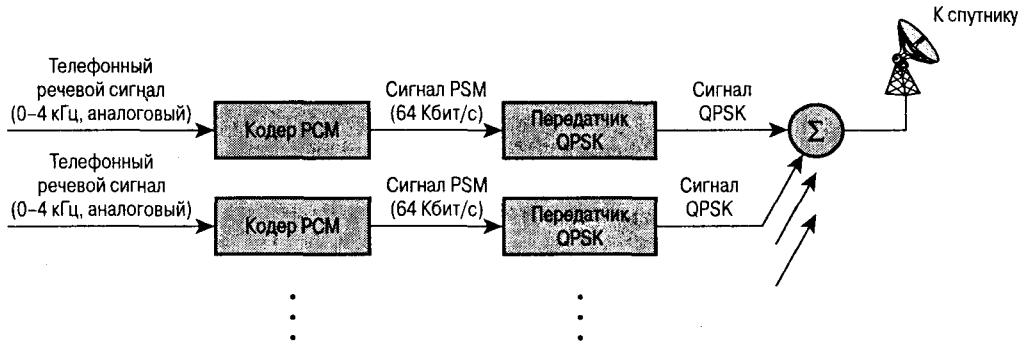
Технология SCPC привлекательна для удаленных областей, в которых находится хотя бы несколько пользовательских станций. Если FDMA используется в качестве основного средства в системах связи на большие расстояния, SCPC предоставляет услуги напрямую конечному пользователю. Хотя в схеме SCPC ширина полосы используется более эффективно, чем в FDMA, эта технология имеет все недостатки фиксированного распределения. Она оказывается особенно неудобной для слишком удаленных областей, в которых наземная станция часто обслуживает всего одну или две пользовательские станции. Повысить эффективность позволяет использование схемы DAMA, согласно которой в канале выделяется набор подканалов, рассматриваемый в качестве резерва связи. Если требуется установить полнодуплексную связь между двумя наземными станциями, из резерва выделяется пара подканалов.

Первой доступной для коммерческих целей системой DAMA SCPC была SPADE (single-channel-per-carrier, pulse code modulation, multiple-access demand assignment equipment — оборудование импульсно-кодовой модуляции с множественным доступом с распределением запросов по требованию и одноканальным использованием несущей), введенная в действие в начале 70-х и в настоящее время доступная через спутники INTELSAT. Каждый подканал этой системы несет QPSK-модулированный сигнал со скоростью 64 Кбит/с, что занимает полосу

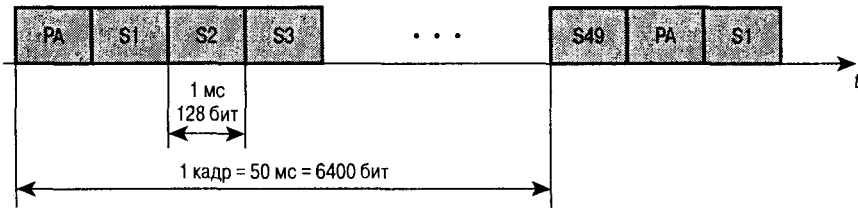
38 кГц, а также защитную полосу 7 кГц. Как правило, этот сигнал используется для передачи голосового сигнала с импульсно-кодовой модуляцией. Вместе со служебными подканалами, которые будут рассмотрены ниже, общее количество доступных подканалов достигает 794 (рис. 9.12, а). Эти подканалы разбиты на пары таким образом, чтобы два канала, отстоящие друг от друга на 18,045 МГц, всегда создавали полнодуплексный контур связи (например, каналы с номерами 3 и 404, 4 и 405, 399 и 800). Кроме того, имеется общий канал передачи сигналов (common-signaling channel — CSC) шириной 160 кГц, по которому передается сигнал PSK со скоростью 128 Кбит/с.



а) Распределение частот



б) Возможная конфигурация передатчика QPSK SCPC



в) Формат кадра TDMA CSC

Рис. 9.12. Система спутниковой связи SPADE для коммутируемой службы SCPC [COUC01]

Выделение каналов по требованию осуществляется наземной станцией с помощью канала CSC. Канал CSC используется для передачи повторяющихся кад-

ров TDM, как показано на рис. 9.12, в (сравните с рис. 2.13). Кадр разделен на 50 интервалов, первый из которых содержит начальную последовательность битов, требуемую для синхронизации. Остальные интервалы закреплены за 49 станциями. Эти станции имеют привилегию при формировании полнодуплексных контуров по требованию. Распределение по требованию осуществляется следующим образом. Предположим, что станция  $S_i$  желает установить канал связи со станцией  $S_j$ .  $S_i$  случайным образом выбирает подканал из доступных незанятых каналов и передает идентификатор подканала и адрес станции  $S_j$  во временной интервал станции  $S_i$ . Примерно через 0,25 с станция  $S_j$  получит запрос по нисходящему каналу связи. Если предположить, что подканал остался доступным и что доступна станция  $S_j$ , то она передаст подтверждение в свой собственный временной интервал, который станция  $S_i$  получит еще через четверть секунды. Когда сеанс связи завершится, во временном интервале одной из станций будет передана информация о разединении, чтобы уведомить другие станции о том, что подканал снова свободен.

Так как в описанной схеме могут участвовать только 49 станций, для распределения по требованию большинство подканалов не нужны и используются по обычной схеме FAMA SCPC.

### 9.3. РАСПРЕДЕЛЕНИЕ ПРОПУСКНОЙ СПОСОБНОСТИ — ВРЕМЕННОЕ РАЗДЕЛЕНИЕ

Хотя уплотнение с частотным разделением до сих пор широко используется в спутниковой связи, все большую популярность приобретают схемы уплотнения с временным разделением. Перечислим причины, по которым это происходит.

- Непрерывное падение стоимости цифровых составляющих.
- Преимущества цифровых методов обработки, включая возможность исправления ошибок.
- Повышенная эффективность схем TDM, которая обусловлена отсутствием комбинационных помех.

Как и при частотном разделении, все схемы временного деления обеспечивают множественный доступ и включают схемы FAMA-TDMA и DAMA-TDMA. Метод FAMA-TDMA, по сути, не отличается от описанного в разделе 2.5 синхронного временного деления. Передача осуществляется в форме повторяющейся последовательности кадров, каждый из которых делится на несколько временных интервалов. Каждое положение интервала в последовательности кадров соответствует определенному передатчику. Периоды кадров лежат в диапазоне от 10 мкс до более чем 2 мс и состоят из 3–100 интервалов и более. Скорость передачи данных составляет от 10 Мбит/с до более чем 100 Мбит/с.

Типичный формат кадра показан на рис. 9.13 (сравните с рис. 2.13, б). Как правило, кадр начинается с двух опорных пакетов, определяющих начало кадра. Эти два пакета предоставляются двумя разными наземными станциями, так что система может продолжать функционировать даже при потере связи с одной из опорных станций вследствие неисправности. Каждый опорный пакет начинается



с последовательности для восстановления тактовой синхронизации и несущей, являющейся уникальной и позволяющей синхронизироваться с центральным тактовым генератором. Каждой из  $N$  станций выделяется один или несколько интервалов в кадре. Станция использует выделенный ей интервал для передачи пакета данных, состоящего из предварительной последовательности и пользовательской информации. Предварительная последовательность содержит управляющую информацию и информацию о синхронизации плюс данные, идентифицирующие станцию назначения. Отдельные пакеты разделены защитными интервалами, предотвращающими наложение данных.

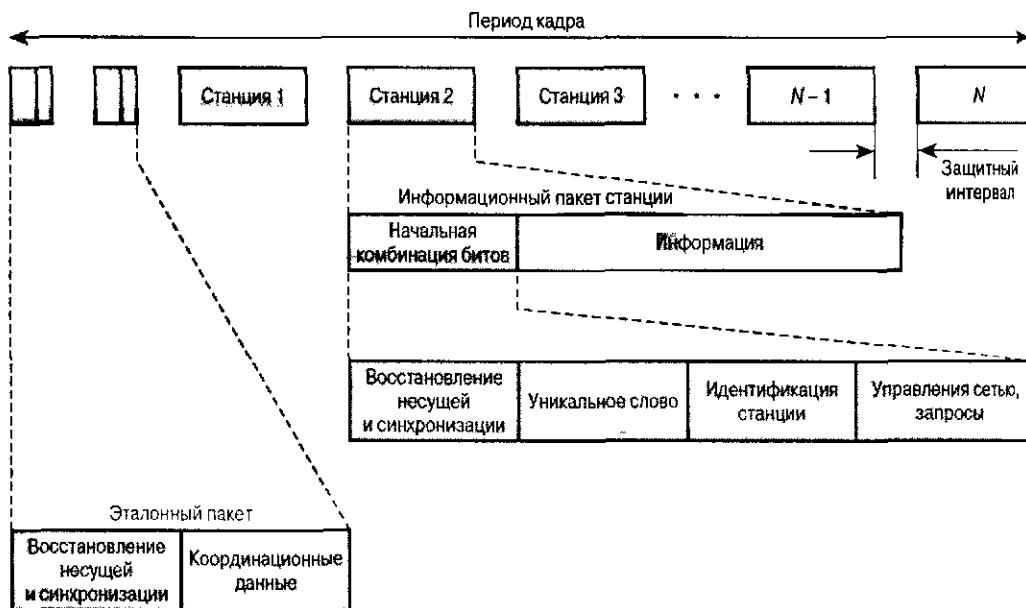
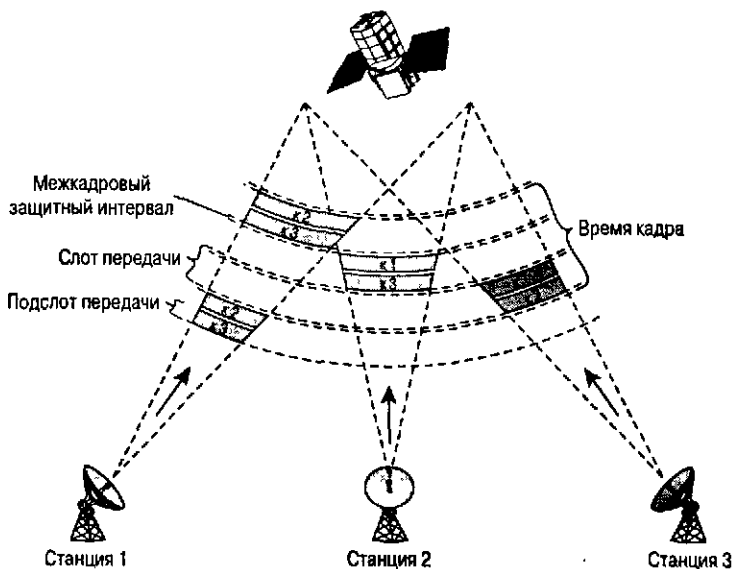


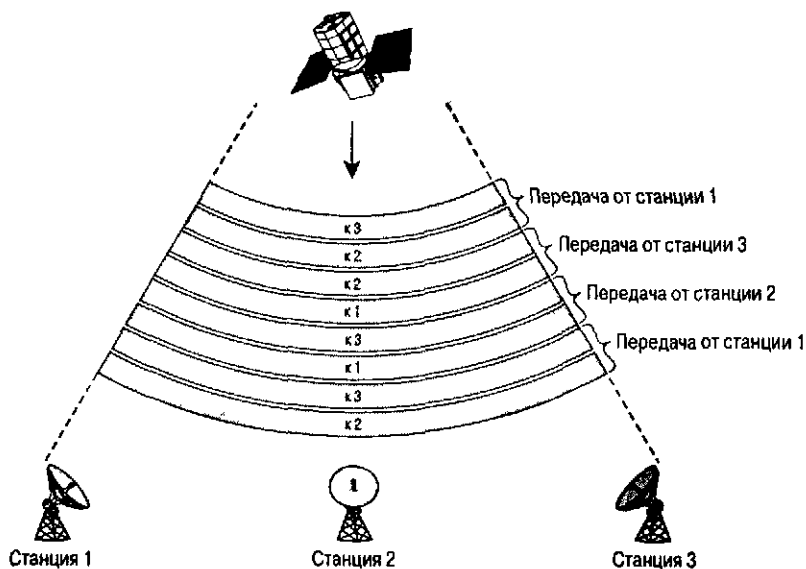
Рис. 9.13. Пример формата кадра TDMA

На рис. 9.14 показан принцип действия систем FAMA-TDMA. Отдельные наземные станции по очереди используют восходящий канал связи и могут помещать пакеты данных в выделенные для них временные интервалы. Спутник ретранслирует входящие данные всем станциям. Таким образом, передача и прием информации каждой станцией осуществляется в определенный временной интервал. Спутник также повторяет опорные пакеты данных, поэтому все станции, принимающие эти пакеты, имеют возможность синхронизировать свою работу.

Каждый из повторяющихся временных интервалов является каналом и не зависит от других каналов. Поэтому его можно использовать любым способом, по выбору передающей станции. Можно, например, организовать коммутацию, включив в каждый временной интервал поле адреса. Тогда, хотя временные интервалы и останутся закрепленными за станциями, в каждом нисходящем интервале несколько станций смогут ожидать данные, адресованные именно им. С помощью еще одной схемы передающая наземная станция может разделить свой временной интервал на подынтервалы и, таким образом, отправлять данные по нескольким подканалам в одном канале TDMA.



а) Восходящий канал связи



б) Нисходящий канал связи

Рис. 9.14. Функционирование системы FAMA-TDMA

Обычная схема TDMA эффективнее традиционной схемы FDMA, так как на защитные интервалы и управляющие биты TDMA расходуется меньшая пропускная способность, чем на защитные полосы FDMA. Это проиллюстрировано на рис. 9.15. Обратите внимание на впечатляющее падение пропускной способности для FDMA по мере увеличения числа каналов. В случае использования схемы TDMA при увеличе-

нии числа временных интервалов (тех же каналов) пропускная способность уменьшается гораздо медленнее. Кроме того, при использовании более длинных кадров также увеличивается эффективность связи. Для сравнения, система SCPC обеспечивает постоянную пропускную способность в 800 каналов, независимо от того, поделена ее полоса частот между множеством станций или между всего несколькими наземными станциями.

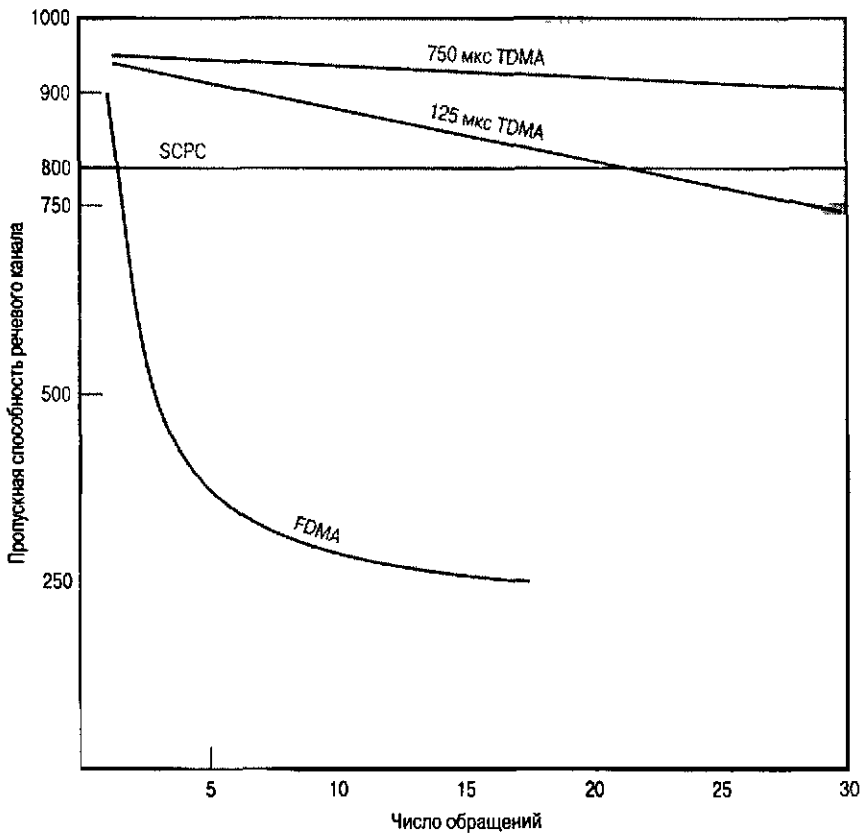


Рис. 9.15. Относительная эффективность различных спутниковых схем распределения пропускной способности [EDEL82]

В полосах с более высокой частотой (Ku и K) можно достичь даже большей эффективности. При таких частотах лучи, передаваемые со спутника, можно довольно точно сфокусировать, что позволит передавать много лучей одной и той же частоты в разные зоны. Таким образом, спутник сможет обслуживать довольно много зон, в каждой из которых может находиться много наземных станций. Сообщение между этими станциями в пределах одной зоны осуществляется по обычной схеме FDMA-TDMA. Более того, можно организовать сообщение между станциями, принадлежащими различным областям, если спутник имеет возможность переключать временные интервалы с одного луча на другой. Этот метод известен как TDMA со спутниковой коммутацией (satellite-switched TDMA — SS/TDMA).

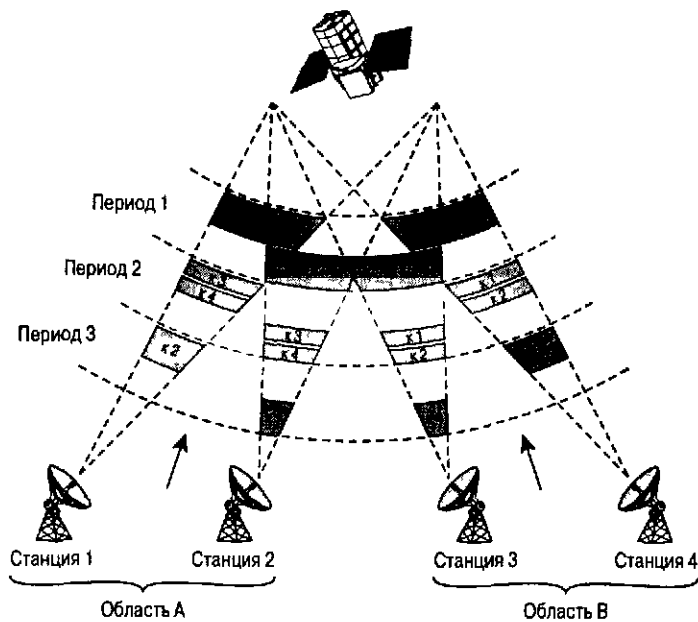
На рис. 9.16 показана простая система SS/TDMA, обслуживающая две зоны, в каждой из которых расположено по две станции. Как и в обычной схеме TDMA, в пределах одной зоны в каждый момент времени передавать может только одна станция. Таким образом, в пределах зоны А в любой конкретный временной интервал может передавать либо станция 1, либо станция 2. Так же и в пределах зоны В передавать в данный момент времени может либо станция 3, либо станция 4. Сигналы со станций из разных зон не интерферируют, благодаря использованию по-разному поляризованных сигналов либо использованию разных частот. На спутнике принимаемые данные тут же ретранслируются на частоте нисходящего канала связи. При этом используются два различных нисходящих луча. На спутнике имеется коммутатор для соединения входящих и выходящих лучей. Варианты соединений, выполняемых коммутатором, могут меняться со временем. На рисунке нисходящий луч А повторяет восходящий луч А в течение периодов 1 и 3 и повторяет восходящий луч В в течение периода 2. Таким образом, любая станция в любой зоне может отправлять данные любой другой станции.

Для спутника, обслуживающего  $N$  зон, существует  $N$  входящих потоков TDM. В любой конкретный момент времени переключатель настроен на особый вид преобразования этих восходящих лучей в  $N$  нисходящих лучей. Каждая конфигурация называется *режимом*, и для полной связности системы требуется  $N!$  различных режимов. В табл. 9.2 приведены режимы для системы с тремя лучами. Например, станции в зоне А могут общаться друг с другом в режимах 1 и 2, общаться со станциями зоны В в режимах 3 и 5 и т.д. Спутник периодически меняет режим. Скорее всего, смена режимов будет происходить один раз за интервал. Структуру режима и его длительность обычно можно настраивать командами с Земли, чтобы вовремя реагировать на изменившиеся требования.

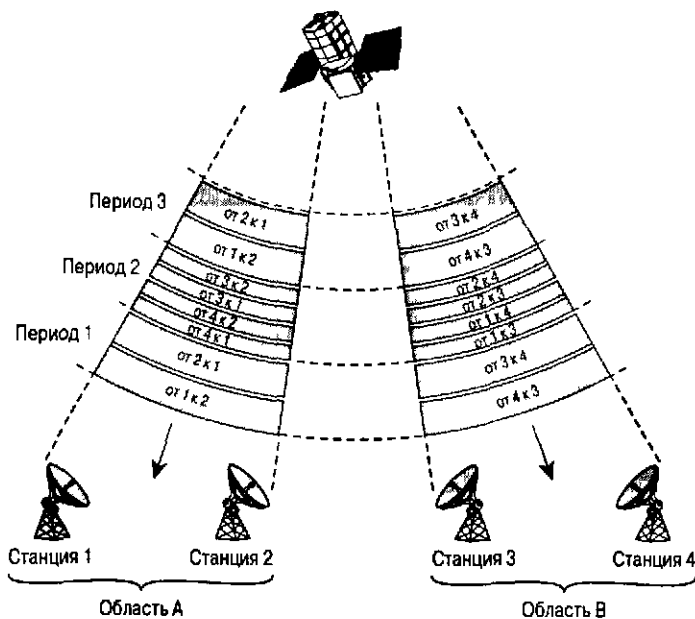
Таблица 9.2. Режимы системы SS/TDMA (три луча)

Вход	Выход					
	Режим 1	Режим 2	Режим 3	Режим 4	Режим 5	Режим 6
А	А	А	В	С	В	С
В	В	С	С	А	С	В
С	С	В	А	В	А	А

Наконец, со спутниковыми схемами TDMA используются методы DAMA. Система SS/TDMA является в определенном смысле системой DAMA, если структура режима может быть изменена командой с Земли. Довольно часто, если речь заходит о временном разделении, схему DAMA называют схемой множественного доступа с совместным использованием одного канала. Такие методы обычно служат основой для кооперации наземных станций с целью совместного использования канала.



а) Восходящий канал связи



б) Нисходящий канал связи

Рис. 9.16. Функционирование системы SS/TDMA

## 9.4. РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

Понятный и не перегруженный техническими деталями обзор спутниковой связи дан в [INGL97]. В [ELBE99] содержится больше технических подробностей. Прекрасная техническая трактовка дана также в [FREE97].

ELBE99 Elbert B. *Introduction to Satellite Communication*. — Boston: Artech House, 1999.

FREE97 Freeman R. *Radio System Design for Telecommunications*. — New York: Wiley, 1997.

INGL97 Inglis A., Luther A. *Satellite Technology: An Introduction*. — Boston: Focal Press, 1997.



### Рекомендуемые Web-сайты

- **Lloyd Wood's satellite Web page** ([www.ee.surrey.ac.uk/Personal/L.Wood/constellations/](http://www.ee.surrey.ac.uk/Personal/L.Wood/constellations/)). Прекрасный обзор спутниковой связи, снабженный большим количеством ссылок.
- **Satellite Industry Assosiation** ([www.sia.org/](http://www.sia.org/)). Торговая организация, представляющая интересы космических и коммуникационных компаний США в области коммерческих спутников. Большое количество ссылок.

## 9.5. ТЕРМИНЫ, ВОПРОСЫ И ЗАДАЧИ

### Основные термины

“один канал — одна несущая” (SCPC)	множественный доступ с временным разделением (TDMA)	нисходящий канал связи орбита спутника
TDMA со спутниковой коммутацией (SS/TDMA)	множественный доступ с распределением по запросу (DAMA)	поляризация спутник
восходящий канал связи геосинхронная орбита (GEO)	множественный доступ с фиксированным распределением (FAMA)	спутник связи спутник мобильной службы связи (MSS)
геостационарная орбита (GEO)	множественный доступ с частотным разделением (FDMA)	спутник службы радиовещания (BSS)
зона обслуживания спутника	наземная станция	спутник стационарной службы (FSS)
многократное использование частоты	низкая околоземная орбита (LEO)	средняя околоземная орбита (MEO)
		транспондер

## Вопросы

1. Перечислите три способа классификации спутников связи.
2. Какие существуют ключевые различия между спутниковыми и наземными беспроводными системами связи?
3. Перечислите три способа классификации орбит спутников.
4. Опишите спутники GEO, LEO и MEO (расшифровав также аббревиатуры). Сравните такие параметры этих спутников, как размер и форма орбит, мощность сигнала, возможность многократного использования частоты, задержка распространения сигнала, количество спутников, необходимое для покрытия земного шара, и укажите, насколько часто требуется передача сигнала с одного спутника на другой.
5. При каких обстоятельствах вы бы использовали спутники GEO, LEO и MEO, соответственно?
6. Перечислите три ключевых фактора, от которых зависит производительность спутниковой связи.
7. Каковы основные причины затухания спутниковых сигналов в атмосфере?
8. Укажите три фактора, ограничивающих количество подканалов, на которые можно разделить спутниковый канал по схеме FDMA?

## Задачи

1. Какой процент от ширины полосы используется для защитных полос в транспондере, схема которого показана на рис. 9.10?
2. Пусть кадр TDMA, схема которого показана на рис. 9.13, имеет следующие параметры: длина кадра равна 2 мс. Модуляция QPSK при скорости передачи элементарных сигналов — 60,136 Мбод; все пакеты имеют одинаковую длину, которая составляет 16 512 бит; опорные пакеты имеют длину 579 бит, начальная комбинация битов имеет длину 560 бит, а защитный интервал составляет 24 бит. Пусть между собой общаются две опорных станции, так что требуется два опорных пакета.
  - а. Определите максимальное количество наземных станций, которые может обслужить подобная система.
  - б. Чему равна эффективность кадра (имеется в виду часть кадра, несущая пользовательские данные, а не биты служебной информации)?

# ГЛАВА 10

## БЕСПРОВОДНЫЕ СОТОВЫЕ СЕТИ

- 10.1. Принципы сотовой связи
- 10.2. Аналоговые системы первого поколения
- 10.3. Системы TDMA второго поколения
- 10.4. Системы CDMA второго поколения
- 10.5. Системы третьего поколения
- 10.6. Рекомендуемая литература и Web-сайты
- 10.7. Термины, вопросы и задачи



**И**з всех потрясающих воображение успехов, достигнутых в сфере телекоммуникаций и обмена данными, возможно, самым революционным является создание сотовой сети. Именно на сотовой технологии основана мобильная беспроводная связь, которая открыла доступ к связи даже тем пользователям, которых трудно было обслужить с помощью обычной проводной сети. Сотовая технология лежит в основе мобильных телефонов, систем персональной коммуникации, беспроводной работы в Internet, беспроводных Web-приложений и многого другого.

Эта глава начнется с обзора основных принципов, на которых построены все сотовые сети. Затем будут рассмотрены конкретные сотовые технологии и стандарты, удобно группируемые в три поколения. Технологии первого поколения основаны на аналоговых сигналах и в настоящее время все еще широко используются, хотя постепенно сходят со сцены. Доминирующей технологией на сегодня являются цифровые системы, принадлежащие второму поколению. Наконец, высокоскоростные цифровые системы третьего поколения только начали появляться.

## 10.1. ПРИНЦИПЫ СОТОВОЙ СВЯЗИ

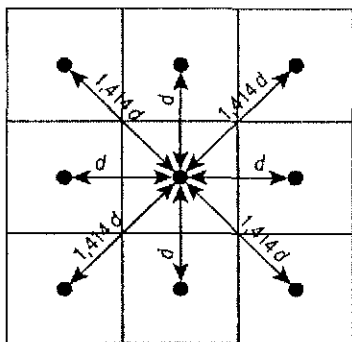
Сотовой радиосвязью называется технология, разработанная для увеличения пропускной способности мобильных радиотелефонных услуг. До введения сотовой радиосвязи для предоставления этих услуг требовались передатчики и приемники высокой мощности. Типичная система связи могла обслуживать около 25 каналов и имела эффективный радиус действия около 80 км. Для увеличения пропускной способности такой системы нужно было использовать оборудование более низкой мощности и, следовательно, меньшего радиуса действия при участии в ней нескольких передатчиков и приемников. В начале данного раздела кратко рассматривается организация сотовой системы, после чего анализируются некоторые особенности ее внедрения.

### Организация сотовой сети

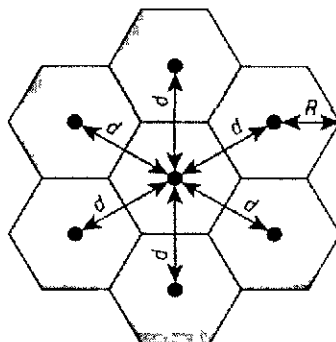
Принцип организации сотовой связи состоит в использовании множества маломощных (100 Вт и ниже) передатчиков. Поскольку диапазон действия таких передатчиков довольно мал, зону обслуживания системы можно разбивать на ячейки, каждая из которых будет обслуживаться собственной антенной. Каждая ячейка, которой выделяется своя полоса частот, обслуживается базовой станцией, состоящей из передатчика, приемника и модуля управления. Смежные ячейки используют разные частоты, чтобы избежать интерференции или перекрестных помех. В то же время ячейки, находящиеся на довольно большом расстоянии друг от друга, могут использовать одинаковые полосы частот.

При проектировании такой системы первое, что нужно сделать, — это решить, какую форму должны иметь ячейки, на которые будет разбита зона обслуживания. Самым простым решением была бы сетка, состоящая из квадратных ячеек (рис. 10.1, а). Однако такая геометрическая форма оказалась не идеальной. Если сторона квадратной ячейки равна  $d$ , тогда ячейка будет иметь четыре соседа на расстоянии  $d$  и четыре — на расстоянии  $\sqrt{2}d$ . В то же время, если пользователь мобильных услуг находится в пределах одной ячейки и движется по направлению к ее границе, было бы лучше, чтобы все смежные антенны находились на равных расстояниях друг от друга. В этом случае проще опре-

делить момент, в который следует переключать пользователя на другую антенну, а также выбирать новую антенну. Равное расстояние между смежными антеннами достигается только в шестиугольной схеме (см. рис. 10.1, б). Радиус шестиугольника определяется как радиус окружности, описанной вокруг него (эта величина равна расстоянию от центра фигуры до каждой из ее вершин, а также длине стороны шестиугольника). Для ячейки с радиусом  $R$  расстояние между центром ячейки и центром любой смежной ячейки равняется  $d = \sqrt{3}R$ .



а) Квадратная схема



б) Гексагональная схема

Рис. 10.1. Геометрические структуры сотовых систем

На практике точная шестиугольная структура не используется. Отклонения от идеальных шестиугольников обусловлены топографическими ограничениями, местными условиями распространения сигнала и соображениями целесообразности расположения антенн.

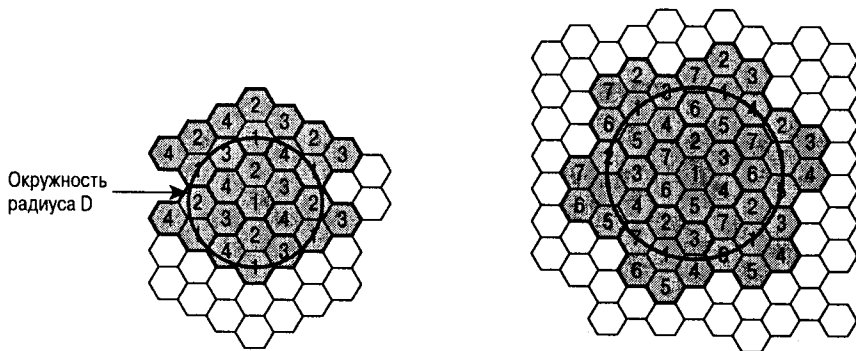
В беспроводной сотовой системе нельзя неограниченно использовать одну и ту же частоту для разных сообщений, так как при передаче на произвольных частотах разные сигналы могут интерферировать, даже если географически они разделены. Поэтому для систем, поддерживающих большое количество одновременных сеансов связи, нужен механизм, определяющий принципы использования спектра.

### Многokратное использование частот

В каждой ячейке сотовой сети имеется базовый трансивер. Мощность передаваемых сигналов тщательно регулируется (несколько это возможно для быстро меняющихся условий сред мобильной связи), поскольку (1) требуется осуществлять связь в пределах одной ячейки, но (2) это не должно приводить к интерференции сигналов данной ячейки с сигналами соседних. Как правило, каждой ячейке выделяется 10–50 частот, в зависимости от планируемой нагрузки. Кроме того, нужен механизм использования одной и той же частоты в ячейках, расположенных недалеко друг от друга, чтобы одну частоту можно было использовать для нескольких одновременных сеансов связи.

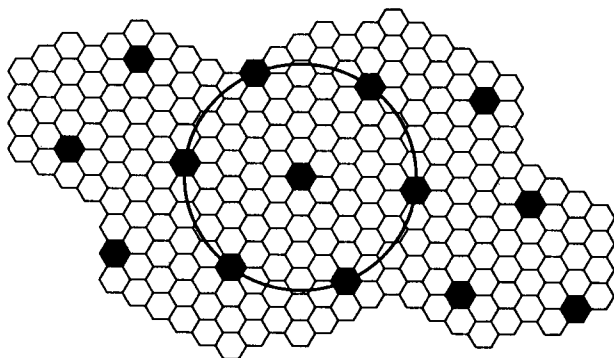
Важным вопросом, разумеется, является определение удаленности двух ячеек, использующих одну частоту, поскольку сигналы этих ячеек не должны интерферировать друг с другом. Были предложены различные модели многократного использования частот, некоторые примеры приведены на рис. 10.2. Если схема состоит из  $N$  ячеек, для которых выделяется одинаковое количество

частот, то каждая ячейка будет иметь  $K/N$  частот, где  $K$  — общее число частот, выделяемых системе. Мобильная телефонная система AMPS, в которой  $K = 395$ , а  $N = 7$ , представляет собой наименьшую систему, в которой можно обеспечить достаточную изоляцию двух сеансов использования одной и той же частоты. Это означает, что в среднем на одну ячейку должно приходиться не более 57 частот.



а) Схема повторного использования частот для  $N = 4$

б) Схема повторного использования частот для  $N = 7$



в) Черные ячейки указывают на повторное использование частот для  $N = 19$

Рис. 10.2. Модели многократного использования частот

Для характеристики повторного использования частоты существуют следующие параметры:

$D$  — минимальное расстояние между центрами ячеек, которые используют одну и ту же полосу частот (называемую группой внутренних каналов);

$R$  — радиус ячейки;

$d$  — расстояние между центрами смежных ячеек ( $d = \sqrt{3}R$ ).

$N$  — число ячеек в минимальном фрагменте, периодическим повторением которого образуется вся схема (каждая ячейка фрагмента использует уникальную полосу частот). Этот параметр еще называют **кратностью использования**.

В шестиугольной схеме возможны только следующие значения  $N$ :

$$N = I^2 + J^2 + (I \times J), \quad I, J = 0, 1, 2, 3, \dots$$

Таким образом, возможными значениями  $N$  являются числа 1, 3, 4, 7, 9, 12, 13, 16, 19, 21 и т.д. Верно следующее соотношение:

$$\frac{D}{R} = \sqrt{3N}.$$

Это можно записать и по-другому:  $D/d = \sqrt{N}$ .

### Увеличение пропускной способности

Со временем, когда система будет обслуживать все больше клиентов, трафик может распределиться таким образом, что какой-нибудь ячейке для обслуживания звонков не хватит выделенных ей частот. Для выхода из такой ситуации используется несколько подходов.

- **Добавление новых каналов.** Обычно, когда система установлена в определенном регионе, используются не все каналы, и с расширением системы можно просто добавлять новые.
- **Займствование частот.** В самом простом случае перегруженные ячейки могут “одалживать” частоты у смежных ячеек.
- **Расщепление ячеек.** На практике распределение трафика и топография местности неоднородны, что также дает возможность увеличения пропускной способности. Ячейки в областях с повышенным спросом на услуги мобильной связи можно расщеплять. Как правило, размеры исходных ячеек колеблются от 6,5 до 13 км. Меньшие ячейки также можно разбивать, однако следует помнить, что на практике радиус 1,5 км считается минимальным (см. ниже обсуждение микроячеек!). При использовании меньших ячеек нужно уменьшать уровень мощности, чтобы сигнал оставался в пределах ячейки. Кроме того, при движении мобильные устройства переходят из одной ячейки в другую, что требует передачи вызова от одного базового трансивера другому. Этот процесс называется *переключением* (handoff). Так вот, по мере уменьшения размера ячейки переключения будут происходить все чаще. На рис. 10.3 изображена схема расщепления ячейки, которая обеспечивает увеличение пропускной способности системы. При уменьшении радиуса ячейки в  $F$  раз размеры покрываемой области уменьшаются в  $F^2$  раз, а требуемое число базовых станций увеличивается в те же  $F^2$  раз.
- **Разбивка ячеек на секторы.** При разбивке на секторы ячейка делится на несколько клиновидных секторов, в каждом из которых остается свой набор каналов. Обычно на ячейку приходится 3–6 секторов. Каждому сектору предоставляется отдельный набор каналов ячейки, а для фокусировки сигнала на отдельных секторах используются направляемые антенны базовой станции.
- **Микроячейки.** По мере уменьшения ячейки антенны перемещаются с крыш высотных зданий и вершин холмов на крыши зданий поменьше или на стены высотных домов и в конце концов оказываются на фонарных столбах, с высоты которых они обслуживают микроячейки. Любое уменьшение разме-

ра ячейки сопровождается уменьшением уровня мощности сигналов, излучаемых базовой станцией. Микроячейки полезно располагать на городских улицах в густо населенных районах, а также внутри больших зданий общественного пользования.

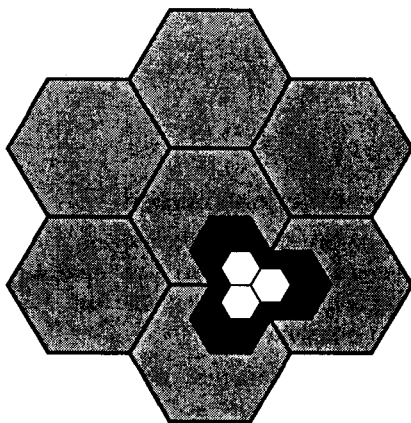


Рис. 10.3. Расщепление ячейки

В табл. 10.1 представлены характерные параметры традиционных ячеек, именуемых макроячейками, и микроячеек, созданных описанным выше способом. Средний разброс задержек — это усреднение разброса задержек по многим трактам распространения (ведь один и тот же сигнал может распространяться несколькими путями, и существует временная задержка между прибытием в приемник самого раннего и самого позднего сигнала). Как отмечалось выше, меньшие ячейки позволяют использовать меньшую мощность и обеспечивают прекрасные условия распространения сигнала.

Таблица 10.1. Характерные параметры макро- и микроячеек [ANDE95]

	Макроячейка	Микроячейка
Радиус ячейки	1–2 км	0,1–1 км
Мощность передаваемого сигнала	1–10 Вт	0,1–1 Вт
Средний разброс задержек	0,1–10 мкс	10–100 нс
Максимальная скорость передачи данных	0,3 Мбит/с	1 Мбит/с

**Пример.** [HAAS00]. Пусть система состоит из 32 ячеек с радиусом 1,6 км каждая, имеет выделенную полосу частот, позволяющую поддерживать 336 информационных каналов, и кратность использования частоты  $N = 7$ . Какую географическую зону обслуживают эти 32 ячейки, сколько каналов приходится на одну ячейку, и какое число одновременно поступающих звонков может обрабатываться? Повторите задачу для 128 ячеек с радиусом 0,8 км.

На рис. 10.4, а показана схема, покрывающая приблизительно квадратную область. Площадь одного шестиугольника с радиусом  $R$  составляет  $1,5R\sqrt{3}$ . Т.е. шестиугольник с радиусом 1,6 км покрывает зону площадью  $6,65 \text{ км}^2$ , а 32

ячейки покрывают зону площадью  $6,65 \times 32 = 213 \text{ км}^2$ . Для  $N = 7$  количество каналов на одну ячейку составляет  $336/7 = 48$ , а общая пропускная способность составляет  $48 \times 32 = 1536$  каналов. Для схемы, показанной на рис. 10.4, б, площадь покрываемой зоны равна  $1,66 \times 128 = 213 \text{ км}^2$ . Количество каналов на одну ячейку составляет  $336/7 = 48$ , а общая пропускная способность —  $48 \times 128 = 6144$  каналов.

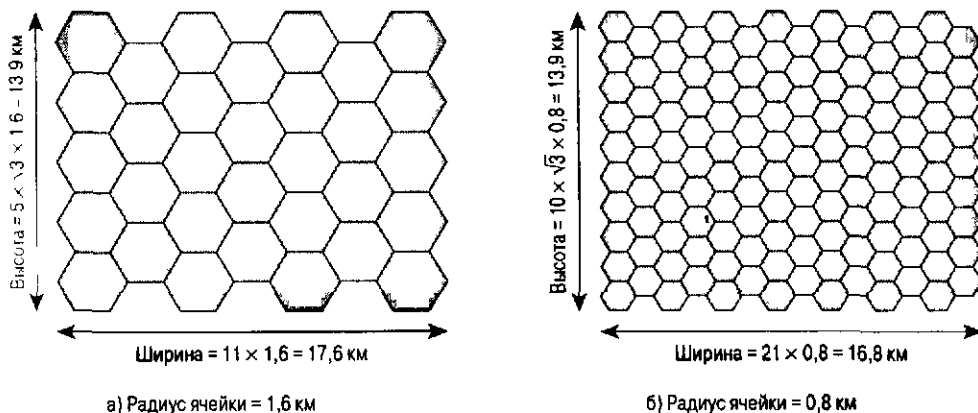


Рис. 10.4. Пример схемы с многократным использованием частот

## Функционирование сотовой системы

На рис. 10.5 показаны основные элементы сотовой системы. Примерно в центре каждой ячейки находится базовая станция. Базовая станция состоит из антенны, контроллера и нескольких трансиверов, которые служат для связи в каналах, выделенных этой ячейке. Контроллер используется для обработки соединений мобильного устройства с остальной сетью. В любой момент в пределах ячейки могут быть активными и перемещаться несколько пользователей мобильной связи, общающихся с базовой станцией. Каждая базовая станция подсоединена к коммутатору мобильных телекоммуникаций (mobile telecommunications switching office — MTSO), причем один коммутатор MTSO может обслуживать несколько базовых станций. Обычно связь между коммутатором MTSO и базовой станцией является проводной, хотя возможна также беспроводная связь. Коммутатор MTSO устанавливает соединение между мобильными устройствами. Кроме того, MTSO соединен также с общественной телефонной или телекоммуникационной сетью и может соединять стационарных абонентов с сетью общего пользования и мобильных абонентов с сотовой сетью. Коммутатор MTSO выделяет для каждого соединения голосовой канал, выполняет переключения (о которых мы поговорим ниже) и контролирует звонки для передачи информации о счетах.

Работа сотовой системы полностью автоматизирована и не требует от пользователя никаких действий, кроме заказа разговоров и ответа на звонки. Между мобильным устройством и базовой станцией можно устанавливать каналы связи двух типов: каналы управления и информационные каналы. Каналы управления используются для обмена информацией, касающейся заказа и поддержания

звонка, а также установления связи между мобильным устройством и ближайшей к нему базовой станцией. Информационные каналы служат для передачи голоса или данных между пользователями. На рис. 10.6 показаны шаги, которые следует предпринять для обычного соединения двух мобильных пользователей, находящихся в зоне действия одного коммутатора MTSO.

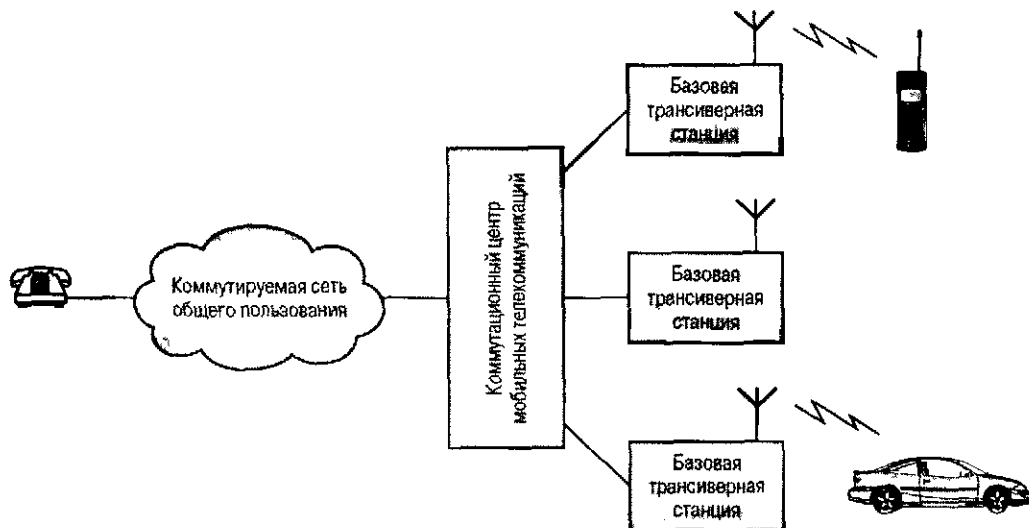


Рис. 10.5. Общий вид сотовой системы

- **Инициализация мобильного устройства.** Включенное мобильное устройство проводит сканирование и выбирает самый сильный настроечный канал управления, используемый данной системой (см. рис. 10.6, а). Ячейки с различными полосами частот периодически транслируют сигналы в различных настроечных каналах. Приемник мобильного устройства выбирает самый сильный настроечный канал и начинает его прослушивать. В результате этой процедуры мобильное устройство автоматически выбирает антенну базовой станции той ячейки, в пределах которой оно будет действовать<sup>1</sup>. Затем выполняется квитирование между мобильным устройством и коммутатором MTSO, контролирующим данную ячейку, что тоже осуществляется через базовую станцию этой ячейки. Квитирование используется для опознания пользователя и для регистрации его местоположения. Все время, пока включено мобильное устройство, эта процедура сканирования периодически повторяется, что позволяет следить за движением устройства. Если устройство входит в новую ячейку, выбирается новая базовая станция. Кроме того, мобильное устройство следит за сигналами избирательного вызова, о чем мы поговорим ниже.

<sup>1</sup> Обычно выбранная антенна, а следовательно, и базовая станция, является ближайшей к мобильному устройству. Однако из-за аномалий в распространении сигнала это не всегда так.

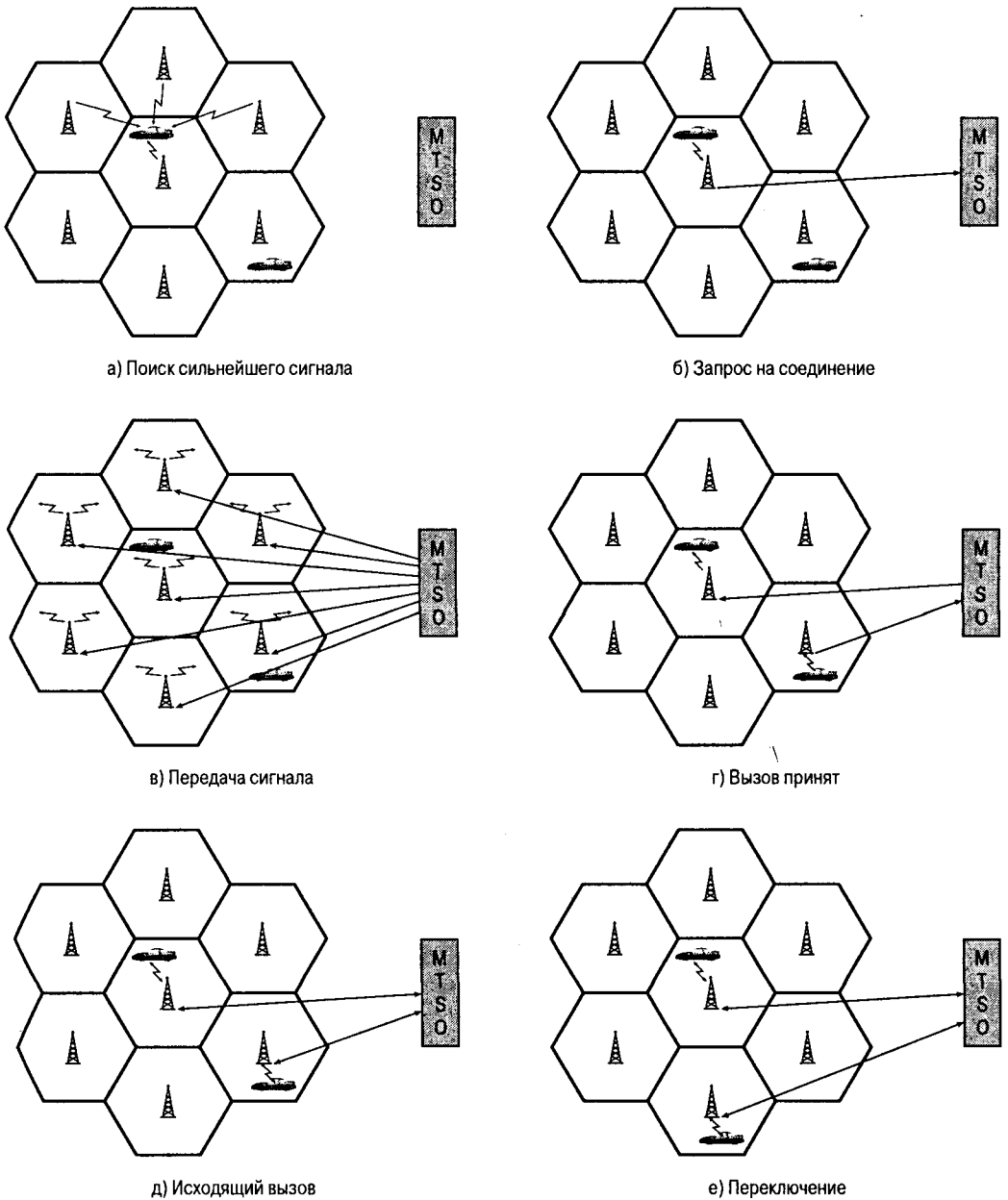


Рис. 10.6. Пример мобильного сотового соединения

- **Звонок с мобильного устройства.** Звонок с мобильного устройства начинается с отправки номера вызываемого устройства по предварительно выбранному каналу (см. рис. 10.6, б). Приемник мобильного устройства сначала проверяет, свободен ли настроенный канал, анализируя информацию в прямом (от базовой станции) канале. Когда обнаруживается, что канал свободен, мобильное устройство может начинать передачу в соответствующем



обратном (к базовой станции) канале. Базовая станция в свою очередь отправляет запрос на коммутатор MTSO.

- **Избирательный вызов.** Далее коммутатор MTSO пытается установить связь с вызываемым устройством. Коммутатор отправляет адресное сообщение определенной базовой станции, в зависимости от номера вызывающего мобильного устройства (см. рис. 10.6, в). Каждая базовая станция передает сигналы избирательного вызова в собственном выделенном настроечном канале.
- **Принятие вызова.** Вызываемое мобильное устройство распознает свой номер в настроечном канале, за которым следит в настоящий момент, и отвечает данной базовой станции. Базовая станция отправляет ответ на коммутатор MTSO, который устанавливает канал связи между вызывающей и вызываемой базовыми станциями. В то же самое время коммутатор MTSO выбирает подходящий канал информационного обмена внутри ячейки каждой базовой станции и уведомляет каждую базовую станцию, которые в свою очередь уведомляют свои мобильные устройства (см. рис. 10.6, г). Оба мобильных устройства настраиваются на выделенные им каналы.
- **Текущий вызов.** Пока поддерживается соединение, два мобильных устройства обмениваются голосовыми сигналами или данными, проходящими через соответствующие базовые станции и коммутатор MTSO (см. рис. 10.6, д).
- **Переключение.** Если мобильное устройство во время соединения выходит за пределы одной ячейки и входит в зону действия другой, то старый информационный канал следует заменить каналом, выделенным новой базовой станции в новой ячейке (см. рис. 10.6, е). Система осуществляет это изменение, не прерывая звонка и не беспокоя пользователя.

Система также выполняет некоторые другие функции, не представленные на рис. 10.6.

- **Блокирование вызова.** Если при звонке с мобильного устройства все информационные каналы, выделенные ближайшей базовой станции, заняты, то мобильное устройство предпринимает предварительно заданное количество последовательных попыток установления связи. После определенного количества неудачных попыток пользователю возвращается сигнал “занято”.
- **Завершение вызова.** Когда один или оба пользователя вешают трубку, об этом узнает коммутатор MTSO и освобождает информационные каналы обеих базовых станций.
- **Потеря вызова.** Если в определенный период соединения из-за интерференции или слабого сигнала базовая станция не может поддерживать минимально требуемую интенсивность сигнала, то информационный канал связи с пользователем прерывается, о чем уведомляется коммутатор MTSO.
- **Звонки стационарным и удаленным мобильным абонентам/от стационарных и удаленных мобильных абонентов.** Коммутатор MTSO подключен к коммутатору общественной телефонной сети. Это означает, что коммутатор MTSO может устанавливать соединение между мобильным пользователем из своей зоны и стационарным абонентом через телефонную сеть. Более того, MTSO может соединяться через телефонную сеть либо через выделенные каналы связи с удаленными MTSO и устанавливать соединение между мобильным пользователем из своей зоны и удаленным мобильным пользователем.

## Эффекты распространения радиоволн в мобильной связи

Радиосвязь с подвижными объектами имеет свои сложности, отсутствующие в проводных или стационарных беспроводных системах связи. Особого внимания требуют два аспекта: интенсивность сигнала и эффекты распространения сигнала.

- **Интенсивность сигнала.** Интенсивность сигнала между базовой станцией и мобильным устройством должна быть достаточно высокой, чтобы поддерживать качество сигнала на приемнике, но при этом не слишком высокой, чтобы не создавать сильной интерференции с каналами других ячеек, которые используют ту же полосу частот. Ситуацию дополнительно усложняют несколько факторов. Уровень искусственного шума бывает очень разным. Например, шум зажигания автомобилей в диапазоне частот сотовой связи в городе гораздо больше, чем в пригороде. Есть и другие источники шумовых сигналов, характеристики которых в большой степени зависят от места. Поэтому интенсивность сигнала является переменной величиной и ведет себя как функция расстояния от базовой станции до точки в пределах ячейки. Кроме того, интенсивность сигнала динамически меняется по мере движения мобильного устройства.
- **Замирание.** Даже если интенсивность сигнала лежит в эффективном диапазоне, разрушить сигнал или вызвать его отклонение могут эффекты, возникающие при распространении сигнала. Замирание сигнала и различные контрмеры были рассмотрены в разделе 5.4.

При проектировании сотовой сети инженеры-связисты должны учитывать различные эффекты распространения сигнала, желаемый максимальный уровень мощности передачи на базовой станции и на мобильном устройстве, обычную характерную высоту антенны мобильного устройства и доступную высоту антенны базовой станции. На основе этих факторов и определяются размеры отдельной ячейки. К сожалению, как было отмечено выше, эффекты распространения являются динамическими и их бывает трудно предсказать. Наилучшее, что можно сделать, — это разработать модель, основываясь на опытных данных, и, применив эту модель к данной окружающей среде, разработать правило определения размера ячейки. Одна из наиболее широко используемых моделей была представлена Окумурой (Okumura) [OKUM68], а ее усовершенствование предложено Хатой (Hata) [HATA80]. В качестве исходных данных был принят подробный анализ района Токио и получена информация о потерях на тракте для городской среды. Модель Хаты является обобщением опытных фактов, в котором учтено много условий и типов сред. В городской среде предлагается следующее выражение для потерь в тракте.

$$L_{\text{дБ}} = 69,55 + 26,16 \lg f_c - 13,82 \lg h_r - A(h_r) + (44,9 - 6,55 \lg h_r) \lg d. \quad (10.1)$$

Здесь

- $f_c$  — частота несущей в мегагерцах, которая варьируется от 150 до 1500 МГц;
- $h_t$  — высота передающей антенны (базовой станции) в метрах, лежит в пределах 30–300 м;
- $h_r$  — высота принимающей антенны (мобильного устройства) в метрах, лежит в пределах 1–10 м;
- $d$  — расстояние между антеннами в километрах, варьируется в пределах 1–20 км;
- $A(h_r)$  — поправочный коэффициент для высоты мобильной антенны.

Для небольших или средних городов поправочный коэффициент вычисляется по такой формуле:

$$A(h_r) = (1,1 \lg f_c - 0,7)h_r - (1,56 \lg f_c - 0,8) \text{ дБ.}$$

Для больших городов множитель вычисляется по следующей формуле:

$$A(h_r) = 8,29[\lg(1,54 h_r)]^2 - 1,1 \text{ дБ для } f_c \leq 300 \text{ МГц,}$$

$$A(h_r) = 3,2[\lg(11,75 h_r)]^2 - 4,97 \text{ дБ для } f_c \geq 300 \text{ МГц.}$$

Чтобы оценить потери на тракте для пригорода, формулу потерь для городской среды (10.1) нужно изменить следующим образом:

$$L_{\text{дБ}}(\text{пригород}) = L_{\text{дБ}}(\text{город}) - 2[\lg(f_c/28)]^2 - 5,4.$$

Для потерь на открытом пространстве формула выглядит следующим образом:

$$L_{\text{дБ}}(\text{открытая местность}) = L_{\text{дБ}}(\text{город}) - 4,78 \lg(f_c)^2 - 18,733 \lg f_c - 40,98.$$

Модель Окумуры/Хаты считается одной из лучших по точности предсказания потерь на тракте и представляет собой практический способ оценки потерь на тракте для множества ситуаций [FREE97, RAPP97].

**Пример.** [FREE97]. Пусть  $f_c = 900$  МГц,  $h_t = 40$  м,  $h_r = 5$  м и  $d = 10$  км. Оцените потери на тракте для города средней величины.

$$\begin{aligned} A(h_r) &= (1,1 \lg 900 - 0,7) 5 - (1,56 \lg 900 - 0,8) \text{ дБ} = \\ &= 12,75 - 3,8 = 8,95 \text{ дБ,} \end{aligned}$$

$$\begin{aligned} L_{\text{дБ}} &= 69,55 + 26,16 \lg 900 - 13,82 \lg 900 - 8,95 + (44,9 - 6,55 \lg 40) \lg 10 = \\ &= 69,55 + 77,28 - 22,14 - 8,95 + 34,4 = 150,14 \text{ дБ.} \end{aligned}$$

## Переключение

Переключение<sup>2</sup> — это процедура изменения “прописки” мобильного устройства с одной базовой станции на другую при переходе мобильного пользователя из одной ячейки в другую. Переключение пользователя осуществляется в разных системах по-разному, при этом задействовано множество факторов. Здесь будет дан только краткий обзор этой процедуры.

Переключение может инициироваться сетью, тогда решение о переключении принимается исключительно на основе сетевых измерений сигналов, принимаемых от мобильного устройства. Существуют альтернативные схемы переключения с участием пользователя, которые позволяют мобильному устройству участвовать в принятии решения о переключении, предоставляя возможность обратной связи с сетью по сигналам, принимаемым мобильным устройством. В любом случае для принятия решения могут использоваться различные метрики производительности. В [HAAS00], например, приводятся такие.

<sup>2</sup> В американской документации стандартов сотовой сети используется термин *handoff*, в документации Международного телекоммуникационного союза (ITU) — термин *handover*. В технической литературе встречаются оба термина; поскольку они означают одно и то же, на русский язык их принято переводить одинаково — *переключение*. — Прим. ред.

- **Вероятность блокирования вызова.** Вероятность того, что новая ячейка окажется заблокированной вследствие превышения возможностей (пропускной способности) базовой станции. В таком случае решение о передаче мобильного устройства соседней ячейке принимается не на основе качества сигнала, а в зависимости от степени использования пропускной способности.
- **Вероятность потери соединения.** Вероятность того, что в процессе переключения звонок будет потерян.
- **Вероятность завершения звонка.** Вероятность того, что разрешенный звонок не будет прерван до его завершения.
- **Вероятность неудачного переключения.** Вероятность того, что переключение будет выполнено при неподходящих условиях приема.
- **Вероятность блокирования переключения.** Вероятность того, что переключение не будет завершено успешно.
- **Вероятность переключения.** Вероятность того, что до завершения разговора произойдет одно переключение.
- **Частота переключений.** Число переключений за единицу времени.
- **Длительность прерывания.** Промежуток времени в ходе переключения, в течение которого мобильный пользователь не будет соединен с базовой станцией.
- **Задержка переключения.** Расстояние, на которое переместится мобильное устройство от точки, где должно было произойти переключение, до точки, в которой оно действительно произошло.

Основным параметром, используемым при принятии решения о переключении, является измеренная интенсивность сигнала от мобильного устройства к базовой станции. Обычно на базовой станции сигнал усредняется по времени движения, чтобы отбросить быстрые флуктуации, возникающие вследствие эффектов многолучевого распространения. На рис. 10.7, а (на основе иллюстраций [POLL96]) показана усредненная мощность, принимаемая на двух смежных базовых станциях от мобильного устройства, которое движется от станции А, находящейся в точке  $L_A$ , к станции В, находящейся в точке  $L_B$ . Этот рисунок служит иллюстрацией различных стратегий переключения, используемых для определения момента переключения.

- **Относительная интенсивность сигнала.** Обслуживание мобильного устройства передается с базовой станции А базовой станции В в тот момент, когда интенсивность сигнала, принимаемого на станции В, впервые превысит интенсивность на станции А. Если в дальнейшем интенсивность сигнала на станции В упадет ниже значения, регистрируемого на станции А, то мобильное устройство снова будет переключено на станцию А. На рис. 10.7, а переключение происходит в точке  $L_1$ . В этой точке интенсивность сигнала, регистрируемого на базовой станции А, все еще достаточна, но имеет тенденцию к снижению. Поскольку вследствие эффектов многолучевого распространения интенсивность сигнала флуктуирует, то даже при усреднении мощности такой подход может дать эффект пинг-понга, когда устройство периодически переключается между двумя базовыми станциями.

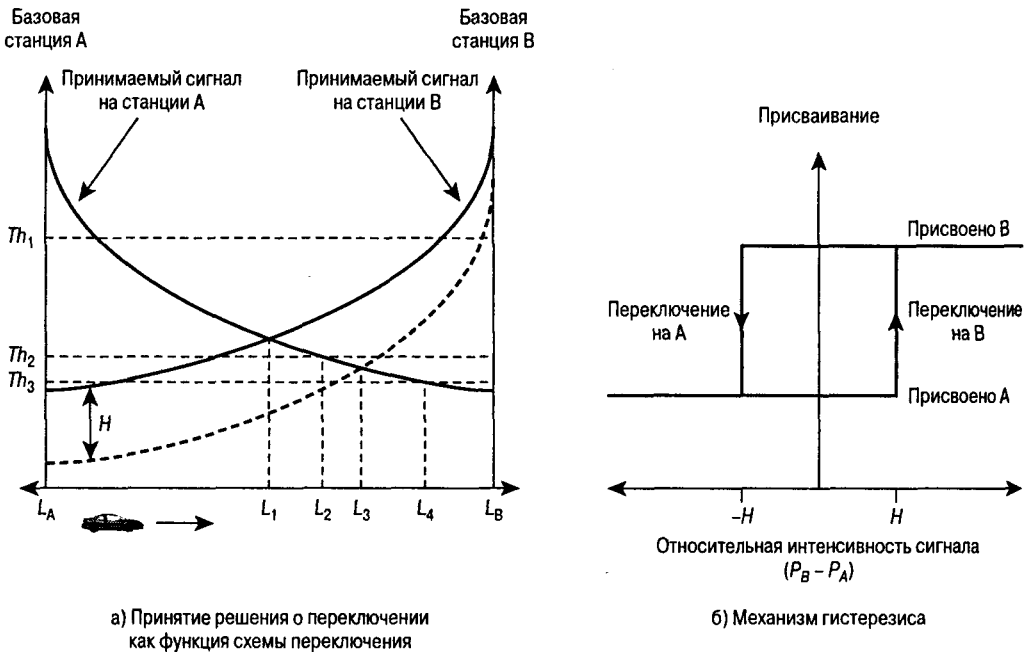


Рис. 10.7. Переключение между двумя ячейками

- Интенсивность сигнала относительно порога.** Переключение происходит, если (1) сигнал на текущей базовой станции становится довольно слабым (ниже установленного порога) и если (2) появляется другой, более сильный сигнал. Смысл этой стратегии состоит в том, что пока сигнал на текущей базовой станции является достаточно сильным, переключать мобильное устройство необязательно. Если использовать большой порог, например  $Th_1$ , эта схема будет равносильна схеме с определением относительной интенсивности сигнала. При пороговом значении  $Th_2$  переключение произойдет в точке  $L_2$ . Если задать достаточно низкий порог по сравнению с критической интенсивностью сигнала (т.е. с интенсивностью сигнала в точке  $L_1$ ), например  $Th_3$ , то мобильное устройство может зайти довольно глубоко в новую ячейку ( $L_4$ ), прежде чем произойдет переключение. Это снизит качество канала связи и способно привести к разрыву соединения. Пороговое значение нельзя определять, не учитывая другие параметры, поскольку его эффективность зависит от предварительных знаний о критической интенсивности сигнала между текущей базовой станцией и потенциальной новой базовой станцией.
- Относительная интенсивность сигнала и гистерезис.** Переключение происходит только в том случае, если сигнал на новой базовой станции значительно сильнее (на величину  $H$ , показанную на рис. 10.7, а), чем на текущей станции. Тогда переключение происходит в точке  $L_3$ . Эта схема предотвращает эффект пинг-понга, так как если переключение уже произошло, знак порогового значения  $H$  меняется на противоположный.

Термином *гистерезис* обозначается явление, известное как магнитный гистерезис, которое можно пояснить с помощью рис. 10.7, б. Механизм переключения можно представить состоящим из двух состояний. Пока мобильное устройство относится к станции А, переключение произойдет только тогда, когда относительная интенсивность сигнала достигнет порога  $H$  или превысит его. Мобильное устройство, переданное станции В, останется там до тех пор, пока относительная интенсивность сигнала не упадет ниже  $-H$ , и только тогда устройство будет передано обратно станции А. Единственный недостаток этой схемы: первая передача может все еще быть необязательной, если базовая станция А имеет достаточно интенсивный сигнал.

- **Относительная интенсивность сигнала, гистерезис и порог.** Переключение происходит, только если (1) текущий уровень сигнала упадет ниже порога и (2) базовая станция-кандидат окажется сильнее, чем текущая станция, на величину  $H$ . В нашем примере переключение мобильного устройства произойдет в точке  $L_3$ , если пороговое значение будет либо  $Th_1$ , либо  $Th_2$ , и в точке  $L_4$  — если пороговое значение будет равно  $Th_3$ .
- **Технологии прогнозирования.** Решение о переключении принимается на основе ожидаемого в будущем значения интенсивности принимаемого сигнала.

Принятие решения о переключении осложняется использованием схем регулирования мощности, которые позволяют базовым станциям динамически настраивать мощность сигналов, передаваемых мобильными устройствами. На этом вопросе мы остановимся подробно в следующем разделе.

## Регулирование мощности

Учитывая наличие многих спорных моментов, при проектировании в сотовой системе желательно предусмотреть возможность динамической регулировки мощности.

1. Чтобы сообщение проходило эффективно, мощность принимаемого сигнала должна существенно превышать фоновые шумы, в связи с этим предъявляются повышенные требования к мощности передатчика. По мере того как мобильное устройство удаляется от передатчика, принимаемая мощность уменьшается из-за обычного затухания. Кроме того, эффекты отражения, дифракции и рассеяния могут вызвать быструю смену уровней принимаемой мощности даже на небольших расстояниях. Ведь уровень мощности является суммой сигналов, приходящих по множеству различных путей, а фазы, соответствующие этим путям, являются случайными, так что при их сложении сигнал иногда усиливается, а иногда ослабляется. Отметим также, что при движении мобильного устройства воздействие сигналов, поступающих разными путями, может меняться.
2. В то же время желательно минимизировать мощность сигнала, передаваемого мобильным устройством, чтобы уменьшить интерференцию с каналами удаленных ячеек, использующих одну и ту же полосу частот, уменьшить вред, наносимый здоровью людей, и сэкономить энергию батарей.

3. В системах с расширенным спектром (SS), использующих множественный доступ с кодовым разделением каналов (CDMA), желательно выравнивать уровни мощности сигналов, принимаемых базовой станцией от разных мобильных устройств. Это влияет на производительность системы, так как всем пользователям выделены одни и те же частоты.

На рис. 10.8 (на основе иллюстраций из [PICH97]) показано два варианта регулирования мощности. Регулирование мощности по разомкнутому циклу зависит исключительно от мобильного устройства, без учета ответной реакции базовой станции, и используется в некоторых системах расширенного спектра. В таких системах базовая станция непрерывно передает немодулированный пилот-сигнал. Этот сигнал позволяет мобильному устройству синхронизироваться с прямым каналом (от базовой станции к мобильному устройству) CDMA и дает эталонную фазу для демодуляции. Его можно также использовать для контроля мощности. Мобильное устройство следит за уровнем мощности принимаемого контрольного сигнала и устанавливает передаваемую мощность в обратном канале (от мобильного устройства к базовой станции), обратно пропорциональную мощности сигнала. При таком подходе предполагается, что интенсивности сигналов в прямом и обратном каналах связи сильно коррелируют, как обычно и бывает. Регулирование по разомкнутому циклу не такое точное, как регулирование по замкнутому циклу. При этом первая схема позволяет оперативнее реагировать на быстрые флуктуации интенсивности сигнала, например, подобные тем, что возникают при появлении мобильного устройства из-за большого здания. Такая быстрая реакция требуется в обратном канале связи систем CDMA, где при случайном увеличении интенсивности принимаемого сигнала на базовой станции могут подавляться все прочие сигналы.

При регулировании мощности по замкнутому циклу выравнивается интенсивность сигнала в обратном канале (от мобильного устройства к базовой станции). При этом учитываются характеристики этого обратного канала, такие, как уровень мощности принимаемого сигнала, отношение сигнал/шум или частота появления в принимаемом сигнале ошибочных битов. Базовая станция принимает решение о регулировании мощности и передает команды регулирования мощности в канал управления мобильного устройства. Регулирование по замкнутому циклу также используется для выравнивания мощности в прямом канале. В этом случае мобильное устройство предоставляет базовой станции информацию о качестве принимаемого сигнала, а базовая станция затем регулирует передаваемую мощность.

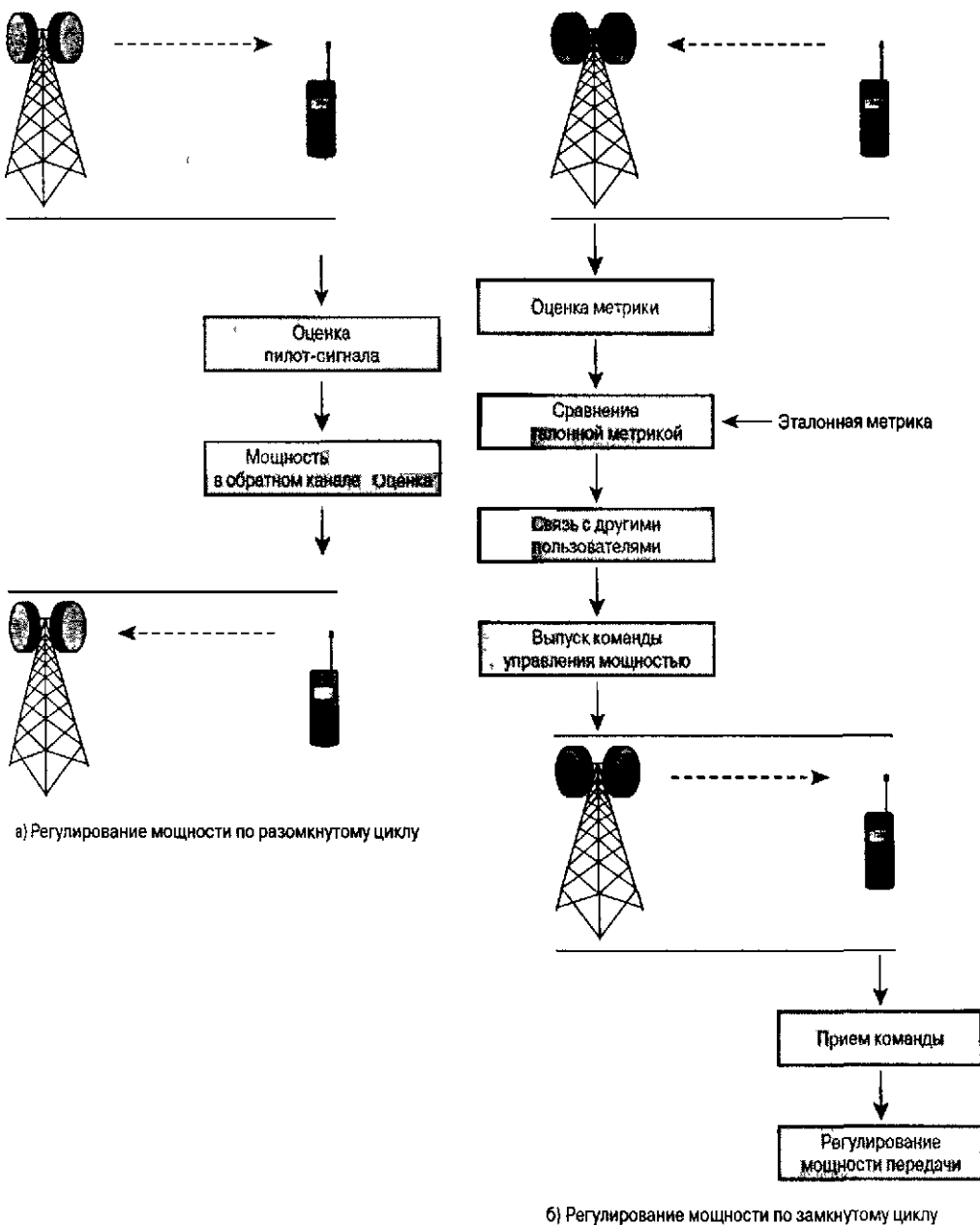


Рис. 10.8. Регулирование мощности в обратном канале

В табл. 10.2 показаны классы мощности, используемые в стандарте GSM, который является стандартом для систем TDMA и будет рассмотрен в разделе 10.3. В соответствии с мощностью выходных сигналов стандарт GSM определяет восемь классов каналов базовых станций и пять классов мобильных станций. Регулирование в обоих направлениях проводится по схеме замкнутого цикла.



Таблица 10.2. Классы передатчиков согласно стандарту GSM

Класс мощности	Мощность базовой станции, Вт	Мощность мобильной станции, Вт
1	320	20
2	160	8
3	80	5
4	40	2
5	20	0,8
6	10	
7	5	
8	2,5	

## Регулирование трафика

Для систем FDMA пропускная способность ячейки равна количеству выделенных ей частотных каналов. В идеальной ситуации количество доступных каналов в ячейке должно равняться общему числу абонентов, которые могут быть активны в любой момент. На практике же невозможно достичь такой пропускной способности, которая позволяла бы выдержать любую возможную нагрузку в любой момент времени. К счастью, не все абоненты заказывают разговоры одновременно, поэтому разумно иметь сеть такого размера, чтобы она могла выдерживать некоторый ожидаемый уровень трафика. Во многом производительность системы зависит от схемы регулирования трафика.

Концепция регулирования трафика была разработана при проектировании телефонных коммутаторов и телефонных сетей с коммутацией каналов, однако эта концепция в равной степени применима и к сотовым сетям. Рассмотрим ячейку, способную обслуживать  $N$  пользователей одновременно (т.е. пропускная способность которой равна  $N$  каналов), у которой есть  $L$  потенциальных абонентов ( $L$  мобильных устройств). Если  $L < N$ , систему называют *неблокируемой*; одновременно могут быть обработаны все звонки. Если  $L > N$ , систему называют *блокируемой*; абонент, пытающийся заказать разговор, может обнаружить, что пропускная способность использована полностью, т.е. заблокирована для него. Для блокируемых систем основными интересующими нас вопросами производительности являются следующие.

1. Какова степень блокирования, т.е. какова вероятность того, что запрос соединения будет заблокирован? Или какая пропускная способность ( $N$ ) необходима для того, чтобы достигнуть определенной верхней границы вероятности блокирования?
2. Если заблокированные соединения ожидают предоставления услуги, то каково среднее время задержки? Или какая пропускная способность нужна для достижения определенного значения средней задержки?

В этом разделе мы кратко изложим соответствующие концепции регулирования трафика и приведем примеры их применения. В приложении Б этот вопрос рассмотрен более подробно.

Величину загрузки системы определяют два параметра:

$\lambda$  — средняя частота звонков (запросов соединения), поступающих в единицу времени;

$h$  — среднее время разговора.

Основной мерой нагрузки является интенсивность трафика, которая выражается в безразмерных единицах, эрлангах.

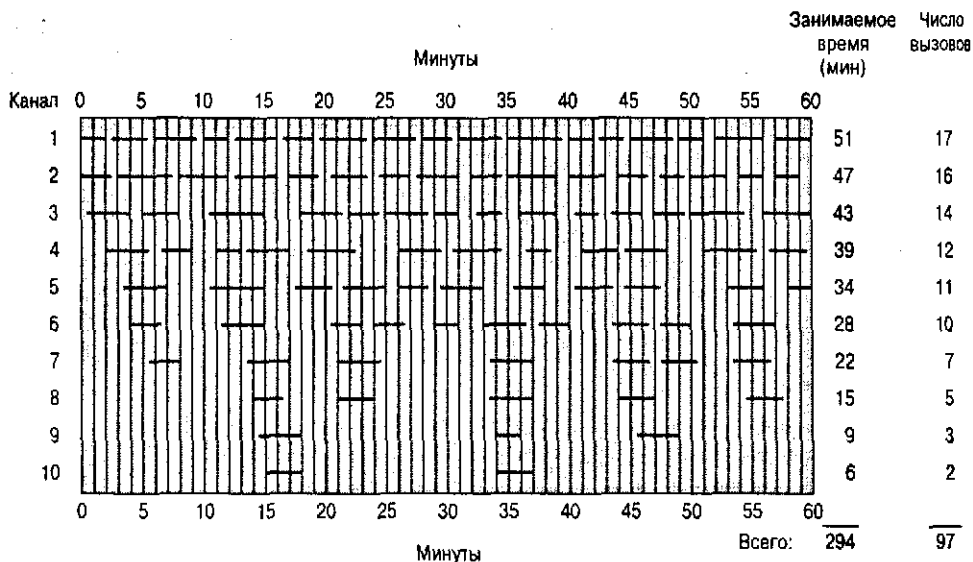
$$A = \lambda h.$$

Величину  $A$  можно интерпретировать несколькими способами. Это нормированная версия  $\lambda$ :  $A$  равно среднему числу звонков, поступающих в течение среднего времени разговора. Можно также рассматривать ячейку как систему обслуживания с несколькими серверами, в которой количество серверов равно пропускной способности  $N$ . Среднее время предоставления услуги на сервере равно  $h$ . Основным соотношением в системе обслуживания с несколькими серверами является  $\lambda h = \rho N$ , где  $\rho$  — использование сервера, или часть времени, в течение которого сервер занят. Отсюда  $A = \rho N$ , и параметр  $A$  является мерой среднего числа требуемых каналов.

**Пример.** Если частота вызовов в среднем составляет 20 вызовов в минуту, а среднее время разговора — 3 минуты, тогда  $A = 60$ . Можно ожидать, что ячейка с пропускной способностью 120 каналов будет использоваться наполовину в любой момент времени. В то же время коммутатора с пропускной способностью 50 будет явно недостаточно. Пропускная способность в 60 каналов будет удовлетворять средний спрос, однако из-за флуктуаций около среднего значения  $A$  временами ее будет не хватать.

**Пример.** Чтобы прояснить эти концепции, рассмотрим рис. 10.9, на котором показана модель функционирования ячейки с пропускной способностью 10 каналов в течение одного часа. Частота вызовов в минуту составляет 97/60. Среднее время разговора на один вызов в минутах составляет 294/97. Таким образом,  $A = (97/60) \times (294/97) = 4,9$  эрлангов. Другой способ трактовки параметра  $A$  — рассматривать его как среднее число обрабатываемых вызовов. Т.е. в среднем задействовано 4,9 канала. В то же время последняя интерпретация соответствует действительности только для неблокируемых систем. Параметр  $\lambda$  был определен как частота попыток соединения, а не обслуживаемого трафика.

Обычная блокируемая система ограничивается некоторым верхним пределом интенсивности трафика. Считается неразумным проектировать систему с учетом трафика, ожидаемого в пиковое время. Чаще всего системы проектируются в расчете на среднюю частоту попыток соединения, имеющую место в час наибольшей нагрузки. Час наибольшей нагрузки определяется как 60-минутный период в течение дня, когда трафик наибольший. Союз ITU-T рекомендует усреднять трафик в часы пик по 30 самым загруженным дням года; данный параметр называется “средним трафиком в часы пик”. В Северной Америке практикуют усреднение по 10 наиболее загруженным дням. Отметим, что это типичные измерения обработанного, а не запрашиваемого трафика, и они являются лишь оценкой истинной нагрузки.



Примечание: Горизонтальные линии указывают занимаемые периоды времени, округленные до 1/2 мин.

Рис. 10.9. Пример распределения нагрузки в ячейке с пропускной способностью 10 каналов

Параметр  $A$  — мера трафика в часы пик — является входным при моделировании нагрузки. Затем, исходя из конкретной модели, ищутся ответы на вопросы, подобные тем, что приведены в начале раздела. Природу модели определяют два ключевых фактора:

- способ обработки заблокированных звонков;
- число источников трафика.

Блокированные звонки можно обрабатывать двумя способами. Во-первых, их можно помещать в очередь ожидания свободного канала. Эта практика называется задержкой неудачного вызова (LCD), хотя в действительности вызов не является неудачным, он просто отложен. Во-вторых, заблокированный вызов можно отклонить или прервать. Это, в свою очередь, приводит к двум предположениям по поводу действий пользователя. Если пользователь вешает трубку, ожидает в течение некоторого случайного интервала времени и только потом возобновляет попытку дозвониться, то это называется очисткой неудачного вызова (LCC). Если же пользователь периодически пытается дозвониться, то это называется удержанием неудачного вызова (LCH). Для каждой из этих двух возможностей блокирования выведены формулы, которые характеризуют производительность системы. Для сотовых систем обычно используется модель LCC, которая, как правило, более точно описывает систему.

Вторым ключевым элементом модели трафика является предположение о том, конечно или бесконечно количество пользователей. В модели с бесконечным количеством источников нагрузки предполагается фиксированная частота прибытия звонков. Для конечного числа источников частота прибытия звонков зависит от числа уже занятых источников. В частности, если общее число пользователей равно  $L$  и каждый пользователь заказывает соединения со средней час-

тойой  $\lambda/L$ , тогда, если ячейка совсем не занята, частота прибытия звонков равна  $\lambda$ . В то же время если в момент времени  $t$  обслуживаются  $K$  пользователей, тогда мгновенная частота прибытия звонков в этот момент равна  $\lambda(L - K)/L$ . Отметим, что модели с бесконечным числом источников аналитически проще, а применять их стоит, если число источников по крайней мере в 5–10 раз превышает пропускную способность системы.

### Модель с бесконечным числом источников и очисткой неудачных вызовов

Для модели LCC с бесконечным числом источников ключевым интересующим нас параметром является вероятность потери вызова, или **уровень обслуживания**. Итак, уровень обслуживания 0,01 означает, что в течение часа повышенной нагрузки вероятность блокирования вызова равна 0,01. Вполне приемлемыми считаются значения вероятности в пределах 0,01–0,001.

Уравнение для модели LCC с бесконечным числом источников имеет следующий вид:

$$P = \frac{A^N}{N!} \cdot \frac{1}{\sum_{x=0}^N \frac{A^x}{x!}}$$

Здесь

- $A$  — поступающий трафик в эрлангах;
- $N$  — количество серверов;
- $P$  — вероятность блокирования (уровень обслуживания).

Это уравнение легко ввести в программу, кроме того, доступно множество таблиц с уже готовыми значениями параметра  $A$  (фрагмент одной из них приведен в табл. 10.3). Задавая поступающую нагрузку и количество серверов, можно подсчитать уровень обслуживания по формуле или определить его из таблицы. Чаще всего, правда, стоит обратная задача: определить трафик, который можно обслужить при данной пропускной способности, чтобы достичь данного уровня обслуживания. Еще одной задачей является определение пропускной способности, требуемой для обработки данного трафика при данном уровне обслуживания. Для решения таких задач нужны либо таблицы, либо подходящие программы, способные определить нужное значение методом перебора.

**Таблица 10.3. Качество обслуживания**

Пропускная способность (в эрлангах) для определенных значений уровня обслуживания					
Число серверов ( $N$ )	$P = 0,02$ (1/50)	$P = 0,01$ (1/100)	$P = 0,005$ (1/200)	$P = 0,002$ (1/500)	$P = 0,001$ (1/1000)
1	0,02	0,01	0,005	0,002	0,001
4	1,09	0,87	0,7	0,53	0,43
5	1,66	1,36	1,13	0,9	0,76
10	5,08	4,46	3,96	3,43	3,09
10	13,19	12,03	11,10	10,07	9,41

Число серверов ( $N$ )	$P = 0,02$ (1/50)	$P = 0,01$ (1/100)	$P = 0,005$ (1/200)	$P = 0,002$ (1/500)	$P = 0,001$ (1/1000)
24	16,64	15,27	14,21	13,01	12,24
40	31,0	29,0	27,3	25,7	24,5
70	59,13	56,1	53,7	51,0	49,2
100	87,97	84,1	80,9	77,4	75,2

Изучая табл. 10.3, можно сделать два важных вывода.

1. При одинаковом уровне обслуживания более эффективны системы с большей пропускной способностью.
2. Системы с большей пропускной способностью более восприимчивы к снижению уровня обслуживания.

**Пример.** Для иллюстрации первого утверждения рассмотрим две ячейки, каждая из которых имеет пропускную способность 10 каналов. Их общая пропускная способность равна 20 каналам, и вместе они могут справляться с поступающим трафиком интенсивностью 6,86, обеспечивая уровень обслуживания, равный 0,002. В то же время одна ячейка с пропускной способностью 20 каналов при уровне обслуживания 0,002 может обрабатывать трафик в 10,07 эрлангов. Для иллюстрации второго утверждения рассмотрим ячейку с 10 каналами, имеющую уровень обслуживания 0,002 и трафик 3,43 эрланга. При увеличении трафика на 30% уровень обслуживания уменьшается до 0,01. В то же время для ячейки с пропускной способностью 70 каналов уровень обслуживания падает с 0,002 до 0,01 всего лишь при 10%-ном увеличении трафика.

В приведенном примере фигурировал поступающий трафик, но если изменения различных метрик производятся в системе, то в них входит обработанный трафик. Несложно разработать программу, которая в качестве входных данных будет принимать обработанный трафик и превращать его в поступающий. Связь между обработанным трафиком  $C$  и поступающим трафиком  $A$  выглядит так:

$$C = A(1 - P).$$

Для малых значений  $P$   $A$  является хорошим приближением  $C$ .

### Эффект переключения

Одной из сложностей моделей трафика сотовой системы, отсутствующей в моделях других систем, является влияние переключения. Указанный эффект продемонстрирован на рис. 10.10. Частота поступления вызовов — это сумма двух величин: новые звонки, поступившие с мобильных устройств ячейки ( $\lambda_1$ ), и звонки мобильных устройств, переданные ячейке при входе этих устройств в ячейку ( $\lambda_2$ ). Общая частота поступления звонков:  $\lambda = \lambda_1 + \lambda_2$ . Точно так же, частота завершения звонков состоит из собственно прекращенных звонков и звонков, переключенных на другую ячейку. Таким образом, требуется соответствующая коррекция модели.

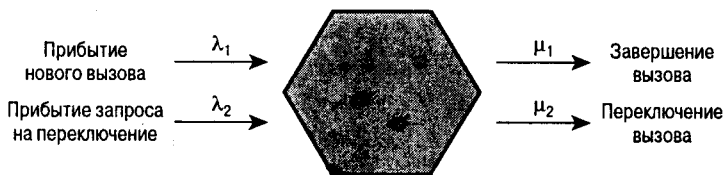


Рис. 10.10. Модель нагрузки ячейки

## 10.2. АНАЛОГОВЫЕ СИСТЕМЫ ПЕРВОГО ПОКОЛЕНИЯ

Первоначально сотовые телефонные сети предоставляли для обмена информацией аналоговые каналы. Теперь такие сети называют системами первого поколения. С начала 80-х годов самой популярной технологией первого поколения в Северной Америке была AMPS (Advanced Mobile Phone System — усовершенствованная система мобильной телефонной связи), разработанная компанией AT&T. Этот подход также получил распространение в Южной Америке, Австралии и Китае. Несмотря на то что на смену этим системам постепенно приходят системы второго поколения, технология AMPS используется все еще довольно широко. Краткий обзор этой технологии приводится ниже.

### Распределение спектра

В Северной Америке для системы AMPS выделены две полосы частот по 25 МГц (табл. 10.4): одна для передачи сигналов с базовой станции на мобильные устройства (869–894 МГц), вторая — для передачи сигналов с мобильных устройств на базовую станцию (824–849 МГц). Каждая полоса с целью создания условий для конкуренции разделена на две (т.е. чтобы на каждом рынке можно было разместить двух операторов). Оператору для его системы выделяется только 12,5 МГц в каждом направлении. Каналы разделены полосами по 30 кГц, что позволяет каждому оператору поддерживать по 416 каналов. Двадцать один канал выделен для управления, оставшиеся 395 — для передачи вызовов. Каналы управления представляют собой каналы обмена данными, в которых поддерживается скорость передачи данных 10 Кбит/с. Разговоры передаются по каналам для переговоров в аналоговом виде, с использованием частотной модуляции. Управляющую информацию также можно отправлять как данные по каналам для переговоров, в форме пакетов. Поскольку указанного количества каналов недостаточно для большинства основных рынков, приходится либо довольствоваться меньшей полосой частот на один разговор, либо внедрять многократное использование частот. В различных подходах к мобильной телефонии применяются оба варианта. Стандарт AMPS предусматривает многократное использование частот.

Таблица 10.4. Параметры системы AMPS

Полоса частот для передачи сигналов с базовой станции	869–894 МГц
Полоса частот для передачи сигналов с мобильного устройства	824–849 МГц

Расстояние между прямым и обратным каналами	45 МГц
Ширина канала	30 кГц
Количество полнодуплексных голосовых каналов	790
Количество полнодуплексных каналов управления	42
Максимальная мощность мобильного устройства	3 Вт
Размер ячейки, радиус	2–20 км
Модуляция в голосовом канале	Частотная модуляция, максимальное отклонение 12 кГц
Модуляция в канале управления	Частотная манипуляция, максимальное отклонение 8 кГц
Скорость передачи данных	10 Кбит/с
Код защиты от ошибок	БХЧ (48, 36, 5) и (40, 28, 5)

### Функционирование систем первого поколения

В постоянной памяти каждого сотового телефона, поддерживающего стандарт AMPS, имеется *модуль числового распределения* (numeric assignment module — NAM). В модуле NAM хранится номер телефона, выделенный ему поставщиком услуг, а также регистрационный номер телефона, присвоенный мобильному устройству производителем. Когда телефон включается, он начинает передавать коммутатору MTSO свой регистрационный номер и номер телефона (см. рис. 10.5). Коммутатор MTSO имеет базу данных с информацией о мобильных устройствах, которые заявлены украденными, поэтому зная регистрационный номер, можно заблокировать украденный телефон. Телефонный же номер используется коммутатором MTSO для операций со счетом. Если телефоном пользуются в другом городе, то счет за услуги все равно выставляется местному поставщику услуг.

При заказе разговора имеют место такие действия [COUC01].

1. Абонент инициирует вызов, вводя телефонный номер вызываемой стороны, и нажимает кнопку “send” (отправить).
2. Коммутатор MTSO проверяет, действителен ли номер телефона и уполномочен ли пользователь заказывать разговор; некоторые поставщики услуг требуют, чтобы пользователь вводил не только номер телефона вызываемого абонента, но и PIN-код (личный номер) во избежание использования украденных устройств.
3. Коммутатор MTSO посылает на сотовый телефон пользователя сообщение, в котором указано, какой информационный канал следует использовать для отправления и приема сигналов.
4. Коммутатор MTSO также отправляет сигнал вызываемой стороне. Действия 2–4 происходят в течение 10 с процесса инициации вызова.
5. После ответа вызываемой стороны коммутатор MTSO устанавливает канал связи между двумя абонентами и начинает считать стоимость разговора.
6. Когда один из абонентов вешает трубку, коммутатор MTSO освобождает канал связи и радиоканалы передачи данных и вычисляет окончательную стоимость разговора.

## Каналы управления в системе AMPS

Система AMPS состоит из 21 полнодуплексного канала управления шириной 30 кГц. Каждый из них включает в себя 21 обратный канал (reverse control channel — RCC) от абонента к базовой станции и 21 прямой канал от базовой станции к абоненту. По этим каналам передаются (в форме кадров) данные с использованием схемы частотной манипуляции (FSK).

На рис. 10.11, а показана структура кадра канала RCC. Кадр начинается с 48-битового поля-предтечи, состоящего из 30-битового поля тактовой синхронизации (чередующиеся нули и единицы), 11-битовой последовательности синхронизации слов (11100010010) и 7-битового цифрового кода цвета (digital color code — DCC). Код DCC используется для распознавания данных, передаваемых в ячейках внутренних каналов; он представляет собой уникальный идентификатор базовой станции и служит адресом назначения для кадра RCC. За предтечей следует собственно информационный кадр, который содержит от одного до шести слов данных. Каждое слово вмещает 36 бит данных и закодировано с помощью укороченной версии кода БХЧ (63, 51, 5) (см. табл. 8.4). В этой укороченной версии к 36 битам данных добавляется 12 контрольных разрядов и образуется 48-битовое слово. Для дальнейшего повышения надежности каждое слово передается пять раз в одном и том же кадре, а для восстановления слова на базовой станции используется мажоритарная логика. Если учесть все служебные сигналы, то скорость передачи данных будет порядка нескольких сотен битов в секунду. Примерами сообщений RCC являются сообщения инициации связи, отклик на избирательный вызов и подтверждение порядка.

Структура кадра FCC (см. рис. 10.11, б) начинается с 10-битовой последовательности тактовой синхронизации и 11-битовой последовательности синхронизации слов. Каждый кадр состоит из двух слов данных. Каждое слово закодировано с помощью кода БХЧ и содержит 28 бит данных и 12 контрольных битов. Как и ранее, из соображений надежности каждое слово повторяется пять раз. Кроме того, в каждом кадре FCC предоставляется информация о состоянии соответствующего кадра RCC (свободен или занят), для чего применяется бит “занято/свободно”, который вводится в каждый десятый бит кадра. Таким образом, общий объем кадра равен 463 бит. При скорости передачи сигналов 10 Кбит/с скорость передачи данных (за вычетом служебных сигналов) составляет около 1,2 Кбит/с. Сообщения FCC включают в себя сообщения избирательного вызова и сообщения о распределении частот.

Наконец, управляющую информацию можно передавать по голосовым каналам в течение разговора. Мобильное устройство или базовая станция могут составлять пакеты данных, прекращая передачу голоса примерно на 100 мс и заменяя его FSK-модулированным сообщением. Этот вариант используется для обмена срочными сообщениями, например связанными с изменением мощности или переключением.





и большей пропускной способности. В [BLAC99] перечислены следующие ключевые различия этих двух поколений.

- **Цифровые информационные каналы.** Наиболее заметным отличием систем двух поколений является то, что системы первого поколения практически полностью аналоговые, в то время как системы второго поколения являются цифровыми. В частности, системы первого поколения спроектированы для поддержки голосовых каналов с использованием частотной модуляции; цифровые данные можно передавать только с использованием модема, который преобразует цифровые данные в аналоговую форму. Системы второго поколения предоставляют цифровые информационные каналы. Эти каналы уже поддерживают передачу цифровых данных, а голосовые каналы сначала нужно закодировать в цифровую форму, а лишь затем передавать. Разумеется, в системах второго поколения пользовательские данные (информация и оцифрованная речь) нужно вначале преобразовать в аналоговые сигналы, а лишь потом передавать их между мобильным устройством и базовой станцией (см., например, рис. 6.14).
- **Шифрование.** Поскольку весь обмен пользовательскими сообщениями и управляющими данными в системах второго поколения является цифровым, то для защиты от подслушивания данные довольно легко зашифровать. Эту возможность предоставляют все системы второго поколения, а системы первого поколения отправляют пользовательские данные в чистом виде, не обеспечивая никакой защиты.
- **Обнаружение и исправление ошибок.** В системах второго поколения применяются схемы обнаружения и исправления ошибок наподобие тех, что были рассмотрены в главе 8. В результате можно обеспечить довольно чистый прием речи.
- **Доступ к каналам.** В системах первого поколения каждая ячейка поддерживает несколько каналов. В любой момент времени канал может быть выделен только одному пользователю. В системах второго поколения ячейкам также выделяется по несколько каналов, однако каждый канал может совместно использоваться несколькими пользователями посредством схем множественного доступа с временным разделением (TDMA) или множественного доступа с кодовым разделением (CDMA). Системы, основанные на TDMA, будут рассмотрены в этом разделе, а системы, основанные на CDMA, — в разделе 10.4.

Начиная с 90-х годов, было внедрено немало различных систем второго поколения. В табл. 10.5 перечислены некоторые ключевые характеристики трех наиболее важных систем этого поколения.

**Таблица 10.5. Сотовые телефонные системы второго поколения**

	GSM	IS-136	IS-95
Год введения	1990	1991	1993
Метод доступа	TDMA	TDMA	CDMA
Полоса частот для передачи сигналов базовой станции	935–960 МГц	869–894 МГц	869–894 МГц
Полоса частот для передачи сигналов мобильного устройства	890–915 МГц	824–849 МГц	824–849 МГц

	GSM	IS-136	IS-95
Расстояние между прямым и обратным каналами	45 МГц	45 МГц	45 МГц
Ширина канала	200 кГц	30 кГц	1250 кГц
Число дуплексных каналов	125	832	20
Максимальная мощность мобильного устройства	20 Вт	3 Вт	0,2 Вт
Количество пользователей на один канал	8	3	35
Схема модуляции	GMSK	$\pi/4$ DQPSK	QPSK
Скорость передачи битов на несущей	270,8 Кбит/с	48,6 Кбит/с	9,6 Кбит/с
Кодер речи	RPE-LTP	VSELP	QCELP
Скорость кодирования битов речи	13 Кбит/с	8 Кбит/с	8, 4, 2, 1 Кбит/с
Размер кадра	4,6 мс	40 мс	20 мс
Код коррекции ошибок	Сверточный код со степенью кодирования 1/2	Сверточный код со степенью кодирования 1/2	Сверточный код со степенью кодирования 1/2 в прямом канале; со степенью кодирования 1/3 в обратном канале

## Множественный доступ с временным разделением каналов

Доступ многих пользователей к сотовой системе первого поколения осуществляется с помощью технологии FDMA (frequency division multiple access — множественный доступ с частотным разделением). Этот термин был введен ранее, при рассмотрении спутниковой связи; для сотовых систем принцип работы схемы не изменился, и его можно описать следующим образом. Для каждой ячейки выделяется  $2M$  каналов, шириной  $\delta$  Гц каждый. Половина каналов (обратные каналы) используются для передачи данных с мобильного устройства на базовую станцию:  $f_c, f_c + \delta, f_c + 2\delta, \dots, f_c + (M - 1)\delta$ , где  $f_c$  — центральная частота в самом низкочастотном канале. Оставшаяся половина каналов (прямые каналы) используется для передачи данных с базовой станции на мобильное устройство:  $f_c, f_c + \delta + \Delta, f_c + 2\delta + \Delta, \dots, f_c + (M - 1)\delta + \Delta$ , где  $\Delta$  — расстояние между обратным и прямым каналами. При установлении соединения с мобильным пользователем для полнодуплексного сообщения выделяются два канала — на частоте  $f$  и на частоте  $f + \Delta$ . Такой порядок предоставления каналов довольно неэкономичен, так как большую часть времени один или оба канала остаются незанятыми.

Технология разделения каналов TDMA уже упоминалась при обсуждении спутниковой связи (см., например, рис. 9.14). Применение схемы TDMA в сотовой системе можно описать следующим образом. Так же, как и при использовании FDMA, каждой ячейке выделяется некоторое количество каналов, половина

из которых используются для обратной связи, а половина — для прямой. И снова для полнодуплексного соединения мобильному устройству выделяется пропускная способность на парных — прямом и обратном — каналах. Кроме того, каждый физический канал подразделяется на несколько логических каналов. Передача данных осуществляется в виде последовательности кадров с повторяющейся структурой: каждый кадр делится на некоторое число слотов. Положение каждого слота в последовательности кадров определяет отдельный логический канал. Пример логического канала приводился на рис. 9.13.

## Вопросы проектирования мобильных беспроводных систем TDMA

Перед тем как переходить к рассмотрению системы GSM, будет полезно изучить некоторые общие принципы, проанализировав данные из [JONE93]. После этого анализа станут понятны мотивы некоторых проектных решений, принятых в GSM. Общая цель состоит в определении такой длины и структуры слота трафика, которые обеспечат эффективную передачу речи и данных, а также эффективное использование спектра радиодиапазона. Рассмотрим следующий набор требований.

- **Количество логических каналов (число слотов в кадре TDMA)** — 8; оказывается, что это минимальное количество слотов, при котором оправдываются затраты на разделение каналов.
- **Максимальный радиус ячейки (R)** — 35 км; он обеспечивает достаточный уровень трафика в сельской местности.
- **Частота** — около 900 МГц; обычно именно такая частота выделяется для мобильных радиоприложений.
- **Максимальная скорость движения ( $V_m$ )** — 250 км/ч или 69,4 м/с, чтобы можно было пользоваться мобильными устройствами в скоростных поездах.
- **Максимальная задержка кодирования** — приблизительно 20 мс, чтобы не приводить к чрезмерным задержкам в стационарных сетях, в которых может использоваться спутниковая связь.
- **Максимальный разброс задержек ( $\Delta_m$ )** — 10 с (в горных районах); эта задержка представляет собой разницу задержек распространения различных компонентов многолучевых сигналов, прибывающих на одну и ту же антенну.
- **Ширина полосы частот** не должна превышать 200 кГц, что соответствует 25 кГц на один канал (именно эта величина в настоящее время принята для европейских аналоговых сотовых систем с частотной модуляцией).

На рис. 10.12 приведены шаги, которые предлагается выполнить при проектировании слота системы TDMA. Этой схемой мы будем руководствоваться при дальнейшем обсуждении.

Кодер речи должен обеспечивать удовлетворительное качество речи при минимальной скорости передачи данных. Традиционным видом речевого кодирования, дающим поток цифровых битов, является импульсно-кодовая модуляция (PCM), которая приводит к скорости передачи данных 64 Кбит/с (см. раздел 6.4). Эта скорость слишком высока для сотовой связи. Для имеющихся в настоящее время технологий разумно использовать скорость передачи данных около 12 Кбит/с, что дает хорошее качество воспроизведения речи.



Рис. 10.12. Проектирование слота для системы TDMA

Если ограничить задержку кодирования величиной 20 мс, тогда закодированную речь можно группировать блоками по 20 мс или выборками речи по 240 бит. При скорости 12 Кбит/с данные также группируются в 240-битовые блоки. Затем к этим блокам можно применять схемы исправления ошибок.

В цифровых системах второго поколения обычно используются сверточные коды коррекции ошибок со степенью кодирования 1/2. При таком кодировании число битов в блоке данных увеличивается до 480. Кроме того, к блоку данных нужно добавить восемь бит, чтобы учесть длину регистра сдвига (см. раздел 8.3). Таким образом, длина блока речи составит 488 битов.

При выбранных параметрах минимальная скорость передачи битов для восьмиканальной системы составит

$$\frac{8 \text{ каналов} \times 488 \text{ бит/канал}}{20 \times 10^{-3} \text{ с}} = 195,2 \text{ Кбит/с.}$$

В действительности потребуется несколько более высокая скорость передачи битов, что обусловлено другими соображениями, о которых мы поговорим ниже. Это означает, что в доступной полосе шириной 200 кГц придется обеспечивать скорость передачи данных выше, чем 200 Кбит/с. На практике таких скоростей передачи данных невозможно достигнуть без использования адаптивного выравнивания. Как обсуждалось в разделе 5.4, для осуществления адаптивного выравнивания

нивания в мобильной среде потребуется включать новую настроечную последовательность всякий раз, когда мобильное устройство переместится на расстояние, достаточное для изменения характеристик тракта передачи. Предположим, что настроечная последовательность включена в каждый слот. Грубый критерий, предложенный в [JONE93], состоит в том, что фаза несущего сигнала после настроечной последовательности не должна превышать  $1/20$  длины волны (т.е.  $\lambda/10$ ). При частоте 900 МГц длина волны составляет 0,333 м. Итак,

$$\text{максимальная длительность передачи} = \frac{\lambda / 20}{V_m} = \frac{0,333 / 20}{69,4} = 0,24 \text{ мс}.$$

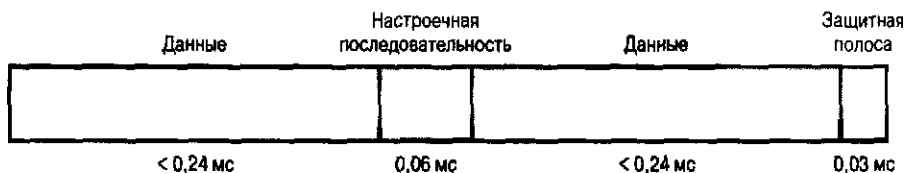
Настроечную последовательность можно использовать более выгодно, если передавать 0,24 мс речи или данных до нее и после нее и использовать настроечную последовательность в сочетании с 0,48 мс данных.

Далее требуется определить длину настроечной последовательности. При проектировании эквалайзера для многолучевого сигнала, ширина полосы которого примерно равна скорости передачи битов (200 кГц, 200 Кбит/с), используют эвристическое правило: число отводов эквалайзера (см. рис. 5.14) должно в шесть раз превышать число битов, передаваемых за максимальное время рассеивания ( $\Delta_m = 0,01$  мс). Таким образом, для настроечной последовательности в слоте выделяется 0,06 мс.

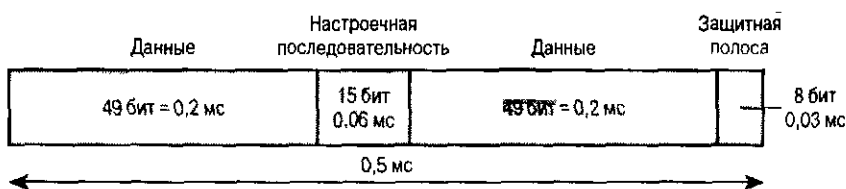
Теперь предположим, что в конце каждого слота нужен защитный интервал. Его вводят из-за различия во времени задержек между разными мобильными устройствами и базовой станцией. Поскольку один и тот же кадр TDMA совместно используют восемь мобильных устройств, требуется синхронизировать передачу данных с мобильных устройств, чтобы данные с одного из них не приходились на чужие слоты. В функции базовой станции входит предоставление каждому мобильному устройству информации о синхронизации, что позволяет выравнять относительные задержки и сохранять структуру кадра TDMA. Однако мобильные устройства могут двигаться и относительно базовой станции, и относительно друг друга, так что для учета этих отклонений после каждого слота вставляется защитный интервал. Когда мобильное устройство впервые заказывает соединение базовой станции, базовая станция предоставляет мобильному устройству информацию о синхронизации, которая зависит от текущего времени задержки распространения сигнала между мобильным устройством и базовой станцией. Избежать частого обновления информации о синхронизации можно путем введения достаточного защитного времени. Приведем пример расчета защитного интервала. Средний телефонный звонок занимает 130 с [JONE93], так что радиальное расстояние до базовой станции, которое может покрыть мобильное устройство, составляет  $(130 \text{ с}) \times (69,4 \text{ м/с}) = 9022 \text{ м}$ . Изменение задержки распространения сигнала, обусловленное перемещением на такое расстояние, составляет  $9022 / (3 \times 10^8 \text{ м/с}) = 0,03 \text{ мс}$ .

На рис. 10.13, а показана полученная схема слота. Следующий этап состоит в распределении закодированного блока данных между несколькими слотами, в которые также войдут настроечная последовательность и защитные биты. Максимальная длительность слота равна приблизительно 0,57 мс. Учитывая, что каждый кадр состоит из 8 слотов, получаем длину кадра 4,6 мс. Пусть данные нужно отправить с задержкой кодирования 20 мс, так что если длину кадра считать равной 4 мс, а длину каждого слота равной 0,5 мс, удобно будет отправить блоки речи в пяти последовательных слотах в одном и том же канале. Блок речи

состоит из 488 бит, так что в каждый слот следует поместить  $488/5$  или около 98 бит данных. Это даст скорость передачи битов  $98/0,4 = 245$  Кбит/с. При такой скорости передачи данных минимальное количество требуемых настроечных битов составляет  $(0,06 \text{ мс}) \times (245 \text{ Кбит/с}) = 14,7$ , что при округлении даст 15 бит. Подобным образом рассчитывается и минимальное количество защитных битов, которое оказывается равным  $(0,03 \text{ мс}) \times (245 \text{ Кбит/с}) = 7,35$ , т.е. при округлении в большую сторону — 8 бит.



а) Временная схема поля



б) Пространственная схема поля

Рис. 10.13. Слот TDMA

Получившаяся в результате структура кадра показана на рис. 10.13, б. По каналу со скоростью передачи 242 Кбит/с передается 121 бит за 0,5 мс.

## Глобальная система мобильной связи

До того как была разработана Глобальная система мобильной связи (GSM), в странах Европы использовалось множество разных несовместимых сотовых телефонных технологий первого поколения. Стандарт GSM был разработан для внедрения в Европе общей технологии второго поколения, чтобы одни и те же абонентские устройства можно было использовать по всему континенту. Эта технология оказалась весьма успешным и, возможно, самым популярным мировым стандартом для систем нового поколения. Впервые стандарт GSM появился в 1990 году в Европе. Теперь подобные системы внедрены в Северной и Южной Америке, Азии, Северной Африке, а также в Средней Азии и в Австралии. Ассоциация GSM заявила, что к концу 2000 года она обслуживала три четверти всех абонентов сотовой сети по всему миру, причем подавляющее большинство абонентов GSM находилось в Европе и на тихоокеанском побережье Азии, 8 миллионов проживает также в Северной Америке.

## Архитектура сети GSM

На рис. 10.14 показаны ключевые функциональные элементы системы GSM. Линии Um, Abis и A обозначают интерфейсы между функциональными элементами, которые стандартизированы в документации GSM. Таким образом, можно

приобретать оборудование у разных поставщиков и ожидать, что оно будет успешно взаимодействовать. В стандарте GSM определены также дополнительные интерфейсы, но здесь они рассматриваться не будут.

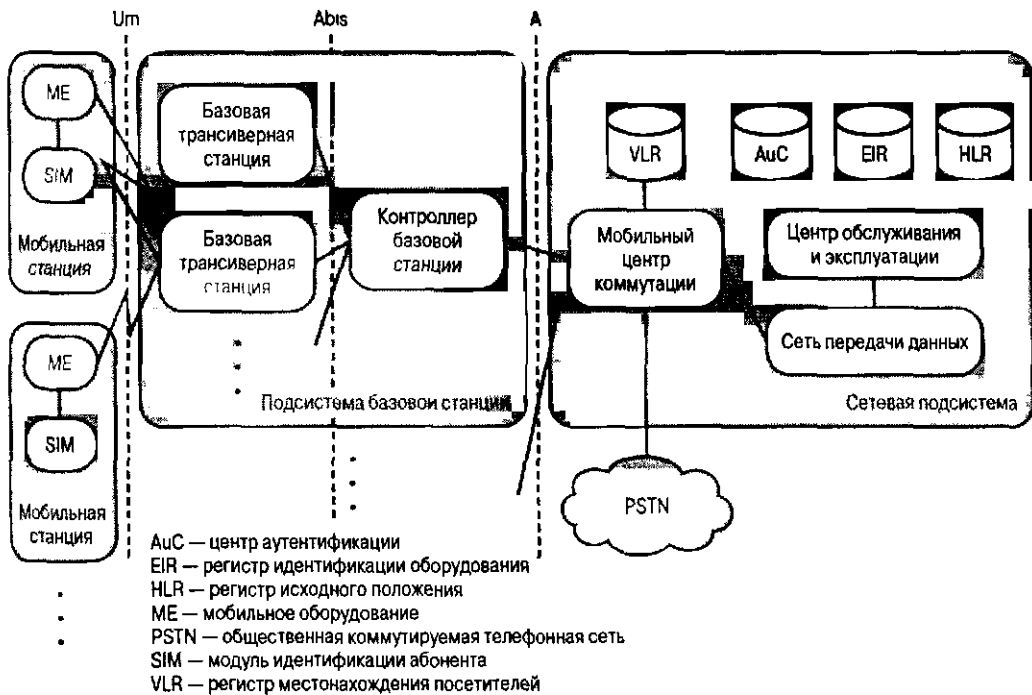


Рис 10 14 Общая архитектура GSM

### Мобильная станция

Через интерфейс Um, называемый также радиointерфейсом, мобильная станция общается с трансивером базовой станции в той ячейке, в которой находится мобильное устройство. Термином мобильное оборудование (mobile equipment — ME) обозначается физический терминал, такой, как телефон или устройство персональной службы связи (personal communication service — PCS), включающее в себя радиотрансивер, процессоры для обработки цифровых сигналов и модуль идентификации абонента (subscriber identity module — SIM). SIM представляет собой портативное устройство, имеющее вид интеллектуальной карточки или встраиваемого модуля, в котором хранится идентификационный номер абонента, координаты сетей, которыми разрешено пользоваться абоненту, ключи шифрования и другая информация об абоненте. Абонентские устройства GSM до вставки модуля SIM абсолютно неотличимы друг от друга. Поэтому путешествующий абонент, захвативший с собой свой модуль SIM, может в разных странах использовать разные устройства, вставляя в них свой модуль. В действительности, за исключением определенных срочных соединений, абонентские устройства не будут работать без вставленного модуля SIM. Поэтому носить с собой везде нужно именно модуль SIM, а само устройство брать с собой необязательно.



## Подсистема базовой станции

Подсистема базовой станции (base station subsystem — BSS) состоит из контроллера базовых станций и одной или нескольких базовых трансиверных станций. Каждая базовая трансиверная станция (base transceiver station — BTS) определяет ячейку, в которую входит радиоантенна, радиотрансивер и канал связи с контроллером базовых станций. Ячейка GSM может иметь радиус от 100 м до 35 км, в зависимости от среды. Контроллер базовой станции (base station controller — BSC) может совмещаться с BTS или управлять работой нескольких устройств BTS, а следовательно, несколькими ячейками. Контроллер BSC резервирует радиочастоты, управляет переключениями мобильных устройств с одной ячейки на другую в пределах одной подсистемы BSS и контролирует избирательное обращение.

## Сетевая подсистема

Сетевая подсистема (network subsystem — NS) обеспечивает связь между сотовой сетью и общественными коммутируемыми телекоммуникационными сетями. Подсистема NS управляет переключениями между ячейками, находящимися в различных подсистемах базовых станций, опознает пользователей и подтверждает достоверность их счетов, а также выполняет функции роуминга мобильных пользователей. Центральным элементом подсистемы NS является **мобильный центр коммутации** (mobile switching center — MSC). Он управляет четырьмя базами данных.

- **База данных регистра исходного положения (home location register — HLR).** В регистре HLR хранится информация, как временная, так и постоянная, о каждом из абонентов, который “принадлежит” системе (т.е. об абонентах, телефонные номера которых связаны с центром коммутации).
- **База данных регистра местонахождения посетителей (visitor location register — VLR).** Одну из важных частей информации составляет местонахождение абонента. Местонахождение определяется из регистра VLR, в который введен абонент. В регистре местонахождения посетителей хранится информация об абонентах, которые в данный момент физически находятся в районе, обслуживаемом данным центром коммутации. В регистре отмечается, является ли абонент активным, а также фиксируются другие параметры, связанные с абонентом. При поступлении звонка абоненту система использует связанный с абонентом телефонный номер для опознания исходного для данного абонента центра коммутации. Этот центр коммутации, в свою очередь, в своем регистре HLR может найти центр коммутации, в зоне действия которого в данный момент физически находится абонент. При поступлении звонка от абонента регистр VLR используется для инициирования звонка. Даже если абонент находится в зоне, принадлежащей его исходному центру коммутации, он может также быть представлен в регистрах VLR других центров коммутации.
- **База данных центра аутентификации (authentication center — AuC).** Эта база данных используется в процессе аутентификации; например, в ней хранятся ключи аутентификации и шифрования для всех абонентов, представленных как в регистрах исходного положения, так и в регистрах местонахождения посетителей. Центр управляет доступом к данным пользователей, а также процессом аутентификации при присоединении абонента к сети. Данные, передаваемые системами GSM, шифруются, поэтому они конфиденциальны. Для шифровки данных, передаваемых от абонента тран-

сиверу базовой станции, используется поточный шифр А5. В то же время переговоры по сети с наземными линиями связи проходят без шифрования. Другой поточный шифр, А3, используется для аутентификации.

- **База данных регистра идентификации оборудования (equipment identity register — EIR).** В этой базе данных хранятся записи о типе оборудования, которое имеется на мобильной станции. Эта база данных также важна для безопасности (например, для блокирования звонков с украденных мобильных устройств и предотвращения использования сети станциями, которым не было дано такого разрешения).

## Аспекты радиосвязи

Системам GSM выделяется 25 МГц для передачи данных с базовых станций (935–936 МГц) и 25 МГц для передачи с мобильных устройств (890–915 МГц). Вне Европы для систем GSM определены также и другие полосы частот. Пользователи получают доступ к сети с помощью комбинации схем FDMA (множественный доступ с частотным разделением) и TDMA (множественный доступ с временным разделением) (обе схемы были рассмотрены в предыдущей главе). Несущие расположены через каждые 200 кГц, т.е. всего имеется 125 полнодуплексных каналов. Каналы модулированы для использования скорости передачи данных 270,833 Кбит/с. Как и в системе AMPS, существует два типа каналов — информационные каналы и каналы управления.

### Формат TDMA

В системах GSM используется сложная иерархия кадров TDMA, определяющая логические каналы (рис. 10.15). По сути, каждая полоса частот шириной 200 кГц делится на 8 логических каналов, определяемых повторяющимися последовательностями слотов.

На низшем уровне иерархии находится слот, длительность которого равна  $15/26$  мс, или примерно 0,577 мс. При скорости передачи битов 270,833 Кбит/с каждый слот будет иметь длину 156,25 бит. Слот состоит из следующих полей.

- **Завершитель.** Позволяет синхронизировать передачу данных с мобильных устройств, находящихся на разных расстояниях от базовой станции (будет рассмотрено ниже).
- **Зашифрованные биты.** Данные шифруются в блоки с помощью обычного шифрования 114 открытых битов в 114 зашифрованных битов. Затем зашифрованные биты помещаются в два 57-битовых поля слота.
- **Захваченные биты.** Используются для указания, содержит ли блок данные или же он “захвачен” для срочного сигнала управления.
- **Настроечная последовательность.** Используется для настройки параметров приемника в соответствии с особенностями текущего тракта распространения сигнала, а при многолучевом распространении еще и для выбора наиболее сильного сигнала. Настроечная последовательность является известной комбинацией битов, индивидуальной для каждой ячейки. Она позволяет мобильным устройствам и базовой станции определять, действительно ли пришел принимаемый сигнал с передатчика или же он возник в результате сильных помех.

Кроме того, настроечная последовательность используется для выравнивания сигналов, распространяющихся по множественным траекториям, которое позволяет отделять нужный сигнал от нежелательных отражений. Зная, как изменяется известная настроечная последовательность вследствие многолучевого замирания, в оставшейся части сигнала этот эффект можно компенсировать.

**Защитные биты.** Используются во избежание наложения пакетов данных вследствие разных задержек на тракте.

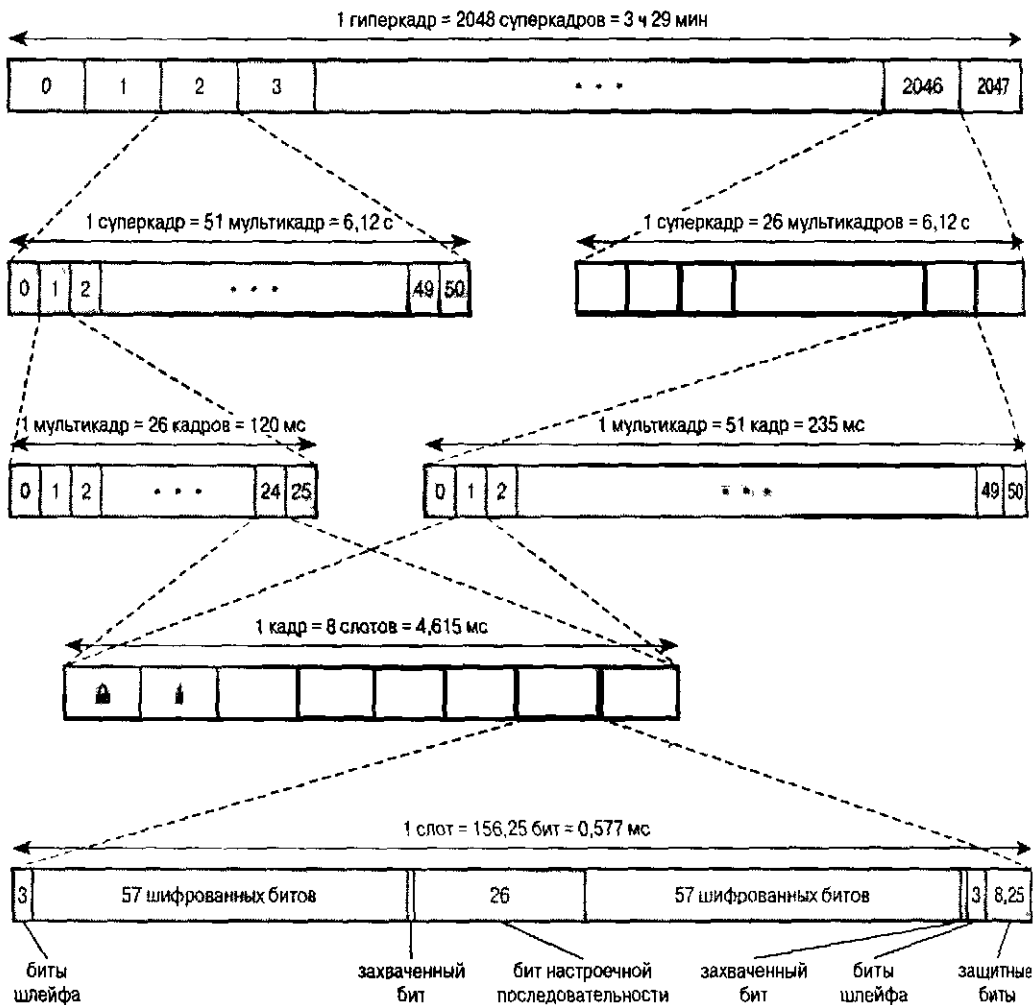


Рис. 10.15. Формат кадра GSM

Формат слота, показанный на рис. 10.15, называется нормальным пакетом и несет пользовательские данные (сравните с рис. 10.13, б). Для управления передачей сигналов используются другие форматы пакетов.

Более крупной иерархической единицей являются мультикадры из 26 кадров, в которые обычно собираются 8-слотовые кадры TDMA. Один из кадров мультикадра

используется для управления передачей сигналов, а другой пока не используется, т.е. для передачи данных используются 24 кадра. Таким образом, каждому информационному каналу выделяется по одному слоту на кадр и по 24 кадра на каждый мультикадр длительностью 120 мс. Вычислим скорость передачи данных.

$$\frac{114 \text{ бит/слот} \times 24 \text{ слота/мультикадр}}{10 \text{ мс/мультикадр}} = 22,8 \text{ Кбит/с}.$$

В спецификации GSM допускаются также информационные каналы половинной скорости передачи данных, когда имеются два информационных канала, каждый из которых занимает один слот в 12 из 26 кадров. При использовании кодеров речи с половинной скоростью передачи данных пропускная способность системы удваивается. Есть также мультикадры, содержащие по 51 кадру, которые используются для управления трафиком.

### Кодирование речи

На рис. 10.16 показана схема обработки речевых сигналов для передачи данных по логическому информационному каналу. Все этапы обработки последовательно рассмотрены ниже.

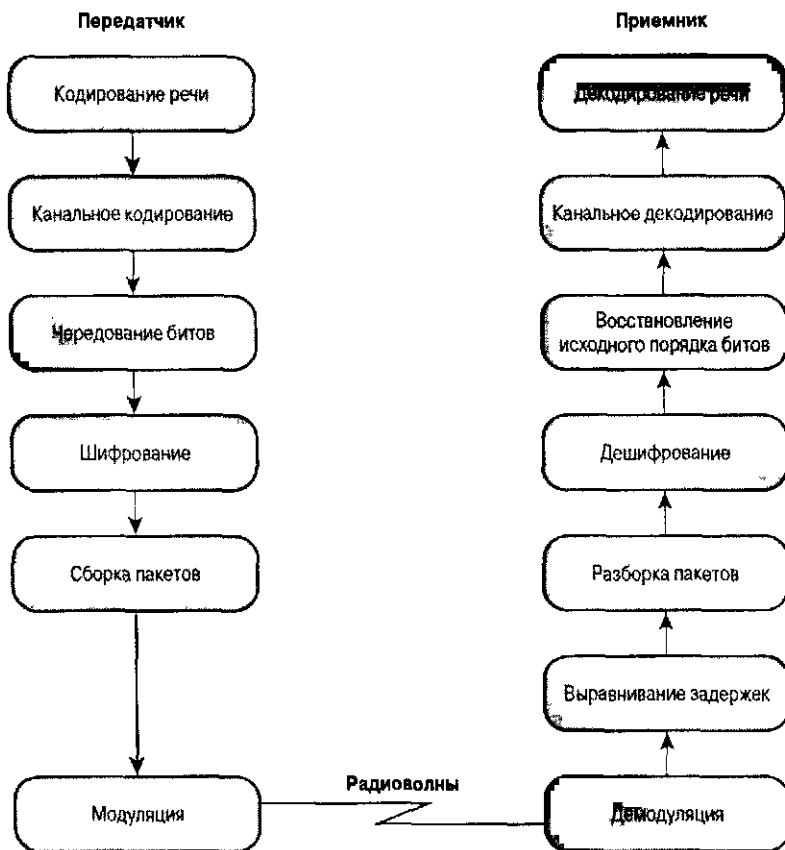


Рис. 10.16. Обработка речевых сигналов в системе GSM

Речевой сигнал сжимается с помощью алгоритма RPE-LPE (Regular Pulse Excited — Linear Predictive Coder, активизация регулярными импульсами — линейный кодер с предсказанием) [KRO086]. Суть алгоритма состоит в том, что для предсказания текущей выборки используются данные из предыдущих выборок. При кодировании каждая выборка представляется в виде линейной комбинации предыдущих выборок и описывается с помощью коэффициентов этой линейной комбинации и закодированной разности предсказанной и действительной выборок. В результате такого кодирования каждые 20 мс получается 260 бит, т.е. приблизительно скорость передачи данных составляет 13 Кбит/с. По качеству речи, воспроизводимой с помощью такого кодирования, биты в 260-битовом блоке можно разделить на три класса.

- **Класс Ia:** 50 бит, наиболее чувствительных к битовым ошибкам.
- **Класс Ib:** 132 бит, умеренно чувствительных к битовым ошибкам.
- **Класс IIa:** 78 бит, минимально чувствительных к битовым ошибкам.

Для защиты первых 50 бит используется 3-битовая циклическая проверка четности с избыточностью (CRC). При обнаружении ошибки вся выборка отбрасывается и заменяется измененной версией предыдущего блока. Эти 53 бит, 132 бит класса Ib, а также 4-битовая остаточная последовательность затем защищаются сверточным кодом (1, 2, 5), что в итоге дает  $189 \times 2 = 378$  бит. Оставшиеся 78 бит остаются незащищенными и прилагаются к защищенным битам для того, чтобы дополнить блок до 456 бит, что в результате даст скорость передачи данных  $456/20 \text{ мс} = 22,8 \text{ Кбит/с}$ , которая является характерной скоростью передачи данных в информационных каналах систем GSM.

Для того чтобы защитить передаваемые данные от пакетов ошибок, каждый 456-битовый блок делится на восемь 57-битовых блоков, которые передаются в восьми последовательных слотах. Так как каждый слот может вмещать в себя два 57-битовых блока, в каждом пакете оказываются данные из двух различных выборок речи.

С помощью описанных шагов речевые данные шифруются по 114 бит за такт, собираются в слоты (сборка пакета) и, наконец, модулируются для передачи. Используемая схема модуляции, GMSK (Gaussian Minimum Shift Keying — гауссова манипуляция с минимальным частотным сдвигом), представляет собой разновидность частотной модуляции (FSK).

### Кодирование данных

Цифровые данные обрабатываются тем же способом, что применяется к речевым сигналам. Данные обрабатываются блоками по 240 бит каждые 20 мс, что дает скорость передачи данных 12 Кбит/с. В зависимости от определения логических каналов реальная поддерживаемая скорость передачи данных может составлять 9,6, 4,8 и 2,4 Кбит/с. Каждый блок дополняется четырьмя остаточными битами. Для получения блока размером  $244 \times 2 = 488$  бит используется сверточный код (1, 2, 5). Затем 32 бит из этого блока опускаются, и в блоке остается 456 бит. Для распределения данных по пакетам используется схема чередования битов, что, опять же, уменьшает влияние импульсных помех. 488 бит распределяются по 22 пакетам следующим образом:

- 1-й и 22-й пакеты содержат по 6 бит каждый;
- 2-й и 21-й пакеты содержат по 12 бит каждый;

- 3-й и 20-й пакеты содержат по 18 бит каждый;
- пакеты с 4-го по 19-й содержат по 24 бит каждый.

В результате каждый пакет переносит информацию из 5 или 6 последовательных блоков данных.

### **Медленная скачкообразная перестройка частоты**

Выше было сказано, что информационному каналу выделяется определенная частота как для передачи, так и для приема данных. Это не совсем верно. В GSM и во многих других сотовых схемах используется технология, известная как медленная скачкообразная перестройка частоты, которая призвана повысить качество сигнала. Каждый следующий кадр TDMA в данном канале переносится на своей несущей частоте. Таким образом, частота передачи изменяется один раз каждые 4,615 мс. Медленная скачкообразная перестройка частоты позволяет компенсировать эффект многолучевого замирания, который зависит от несущей частоты. Медленная скачкообразная перестройка частоты позволяет также уменьшить эффект интерференции соседних каналов. Напомним, что эта технология является разновидностью связи с расширенным спектром.

### **Выравнивание задержек**

Поскольку мобильные устройства находятся на разных расстояниях от базовой станции в пределах одной ячейки, передаваемые ими данные приходят с разными задержками. Это явление создает проблемы при проектировании системы, так как один кадр TDMA одновременно использует до восьми мобильных устройств. Потому решающее значение имеет синхронизация слотов кадра. Базовая станция передает разным мобильным устройствам управляющие сигналы, которые позволяют синхронизировать распределение слотов. В формате кадра предусмотрены остаточные биты и защитные биты, которые обеспечивают зазор, позволяющий предотвратить перекрытие битов данных из разных слотов. Базовая станция может синхронизировать любое активное мобильное устройство, посылая управляющие сигналы, в которых будет указано, следует ускорить или замедлить отсчет времени.

## **Протокольная архитектура передачи сигналов GSM**

Между ключевыми объектами, изображенными на рис. 10.14, связанными с мобильностью, радиоресурсами и управлением соединением, происходит интенсивный обмен управляющими сообщениями. Подробное описание различных форматов и семантики сообщений заняло бы целую книгу. Здесь будет представлен только краткий обзор структуры, из которого станет видна сложность проектирования систем второго поколения.

На рис. 10.17 обобщаются протоколы, используемые между основными элементами архитектуры сети. Нижний уровень архитектуры предназначен для физической связи между объектами. Между мобильной станцией и базовой станцией-трансивером действует радиосвязь, описанная в предыдущих разделах, посредством которой переносятся данные высших уровней. Между другими объектами используется стандартный цифровой канал со скоростью передачи данных 64 Кбит/с.

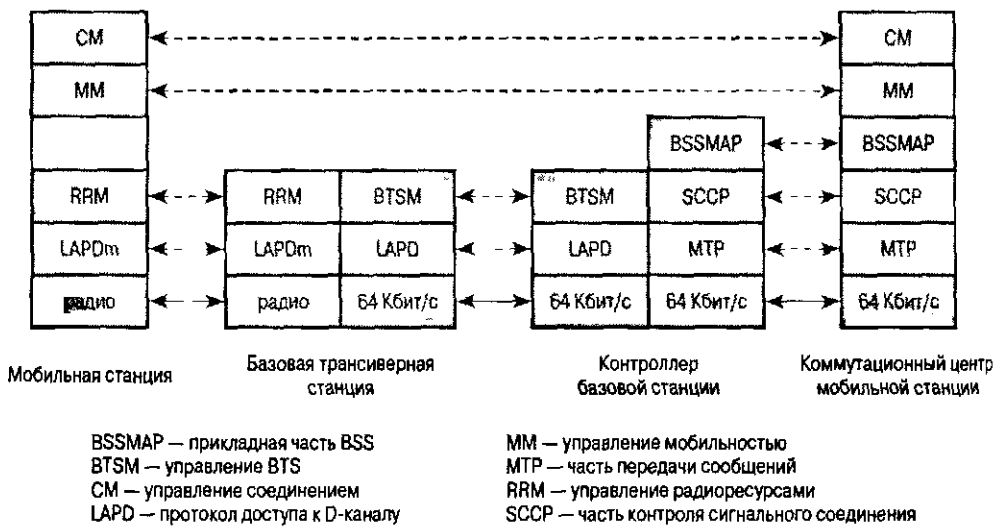


Рис. 10.17. Протокольная архитектура передачи сигналов GSM

На канальном уровне используется протокол управления каналом передачи данных (см. рис. 4.3), известный как LAPDm. Он представляет собой модифицированную версию протокола LAPD, определенного для сети ISDN (Integrated Services Digital Network — цифровая сеть с интеграцией услуг). Оставшиеся каналы связи используют обычный протокол LAPD. По сути, протокол LAPD разработан для преобразования потенциально ненадежных физических каналов связи в надежные каналы передачи данных. Осуществляется это посредством использования циклической проверки четности с избыточностью, которая проводится для обнаружения ошибок (как было описано в разделе 8.1), а также автоматического запроса на повторение (ARQ) для повторной передачи поврежденных кадров (как было описано в разделе 8.4)<sup>3</sup>.

Над канальным уровнем размещено множество протоколов, которые обеспечивают выполнение особых функций, включая перечисленные ниже.

- **Управление радиоресурсами.** Управление установкой, эксплуатацией и удалением радиоканалов, включая переключения.
- **Управление мобильностью.** Процедуры обновления и регистрации местоположения, а также защита и аутентификация.
- **Управление соединением.** Управление заказом, обслуживанием и завершением звонков (соединений между пользователями).
- **Мобильная прикладная часть (mobile application part — MAP).** Осуществление передачи большинства сигналов между различными объектами, находящимися в фиксированных точках сети, например между регистром HLR и регистром VLR.
- **Управление базовой трансиверной станцией.** Выполнение различных функций управления и административных функций на базовой трансиверной станции под руководством контроллера базовых станций.

<sup>3</sup> Рассмотрение протоколов управления передачей данных приведено в приложении Г.

MAP находится не непосредственно над канальным уровнем, а над двумя промежуточными протоколами, SCCP и MTP. Эти протоколы являются частью Системы передачи сигналов номер 7 (SSN7), которая представляет собой набор протоколов, разработанных для обеспечения контроля над передачей сигналов в коммутируемых сетях, таких, как цифровые общественные телекоммуникационные сети. Эти протоколы обеспечивают выполнение общих функций, которые используются различными приложениями, включая MAP.

## 10.4. СИСТЕМЫ CDMA ВТОРОГО ПОКОЛЕНИЯ

Множественный доступ с кодовым разделением каналов (code division multiple access — CDMA) представляет собой основанную на расширении спектра схему, которая была рассмотрена в разделе 7.4 и которая является второй, после TDMA, альтернативой разделения каналов для сотовых сетей второго поколения. Этот раздел начинается с краткого обзора преимуществ подхода CDMA, а затем рассматривается наиболее широко используемая схема такого типа — IS-95.

### Множественный доступ с кодовым разделением каналов

Схему CDMA для сотовых систем можно описать следующим образом. Как и при использовании FDMA, каждой ячейке выделяется некоторая полоса частот, которая делится на две части — половина для обратного канала (от мобильного устройства к базовой станции) и половина для прямого канала (от базовой станции к мобильному устройству). Для полnodуплексной связи мобильное устройство использует оба канала, прямой и обратный. Передача данных осуществляется по схеме расширения спектра методом прямой последовательности (DSSS), при котором используется раздробленный код для увеличения скорости передачи данных, а это дает увеличение полосы сигнала. Множественный доступ обеспечивается путем присваивания нескольким пользователям ортогональных раздробленных кодов, чтобы приемник мог выделить отдельные передачи из общего сигнала.

Использование схемы CDMA для сотовых сетей имеет несколько преимуществ.

- **Частотное разнесение.** Из-за того что передаваемые сигналы рассеяны по широкой полосе частот, искажение передачи на определенной частоте, например вследствие шума или селективного замирания, меньше влияет на сигнал.
- **Снижение негативных эффектов многолучевого распространения.** Кроме способности DSSS противостоять многолучевому замиранию посредством частотного разнесения, раздробленные коды, используемые для CDMA, имеют низкую автокорреляцию. Поэтому версия сигнала, который задерживается более чем на один элементарный интервал, не интерферирует с доминирующим сигналом, как это происходит в других средах многолучевого распространения.
- **Конфиденциальность.** Так как расширенный спектр получается посредством использования шумоподобных сигналов, где каждый пользователь имеет свой код, такой схеме изначально присуща конфиденциальность.
- **Постепенное снижение эффективности функционирования системы.** В схемах TDMA или FDMA к системе может одновременно обращаться только фиксированное число пользователей. В CDMA же по мере того как все больше пользователей получают одновременный доступ к системе, уровень



шума и, следовательно, частота появления ошибок увеличиваются. Система приближается к моменту неприемлемого качества связи постепенно.

Следует упомянуть и о многочисленных недостатках сотовой схемы CDMA.

- **Самоподавление.** Если все мобильные пользователи синхронизированы не идеально, данные, поступающие от разных пользователей, не будут идеально выравнены по границам раздробленных кодов. Расширенные последовательности разных пользователей не будут ортогональными, поэтому будет присутствовать некоторая взаимная корреляция. Это отличает данную схему от FDMA и TDMA, в которых существуют защитные временные или частотные интервалы, соответственно, т.е. сигналы являются хотя бы приближенно ортогональными.
- **Проблема расположения.** Сигналы, поступающие от более близких к передатчику объектов, не так затухают, как сигналы, пришедшие издалека. При отсутствии полной ортогональности данные, передаваемые с удаленного мобильного устройства, бывает довольно трудно восстановить. Поэтому в системах CDMA очень важное значение имеют схемы регулирования мощности.
- **Мягкое переключение.** Как будет показано далее, для плавности переключения с одной ячейки в следующую требуется, чтобы мобильное устройство вошло в новую ячейку до того, как оно оставит старую. Это называется мягким переключением, которое является более сложным, чем жесткое переключение, используемое в схемах FDMA и TDMA.

## Вопросы проектирования мобильных беспроводных систем CDMA

Перед тем как переходить к конкретному примеру — системе GSM, — будет полезно рассмотреть некоторые вопросы проектирования, общие для всех сотовых систем CDMA.

### RAKE-приемник

Если в многолучевой среде, которая является весьма распространенной для сотовых систем, множественные версии сигнала пребывают на расстоянии более одного элементарного интервала друг от друга, то приемник может восстановить сигнал, соотнося элементарные последовательности с доминирующим входящим сигналом. Оставшиеся сигналы рассматриваются как шум. Однако более производительным оказывается восстановление сигналов, прошедших по разным маршрутам, и последующее их суммирование (с подходящими значениями задержек). Этот принцип используется в RAKE-приемниках.

На рис. 10.18 проиллюстрирован принцип действия RAKE-приемника. Исходный двоичный сигнал расширяется с помощью применения к нему и раздробленному коду передатчика операции исключающего ИЛИ (XOR). Затем расширенная последовательность модулируется для передачи по беспроводному каналу связи. Из-за эффектов многолучевого распространения в канале оказывается много копий сигнала, каждая из которых имеет свое время задержки ( $\tau_1$ ,  $\tau_2$ , и т.д.) и свой коэффициент затухания ( $a_1$ ,  $a_2$  и т.д.). В приемнике полученный сигнал демодулируется. Демодулированный поток элементов сигнала затем подается на блок корреляторов, причем каждый задерживается на определенную величину (свою для каждого сигнала). Затем сигналы суммируются с весовыми коэффициентами, подходящими для данного канала.

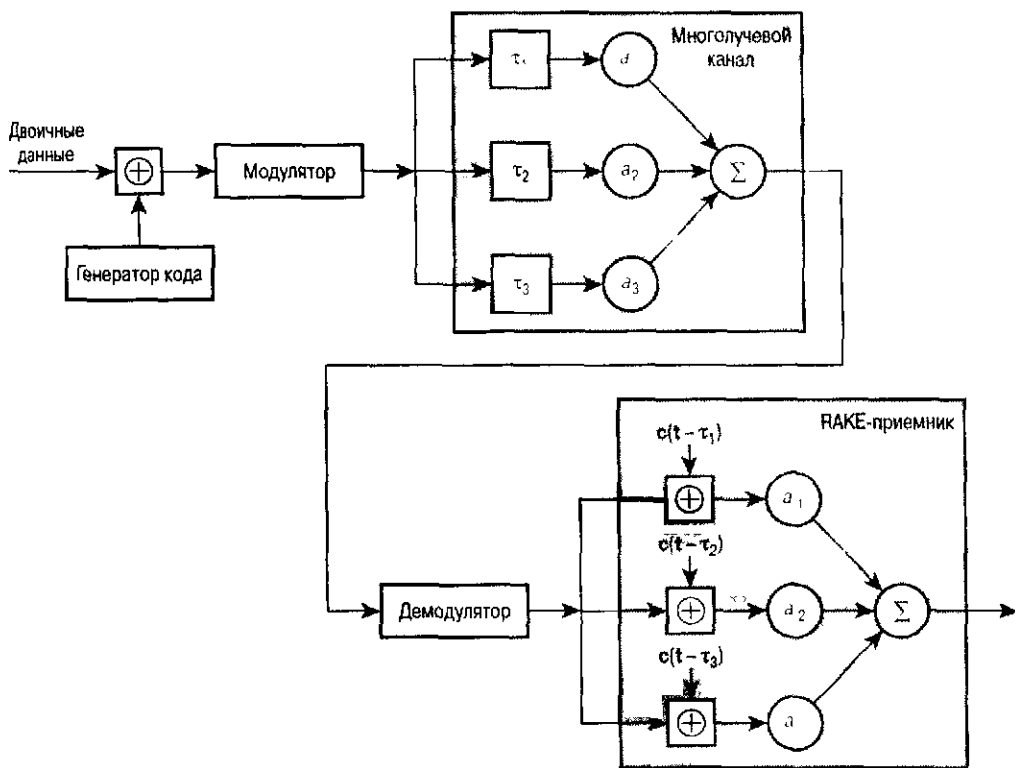


Рис. 10.18. Принцип работы RAKE-приемника [PRAS98]

### Мягкое переключение

В системах FDMA и TDMA соседние ячейки используют различные участки доступного спектра (т.е. кратность использования частоты больше единицы и обычно составляет  $N = 7$ ). Когда интенсивность сигнала в соседней ячейке превышает интенсивность в текущей ячейке, мобильной станции приходит указание переключиться на новую полосу частот, которая выделена для новой ячейки. Такая процедура называется **жестким переключением**. В типичной сотовой системе CDMA пространственное разделение частот не используется (т.е.  $N = 1$ ), так как большую часть времени интерференция с соседними ячейками не будет препятствовать корректному приему сигналов DSSS.

При мягком переключении мобильная станция временно подключается к более чем одной базовой станции одновременно. Мобильное устройство поначалу могло быть соединено с одной станцией. Если затем устройство входит в зону, в которой качество передачи двух базовых станций сравнимо (с некоторым порогом по отношению друг к другу), то мобильное устройство переходит в состояние мягкого переключения, в котором оно подключено к обеим базовым станциям. Мобильное устройство остается в этом состоянии до тех пор, пока одна из базовых станций не станет явно преобладать, и тогда устройство переключится непосредственно на эту ячейку.

Находясь в состоянии мягкого переключения, данные, передаваемые с мобильного устройства, достигают обеих базовых станций и отправляются в центр

коммутации, который оценивает качество обоих сигналов и выбирает один из них. Затем коммутатор отправляет данные или оцифрованную речь обеим станциям, которые передают эти сигналы на мобильное устройство. Мобильное устройство комбинирует два входящих сигнала и восстанавливает информацию.

## IS-95

Наиболее популярной схемой CDMA второго поколения является система IS-95, которая первоначально была внедрена в Северной Америке. В табл. 10.5 перечислены некоторые ключевые параметры системы IS-95. Структура передачи данных в прямом и обратном каналах связи отличается, потому будет описана для каждого направления отдельно.

### Прямой канал системы IS-95

В табл. 10.6 приведены параметры прямого канала связи. Прямой канал связи состоит из 64 логических каналов CDMA, каждый из которых занимает полосу частот шириной 1228 кГц (рис. 10.19, а). При прямой связи поддерживаются четыре типа каналов.

- **Пилот-сигнал (канал 0).** Постоянно передаваемый сигнал. Этот канал позволяет мобильному устройству получать информацию о синхронизации, дает точку отсчета для определения сдвига фаз в процессе демодуляции и средство для сравнения интенсивности сигнала для определения момента переключения. Пилот-канал состоит из всех нулей.
- **Синхронизация (канал 32).** Канал со скоростью передачи данных 1200 бит/с используется мобильной станцией для получения идентификационной информации о сотовой системе (системное время, состояние длинного кода, проверка протоколов и т.д.)
- **Избирательные сообщения (каналы 1–7).** Содержат сообщения для одной или нескольких мобильных станций.
- **Информационный обмен (каналы 8–31 и 33–63).** В прямом канале поддерживается 55 информационных каналов. В исходной спецификации указано, что поддерживаются скорости передачи данных до 9600 бит/с. Позже был добавлен второй набор скоростей до 14400 бит/с.

Обратите внимание, что все эти каналы используют одну и ту же полосу частот. Для распознавания разных каналов используется раздробленный код. Раздробленный код прямого канала включает 64 ортогональных 64-битовых кодов, полученных из матрицы Уолша размером  $64 \times 64$  (эта матрица была рассмотрена в разделе 7.5, см. рис. 7.17).

Таблица 10.6. Параметры прямого канала связи системы IS-95

Канал	Синхронизация	Избирательный вызов	Первый набор каналов	Первый набор информационных каналов	Второй набор информационных каналов	Второй набор информационных каналов					
Скорость передачи данных (бит/с)	1200	4800	9600	1200	2400	4800	9600	1800	3600	7200	14400
Количество повторений кода	2	1	8	4	2	1	8	4	2	1	1
Скорость передачи символов модуляции (символ/с)	4800	19200	19200	19200	19200	19200	19200	19200	19200	19200	19200
Количество чипов на СИМВОЛ МОДУЛЯЦИИ	256	64	64	64	64	64	64	64	64	64	64
Количество чипов на бит	1024	256	128	1024	512	256	128	682,67	341,33	170,67	85,33

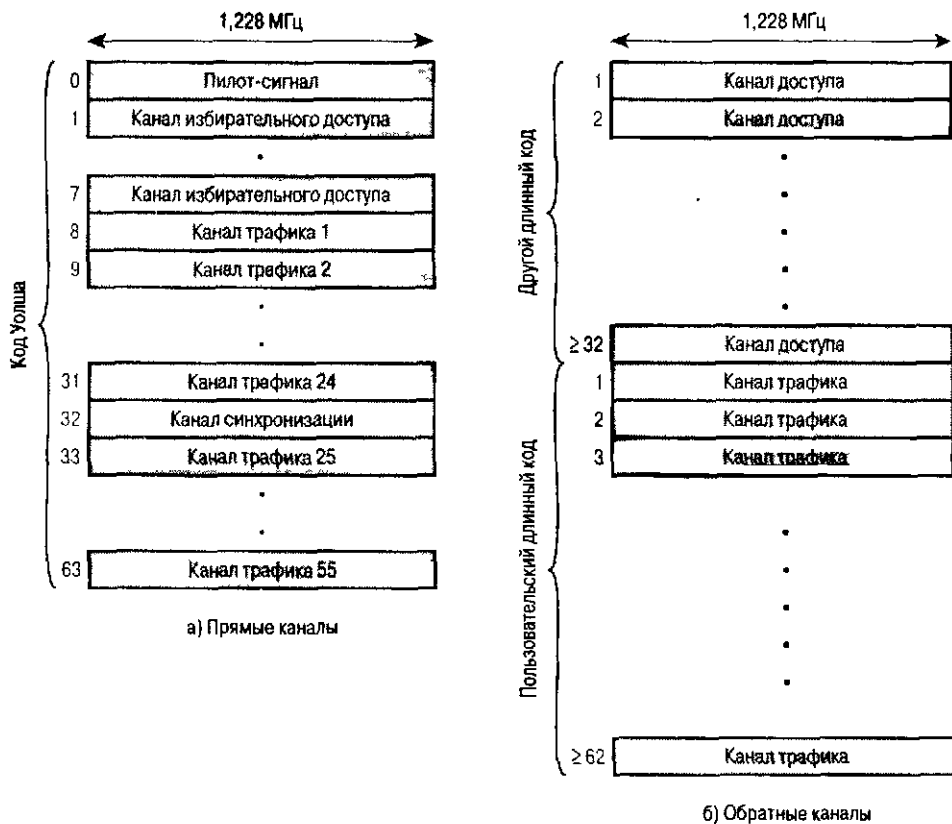


Рис. 10.19. Структура канала системы IS-95

На рис. 10.20 показана схема обработки данных, передаваемых по прямому информационному каналу со скоростью, соответствующей первому набору. Для голосовых передач речь кодируется со скоростью 8550 бит/с. После добавления дополнительных битов для обнаружения ошибок скорость повышается до 9600 бит/с. Когда пользователь молчит, вся пропускная способность канала не используется. В такие периоды тишины скорость передачи данных снижается до 1200 бит/с. Скорость 2400 бит/с используется для передачи переходов в фоновом шуме, а скорость 4800 бит/с — для смешивания оцифрованной речи и сигнальных данных.

Данные или оцифрованная речь передаются блоками по 20 мс с использованием схемы прямого исправления ошибок (сверточный код со степенью кодирования  $1/2$ ). Таким образом, эффективная скорость передачи данных удваивается и достигает максимум 19,2 Кбит/с. Если скорости передачи данных ниже, то биты, выходящие из кодера (называемые кодовыми символами) копируются, что позволяет достичь скорости 19,2 Кбит/с. Затем данные чередуются в блоках для уменьшения влияния длительных помех.

После чередования биты данных скремблируются. Этот шаг необходим для обеспечения конфиденциальности данных, а также для предотвращения отправки одинаковых фрагментов данных, что, в свою очередь, уменьшает вероятность одновременного отправления данных пользователями в момент наибольшей загрузки

системы. Скремблирование осуществляется посредством длинного кода — псевдослучайного числа из 42-битового регистра сдвига. Регистр сдвига устанавливается в исходное положение с помощью электронного регистрационного номера пользователя. Выходные данные генератора длинного кода передаются со скоростью 1,2288 Мбит/с, что в 64 раза выше скорости 19,2 Кбит/с, так что выбирается только один бит из 64 (операция децимации). Получающийся поток сравнивается (операция исключающего ИЛИ) с выходными данными устройства блочного чередования.

Информационные биты  
прямого канала трафика  
(172, 80, 40 или 16 бит/кадр)

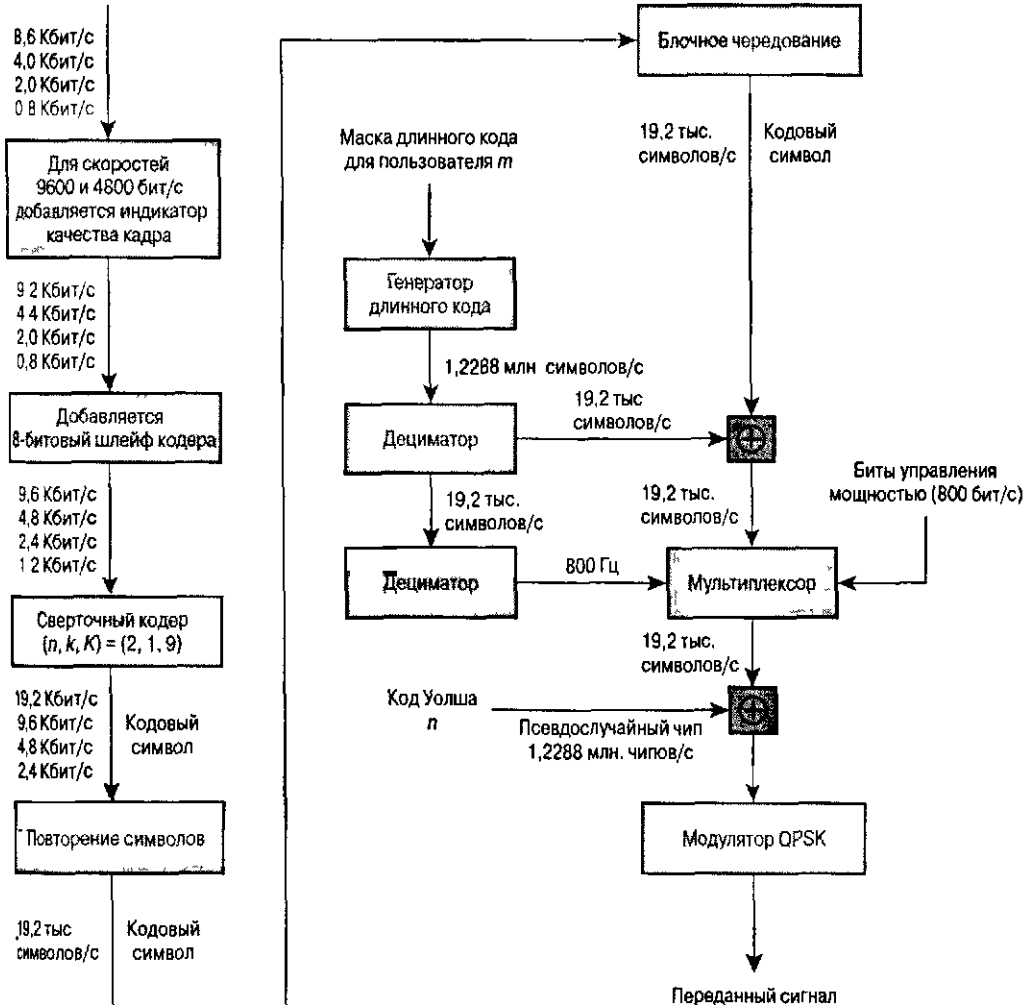


Рис. 10.20. Схема передачи данных по прямому каналу системы IS-95

На следующем этапе обработки вводится информация о регулировании мощности в информационном канале. Этот поток, имеющий скорость 800 бит/с, вводится в данные, передаваемые со скоростью 19,2 Кбит/с посредством замены

некоторых битов кода и кодирования контрольных битов с помощью генератора длинного кода. В таких “захваченных” битах передается информация, указывающая мобильному устройству, что следует увеличить, уменьшить или оставить неизменным текущий уровень мощности.

Следующий шаг — применение схемы DSSS; используется одна строка матрицы Уолша ( $64 \times 64$ ), скорость увеличивается от 19,2 Кбит/с до 1,2288 Мбит/с. Одна строка матрицы выделяется мобильной станции еще во время установления соединения. При наличии в передаваемых данных нуля посылается 64 бит выделенной строки; в противном случае посылается результат применения к строке побитового исключающего ИЛИ. Таким образом, окончательная скорость передачи данных получается равной 1,2288 Мбит/с. Этот цифровой поток битов затем модулируется на несущей с использованием схемы QPSK. Напомним из главы 6, что QPSK предполагает создание двух потоков битов, которые модулируются отдельно (см. рис. 6.6.). В схеме IS-95 данные разделяются на синфазный (I) и квадратурный (Q) каналы, и к данным и уникальному короткому коду применяется операция исключающего ИЛИ. Короткие коды генерируются 15-битовым регистром сдвига как псевдослучайные числа.

## Обратный канал системы IS-95

В табл. 10.7 представлены параметры обратного канала связи. Обратный канал связи может включать в себя до 94 логических каналов CDMA, каждый из которых занимает одну и ту же полосу частот шириной 1228 кГц (см. рис. 10.19, б). Обратный канал связи поддерживает до 32 каналов доступа и до 62 информационных каналов.

Таблица 10.7. Параметры обратного канала связи системы IS-95

Канал	Доступ	Первый набор информационных каналов				Второй набор информационных каналов			
Скорость передачи данных (бит/с)	4800	1200	2400	4800	9600	1800	3600	7200	14400
Степень кодирования	1/3	1/3	1/3	1/3	1/3	1/2	1/2	1/2	1/2
Скорость передачи символов до повторения (символ/с)	14400	3600	7200	14400	28800	3600	7200	14400	28800
Количество повторений символов	2	8	4	2	1	8	4	2	1
Скорость передачи символов после повторения (символ/с)	28800	28800	28800	28800	28800	28800	28800	28800	28800
Коэффициент заполнения	1	1/8	1/4	1/2	1	1/8	1/4	1/2	1
Количество символов кода на символ модуляции	6	6	6	6	6	6	6	6	6

Канал	Первый набор информационных каналов				Второй набор информационных каналов				
Количество чипов на символ модуляции	256	256	256	256	256	256	256	256	256
Количество чипов на бит	256	128	128	128	128	256/3	256/3	256/3	256/3

При обратной связи каждый информационный канал однозначно связан с одним мобильным устройством. Каждая станция имеет уникальную маску длинного кода, основанную на ее электронном регистрационном номере. Маска длинного кода представляет собой 42-битовое число, так что всего может быть  $(2^{42} - 1)$  различных масок. Каналы доступа используются мобильным устройством для инициации звонка, ответа на сообщения нисходящего канала избирательного вызова и для обновления сведений о местоположении.

На рис. 10.21 показаны этапы обработки данных, передаваемых по обратному информационному каналу при скорости, соответствующей первому набору. Первые несколько шагов те же, что и для прямого канала. Для обратного канала сверточный кодер имеет степень кодирования  $1/3$ , т.е. эффективная скорость передачи данных утраивается до максимального значения 28,8 Кбит/с. Затем применяется блочное чередование данных.

Следующим этапом является расширение данных с помощью матрицы Уолша. Способ и цель использования матрицы отличаются от способа и цели в прямом канале. В обратном канале связи данные, получающиеся в результате блочного чередования, группируются в модули по 6 бит. Каждый 6-битовый модуль играет роль индекса для выбора строки матрицы Уолша  $64 \times 64$ , и эта строка заменяется входными данными. Таким образом, скорость передачи данных увеличивается в  $64/6$  раз, т.е. до 307,2 Кбит/с. Целью такого кодирования является улучшение приема на базовой станции. Так как 64 возможных варианта кодирования являются ортогональными, блочное кодирование облегчает процесс принятия решения на приемнике, это удобно также с точки зрения вычислений (подробнее см. в [РЕТЕ95]). Модуляцию Уолша можно рассматривать как разновидность блочного кода коррекции ошибок  $(n, k) = (64, 6)$  с  $d_{\text{min}} = 32$ . Фактически здесь все расстояния равны 32.

Для уменьшения интерференции с другими мобильными станциями вводится генератор случайных чисел для пакета данных (см. обсуждение этого вопроса в [BLAC99]). Его действие основано на использовании маски длинного кода для сглаживания данных в каждом 20-миллисекундном кадре.

Следующий шаг — применение схемы DSSS. В обратном канале длинный код, являющийся уникальным для мобильного устройства, с помощью операции исключяющего ИЛИ сравнивается с выходными данными генератора случайных чисел, что позволяет получить окончательную скорость передачи данных, равную 1,2288 Мбит/с. Затем этот цифровой поток модулируется на несущей с использованием ортогональной схемы QPSK. Эта модуляция отличается от использованной в прямом канале, так как здесь в модуляторе используется задержка (см. рис. 6.6), гарантирующая ортогональность. Причина использования разных модуляторов заключается в том, что в прямом канале коды расширения являют-



ся ортогональными, поскольку все получены из матрицы Уолша, в то время как в обратном канале ортогональность кодов расширения не гарантирована.

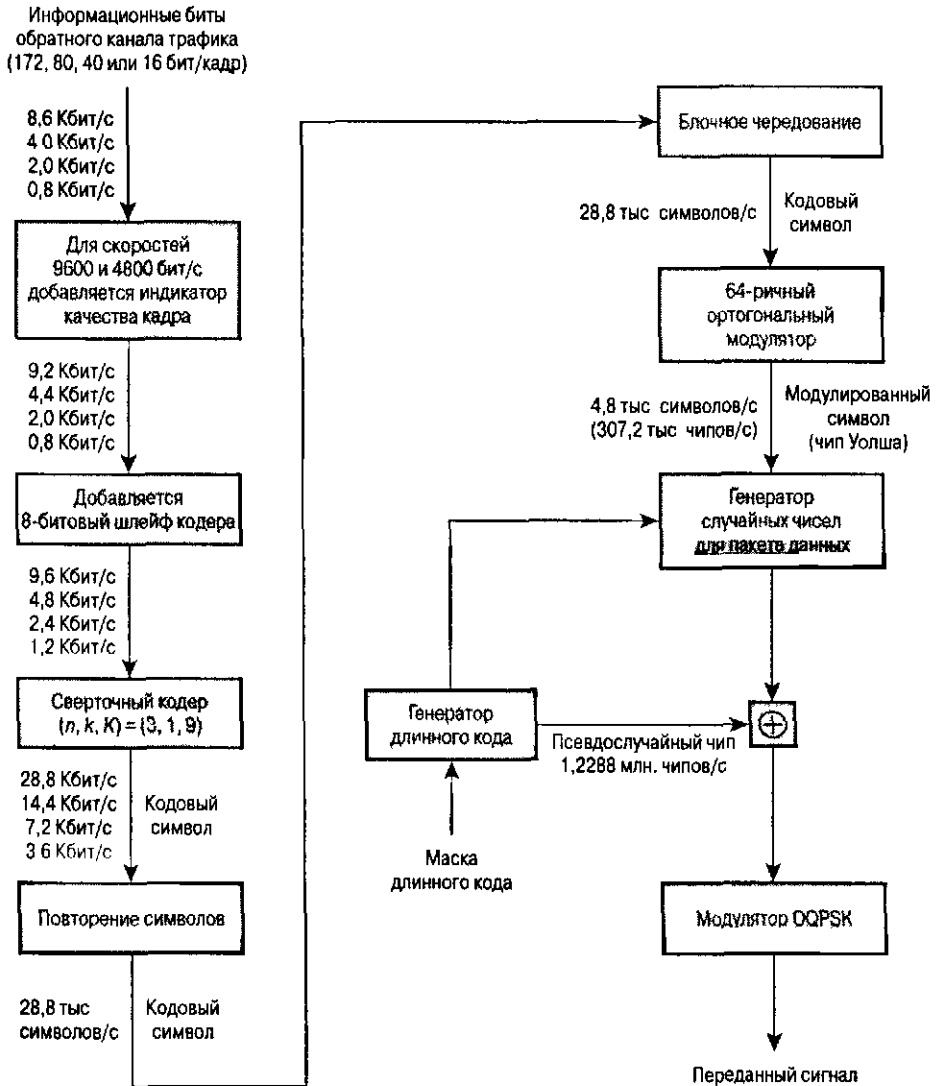


Рис. 10.21. Обратный канал связи системы IS-95

## 10.5. СИСТЕМЫ ТРЕТЬЕГО ПОКОЛЕНИЯ

Беспроводные системы связи третьего поколения разрабатываются с целью получения высокоскоростных беспроводных средств передачи не только речи, но и данных, мультимедиа и видео. По инициативе ITU IMT-2000 определено следующее видение союзом ITU возможностей систем третьего поколения.

- Качество речи сравнимо с качеством речи в общественной коммутируемой телефонной сети.
- Пользователям, имеющим высокоскоростные автомобили, в пределах довольно большой зоны доступна скорость передачи данных 144 Кбит/с.
- Пешеходам, стоящим или медленно идущим в пределах небольших участков, доступна скорость передачи данных 384 Кбит/с.
- Для учреждений будет поддерживаться скорость передачи данных 2,048 Мбит/с.
- Передачи данных могут быть симметричными и асимметричными.
- Поддерживается связь как с коммутацией пакетов, так и с коммутацией каналов.
- Имеется адаптивный интерфейс с Internet, который позволяет эффективно отразить асимметрию прибывающего и отправляемого трафика.
- Как правило, доступный спектр частот используется более эффективно.
- Поддерживается разнообразное мобильное оборудование.
- Система достаточно гибка для введения новых услуг и технологий.

Вообще одной из движущих сил современных технологий коммуникации является тенденция к универсализации персональных средств телекоммуникации и доступа к сообщениям. Первая концепция означает, что человек, используя один персональный счет, сможет легко идентифицировать себя и использовать с максимальным удобством любую систему коммуникации по всей стране, континенту или даже по всему миру. Второе понятие означает возможность использовать один терминал для соединения с информационными службами в разнообразных средах (например, можно было бы иметь портативный терминал, который будет одинаково хорошо работать в учреждении, на улице и на борту самолета). Осуществлению этих целей будет, несомненно, способствовать революция, происшедшая в области персональной вычислительной техники. Большим шагом вперед являются, например, модули идентификации абонента (SIM), использующиеся в сотовых телефонных системах GSM.

Эти концепции глобальной беспроводной связи были названы персональными службами связи (PCS) и персональными сетями связи (PCN), а их воплощение в жизнь является задачей беспроводных систем третьего поколения.

Вообще планируется, что в дальнейшем технология будет цифровой с возможностью множественного доступа с временным или кодовым разделением каналов, чтобы эффективно использовать спектр частот и обеспечивать высокую пропускную способность.

Телефоны PCS, согласно проекту, должны иметь меньшую мощность и быть относительно маленькими и легкими. Во всех странах сейчас ведутся работы, по завершении которых одни и те же терминалы можно будет использовать повсеместно.

## **Альтернативные интерфейсы**

На рис. 10.22 показаны альтернативные схемы, которые были разработаны как часть IMT-2000. В спецификации охвачен набор радиointерфейсов для оптимизации производительности в разных радиосредах. Для обеспечения постепенного

перехода от существующих систем первого и второго поколения к системам третьего поколения понадобились пять альтернативных схем.

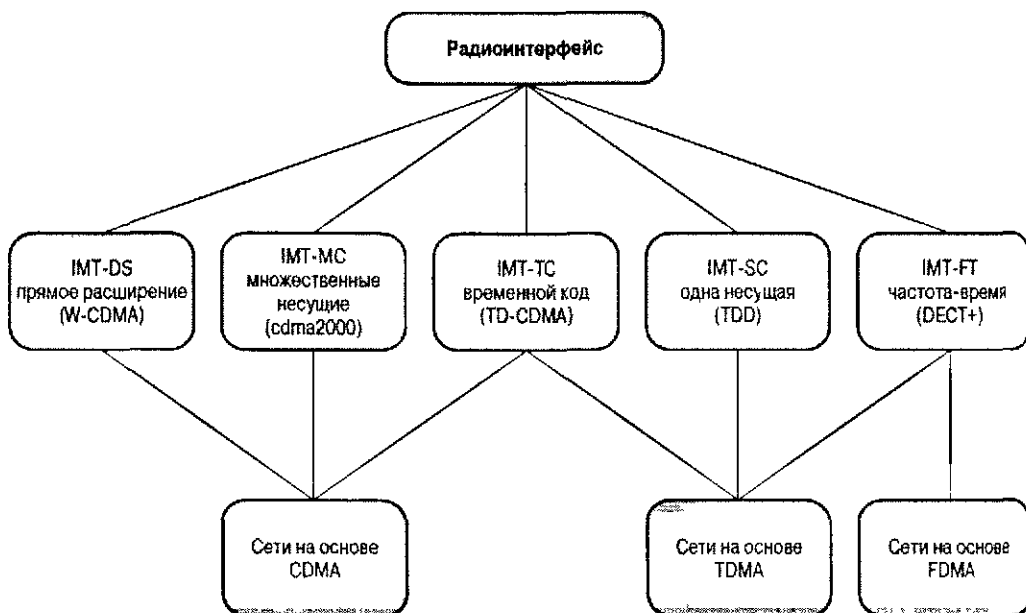


Рис. 10.22. Наземные радиоинтерфейсы IMT-2000

Эти пять альтернатив являются эволюцией систем второго поколения. Две спецификации были созданы на основе работ Европейского института стандартов по телекоммуникациям (European Telecommunications Standards Institute — ETSI) по разработке универсальной мобильной телекоммуникационной системы (universal mobile telecommunications system — UMTS) как стандарта для беспроводных систем третьего поколения в Европе. Системы UMTS включают в себя два стандарта. Один из этих стандартов известен как широкополосная схема CDMA, или W-CDMA. В этой схеме для обеспечения высокой скорости передачи данных при эффективном использовании полосы частот применяется технология CDMA. В табл. 10.8 приведены некоторые основные параметры схемы W-CDMA. Еще одна схема систем UMTS, предложенная в Европе, называется IMT-TC, или CDMA с временным разделением (TD-CDMA). Этот подход представляет собой сочетание технологий W-CDMA и TDMA. Технология IMT-TC должна обеспечить обновление систем GSM, основанных на технологии TDMA.

Таблица 10.8. Параметры широкополосной схемы CDMA

Ширина полосы канала	5 МГц
Структура прямого радиоканала	Прямое расширение
Скорость передачи чипов	$3,84 \cdot 10^6$ чипов/с
Длина кадра	10 мс
Количество слотов в кадре	15

Модуляция расширения	Сбалансированная QPSK (прямой канал), двухканальная QPSK (обратный канал), сложные схемы расширения
Модуляция данных	QPSK (прямой канал), BPSK (обратный канал)
Когерентное обнаружение	Пилот-символы
Уплотнение обратного канала	Временное уплотнение каналов управления и пилот-каналов. Синфазный и квадратурный каналы разделяются на информационные каналы и каналы управления
Передача на нескольких скоростях	Различные схемы расширения и множественные коды
Коэффициенты расширения	4–256
Расширение (в прямом канале)	Ортогональные последовательности переменной длины для разделения каналов. Последовательности Голда $2^{18}$ для разделения ячеек и пользователей.
Расширение (в обратном канале)	То же, что и в прямом, отличаются только временные сдвиги в синфазных и квадратурных каналах
Тип переключения	Мягкое переключение

Еще одна система, основанная на CDMA, известная как cdma2000, имеет североамериканское происхождение. Эта схема похожа на широкополосную схему CDMA, но несовместима с ней. Отчасти это потому, что в стандартах используются разные скорости передачи элементов сигнала (чипов). Кроме того, в системе cdma2000 используется технология, именуемая работой с несколькими несущими, которая не применяется в схеме W-CDMA.

В других разделах книги мы уже рассматривали некоторые технологии, которые появляются в системах третьего поколения. Это турбокоды (раздел 8.3), коды Голда и коды расширения переменной длины (раздел 7.5).

Еще две спецификации интерфейса показаны на рис. 10.22. Схема IMT-SC разработана преимущественно для сетей, использующих исключительно TDMA. Использование IMT-FC с несущими TDMA и CDMA позволяет предложить некоторые услуги третьего поколения. Эта схема является частью стандарта DECT (Digital European Cordless Telecommunications — цифровые европейские беспроводные телекоммуникации), который будет рассмотрен в главе 11.

## Вопросы проектирования систем CDMA

Доминирующей технологией для систем третьего поколения является CDMA. И хотя были разработаны три различные схемы CDMA, их проекты имеют некоторые общие особенности. Например, в [OJAN98] перечислено следующее.

- **Ширина полосы частот.** Важной целью при проектировании систем третьего поколения является ограничение использования каналов до 5 МГц. Существует несколько причин принятия такого решения. С одной стороны, полоса шириной 5 МГц и более повышает шансы приемника на восстановление многолучевого сигнала. С другой стороны, доступный спектр ограничен из соображений конкуренции, и 5 МГц являются разумной верхней границей, которая может быть выделена системам третьего поколения. И,

наконец, 5 МГц подходят для поддержки скоростей передачи данных 144 и 384 Кбит/с, что является одной из основных услуг третьего поколения.

- **Скорость передачи элементов сигнала.** При заданной ширине полосы скорость передачи элементов сигнала (чипов) зависит от желаемой скорости передачи данных, потребности в обработке ошибок и ограничений, накладываемых на ширину полосы частот. При заданных параметрах проекта разумной будет скорость передачи элементов сигнала  $3 * 10^6$  элементов/с и выше.
- **Передача на нескольких скоростях:** Термин *передача на нескольких скоростях* (multirate) означает предоставление определенному пользователю нескольких логических каналов с фиксированной скоростью передачи данных, причем разные скорости передачи данных предоставляются в разных логических каналах. Далее, логические каналы могут независимо обрабатываться коммутатором и направляться через беспроводные и стационарные сети различным адресатам. Преимуществом передачи на нескольких скоростях является то, что система становится достаточно гибкой, чтобы поддерживать несколько приложений данного пользователя, и может эффективно использовать пропускную способность, предоставляя каналы, требуемые только для данной услуги.

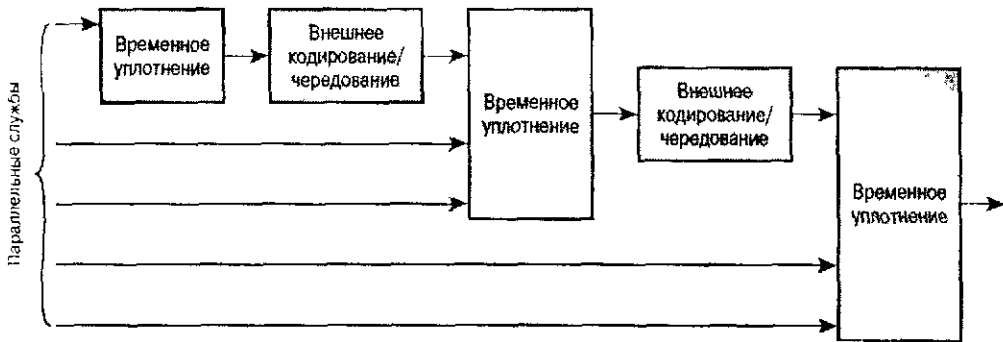
Передачи на нескольких скоростях можно добиться в одном канале CDMA с использованием схемы TDMA. Для достижения различных скоростей передачи данных в кадре выделяется различное количество слотов. Все подканалы при данной скорости передачи данных будут защищены схемами исправления ошибок и чередования (рис. 10.23, а). Альтернативой является использование нескольких кодов CDMA с отдельным кодированием и чередованием и наложение их на отдельные каналы CDMA (см. рис. 10.23, б).

## 10.6. РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА И WEB-САЙТЫ

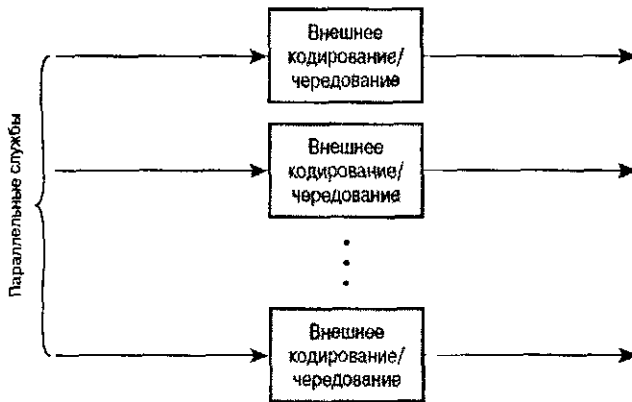
В [BERT94] и [ANDE95] содержится обзор эффектов распространения в сотовых беспроводных сетях. В [POLL96] подробно рассмотрены проблемы переключения. [EVER94] и [ORLI98] содержат расчеты аналитических параметров трафика сотовой сети. [BLAC99] является одним из лучших технических обзоров сотовых систем второго поколения. Понятия и идеи стандарта GSM изложены в [RAHN93], а более подробно стандарт GSM рассмотрен в [GARG99].

В [TANT98] собраны многие важные статьи, имеющие отношение к использованию технологии CDMA в сотовых сетях. В [DINA98] дан обзор как псевдослучайных, так и ортогональных кодов расширения для сотовых сетей CDMA.

В [OJAN98] речь идет о ключевых технических особенностях проектирования систем третьего поколения. Небезынтересен также обзор этой темы, содержащийся в [ZENG00], хотя более подробный анализ вопроса дан в [PRAS00].



а) Временное уплотнение



б) Кодовое уплотнение

Рис. 10.23. Принципы временного и кодового разделения каналов [OJAN98]

- ANDE95 Anderson J., Rappaport T., Yoshida S. Propagation Measurements and Models for Wireless Communications Channels. — *IEEE Communications Magazine*, January 1995.
- BERT94 Bertoni H., Honcharenko W., Maciel L., Xia H. UHF Propagation Prediction for Wireless Personal Communications. — In: *Proceedings of the IEEE*, September 1994.
- BLAC99 Black U. *Second-generation Mobile and Wireless Networks*. — Upper Saddle River, NJ: Prentice Hall, 1999.
- DINA98 Dinan E., Jabbari B. Spreading codes for Direct Sequence CDMA and Wideband CDMA Cellular Networks. — *IEEE Communications Magazine*, September 1998.
- EVER94 Everitt D. Traffic Engineering of the Radio Interface for Cellular Mobile Networks. — In: *Proceedings of the IEEE*, September 1994.
- GARG99 Garg V., Wilkes J. *Principles and Applications of GSM*. — Upper Saddle River, NJ: Prentice Hall, 1999.
- OJAN98 Ojanpera T., Prasad G. An overview of Air Interface Multiple Access for IMT-2000/UMTS. — *IEEE Communications Magazine*, September 1998.
- ORLI98 Orlik P., Rappaport S. Traffic Performance and Mobility Modeling of Cellular Communications with Mixed Platforms and Highly Variable Mobilities. — In: *Proceedings of the IEEE*, July 1998.

- POLL96 Pollini G. Trends in Handover Design. — *IEEE Communications Magazine*, March 1996.
- PRAS00 Prasad R., Mohr W., Konhauser W., eds. *Third-Generation Mobile Communications Systems*. — Artech House, 2000.
- RAHN93 Rahnema M. Overview of the GSM System and Protocol Architecture. — *IEEE Communications Magazine*, April 1993.
- TANT98 Tantaratana S., Ahmed K., eds. *Wireless Applications of Spread Spectrum Systems: Selected Readings*. — Piscataway, NJ: IEEE Press, 1998.
- ZENG00 Zeng M., Annamalai A., Bhargava V. Harmonization of Global Third-generation Mobile Systems. — *IEEE Communications Magazine*, December 2000.



### Рекомендуемые Web-сайты

- **Cellular Telecommunications and Internet Association** ([www.wow-com.com/](http://www.wow-com.com/)). Web-сайт промышленного консорциума, на котором представлена информация об успешных применениях беспроводной технологии.
- **GSM World** ([www.gsmworld.com/index1.html](http://www.gsmworld.com/index1.html)). Содержится много сведений и ссылок, касающихся стандарта GSM.
- **CDMA Development Group** ([www.cdg.org/](http://www.cdg.org/)). Сведения и ссылки для системы IS-95 и для систем CDMA вообще.

## 10.7. ТЕРМИНЫ, ВОПРОСЫ И ЗАДАЧИ

### Основные термины

базовая станция	множественный доступ с	регулирование мощности
блокируемая сеть	кодовым разделением	по разомкнутому
глобальная система	каналов (CDMA)	циклу
мобильной связи	мягкое переключение	сеть второго
(GSM)	неблокируемая сеть	поколения (2G)
жесткое переключение	обратный канал	сеть первого
кратность использования	передача	поколения (1G)
частоты	переключение	сеть третьего
многократное	прямой канал	поколения (3G)
использование частот	радиосвязь с	сотовая сеть
множественный доступ с	мобильными	усовершенствованная
временным	объектами	система мобильной
разделением каналов	регулирование мощности	телефонной связи
(TDMA)	регулирование мощности	(AMPS)
	по замкнутому циклу	

### Вопросы

1. Какая геометрическая форма ячеек используется при проектировании сотовых систем?
2. В чем заключается принцип многократного использования частот в сотовых сетях?

3. Перечислите пять способов повышения пропускной способности сотовой сети.
4. Поясните, в чем состоит функция избирательного вызова сотовых систем.
5. Перечислите и кратко опишите различные наборы параметров, которые можно использовать для принятия решения о переключении.
6. Если во время связи с базовой станцией мобильное устройство перемещается, какие факторы определяют необходимость и степень регулирования мощности?
7. Объясните разницу между регулированием мощности по замкнутому и разомкнутому циклу.
8. Чем отличается интенсивность трафика от средней частоты вызовов в системе?
9. Каковы ключевые отличия между сотовыми системами первого и второго поколений?
10. Каковы преимущества использования CDMA для сотовых сетей?
11. Каковы недостатки использования CDMA для сотовых сетей?
12. Объясните разницу между жестким и мягким переключением.
13. Приведите некоторые ключевые характеристики и отличительные особенности сотовых систем третьего поколения, сравните их с системами второго поколения.

## Задачи

1. Рассмотрим четыре различные сотовые системы со следующими характеристиками. Полосы частот лежат между 825 и 845 МГц для передачи сигналов с мобильного устройства и между 870 и 890 МГц — для передачи сигналов с базовой станции. Дуплексный контур содержит по одному каналу шириной 30 кГц в каждом направлении. Системы отличаются только кратностью использования частоты, которая равна 4, 7, 12 и 19, соответственно.
  - а. Пусть в каждой системе кластеры ячеек (4, 7, 12, 19) дублируются 16 раз. Найдите количество одновременных соединений, которые может поддерживать каждая система.
  - б. Найдите количество одновременных соединений, которые может поддерживать одна ячейка каждой системы.
  - в. Оцените площадь зоны, покрываемой каждой системой, в ячейках.
  - г. Пусть размеры ячеек всех четырех систем одинаковы и каждой системой покрывается фиксированная зона, площадь которой равна 100 ячейкам. Найдите количество одновременных соединений, которые может поддерживать каждая система.
2. Опишите последовательность событий подобную той, что изображена на рис. 10.6:
  - а) для звонка с мобильного устройства неподвижному абоненту;
  - б) для звонка неподвижного абонента на мобильное устройство.
3. При обсуждении процедуры переключения на основе отношения интенсивности сигнала к пороговому значению указывалось, что если задать слишком низкий порог, такой, как  $Th_3$ , то мобильный пользователь может успеть переместиться далеко вглубь новой ячейки ( $L_4$ ). Это снизит качество канала связи и может привести к перебоям в соединении. Можете ли вы указать еще один недостаток этой схемы?
4. В системах управления обычно используется схема гистерезиса. В качестве примера опишите механизм гистерезиса, используемый в термореле домашних приборов.
5. Длительность телефонного соединения некоего абонента равна 23 минуты. В течение часа от этого абонента других заказов на соединение не поступало. Каков трафик этого соединения в эрлангах?



6. С помощью табл. 10.3 найдите приближенные ответы на следующие вопросы. В каждом случае дайте словесное описание общей решаемой задачи. *Подсказка:* можно использовать интерполяцию прямой линией.
- Дано  $N = 20$ ,  $A = 10,5$ . Найдите  $P$ .
  - Дано  $N = 20$ ,  $P = 0,015$ . Найдите  $A$ .
  - Дано  $P = 0,005$ ,  $A = 6$ . Найдите  $N$ .
7. Ширина полосы частот, выделенной аналоговой сотовой системе, равна 33 МГц. Для обеспечения полнодуплексной передачи голоса и управляющих сигналов используются два симплексных (односторонних) канала.
- Какое количество каналов доступно в каждой ячейке, если кратность использования частоты равна (1) 4 ячейкам, (2) 7 ячейкам и (3) 12 ячейкам?
  - Пусть для каналов управления выделена полоса 1 МГц, но одной ячейке нужен только один такой канал. Определите, как разумно распределить каналы управления и голосовые каналы в каждой ячейке для тех трех значений кратности использования частоты, что были даны в пункте а.
8. Как уже упоминалось, односторонняя полоса частот, доступная одному оператору в системе AMPS, равна 12,5 МГц, причем ширина канала — 30 кГц, а количество каналов управления — 21. Нужно вычислить эффективность, с которой система использует свою полосу частот для конкретной системы. Даны следующие параметры:
- площадь ячейки — 8 км<sup>2</sup>;
  - общая площадь покрываемой зоны — 4000 км<sup>2</sup>;
  - кратность использования частоты — 7;
  - среднее количество звонков в секунду в час пик — 1,2;
  - среднее время разговора — 100 с;
  - вероятность блокирования звонка — 2%.
- Сколько голосовых каналов приходится на одну ячейку?
  - Используя данные табл. 10.3 и интерполяцию прямой линией, определите общий трафик на одну ячейку в эрлангах на ячейку. Затем преобразуйте это значение в эрланги на квадратный километр.
  - Вычислите количество звонков, приходящих в одну ячейку за один час, а также количество звонков, приходящих за один час на один квадратный километр.
  - Вычислите количество пользователей, обращающихся в одну ячейку за один час, а также количество пользователей, обслуженных за один час в одном квадратном километре.
  - Ниже приводится принятое определение спектральной эффективности модуляции, или эффективности модуляции, в эрлангах на 1 МГц на квадратный километр
- $$\eta = \frac{\text{(общий трафик, обслуживаемый системой)}}{\text{(ширина полосы частот)(общая площадь покрываемой зоны)}}$$
- Определите эффективность модуляции этой системы.

9. Сотовой системе, использующей технологию FDMA, в каждом направлении выделено по 12,5 МГц. Защитная полоса на краях выделенной части спектра имеет ширину 10 кГц, а ширина канала равна 30 кГц. Сколько каналов доступно?
10. Если одним радиоканалом поддерживаются 8 речевых каналов и если не предполагается никакой защитной полосы, то сколько пользователей смогут одновременно обслуживаться системой GSM?
11. а. Чему равно время передачи бита в стандарте GSM?  
 б. Если пользователю в каждом кадре выделено по одному слоту, то чему будет равно время задержки между последовательными передачами в последовательных кадрах?
12. Если рассматривать завершающие биты, захваченные биты, защитные биты и настроечные биты стандарта GSM как служебные сигналы, а остальные биты рассматривать как данные, то каков процент служебных битов в кадре GSM?
13. Используя определение медленной скачкообразной перестройки частоты, приведенное в главе 7, покажите, что в стандарте GSM используется медленная скачкообразная перестройка частоты.
14. Спектральная эффективность сотовой системы FDMA определяется по формуле

$$\eta_a = (B_c N_T) / B_w,$$

где

$B_c$  — ширина полосы частот канала;

$B_w$  — общая ширина полосы в одном направлении;

$N_T$  — общее количество голосовых каналов в покрываемой зоне.

- а. Какова верхняя граница  $\eta_a$ ?
- б. Определите  $\eta_a$  для системы из задачи 8.
15. Рассмотрим систему из 7 ячеек, покрывающую область площадью 2100 км<sup>2</sup>. Трафик в семи каналах распределен следующим образом:
- | Номер ячейки     | 1    | 2    | 3    | 4    | 5    | 6    | 7    |
|------------------|------|------|------|------|------|------|------|
| Трафик (эрланги) | 30,8 | 66,7 | 48,6 | 33,2 | 38,2 | 37,8 | 32,6 |
16. Каждый пользователь производит в среднем 0,03 эрланга трафика в час со средним временем разговора, равным 120 с. Система состоит из 395 каналов и спроектирована для обеспечения уровня обслуживания 0,02.
- а. Определите количество абонентов в каждой ячейке.  
 б. Определите количество звонков, приходящихся в час на одного абонента.  
 в. Определите количество звонков, приходящихся в час на одну ячейку.  
 г. Определите, сколько каналов требуется каждой ячейке. *Подсказка:* нужно выполнить экстраполяцию на основе значений, приведенных в табл. 10.3.  
 д. Определите число абонентов.  
 е. Определите среднее количество абонентов на один канал.  
 ж. Определите плотность абонентов на квадратный километр.  
 з. Определите общий трафик (в эрлангах).  
 и. Определите плотность трафика в эрлангах на квадратный километр.  
 к. Каков радиус ячейки?

# ГЛАВА 11

## БЕСШНУРОВЫЕ СИСТЕМЫ И БЕСПРОВОДНЫЕ АБОНЕНТСКИЕ ЛИНИИ

**11.1. Беспроводные системы**

**11.2. Беспроводные абонентские линии связи**

**11.3. Стандарт IEEE 802.16**

**11.4. Рекомендуемая литература и Web-сайты**

**11.5. Термины, вопросы и задачи**

**Приложение 11А. Линейный фильтр с предсказанием**

**В** этой главе рассмотрены две технологии, позволяющие оборудовать беспроводными системами жилые дома и учреждения: бесшнуровые системы и беспроводные абонентские линии связи (wireless local loop — WLL).

## 11.1. БЕСШНУРОВЫЕ СИСТЕМЫ

Стандартизованные бесшнуровые системы основаны на технологии беспроводного телеграфа. Первоначально радиотелефоны разрабатывались для предоставления пользователям свободы передвижения в пределах их жилья или небольшого учреждения. Все началось с отделения телефонной трубки от остальной части телефона (именуемой базовой станцией, или базой) и установления простого беспроводного аналогового канала связи. По мере совершенствования технологии были разработаны и цифровые радиотелефоны. Разные устройства имели разные беспроводные интерфейсы. Впрочем, поскольку производители продавали и базу и телефонную трубку как одно устройство, стандарты вводить не требовалось.

Организации по стандартизации заинтересовались бесшнуровыми технологиями, когда встал вопрос о расширении диапазона их применения, причем в двух направлениях. Во-первых, в бесшнуровых системах одна базовая станция может обслуживать нескольких пользователей и взаимодействовать не только с несколькими телефонными трубками, но и с устройствами обработки речи и данных (например, факсом или принтером). Во-вторых, бесшнуровые системы можно использовать в различных средах.

- **По месту жительства.** Установленная в жилом доме базовая станция может обеспечивать передачу речи и данных, что предоставит пользователю не только местную связь, но и связь с общественной телефонной сетью.
- **Учреждения.** Одна базовая станция может поддерживать связь в пределах небольшого учреждения, обслуживая при этом несколько телефонных трубок и устройств обработки данных. В более крупном учреждении несколько базовых станций можно объединить в некое подобие сотовой конфигурации, подсоединив при этом все базовые станции к коммутатору PBX (private branch exchange — телефонная сеть частного использования). Такая конфигурация может обслуживать сотни и даже тысячи пользователей.
- **Телеточки.** Телеточкой называется размещение базовой станции в общественном месте, например в торговом пассаже или в аэропорту. Впрочем, этот вид услуг не пользуется успехом на рынке.

В стандартах бесшнуровых систем должны учитываться множество технических моментов. Перечислим некоторые из них ([WEBB00]).

1. Радиус удаления телефонной трубки от базовой станции является небольшим, до 200 м, поэтому используются устройства малой мощности. Как правило, выходная мощность бесшнуровой системы должна быть на два-три порядка ниже мощности сотовых систем.
2. Телефонная трубка и базовая станция должны быть недорогими. Поэтому в таких областях, как кодирование речи и выравнивание каналов, следует использовать простые технические решения.

3. Большой выбор частот будет доступен не всегда, поскольку пользователи являются владельцами как базовой станции, так и мобильной части телефона и могут устанавливать свои телефоны где угодно. Следовательно, где бы ни использовалась система, она должна уметь отыскивать канал с наименьшей интерференцией.

Для бесшнуровых систем было предложено много разных стандартов, самым известным из которых является разработанный в Европе стандарт DECT (Digital Enhanced Cordless Telecommunications — цифровые расширенные беспроводные телекоммуникации)<sup>1</sup>. Эквивалент этого стандарта, разработанный в США, известен как PWT (personal wireless communications — персональные беспроводные коммуникации). В табл. 11.1 приведены некоторые ключевые параметры стандартов DECT и PWT. В этих системах используется схема, известная как дуплекс с временным разделением (time division duplex — TDD). Поэтому сначала в общих чертах будет рассмотрена схема TDD, а затем мы перейдем к подробностям стандарта DECT.

Таблица 11.1. Параметры систем DECT и PWT

	DECT	PWT
Ширина полосы частот	20 МГц	20 МГц
Полоса частот	от 1,88 до 1,9 ГГц	от 1,91 до 1,92 ГГц
Схема доступа	TDD/TDMA/FDMA	TDD/TDMA/FDMA
Ширина несущей полосы частот	1,728 МГц	1,25 МГц
Количество несущих	10	8
Количество каналов на одну несущую	12	12
Общее количество каналов	120	120
Переключение	Возможно	Возможно
Скорость передачи данных	1,152 Мбит/с	1,152 Мбит/с
Скорость передачи речи	32 Кбит/с	32 Кбит/с
Схема кодирования речи	ADPCM	ADPCM
Схема модуляции	Гауссова FSK	$\pi/4$ DQPSK
Максимальная выходная мощность	250 мВт	90 мВт
Средняя выходная мощность	10 мВт	10 мВт
Максимальный радиус ячейки	30–100 м	30–100 м

## Дуплекс с временным разделением

В схеме TDD, известной также как уплотнение с временным сжатием (time-compression multiplexing — TCM), в любой момент времени передача данных ведется только в одном направлении, а направления передачи чередуются. Ниже последовательно рассматриваются схема TDD и TDD-TDMA.

<sup>1</sup> Ранее аббревиатура DECT расшифровывалась как Digital European Cordless Telecommunications — цифровые европейские беспроводные телекоммуникации.

## Простая схема TDD

Чтобы с помощью обычной схемы TDD получить необходимую абоненту скорость передачи данных поток битов сначала разделяется передатчиком на равные сегменты, которые сжимаются во времени до более высоких скоростей передачи данных и передаются в форме пакетов, а принимающая сторона затем восстанавливает исходную скорость передачи данных. Между передачей пакетов в противоположных направлениях имеется короткий перерыв, нужный для установления канала связи. Таким образом, реальная скорость передачи данных в канале должна быть более чем в два раза выше, чем скорость передачи данных, требуемая двумя конечными системами.

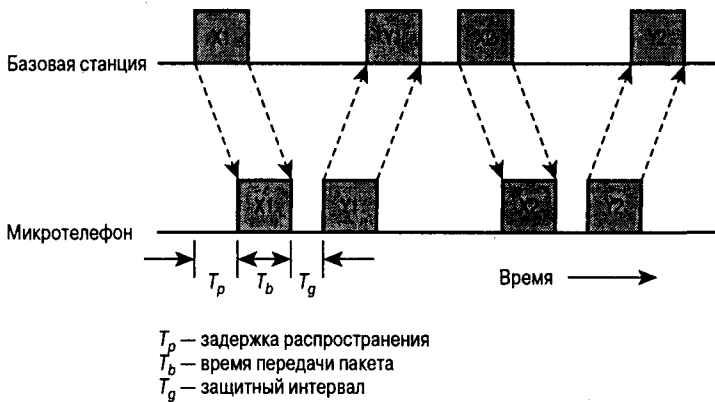


Рис. 11.1. Передача данных с помощью дуплекса с временным разделением

На рис. 11.1 показана хронологическая схема передачи. Право передачи данных периодически переходит от одной стороны к другой. При этом каждая сторона отправляет блоки некоторой фиксированной длины, передача которых занимает время  $T_b$  (линейная функция числа битов в блоке). Для распространения сигнала от передатчика к приемнику требуется время  $T_p$  (линейная функция расстояния между передатчиком и приемником). Наконец, для обращения канала введено защитное время  $T_g$ . Таким образом, время, требуемое для отправки одного блока, равно  $(T_p + T_b + T_g)$ . Кроме того, поскольку две стороны должны передавать данные по очереди, скорость передачи блоков любой стороной можно записать как  $1/2(T_p + T_b + T_g)$ . Эту величину можно связать с эффективной скоростью передачи данных,  $R$ , с которой ведут свои передачи обе стороны. Пусть размер блока равен  $B$  бит, а требуемое значение скорости передачи равно  $R$  бит/с. Тогда эффективное число битов, передаваемых в секунду, будет

$$R = B/2(T_p + T_b + T_g).$$

Реальную скорость передачи данных в среде,  $A$ , легко записать как

$$A = B/T_b.$$

Из двух последних формул получаем

$$A = 2R \left( 1 + \frac{T_p + T_g}{T_b} \right).$$

Таким образом, реальная скорость передачи данных по каналу связи более чем вдвое превышает эффективную скорость передачи, которую наблюдают общающиеся стороны. Выбор размера блока,  $B$ , диктуется противоречивыми требованиями к системе. По мере увеличения значения  $B$  снижается реальная скорость передачи данных,  $A$ . В этом случае упрощается реализация системы. С другой стороны, повышение размеров блока сопровождается увеличением времени задержки из-за необходимости буферизации, что нежелательно при передаче речи.

**Пример.** Среди стандартных интерфейсов, определенных для сетей ISDN (Integrated Services Digital Network — цифровая сеть с интеграцией услуг), есть базовый интерфейс, обеспечивающий скорость 192 Кбит/с при передаче 48-битовых кадров. Пусть используется технология TDD и размер блока равен длине кадра. Предположим, что расстояние между абонентом и коммутатором сети равно 1 км, а защитный интервал равен 10 мкс. Какова реальная скорость передачи данных?

Время передачи пакета —  $(48 \text{ бит}) / (192 \text{ Кбит/с}) = 250 \text{ мкс}$ . Задержка распространения —  $(1 \text{ км}) / (3 \times 10^8 \text{ м/с}) = 3,33 \text{ мкс}$ . Таким образом, получаем

$$A = 2 \times 192 \times [1 + (3,33 + 10) / 250] = 404 \text{ Кбит/с}.$$

## TDD-TDMA

Схема TDD первоначально предназначалась для проводных абонентских систем, однако в настоящее время она нашла применение и в системах с беспроводной конфигурацией. В беспроводной системе TDD позволяет передавать и принимать сигналы на одной и той же несущей частоте, но в разные моменты времени. Беспроводная схема TDD, как правило, используется в сочетании с технологией TDMA, в которой несколько абонентов могут по очереди принимать сигналы в прямом канале (от базовой станции к трубке) и затем передавать сигналы в обратном канале (от трубки к базовой станции) по очереди, используя при этом одну и ту же несущую частоту.

Использование дуплекса с временным разделением имеет два важных преимущества перед использованием системы TDMA с двумя разными несущими частотами, по одной на каждое направление. Первым преимуществом является возможность устранения быстрого замирания, а вторым — улучшенное распределение пропускной способности.

Что касается быстрого замирания, то его эффект меняется с изменением частоты<sup>2</sup>. Поэтому связь с мобильным пользователем, принимающим на одной частоте, а передающим на другой, может ухудшаться только в одном направлении. Одним из способов преодоления быстрого замирания является пространственное разнесение антенн. В такой схеме приемник использует две антенны, расстояние между которыми не меньше длины волны. При такой конфигурации довольно велика вероятность того, что при сильном замирании сигнала от одной из антенн сигнал от другой антенны будет распространяться нормально. В этом случае приемник сможет выбрать для

<sup>2</sup> Быстрое замирание рассмотрено в разделе 5.4.

демодуляции более сильный из поступающих сигналов. Такой метод подходит для крупной базовой станции, однако непрактичен для маленьких, недорогих. Более того, желательно избегать использования сложных алгоритмов выравнивания или прямой защиты от ошибок. Впрочем, в схеме TDD все же можно использовать пространственное разнесение антенн. Для приема каждого пакета, прибывающего на базовую станцию по прямому каналу связи, базовая станция выбирает ту антенну, сигнал которой сильнее. Затем базовая станция использует ту же антенну для последующей передачи в соответствующем прямом канале. Ввиду того что для обоих направлений используется одна и та же частота, эта антенна должна давать самый сильный сигнал и в обратном направлении, к трубке.

Второе преимущество схемы TDD-TDMA перед схемой FDMA-TDMA проявляется в механизме **распределения пропускной способности**. Для многих приложений, ориентированных на работу с данными (в отличие от приложений, предназначенных для работы с речью), наибольший объем нагрузки приходится, как правило, на прямое направление, а не на обратное. При использовании обычной схемы FDMA с предоставлением прямому и обратному каналам полос частот одинаковой ширины, системы проектируются с учетом нагрузки в прямом канале, так что большая часть полосы обратного канала обычно не используется. Если же использовать схему TDD, то контроллер системы динамически распределяет слоты TDMA в каждом направлении, так что при необходимости прямому каналу может предоставляться большее число слотов.

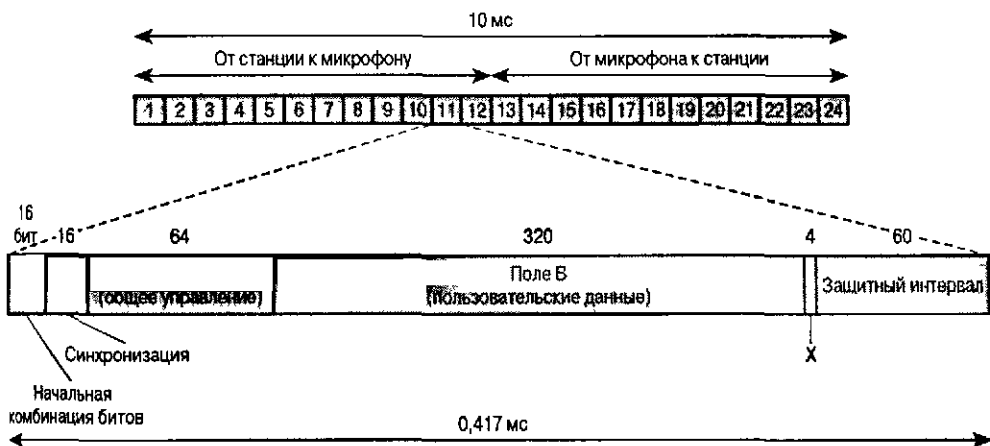
## Формат кадра DECT

При объяснении стандарта DECT лучше всего начинать с формата кадра TDD-TDMA, схема которого показана на рис. 11.2. Один кадр занимает 10 мс и состоит из 24 временных интервалов (слотов) TDMA. По умолчанию первые 12 слотов определяют 12 логических каналов, по которым осуществляется передача данных от базовой станции к трубке (прямое направление), а оставшиеся 12 определяют 12 логических каналов для передачи от трубки к базовой станции (обратное направление). Таким образом, на одной несущей поддерживаются 12 полнодуплексных логических каналов. Направления передачи чередуются, сначала идут 12 слотов в прямом направлении, за которыми следуют 12 слотов в обратном направлении.

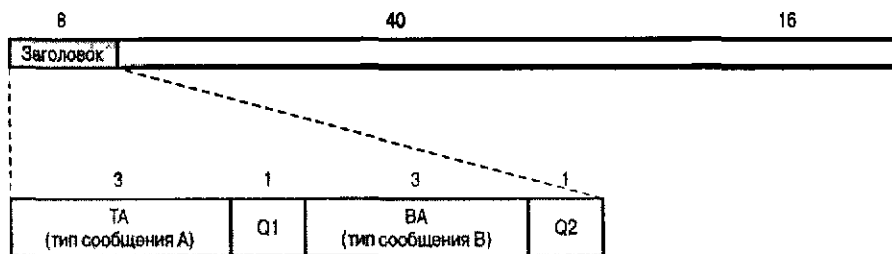
Каждый слот имеет длину  $10/24 = 0,417$  мс и всего вмещает 480 бит, распределенных по 6 полям. Таким образом, скорость передачи данных составляет  $480/0,417 = 1,152$  Мбит/с. Перечислим названные шесть полей.

- **Предварительная последовательность (16 бит)**. Предупреждает приемник о поступлении информации.
- **Синхронизация (16 бит)**. Позволяет синхронизировать прием, предоставляя в качестве точки отсчета начало временного интервала.
- **Поле А (64 бит)**. Используется для управления сетью.
- **Поле В (320 бит)**. Содержит пользовательские данные.
- **Поле Х (4 бит)**. Состоит из четырех битов контроля четности, которые позволяют терминалам и базовой станции проверять качество передачи сигнала.
- **Защитное поле (60 бит)**. Представляет собой защитный интервал длиной 52 мкс, соответствующий интервалу  $T_s$  на рис. 11.1.





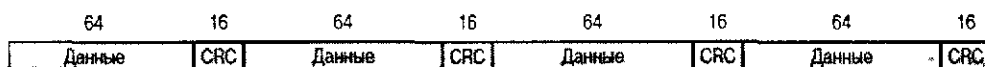
а) Формат кадра и временного слота



б) Формат поля А



Незащищенный режим (речь = 32 Кбит/с)



Защищенный режим (данные = 25,6 Кбит/с)

в) Формат поля В

Рис. 11.2. Форматы DECT

## Поле А

Поле А, которое служит для передачи управляющей информации, состоит из трех подполей.

- **Заголовок (8 бит).** Указывается тип информации, передаваемой в полях А и В.
- **Данные (40 бит).** Здесь находится сообщение управления сетью.
- **CRC (16 бит).** 16-битовый код обнаружения ошибок в заголовке и данных (код CRC рассмотрен в разделе 8.1).

В поле А уплотняются пять логических каналов управления, обозначаемых С, М, N, Р и Q (см. табл. 11.2). Первые три бита заголовка определяют тип информации, содержащейся в 40-битовом поле данных. Биты Q1 и Q2 образуют сигнальный канал, предоставляющий информацию о качестве сигнала на устройстве с другой стороны.

Таблица 11.2. Логические каналы DECT [GOOD97]

Канал	Скорость передачи данных, Кбит/с		Назначение
	Передача от базы	Передача от терминала	
I <sub>N</sub>	2	2	Незащищенная пользовательская информация (телефон)
I <sub>P</sub>	Переменная	Переменная	Защищенная пользовательская информация (данные)
С	0-2	0-2	Управление соединением
М	0-2	0-2	Управление физическим уровнем
N	0,25-3,75	2-4	Квитирование
Р	0-1,5	0	Избирательное обращение
Q	0,25	0	Системная информация

Канал Q используется для широковещательной передачи общей системной информации от базовой станции всем терминалам. Информация канала Q содержится в одном кадре 16-кадрового мультикадра. Поэтому скорость передачи битов в канале Q составляет 40 бит за 160 мс, или 250 бит/с. Канал Р обеспечивает передачу сообщений от базовой станции к некоторым терминалам. Так же, как и в сотовой системе, избирательное сообщение служит для извещения определенного пользователя о входящем звонке. Канал Р может занимать до 6 кадров мультикадра, что дает максимальную скорость передачи 1,5 Кбит/с. В ответ на избирательное сообщение или во время переключения терминалы используют двусторонний канал М, служащий для обмена с базовой станцией сообщениями, касающимися управления доступом к среде. Эти сообщения позволяют координировать распределение нагрузки и переключение в канале. Для сообщений канала М могут использоваться до 8 кадров мультикадра в каждом направлении, или половина пропускной способности поля А. Если соединение уже установлено, канал N предоставляет протокол квитирования. Базовая станция передает идентификационный код в прямом канале, а терминалы отражают этот код в соответствующих обратных N-каналах. Это позволяет обеим сторонам обнаруживать сильные сигналы, прибывающие из интерферирующих источников по одному и тому же физическому каналу. Канал N использует от 1 до 15 из 16 кадров в прямом мультикадре, в зависимости от потребностей каналов М, С и Р. В обратном направлении канал N занимает от 8 до 16 кадров каждого мультикадра, в зависимости от потребностей каналов М и С. Наконец, канал С предоставляет возможность управления активным соединением и может занимать до 8 кадров мультикадра в каждом направлении.

## Поле В

Данные в поле В передаются в одном из двух режимов. Незащищенный режим используется для передачи оцифрованной речи. Скорость передачи данных в каждом канале каждого направления равна примерно 320 бит за 10 мс (32 Кбит/с). Следовательно, можно использовать только такую схему кодирования речи, которая соответствует скорости 32 Кбит/с (см. последующее обсуждение).

В защищенном режиме передаются неголосовые данные. В этом случае поле В разбито на четыре блока по 64 бит, причем каждый блок защищен 16-битовым полем CRC. Отметим, что коррекция ошибок не производится, выполняется только их обнаружение. Восстановление данных, в которых обнаружены ошибки, происходит на более высоком уровне с использованием протокола ARQ (см. раздел 8.4). В защищенном режиме достигается скорость передачи данных до 25,6 Кбит/с для канала, который занимает один слот кадра. Схема DECT позволяет динамически распределять слоты между логическими каналами в обоих направлениях, так что указанная величина — это базовая скорость передачи данных.

## Работа схемы DECT

На рис. 11.3 показана архитектура протокола, который поддерживает операции DECT. На физическом уровне данные передаются в кадрах TDD-TDMA (см. рис. 11.2) на одной из 10 несущих радиочастот. Схема модуляции — гауссова FSK с номинальным отклонением 288 кГц. По сути, это та же схема GMSK, что используется в системе GSM (глава 10). Скорость передачи данных и ширина полосы частот относятся друг к другу как 2:1 (сравните с табл. 6.2), скорость передачи данных — 1,152 Мбит/с.

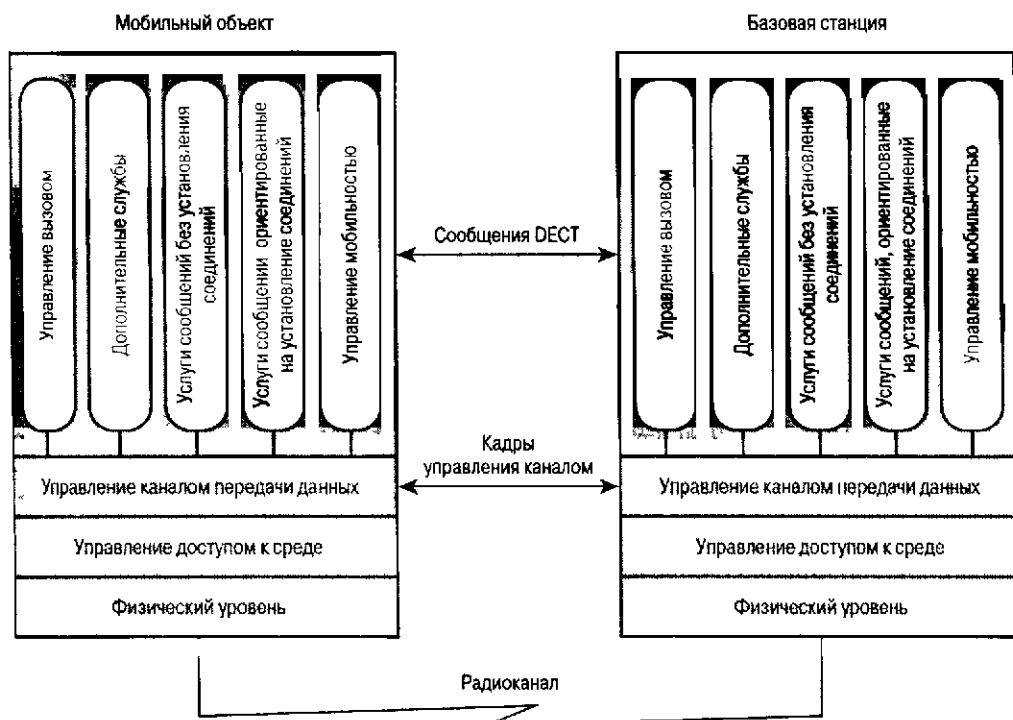


Рис. 11.3. Архитектура протоколов DECT

На уровне управления доступом к среде (MAC) осуществляется выбор физических каналов и устанавливается или освобождается соединение в этих каналах. Кроме того, на этом уровне информация уплотняется в формат кадров TDD-TDMA. На уровне MAC предоставляются три типа услуг.

- **Широковещание.** Передача широковещательных сообщений в поле А.
- **Услуги с установлением соединения.** Передача пользовательских данных в поле В.
- **Услуги без установления соединения.** Поддержка отдельных сообщений DECT, отправляемых в поле А.

Уровень управления каналом передачи данных обеспечивает надежную передачу сообщений с помощью традиционных процедур управления каналом передачи данных, включая обнаружение ошибок и автоматический запрос повторной передачи.<sup>3</sup>

Выше уровня управления каналом передачи данных предлагается такой набор услуг.

- **Управление вызовами.** Управление звонками с коммутацией каналов, включая установление и освобождение соединения.
- **Дополнительные услуги.** Услуги, для предоставления которых не нужны звонки.
- **Службы сообщений без установления соединения.** Поддержка передачи сообщений без установления соединения. Длинные сообщения на время передачи разбиваются на блоки.
- **Службы сообщений с установлением соединения.** Поддержка передачи сообщений с установлением соединения.
- **Управление мобильностью.** Функции, необходимые для конфиденциального предоставления услуг DECT. Управление мобильностью разделено на семь групп услуг.
  - **Процедура идентификации.** Используется для распознавания базовой станцией мобильного устройства.
  - **Процедура аутентификации.** Определяется, является ли мобильное устройство легальным пользователем.
  - **Процедура определения местонахождения.** Используется в системах, в которых местоположение мобильного устройства отслеживается несколькими базовыми станциями.
  - **Процедура предоставления прав доступа.** Мобильному устройству предоставляется право доступа к локальной или глобальной сети определенного типа.
  - **Процедура назначения ключа.** Распределяются ключи шифрования для защиты сетевой управляющей информации, а также пользовательской информации.
  - **Процедура запроса параметров.** Используется для обмена информацией о мобильном устройстве и функционировании сети.
  - **Процедура шифрования данных.** Операции шифрования и дешифрования.

---

<sup>3</sup> Протоколы управления каналом передачи данных подробно рассмотрены в приложении Г.

## Адаптивная дифференциальная импульсно-кодовая модуляция

Для оцифровки речи в схеме DECT используется процедура ADPCM (adaptive differential pulse code modulation — адаптивная дифференциальная импульсно-кодовая модуляция). Эта процедура была стандартизована ITU-T и используется во многих приложениях, включая цифровые сети с коммутацией каналов, такие, как ISDN, и при реализации многих беспроводных абонентских линий связи. Далее будут рассмотрены принципы, лежащие в основе модуляции, будет описана более простая дифференциальная схема PCM (DPCM) и, наконец, рассмотрена собственно схема ADPCM.<sup>4</sup>

### Дифференциальное квантование

Дифференциальное квантование основано на принципе, согласно которому речевые сигналы незначительно изменяются от выборки к выборке. При обычной импульсно-кодовой модуляции квантованные значения смежных выборок будут, в общем случае, близки друг к другу, поэтому передаваемые модулированные значения будут содержать много избыточной информации. Следовательно, имеет смысл передавать только значение разности двух смежных выборок, а не их абсолютные значения, тогда для передачи одной выборки потребуется меньшее число битов (см. рис. 11.4). Если значение  $k$ -й выборки равно  $m(k)$ , то передается лишь разность  $d(k) = m(k) - m(k - 1)$ . Тогда если приемник будет располагать точным начальным значением, то по последовательности разностей  $d(k)$  будет восстановлена последовательность абсолютных значений выборок  $m(k)$ .

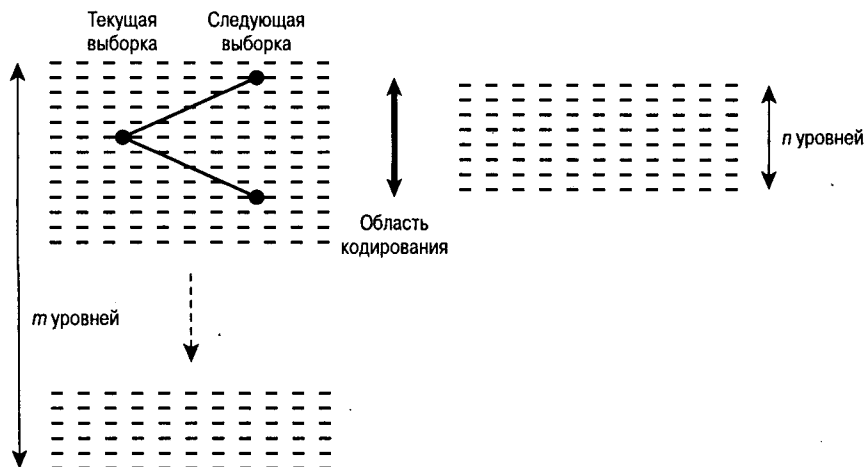


Рис. 11.4. Обычное и дифференциальное квантование [ВУСН99]

В то же время, если передавать только значения разностей между текущими и непосредственно предшествующими им выборками (для чего потребуется меньшее число битов), существует опасность постепенного нарастания от-

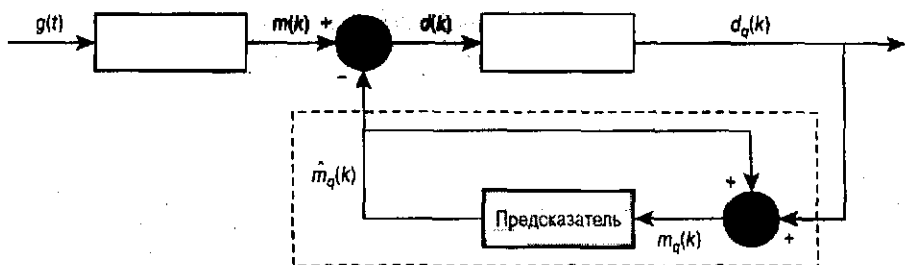
<sup>4</sup> Возможно, перед чтением данного раздела читателю стоит просмотреть обзор модуляции PCM в разделе 6.4.

клонения выходных данных приемника от истинных значений. Если значение разности двух выборок превысит значение, которое можно представить с помощью передаваемых битов, приемник не сможет правильно воспроизвести входные данные и не сможет впоследствии исправить ошибку. Для решения этой проблемы передатчик должен не только передавать значения разностей, но также дублировать декодирующую функцию, которая будет использоваться приемником. Тогда кодер будет выдавать разность между текущей выборкой и той выборкой, которая, по сведениям кодера, получена на приемнике при предыдущей передаче. В результате кодер будет инструктировать декодер, как вносить изменения в уже полученные выходные данные и таким образом автоматически производить коррекцию.

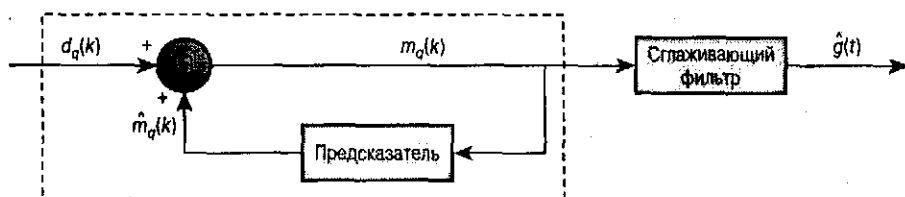
### Дифференциальная РСМ

Описанная выше схема основана на предположении о том, что значения разностей будут меньшими, чем абсолютные значения выборок. Еще большую выгоду можно извлечь, если предположить, что голосовые сигналы изменяются относительно медленно, поэтому можно довольно точно оценить или предсказать значение  $k$ -й выборки  $m(k)$  на основе значений предыдущих выборок. Для оценки  $k$ -й выборки,  $\hat{m}(k)$ , нужно будет передать разность  $d(k) = m(k) - \hat{m}(k)$ . Если использовать достаточно точную оценочную функцию, то эта разность будет меньше разности значений двух последовательных выборок. На приемнике, где используется та же оценочная функция, входящее значение разности будет добавлено к оценке предыдущей выборки и таким образом будет оценена текущая выборка. Этот принцип положен в основу дифференциальной импульсно-кодовой модуляции (DPCM). Отметим, что простая разностная схема, описанная в предыдущем параграфе, является всего лишь частным случаем DPCM, в котором для оценки выбрана функция  $\hat{m}(k) = m(k - 1)$ .

На рис. 11.15, *a* показана блок-схема передатчика. Чтобы понять, что представляет собой эта диаграмма, рассмотрим подробно все ее составляющие. На вход передатчика подается голосовой сигнал  $g(t)$ . Сначала из сигнала извлекается аналоговая выборка  $m(k)$ , представляющая собой выборку амплитудно-импульсной модуляции (РАМ). Уже на этом этапе начинаются трудности. Казалось бы, ничто не мешает построить оценочное значение  $\hat{m}(k)$ , найти разность и передать значение этой разности. Однако принимающая сторона вместо предыдущих аналоговых выборок  $m(k - 1)$ ,  $m(k - 2)$  и т.д., которые находились на передатчике, будет располагать последовательностью квантованных разностей, на основе которых можно воссоздать только квантованную выборку  $m_q(k)$ . Поэтому на приемнике воссоздать  $\hat{m}(k)$  не удастся, а можно будет, основываясь на значениях предыдущих квантованных выборок  $m_q(k - 1)$ ,  $m_q(k - 2)$  и т.д., определить только  $\hat{m}_q(k)$ , оценку квантованной выборки  $m_q(k)$ . Если передатчик будет основывать предсказания на значениях  $m(k)$ , а приемник — на значениях  $m_q(k)$ , то выборка на приемнике будет восстановлена неправильно. Передатчик должен, как и приемник, определять значения  $m_q(k)$  и передавать разность  $d(k) = m(k) - \hat{m}_q(k)$ . Тогда приемник на основе полученных разностей  $d(k)$  сможет правильно восстановить выборку  $m_q(k)$ .



а) Кодер DPCM



б) Декодер DPCM

Рис. 11.5. Передатчик и приемник DPCM

Осталось только показать, что схема, изображенная на рис. 11.5, а, воспроизводит нужные значения квантованных разностей. На выходе предсказателя определяется значение  $\hat{m}_q(k)$ , которое является предсказанным значением выборки  $m(k)$ . Затем берется разность

$$d(k) = m(k) - \hat{m}_q(k).$$

Далее, полученная разность квантуется, чтобы получить значение квантованной разности  $d_q(k)$ , которое можно представить следующим образом:

$$d_q(k) = d(k) - e(k),$$

где  $e(k)$  представляет собой ошибку квантования, получающуюся вследствие аппроксимации аналоговой величины  $d(k)$  цифровой величиной  $d_q(k)$ . Таким образом, входные данные предсказателя состоят из квантованных разностей и выходных данных предыдущего предсказания:

$$\begin{aligned} m_q(k) &= \hat{m}_q(k) + d_q(k) = \\ &= [m(k) - d(k)] + [d(k) + e(k)] = \\ &= m(k) + e(k). \end{aligned}$$

Видно, что  $m_q(k)$  является квантованным аналогом  $m(k)$ . Таким образом, как и было задумано, на вход предсказателя подаются значения  $m_q(k)$ , необходимые для работы приемника. Логика приемника показана на рис. 11.5, б. Часть приемника, которая заключена в пунктирную рамку, ничем не отличается от заключенного в пунктирную рамку участка передатчика. На вход этих участков подается одно и то же значение,  $d_q(k)$ , следовательно, на выходе получаются оди-

наковые значения  $m_q(k)$ . Затем на приемнике выходные данные этого участка пропускаются через фильтр, вследствие чего получается аналоговый сигнал, являющийся приближением исходного аналогового сигнала.

Как правило, предсказатель — это реализованная с использованием регистров сдвига линейная взвешенная сумма предыдущих выборок и задержек, равных интервалу выборки (рис. 11.6).

$$\hat{m}_q(k) = \sum_{i=1}^N A_i m_q(k-i). \quad (11.1)$$

Более подробно устройство линейного предсказателя обсуждается в приложении 11А.

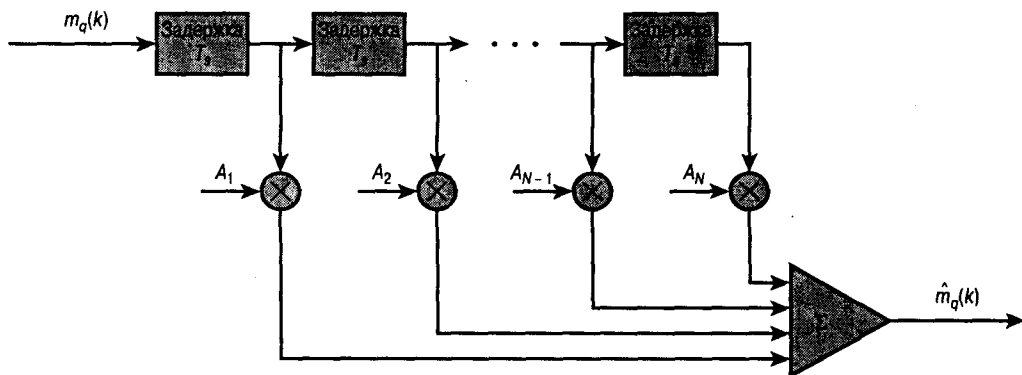


Рис. 11.6. Схема линейного предсказателя

### Адаптивная дифференциальная РСМ

Производительность схемы DPCM можно довести, если использовать адаптивные методы предсказания и квантования, которые позволят предсказателю и устройству квантования адаптироваться к изменяющимся характеристикам кодируемой речи. В этом разделе описывается схема ADPCM, стандартизованная ИТУ-Т в Рекомендации G.726<sup>5</sup>, которая используется в системах DECT и во многих абонентских линиях связи.

На рис. 11.7 показана базовая логика ADPCM. Передатчик преобразует поток оцифрованной речи со скоростью 64 Кбит/с в сжатый поток со скоростью 32 Кбит/с. Общая блок-схема аналогична схеме, приведенной для модуляции DPCM. В этом случае адаптивный предсказатель использует не только значения своих предыдущих выходных данных, но и значения предыдущих квантованных разностей. В предсказателе и в устройстве квантования параметры алгоритмов регулируются динамически, в зависимости от статистических свойств предыдущих выборок. Как и при использовании схемы DPCM, ядро декодера — копия части логики кодера.

<sup>5</sup> *General Aspects of Digital Transmission Systems; Terminal Equipments: 40, 32, 20, 16 kbit/s ADPCM. 1990.*



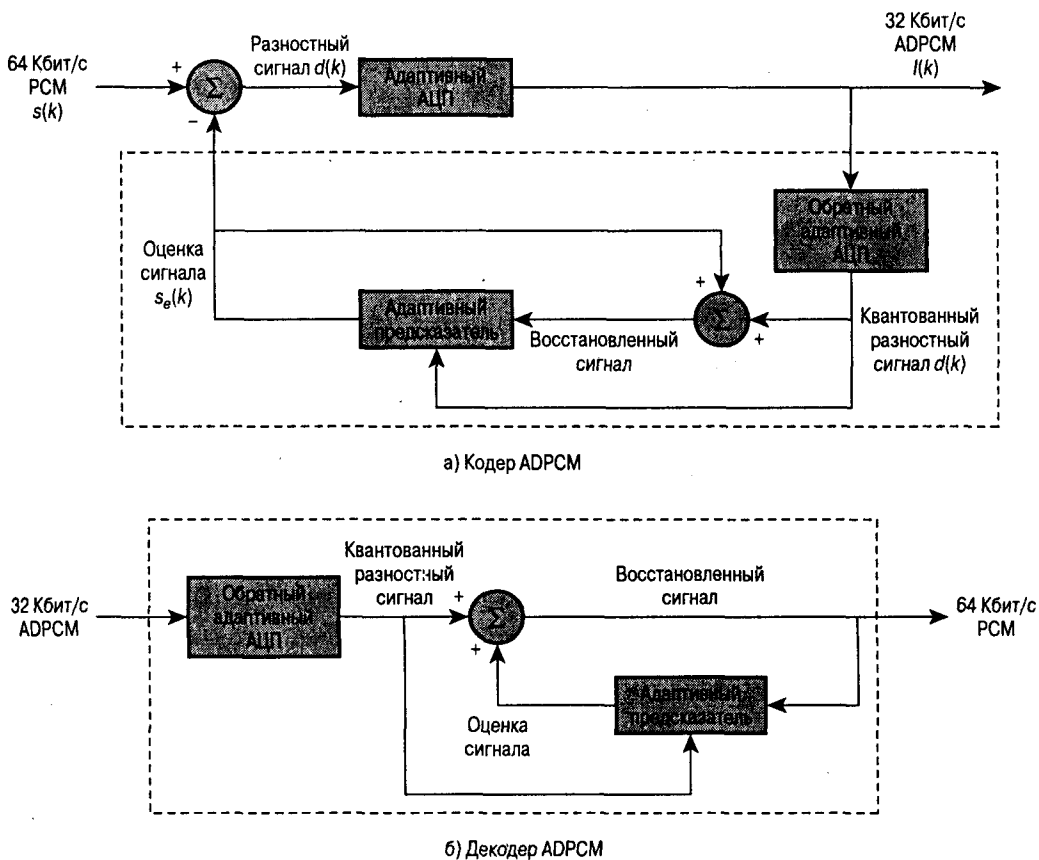


Рис. 11.7. Упрощенная схема передатчика и приемника ADPCM

Адаптивное устройство квантования (аналого-цифрового преобразователя — АЦП) в качестве входных данных принимает значение разности  $d(k)$  между модулированным сигналом  $s(k)$  и оценкой этого сигнала  $s_e(k)$ , а на выходе выдает 4-битовый модулированный сигнал ADPCM  $I(k)$ . Соответствующее отображение значений приведено в первых двух столбцах табл. 11.3. До квантования  $d(k)$  представляется в виде логарифма по основанию 2, масштабируется на величину  $y(k)$  — множитель адаптации, основанный на скорости изменения сигнала от выборки к выборке. Полученный результат представляется как 4 бит, которые и передаются как сигнал ADPCM.

Таблица 11.3. Нормированные характеристики входа/выхода устройства квантования ADPCM

Диапазон нормированных входных значений АЦП, $\log_2 d(k)  - y(k)$	Выход адаптивного АЦП, $I(k)$	Выход обратного адаптивного АЦП, $d_q(k)$
$[3,12, +\infty)$	7	3,32
$[2,72, 3,12)$	6	2,91

Диапазон нормированных входных значений АЦП, $\log_2 d(k)  - y(k)$	Выход адаптивного АЦП, $I(k)$	Выход обратного адаптивного АЦП, $d_q(k)$
[2,34, 2,72)	5	2,52
[1,91, 2,34)	4	2,13
[1,38, 1,91)	3	1,66
[0,62, 1,38)	2	1,05
[-0,98, 0,62)	1	0,031
$[-\infty, -0,98)$	0	$-\infty$

На вход обратного адаптивного АЦП подается 4-битовый сигнал  $I(k)$ , а на выходе получается разностный сигнал  $d_q(k)$ , который затем используется в качестве входного сигнала для адаптивного предсказателя. Это отображение представлено во втором и третьем столбцах табл. 11.3.

Адаптивный предсказатель имеет структуру линейного предсказателя (см. рис. 11.6), который, впрочем, сложнее, чем предсказатель схемы DPCM. Предсказатель для схемы ADPCM является комбинацией рекурсивного фильтра второго порядка и нерекурсивного фильтра шестого порядка, что можно представить следующим уравнением (сравните с уравнением 11.1):

$$s_e(k) = \sum_{i=1}^2 A_i(k-i)s_e(k-i) + \sum_{i=1}^6 B_i(k-i)d_q(k-i), \quad (11.2)$$

где коэффициенты  $A$  и  $B$  обновляются с помощью алгоритмов адаптации к скорости изменения выборок  $s_e(k)$  и  $d_q(k)$ . Таким образом, адаптивный предсказатель настраивается на скорость изменения разностей выборок, а также на скорость изменения оценочных значений сигнала.

### Субъективное измерение производительности кодера

Целью кодирования речи является использование избыточности сигнала, структуры сигнала и знаний об особенностях человеческого восприятия для предоставления различных скоростей передачи данных с сохранением при этом максимально возможного качества сигнала. При оценке производительности кодера речи более подходят субъективные измерения качества, чем объективные, например, по отношению сигнал/шум или среднеквадратической ошибке. Впрочем, надежные значения субъективных измерений, как правило, труднее получить.

Из субъективных измерений, используемых как стандарт качества услуг, наиболее распространенным является усредненная оценка разборчивости речи (mean opinion score — MOS), которая устанавливается на основе субъективного тестирования [DAUM82]. Для измерений параметра MOS группу субъектов просят прослушать образец кодированной речи и оценить его по пятибалльной шкале. Для повышения надежности результатов тест проводится несколько раз, с разными образцами речи и разными группами слушателей. Шкала MOS используется во множестве спецификаций как стандарт качества (например, в спецификации IEEE 802.16, описанной в разделе 11.3).

Используемая шкала приведена в табл. 11.4, где также даны две ее альтернативные (но при этом эквивалентные) интерпретации. На рис. 11.8 показаны

результаты тестов MOS из отчета [RAB195] для различных схем кодирования речи, включая ADPCM. Обратите внимание, что даже некодированная речь не получает наивысшей оценки 5. Это происходит потому, что слушатели иногда склонны ставить оценку 4 даже тем образцам речи, которые вполне заслуживают оценки 5 [JAYA84]. Учитывая это явление, оценки от 4 до 4,5 можно считать признаком высокого качества оцифровки.

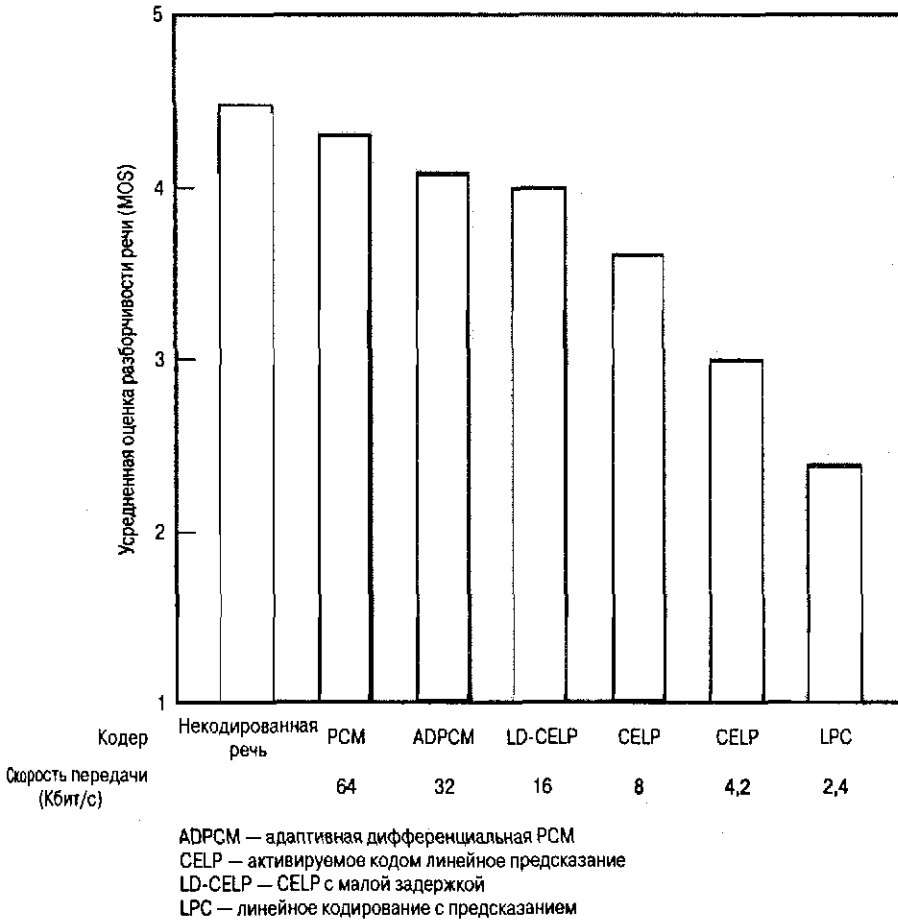


Рис. 11.8. Субъективное качество речи для различных кодеров

Таблица 11.4. Усредненная оценка разборчивости речи (MOS)

Оценка	Качество	Недостатки
5	Отличное	Незначительные
4	Хорошее	(Еле) ощутимые, но не раздражающие
3	Достаточное	(Ощутимые и) слегка раздражающие
2	Плохое	Раздражающие (но не неприятные)
1	Неудовлетворительное	Очень раздражающие (неприятные)

## 11.2. БЕСПРОВОДНЫЕ АБОНЕНТСКИЕ ЛИНИИ СВЯЗИ

Традиционно услуги передачи речи и данных предоставляются конечным пользователям посредством местной (или абонентской) линии связи, т.е. с помощью проводных систем. Для абонентов, находящихся в жилых домах, в качестве стандартного средства соединения использовались и продолжают использоваться витые пары. Для организаций и государственных учреждений используются витые пары, коаксиальный кабель и оптоволокно.

По мере роста спроса на пропускную способность, особенно после возникновения Internet, традиционной технологии витых пар становилось недостаточно. Поставщики телекоммуникационных услуг разработали множество удовлетворяющих предъявляемым требованиям технологий, включая ISDN (Integrated Services Digital Network — цифровая сеть с интеграцией услуг), а также семейство технологий цифровых абонентских линий, известное как xDSL. Кроме того, фирмы, занимающиеся кабельным телевидением, ввели двусторонние высокоскоростные услуги с использованием технологии кабельного модема. Таким образом, надежный и высокоскоростной доступ к средствам связи в жилых домах, частных и государственных конторах вполне может быть обеспечен и проводными технологиями.

Тем не менее растет интерес и к конкурирующим технологиям беспроводного абонентского доступа. Эти технологии объединяют под общим названием **беспроводные абонентские линии связи** (wireless local loop — WLL), или **стационарный беспроводный доступ**. Существует два вида беспроводных абонентских линий связи: узкополосные линии, которыми предлагается заменить существующие телефонные линии, и широкополосные линии<sup>6</sup>, которые обеспечивают двусторонние высокоскоростные услуги передачи речи и данных. Табл. 11.5, составленная на основе данных [WEBB00], позволяет сравнить линии WLL с другими технологиями, предоставляющими беспроводный доступ стационарным абонентам.

**Таблица 11.5. Альтернативные поставщики доступа конечным пользователям**

Объект	Технология	Приложение		
		Услуги телефонии	Широковещание	Компьютерные приложения
Общественная телефонная сеть	Витая пара, ISDN, xDSL	Одна и две линии	Видео по заказу	Высокоскоростной асимметричный доступ
Поставщик кабельных услуг	Коаксиальный кабель	Одна и две линии	50 и более каналов	Высокоскоростной асимметричный доступ
Поставщик услуг сотовой связи	Сотовая и беспроводная связь	Одна линия	Не предлагается	Возможна поддержка мобильности

<sup>6</sup> Термин *широкополосный* не совсем точен. Как правило, так называют системы, которые обеспечивают скорость передачи данных от 2 Мбит/с до сотен Мбит/с.

Объект	Технология	Приложение		
		Услуги телефонии	Широковещание	Компьютерные приложения
Поставщик услуг сотовой связи третьего поколения	Сотовая связь	Одна линия	Не предлагается	Высокоскоростной асимметричный доступ
Оператор узкополосных линий WLL	Беспроводная связь	Две линии	Не предлагается	Доступ к линии 64 Кбит/с
Оператор широкополосных линий WLL	Беспроводная связь	Предлагаются	50 и более каналов	Высокоскоростной асимметричный доступ
Компания наземного вещания	Аналоговое и цифровое телевидение	Не предлагаются	5–10 каналов	Небольшие возможности загрузки
Компания спутникового вещания	Аналоговое и цифровое телевидение	Не предлагаются	50 и более каналов	Не предлагаются

После обзора беспроводных абонентских линий связи будут рассмотрены некоторые технические вопросы и приведены две разновидности линий WLL. Раздел 11.3 посвящен развивающемуся стандарту для линий WLL, известному как IEEE 802.16.

## Роль линий WLL

На рис. 11.9 показана конфигурация простейшей беспроводной абонентской линии связи. Один поставщик WLL обслуживает от одной до нескольких ячеек. В каждой ячейке имеется антенна базовой станции, установленная на вершине высокого здания или на башне. У каждого отдельного абонента также имеется стационарная антенна, установленная на крыше здания или на столбе, откуда видна антенна базовой станции. Между базовой станцией и коммутатором существует канал связи, который может быть как проводным, так и беспроводным. Коммутатор обычно находится в здании местной телефонной компании, откуда имеется доступ к локальной и междугородной телефонным сетям. Поставщики услуг Internet (Internet service provider — ISP) могут располагаться там же, где и коммутатор, или подключаться к коммутатору посредством высокоскоростного канала связи.

На рис. 11.9 показана система с двухуровневой иерархией. Существуют также и более сложные конфигурации, в которых базовая станция может обслуживать несколько антенн низшего уровня, каждая из которых обслуживает множество абонентов.

Беспроводные абонентские линии связи имеют ряд преимуществ перед проводными схемами поддержки абонентов.

- **Стоимость.** Беспроводные системы обходятся не так дорого, как проводные. Хотя электронное оборудование передатчика и приемника беспроводной сети стоит дороже аналогичного оборудования для проводного соединения, отсутствуют затраты на прокладывание километров кабеля (под землей или от столба к столбу) и содержание проводной инфраструктуры.

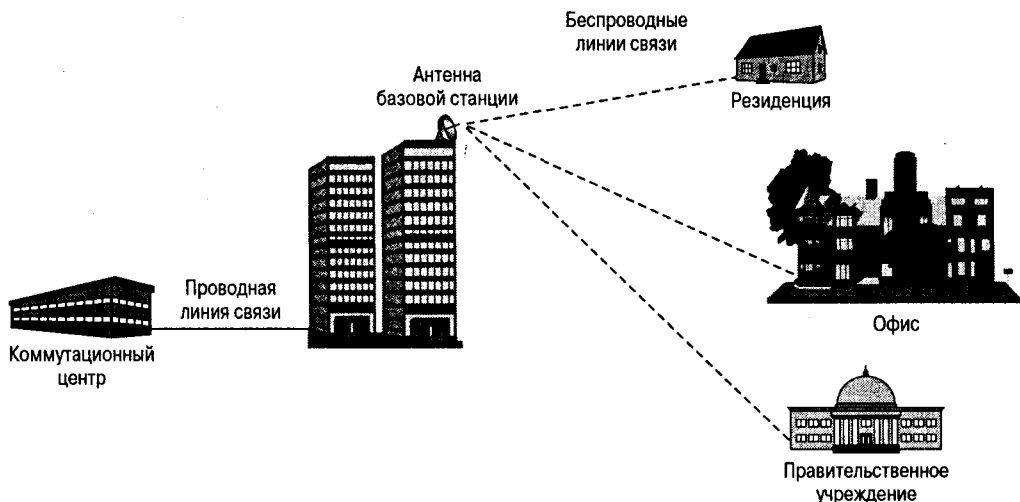


Рис. 11.9. Конфигурация беспроводной абонентской линии связи

- **Время установки.** Беспроводные системы, как правило, устанавливаются довольно быстро. Время в основном тратится только на получение разрешения на использование полосы частот и на поиск подходящей возвышенности для антенн базовой станции. Если решены эти два вопроса, то установка системы WLL займет малую толику того времени, которое потребовалось бы для установки новой проводной системы.
- **Выборочная установка.** Радиоустройства устанавливаются только для тех абонентов, которым нужны услуги в настоящее время. При прокладывании кабелей проводной системы, как правило, предполагается, что придется обслуживать всех абонентов в данном районе.

Целесообразность установки абонентских беспроводных линий связи можно оценить, сравнив схему WLL с двумя другими технологиями.

- **Проводная схема с использованием проложенного кабеля.** Большая часть жителей Земли не имеет телефонов. Многие абоненты телефонных линий либо не имеют доступа к линиям достаточно высокого качества либо находятся слишком далеко от центральной станции, чтобы эффективно использовать xDSL, что не позволяет поддерживать высокоскоростные приложения. Кроме того, многие абоненты не имеют кабельного телевидения либо поставщики не могут обеспечить услуг двусторонней передачи данных. Наконец, по стоимости линии WLL вполне способны конкурировать с проводными схемами. Поэтому при установке новой системы приходится призадуматься, какую из схем выбрать: проводную или беспроводную.
- **Мобильные сотовые технологии.** Современные сотовые системы слишком дороги и слишком ограничены в средствах, чтобы представлять собой реальную альтернативу линиям WLL. Даже тогда, когда станут доступны сотовые системы третьего поколения, они, скорее всего, будут более дорогими и менее функциональными, чем широкополосные системы WLL. Основным преимуществом систем WLL перед мобильными сотовыми системами явля-

ется то, что абонентское устройство стационарно, поэтому абонентская антенна направляется на антенну базовой станции и обеспечивает максимальное возможное качество передачи в обоих направлениях.

В США Федеральная комиссия по средствам связи выделила для коммерческих стационарных беспроводных услуг 15 полос частот, находящихся в диапазоне 2–40 ГГц. В других странах выделены подобные полосы частот. Следует отметить, что эти частоты значительно выше тех, что используются для сотовых систем. На этих частотах, которые часто называют миллиметровыми, характеристики распространения сигналов существенно отличаются от характеристик мегагерцового диапазона. Далее будут проанализированы характеристики распространения сигналов, а затем представлена технология, используемая во множестве систем WLL и известная как ортогональное FDM (OFDM). В заключение будут рассмотрены два наиболее интересных подхода к приложениям WLL: локальные многоточечные распределительные услуги (LMDS) и многоканальные многоточечные распределительные услуги (MMDS).

## Распространение сигнала в системах WLL

В большинстве высокоскоростных схем WLL используется диапазон частот, называемый миллиметровым. Хотя четкого определения термина *миллиметровая волна* нет, как правило, считается, что она имеет частоту не менее 10 ГГц. Т.е. миллиметровым считается диапазон 10–300 ГГц<sup>7</sup>. Этот диапазон был выбран для систем WLL по следующим причинам.

1. Выше частоты 25 ГГц находятся широкие полосы частот, которые практически не используются.
2. При таких высоких частотах можно использовать широкие каналы, что приводит к увеличению скорости передачи данных.
3. Можно использовать трансиверы малых размеров и блоки адаптивных антенн.

В то же время миллиметровые системы имеют ряд нежелательных характеристик распространения.

1. Потери в свободном пространстве растут как квадрат частоты (уравнение (5.2)); т.е. в этом диапазоне потери оказываются значительно больше, чем в диапазонах, используемых для обычных микроволновых систем.
2. Для частот ниже 10 ГГц, как правило, можно игнорировать эффекты затухания сигнала, обусловленные осадками и атмосферным или газовым поглощением. При частотах выше 10 ГГц эти эффекты затухания становятся довольно заметными.
3. Существенно возрастают потери вследствие многолучевого распространения. Как отмечалось в главе 5, при падении электромагнитной волны на поверхность, характерные размеры которой значительно превышают длину волны сигнала, происходит отражение. Если размеры препятствия сравнимы с длиной волны или меньше ее, происходит рассеяние, если же волновой

---

<sup>7</sup> Длина волны с частотой 10 ГГц в свободном пространстве составляет 30 мм, а длина волны с частотой 300 ГГц — 1 мм.

фронт попадает на края препятствия, которые велики по сравнению с длиной волны, то происходит дифракция.

Ввиду перечисленных негативных характеристик системы WLL могут обслуживать только ячейки ограниченного радиуса, который обычно не превышает одного километра. Кроме того, нужно обходить барьеры, встречающиеся на линии прямой видимости, включая листву. Наконец, на диапазон применения и доступность систем WLL влияют осадки и уровень влажности.

### Зоны Френеля

Для эффективной связи с помощью миллиметровых волн нужно обеспечить беспрепятственную линию прямой видимости между передатчиком и приемником. Возникает вопрос: сколько же пространства вокруг прямого тракта между передатчиком и приемником должно быть свободно от преград? При ответе на него удобно использовать такое понятие, как зоны Френеля.

Понятие зон Френеля основано на принципе Гюйгенса, согласно которому любой малый элемент пространства на пути электромагнитной волны может рассматриваться как источник вторичных волн, и поле излучения может рассматриваться как суперпозиция всех вторичных волн. На основе этого принципа можно показать, что объекты, лежащие внутри концентрических окружностей, проведенных вокруг линии прямой видимости двух трансиверов, могут влиять на качество связи как положительно, так и отрицательно. Все препятствия, попадающие внутрь первой окружности, первой зоны Френеля, оказывают наиболее негативное влияние.

Рассмотрим точку, находящуюся на прямом тракте между передатчиком и приемником, причем расстояние от точки до передатчика равно  $S$ , а расстояние от точки до приемника равно  $D$ , т.е. расстояние между передатчиком и приемником равно  $S + D$  (рис. 11.10). Вычислим радиус первой зоны Френеля в этой точке:

$$R = \sqrt{\frac{\lambda SD}{S + D}},$$

где  $R$ ,  $S$  и  $D$  измеряются в одних и тех же единицах, а  $\lambda$  обозначает длину волны сигнала вдоль тракта. Для удобства формулу можно переписать следующим образом:

$$R_m = 17,3 \sqrt{\frac{1}{f_{\text{ГГц}}} \frac{S_{\text{км}} D_{\text{км}}}{S_{\text{км}} + D_{\text{км}}}},$$

где  $R$  выражается в метрах, два остальных расстояния — в километрах, а частота сигнала — в гигагерцах.

**Пример.** Пусть расстояние между двумя трансиверами равно 10 км, а частота несущей — 2,4 ГГц. Тогда радиус первой зоны Френеля в точке, расположенной посередине между трансиверами, равен 17,66 м. Если частота сигнала — 28 ГГц, радиус зоны в той же точке равен 5,17 м. Рассмотрим теперь точку тракта, которая отстоит на 1 км от одного из трансиверов. Для сигнала частотой 28 ГГц радиус первой зоны Френеля в этой точке будет равен 3,1 м.



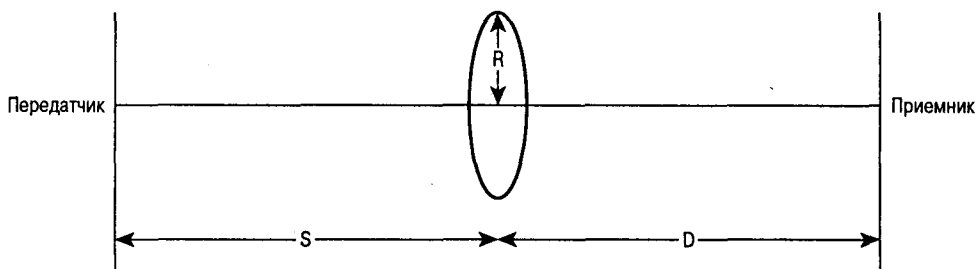


Рис. 11.10. Первая зона Френеля

Было установлено, что если внутри окружности, радиус которой составляет примерно 0,6 радиуса первой зоны Френеля, проведенной вокруг любой точки между двумя трансиверами, нет никаких преград, то затуханием сигнала, обусловленным наличием преград, можно пренебречь [FREE97]. Одной из преград является земля. Следовательно, высота двух антенн должна быть такой, чтобы вдоль тракта не было ни одной точки, расстояние от которой до земли было бы меньше, чем 0,6 радиуса первой зоны Френеля.

### Атмосферное поглощение

При распространении радиоволн с частотой свыше 10 ГГц через атмосферу происходит молекулярное поглощение. Интенсивность этого поглощения является нечетной функцией частоты, которая показана на рис. 11.11. Приблизительно на частоте 22 ГГц находится пик, соответствующий поглощению водяным паром, а на частоте около 60 ГГц — пик, соответствующий поглощению кислородом. На рисунке видно, что существует два благоприятных окна для сообщения, одно расположено в диапазоне примерно от 28 ГГц до 42 ГГц, и затухание в нем не превышает 0,13 дБ/км, другое находится в диапазоне 75–95 ГГц, и здесь затухание составляет порядка 0,4 дБ/км.

Впрочем, следует отметить, что данные о поглощении, представленные на рис. 11.11, соответствуют определенным значениям температуры, относительной влажности и атмосферного давления. Для других значений этих параметров форма кривых остается той же, однако изменяются абсолютные значения, особенно при изменении температуры и относительной влажности. В табл. 11.6, составленной на основе данных [DALK96], показано влияние температуры и влажности.

Таблица 11.6. Поглощение сигнала с частотой 28 ГГц чистым воздухом (дБ/км)

	Относительная влажность		
	0%	50%	100%
0°	0,02	0,05	0,08
10°	0,02	0,08	0,14
20°	0,02	0,12	0,25
30°	0,02	0,20	0,44
40°	0,01	0,33	0,79

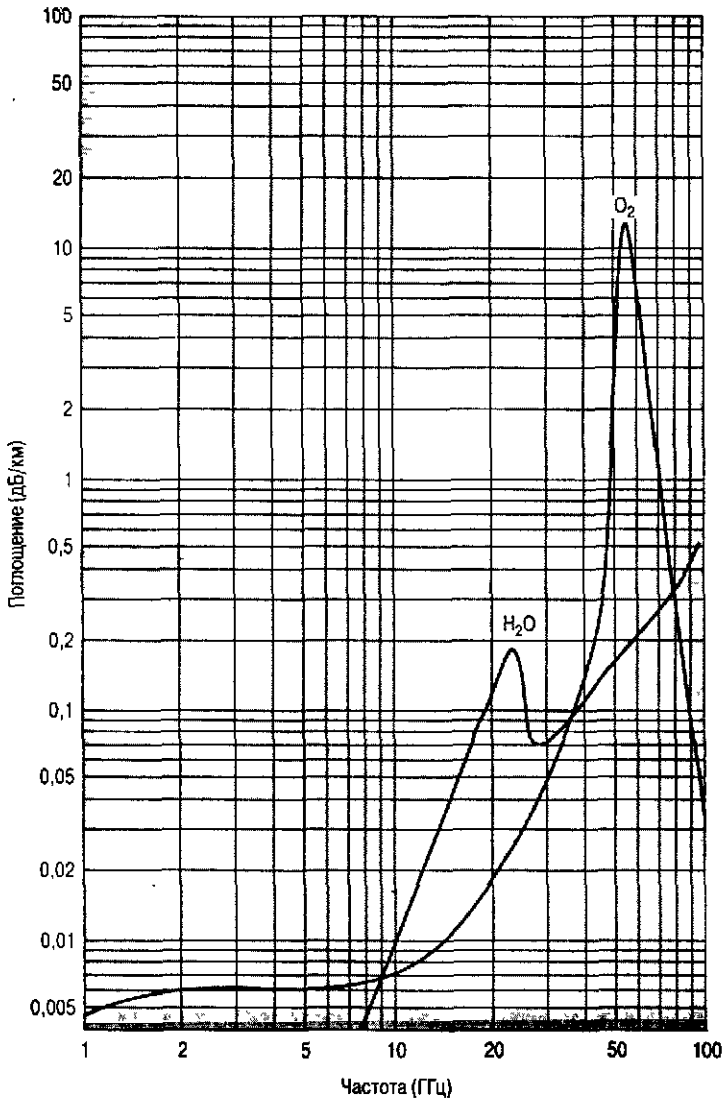


Рис. 11.11. Поглощение атмосферным газом  
(атмосферное давление — 1013 мбар; температура — 15 °С; плотность водяного пара — 7,5 г/м<sup>3</sup>)

**Пример.** Рассмотрим линию связи длиной 6 км с частотой передачи 25 ГГц. Каковы будут потери в свободном пространстве? Чему будут равны дополнительные потери из-за атмосферного затухания в жаркий сырой день при температуре 30 °С и относительной влажности 100%?

Из уравнения (5.2) находим потери в свободном пространстве:  $20 \lg(f) + 20 \lg(d) - 147,56 = 136,95$  дБ. Дополнительные потери, обусловленные атмосферным затуханием, составляют  $6 \times 0,44 = 2,64$  дБ.

## Влияние осадков

Одной из самых серьезных проблем при распространении миллиметровых волн является поглощение, вызванное осадками. Наличие капель дождя может привести к значительному снижению надежности и производительности канала связи, а в периоды сильных дождей может являться решающим фактором. Учесть влияние дождя на распространение миллиметровых волн довольно сложно, так как нужно определять форму и размеры капли, частоту падения капель, интенсивность выпадения осадков и т.п. Наиболее распространена формула оценки поглощения вследствие дождя, согласно которой

$$A = aR^b, \quad (11.3)$$

где поглощение измеряется в децибелах на 1 км, интенсивность дождя измеряется в миллиметрах в час, а параметры  $a$  и  $b$  зависят от распределения размеров капель и от частоты. Более того, поглощение зависит также от поляризации электромагнитной волны. В табл. 11.7, составленной на основе данных [FREE97], приведены типичные значения параметров  $a$  и  $b$  как функций частоты для горизонтально и вертикально поляризованных волн.

Таблица 11.7. Коэффициенты для оценки поглощения вследствие дождя

Частота, ГГц	$a_h$	$a_v$	$b_h$	$b_v$
1	0,0000387	0,0000352	0,912	0,880
2	0,000154	0,000138	0,963	0,923
6	0,00175	0,00155	1,308	1,265
10	0,0101	0,00887	1,276	1,264
20	0,0751	0,0691	1,099	1,065
30	0,187	0,167	1,021	1,000
40	0,350	0,310	0,939	0,929
50	0,536	0,479	0,873	0,868

При определенных значениях  $a$  и  $b$  поглощение зависит от интенсивности дождя,  $R$ . Основное беспокойство вызывают те периоды времени, в течение которых интенсивность дождя превышает некоторый порог. Здесь все зависит от климатической зоны. В табл. 11.8, разработанной Международным телекоммуникационным союзом (ITU), Земля разделена на 15 климатических зон в зависимости от выпадения осадков и показаны значения  $R$ , которые в течение года превышаются на различные периоды времени. Эту информацию можно использовать для определения доступности канала радиосвязи.

**Пример.** Для вертикально поляризованной волны с частотой 30 ГГц, распространяющейся в климатической зоне P, 0,01% времени распространения интенсивность осадков превышает 145 мм/ч. Определим поглощение, которое следует добавить к потерям в свободном пространстве:  $A = 0,167 \times 145 = 24,2$  дБ/км.

**Таблица 11.8. Превышаемая интенсивность осадков (в мм/ч) для разных регионов**

		Дождевая климатическая зона														
		A	B	C	D	E	F	G	H	J	K	L	M	N	P	Q
Процент времени, в течение которого превышает указанное значение	1	0,1	0,5		2,1	0,6	2	3	2	8	1,5	2	4	5	12	24
	0,3	0,8	2	13	4,5	2,4	5	7	4	13	4,2	7	11	15	34	49
	0,1	2	8	5	8	6	8	12	10	20	12	15	22	35	64	72
	0,01	8	12	15	19	22	28	30	32	35	42	60	63	95	145	115

### Влияние растительности

Часть тракта канала связи WLL может проходить сквозь растительность, по большей части сквозь листву высоких деревьев. В некоторых пригородных областях и в маленьких городах такие преграды, скорее всего, устранить не удастся, даже устанавливая антенны на крышах. Исследование, описанное в [PAPA97], привело к следующим выводам.

1. Наличие деревьев вблизи местоположения абонента может привести к замиранию вследствие многолучевого распространения.
2. Основными многолучевыми эффектами, к которым приводит наличие лиственного покрова, являются дифракция и рассеяние.
3. Измерения, проведенные в садах с периодической структурой, дали такие результаты: поглощение 12–20 дБ на одно дерево для лиственных пород и до 40 дБ для группы из 1–3 хвойных деревьев, когда листва находится внутри 60% первой зоны Френеля.
4. Эффекты многолучевого распространения находятся в сильной зависимости от ветра.

Таким образом, при установке систем WLL для каждого абонента нужно постараться, чтобы в 60% первой зоны Френеля не было листвы. Тем не менее наличие деревьев не делает связь невозможной, оно просто означает, что требуются некоторые адекватные контрмеры, например применение схем прямого исправления ошибок.

### Ортогональное уплотнение с частотным разделением

В ортогональной схеме FDM (OFDM), именуемой также модуляцией с множественными несущими, используется несколько несущих сигналов с разными частотами, посредством каждого из которых отправляется некоторое количество битов. Схема напоминает уплотнение FDM, однако в схеме OFDM все подканалы предоставлены единственному источнику данных.

На рис. 11.12 показана схема OFDM. Предположим, имеется поток данных со скоростью  $R$  бит/с и доступной полосой  $N\Delta f$ , центрированной на частоте  $f_0$ . Для отправки потока данных можно использовать полосу частот целиком, и в таком случае время передачи одного бита будет равно  $1/R$ . Кроме того, можно разделить поток данных на  $N$  подпотоков с помощью преобразователя последовательного потока в параллельный. Тогда каждый подпоток данных будет иметь

скорость передачи  $R/N$  бит/с и будет передаваться на собственной несущей частоте, причем расстояние между смежными несущими частотами будет равно  $\Delta f$ . Теперь время передачи бита будет составлять  $N/R$ .

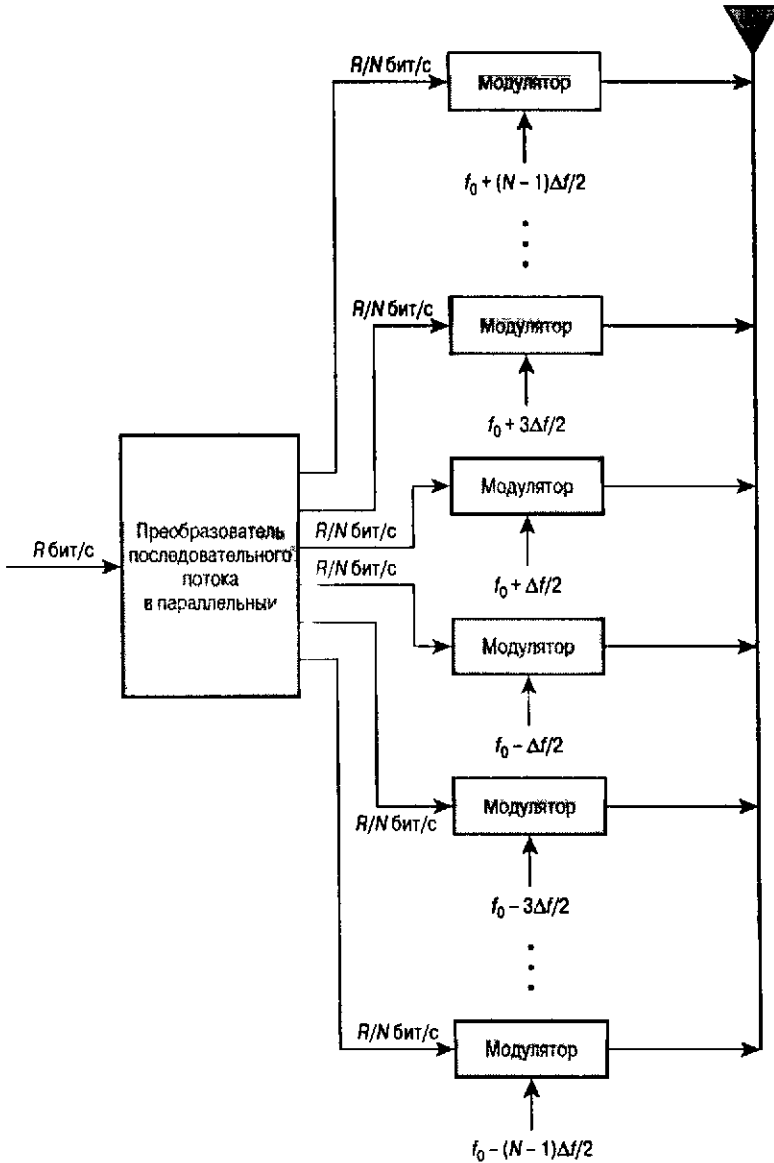


Рис. 11.12. Ортогональное уплотнение с частотным разделением

Схема OFDM имеет несколько преимуществ. Во-первых, селективному замиранию будут подвержены только некоторые подканалы, а не весь сигнал. Если поток данных защищен кодом прямого исправления ошибок, то с тем замиранием легко бороться. Но что более важно, OFDM позволяет по-

давить межсимвольную интерференцию (ISI) во многолучевой среде. Как уже отмечалось в главе 5, ISI оказывает значительное влияние при высоких скоростях передачи данных, так как расстояние между битами (или символами) является малым. В схеме OFDM скорость передачи данных уменьшается в  $N$  раз, что позволяет увеличить время передачи символа в  $N$  раз. Таким образом, если время передачи символа для исходного потока составляет  $T_s$ , то период сигнала OFDM будет равен  $NT_s$ . Это позволяет существенно снизить влияние межсимвольных помех. При проектировании системы  $N$  выбирается таким образом, чтобы величина  $NT_s$  значительно превышала среднеквадратичный разброс задержек канала.

Как видно из предыдущих рассуждений, при использовании схемы OFDM можно обойтись без эквалайзеров. Напомним, что эквалайзеры представляют собой сложные устройства, причем их сложность возрастает с увеличением количества символов, на которых приходится ISI.

Чаще всего со схемой OFDM используется модуляция QPSK (quadrature phase-shift keying — квадратурная фазовая манипуляция). В рассматриваемом случае каждый передаваемый символ представляет два бита. Вариант схемы OFDM/QPSK, используемой в системе MMDS [CISC00], занимает полосу 6 МГц, имеет 512 отдельных несущих частот, причем расстояние между соседними несущими немного меньше 12 кГц. Для минимизации межсимвольных помех данные передаются в пакетах, причем каждый пакет включает циклический префикс, за которым следуют символы данных. Циклический префикс используется для сглаживания переходов между соседними пакетами, обусловленных многолучевыми эффектами. Для рассматриваемой системы префикс состоит из 64 символов, за которыми в каждом пакете следуют 512 символов QPSK. Таким образом, в каждом подканале символы QPSK разделены префиксом длительностью 64/512 от длительности пакета. В общем случае, к моменту окончания префикса сигнал, полученный путем комбинирования сигналов, пришедших по различным трактам, не зависит от выборок предыдущих пакетов. Значит, в таком сигнале межсимвольная интерференция отсутствует.

## Многоканальные многоточечные распределительные услуги (MMDS)

В табл. 11.9 перечислены 15 полос частот, выделенных в Соединенных Штатах для беспроводного доступа. Пять из них, от 2,15 ГГц до 2,68 ГГц, используются в системах MMDS. Лицензии на первые две полосы были выданы в 70-х годах на вещание 6-мегагерцовых телевизионных каналов, или, как тогда говорилось, на многоточечные распределительные услуги (multipoint distribution service — MDS). В 1996 году Федеральная комиссия по средствам связи США увеличила долю таких полос до нынешнего уровня и разрешила предоставлять многоканальные услуги, называемые MMDS. Системы MMDS конкурировали с поставщиками телевизионных услуг и обслуживали сельскохозяйственные районы, до которых не доходило широкоэмитательное или кабельное телевидение. По этим причинам MMDS также называют *беспроводными кабельными системами*.

**Таблица 11.9. Полосы стационарной беспроводной связи, выделенные Федеральной комиссией по средствам связи США**

Частота, ГГц	Использование
2,1500–2,1620	Лицензируемые услуги MDS и MMDS; две полосы по 6 МГц каждая
2,4000–2,4835	Нелицензируемые услуги ISM
2,5960–2,6440	Лицензируемые услуги MMDS; восемь полос по 6 МГц каждая
2,6500–2,6560	Лицензируемые услуги MMDS
2,6620–2,6680	Лицензируемые услуги MMDS
2,6740–2,6800	Лицензируемые услуги MMDS
5,7250–5,8750	Нелицензируемые услуги ISM-UNII
24,000–24,250	Нелицензируемые услуги ISM
24,250–25,250	Лицензируемые услуги
27,500–28,350	Лицензируемые услуги LMDS (Блок А)
29,100–29,250	Лицензируемые услуги LMDS (Блок А)
31,000–31,075	Лицензируемые услуги LMDS (Блок Б)
31,075–31,225	Лицензируемые услуги LMDS (Блок А)
31,225–31,300	Лицензируемые услуги LMDS (Блок Б)
38,600–40,000	Лицензируемые услуги

ISM — промышленные, научные и медицинские приложения

LMDS — локальные многоточечные распределительные услуги

MDS — многоканальные распределительные услуги

MMDS — многоканальные многоточечные распределительные услуги

UNII — нелицензируемая национальная информационная инфраструктура

Передаваемая мощность, разрешенная Федеральной комиссией по средствам связи, позволяет базовой станции MMDS обслуживать зону радиусом 50 км, однако абонентские антенны должны находиться на линии прямой видимости. Систему MMDS можно использовать для поддержки двусторонних услуг; в других странах система MMDS используется также для организации двустороннего доступа. Таким образом, MMDS представляет собой альтернативу широкополосным службам обмена данными, таким, как доступ к Internet.

В отличие от систем LMDS, которые будут рассмотрены ниже, основным недостатком систем MMDS является то, что они работают на более низких частотах и могут предложить гораздо меньшую ширину полосы частот, чем системы LMDS. Даже при использовании современных технологий один канал MMDS может обеспечить в восходящем направлении (от абонента к базовой станции) скорость передачи данных 27 Мбит/с. Т.е. для отдельных абонентов скорость передачи данных будет лежать в диапазоне от 300 Кбит/с до 3 Мбит/с [ORT100]. Предоставить более высокие скорости передачи данных или поддерживать больше пользователей в восходящем направлении может позволить внедрение ортогонального FDM. И все же, скорее всего, из-за меньшей ширины полосы, системы MMDS будут использоваться в основном абонентами в жилых домах и в небольших компаниях, тогда как системы LMDS окажутся более привлекательными для компаний побольше, нуждающихся в полосах частот большей ширины.

Перечислим преимущества систем MMDS перед системами LMDS.

1. У сигналов MMDS больше длина волны (более 10 см), и они могут распространяться на большие расстояния без существенного снижения мощности. Следовательно, системы MMDS могут работать в значительно больших ячейках, что позволяет снизить затраты на оборудование для дополнительных базовых станций.
2. Оборудование для работы на более низких частотах стоит дешевле, что выгодно как для абонента, так и для владельца базовой станции.
3. Сигналы MMDS труднее блокировать объектами, и они не так подвержены поглощению осадками.

## Локальные многоточечные распределительные услуги (LMDS)

Локальные многоточечные распределительные услуги (local multipoint distribution service — LMDS) представляют собой относительно новый тип услуг WLL по доставке телевизионных сигналов и обеспечению двусторонней широкополосной связи на миллиметровых частотах. В Соединенных Штатах для систем LMDS будут выделены частоты порядка 30 ГГц, в Европе и некоторых других областях земного шара — порядка 40 ГГц. В табл. 11.9 показаны полосы частот, используемые в Соединенных Штатах для стационарного беспроводного доступа с помощью LMDS.

Системы LMDS имеют следующие преимущества.

1. Относительно высокая скорость передачи данных, измеряемая мегабитами в секунду.
2. Возможность передачи видеосигналов, телефонной связи и обмена данными.
3. Относительно низкая стоимость по сравнению с альтернативными кабельными услугами.

Основным недостатком LMDS является малый радиус зоны обслуживания для одной базовой станции. Поэтому для обслуживания одного района приходится вводить большое количество базовых станций.

В типичной системе LMDS антенна базовой станции располагается на крыше высотного здания или на высоком столбе, с которого видна вся зона обслуживания, причем линия прямой видимости, по возможности, не должна проходить через листву. Одна антенна базовой станции покрывает угловой сектор от 60° до 90°. Таким образом, для полного охвата зоны обслуживания необходимо иметь на базовой станции от 4 до 6 антенн. Радиус же зоны обслуживания одной базовой станции обычно лежит в пределах 2–4 км. Если пропускная способность нисходящего канала связи составляет 36 Мбит/с, то, учитывая доступную ширину полосы, вполне разумно установить скорость передачи данных в восходящем канале от абонента равной 1 Мбит/с.

Немаловажно то, что коротковолновые сигналы систем LMDS не могут огибать такие объекты, как здания, стены или густая листва, или проходить сквозь них. Как говорилось выше, допустимым является только некоторое количество листвы, а при наличии густой листвы придется применять специальные контрмеры. Вводя базовые станции с перекрывающимися зонами обслуживания, можно обслуживать экранированные области из ячейки одной базовой станции с базовой станции соседней ячейки. Кроме того, можно использовать ретрансляторы и отражатели.



## 11.3. СТАНДАРТ IEEE 802.16

С ростом интереса к услугам LMDS возникла потребность в разработке стандартов в этой области. В результате в 1999 году комитет IEEE 802 образовал рабочую группу 802.16 для разработки стандартов для беспроводного широкополосного доступа, которые должны были обеспечить удовлетворение следующих требований [MARK99].

- Использование беспроводной связи в СВЧ-диапазоне или миллиметровом диапазоне.
- Использование лицензируемого спектра (в основном).
- Охват в региональном масштабе.
- Наличие доступа к сети общего пользования (в основном).
- Использование точечно-многоточечной архитектуры со стационарными антеннами, размещаемыми на крышах или башнях.
- Возможность эффективной передачи неоднородного трафика при поддержании приемлемого качества обслуживания (QoS).
- Наличие возможности широкополосной передачи (>2 Мбит/с).

По сути, рабочая группа IEEE 802.16 стандартизовала радиointерфейс и функции, связанные с LMDS. На момент написания этой книги разработкой стандарта занимались три рабочие группы.

- IEEE 802.16.1. Радиointерфейс для диапазона 10–66 ГГц.
- IEEE 802.16.2. Сосуществование систем с широкополосным беспроводным доступом.
- IEEE 802.16.3. Радиointерфейс для лицензируемых частот 2–11 ГГц.

В табл. 11.10–11.12 определены область деятельности и цели этих трех групп, сформулированные при их образовании. На рис. 11.13 показано, какое место занимают стандарты 802.16.1 и 802.16.3 в общей структуре беспроводных услуг.

**Таблица 11.10. IEEE 802.16.1. Радиointерфейс для полосы 10–66 ГГц**

### Область деятельности

Стандарт определяет физический уровень и уровень управления доступом к среде радиointерфейса точечно-многоточечных систем, способных к взаимодействию и имеющих стационарный широкополосный беспроводный доступ. Спецификация позволяет передавать данные, видео и речь. Стандарт применяется к системам, действующим на частотах порядка 30 ГГц, а также к системам, работающим на частотах 10–66 ГГц.

### Цель

Проект призван содействовать быстрому внедрению инновационных, рентабельных продуктов с широкополосным беспроводным доступом, обеспечивая возможность взаимодействия продуктов от разных производителей. Предоставляя альтернативы проводному широкополосному доступу, проект будет стимулировать конкуренцию в сфере предоставления широкополосного доступа. Кроме того, проект будет стимулировать изучение сосуществования систем, поощрять глобальное регулирование и содействовать ускорению коммерциализации спектра беспроводных систем с широкополосным доступом.

## Табл. 11.11. IEEE 802.16.2. Сосуществование систем с широкополосным беспроводным доступом

### Область деятельности

В рамках этого проекта разрабатываются рекомендованные правила проектирования и координирования систем с широкополосным беспроводным доступом (BWA). Эти правила минимизируют интерференцию, а также максимизируют производительность системы и качество предоставляемых услуг. Правила должны обеспечить сосуществование систем посредством использования частотного и пространственного разделения и будут охватывать три сферы. Во-первых, через излучаемую мощность, спектральные маски и диаграммы направленности антенн будут установлены рекомендуемые границы внутрисполосного и внеполосного излучения передатчиков систем BWA. Во-вторых, будут определены рекомендуемые значения допустимых параметров приемника, включая понижение уровня шума и блокирование пропускной способности, которое возникает вследствие интерференции с другими системами BWA, а также с другими наземными или спутниковыми системами. В-третьих, будут предоставлены параметры согласования, включая схемы распределения частот, расстояния разнесения и пределы плотности потока мощности, что обеспечит успешное внедрение систем BWA с допустимой интерференцией. В область деятельности стандарта попадает интерференция систем с одинаковой полосой частот, систем, внедренных в одной географической области на разных полосах частот (включая различные системы, внедренные одним владельцем лицензии в подполосах частот, определенных в лицензии). В область деятельности стандарта не входят проблемы сосуществования, вызванные внутрисистемным многократным использованием частот в пределах полосы частот, указанной в лицензии оператора, и не будет рассматриваться влияние интерференции систем BWA с наземными или спутниковыми не BWA-системами.

### Цель

Целью рекомендуемых правил является обеспечение условий сосуществования владельцев лицензий, поставщиков услуг, групп внедрения и системотехнических предприятий. Параметры оборудования, определенные в этих правилах, помогут производителям оборудования и комплектующих, а также промышленным ассоциациям действовать в едином русле. Польза разрабатываемых правил заключается в следующем.

- Будет обеспечено сосуществование различных систем с надежной гарантией того, что будут удовлетворены требования к производительности систем.
- Будет минимальной потребность в изучении случайной интерференции и координации действий операторов для решения проблем интерференции.
- Будет обеспечена благоприятная электромагнитная среда для внедрения и работы систем BWA, включая будущие системы, удовлетворяющие стандарту 802.16.
- Будут оптимизированы покрытие зоны обслуживания и использование спектра.
- Системы будут внедрены максимально экономно.

## Таблица 11.12. IEEE 802.16.3. Радиоинтерфейс для лицензируемых частот 2–11 ГГц

### Область деятельности

Стандарт определяет физический уровень и уровень управления доступом к среде радиоинтерфейса точно-многоточечных систем, способных к взаимодействию и имеющих стационарный широкополосный беспроводный доступ (например, систем, поддерживающих скорости передачи данных DS1/E1 и выше). Спецификация позволяет передавать данные, видео и речь с определенным качеством в лицензируемых полосах частот, выделенных для доступа к общественной сети. Спецификация применяется к системам, действующим на частотах 2–11 ГГц.

### Цель

Проект призван содействовать быстрому внедрению инновационных, рентабельных продуктов с широкополосным беспроводным доступом, обеспечивая возможность взаимодействия продуктов

от разных производителей. Предоставляя альтернативы проводному широкополосному доступу, проект будет стимулировать конкуренцию в сфере предоставления широкополосного доступа. Кроме того, проект будет стимулировать изучение сосуществования систем, поощрять глобально согласованное распределение и содействовать ускорению коммерциализации спектра беспроводных систем с широкополосным доступом. Использование полос частот в диапазоне 2–11 ГГц охватывает рынок, который включает жилые дома, малые и домашние офисы (SOHO), а также малые и средние предприятия (SME).

Характеристики и приложения		Опции и вопросы	Беспроводное решение
2–644 Мбит/с Каналы связи ячеек Замена кабельных соединений	1 Транспорт	СВЧ TDM; ядро IP Дорогая кабелепроводка	Двухточечное радио
2–155 Мбит/с Малый/средний бизнес Выделение полосы по требованию	2 Привилегированный доступ	Микрометрические многоточечные системы Ограничены линией визирования Средние цены	Многоточечное радио IEEE 802.16.1
< 2 Мбит/с Домашний/малый бизнес Доступ с высоким совместным использованием	3 Доступ общего рынка	Многоточечные СВЧ-системы Сильный разброс цен Должны работать за пределами линии визирования	Многоточечное радио IEEE 802.16.3
Беспроводная Ethernet Дом "IP"	4 Беспроводные сети	Множественное применение Высокий разброс цен Зашумленные среды	Беспроводные ЛВС IEEE 802.11

Рис. 11.13. Многоуровневая перспектива беспроводной связи и распределительных систем [IEEE 802.16]

На настоящий момент больше других преуспела группа 802.16.1, работа которой, по-видимому, вызовет наиболее живой интерес в отрасли, так как касается определения доступных полос частот для систем LMDS. В этом разделе будет представлен обзор архитектуры и услуг стандарта 802.16, а затем более подробно рассмотрена спецификация стандарта 802.16.1.

## Архитектура IEEE 802.16

### Эталонная архитектура системы

Стандарты 802.16 разрабатываются по отношению к абстрактной эталонной модели системы, показанной на рис. 11.14. Согласно данным стандартам, беспроводная услуга — это предоставление тракта передачи сообщений между абонентом, которым может являться как одно абонентское устройство, так и абонентская сеть (например, ЛВС, РВХ, сеть на основе IP), и базовой сетью (доступ к которой определяется протоколом 802.16). Примерами базовых сетей являются общественная телефонная сеть и Internet. В этой модели определены три типа интерфейса. В стандарте IEEE 802.16 особое значение уделяется радиointерфейсу между трансиверной станцией абонента и базовой трансиверной станцией. В стандарте указаны все детали этого интерфейса, которые будут последовательно описаны ниже. В эталонной модели системы также представлены интерфейсы между трансиверными станциями и сетями, расположенными за ними (SNI и BNI). Подробное описание этих интерфейсов не является целью стандартов 802.16. В эталонной модели

системы данные интерфейсы представлены только для того, чтобы показать, что технологии, используемые в абонентских и базовых сетях (речь, АТМ и т.д.), влияют на технологии, используемые в радиоинтерфейсе, и на услуги, предоставляемые трансиверными станциями через этот интерфейс.

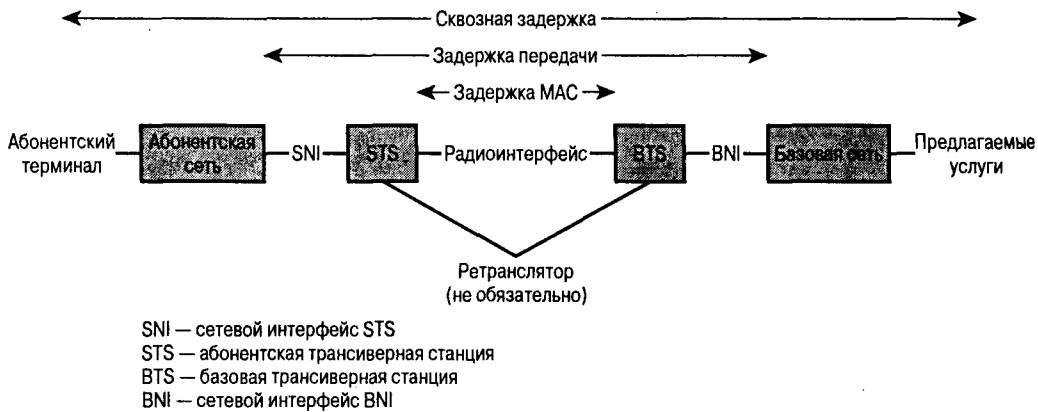


Рис. 11.14. Основные элементы системы IEEE 802.16

Наконец, в эталонную модель системы включен некий ретранслятор. В спецификации радиоинтерфейса предусмотрена возможность использования ретрансляторов или отражателей для обхода препятствий и расширения зоны обслуживания ячейки.

### Архитектура протоколов

Протоколы, определенные для беспроводной передачи, решают вопросы, связанные с передачей блоков данных по сети. В терминах OSI высокоуровневые протоколы (уровни 3, 4 и выше; см. рис. 4.3) не зависят от архитектуры сети и применимы к различным сетям и коммуникационным интерфейсам. Поэтому рассмотрение протоколов 802.16 затронет только два первых уровня модели OSI.

На рис. 11.15 показана взаимосвязь четырех уровней протоколов, определенных в архитектуре 802.16. Рассмотрим уровни снизу вверх. Два нижних уровня модели 802.16 соответствуют физическому уровню модели OSI и включают такие функции:

- кодирование и декодирование сигналов;
- создание и удаление начальной последовательности битов (используется для синхронизации);
- передача и прием битов.

Кроме того, физический уровень модели 802 включает в себя спецификацию среды передачи и полосы частот. Вообще считается, что эти вопросы рассматриваются "под" нижним уровнем модели OSI. Тем не менее выбор среды передачи и полосы частот играет решающую роль при проектировании беспроводной связи, поэтому включено также и описание среды. На самом деле все вопросы, касающиеся среды, относят к физическому уровню протокола 802.16, а приведенный выше перечень вопросов рассматривают на уровне передачи.

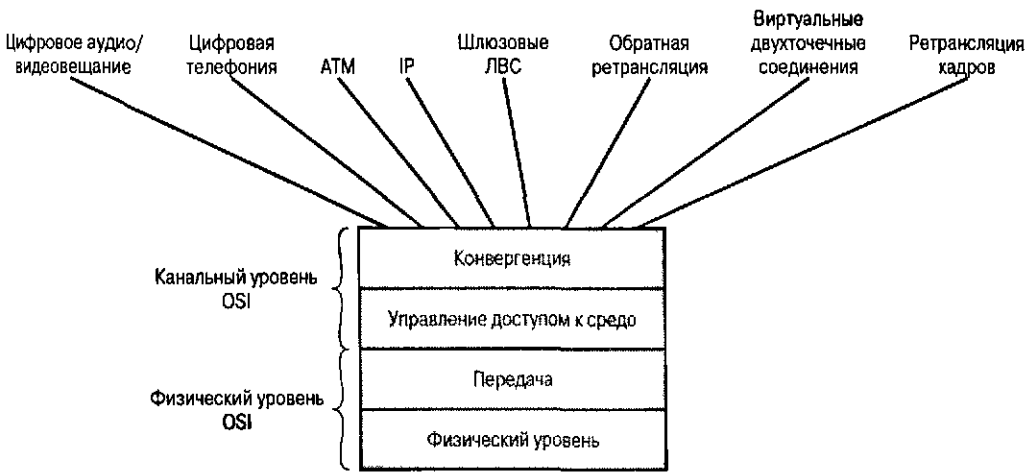


Рис. 11.15. Архитектура протоколов IEEE 802.16

Над физическим уровнем и уровнем передачи находятся функции, связанные с предоставлением услуг абонентам. Приведем примеры таких функций.

- При передаче: сбор данных в кадр, имеющий адресное поле и поле обнаружения ошибок.
- При приеме: разбор кадра, распознавание адреса и обнаружение ошибок.
- Управление доступом к беспроводной среде передачи.

Эти функции собраны на уровне управления доступом к среде (MAC). На этом уровне протокол между базовой станцией и абонентской станцией отвечает за предоставление доступа к радиоканалу. В частности, в протоколе MAC определяется, как и когда базовая станция и абонентская станция могут начинать передачу в канале. Поскольку для некоторых уровней, находящихся выше уровня MAC (например, ATM), нужно определять уровень предоставляемых услуг (QoS), протокол MAC должен уметь распределять пропускную способность таким образом, чтобы удовлетворить требованиям, предъявляемым к этим услугам. В нисходящем направлении (от базовой станции к абонентским станциям) находится только один передатчик, потому протокол MAC для этого направления является достаточно простым. В восходящем направлении за доступ состязаются несколько абонентских станций, в результате чего протокол MAC усложняется.

Над уровнем MAC находится уровень сходимости, в котором собраны функции, характерные для конкретного вида предоставляемых услуг. Протокол уровня сходимости должен выполнять такие действия:

- инкапсулировать PDU (модули данных протокола) верхних уровней в естественные для 802.16 кадры MAC/PHY;
- отображать адреса верхнего уровня в адреса 802.16;
- преобразовывать параметры QoS в естественный формат 802.16 MAC;
- адаптировать временные зависимости трафика верхнего уровня в эквивалентную услугу MAC.

В некоторых случаях, таких, как передача цифрового аудио и видео, уровень сходимости не нужен, и поток цифровых данных передается уровню передачи. Уровень сходимости требуется для услуг высшего уровня, в которых используется структура PDU.

Пример структуры протокола с уровнем сходимости для управления трафиком на основе протокола TCP/IP показан на рис. 11.16 (сравните с рис. 4.2). Данные высшего уровня передаются на подуровень управления логическим каналом (LLC), на котором управляющая информация прилагается как заголовок, создавая *модуль данных протокола (PDU) LLC*. Эта управляющая информация используется для работы протокола LLC, представляющего собой разновидность протокола управления каналом передачи данных (см. приложение Г). Затем PDU LLC передается на уровень MAC, на котором перед модулем и за модулем присоединяется управляющая информация, образуя *кадр MAC*. Как и раньше, управляющая информация нужна для работы протокола MAC. На рисунке также показано использование протокола TCP/IP и уровень приложений, который находится над протоколами 802.16.

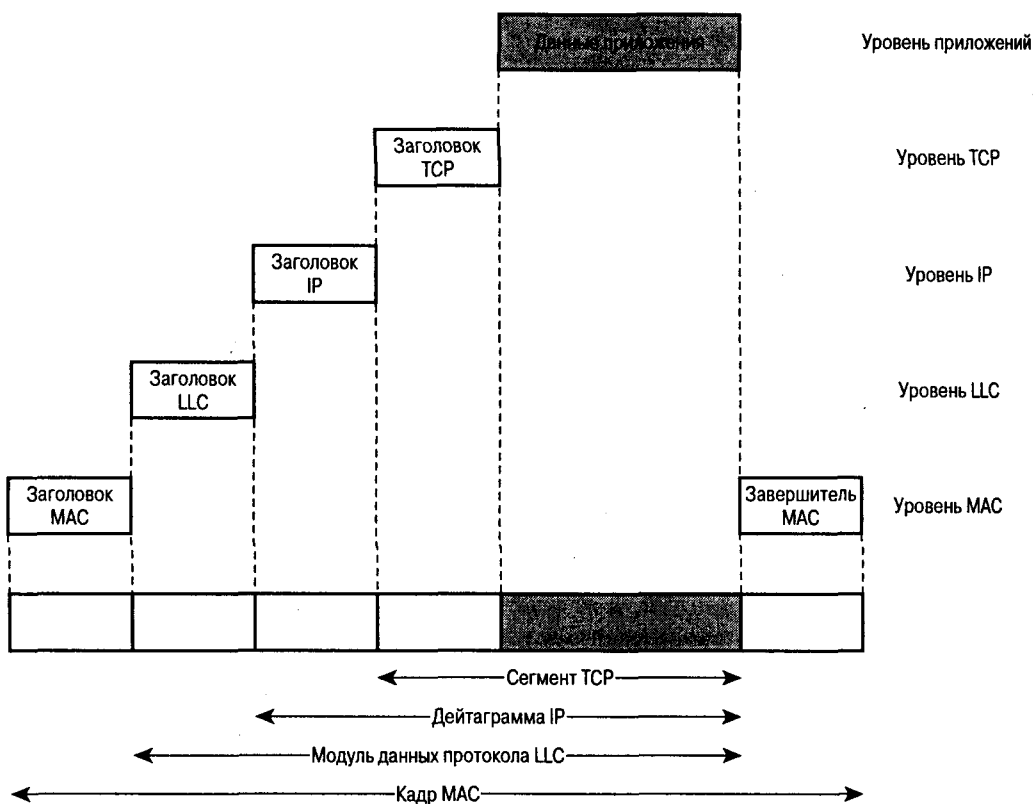


Рис. 11.16. Протоколы IEEE 802.16 в контексте

## Услуги

Требования, предъявляемые к стандартам IEEE 802.16, определяются в терминах услуг-носителей, которые должны поддерживать системы 802.16. Под услугами-носителями понимаются разновидности трафика, создаваемые абонентской или базовой сетью, которые изображены на рис. 11.14. Например, интерфейс 802.16 должен быть способен поддерживать скорость передачи данных и качество передачи, требуемые сетью ATM или сетью, основанной на протоколе IP, либо интерфейс должен поддерживать скорость передачи данных и удовлетворять требованиям, предъявляемым к задержкам при передаче аудио- и видеосигналов. Требования для отдельных услуг-носителей определены в стандартах 802.16.1 и 802.16.3.

### Услуги IEEE 802.16.1

Протокол IEEE 802.16.1 разработан для поддержки следующих услуг-носителей.

- **Многоадресная передача цифрового аудио и видео.** Односторонняя передача потоков цифровых аудио- и видеосигналов абонентам. Главным примером этой услуги является радиовещание и передача видеоданных, подобная цифровому кабельному телевидению и цифровому спутниковому телевидению. Частным случаем этой услуги является двустороннее видео, подобное телеконференциям. В таких случаях задержка неприемлема, так как здесь важно обеспечить интерактивность соединения.
- **Цифровая телефония.** Поддержка разделенных цифровых телефонных потоков. Эта классическая услуга WLL является альтернативой проводному доступу к общественной телефонной сети.
- **ATM.** Предоставление канала связи, который поддерживает передачу ячеек ATM как части общей сети ATM. Канал связи в стандарте 802.16 должен поддерживать различные услуги QoS, определенные для систем ATM.
- **Протокол Internet.** Поддержка передачи дейтаграмм IP. Канал связи 802.16 должен предоставлять эффективные своевременные услуги. Кроме того, для сетей, основанных на протоколе IP, в настоящее время определены различные виды услуг и стандарт 802.16 должен поддерживать их все.
- **ЛВС с соединением через мосты.** Услуга подобна предыдущей и позволяет передавать данные между двумя ЛВС с коммутацией на уровне MAC.
- **Магистраль.** Реализуется для сотовых или цифровых беспроводных телефонных сетей. Система 802.16 может быть удобным средством предоставления беспроводных каналов связи для беспроводных телефонных базовых станций.
- **Ретрансляция кадров.** Подобна услуге ATM, но при ретрансляции кадров используются кадры переменной длины, тогда как ячейки ATM имеют фиксированный размер.

Требования, предъявляемые услугами к стандарту 802.16, также можно представить в табличной форме — табл. 11.13 взята непосредственно из документа функциональных требований. Услуги-носители группируются в три основных категории.

Таблица 11.13. Требования IEEE 802.16.1 к услугам и качеству обслуживания

	Услуги-носители	Скорость передачи полезной нагрузки MAC	Максимальная вероятность ошибки	Максимальная задержка (односторонняя)
Основанные на каналах	Высококачественные узкополосные, телефония речевого диапазона (MOS вокодера $\geq 4,0$ )	32–64 Кбит/с	BER: $10^{-6}$	5 мс
	Низкокачественные узкополосные, телефония речевого диапазона (MOS вокодера $< 4,0$ )	6–16 Кбит/с	BER: $10^{-4}$	10 мс
	Магистраль	$\leq 155$ Мбит/с	BER: $10^{-6}$	5 мс
Пакеты переменной длины	Срочные услуги передачи пакетов	4–13 Кбит/с (речь) от 32 Кбит/с до 1,5 Мбит/с (видео)	BER: $10^{-6}$	10 мс
	Несрочные услуги: IP, IPX, ретрансляция кадров, непрерывное аудио и видео, групповая передача данных и т.д.	$\leq 155$ Мбит/с	BER: $10^{-8}$	Не определена
	видео в формате MPEG	$\leq 8$ Мбит/с	BER: $10^{-11}$	TBD
Ячейки и пакеты фиксированной длины	Ретрансляция кадров ATM — CBR	От 16 Кбит/с до 155 Мбит/с	CLR: $3 \times 10^{-8}$ CER: $10^{-6}$ CMR: 1 ошибка в сутки SEBCR: $10^{-4}$	10 мс
	Ретрансляция кадров ATM — rt-VBR	От 16 Кбит/с до 155 Мбит/с	CLR: $10^{-5}$ CER: $4 \times 10^{-6}$ CMR: 1 ошибка в сутки SEBCR: $10^{-4}$	10 мс
	Ретрансляция кадров ATM — другие услуги	$\leq 155$ Мбит/с	CLR: $10^{-5}$ CER: $4 \times 10^{-6}$ CMR: 1 ошибка в сутки SEBCR: $10^{-4}$	Не определена

- **Основанные на каналах.** Эти услуги предоставляют пропускную способность с возможностью коммутации каналов, в которой соединения с абонентами устанавливаются через базовую сеть.
- **Пакеты переменной длины.** Примерами услуг, в которых используются PDU переменной длины, являются передача пакетов IP и ретрансляция



кадров. Еще одним примером является видео в формате MPEG, представляющем собой схему сжатия видеоданных, в которой последовательные блоки цифровой видеоинформации могут иметь разные размеры.

- **Ячейки и пакеты фиксированной длины.** Эта услуга предоставляется системам ATM.

Требования в таблице сгруппированы по трем пунктам. Первая категория — это скорость передачи данных, которую должны поддерживать системы. Вторая — частота появления ошибок. Для большинства услуг установлен верхний предел частоты появления ошибочных битов (bit error ratio — BER). Для систем ATM используются также различные параметры QoS.

Последняя категория — максимальная односторонняя задержка. Для отображения этого параметра в контексте на рис. 11.14 показаны три категории задержек, определенных в стандарте 802.16.

- **Средняя задержка доступа.** Когда трансиверная станция готова к передаче, средняя задержка доступа означает время ожидания станции перед началом передачи.
- **Задержка передачи.** Задержка распространения сигнала от SNI к BNI или от BNI к SNI. В нее входит средняя задержка доступа плюс время обработки сигнала на уровне MAC для подготовки его к передаче (с STS или BTS) и время подготовки к приему на уровне MAC (на BTS или STS).
- **Сквозная задержка.** Суммарная задержка передачи между терминалом в сети абонента и конечной услугой за пределами базовой сети. В нее входит и задержка передачи.

Максимальная односторонняя задержка в табл. 11.13 — это задержка передачи.

### Услуги IEEE 802.16.3

Стандарт IEEE 802.16.3 разрабатывался для поддержки следующих услуг-носителей.

- **Передача речи.** Основанные на передаче пакетов (в противоположность коммутации каналов) услуги, эквивалентные предлагаемым общественной коммутируемой телефонной сетью.
- **Передача данных.** Поддержка трафика, основанного на протоколе IP, включая определяемые IP требования к качеству обслуживания.
- **ЛВС с соединением через мосты.** Услуга подобна предыдущей и позволяют передавать данные между двумя ЛВС с коммутацией на уровне MAC.

Цифровой эквивалент перечисленных требований дан в табл. 11.14.

### Уровень MAC стандарта IEEE 802.16.1

Структура данных, передаваемых абоненту или абонентом по радиоинтерфейсу 802.16.1, — это последовательность кадров MAC. В данном контексте термин кадр MAC обозначает PDU, который содержит управляющую информацию MAC и данные более высокого уровня. Не стоит путать этот кадр с кадром TDMA, который состоит из последовательности слотов, каждый из которых выделен для конкретного абонента. Слот кадра TDMA может содержать один кадр MAC, ка-

кую-то часть кадра MAC или же несколько кадров MAC. Последовательность слотов в группе кадров TDMA, которая выделена одному абоненту, образует логический канал, по которому передаются кадры MAC.

**Таблица 11.14. Требования IEEE 802.16.3 к услугам и качеству обслуживания**

Услуга	Максимальная вероятность ошибки	Максимальная задержка (односторонняя)
Телефония высокого качества (MOS вокодера $\geq 4.0$ )	BER: $10^{-6}$	20 мс
Телефония стандартного качества (MOS вокодера $< 4.0$ )	BER: $10^{-4}$	40 мс
Срочные услуги передачи пакетов	BER: $10^{-6}$	20 мс
Несрочные услуги	BER: $10^{-9}$	Не определена

### Соединения и поток услуги

Протокол MAC стандарта 802.16.1 является ориентированным на соединение. Т.е. перед обменом данными между одноранговыми объектами (пользователями MAC) устанавливается логическое соединение. Каждый кадр MAC включает в себя идентификатор соединения, который используется протоколом MAC для доставки поступающих данных нужному пользователю MAC. Кроме того, существует однозначное соответствие между идентификатором соединения и потоком услуги. Поток услуги определяет параметры качества обслуживания (QoS) для модулей PDU, которыми обмениваются сообщающиеся объекты.

Понятие потока услуги является центральным моментом в работе протокола MAC. Потоки услуг предоставляют механизм управления качеством обслуживания в нисходящем и восходящем направлениях. В частности, они являются важной составляющей процесса распределения полосы частот. Базовая станция распределяет полосу частот для передачи в восходящем и нисходящем направлениях на основании потоков услуг для каждого активного соединения. Примерами параметров потоков услуг являются время ожидания (максимально допустимая задержка), дрожание (максимально допустимая вариация задержки) и пропускная способность (минимальная приемлемая скорость передачи битов).

### Формат кадра

Чтобы разобраться в работе протокола MAC, лучше всего рассмотреть формат его кадра (рис. 11.17). Кадр делится на три секции.

- **Заголовок.** Содержит информацию управления протоколом, необходимую для функционирования протокола MAC.
- **Полезная нагрузка.** В этой секции могут находиться данные более высокого уровня (например, ячейки ATM, пакеты IP, блоки оцифрованной речи) или управляющее сообщение MAC.
- **CRC.** Поле циклической проверки четности с избыточностью содержит код обнаружения ошибок (рассмотрен в разделе 8.1).

Определены три формата заголовка. Существуют общие форматы заголовков в восходящем (к базовой станции) и в нисходящем (к абоненту) направлениях. Эти форматы используются для кадров, в которых содержатся как данные более

высокого уровня, так и управляющие сообщения MAC. Третий формат используется для кадра, содержащего запрос на предоставление полосы частот.

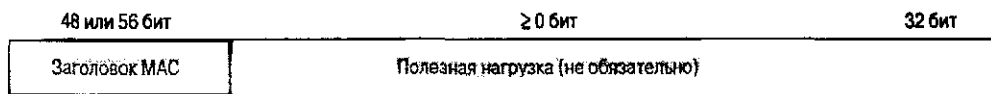


Рис. 11.17. Формат кадра IEEE 802.16.1

Формат нисходящего заголовка показан на рис. 11.18, а. Он состоит из следующих полей.

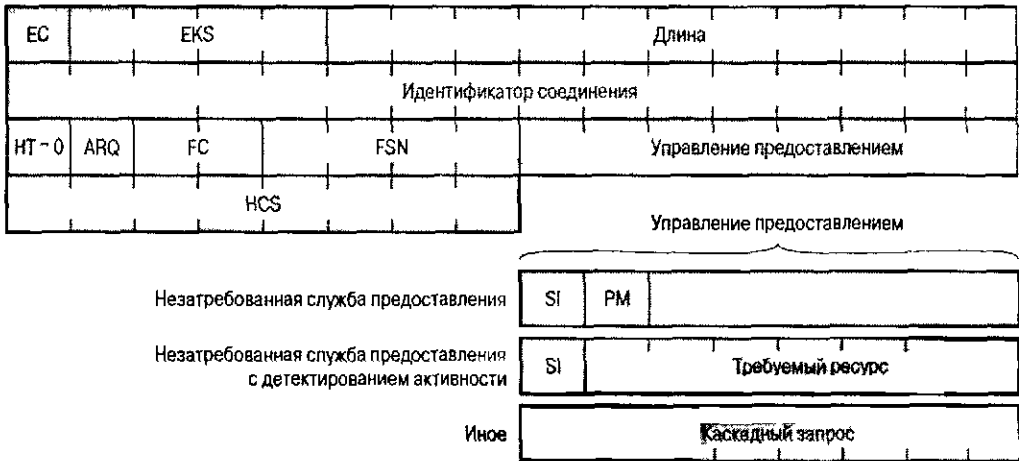
- **Управление шифрованием (1 бит).** Указывает, зашифрована ли полезная нагрузка.
- **Последовательность ключа шифрования (4 бит).** Указатель на вектор информации о ключе шифрования, который используется при шифровании полезной нагрузки.
- **Длина (11 бит).** Длина в байтах всего кадра MAC.
- **Идентификатор соединения (16 бит).** Однонаправленный адрес уровня MAC, идентифицирующий соединение с одноранговыми объектами на уровнях MAC абонента и базовой станции. Данный идентификатор отображается в идентификатор SFID, который определяет параметры качества обслуживания потока услуг, связанного с данным соединением.
- **Тип заголовка (1 бит).** Позволяет определить, является заголовок обычным или же он соответствует кадру с запросом на предоставление полосы частот.
- **Индикатор ARQ (1 бит).** Указывает, относится ли кадр соединению, допускающему ARQ. Если да, то используется механизм ARQ, описанный в разделе 8.4, и в начало кадра добавляется 2-байтовое контрольное поле. Данная структура состоит из 4-битового числа, обозначающего количество попыток повторной передачи, и 12-битового порядкового номера. Значение в поле числа попыток сбрасывается при первой отправке пакета и возрастает на единицу всякий раз, когда пакет передается повторно (до последнего значения 15). Поле порядкового номера присоединяется к каждому пакету при его первой передаче, а затем также увеличивается на единицу при последующих передачах.
- **Управление фрагментами (2 бит).** Используется при фрагментации и повторной сборке, что будет рассмотрено ниже.
- **Порядковый номер фрагмента (4 бит).** Номер текущего фрагмента в последовательности.
- **Последовательность проверки заголовка (8 бит).** Для обнаружения ошибок в заголовке используется 8-битовый код CRC.

Фрагментация используется для разделения блоков данных более высокого уровня на два или большее количество фрагментов с целью уменьшения размера кадров MAC. Это позволяет увеличить эффективность использования доступной полосы частот в соответствии с требованиями качества обслуживания потока услуг. Если фрагментация не проводится, значение в поле управления фрагментами (FC) равно 00. Если же кадр делится на фрагменты, тогда всем фрагментам будет выде-

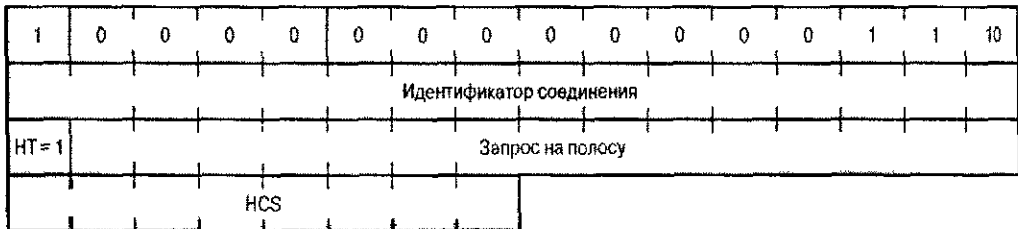
лен один и тот же порядковый номер (FSN), а значение в поле FC интерпретируется следующим образом: 01 — первый фрагмент, 11 — промежуточный фрагмент, 10 — последний фрагмент. Пользователь MAC, которому передаются эти кадры, должен будет собрать воедино все фрагменты с одним и тем же номером FSN.



а) Общий нисходящий заголовок



б) Общий восходящий заголовок



в) Заголовок запроса на полосу

ARQ — автоматический запрос повторной передачи  
 EC — управление шифрованием  
 EKS — ключевая последовательность шифрования  
 FC — управление фрагментами  
 FSN — порядковый номер фрагмента

HCS — контрольная последовательность заголовка  
 HT — тип заголовка  
 PM — бит "опроси меня"  
 SI — индикатор передвижения очереди

Рис. 11.18. Форматы заголовка MAC стандарта IEEE 802.16

Формат восходящего заголовка показан на рис. 11.18, б. Этот заголовок содержит те же поля, что и нисходящий заголовок, плюс 8-битовое поле предоставления управления (GM). Это поле используется абонентом для передачи управления полосой частот базовой станции. Для этого поля существуют три различных типа кодирования, в зависимости от типа соединения. В поле GM находятся такие подполя.

- **Индикатор продвижения очереди (1 бит).** Если этот бит установлен, он определяет продвижение права передачи относительно длины очереди восходящего канала.
- **Бит опроса (1 бит).** Если этот бит установлен, базовую станцию просят произвести опрос.
- **Требуемый ресурс (7 бит).** Запрос пропускной способности, требуемой для соединения.
- **Каскадный запрос (8 бит).** Число байтов пропускной способности восходящего канала, требуемых абоненту для этого соединения.

Первые два формата поля GM связаны с незапрашиваемым предоставлением услуги (UGS). Эта услуга разработана для поддержки потока услуг в реальном времени. По сути, базовая станция периодически, используя управляющие сообщения MAC, предоставляет абоненту право передачи некоторого количества байтов на данное соединение. Распределение происходит в соответствии с требованиями передачи в реальном времени. Если абонент обнаруживает, что очередь передаваемых им данных превышает пороговый размер, он отправляет поле GM, в котором установлен бит SI, и либо просит провести опрос на желание использовать полосу частот, задавая бит PM, либо заказывает требуемую пропускную способность в следующем слоте. Последний метод используется для услуг UGS с обнаружением активности; это означает, что поток может просто стать неактивным на значительный период времени. Для других типов услуг поле GM можно использовать для запроса пропускной способности. Это называется каскадным запросом, поскольку запрос не является отдельным кадром MAC с запросом управления полосой частот, а является частью кадра MAC, содержащего также пользовательские данные.

Наконец, заголовок запроса предоставления полосы частот используется абонентом для запроса дополнительной полосы частот. Этот заголовок используется для кадров MAC, не содержащих полезной нагрузки. В 15-битовом поле запроса указывается число байтов пропускной способности, требуемое для передачи данных по восходящему каналу.

## Управляющие сообщения MAC

В стандарте IEEE 802.16.1 определено несколько типов управляющих сообщений, которые могут использоваться базовой станцией и абонентом для управления радиointерфейсом и обменом данными по различным соединениям. Сообщения используются для обмена рабочими параметрами и информацией, относящейся к состоянию или шифрованию, а также для управления пропускной способностью. Ниже кратко описаны все типы управляющих сообщений, определенные в настоящее время.

- **Дескриптор восходящего и нисходящего каналов.** Передает характеристики физического канала.

- **Определение доступа к восходящему и нисходящему каналам.** Предоставляет доступ к восходящему и нисходящему каналам.
- **Зондирующий запрос и ответ на запрос.** Запрос используется абонентом для определения задержки сети и для запроса регулирования мощности и/или модуляции. Ответ базовой станции содержит запрашиваемые параметры.
- **Запрос регистрации, ответ на запрос и подтверждение регистрации.** Запрос передается абонентом во время инициализации и содержит соответствующие параметры. Базовая станция посылает ответ на запрос, а абонент для завершения квитирования отправляет подтверждение получения ответа.
- **Запрос управления ключом секретности и ответ на запрос.** Используется для обмена информацией о ключах шифрования.
- **Запрос динамического добавления услуги, ответ и подтверждение.** Запрос отправляется абонентом с целью заказа нового потока услуг.
- **Запрос динамического изменения услуги, ответ и подтверждение.** Запрос отправляется абонентом или базовой станцией с целью динамического изменения параметров существующего потока услуг.
- **Запрос динамического удаления услуги, ответ и подтверждение.** Запрос отправляется базовой станцией с целью удаления существующего потока услуг.
- **Запрос предоставления ширококвещательного опроса и ответ на запрос.** Отправляется абонентом с просьбой включить его в группу ширококвещательного опроса.
- **Запрос типа предоставления права передачи по нисходящему каналу.** Отправляется базовой станцией абоненту для указания метода модуляции и схемы FEC определенному соединению с целью повышения производительности.
- **Подтверждение ARQ.** Используется для подтверждения успешного приема одного или нескольких пакетов от однорангового объекта MAC.

## **Физический уровень стандарта IEEE 802.16.1**

На физическом уровне стандарта 802.16.1 поддерживаются различные структуры точно-многоточечных нисходящих каналов и многоточечно-точечных восходящих каналов (табл. 11.15). Эти структуры отражают расхождения в требованиях, предъявляемых к этим двум направлениям. Как правило, большинству систем требуется большая пропускная способность в нисходящем направлении, чтобы можно было поддерживать асимметричное соединение с отдельными абонентами для использования, например, Web-приложений в Internet. Для восходящего направления нужно рассмотреть также вопрос доступа к передающей среде, так как за доступную пропускную способность состязаются множество абонентов. Эти требования и отражены в спецификации физического уровня.

### **Восходящая передача**

Для передачи в восходящем направлении используется технология DAMA-TDMA (множественный доступ с распределением по запросу; множественный доступ с временным разделением). Как было определено в главе 9, DAMA — это схема распределения пропускной способности, адаптируемая для оптимального соответствия

меняющимся требованиям различных станций. TDMA — это просто схема разделения времени в канале на последовательность кадров, каждый из которых состоит из нескольких слотов, и один или несколько слотов в каждом кадре выделяется для создания логического канала. При использовании комбинированной схемы DAMA-TDMA распределение слотов в канале происходит динамически.

Для исправления ошибок при передаче в восходящем направлении используется код Рида-Соломона (Reed-Solomon) и схема модуляции, основанная на QPSK.

**Таблица 11.15. Модели физического уровня стандарта IEEE 802.16.1**

	Восходящий канал связи	Непрерывный нисходящий канал связи (режим А)	Пакетный нисходящий канал связи (режим В)
Схема доступа	DAMA-TDMA	TDM	DAMA-TDMA
Схема дуплексной передачи	Согласовывается с нисходящим каналом	FDD	FDD с адаптивной модуляцией; FSDD; TDD

DAMA-TDMA — множественный доступ с распределением по запросу, множественный доступ с временным разделением

FDD — дуплекс с частотным разделением каналов

FSDD — дуплекс с разделением каналов методом сдвига частоты

TDD — дуплекс с временным разделением каналов

TDM — уплотнение с временным разделением

### Нисходящая передача

В нисходящем направлении стандартом определено два режима работы: один для передачи непрерывного потока (режим А), такого, как аудио- или видеоданные, а другой — для передачи отдельных пакетов (режим В), таких, как трафик на основе протокола IP.

Для доступа к каналу в режиме передачи непрерывного потока используется простая схема TDM. Для распределения пропускной способности между восходящей и нисходящей нагрузкой используется дуплекс с частотным разделением каналов (FDD). Схема FDD означает, что для передачи данных в разных направлениях используются разные полосы частот. Эта схема эквивалентна схеме FAMA-FDMA (множественный доступ с фиксированным распределением; множественный доступ с частотным разделением). Для FDD предполагается, что все абоненты могут одновременно передавать и принимать данные, каждый на собственной частоте.

В режиме передачи пакетов для доступа к каналу используется схема DAMA-TDMA. Для разделения нагрузки между восходящим и нисходящим потоками можно использовать три альтернативные схемы.

- FDD с адаптивной модуляцией. Это та же схема FDD, что используется для передачи восходящего потока, но дополненная механизмами динамического регулирования пропускной способности для изменения модуляции и прямого исправления ошибок.
- Дуплекс с разделением каналов методом сдвига частоты (FSDD). Эта схема похожа на FDD, только не все абоненты могут передавать и принимать сигналы одновременно.

- Дуплекс с временным разделением каналов (TDD). Эта схема обсуждалась в разделе 11.1. Используется кадр TDMA, одна часть слотов в котором выделена для передачи восходящего потока, а другая — для передачи нисходящего потока.

Наличие этих альтернативных схем обеспечивает значительную гибкость при проектировании системы, что позволяет оптимизировать использование пропускной способности.

## 11.4. РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА И WEB-САЙТЫ

В [PHIL98] подробно рассмотрены стандарты DECT и PWT.

Обстоятельный анализ эффектов распространения волн миллиметрового диапазона можно найти в [FREE97]. Более короткий, но не менее полезный обзор по этой же теме приведен в [DALK96].

[NORD00] представляет собой общий обзор технологии LMDS и ее приложений. В [FUNG98] рассмотрены некоторые специальные вопросы, касающиеся проектирования приемников/передатчиков MMDS. В [WEBB00] содержится детальный обзор технологии WLL. В [KELL00] дан тщательный анализ схемы OFDM. В [BOLC01] рассмотрены особенности проектирования систем с широкополосным беспроводным доступом.

**BOLC01** Bolcskei H. et al. Fixed Broadband Wireless Access: State of the Art, Challenges, and Future Directions. — *IEEE Communications Magazine*, January 2001.

**DALK96** Dalke R., Hufford G., Ketchum R. *Radio Propagation Considerations for Local Multipoint Distribution Systems*/ National Telecommunications and Information Administration Publication PB97116511, August 1996.

**FREE97** Freeman R. *Radio System Design for Telecommunications*. — New York: Wiley, 1997.

**FUNG98** Fung P. A Primer on MMDS Technology. — *Communication System Design*, April 1998.

**KELL00** Keller T., Keller H. Adaptive Multicarrier Modulation: A Convenient Framework for Time-Frequency Processing in Wireless Communication. — In: *Proc. of the IEEE*, May 2000.

**NORD00** Nordbotten A. LMDS Systems and Their Application. — *IEEE Communications Magazine*, June 2000.

**PHIL98** Phillips J., Namee G. *Personal Wireless Communications with DECT and PWT*. — Boston: Artech House, 1998 (csdmag.com).

**WEBB00** Webb W. *Introduction to Wireless Local Loop: Broadband and Narrowband Systems*. — Boston: Artech House, 2000.



### Рекомендуемые Web-сайты.

- **Wireless Communications Association International** (wcai.com/). Представлена глобальная мировая индустрия стационарного широкополосного доступа. Хороший набор ссылок.



- **The IEEE 802.16 Working Group on Broadband Wireless Access Standards** ([grouper.ieee.org/groups/802/16/index.html](http://grouper.ieee.org/groups/802/16/index.html)). Представлена документация рабочей группы, а также архивы обсуждений.
- **U.S. International Wireless Electronic Systems Testbed** ([nwest.nist.gov/rscwc.html](http://nwest.nist.gov/rscwc.html)). Представлены результаты измерений и стандарты в сфере широкополосного беспроводного доступа. Имеются полезные документы и ссылки.

## 11.5. ТЕРМИНЫ, ВОПРОСЫ И ЗАДАЧИ

### Основные термины

адаптивная	линейный фильтр с	частотным
дифференциальная	предсказанием	разделением (OFDM)
импульсно-кодированная	локальные	стационарный
модуляция (ADPCM)	многоточечные	беспроводный доступ
беспроводная	распределительные	стационарный
абонентская линия	услуги (LMDS)	широкополосный
связи (WLL)	многоканальные	беспроводный доступ
беспроводная кабельная	многоточечные	управление доступом к
система	распределительные	среде (MAC)
бесшнуровая система	услуги (MMDS)	управление логическим
дифференциальное	модуляция с	каналом (LLC)
квантование	множественными	цифровые расширенные
дуплекс с временным	несущими	беспроводные
разделением (TDD)	ортогональное	телекоммуникации
зона Френеля	уплотнение с	(DECT)
кадр MAC		

### Вопросы

1. В каких направлениях проводится стандартизация традиционных бесшнуровых телефонных сетей?
2. В чем заключается различие между TDD и TDM?
3. Перечислите и кратко определите логические каналы DECT.
4. Какие ключевые преимущества имеют линии WLL перед проводными линиями связи?
5. Назовите основные преимущества передачи волн миллиметрового диапазона по сравнению с передачей волн большей длины.
6. Назовите основные недостатки передачи волн миллиметрового диапазона по сравнению с передачей волн большей длины.
7. Каковы основные достоинства схемы OFDM?
8. В чем состоят преимущества MMDS по сравнению с LMDS?
9. Назовите ключевые различия стандартов IEEE 802.16.1, IEEE 802.16.2 и IEEE 802.16.3.

## Задачи

1. В системе GSM, описанной в главе 10, использована схема TDMA с некоторым количеством дуплексных каналов, каждый из которых состоял из двух выделенных полос частот, по одному для передачи данных в каждом направлении. Предположим, что требуется спроектировать подобную систему с полосой частот той же ширины, но с использованием схемы TDD. Перечислите ключевые параметры системы TDD-GSM и сравните их с параметрами реальной системы.
2. Канал связи в СВЧ-диапазоне вдоль линии прямой видимости с длиной 12 км планируется установить в Германии, которая принадлежит климатической зоне Н (см. табл. 11.8). Рабочая частота канала равна 28 ГГц, требуется обеспечить доступность системы 99,99% всего времени ее работы. Местная статистика выпадения осадков неизвестна. Антенна имеет горизонтальную поляризацию.
  - а. Какую степень поглощения сигнала можно ожидать?
  - б. Повторите вычисления для 99,9% и 99% доступности и прокомментируйте практическую применимость требований для разных значений доступности.

## ПРИЛОЖЕНИЕ 11А. ЛИНЕЙНЫЙ ФИЛЬТР С ПРЕДСКАЗАНИЕМ<sup>1</sup>

Линейный фильтр с предсказанием основан на использовании аппроксимации функции рядом Тейлора. Функцию  $g(t)$ , имеющую производные всех порядков по  $t$ , можно разложить в степенной ряд Тейлора:

$$g(x) = \sum_{n=0}^{\infty} \frac{(x-t)^n}{n!} g^{(n)}(t), \quad (11.4)$$

где  $g^{(n)}(t)$  — это  $n$ -я производная функции  $g(t)$ . Предположим теперь, что требуется производить выборку значений функции  $g(t)$  с периодом  $T_s$ . Тогда, определив  $T_s = x - t$ , уравнение (11.4) можно переписать следующим образом:

$$g(t + T_s) = \sum_{n=0}^{\infty} \frac{(T_s)^n}{n!} g^{(n)}(t). \quad (11.5)$$

Из уравнения (11.5) следует, что значение сигнала в момент времени  $t + T_s$  можно предсказать, зная значения сигнала и его производных в момент времени  $t$ . В действительности для малых значений  $T_s$  (намного меньше единицы) хорошего приближения можно добиться, учитывая только первую производную:

$$g(t + T_s) \approx g(t) + T_s g'(t). \quad (11.6)$$

Обозначим значение  $k$ -й выборки  $g(t)$  в момент времени  $t = kT_s$  как  $m(k)$ ; тогда  $m(k \pm 1) = g(kT_s, \pm T_s)$ . Далее можно аппроксимировать производную  $g'(kT_s)$  выражением  $[g(kT_s) - g(kT_s - T_s)]/T_s$ . Следовательно, из уравнения (11.6) получаем

<sup>1</sup> При изложении материала данного приложения была использована работа [LATH98].

$$\begin{aligned}
 m(k+1) &\approx m(k) + T_s \left[ \frac{m(k) - m(k-1)}{T_s} \right] = \\
 &= 2m(k) - m(k-1).
 \end{aligned}$$

Таким образом, зная две предыдущие выборки, можно вычислить приближенное значение выборки в момент  $(k+1)$ . Приближение становится более точным при учете более ранних выборок. При использовании  $N$  выборок приближенная формула принимает такой вид:

$$m(k) \approx A_1 m(k-1) + A_2 m(k-2) + \dots + A_N m(k-N). \quad (11.7)$$

Правая часть уравнения (11.7) — это  $\hat{m}(k)$ , предсказанное значение выборки  $m(k)$ . Таким образом,

$$\hat{m}(k) = A_1 m(k-1) + A_2 m(k-2) + \dots + A_N m(k-N). \quad (11.8)$$

Данная формула описывает линейный предсказатель, который показан на рис. 11.6. Коэффициенты предсказателя  $A_i$  определяются из статистических свойств выборок.

# ГЛАВА 12

## ПРОТОКОЛЫ MOBILE IP И WAP

### 12.1. Mobile IP

Работа Mobile IP

Обнаружение

Регистрация

Туннелирование

### 12.2. Протокол беспроводных приложений

Обзор архитектуры

Язык разметки для беспроводных приложений

WMLScript

Среда беспроводных приложений

Беспроводный сеансовый протокол

Протокол беспроводных транзакций

Протокол безопасности беспроводного  
транспортного уровня

Беспроводный протокол дейтаграмм

### 12.3. Рекомендуемая литература и Web-сайты

### 12.4. Термины, вопросы и задачи

#### Приложение 12А. Протокол управляющих сообщений Internet

Обнаружение маршрутизатора

#### Приложение 12Б. Аутентификация сообщений

Односторонняя хэш-функция

Ключевой хэш-код

#### Приложение 12В. Примитивы и параметры служб

**В** этой главе рассмотрены два стандарта, которые поддерживают организацию беспроводных сетей на уровне приложений: Mobile IP и WAP (Wireless Application Protocol — протокол беспроводных приложений).

## 12.1. MOBILE IP

Протокол Mobile IP был разработан в связи с увеличением интереса к портативным и мобильным компьютерам, этот стандарт должен дать компьютерам возможность поддерживать связь с Internet при перемещении между разными точками подключения к Internet. Хотя Mobile IP может использоваться и с проводными соединениями, в которых компьютер отключается от одной физической точки подключения и подключается к другой, особенно он удобен именно при беспроводных соединениях.

В данном контексте термин *мобильный* означает, что пользователь соединяется с одним или несколькими приложениями Internet, что точка подключения пользователя меняется динамически и что, несмотря на изменения, все соединения автоматически поддерживаются. Данная ситуация отличается от той, когда пользователь (например, бизнесмен в пути) через свой ноутбук связывается с поставщиком услуг Internet (Internet service provider — ISP) из разных точек своего маршрута. В этом случае соединение пользователя с Internet прерывается каждый раз, когда он начинает двигаться, для нового подключения к Internet, пользователь должен инициировать новое соединение с поставщиком. Всякий раз при соединении с Internet пользователю выделяется новый, временный, IP-адрес (как правило, за это отвечает программное обеспечение со стороны поставщика). Этот временный IP-адрес используется в каждом соединении уровня приложений (например, FTP, Web-соединение). Отметим, что для такого типа доступа существует более удачный термин — *кочевой* (nomadic).

Ниже приводится общий обзор Mobile IP, после чего мы углубимся в некоторые детали.

### Работа Mobile IP

В главе 4 уже упоминалось, что маршрутизаторы используют IP-адрес в IP-дейтаграмме для выполнения маршрутизации. В частности, сетевая часть IP-адреса (рис. 4.13) используется маршрутизаторами для перемещения дейтаграмм от исходного компьютера в сеть, к которой подключен целевой компьютер. Затем последний маршрутизатор тракта, который подключен к той же сети, что и целевой компьютер, использует хостовую часть IP-адреса для доставки дейтаграмм IP по месту назначения. Кроме того, этот IP-адрес известен вышестоящему уровню протокольной архитектуры (рис. 4.4). В частности, большинство приложений Internet поддерживаются соединениями TCP. При установке TCP-соединения объекты TCP с каждой стороны соединения знают IP-адрес корреспондентского хоста. Когда сегмент TCP передается для доставки уровню IP, протокол TCP дает IP-адрес, а протокол IP создает дейтаграмму IP с IP-адресом в заголовке и передает дейтаграмму далее для доставки по маршруту целевому адресату. В то же время, если речь идет о мобильном хосте, IP-адрес может меняться при активных одном или нескольких соединениях TCP.

На рис. 12.1 схематически изображено обращение протокола Mobile IP с динамическими IP-адресами. Мобильный узел прикрепляется к определенной сети, известной как его базовая сеть (home network). IP-адрес в этой сети, именуемый базовым адресом, является статическим. Когда мобильный узел перемещает свою точку подключения в другую сеть, эта новая сеть называется чужой (foreign network). Как только мобильный узел вновь соединяется с сетью, он объявляет о своем присутствии, регистрируясь на узле сети (обычно маршрутизаторе), принадлежащем чужой сети и известном как чужой агент (foreign agent). Далее мобильный узел связывается с подобным агентом в своей базовой сети (базовым агентом) и предоставляет ему свой адрес передачи (care-of address); этот адрес определяет местоположение чужого агента. Как правило, для реализации функций базового и чужого агентов используются один или несколько маршрутизаторов сети.

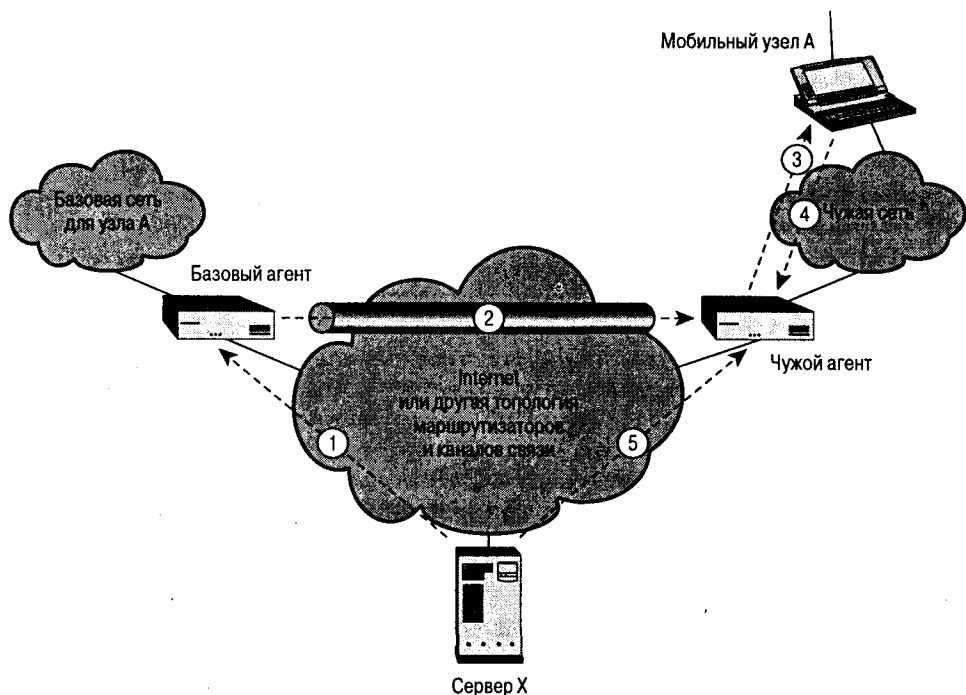


Рис. 12.1. Сценарий Mobile IP

Во время обмена дейтаграммами IP через соединение между мобильным узлом и другим хостом (на рис. 12.1 — сервером) происходят следующие действия.

1. Сервер X передает дейтаграмму IP, предназначенную мобильному узлу А, с базовым адресом А в заголовке. Дейтаграмма IP направляется по маршруту в базовую сеть А.
2. Поступившая дейтаграмма IP перехватывается в базовой сети базовым агентом. Базовый агент инкапсулирует всю дейтаграмму в новую дейтаграмму IP, у которой в заголовке находится адрес передачи А, и повторно передает дейтаграмму. Использование внешней дейтаграммы IP с иным адресом назначения называется туннелированием. Эта дейтаграмма IP передается чужому агенту.

3. Чужой агент отсекает внешний заголовок IP, инкапсулирует исходную дейтаграмму IP в PDU сетевого уровня (например, кадр LLC) и доставляет исходную дейтаграмму A через чужую сеть.
4. Когда станция A передает IP-трафик серверу X, она использует IP-адрес X. В нашем примере это стационарный адрес; т.е. X не является мобильным узлом. Каждая дейтаграмма IP передается станцией A маршрутизатору чужой сети с целью доставки ее серверу X. Как правило, этот маршрутизатор также является чужим агентом.
5. Дейтаграмма IP перемещается от A к X непосредственно через Internet, используя IP-адрес сервера X.

Поддержка операций, показанных на рис. 12.1, возможна благодаря трем основным характеристикам Mobile IP.

- **Обнаружение.** Мобильный узел использует процедуру обнаружения для идентификации предполагаемых базовых и чужих агентов.
- **Регистрация.** Мобильный узел использует процедуру регистрации с аутентификацией для оповещения базового агента о своем адресе передачи.
- **Туннелирование.** Туннелирование используется для пересылки дейтаграмм IP от базового агента по адресу передачи.

На рис. 12.2 показана протокольная поддержка возможностей Mobile IP. Протокол регистрации устанавливает связь приложения мобильного узла с приложением базового агента, поэтому использует протокол транспортного уровня. Поскольку регистрация — это простая операция “запрос-ответ”, услуги TCP на основе соединения не требуются, следовательно, как транспортный протокол используется UDP. При обнаружении используется протокол ICMP (Internet Control Message Protocol — протокол управляющих сообщений Интернет), при этом к заголовку ICMP добавляются подходящие расширения. Протокол ICMP (описан в приложении 12A) — это весьма удачный для операции обнаружения протокол без организации соединения. Наконец, туннелирование выполняется на уровне IP.

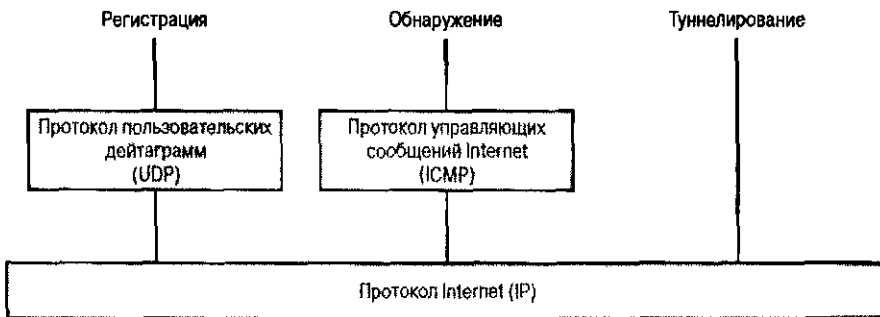


Рис. 12.2. Протокольная поддержка Mobile IP

Mobile IP описывается в нескольких документах RFC. Основной из них — RFC 2002; некоторая полезная терминология из этого документа представлена в табл. 12.1.

**Таблица 12.1. Терминология Mobile IP (RFC 2002)**

<b>Мобильный узел</b>	Хост или маршрутизатор, который изменяет свою точку подключения, перенося ее из одной сети или подсети в другую. Мобильный узел может изменять свое местоположение без смены IP-адреса; он может продолжать поддерживать связь с другими узлами Internet в любом месте, используя (постоянный) IP-адрес, пока на канальном уровне доступно соединение с точкой подключения
<b>Базовый адрес</b>	IP-адрес, который присваивается мобильному узлу на длительный период времени. Он остается неизменным, независимо от места присоединения узла к Internet
<b>Базовый агент</b>	Маршрутизатор в базовой сети мобильного узла, который туннелирует дейтаграммы для доставки к мобильному узлу, если тот находится не в базовой сети, и поддерживает информацию о текущем местоположении мобильного узла
<b>Базовая сеть</b>	Сеть (возможно, виртуальная), сетевой префикс которой совпадает с префиксом базового адреса мобильного узла. Отметим, что согласно стандартному механизму маршрутизации IP дейтаграммы, направленные по базовому адресу мобильного узла, передаются в базовую сеть мобильного узла
<b>Чужой агент</b>	Маршрутизатор в сети, посещаемой мобильным узлом, который после регистрации предлагает мобильному узлу услуги маршрутизации. Чужой агент детуннелирует и доставляет мобильному узлу дейтаграммы, туннелированные базовым агентом мобильного узла. Для дейтаграмм, посылаемых мобильным узлом, чужой агент может использоваться как маршрутизатор по умолчанию, направляющий дейтаграммы адресатам
<b>Чужая сеть</b>	Любая сеть, которая для мобильного узла не является базовой
<b>Адрес передачи</b>	Для дейтаграмм, которые передаются мобильному узлу, покинувшему базовую сеть, это конечная точка туннеля к мобильному узлу. Протокол может использовать два разных типа адресов передачи: "адрес передачи чужого агента" — это адрес чужого агента, на котором зарегистрировался мобильный узел; "внутренний адрес передачи" — полученный извне локальный адрес, который мобильный узел сопоставляет с одним из собственных сетевых интерфейсов
<b>Корреспондентский узел</b>	Одноранговый объект, с которым связывается мобильный узел. Корреспондентский узел может быть как мобильным, так и стационарным
<b>Канал</b>	Устройство или среда, посредством которой узлы могут общаться на канальном уровне. Канал составляет основу сетевого уровня
<b>Узел</b>	Хост или маршрутизатор
<b>Туннель</b>	Тракт, по которому следует инкапсулированная дейтаграмма. Принцип работы туннеля: инкапсулированная дейтаграмма передается по маршруту известному агенту, который извлекает дейтаграмму и доставляет ее конечному адресату

## Обнаружение

Процесс обнаружения в Mobile IP очень похож на процесс объявления маршрутизатора, определенный в ICMP (см. приложение 12А). Соответственно, при обнаружении агента используются сообщения объявления маршрутизатора ICMP с одним или несколькими расширениями, характерными для Mobile IP.

Мобильный узел отвечает за постоянный процесс обнаружения. Он определяет, подключен ли он к базовой сети (в этом случае дейтаграммы IP могут пе-



редаваться без пересылки между маршрутизаторами) или к чужой сети. Поскольку переключение из одной сети в другую происходит на физическом уровне, переход из базовой в чужую сеть может происходить в любое время без уведомления сетевого уровня (т.е. уровня IP). Таким образом, обнаружение для мобильного узла — это непрерывный процесс.

С целью обнаружения маршрутизатор или другой узел сети, который может действовать как агент, периодически передает сообщения с объявлением маршрутизатора (см. рис. 12.20, *г* в приложении 12А) с расширением объявления. Часть сообщения, относящаяся к объявлению маршрутизатора, включает IP-адрес маршрутизатора. Расширение объявления содержит дополнительную информацию о действиях маршрутизатора как агента, о чем будет рассказано ниже. Мобильный узел ожидает эти сообщения с объявлением агента. Поскольку чужой агент может находиться в базовой сети мобильного узла (настроенный на обслуживание гостевых мобильных узлов), прибытие объявления агента не обязательно сообщает мобильному узлу, что он находится в чужой сети. Мобильный узел должен сравнить сетевую часть IP-адреса маршрутизатора с сетевой частью своего базового адреса. В случае несоответствия мобильный узел находится в чужой сети.

Расширение объявления агента, которое следует за полями объявления маршрутизатора, состоит из таких полей (рис. 12.3).

		Биты: 0	8	16	31						
Объявление маршрутизатора (SMR)		Тип	Код	Контрольная сумма							
		Число адресов	Размер адреса	Время жизни							
		Адрес маршрутизатора 1									
		Уровень предпочтения 1									
		⋮									
		Адрес маршрутизатора <i>l</i>									
	Уровень предпочтения <i>l</i>										
Расширение объявления агента		Тип	Длина	Порядковый номер							
		Время жизни		R	V	H	F	M	G	V	Зарезервировано
		Ноль или более адресов передачи									
Необязательное расширение		Тип	Длина	Длина префикса [1]	Длина префикса [2]						
		Дополнительные длины префикса (если требуются)									

Рис. 12.3. Сообщение Mobile IP с объявлением агента

- **Тип. 16** — объявление агента.
- **Длина.**  $(6 + 4N)$ , где  $N$  — число объявленных адресов передачи.

- **Порядковый номер.** Счетчик сообщений с объявлениями агента с момента инициализации агента.
- **Время жизни.** Наибольший промежуток времени (в секундах), в течение которого агент может принимать запрос регистрации от мобильного узла.
- **R.** Требуется регистрация на данном чужом агенте (или другом агенте этой сети). Повторную регистрацию должны пройти даже мобильные узлы, уже получившие адреса передачи от данного чужого агента.
- **V.** Занято. Чужой агент не примет запросы на регистрацию от дополнительных мобильных узлов.
- **H.** Этот агент в данной сети предоставляет услуги как базовый агент.
- **F.** Этот агент в данной сети предоставляет услуги как чужой агент.
- **M.** Этот агент может получать туннелированные дейтаграммы IP, которые используют минимальную инкапсуляцию (объясняется ниже).
- **G.** Этот агент может получать туннелированные дейтаграммы IP, которые используют общую маршрутную инкапсуляцию (объясняется ниже).
- **V.** Этот агент поддерживает использование сжатия заголовка по алгоритму Ван Якобсона. Данный алгоритм определен в RFC 1144 и предназначен для сжатия полей в заголовках TCP и IP.
- **Адрес передачи.** Адрес или адреса передачи, которые поддерживаются данным агентом в данной сети. При установленном бите F в этом поле должен присутствовать хотя бы один адрес. Возможно также наличие нескольких адресов.

Кроме перечисленных полей может присутствовать необязательное **расширение**, размер которого равен длине префикса. Здесь указывается, сколько битов в адресе маршрутизатора определяют номер сети (рис. 4.13). Мобильный узел использует эту информацию для сравнения сетевой части своего IP-адреса с сетевой частью адреса маршрутизатора. Перечислим поля расширения.

- **Тип. 19** — объявление с размером, равным длине префикса.
- **Длина. N**, где  $N$  — значение в поле числа адресов в блоке объявления маршрутизатора данного сообщения ICMP. Иными словами, это количество адресов маршрутизаторов, перечисленных в данном сообщении ICMP.
- **Длина префикса.** Число начальных битов, которые определяют сетевой номер соответствующего адреса маршрутизатора, указанного в блоке объявления маршрутизатора. Число таких полей совпадает с числом полей адресов маршрутизаторов ( $N$ ).

## Запрос агента

Ожидается, что чужие агенты будут периодически передавать сообщения с объявлением агента. Если информация об агенте требуется мобильному узлу немедленно, он может передать сообщение с запросом маршрутизатора (см. рис. 12.20, *д* в приложении 12А). Любой агент, принявший это сообщение, передает в ответ объявление агента.

## Обнаружение движения

Как упоминалось ранее, мобильный узел может перемещаться из одной сети в другую с помощью некоторого механизма переключения, не уведомляя об этом уровень IP. Чтобы агент мог обнаружить такое перемещение, используется процесс обнаружения агента. Имеются два алгоритма действий агента.

- **Использование поля времени жизни.** Когда мобильный узел получает сообщение с объявлением агента от чужого агента, который он в настоящее время использует или на котором он регистрируется, он записывает значение поля времени жизни и использует его как таймер. Если время таймера истечет до получения узлом объявления от другого агента, то узел предполагает, что контакт с агентом утрачен. Если же до истечения времени таймера мобильный узел принимает от другого агента объявление с приемлемым временем жизни, он может зарегистрироваться на новом агенте. В противном случае для поиска агента мобильному узлу придется использовать запрос агента.
- **Использование префикса сети.** Мобильный узел проверяет, имеет ли какое-либо из недавно полученных объявлений агента такой же префикс сети, что и текущий адрес передачи узла. Если нет, мобильный узел предполагает, что он переместился и может регистрироваться на агенте, объявление которого было получено только что.

## Внутренние адреса

До настоящего момента в обсуждении фигурировал адрес передачи, соотнесенный с чужим агентом; т.е. адрес передачи — это IP-адрес чужого агента. По данному адресу передачи данный чужой агент будет получать дейтаграммы, предназначенные мобильному узлу, и передавать их мобильному узлу через чужую сеть. В то же время в некоторых случаях мобильный узел может переходить в сеть, не имеющую чужих агентов или имеющую только занятые чужие агенты. В качестве альтернативы этот мобильный узел может сам выступать в роли чужого агента, используя внутренний адрес передачи. Внутренний адрес — это полученный мобильным узлом IP-адрес, который соотнесен с текущим интерфейсом мобильного узла к сети.

Для протокола Mobile IP не важно, как мобильный узел получает смежный адрес. Один вариант — динамически получить временный IP-адрес через службу Internet, такую, как DHCP (Dynamic Host Configuration Protocol — протокол динамической конфигурации хоста); второй — смежный адрес может назначаться мобильному узлу как долгосрочный адрес, использующийся только во время посещения данной чужой сети.

## Регистрация

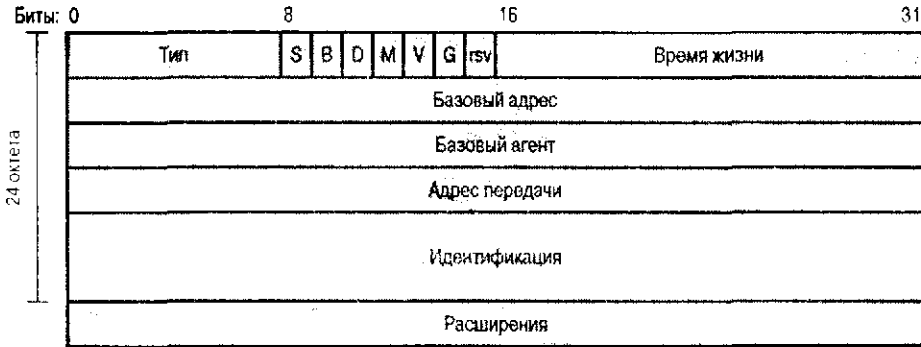
Как только мобильный узел обнаружит, что он находится в чужой сети, и получит адрес передачи, он должен предупредить об этом свой базовый агент и запросить у него пересылки IP-трафика, предназначенного узлу. Процесс регистрации состоит из таких четырех этапов.

1. Мобильный узел запрашивает услугу передачи, отправляя запрос регистрации чужому агенту, который мобильный узел желает использовать.
2. Чужой агент передает данный запрос базовому агенту мобильного узла.

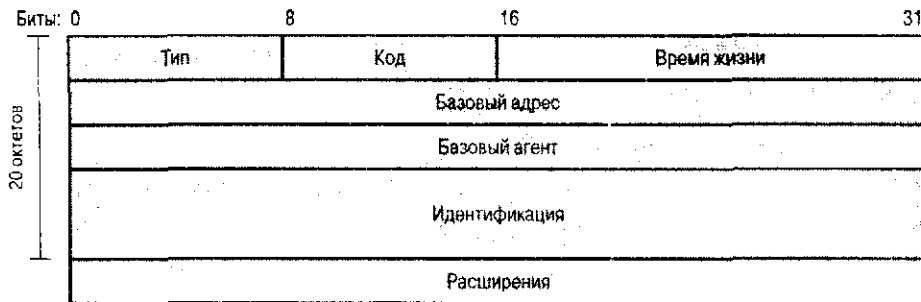
3. Базовый агент либо принимает, либо отклоняет запрос и отправляет ответ регистрации чужому агенту.
4. Чужой агент пересылает этот ответ мобильному узлу.

Если мобильный узел использует смежный адрес передачи, тогда он регистрируется непосредственно на базовом, а не на чужом агенте.

При операции регистрации используются два типа сообщений, располагающихся в сегментах UDP (рис. 12.4). Сообщение с запросом регистрации состоит из следующих полей.



а) Запрос регистрации



б) Ответ на запрос регистрации

Рис. 12.4. Сообщения регистрации Mobile IP

- **Тип. 1** — запрос регистрации.
- **S.** Одновременные привязки. Мобильный узел просит, чтобы базовый агент сохранил его прежние привязки мобильности. При одновременных активных нескольких привязках базовый агент пересылает несколько копий дейтаграмм IP, по каждому адресу передачи, зарегистрированному для данного мобильного узла. Множественные одновременные привязки могут быть полезны, например, при переключениях в беспроводных приложениях, где они позволяют увеличить надежность.

- **В.** Широковещательные дейтаграммы. Поле указывает, что мобильный узел желает получить копии широковещательных дейтаграмм, которые он должен был бы получить, если бы был присоединен к базовой сети.
- **D.** Декапсуляция мобильным узлом. Мобильный узел использует смежный адрес и декапсулирует собственные туннелированные дейтаграммы IP.
- **M.** Показывает, что базовый агент должен использовать минимальную инкапсуляцию (объясняется ниже).
- **V.** Показывает, что базовый агент должен использовать алгоритм сжатия заголовков Ван Якобсона.
- **G.** Показывает, что базовый агент должен использовать общую маршрутную инкапсуляцию (объясняется ниже).
- **Время жизни.** Время (в секундах), в течение которого действительна регистрация. Значение “нуль” является запросом на отмену регистрации.
- **Базовый адрес.** Базовый IP-адрес мобильного узла. Возможно, базовый агент получит дейтаграммы IP с таким целевым адресом; в этом случае он должен будет переслать их по адресу передачи.
- **Базовый агент.** IP-адрес базового агента мобильного узла. Информировать чужой агент об адресе, по которому должен передаваться этот запрос.
- **Адрес передачи.** IP-адрес на этом конце туннеля. Именно по этому адресу базовый агент должен пересылать дейтаграммы IP, пришедшие по базовому адресу мобильного узла.
- **Идентификация.** 64-битовое число, генерируемое мобильным узлом, которое используется для согласования запросов регистрации с ответами регистрации и в целях безопасности (объясняется ниже).
- **Расширения.** На настоящий момент определено только одно расширение — расширение аутентификации (объясняется ниже).  
**Ответ на запрос регистрации** состоит из следующих полей.
- **Тип. 3** — ответ на запрос регистрации.
- **Код.** Результат запроса регистрации (см. табл. 12.2).
- **Время жизни.** Если из поля кода следует, что регистрация принята, то в этом поле указывается время (в секундах) до окончания срока действия регистрации. Значение “нуль” сообщает об отмене регистрации мобильного узла.
- **Базовый адрес.** Базовый IP-адрес мобильного узла.
- **Базовый агент.** IP-адрес базового агента мобильного узла
- **Идентификация.** 64-битовое число, которое используется для согласования запросов регистрации с ответами на них.
- **Расширения.** На настоящий момент определено только одно расширение — расширение аутентификации (объясняется ниже).

**Таблица 12.2. Значения кодов ответа на запрос регистрации Mobile IP**

<b>Успешная регистрация</b>	<b>Отказ в регистрации чужим агентом</b>
0 — регистрация принята	64 — причина не уточняется
1 — регистрация принята, но одновременная привязка мобильности не поддерживается	65 — административный запрет
<b>Отказ в регистрации базовым агентом</b>	66 — недостаточно ресурсов
128 — причина не уточняется	67 — сбой аутентификации на мобильном узле
129 — административный запрет	68 — сбой аутентификации на базовом агенте
130 — недостаточно ресурсов	69 — затребованное время жизни слишком велико
131 — сбой аутентификации на мобильном узле	70 — неверный формат запроса
132 — сбой аутентификации на чужом агенте	71 — неверный формат ответа
133 — несоответствие идентификаторов регистрации	72 — запрошенная инкапсуляция недоступна
134 — неверный формат запроса	73 — запрошенное сжатие по Ван Якобсону недоступно
135 — слишком много одновременных привязок мобильности	80 — недоступна базовая сеть (получена ошибка ICMP)
136 — неизвестный адрес базового агента	81 — недоступен хост базового агента (получена ошибка ICMP)
	82 — недоступен порт базового агента (получена ошибка ICMP)
	88 — недоступен базовый агент (получена ошибка ICMP)

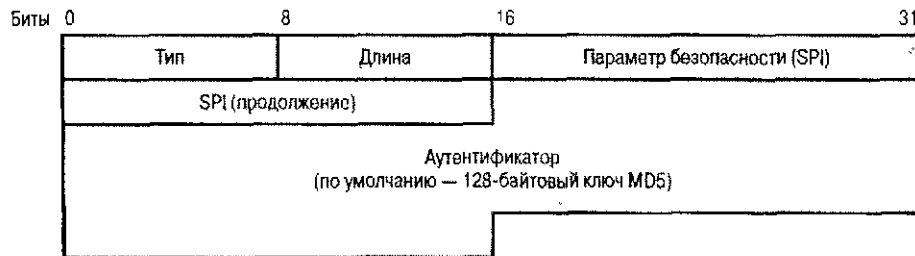
### Защита процедуры регистрации

В первую очередь при регистрации следует обратить внимание на безопасность. Протокол Mobile IP противостоит двум типам атак.

1. Узел может выдавать себя за чужой агент и отправлять базовому агенту запрос регистрации с целью присвоения трафика, предназначенного мобильному узлу.
2. Несанкционированный агент может отвечать старыми сообщениями регистрации, эффективно удаляя узел из сети.

Схема, используемая для защиты от таких атак, включает применение аутентификации сообщений и надлежащее использование поля идентификации в сообщениях с запросом регистрации и ответом на него (рис. 12.4).

Для аутентификации сообщений каждый запрос регистрации и ответ на него содержат **расширение аутентификации** (рис. 12.5) со следующими полями.



*Рис. 12.5. Расширение аутентификации Mobile IP*

- **Тип.** Используется для выделения типа данного расширения аутентификации.
- **Длина.** 4 плюс количество байтов в аутентификаторе.
- **Индекс параметра защиты (security parameter index — SPI).** Индекс, который определяет среду безопасности между парой узлов. Среда конфигурируется так, чтобы два узла сообща использовали секретный ключ и параметры, связанные с этим соединением (например, алгоритм аутентификации).
- **Аутентификатор.** Код, используемый для аутентификации сообщения. Отправитель вводит этот код в сообщение, используя общий секретный ключ. Получатель использует код для обеспечения неизменности и своевременной доставки сообщения. Аутентификатор защищает весь запрос регистрации или ответ на него, любое расширение перед этим расширением и поля типа и длины данного расширения.

По умолчанию алгоритм аутентификации использует алгоритм MD5 для генерации 128-битового профиля сообщения (алгоритм описан в приложении 12Б). В Mobile IP используется режим работы “префикс + суффикс”. Профиль MD5 вычисляется с помощью общего секретного ключа, за которым следуют защищенное поле из сообщения регистрации и снова общий секретный ключ.

*Определено три типа расширений аутентификации.*

- **Мобильный-базовый.** Это расширение должно присутствовать и обеспечивать аутентификацию сообщений регистрации между мобильным узлом и базовым агентом.
- **Мобильный-чужой.** Расширение может присутствовать при наличии безопасной связи между мобильным узлом и чужим агентом. Чужой агент отсекает это расширение перед ретрансляцией запроса базовому агенту и прибавит его к ответу, прибывающему от базового агента.
- **Чужой-базовый.** Расширение может присутствовать при наличии безопасной связи между чужим и базовым агентами.

Отметим, что аутентификатор защищает поле идентификации в сообщениях запроса и ответа. В результате идентификационное значение может использоваться для защиты от атак с использованием ответных сообщений. Как говорилось выше, идентификационное значение позволяет мобильному узлу сопоставлять ответ с запросом. Более того, если мобильный узел и базовый агент синхронизированы, так что базовый агент может отличить правильное значение от подозрительного, то базовый агент может отклонять подозрительные сообщения. Это можно реализовать, например, с использованием временной метки. До тех пор, пока мобильный узел и базовый агент в разумных пределах синхронизованы, временная метка будет давать требуемый результат. Возможен альтернативный вариант, когда мобильный узел генерирует значения с использованием генератора псевдослучайных чисел. Если базовому агенту известен алгоритм генерации, он знает следующее идентификационное значение.

## Туннелирование

Как только мобильный узел зарегистрировался на базовом агенте, последний должен иметь возможность перехватывать IP-дейтаграммы, отправленные по базовому адресу мобильного узла, и пересылать их посредством туннелирования.

Стандарт не уточняет, как это должно реализовываться, но в качестве одного из возможных механизмов называет ARP (Address Resolution Protocol — протокол разрешения адреса). Базовый агент должен проинформировать другие узлы той же сети (базовой сети), что дейтаграммы IP, направленные по адресу рассматриваемого мобильного узла, должны доставляться (на канальном уровне) данному агенту. Фактически базовый агент “притворяется” мобильным узлом, перехватывает адресованные ему пакеты и передает их через базовую сеть.

Предположим, например, станция R3 на рис. 12.6 используется как базовый агент мобильного узла, который подключен к чужой сети где-либо в Internet. Т.е. существует хост Н, базовая сеть которого — ЛВС Z, который в настоящее время подключен к некоторой чужой сети. Если у хоста D имеется трафик для Н, он создает дейтаграмму IP с базовым адресом Н в поле адреса назначения. IP-модуль хоста D обнаруживает, что данный адресат находится в ЛВС Z и передает дейтаграмму нижестоящему канальному уровню с указанием доставить ее по конкретному адресу MAC-уровня в сети Z. До этого узел R3 должен проинформировать IP-уровень узла D, что дейтаграммы, посланные по этому конкретному адресу, должны переправляться R3. Таким образом, MAC-адрес R3 вносится узлом D в поле назначения протокола MAC исходящего кадра MAC. Подобным образом, если дейтаграмма IP, направленная по базовому адресу мобильного узла, приходит на маршрутизатор R2, он обнаруживает, что адрес назначения принадлежит ЛВС Z и пытается доставить дейтаграмму по MAC-адресу в сети Z. Как и ранее, маршрутизатор R2 должен был перед этим получить информацию, что требуемый ему адрес уровня MAC соответствует R3.

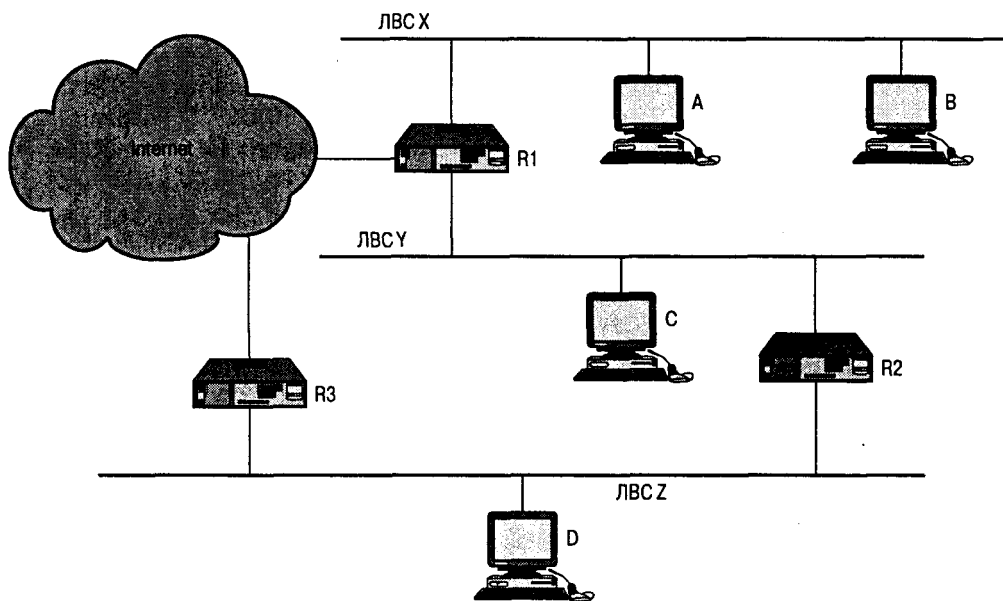


Рис. 12.6. Простой пример межсетевого обмена

Отметим, что для трафика, передаваемого через Internet и попавшего на R3 из Internet, узел R3 должен просто знать, что для такого адреса назначения дейтаграмма должна получаться и переправляться далее.



Для пересылки дейтаграммы IP по адресу передачи базовый агент помещает ее во внешнюю дейтаграмму IP. Данная инкапсуляция заключается в простом помещении заголовка IP перед сегментом TCP, т.е. сегмент TCP инкапсулируется в дейтаграмму IP. В Mobile IP возможны три разновидности инкапсуляции.

- **Инкапсуляция IP-в-IP.** Это простейший подход, определенный в RFC 2003.
- **Минимальная инкапсуляция.** При этом подходе используется меньше полей, которые определены в RFC 2004.
- **Общая маршрутная инкапсуляция (Generic Routing Encapsulation — GRE).** Общая процедура инкапсуляции, которая была разработана до протокола Mobile IP, определена в RFC 1701.

Краткий обзор первых двух методов приводится ниже.

### **Инкапсуляция IP-в-IP**

При этом подходе вся дейтаграмма IP используется как полезная нагрузка новой дейтаграммы IP (рис 12.7, а). Внутренний, исходный заголовок IP не изменяется (за исключением поля времени жизни, значение которого уменьшается на 1). Внешний заголовок — это заголовок всего фрагмента IP. Два поля (на рисунке не заштрихованы) копируются из внутреннего заголовка: номер версии (идентификатор протокола IPv4) — 4 и тип услуги, запрашиваемой для внешней дейтаграммы IP, — тот же, что во внутренней дейтаграмме IP.

Во внутреннем заголовке IP адрес источника указывает хост, который отправил исходную дейтаграмму, а адрес назначения — это базовый адрес нужного получателя. Во внешнем заголовке IP адреса источника и назначения соответствуют входу и выходу туннеля. Таким образом, адрес источника — это обычно IP-адрес базового агента, а адрес назначения — это адрес передачи нужного получателя.

**Пример.** Рассмотрим дейтаграмму IP, созданную на сервере X (рис. 12.1) и предназначенную мобильному узлу А. Адрес источника исходной дейтаграммы IP соответствует IP-адресу станции X, а адрес назначения — базовому IP-адресу станции А. Сетевая часть базового адреса А определяет базовую сеть А, так что дейтаграмма передается через Internet в базовую сеть узла А, где она перехватывается базовым агентом. Базовый агент инкапсулирует входящую дейтаграмму: добавляет внешний заголовок IP, который содержит адрес источника, равный IP-адресу базового агента, и адрес назначения, равный IP-адресу чужого агента в чужой сети, к которой в настоящее время присоединен узел А. Когда новая дейтаграмма достигнет чужого агента, он отсечет внешний заголовок IP и доставит исходную дейтаграмму станции А.

Старый заголовок IP  
Исходный заголовок IP

Версия — 4	IHL	Тип услуги	Общий размер	
Идентификация			Флаги	Смещение фрагмента
Время жизни		Протокол — 4	Контрольная сумма заголовка	
Исходный адрес (адрес базового агента)				
Адрес назначения (адрес передачи)				
Версия — 4	IHL	Тип услуги	Общий размер	
Идентификация			Флаги	Смещение фрагмента
Время жизни		Протокол	Контрольная сумма заголовка	
Исходный адрес (исходный отправитель)				
Адрес назначения (базовый адрес)				
Полезная нагрузка IP (например, сегмент TCP)				

Незатененные поля копируются из внутреннего заголовка IP во внешний

а) Инкапсуляция IP-в-IP

Заголовок минимальной маршрутизации  
Модифицированный заголовок IP

Версия — 4	IHL	Тип услуги	Общий размер	
Идентификация			Флаги	Смещение фрагмента
Время жизни		Протокол — 55	Контрольная сумма заголовка	
Исходный адрес (адрес базового агента)				
Адрес назначения (адрес передачи)				
Протокол	S	Зарезервировано	Контрольная сумма заголовка	
Адрес назначения (базовый адрес)				
Адрес отправителя (исходный отправитель; может отсутствовать)				
Полезная нагрузка IP (например, сегмент TCP)				

Незатененные поля внутреннего заголовка IP копируются из исходного заголовка IP  
Незатененные поля внешнего заголовка IP модифицируются по сравнению с исходным заголовком IP

б) Минимальная инкапсуляция

Рис. 12.7. Инкапсуляция Mobile IP

## Минимальная инкапсуляция

При минимальной инкапсуляции служебные издержки передачи меньше, а использование ее возможно при согласии мобильного узла, базового агента и чужого агента. Новый заголовок помещается между исходным заголовком IP и исходной полезной нагрузкой IP (рис. 12.7, б). Минимальная инкапсуляция содержит следующие поля.

- **Протокол.** Копируется из поля адреса назначения исходного заголовка IP. Данное поле определяет тип протокола исходной полезной нагрузки IP и таким образом определяет тип заголовка, с которого начинается исходная полезная нагрузка IP.
- **S.** Значение 0 означает, что исходный адрес источника отсутствует, а длина данного заголовка — 8 октетов. Значение 1 означает, что исходный адрес источника присутствует, а длина данного заголовка — 12 октетов.
- **Контрольная сумма заголовка.** Вычисляется по всем полям данного заголовка.
- **Исходный адрес назначения.** Копируется из поля адреса назначения исходного заголовка IP.
- **Исходный адрес источника.** Копируется из поля адреса источника исходного заголовка IP. Данное поле присутствует, только если значение бита S равно 1. Если инкапсулирующий узел является создателем дейтаграммы (т.е. дейтаграмма создана на базовом агенте), данное поле отсутствует.

Ниже перечислены поля исходного заголовка, которые модифицируются при формировании нового внешнего заголовка IP.

- **Общая длина.** Увеличивается на размер минимального заголовка передачи (8 или 12).
- **Протокол.** 55 — минимальная инкапсуляция IP.
- **Контрольная сумма заголовка.** Вычисляется по всем полям данного заголовка; значение должно рассчитываться повторно, поскольку некоторые поля были модифицированы.
- **Адрес источника.** IP-адрес инкапсулирующего узла, как правило, базового агента.
- **Адрес назначения.** IP-адрес выхода туннеля. Это адрес передачи, который может быть либо IP-адресом чужого агента, либо IP-адресом мобильного узла (если адрес передачи является внутренним).

Обработка при минимальной инкапсуляции происходит следующим образом. Инкапсулирующий узел (базовый агент) создает дейтаграмму с форматом, представленным на рис. 12.7, б. Данная дейтаграмма теперь может туннелироваться, и она доставляется через Internet по адресу передачи. После этого поля в минимальном заголовке передачи восстанавливаются до значений исходного заголовка, а сам заголовок передачи удаляется из дейтаграммы. Поле общей длины в заголовке IP уменьшается на размер минимального заголовка передачи (8 или 12), а поле контрольной суммы заголовка вычисляется заново.

## 12.2. ПРОТОКОЛ БЕСПРОВОДНЫХ ПРИЛОЖЕНИЙ

Протокол беспроводных приложений (Wireless Application Protocol — WAP) — это универсальный, открытый стандарт, разработанный Форумом WAP для предоставления мобильным пользователям беспроводных телефонов и других беспроводных терминалов, таких, как пейджеры и персональные цифровые помощники (Personal Digital Assistant — PDA), доступа к информационным услугам и услугам телефонии, включая Internet и Web. WAP предназначен для работы со всеми технологиями беспроводных сетей (например, GSM, CDMA и TDMA) и опирается на существующие стандарты Internet, такие, как IP, XML, HTML и HTTP. Кроме того, WAP содержит средства безопасности. Форум WAP был основан компаниями Ericsson, Motorola, Nokia и Phone.com в 1997 году и сейчас насчитывает несколько сот членов. Версия WAP v1.2 была выпущена Форумом в июне 2000 года.

Использование мобильных телефонов и терминалов для услуг передачи данных существенно ограничивается особенностями конструкции устройств и связующими сетями. Устройства содержат ограниченные процессоры, память и батареи. Ограничен также пользовательский интерфейс, дисплеи малы. При использовании беспроводных сетей, в отличие от проводных соединений, предполагается относительно малая ширина полосы, высокая задержка передачи и неопределенные доступность и устойчивость. Более того, все эти характеристики существенно меняются при переходе от терминала к терминалу или из сети в сеть. Наконец, мобильные пользователи с беспроводным доступом имеют разные требования к другим пользователям информационных систем. Кроме того, мобильные терминалы должны быть чрезвычайно легки в использовании, намного легче, чем рабочие станции и персональные компьютеры. Для удовлетворения всех этих разнообразных требований и разработан протокол WAP. Спецификация WAP включает:

- программную модель, основанную на программной модели WWW;
- язык разметки, Wireless Markup Language, являющийся разновидностью XML;
- спецификацию небольшого броузера, подходящего для мобильных, беспроводных терминалов;
- упрощенный стек протоколов связи;
- оболочку для приложений беспроводной телефонии (Wireless Telephony Applications — WTA).

Спецификация WAP состоит из нескольких разных протоколов и модулей, связь между которыми показана на рис. 12.8.

### Обзор архитектуры

Программная модель WAP основана на трех элементах: клиент, шлюз и исходный сервер (рис. 12.9). Для передачи между шлюзом и исходным сервером используется протокол HTTP. Шлюз действует как прокси-сервер беспроводного домена. Его процессором (процессорами) предлагаются услуги, снимающие часть нагрузки с ограниченных ручных, мобильных, беспроводных терминалов. На-

пример, шлюз предоставляет услуги DNS, состыковывает стек протоколов WAP и стек WWW (HTTP и TCP/IP), кодирует информацию из Web в более компактную форму, которая минимизирует беспроводную передачу, а в обратном направлении декодирует из сжатой формы в стандарт Web. Кроме того, шлюз помещает в кэш часто запрашиваемую информацию.

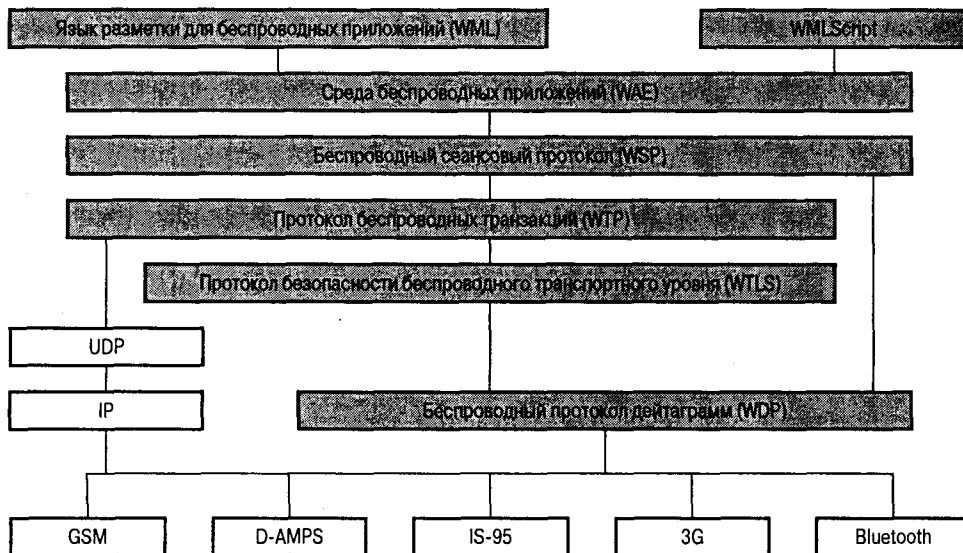


Рис. 12.8. Стек протоколов WAP

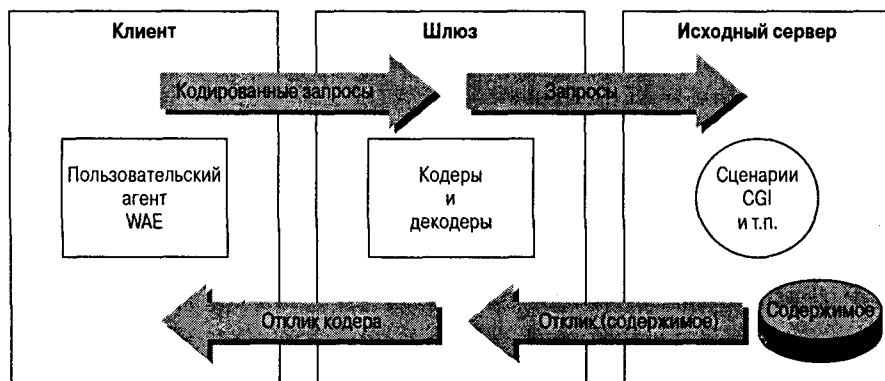


Рис. 12.9. Программная модель WAP

На рис. 12.10 иллюстрируются ключевые компоненты в среде WAP. Используя WAP, мобильный пользователь может просматривать содержимое Web на обычном Web-сервере. Web-сервер предоставляет содержимое в формате страниц HTML, которые передаются с использованием стандартного стека протоколов Web (HTTP/TCP/IP). Содержимое HTML должно проходить через фильтр HTML, который может располагаться там же, где прокси-сервер WAP, либо в отдельном физическом модуле. Фильтр преобразовывает содержимое HTML в содержимое WML. Если фильтр отделен от прокси-сервера, для доставки WML на прокси-сервер ис-

пользуется набор HTTP/TCP/IP. Прокси-сервер преобразовывает WML в более компактную форму, известную как двоичный WML и, используя стек протоколов WAP, доставляет ее мобильному пользователю через беспроводную сеть.

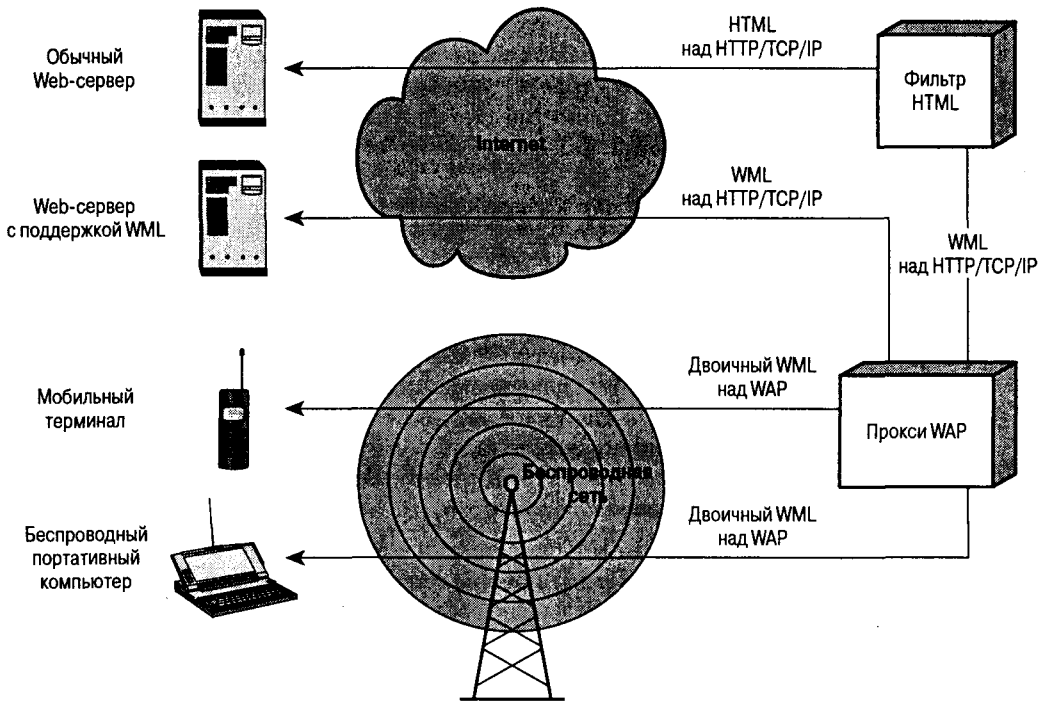


Рис. 12.10. Инфраструктура WAP

Если Web-браузер способен непосредственно генерировать содержимое WML, то это содержимое доставляется прокси-серверу с использованием HTTP/TCP/IP, после чего WML преобразуется в двоичный WML и доставляется мобильному узлу с использованием протоколов WAP.

Архитектура WAP разрабатывалась с учетом двух основных ограничений беспроводного Web-доступа: ограничений мобильного узла (малый размер экрана, ограниченные возможности ввода) и низкой скорости передачи беспроводных цифровых сетей. Даже с введением беспроводных сетей третьего поколения, которые будут обеспечивать широкополосную передачу данных, у ручных мобильных узлов останутся ограниченные возможности ввода и малые дисплеи. Таким образом, протокол WAP и подобные ему средства будут необходимы еще очень долго.

## Язык разметки для беспроводных приложений

Язык разметки для беспроводных приложений (Wireless Markup Language — WML) предназначен для описания содержимого и формата с целью представления данных на устройствах с ограниченными пропускной способностью, размером экрана и возможностями ввода данных пользователем. Он предназначен для работы с телефонными клавиатурами, перьями и другими устройствами ввода, обычными в мобильной беспроводной связи. WML допускает масштабирование дисплеев для использо-

вания двухстрочных экранов, имеющих в некоторых малых устройствах, и больших экранов, имеющих в интеллектуальных телефонах.

В обычном персональном компьютере Web-браузер предоставляет содержимое в форме Web-страниц, закодированных с использованием языка HTML (Hypertext Markup Language — язык гипертекстовой разметки). Для преобразования Web-страницы в формат WML, приемлемый для беспроводных устройств, нужно удалить значительный объем информации (особенно, графику и анимацию). WML представляет преимущественно текстовую информацию, пытаясь отразить суть Web-страницы и максимально облегчить доступ к ней с мобильных устройств.

Стоит выделить следующие важные особенности WML.

- **Поддержка текста и графики.** Команды форматирования и размещения предоставляют возможность отображения текста и ограниченные графические возможности.
- **Модель “колода/карта”.** Документы WML делятся на маленькие, однозначные пользовательские модули, называемые *картами*, между которыми перемещаются пользователи. Карта определяет один или несколько блоков взаимодействия (меню, экран текста или поле ввода текста). Колода WML подобна странице HTML в том, что она определяется Web-адресом (URL) и является единицей передачи содержимого.
- **Поддержка перемещения между картами и колодами.** WML предоставляет возможности обработки событий, которые используются для навигации или выполнения сценариев.

В Web-браузере на основе HTML пользователь перемещается между страницами, щелкая на ссылках. В мобильном устройстве с возможностями WML пользователь взаимодействует с картами, перемещаясь вперед и назад по колоде.

WML — это подобный HTML язык тегов, в котором отдельные элементы языка обособляются взятыми в скобки тегами, состоящими из строчных букв. Как правило, определение карт WML начинается с неотображаемой части, где содержатся выполняемые элементы, за которой следует видимое содержимое. Как пример рассмотрим следующую простую колоду с одной картой (пример взят из [MANN99]):

```
<wml>
  <card id='card1'>
    <p>
      Hello WAP World.
    </p>
  </card>
</wml>
```

Теги `<wml>`, `<card>` и `<p>` замыкают колоду, карту и абзац, соответственно. Подобно HTML большинство элементов заканчиваются закрывающим тегом, который идентичен открывающему с точностью до знака “/”. Когда беспроводное устройство получает такой код, оно выводит на экран терминала сообщение “Hello WAP World.”.

Все теги WML перечислены в табл. 12.3, где они разбиты на восемь функциональных групп.

**Таблица. 12.3. Теги WML**

<b>Метка</b>	<b>Описание</b>
<b>Структура колоды</b>	
<access>	Управление доступом
<card>	Определение карты
<head>	Информация на уровне колоды (метаинформация, доступ, шаблон)
<meta>	Метаинформация
<template>	Привязки событий на уровне колоды
<wml>	Определение колоды
<b>Содержимое</b>	
<img>	Изображение
<p>	Абзац, видимое содержимое
<table>	Таблица
<td>	Данные таблицы
<tr>	Строка таблицы
<b>Форматирование</b>	
<b>	Полужирный
<big>	Большой шрифт
 	Разрыв строки
<em>	Выделение
<i>	Курсив
<small>	Малый шрифт
<strong>	Выделенный шрифт
<u>	Подчеркивание
<b>Пользовательский ввод</b>	
<fieldset>	Группировка информационных элементов
<input>	Информационный элемент
<optgroup>	Подмножество списка позиций
<option>	Одна позиция списка
<select>	Список позиций
<b>Переменные</b>	
<postfield>	Устанавливает переменную запроса http
<setvar>	Устанавливает переменную задания
<b>Таймеры</b>	
<Timer>	Устанавливает таймер
<b>Задания</b>	
<go>	Переход по адресу URL
<noop>	Отсутствие действия
<prev>	Переход к предыдущей карте
<refresh>	Обновление изображения на экране
<b>Привязка событий/заданий</b>	
<a>	Сокращенный анкер
<anchor>	Анкер
<do>	Реакция на нажатие пользователем кнопки
<onevent>	Внутренняя привязка событий



## WMLScript

WMLScript — язык подготовки сценариев, подобный JavaScript. Он предназначен для определения сценариев на устройствах с ограниченными вычислительными возможностями и памятью. Возможные операторы WMLScript показаны в табл. 12.4. WMLScript включает:

- проверку достоверности пользовательского ввода перед отправкой его серверу содержимого;
- средства доступа к устройствам (в том числе периферийным);
- взаимодействие с пользователем без обращения к исходному серверу (например, отображение сообщения об ошибке).

Таблица 12.4. Операторы WMLScript

Оператор	Описание
=	Присваивание
break	Прекращение текущего цикла
continue	Повторение текущего цикла
for	Индексированный цикл
function	Объявление функции
if ... else	Проверка условия
return	Выход из текущей функции
var	Объявление переменной
while	Цикл, управляемый булевой переменной

Среди основных особенностей WMLScript стоит выделить следующие [WAP98].

- **Язык подготовки сценариев, основанный на JavaScript.** WMLScript — это сокращенный вариант JavaScript с некоторыми дополнениями.
- **Процедурная логика.** WMLScript вносит возможности процедурной логики в рассмотренную ниже среду беспроводных приложений (Wireless Application Environment — WAE).
- **Основанный на событиях.** WMLScript может активизироваться в ответ на запрос пользователя или некоторое событие в среде.
- **Компилируемая реализация.** WMLScript может компилироваться в более эффективный байт-код, который передается клиенту.
- **Интегрированный в WAE.** WMLScript интегрирован с WML-броузером. Это позволяет создавать службы с использованием как WML, так и WMLScript.
- **Эффективная поддержка расширяемых библиотек.** WMLScript может использоваться для расширения функциональных возможностей устройств без изменений в программном обеспечении устройства.

## Среда беспроводных приложений

Среда беспроводных приложений (Wireless Applications Environment — WAE) устанавливает прикладную основу для беспроводных устройств, подобных мо-

бильным телефонам, пейджером и PDA. По сути, WAE состоит из инструментов и форматов, предназначенных для облегчения задачи разработки приложений и устройств, поддерживаемых WAP. Основными элементами модели WAE являются следующие (рис. 12.11).

- **Пользовательские агенты WAE.** Программное обеспечение, выполняемое на пользовательском беспроводном устройстве, которое дает конечному пользователю определенные функциональные возможности (например, отображение содержимого на дисплее).
- **Генераторы содержимого.** Приложения (или услуги) на исходных серверах (например, сценарии CGI), которые в ответ на запросы пользовательских агентов с мобильных терминалов выдают содержимое в стандартном формате. WAE не задает каких-либо стандартных генераторов содержимого, но при этом предполагается, что на существующих HTTP-серверах, широко используемых в WWW, таких генераторов достаточно.
- **Стандартное кодирование содержимого.** Определено для удобства навигации пользовательского агента WAE по Web-содержимому.
- **Приложения беспроводной телефонии (Wireless Telephony Application — WTA).** Совокупность связанных с телефонией дополнений (механизмы управления вызовами и функциями), которые предлагают усовершенствованные услуги мобильных сетей. С помощью WTA разработчики приложений могут использовать микробраузер для организации телефонных звонков и реагирования на события в телефонной сети.

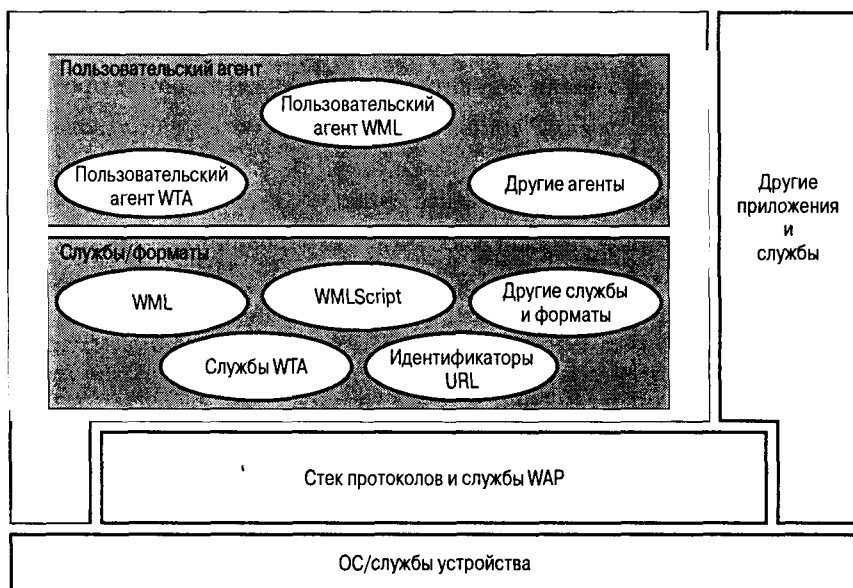


Рис. 12.11. Компоненты клиента WAE [WAPF98]

## Беспроводный сеансовый протокол

Протокол WSP (Wireless Session Protocol) предоставляет приложениям интерфейсы к двум сеансовым службам. Сеансовая служба на основе соединения функционирует над надежным транспортным протоколом WTP, а сеансовая служба без установления соединения используется над ненадежным транспортным протоколом WDP. По сути, протокол WSP основан на HTTP, но имеет некоторые дополнения и модификации, направленные на оптимизацию его применения в беспроводных каналах. Данный протокол противостоит двум основным ограничениям таких каналов: низкой скорости передачи данных и большой вероятности потери соединения вследствие недостаточного охвата или перегрузки ячейки.

WSP — это протокол на основе транзакций, оперирующий по схеме запрос/ответ. Каждый модуль данных протокола (PDU) WSP состоит из тела, которое может содержать данные WML, WMLScript или изображения, и заголовка, который содержит информацию о данных в теле и о транзакции. В WSP также определена операция незапрашиваемой поставки данных, когда сервер передает клиентскому устройству содержимое, на которое не был получен запрос. Эта возможность может использоваться для передачи широковещательных сообщений или в услугах, подобных предоставлению клиентскому устройству заголовков новостей или биржевых котировок.

### Служба WSP

В общем случае режим WSP на основе соединения предлагает следующие услуги.

- Установление надежного сеанса связи клиент/сервер и разрыв его обычным способом.
- Согласование общего уровня функциональных возможностей протокола.
- Обмен содержимого между клиентом и сервером с использованием компактного кодирования.
- Приостановка и возобновление сеанса связи.
- Несинхронизированная поставка клиенту незапрошенных данных.

На уровне служб WSP определяется через набор примитивов служб с соответствующими параметрами. Эти примитивы служб определяют интерфейс между WSP и пользователями WSP в WAE<sup>1</sup>. На уровне протокола спецификация WSP определяет формат PDU, используемый для обмена данными между одноранговыми объектами WSP (см. рис. 4.5 и 4.6).

На рис. 12.12 показаны ключевые типы транзакций WSP, выраженные через обмен примитивами и параметрами. Существуют и другие типы транзакции, но для понимания работы WSP достаточно приведенных.

Установление сеанса включает обмен примитивами S-Connect (рис. 12.12, а). Пользователь WSP действует как клиент (при транзакции клиент — это мобильный узел), запрашивающий сеанс у пользователя WSP, действующего как сервер (Web-сервер) в удаленной системе, для чего серверу WSP передается примитив S-Connect.req. Запрос сопровождаются четыре параметра.

---

<sup>1</sup>Подробнее о концепции примитивов службы и параметрах см. в приложении 12В.

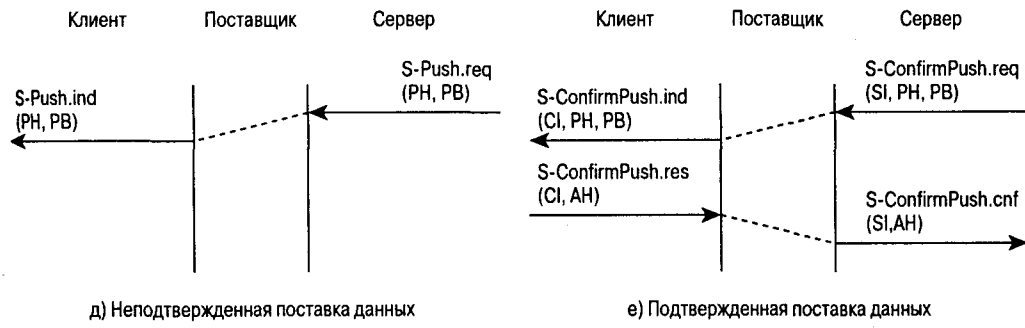
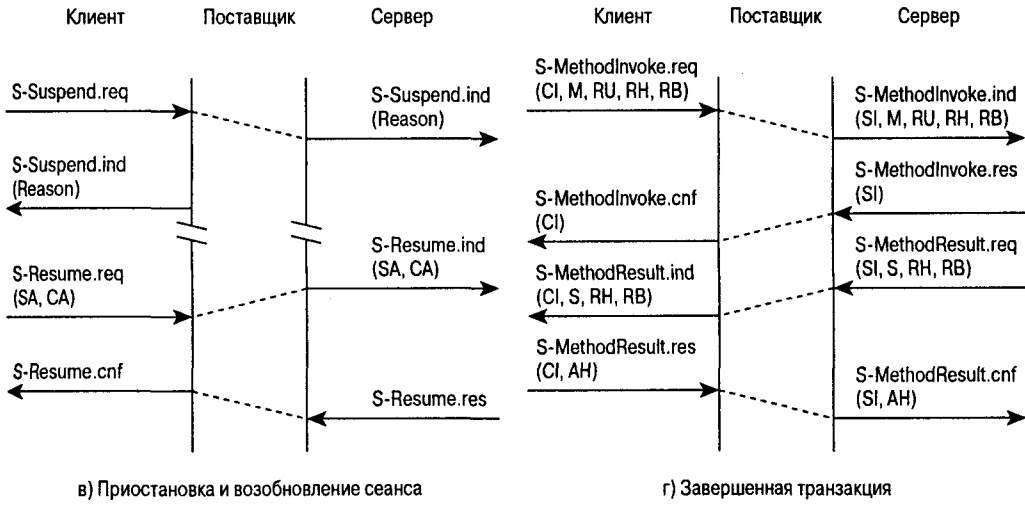
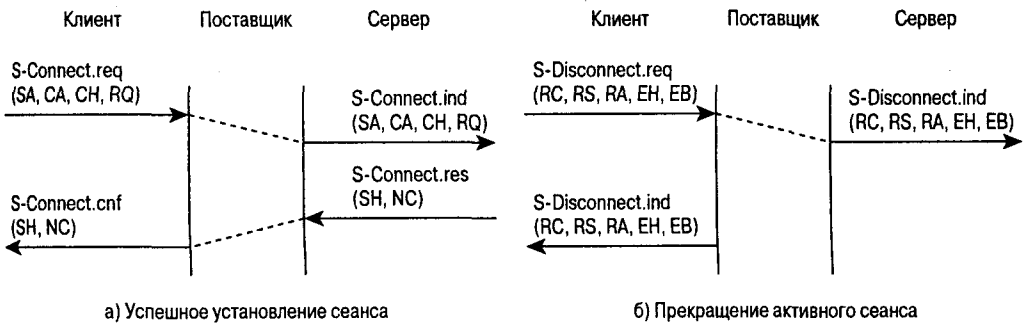


Рис. 12.12. Примитивы и параметры WSP

- **Адрес сервера (SA).** Одноранговый уровень, с которым устанавливается сеанс.
- **Адрес клиента (CA).** Инициатор сеанса.
- **Заголовки клиента (CH).** Содержат атрибутивную информацию, которая может использоваться для передачи между объектами параметров уровня приложений. Эта информация передается WSP без изменений и не обрабатывается WSP.

- **Запрашиваемые возможности (RQ).** Набор возможностей, запрашиваемых клиентом для данного сеанса; перечислены в табл. 12.5.

**Таблица 12.5. Возможности беспроводного сеансового протокола**

Название	Класс	Тип	Описание
Псевдонимы	I	Список адресов	Указывает, какие альтернативные адреса может использовать одноранговый пользователь для доступа к данному пользователю сеансовой службы. Может использоваться для облегчения переключения на новую службу-носитель при возобновлении сеанса
Размер SDU клиента	N	Положительное целое число	Размер самого большого модуля данных службы транзакции, который может пересылаться клиенту в течение сеанса
Расширенные методы	N	Набор имен методов	Набор расширенных (относительно HTTP/1.1) методов, которые поддерживаются одноранговыми объектами клиента и сервера
Кодовые страницы заголовка	N	Набор имен кодовых страниц	Множество расширенных кодовых страниц заголовка, которые поддерживаются одноранговыми объектами клиента и сервера
Максимальное число текущих запросов методов	N	Положительное целое число	Максимальное число вызовов метода, которые могут быть одновременно активны на протяжении сеанса
Максимальное число текущих запросов передачи	N	Положительное целое число	Максимальное количество вызовов подтвержденной поставки данных, которые могут быть одновременно активны на протяжении сеанса
Опции протокола	N	Набор средств и свойств	Это могут быть заголовки поставки данных, подтвержденной поставки, возобновления сеанса и подтверждения
Размер SDU сервера	N	Положительное целое число	Размер самого большого модуля данных службы транзакции, который может пересылаться серверу в течение сеанса

I — информационный класс

N — класс договорных возможностей

Далее WSP клиента подготавливает PDU WSP (который содержит указанные параметры) для передачи запроса одноранговому объекту WSP на сервере. Адрес сервера, адрес клиента и заголовки клиента при этом не меняются. В то же время поставщик службы WSP на стороне клиента, поставщик службы WSP на сервере или оба могут изменять набор запрошенных возможностей с тем, чтобы они не требовали от поставщика большего, чем он может предоставить. Далее примитив S-Connect.ind (возможно, модифицированный, как сказано выше, но с теми же параметрами, что указаны в запросе) передается пользователю WSP на стороне сервера. Если пользователь WSP на стороне сервера принимает запрос сеанса, он отвечает путем активизации WSP посредством примитива S-Connect.rsp, содержащего заголовки сервера и параметр договорных возможностей, причем последний является необязательным. При отсутствии параметра договорных возможностей пользователь WSP соглашается с набором возможно-

стей, предложенных в S-Connect.ind. Если этот параметр есть, он отражает уровень функциональных возможностей, приемлемый для пользователя WSP. Наконец, примитив S-Connect.ind доставляется исходному запрашивающему устройству, а с ним прибывают заголовки сервера и окончательный набор договорных возможностей.

Для завершения сеанса используется примитив S-Disconnect. На рис. 12.12, б показано, как завершение инициирует клиентский пользователь WSP. Примитив запроса содержит следующие параметры.

- **Код причины (RC).** Причина разъединения. Если разрыв соединения вызван перенаправлением клиента на другой сервер, два следующих параметра являются обязательными.
- **Перенаправить защиту (RS).** Указывает, будет ли клиент использовать при перенаправлении текущий безопасный сеанс.
- **Перенаправить адрес (RA).** Альтернативные адреса, которые будут использованы для установления сеанса.
- **Заголовки (EH) и тело ошибок (EB).** Если разъединение вызвано ошибкой, то эти параметры могут использоваться для предоставления серверному пользователю WSP информации об ошибке.

WSP подтверждает прием запроса, возвращая примитив S-Disconnect.ind клиентскому и серверному пользователям WSP.

WSP поддерживает приостановки и возобновление сеанса. Это может использоваться, например, если клиент знает, что он может временно стать недоступным (вследствие разрыва соединения с последующим его восстановлением или роуминга). При приостановке состояние сеанса сохраняется и клиентом, и сервером, при этом теряются любые передающиеся данные. Последовательность действий при приостановке и возобновлении сеанса показана на рис. 12.12, в (инициатором является клиент). При приостановке сеанса единственный параметр в примитиве S-Suspend.ind — код причины. При передаче запроса на возобновление в примитивах запроса и индикации указываются адреса сервера (SA) и клиента (CA).

**Транзакция** включает обмен данными между клиентом и сервером с использованием примитивов S-MethodInvoke и S-MethodResult (рис. 12.12, г). S-MethodInvoke используется для запроса выполнения операции на сервере. Запрос содержит перечисленные ниже параметры.

- **Идентификатор транзакции клиента (CI).** Позволяет отличать незаконченные транзакции.
- **Метод (M).** Идентифицирует запрошенную операцию.
- **URI (Uniform Resource Identifier — универсальный идентификатор ресурса) запроса (RU).** Определяет, к какому объекту будет применена операция.
- **Заголовки (RH) и тело запроса (RB).** Атрибутная информация и данные, связанные с запросом.

Такие же параметры содержит примитив индикации, передаваемый серверу, только в нем идентификатор транзакции клиента заменяется идентификатором транзакции сервера (SI). Указанные на рисунке примитивы ответа и подтверждения используются для подтверждения того, что запрос был доставлен по назначению и содержал идентификатор транзакции.

В ответ на запрос операции передается примитив S-MethodResult. Запрос, переданный сервером, включает идентификатор транзакции сервера (SI), статус этого ответа (S) и заголовки (RH) и тело (RB) ответа, которые содержат атрибутивную информацию и данные, связанные с ответом. Примитивы ответа и подтверждения используются для подтверждения того, что запрос был доставлен по назначению и содержит идентификатор транзакции. Эти два примитива могут также содержать заголовки подтверждения (AH), используемые для возврата серверу некоторой информации.

**Неподтвержденная поставка данных** используется для отправки незатребованной информации от сервера клиенту (рис. 12.12, *д*). С этими примитивами связаны только параметры заголовка (PH) и тела (PB) поставки данных, которые содержат атрибуты и передаваемую информацию. При **подтвержденной поставке данных** (рис. 12.12, *е*) сервер получает подтверждение (AH) того, что данные достигли адресата. В дополнение к заголовкам и телу поставки данных примитив подтвержденной поставки содержит идентификатор поставки; примитивы ответа и подтверждения также могут содержать заголовки подтверждения (AH).

Сеансовая служба без организации соединения дает пользователям WSP возможность обмениваться содержимым без подтверждения. При этом доступны только средства вызова метода и поставки незатребованных данных.

## Модули данных протокола WSP

WSP передает запросы и ответы служб в модулях данных протокола (PDU) WSP. Каждый PDU опускается на транспортный уровень и включается в PDU транспортного уровня как тело. На высшем уровне PDU WSP состоит из трех полей. Поле идентификатора транзакции (TID) используется для соотнесения запросов с ответами в службе WSP без организации соединения и отсутствует в службе на основе соединений. Поле **Тип** определяет тип и функцию PDU и по сути соответствует типу примитива службы, который передается WSP. Наконец, поле **Содержимого, зависящего от типа**, содержит всю информацию, передаваемую в ответ на примитив WSP.

## Протокол беспроводных транзакций

Протокол беспроводных транзакций (Wireless Transaction Protocol — WTP) управляет транзакциями путем передачи запросов и ответов между пользовательским агентом (таким, как браузер WAP) и сервером приложений, при этом производится быстрый просмотр, транзакции электронной коммерции и другие действия. WTP предоставляет надежную транспортную службу, но обходится без многих служебных издержек TCP, что дает упрощенный протокол, удобный для реализации на “тонких” клиентах (например, мобильных узлах) и использования в беспроводных каналах с малой пропускной способностью. Перечислим отличительные особенности WTP.

- Наличие трех классов служб транзакции.
- Возможность надежной передачи данных между пользователями: пользователь WTP передает подтверждение каждого полученного сообщения.
- Возможность передачи в подтверждении дополнительных данных.
- Объединение модулей PDU и задержанных подтверждений с целью уменьшения общего числа отправленных сообщений.
- Использование асинхронных транзакций.

WTP — это протокол на основе транзакции, а не на основе соединения. Не существует явного установления или разрыва соединения, вместо этого имеется надежная служба без установления соединения.

### **Классы транзакций WTP**

WTP имеет три класса транзакций, которые могут инициироваться WSP или другим протоколом более высокого уровня.

- **Класс 0.** ненадежное сообщение инициации без результирующего сообщения.
- **Класс 1.** надежное сообщение инициации без результирующего сообщения.
- **Класс 2.** ненадежное сообщение инициации с одним надежным результирующим сообщением.

**Класс 0** предлагает ненадежную службу дейтаграмм, которая может использоваться для ненадежной операции незатребованной поставки данных. WTP (инициатор, или клиент) инкапсулирует данные пользователя WTP в PDU инициации и передает целевому WTP (отвечающая сторона, или сервер) без подтверждения. Ответчик WTP передает данные целевому пользователю WTP.

**Класс 1** предлагает надежную службу дейтаграмм, которая может использоваться для надежной операции незатребованной поставки данных. Данные инициатора инкапсулируются в PDU инициации и передаются отвечающей станции. Она передает данные целевому пользователю WTP и подтверждает прием данных, возвращая PDU подтверждения (ACK) объекту WTP со стороны инициатора, который подтверждает транзакцию исходному пользователю WTP. Отвечающая станция WTP хранит информацию о состоянии еще некоторое время после передачи ACK, поскольку подтверждение придется передавать повторно, если, например, оно будет утеряно или клиент повторно передаст PDU инициации.

**Класс 2** предлагает службу транзакций типа запрос/ответ и поддерживает выполнение множественных транзакций за один сеанс WSP. Данные клиента инкапсулируются в PDU инициации и передаются серверу, который передает данные целевому пользователю WTP. Этот пользователь подготавливает ответные данные, которые опускаются локальному объекту WTP. Отвечающий объект WTP возвращает эти данные в результирующем PDU. Если на подготовку ответных данных требуется время, превышающее определенное время ожидания, ответчик может отправить PDU подтверждения до отправления результирующего PDU. Это позволяет избежать передачи повторного сообщения инициации.

### **Форматы и функционирование протокола**

В WTP используется шесть типов PDU. Каждый PDU начинается с заглавной части фиксированного размера (рис. 12.13), за которой может следовать заголовок переменного размера, который содержит дополнительную управляющую информацию. Формат этой дополнительной информации — один или несколько элементов протокола транзакций (transaction protocol item — TPI).



CON	Тип PDU — активизация	GTR	TTR	RID
Идентификатор транзакции (TID)				
Версия	TIDnew	U/P	Зарезервировано	TCL

а) PDU активизации

CON	Тип PDU — результат	GTR	TTR	RID
Идентификатор транзакции (TID)				

б) PDU подтверждения

CON	Тип PDU	GTR	TTR	RID
Идентификатор транзакции (TID)				
Порядковый номер пакета				

в) PDU результата

CON	Тип PDU — подтверждение	Tve/Tok	rsv	RID
Идентификатор транзакции (TID)				

г) PDU разрыва

CON	Тип PDU — разрыв	Тип разрыва
Идентификатор транзакции (TID)		
Причина разрыва		

д) Сегментированный PDU активизации или результата

CON	Тип PDU — отрицательное подтверждение	Зарезервировано	RID
Идентификатор транзакции (TID)			
Число потерянных пакетов $N$			
5	Порядковые номера пропущенных пакетов		
$4 + N$			

е) PDU отрицательного подтверждения

CON — флаг продолжения  
 GTR — завершитель группы  
 TTR — завершитель передачи  
 RID — индикатор повторной передачи  
 TIDnew — флаг TIDnew  
 U/P — флаг подтверждения пользователя/ поставщика  
 TCL — класс транзакции  
 Tve/Tok — флаг подтверждения TID/TID OK

Рис. 12.13. Форматы заголовков PDU WTP фиксированного размера

PDU инициации используется для передачи запроса от клиента осерверу; этот 4-байтовый модуль содержит следующие поля.

- **Флаг продолжения.** Если установлен этот флаг, за данным заголовком фиксированного размера следует один или несколько TPI. В свою очередь каждый TPI также начинается с бита продолжения, посредством которого указывается, является ли данный TPI последним, или за ним следуют другие.
- **Тип PDU.** Показывает, что это PDU инициации.
- **Флаг группового завершителя.** Используется при сегментации и повторной сборке (объясняется ниже).
- **Флаг завершителя передачи.** Используется при сегментации и повторной сборке.
- **Идентификатор повторной передачи.** Показывает, является ли данный элемент повторной передачей. Если в течение определенного времени инициатор не получит подтверждения, он снова передаст PDU инициации.
- **Идентификатор транзакции.** Используется для соотнесения PDU с конкретной транзакцией.
- **Версия.** Версия WTP.
- **Флаг TIDnew.** Устанавливается, если инициатор переходит на новую последовательность идентификаторов TID; т.е. значение следующего TID будет меньше текущего.
- **Флаг U/P.** Если этот флаг установлен, он показывает, что клиенту требуется пользовательское уведомление от серверного пользователя WTP. Это означает, что пользователь WTP подтверждает каждое принятое сообщение. Если значение флага не установлено, отвечающий объект WTP может подтверждать поступающий PDU без подтверждения от пользователя.
- **Класс транзакции.** Показывает класс транзакции, требуемый для обработки данного PDU.

Если сообщение (блок данных от WSP), которое передает WTP, слишком велико для текущей службы-носителя, WTP может сегментировать это сообщение и передать его в нескольких пакетах, по одному на каждый PDU инициации. При передаче сообщения в многочисленных маленьких пакетах их отправка и подтверждение получения могут осуществляться группами. В табл. 12.6 показано, как для управления этим процессом используются флаги GTR и TTR.

**Таблица 12.6. Комбинации группового завершителя (Group Trailer — GTR) и завершителя передачи (Transmission Trailer — TTR)**

GTR	TTR	Описание
0	0	Не последний пакет
0	1	Последний пакет сообщения
1	0	Последний пакет группы пакетов
1	1	Сегментация и повторная сборка не поддерживаются

PDU подтверждения (ACK) используется для подтверждения результирующего PDU или PDU инициации. Этот 3-битовый PDU содержит флаг Tve/Tok, интерпретация которого зависит от направления PDU. В направлении от ответчика к инициатору — это флаг Tve, который интерпретируется как: “Есть ли у тебя текущая тран-

закция с данным идентификатором TID?”. В обратном направлении — это флаг Tok, который означает: “У меня есть текущая транзакция с данным TID”.

**Результирующий PDU** — это 3-байтовый PDU, который используется для передачи клиенту ответа сервера.

**PDU разрыва** используется для прекращения транзакции. Существует два типа разрыва: инициированный пользователем и поставщиком. Если разрыв инициируется пользователем WTP (например, WSP), то причина разрыва указывается в теле PDU и передается пользователю WTP в пункт назначения PDU разрыва. Если разрыв инициируется поставщиком WTP (объектом WTP, который отправляет данный PDU разрыва), то в поле причины разрыва указывается одна из следующих причин.

- **Неизвестна.** Ошибка не уточняется.
- **Ошибка протокола.** Принятый PDU не поддается интерпретации.
- **Неверный TID.** Сбой при верификации TID.
- **Класс 2 не реализован.** Сервер не поддерживает запрошенной службы класса 2.
- **Сегментация и повторная сборка не реализованы.** Сервер не поддерживает сегментацию и повторную сборку.
- **Пользовательское подтверждение не реализовано.** Сервер не поддерживает пользовательское подтверждение.
- **WTP версии 1.** Инициатор запросил не поддерживаемую версию WTP; текущая версия — 1.
- **Пропускная способность временно превышена.** Вследствие переполнения транзакция не может быть завершена.

Для организации сегментации и повторной сборки могут использоваться сегментированный PDU инициации и сегментированный результирующий PDU. В этом случае каждый пакет последовательно нумеруется. Если один или несколько пакетов последовательности не прибыли, используется PDU отрицательного подтверждения.

На рис. 12.14 показаны некоторые простые примеры временного упорядочения PDU WTP для указанных трех классов операций.

## Служба WTP

Служба WTP определяется тремя примитивами. TR-Invoke используется для инициации новой транзакции. TR-Result используется для возврата результата ранее инициированной транзакции. Tr-Abort используется для разрыва имеющейся транзакции. На рис. 12.15 показана связь между этими примитивами и транзакцией класса 2 с удержанием подтверждения. Сравните с рис. 12.14, г.

На рис. 12.15 также показано, как служба WTP поддерживает службу WSP. Последовательность примитивов службы WSP предусматривает завершенную транзакцию и идентична последовательности, показанной на рис. 12.12, г.

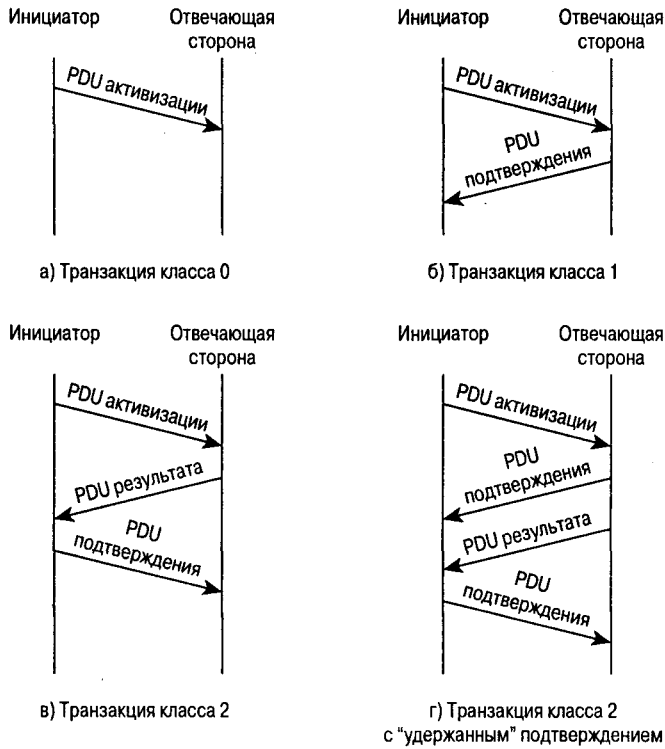


Рис. 12.14. Примеры операций WTP

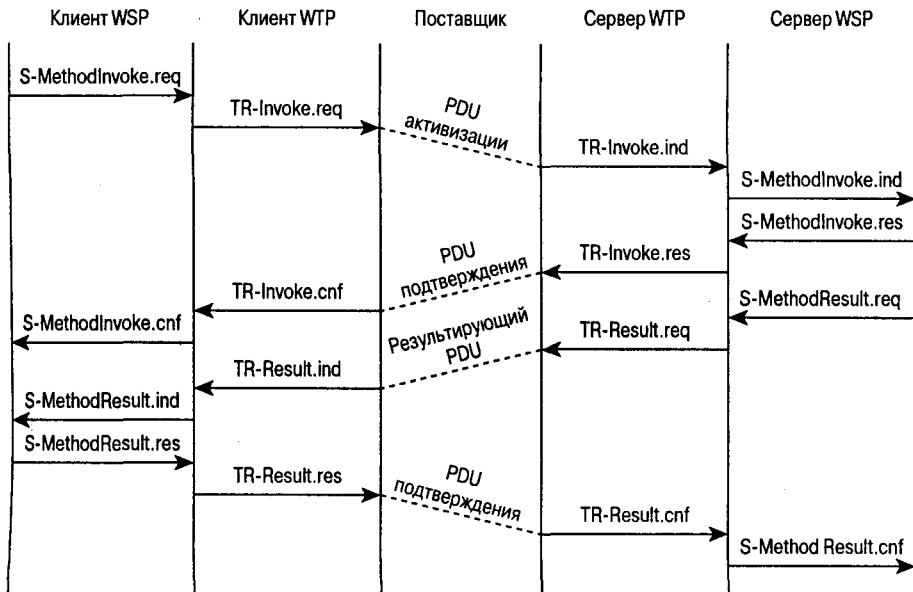


Рис. 12.15. Временная диаграмма WSP-WTP

## Протокол безопасности беспроводного транспортного уровня

Протокол WTLS (Wireless Transport Layer Security) предлагает безопасные службы между мобильным устройством (клиентом) и шлюзом WAP. WTLS основан на промышленном стандарте TLS (Transport Layer Security — протокол безопасности транспортного уровня), который является переработанной версией протокола защищенных сокетов (Secure Sockets Layer — SSL). TLS — это стандартный протокол безопасности, использующийся между Web-браузерами и Web-серверами<sup>2</sup>. WTLS эффективнее TLS, требует меньшего числа передач сообщений. Для обеспечения сквозной безопасности WTLS используется между клиентом и шлюзом, а TLS — между шлюзом и целевым сервером. Системы WAP выполняют переход между WTLS и TLS на шлюзе WAP. Таким образом, шлюз — это уязвимая точка и ее следует хорошо защитить от внешних атак.

WTLS обеспечивает следующее.

- **Целостность данных.** Данные, передаваемые между клиентом и шлюзом, не меняются; для гарантии этого используется аутентификация сообщения.
- **Конфиденциальность.** Шифрование позволяет защитить данные от чтения третьей стороной.
- **Аутентификация.** Аутентификация двух сторон с использованием цифровых сертификатов.
- **Защита типа “отказ в обслуживании”.** Дублирующиеся сообщения или сообщения, не прошедшие верификацию, обнаруживаются и отклоняются.

WTLS — это не единый протокол, а два уровня протоколов, как показано на рис. 12.16. Протокол записи WTLS предлагает базовый набор средств защиты, применяемых протоколами более высоких уровней. В частности, протокол передачи гипертекстовых файлов (Hypertext Transfer Protocol — HTTP), определенный в RFC 2068 и предлагающий службу передачи для взаимодействий Web-клиент/сервер, может использоваться над уровнем WTLS. Частью WTLS считаются три протокола высокого уровня: протокол квитирования (Handshake Protocol), протокол изменения параметров шифрования (Change Cipher Spec Protocol) и протокол извещения (Alert Protocol).

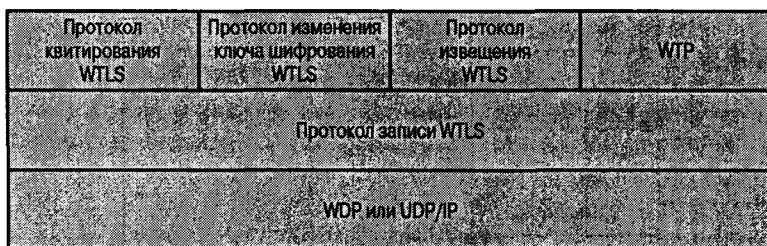


Рис. 12.16. Стек протоколов WTLS

<sup>2</sup> Более подробно о TLS см. в [STAL99].

## Протокол записи WTLS

Протокол записи WTLS берет данные приложения от приложения более высокого уровня WTLS (WTP, протокола квитирования WTLS, протокола извещения WTLS, протокола изменения параметров шифрования WTLS) и инкапсулирует эти данные в PDU. Проследим, как это происходит (рис. 12.17).

1. Полезная нагрузка сжимается с использованием алгоритма сжатия без потерь.
2. Для сжатых данных вычисляется код аутентификации сообщения (Message Authentication Code — MAC), для чего используется код HMAC. HMAC — это хэш-код с ключом, подобный тем (но сложнее их), что описаны в приложении 12Б. С кодом HMAC может использоваться один из многих хэш-алгоритмов, например MD-5 и SHA-1. Длина хэш-кода равна 0, 5 или 10 байт. Поле MAC следует за полем сжатых данных.
3. Сжатое сообщение плюс код MAC шифруется с использованием симметричного алгоритма шифрования. На этом этапе могут применяться схемы DES, RC5 и IDEA.
4. Протокол записи вводит полученный заголовок перед зашифрованной полезной нагрузкой.

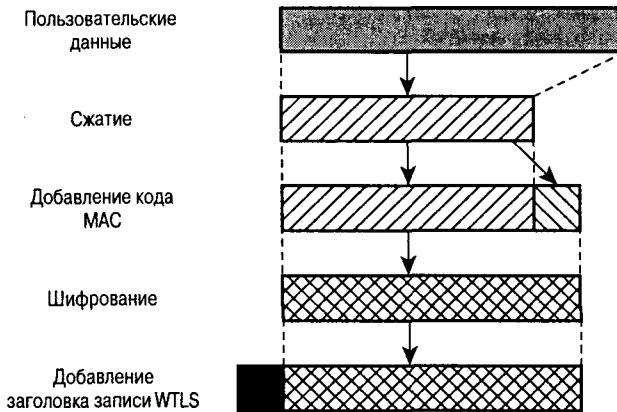
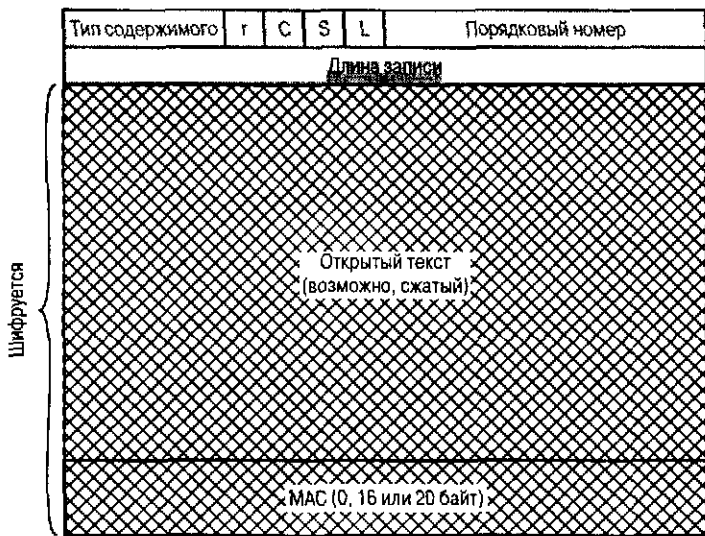


Рис. 12.17. Операция протокола записи WTLS

Заголовок протокола записи состоит из следующих полей (рис. 12.18).

- **Тип содержимого (8 бит).** Протокол более высокого уровня над протоколом записи WTLS.
- **Индикатор параметров шифрования (1 бит).** 0 — сжатие, защита MAC и шифрование не используются.
- **Индикатор поля порядкового номера (1 бит).** Показывает, есть ли поле порядкового номера.
- **Индикатор поля длины записи (1 бит).** Показывает, есть ли поле длины записи.
- **Порядковый номер (16 бит).** Порядковый номер, соотношенный с PDU. Это обеспечивает надежность в ненадежных транспортных службах.
- **Длина записи (32 бит).** Длина в байтах данных открытого текста (или, если используется сжатие, сжатых данных).



- r — зарезервировано
- C — индикатор изменения параметров шифрования
- S — индикатор поля порядкового номера
- L — индикатор поля длины записи
- MAC — код аутентификации сообщения

Рис. 12.18. Формат записи WTLS

### Протокол изменения параметров шифрования

С текущей транзакцией соотнесены определенные параметры шифрования, которые задают алгоритм шифрования, хэш-алгоритм, который используется как часть HMAC, и криптографические параметры, такие, как размер кода MAC. Фактически с каждым сеансом можно сопоставить два состояния. Как только сеанс установлен, имеется текущее рабочее состояние для чтения и записи (т.е. приема и передачи). Кроме того, при работе протокола квитирования создаются состояния ожидания чтения и записи.

Протокол изменения параметров шифрования — это один из трех протоколов, использующих протокол записи WTLS, и он является простейшим. Этот протокол состоит из одного сообщения, которое содержит один бит со значением 1. Единственной задачей этого сообщения является указание начать копирование параметров состояния ожидания в текущее состояние, что приводит к обновлению шифров, используемых для данного соединения.

### Протокол извещения

Протокол извещения предназначается для передачи одноранговому объекту извещений, касающихся работы WTLS. Как и данные любого другого приложения, использующего WTLS, сообщения протокола извещения сжимаются и шифруются в соответствии с параметрами текущего состояния.

Любое сообщение, генерируемое данным протоколом, состоит из двух байтов. Первый байт содержит значение, обозначающее, соответственно, предупреждение (1), критическое сообщение (2) или неустранимую ошибку (3). Если указано наличие неустранимой ошибки, протокол WTLS немедленно разрывает со-

единение. Другие соединения данного сеанса могут продолжать существовать, но установить новое соединение для данного сеанса будет уже невозможно. Второй байт содержит код, обозначающий конкретный смысл извещения. Приведем примеры неустраняемых ошибок.

- **unexpected\_message.** Получено непригодное для обработки сообщение.
- **bad\_record\_mac.** Получено неправильное значение MAC.
- **decompression\_failure.** Функция распаковки сжатых данных получила не правильные входные данные (т.е. данные невозможно распаковать или длина получаемых в результате распаковки данных оказывается больше максимально допустимой).
- **handshake\_failure.** Отправитель не смог согласовать с получателем параметры защиты на основе имеющихся возможностей.
- **illegal\_parameter.** Значение поля в сообщении квитирования выходит за рамки допустимого диапазона или не согласуется со значениями других полей.

Кроме указанных, имеются следующие извещения.

- **bad\_certificate.** Полученный сертификат поврежден (т.е. содержит подпись, верификацию которой выполнить не удалось).
- **unsupported\_certificate.** Тип полученного сертификата не поддерживается.
- **certificate\_revoked.** Сертификат был отозван подписавшим его объектом.
- **certificate\_expired.** Срок годности сертификата истек.
- **certificate\_unknown.** При обработке сертификата возникли некоторые другие проблемы, вследствие чего его использование оказывается невозможным.

## Протокол квитирования

Самой сложной частью WTLS является протокол квитирования. Этот протокол позволяет серверу и клиенту выполнять взаимную аутентификацию и согласовывать алгоритмы шифрования, вычисления MAC и криптографические ключи, которые будут служить для защиты данных, пересылаемых в записи WTLS. Протокол квитирования должен использоваться перед началом пересылки данных приложения.

Протокол квитирования состоит из серии сообщений, которыми обмениваются клиент и сервер. На рис. 12.19 показана схема обмена, необходимая для установления логического соединения между клиентом и сервером. Процесс такого обмена состоит из четырех этапов.

**Этап первый** используется для инициализации логического соединения и для определения связанных с ним характеристик защиты. Процесс инициируется клиентом, отправляющим серверу сообщение `client_hello`, которое содержит идентификатор сеанса и перечень шифровальных алгоритмов и алгоритмов сжатия, поддерживаемых клиентом (в порядке уменьшения степени предпочтения). После отправки сообщения `client_hello`, клиент ждет ответное сообщение `server_hello`. Это сообщение показывает, какой алгоритм шифрования и сжатия будет использован при обмене.

**Этап второй** используется для аутентификации сервера и передачи ключей сервера. Если требуется аутентификация сервера, то данный этап начинается с



отправки сервером его сертификата открытого ключа. Затем при необходимости может быть отправлено сообщение `server_key_exchange`. Данное сообщение используется в некоторых алгоритмах шифрования с открытым ключом, применяемых для симметричного обмена ключами. Далее сервер может запросить сертификат открытого ключа клиента, используя сообщение `certificate_request`. Последним (необходимым) сообщением данного этапа является сообщение `server_hello_done`, которым сервер извещает о завершении фазы приветствия. После отправки этого сообщения сервер ожидает ответ клиента. Это сообщение параметров не имеет.

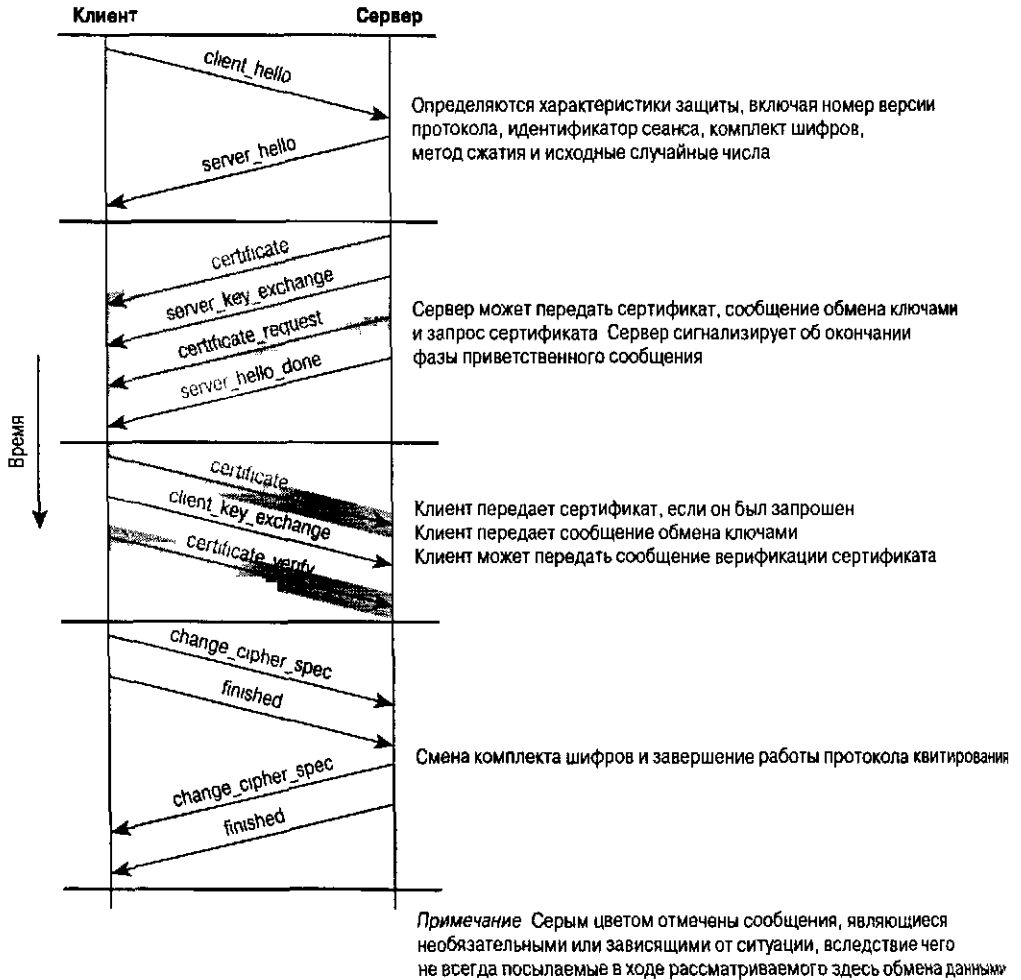


Рис. 12.19. Схема работы протокола квитирования *WTLS*

Этап третий используется для аутентификации клиента и передачи ключей клиента. Получив сообщение `server_hello_done`, клиент должен убедиться в том, что сервер предоставил действительный сертификат (если это требуется) и что параметры сообщения `server_hello` являются приемлемыми. Если проверка завершена успешно, клиент отправляет серверу одно или несколько сообщений.

Если сервер запрашивает сертификат, клиент отправляет сообщение `certificate`. Следующее необходимое сообщение — `client_key_exchange`. Содержимое сообщения зависит от выбранной схемы обмена ключами. В завершение данного этапа клиент может отправить сообщение `certificate_verify`, предоставляя явную верификацию сертификата клиента.

Этап четвертый завершает настройку защиты соединения. Клиент отправляет сообщение `change_cipher_spec` и копирует параметры `CipherSpec` состояния ожидания в поле `CipherSpec` текущего состояния. Отметим, что это сообщение не считается частью протокола квитирования, а отсылается в рамках протокола изменения параметров шифрования. В ответ клиент немедленно отправляет сообщение `finished`, шифрованное новыми алгоритмами с новыми ключами и секретными значениями. Сообщение `finished` подтверждает, что процессы обмена ключами и аутентификация завершились успешно. В ответ на эти два сообщения сервер посылает свое сообщение `change_cipher_spec`, копирует параметры состояния ожидания в текущее состояние и посылает свое сообщение `finished`. На этом процесс квитирования завершается, и теперь клиент и сервер могут начать обмен данными на уровне приложений.

## Беспроводный протокол дейтаграмм

Протокол WDP (`Wireless Datagram Protocol`) используется для адаптации вышестоящего протокола WAP к механизму связи (именуемому носителем), который используется между мобильным узлом и шлюзом WAP. Адаптация может включать сегментацию данных на фрагменты удобного (для носителя) размера и сопряжение с сетью носителя. WDP скрывает детали различных сетей носителей от других уровней WAP. В некоторых случаях протокол WAP реализуется над уровнем IP.

### Служба WDP

Служба WDP определяется двумя примитивами. Примитив `T-DUnitdata` предоставляет службу без подтверждения с такими параметрами.

- **Адрес источника.** Адрес устройства, которое производит запрос к уровню WDP.
- **Порт источника.** Адрес приложения, соответствующий адресу источника.
- **Адрес назначения.** Адресат WDP-данных.
- **Порт назначения.** Адрес приложения, соответствующий адресу назначения.
- **Данные пользователя.** Пользовательские данные с вышестоящего уровня, предоставляемые WDP для передачи порту назначения.

Примитив `T-Derror.ind` используется для оповещения пользователя WDP о бое при доставке дейтаграммы WDP. Кроме таких параметров, как адрес и порт источника и адрес и порт назначения, `T-Derror.ind` содержит параметр кода ошибки с указанием локальной значимости.

### Протокол WDP

В PDU протокола WDP необходимы следующие поля:

- порт назначения;
- порт источника.

Если базовый носитель не поддерживает сегментацию и повторную сборку, эта возможность реализуется в WDP так, чтобы не зависеть от носителя.

Например, в GSM (Global System for Mobile Communications — глобальная система мобильной связи) предписан обязательный формат заголовка пользовательских данных: заголовок состоит из последовательности элементов информации, а каждый элемент определяется идентификатором, длиной и одним или несколькими байтами значения. PDU WDP для GSM имеет следующую структуру.

- **Длина заголовка (1 байт).**
- **Идентификатор числа портов (1 байт).** Значение 5 показывает, что данный элемент информации состоит из двух номеров порта.
- **Длина чисел портов (1 байт).** Значение 4 показывает, что часть данного элемента информации, относящаяся к значению, имеет длину 4 байт.
- **Порт назначения (2 байт).**
- **Порт источника (2 байт).**
- **Идентификатор сегментации и повторной сборки (1 байт).** Значение 0 показывает, что данный элемент информации состоит из информации, касающейся сегментации и повторной сборки.
- **Длина сегментации и повторной сборки (1 байт).** Значение 3 показывает, что часть данного элемента информации, относящаяся к значению, имеет длину 3 байт.
- **Эталонный номер дейтаграммы (1 байт).** Идентификатор, присвоенный всем сегментам, которые составляют один блок пользовательских данных.
- **Число сегментов (1 байт).** Общее число сегментов, которые нужно собрать повторно.
- **Номер сегмента.** Порядковый номер, определяющий место данного сегмента в последовательности всех сегментов блока.
- **Данные пользователя (от 1 до  $n$  байт)**

### **Беспроводный протокол управляющих сообщений**

Протокол WCMP (Wireless Control Message Protocol) выполняет ту же функцию поддержки для WDP, что ICMP выполняет для IP (см. приложение 12А). WCMP используется в средах, которые не имеют носителя IP, следовательно, не могут сами пользоваться ICMP. WCMP используется беспроводными узлами и шлюзами WAP для уведомления о ошибках, случившихся при обработке дейтаграмм WDP. WCMP может также использоваться в информационных и диагностических целях.

В табл. 12.7 перечислены сообщения WCMP и соответствующие коды ошибок.

Таблица 12.7. Сообщения WCMP

Сообщение WCMP	Тип WCMP	Код WCMP	Узел WDP	Шлюз WAP
<b>Место назначения недоступно</b>	51			
• Нет маршрута к месту назначения		0	N/A	O
• Связь запрещена административно		1	N/A	O
• Адрес недоступен		3	N/A	O
• Порт недоступен		4	M	N/A
<b>Проблемы параметра</b>	54			
• Ошибочное поле заголовка		0	O	O
<b>Сообщение слишком большое</b>	60	0	M	N/A
<b>Сбой повторной сборки</b>	61			
• Превышено время повторной сборки		1	O	N/A
• Буфер переполнен		2	O	N/A
<b>Эхо-запрос</b>	178	0	O	N/A
<b>Ответ на эхо-запрос</b>	179	0	M	N/A

M — обязательно  
 O — не обязательно  
 N/A — не применяется

## 12.3. РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА И WEB-САЙТЫ

Хорошей обзорной статьей по Mobile IP является [PERK97]; коммерчески ориентированный труд того же автора — [PERK98]. Желая ознакомиться с данным вопросом еще подробнее можно порекомендовать работу [SOLO98]. Несколько статей о модернизации текущего стандарта Mobile IP содержит выпуск *IEEE Personal Communications* за август 2000 года.

[MANN99] — это обзор WAP, а в [PASS00] описано создание с использованием WAP приложений для беспроводной связи. Подробный обзор WAP представлен в [HEIJ00].

HEIJ00 Heijden M., Taylor M., eds. *Understanding WAP: Wireless Applications, Devices, and Services*. — Boston: Artech House, 2000.

MANN99 Mann S. The Wireless Application Protocol. — *Dr. Dobb's Journal*, October 1999.

PASS00 Passani L. Creating WAP Services. — *Dr. Dobb's Journal*, July 2000.

PERK97 Perkins C. Mobile IP. — *IEEE Communications Magazine*, May 1997.

PERK98 Perkins C. Mobile Networking Through Mobile IP. — *IEEE Internal Computing*, January-February 1998.

SOLO98 Solomon J. *Mobile IP: The Internet Unplugged*. — Upper Saddle River, NJ: Prentice Hall PTR, 1998.



## Рекомендуемые Web-сайты

- **IP Routing for Wireless/Mobile Hosts** ([ietf.org/html.charters/mobileip-charter.html](http://ietf.org/html.charters/mobileip-charter.html)). Рабочая группа IETF по Mobile IP. Содержит текущие публикации RFC и проекты стандартов Internet.
- **Wireless Application Protocol Forum** ([www.wapforum.org](http://www.wapforum.org)). Технические спецификации WAP и ссылки на производителей.

## 12.4. ТЕРМИНЫ, ВОПРОСЫ И ЗАДАЧИ

### Основные термины

Mobile IP	беспроводного	среда беспроводных
адрес передачи	транспортного	приложений
базовая сеть	уровня (WTLS)	(WAE)
базовый агент	протокол	туннелирование
базовый адрес	беспроводных	хэш-функция
беспроводный	приложений	чужая сеть
протокол	(WAP)	чужой агент
дейтаграмм (WDP)	протокол	язык разметки для
беспроводный	беспроводных	беспроводных
сеансовый	транзакций (WTP)	приложений
протокол (WSP)	протокол	(WML)
инкапсуляция	управляющих	
протокол	сообщений	
безопасности	Internet (ICMP)	

### Вопросы

1. Объясните разницу между мобильным пользователем и пользователем с **кочевым** доступом.
2. Что такое “туннелирование”?
3. Перечислите и кратко охарактеризуйте возможности, предлагаемые Mobile IP.
4. Какая связь между обнаружением Mobile IP и ICMP?
5. Назовите два разных типа адресов назначения, которые могут присваиваться **мобильному узлу, подключенному к чужой сети.**
6. В какой ситуации мобильный узел выбирает для использования каждый **адрес** из названных в вопросе 5?
7. Какая разница между фильтром HTML и прокси-сервером WAP?
8. В каких функциональных областях требуется поддержка WML?
9. Какие службы предлагает WSP?
10. При каких обстоятельствах используется каждый из трех классов **транзакций** WTP?
11. Перечислите и кратко охарактеризуйте службы безопасности, предлагаемые **WTLS.**

## Задачи

1. В этой задаче использован рис. 12.6. Пусть ЛВС Z — это базовая сеть для хоста E, а D посылает блок данных E посредством IP.
  - а. Покажите структуру PDU, включая поля заголовка IP и низкоуровневые заголовки (MAC, LLC) с содержанием адресных полей, если E находится в базовой сети.
  - б. Повторите пункт а для E, находящегося в чужой сети, достичь которой можно посредством Internet через R3. Покажите форматы кадра MAC, покидающего хост D, и дейтаграммы IP, покидающей хост R3. Считайте, что используется инкапсуляция IP-в-IP.
  - в. Повторите пункт б для дейтаграммы IP, покидающей R3, но на этот раз считайте, что имеет место минимальная инкапсуляция.
2. Снова обратимся к рис. 12.6; пусть A — мобильный узел, а ЛВС X — чужая сеть для A. Пусть дейтаграмма IP прибывает на хост R1 из Internet и должна проследовать к A. Покажите формат дейтаграммы IP, прибывшей на R1, и кадр MAC, покидающий R1 (включая заголовок или заголовки IP) для следующих случаев.
  - а. Используется инкапсуляция IP-в-IP; адрес передачи — R1.
  - б. Используется минимальная инкапсуляция; адрес передачи — R1.
  - в. Используется инкапсуляция IP-в-IP; адрес передачи — A.
  - г. Используется минимальная инкапсуляция; адрес передачи — A.
3. При типичной реализации Mobile IP на базовом агенте агент поддерживает таблицу мобильных привязок, требуемую для отображения базовых адресов мобильных узлов в их адреса передачи. Какие позиции важны в каждой строке таблицы?
4. При типичной реализации Mobile IP на чужом агенте агент поддерживает таблицу посетителей, в которой содержится информация о мобильных узлах, в настоящее время посещающих данную сеть. Какие позиции важны в каждой строке таблицы?
5. Почему бы в WTLS вместо отдельного протокола изменения параметров шифрования не ввести сообщение `change_cipher_spec` в протокол квитирования?

## ПРИЛОЖЕНИЕ 12А. ПРОТОКОЛ УПРАВЛЯЮЩИХ СООБЩЕНИЙ INTERNET

В стандарте IP задано, что приемлемая реализация этого протокола должна также определять протокол ICMP (RFC 792, RFC 950, RFC 1256). ICMP предлагает средства передачи данному хосту сообщений от маршрутизатора или других хостов. По сути, ICMP обеспечивает обратную связь по вопросам, касающимся проблем в среде связи. Пример такого использования: дейтаграмма, которая не может достичь адресата, если маршрутизатор не может занести пересылаемую дейтаграмму в буфер или может указать станции более короткий путь к адресату. В большинстве случаев в ответ на дейтаграмму посылается сообщение ICMP, что делает либо маршрутизатор, либо намеченный целевой хост.

ICMP — это пользователь IP. Сообщения ICMP создаются и опускаются на уровень IP, где оно инкапсулируется с заголовком IP и передается обычным образом. Поскольку сообщения ICMP передаются в дейтаграммах IP, их доставка не гарантируется, а использование не может считаться надежным.

На рис. 12.20 показан формат разных типов сообщений ICMP. Сообщение ICMP начинается с 64-битового заголовка, который включает такие поля.

- **Тип (8 бит).** Задаёт тип сообщения ICMP.
- **Код (8 бит).** Используется для задания параметров сообщения, которые могут кодироваться одним или несколькими битами.
- **Контрольная сумма (16 бит).** Контрольная сумма всего сообщения ICMP. Используется тот же алгоритм вычисления, что и в протоколе IP.
- **Параметры (32 бит).** Используется для установления более длинных параметров.

Тип	Код	Контрольная сумма
Не используется		
Заголовок IP и 64 бит исходной дейтаграммы		

а) Не доставлено, превышено время ожидания, источник выключен

Тип	Код	Контрольная сумма
Указатель	Не используется	
Заголовок IP и 64 бит исходной дейтаграммы		

б) Ошибки параметров

Тип	Код	Контрольная сумма
IP-адрес маршрутизатора		
Заголовок IP и 64 бит исходной дейтаграммы		

в) Перенаправлено

Тип	Код	Контрольная сумма
Число адресов	Размер адреса	Время жизни
Адрес маршрутизатора 1		
Уровень предпочтения 1		
⋮		
Адрес маршрутизатора n		
Уровень предпочтения n		

г) Объявление маршрутизатора

Тип	Код	Контрольная сумма
Не используется		

д) Запрос маршрутизатора

Тип	Код	Контрольная сумма
Идентификатор	Порядковый номер	
Временная метка отправления		

е) Запрос о временной метке

Тип	Код	Контрольная сумма
Идентификатор	Порядковый номер	
Временная метка отправления		
Временная метка получения		
Временная метка передачи		

ж) Ответ на запрос временной метки

Тип	Код	Контрольная сумма
Идентификатор	Порядковый номер	
Необязательные данные		

з) Эхо-запрос и ответ на него

Тип	Код	Контрольная сумма
Идентификатор	Порядковый номер	

и) Запрос об адресной маске

Тип	Код	Контрольная сумма
Идентификатор	Порядковый номер	
Адресная маска		

к) Ответ на запрос об адресной маске

Рис. 12.20. Форматы сообщений ICMP

Для некоторых типов сообщений за данными полями следуют дополнительные информационные поля, которые еще больше уточняют содержимое сообщения.

Если сообщение ICMP относится к предшествующей дейтаграмме, информационное поле содержит весь заголовок IP плюс первые 64 бит поля данных исходной дейтаграммы. Это позволяет исходному хосту сопоставить входящее сообщение ICMP с предыдущей дейтаграммой. 64 бит поля данных включаются потому, что это позволяет IP-модулю хоста определить, какие протоколы высших уровней используются. В частности, первые 64 бит могут содержать часть заголовка TCP или заголовка другого транспортного уровня.

Сообщение о **недостижимости адресата** появляется в нескольких случаях. Маршрутизатор может вернуть данное сообщение, если он не знает, как достичь сети получателя. Маршрутизаторы некоторых сетей могут определять, является ли определенный хост недостижимым, и возвращать указанное сообщение. Его может возвращать и сам хост-получатель, если протокол пользователя или точка доступа к службе какого-то более высокого уровня недоступна. Это может произойти, если соответствующее поле в заголовке IP было установлено неверно. Данное сообщение также возвращается, если дейтаграмма предписывает использовать маршрут, который в данный момент недостижим. Кроме того, если маршрутизатор должен фрагментировать дейтаграмму, но в ней установлен флаг “не фрагментировать”, дейтаграмма отклоняется и возвращается все то же сообщение.

Маршрутизатор будет возвращать сообщение о **превышении времени ожидания**, если истекло время жизни дейтаграммы. Хост будет посылать это сообщение, если он не может завершить повторную сборку в срок.

Синтаксическая или семантическая ошибка в заголовке IP может стать причиной сообщения об **ошибках параметров**, которое возвращается маршрутизатором или хостом. Например, в опции может быть указано неправильное значение. Поле параметра содержит указатель на октет исходного заголовка, где была обнаружена ошибка.

Сообщение **источник выключен** позволяет ввести элементарную форму управления потоком. Маршрутизатор или хост назначения могут посылать данное сообщение хосту источника, запрашивая уменьшение скорости передачи трафика к адресату. Получив такое сообщение, хост-отправитель должен уменьшить скорость, с которой он передает трафик указанному адресату, пока поступление сообщений “источник выключен” не прекратится. Сообщение “источник выключен” может использоваться маршрутизатором или хостом, вынужденным отклонять дейтаграммы из-за переполнения буфера. В этом случае маршрутизатор или хост возвращает сообщение для каждой дейтаграммы, которую он отклоняет. Кроме того, система может предупреждать перегрузку и выдавать такие сообщения, если ее буфера близки к переполнению. Однако в этом случае дейтаграмма, указанная в сообщении, может быть доставлена благополучно. Таким образом, получение сообщения “источник выключен” не обязательно означает, что соответствующая дейтаграмма не доставлена.

Сообщение **перенаправления** маршрутизатор отправляет хосту для уведомления последнего о наличии лучшего маршрута к месту назначения. Ниже приведен пример такой ситуации (см. рис. 12.6). Маршрутизатор R1 принимает дейтаграмму, предназначенную для D от хоста C в сети Y, к которой присоединен R1. R1 проверяет свою таблицу маршрутов и получает адрес следующего маршрутизатора, R2, по пути следования дейтаграммы к сети Z. Поскольку R2 и



хост, как определяется из их исходных адресов, находятся в одной сети, R1 посылает хосту С сообщение перенаправления. В этом сообщении хосту предлагается послать свой трафик для сети Z непосредственно маршрутизатору R2, поскольку такой путь к месту назначения является кратчайшим, после чего маршрутизатор пересылает исходную дейтаграмму по назначению (через R2). Адрес R2 указывается в поле параметра сообщения перенаправления.

**Эхо-запрос и ответ на эхо-запрос** предлагают механизм проверки возможности связи между объектами. Получатель запроса обязан вернуть ответное сообщение. С запросом связаны идентификатор и порядковый номер, которые позволяют сопоставить его с ответным сообщением. Идентификатор может использоваться наподобие точки доступа к службе, чтобы идентифицировать определенный сеанс связи, а порядковый номер может увеличиваться с каждым посланным эхо-сообщением.

**Запрос о временной метке и ответ на этот запрос** предлагают механизм измерения параметров задержки в объединенной сети. Отправитель запроса может включать в сообщение (в поле параметров) идентификатор, порядковый номер и время передачи сообщения (как оно определяется на хосте-отправителе). Получатель записывает в ответном сообщении время принятия сообщения и время, в которое он передает ответное сообщение. Если данные сообщения передаются с использованием жесткой маршрутизации, то можно измерить задержку передачи сообщения по конкретному маршруту.

Сообщения запроса об адресной маске и ответа на запрос об адресной маске полезны в среде с подсетями. Сообщения запроса и ответа позволяют хосту узнать адресную маску локальной сети, с которой он соединяется. Хост пересылает широковещательное сообщение запроса во всю ЛВС. Маршрутизатор в этой локальной сети направляет ответное сообщение, которое содержит адресную маску

## Обнаружение маршрутизатора

Возможность обнаружения маршрутизатора была добавлена к ICMP после выхода RFC 1256 с целью автоматизации процесса определения хостом адреса маршрутизатора. Чтобы хост послал дейтаграмму IP за пределы сети, к которой он подсоединен, хост должен знать адрес по крайней мере одного маршрутизатора, подсоединенного к его сети. Эти адреса могут быть заранее определены на хостах, но данный подход имеет свои недостатки. В частности, в только что подключившихся хостах (например, мобильных узлах) такие поля могут отсутствовать. RFC 1256 предлагает метод, с помощью которого хосты могут обнаруживать адреса маршрутизаторов. Он применим в сетях, имеющих возможность групповой и/или широковещательной передачи<sup>1</sup>.

В RFC 1256 определены два новых типа сообщения ICMP: объявление маршрутизатора и запрос маршрутизатора. Согласно RFC 1256 каждый маршрутизатор периодически транслирует сообщение с объявлением маршрутизатора, которое содержит следующие поля.

---

<sup>1</sup> Групповой адрес — это адрес, который обозначает группу объектов в пределах домена (например, сеть, объединенная сеть). Широковещательный адрес — это адрес, который обозначает все объекты в пределах домена (например, сеть, объединенная сеть) Групповое и широковещательное сообщение легко реализуется в локальной сети, поскольку

- **Число адресов.** Число адресов маршрутизаторов, объявленных в данном сообщении.
- **Размер поля адреса.** Число 32-битовых слов в каждом адресе маршрутизатора; должно быть равно 2.
- **Время жизни.** Максимальное время (в секундах), в течение которого объявление маршрутизатора считается действительным; значение по умолчанию — 1800 (30 мин).
- **Адрес маршрутизатора  $i$ ,** для  $1 \leq i \leq$  числа адресов. IP-адрес (адреса) передающего маршрутизатора, указывающий интерфейс, с которого было послано данное сообщение.
- **Уровень предпочтения  $i$ ,** для  $1 \leq i \leq$  числа адресов. Приоритет каждого адреса маршрутизатора  $i$  относительно других адресов в данной сети. Это значение представлено в форме дополнительного кода; большее значение указывает на более высокий уровень предпочтения.

Обычно маршрутизатор имеет один IP-адрес в сети, хотя допускаются и множественные IP-адреса. У маршрутизатора будет несколько IP-адресов, если он имеет несколько физических соединений (интерфейсов) сети. Множественные IP-адреса могут также присваиваться одному интерфейсу с целью обслуживания нескольких подсетей; последний случай мы здесь не рассматриваем. Уровень предпочтения используется хостом для определения маршрутизатора по умолчанию, который будет использоваться, если имеющаяся у хоста информация не позволяет определить, какой маршрутизатор является лучшим для достижения заданного адресата. Например (см. рис. 12.6), дейтаграмму IP от хоста D, которая направляется хосту C, лучше пересылать через маршрутизатор R2, тогда как дейтаграмма, предназначенная хосту в другом месте сети, должна посылаться через R3. Но изначально хост D может не знать, какой маршрутизатор использовать; в этом случае ему требуется послать дейтаграмму маршрутизатору по умолчанию. Если позже администратор сети определит, что большая часть трафика из ЛВС Z является локальной (к другим ЛВС в данной ячейке), то R2 может присваиваться больший уровень предпочтения; если же большая часть трафика направлена удаленным адресатам, более высокий уровень предпочтения будет присвоен R3.

Маршрутизаторы по умолчанию должны передавать сообщения с объявлением маршрутизатора каждые 7–10 мин. Если все хосты данной сети имеют возможность многоадресной передачи IP, сообщения должны посылаться по групповому адресу всей системы, 224.0.0.1. В противном случае должен использоваться широковещательный адрес 255.255.255.255.

Если хост активизировался в сети только что, он может запрашивать объявление маршрутизатора от всех подключенных маршрутизаторов, передавая сообщение с запросом маршрутизатора. Отметим, что на рис. 12.20, д это сообщение не содержит информации, отличной от определения типа сообщения. Его цель: заставить все маршрутизаторы объявить себя; таким образом, иной информации и не требуется. Если все маршрутизаторы данной сети поддерживают возможность групповой передачи IP, сообщения должны посылаться по группо-

---

ку все ее элементы используют одну среду передачи; такие сообщения доступны также в многих беспроводных и коммутационных сетевых технологиях.

вому адресу всех маршрутизаторов, 224.0.0.2. В противном случае должен использоваться широковещательный адрес 255.255.255.255.

## ПРИЛОЖЕНИЕ 12Б. АУТЕНТИФИКАЦИЯ СООБЩЕНИЙ

Одним из требований к протоколу регистрации Mobile IP является аутентификация сообщения. В этом приложении представлен краткий обзор данной темы. Более подробно она рассматривается в [STAL99].

Сообщение, файл, документ или другой набор данных называется аутентичным, если он подлинный и поступил от того, кто заявлен как автор данных. Аутентификация сообщения — это процедура, которая позволяет сообщающимся сторонам проверить аутентичность сообщения. Здесь важны оба аспекта: проверка неизменности сообщения и проверка аутентичности источника. Можно также проконтролировать своевременность сообщения (возможно, оно было искусственно задержано и воспроизведено заново) и его место в последовательности сообщений, курсирующих между сторонами.

Общая технология, которая используется для аутентификации сообщения, основывается на ключевой односторонней хэш-функции.

### Односторонняя хэш-функция

Хэш-функция отображает блок данных переменной длины в меньший блок фиксированной длины. Цель хэш-функции — синтезировать “отпечаток пальца” файла, сообщения или блока данных. Для использования при аутентификации сообщения хэш-функция  $H$  должна иметь следующие свойства.

1.  $H$  можно применить к блоку данных любого размера.
2.  $H$  дает выходной сигнал фиксированной длины.
3.  $H(x)$  должна вычисляться относительно просто для любого  $x$ , что позволяет практически реализовать ее на аппаратном и программном уровне.
4. Для любого кода  $h$  путем вычислений невозможно найти такое  $x$ , чтобы  $H(x) = h$ .
5. Для любого блока  $x$  путем вычислений невозможно найти такое  $y \neq x$ , чтобы  $H(y) = H(x)$ .
6. Путем вычислений невозможно найти такую пару  $(x, y)$ , чтобы  $H(x) = H(y)$ .

Первые три свойства являются необходимыми для практического применения хэш-функции к аутентификации сообщений. Четвертое свойство — это свойство односторонности: для данного сообщения легко сгенерировать код, но практически невозможно для данного кода сгенерировать сообщение. Это свойство является важным, если технология аутентификации включает секретные значения, как описано ниже.

Пятое свойство гарантирует невозможность нахождения альтернативного сообщения с идентичным хэш-значением.

Хэш-функция, которая удовлетворяет первые пять свойств из перечисленного выше списка, называется слабой. Если удовлетворяются все шесть свойств,

она называется **сильной**. Шестое свойство является защитой от сложных атак, основанных на задаче о днях рождения.

Помимо аутентификации профиль сообщения обеспечивает обнаружение ошибок. Он выполняет ту же функцию, что и код обнаружения ошибок, такой, как CRC: если при передаче какой-то бит сообщения случайно изменяется, профиль сообщения будет ошибочным.

В качестве хэш-функции широко используется MD5, которая синтезирует 128-битовый профиль сообщения. По умолчанию в протоколе Mobile IP используется именно MD5.

## Ключевой хэш-код

На рис. 12.21 показана технология использования хэш-функции для аутентификации сообщения. Предполагается, что две общающиеся стороны (скажем, А и В) совместно владеют общим секретным значением  $S_{AB}$ . Когда сторона А желает передать В сообщение, она вычисляет хэш-функцию путем конкатенации секретного значения и сообщения:  $MD_M = H(S_{AB} \parallel M)^2$ . Затем стороне В посылается  $[M \parallel MD_M]$ . Поскольку у стороны В есть секретное значение  $S_{AB}$ , она может пересчитать  $H(S_{AB} \parallel M)$  и проверить значение  $MD_M$ . Поскольку само секретное значение не посылается, атакующая сторона не может модифицировать перехваченное сообщение. Пока секретное значение остается секретным, станции не могут принять ложное сообщение за достоверное.

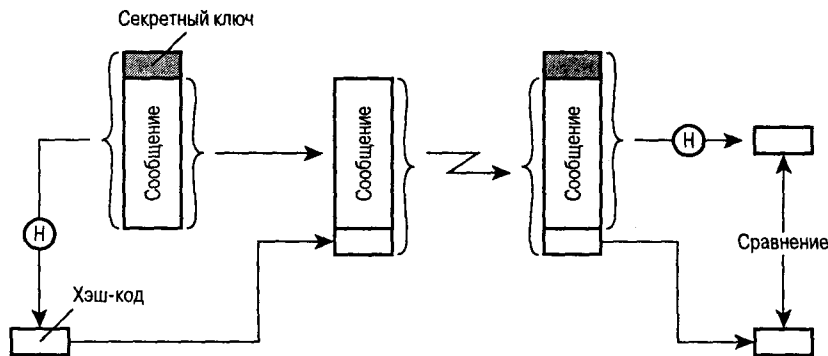


Рис. 12.21. Аутентификация сообщения с использованием односторонней хэш-функции

Схема ключевого хэш-кода зависит от односторонности хэш-функции. Если хэш-функция не является односторонней, атакующая сторона может легко вычислить секретное значение: если имеется возможность перехвата или подслушивания передачи, атакующая сторона получает сообщение  $M$  и хэш-код  $MD_M = H(S_{AB} \parallel M)$ . Далее атакующая сторона находит функцию, обратную хэш-функции, и получает  $S_{AB} \parallel M = H^{-1}(MD_M)$ . Поскольку теперь противник имеет и  $M$ , и  $S_{AB} \parallel M$ , задача восстановления  $S_{AB}$  становится тривиальной.

<sup>2</sup> Символ  $\parallel$  обозначает конкатенацию.

## ПРИЛОЖЕНИЕ 12В. ПРИМИТИВЫ И ПАРАМЕТРЫ СЛУЖБ

В архитектурах связи, таких, как модель OSI или архитектура TCP/IP (рис. 4.4), каждый уровень определяется в двух частях: протокол между одноранговыми (на том же уровне) объектами в разных системах и служба, предоставляемая уровнем вышестоящему уровню той же системы.

Службы между соседними уровнями в протокольной архитектуре выражаются через примитивы и параметры. Примитив устанавливает выполняемую функцию, а параметры используются для передачи данных и управляющей информации. Действительный вид примитива зависит от реализации, один из примеров — вызов процедуры.

Для определения взаимодействия смежных уровней архитектуры в стандартах используются четыре типа примитивов. Все они определены в табл. 12.8. На рис. 12.22, а представлена хронология происходящих событий. Рассмотрим, например, передачу запроса на соединение от WSP-пользователя А к одноранговому объекту В другой системы. Это происходит следующим образом.

**Таблица 12.8. Типы примитивов служб**

---

**ЗАПРОС.** Примитив, выпускаемый пользователем службы для инициации некоторой службы и передачи параметров, нужных для полного определения затребованной службы.

**ИНДИКАЦИЯ.** Примитив, выпускаемый поставщиком службы для одной из двух целей:

- показать, что процедура иницируется пользователем одноранговой службы, и предоставить нужные параметры;
- известить пользователя службы о действиях, иницированных поставщиком.

**ОТВЕТ.** Примитив, выпускаемый пользователем службы с целью подтверждения или завершения некоторой процедуры, ранее иницированной по отношению к данному пользователю через примитив индикации.

**ПОДТВЕРЖДЕНИЕ.** Примитив, выпускаемый поставщиком службы с целью подтверждения или завершения некоторой процедуры, ранее иницированной в ответ на запрос пользователя службы.

---

1. А иницирует службы WSP посредством примитива S-Connect.req. С этим примитивом соотнесены необходимые параметры, например адрес назначения.
2. Объект WSP в системе А подготавливает PDU WSP для отправки одноранговому объекту WSP в системе В. Затем данный PDU посылается инкапсулированным в PDU WTP.
3. Целевой объект WSP получает PDU WSP и передает данные В посредством примитива S-Connect.ind, который содержит адрес источника и другие требуемые параметры.
4. В передает объекту WSP примитив S-Connect.res.
5. Объект WSP передает PDU, в котором уведомляет объект WSP системы А о получении данных.
6. Подтверждение приема передается А посредством примитива S-Connect.cnf.

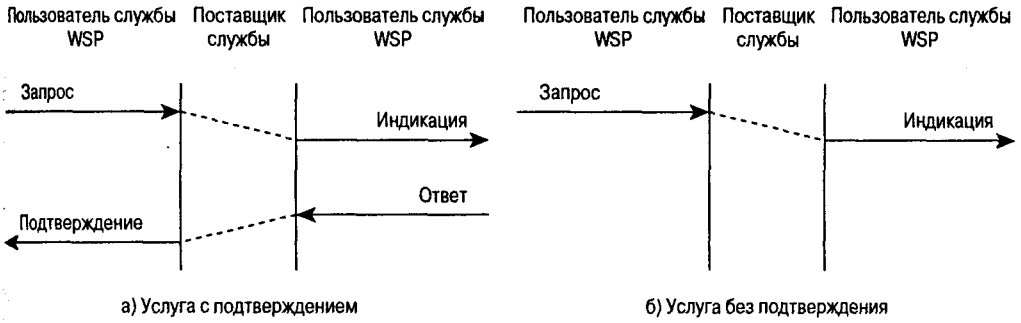


Рис. 12.22. Примитивы служб

Данная последовательность событий называется **услугой с подтверждением**, так как инициатор получает подтверждение, что запрошенная служба дала желаемый результат. Если в процессе фигурируют только примитивы запроса и индикации (что соответствует этапам 1–3), то диалог службы называется **услугой без подтверждения**; инициатор не получает подтверждения того, что запрошенное действие действительно имело место (рис. 12.22, б).

**Часть IV**

# **БЕСПРОВОДНЫЕ ЛОКАЛЬНЫЕ СЕТИ**

**Глава 13. Технология беспроводных локальных сетей**

**Глава 14. Стандарт беспроводных локальных сетей  
IEEE 802.11**

**Глава 15. Bluetooth**

# ГЛАВА 13

## ТЕХНОЛОГИЯ БЕСПРОВОДНЫХ ЛОКАЛЬНЫХ СЕТЕЙ

### 13.1. Обзор

- Применение беспроводных локальных сетей
- Требования к беспроводным локальным сетям
- Технология беспроводных локальных сетей

### 13.2. Инфракрасные локальные сети

- Плюсы и минусы
- Методы передачи

### 13.3. Сети с расширенным спектром

- Конфигурация
- Вопросы передачи

### 13.4. Сети с узкополосной СВЧ-передачей

- Лицензируемые узкополосные системы
- Нелицензируемые узкополосные системы

### 13.5. Рекомендуемая литература и Web-сайты

### 13.6. Термины и вопросы



**В** последние годы беспроводные локальные сети заняли существенную нишу на рынке локальных сетей. Все больше и больше организаций считают, что беспроводные сети являются необходимым дополнением традиционных проводных локальных сетей, удовлетворяющим требованиям мобильности, возможности передислокации, организации эпизодических сетей и охвата мест, труднодоступных при использовании проводных методов.

В данной главе дается обзор беспроводных локальных сетей. Вначале приводятся мотивы использования таких сетей и обобщаются различные подходы. В следующих трех разделах подробнее исследуются три основных типа беспроводных локальных сетей, классифицированных согласно технологии передачи: сети передачи в инфракрасном диапазоне, сети с расширенным спектром и сети с узкополосной СВЧ-передачей.

## **13.1. ОБЗОР**

Как понятно из названия, беспроводная локальная сеть связана с беспроводной средой передачи. До недавнего времени беспроводные сети использовались не очень интенсивно. Среди причин — высокие цены, низкие скорости передачи данных, проблемы профессиональной безопасности и необходимость лицензирования. По мере того как решались эти проблемы, популярность беспроводных локальных сетей быстро росла.

В данном разделе мы вначале рассмотрим требования к беспроводным локальным сетям и преимущества этих сетей, после чего перейдем к обзору ключевых подходов к реализации локальных сетей.

### **Применение беспроводных локальных сетей**

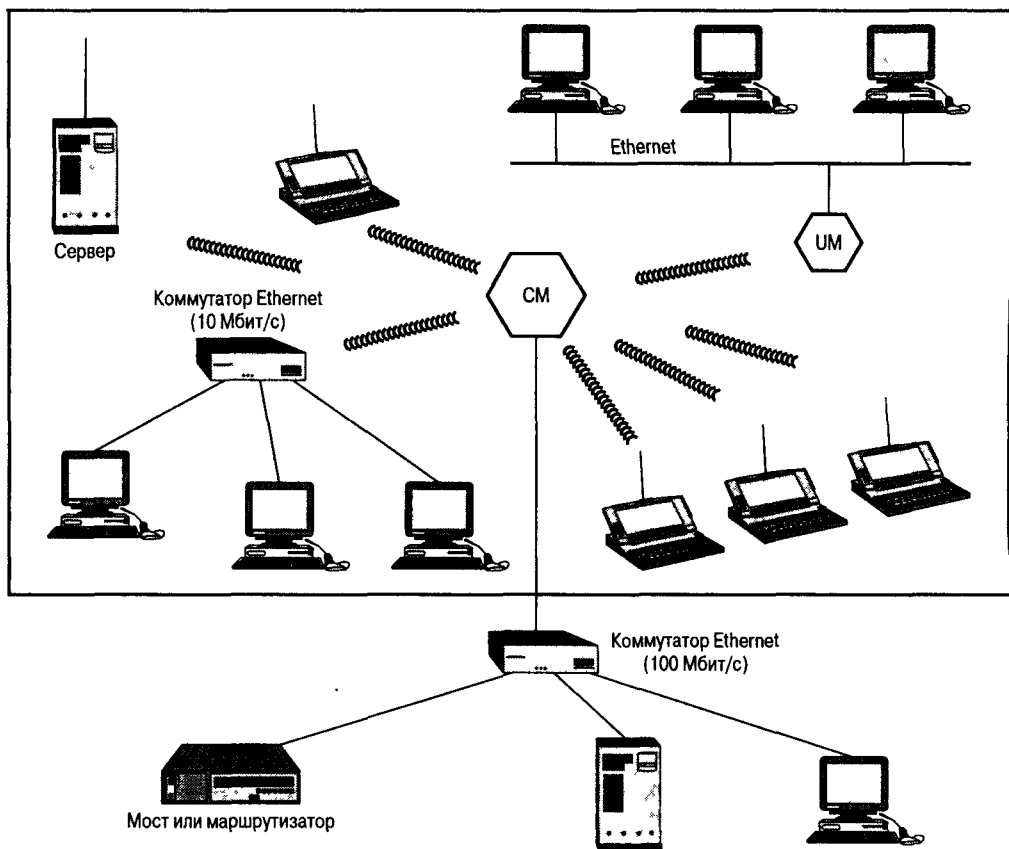
В работе [РАНЛ95] перечислены четыре области применения беспроводных локальных сетей: расширение локальных сетей, связь зданий, кочевой доступ и создание эпизодических сетей.

#### **Расширение локальных сетей**

Ранние продукты рынка беспроводных локальных сетей, появившегося в конце 1980-х, представлялись как замена традиционных проводных локальных сетей. Беспроводные локальные сети экономят средства на прокладку кабеля локальных сетей, облегчают задачу передислокации и другие модификации структуры сети. Впрочем, события опередили этот мотив использования беспроводных локальных сетей. Во-первых, по мере того, как становилась понятнее необходимость использования локальных сетей, архитекторы разрабатывали новые здания, изначально включающие кабелепроводку, которую можно было использовать для нужд сферы информационных технологий. Во-вторых, с прогрессом в технологии передачи данных бурно развилось использование в локальных сетях витой пары. Большинство старых зданий уже в изобилии снабжено кабелями категории 3; многие новые здания имеют кабели категории 5. Следовательно, замены проводных локальных сетей беспроводными в широких масштабах не произошло.

Впрочем, существуют среды, в которых беспроводные локальные сети нашли себя как альтернатива проводным. Примеры: здания с большими открытыми территориями, такие, как предприятия обрабатывающей промышленности, операционные этажи фондовых бирж и склады, исторические строения с недостаточным объемом проложенных витых пар, в которых запрещено сверление каналов под проводку, а также небольшие офисы, в которых установка и эксплуатация проводных локальных сетей неэкономична. Во всех названных случаях беспроводные локальные сети предоставляют эффективную и более привлекательную альтернативу. В большинстве из названных случаев в организации также будет существовать проводная локальная сеть для поддержки серверов и нескольких стационарных рабочих станций. Например, на предприятиях офисная территория обычно отделена от заводских этажей, но должна быть с ними связана с целью организации сети. Таким образом, беспроводные и проводные сети одного здания обычно связываются. Такая область применения беспроводных сетей называется расширением локальной сети.

На рис. 13.1 показан пример конфигурации беспроводной локальной сети, типичной для большинства сред. Имеется магистральная проводная локальная сеть, такая, как Ethernet, поддерживающая серверы, рабочие станции и один или несколько мостов или маршрутизаторов для связи с другими сетями. Имеется также модуль управления (control module — CM), действующий как устройство сопряжения с беспроводной сетью. Модуль управления обладает функциональными возможностями моста или маршрутизатора для связи беспроводной локальной сети и магистрали. Кроме того, он включает некоторую логику управления доступом, такую, как механизм опроса или схему передачи маркера для регулирования доступа от конечных систем. Отметим, что некоторые из конечных систем являются самостоятельными устройствами, такими, как рабочая станция или сервер. Частью конфигурации могут быть также концентраторы или другие пользовательские модули (user module — UM), управляющие несколькими станциями вне проводной локальной сети.



*Рис. 13.1. Пример одноячеечной конфигурации беспроводной локальной сети*

Конфигурацию, представленную на рис. 13.1, можно назвать одноячеечной беспроводной локальной сетью; все беспроводные конечные системы находятся в сфере влияния одного модуля управления. Другая распространенная конфигурация представляет собой многоячеечную беспроводную локальную сеть (рис. 13.2). В этом случае имеем несколько модулей управления, связанных проводной локальной сетью. Каждый модуль управления поддерживает несколько беспроводных конечных систем в пределах своей области связи. Например, в локальной сети, использующей передачу в инфракрасном диапазоне, передача ограничена одной комнатой; следовательно, для каждой комнаты офисного здания, нуждающегося в беспроводной поддержке, нужна отдельная ячейка.

### Связь зданий

Еще одним применением технологии беспроводных сетей является соединение локальных сетей близлежащих зданий, имеющих проводные или беспроводные локальные сети. В этом случае между двумя зданиями используется двухточечный беспроводный канал связи. Как правило, таким образом соединяются мосты или маршрутизаторы. Этот единственный канал передачи данных сам по себе не является локальной сетью, но его принято рассматривать как один из способов применения беспроводных локальных сетей.

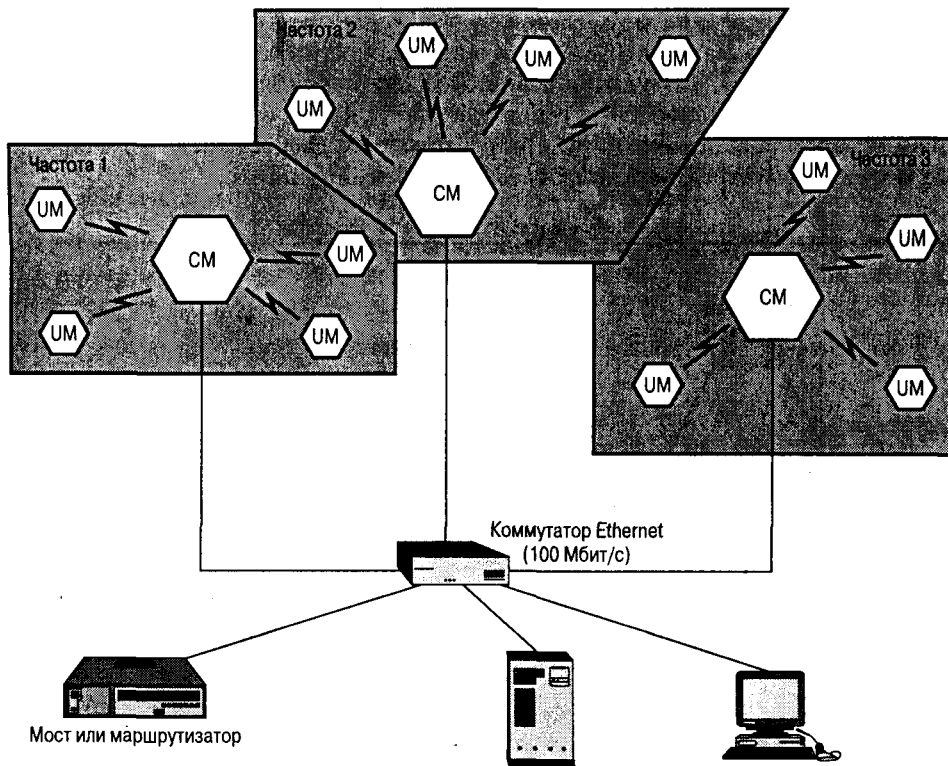


Рис. 13.2. Пример многочастотной конфигурации беспроводной локальной сети

## Кочевой доступ

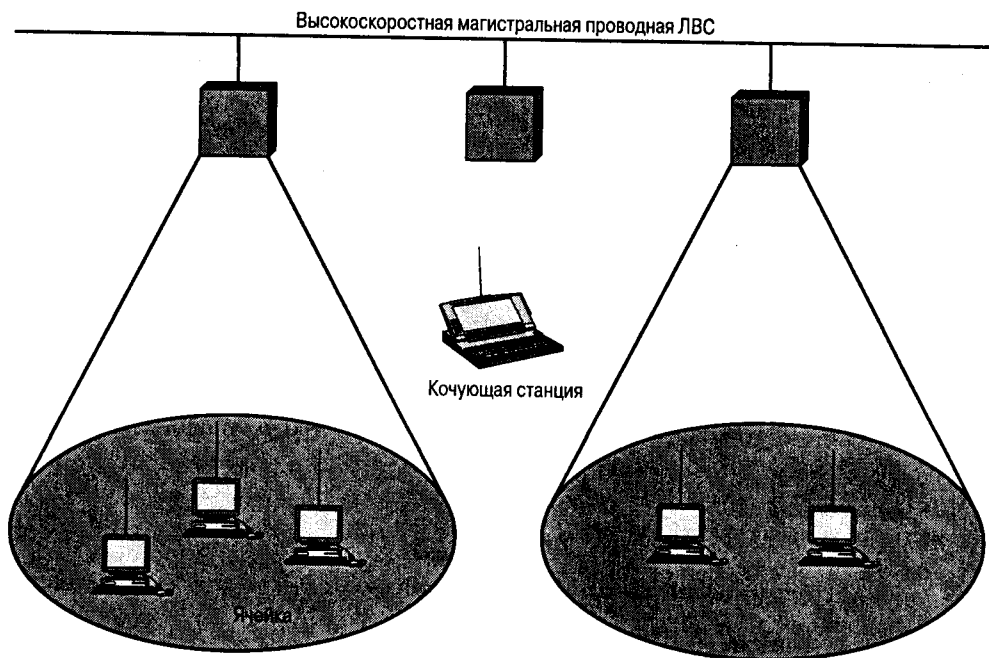
Кочевой доступ предлагает беспроводный канал связи между концентратором локальной сети и мобильным информационным терминалом, снабженным антенной, таким, как лэптоп или ноутбук. Одним из примеров применения такого соединения является передача сотрудником, вернувшимся из командировки, данных из персонального портативного компьютера на сервер в офисе. Кочевой доступ также полезен в обширной среде, такой, как кампус или предприятие, функционирующее за пределами блока зданий. В обоих случаях пользователи со своими портативными компьютерами могут перемещаться и могут связываться с серверами проводной локальной сети из различных мест.

## Организация эпизодических сетей

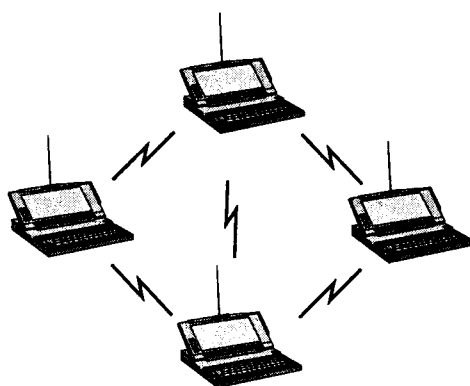
Эпизодическая сеть — это одноранговая сеть (без централизованного сервера), настроенная временно для срочного удовлетворения некоторых потребностей. Например, группа сотрудников, каждый из которых обладает лэптопом или карманным компьютером, может собраться в конференц-зале для деловой или классной встречи. Сотрудники соединяют свои компьютеры во временную сеть только на период встречи.

На рис. 13.3 показаны отличия беспроводной локальной сети, удовлетворяющей требованиям расширения сети и кочевого доступа, от эпизодической беспроводной локальной сети. В первом случае беспроводная сеть формирует

стационарную инфраструктуру, состоящую из одной или нескольких ячеек с модулем управления в каждой. Каждая ячейка может иметь несколько стационарных конечных систем. Кочующие станции могут перемещаться из одной ячейки в другую. В эпизодической сети такая инфраструктура отсутствует. В этом случае имеем набор равноправных станций в пределах досягаемости друг друга, которые могут динамически организоваться во временную сеть.



а) Беспроводная ЛВС с инфраструктурой



б) Эпизодическая сеть

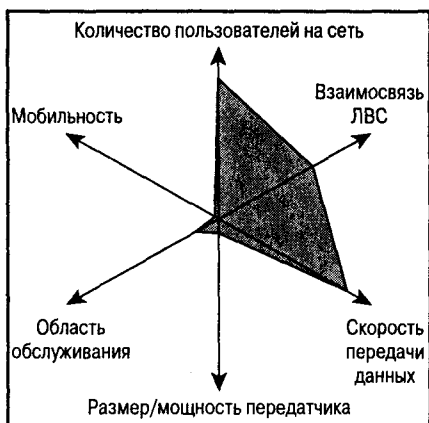
Рис. 13.3. Конфигурации беспроводных локальных сетей

## Требования к беспроводным локальным сетям

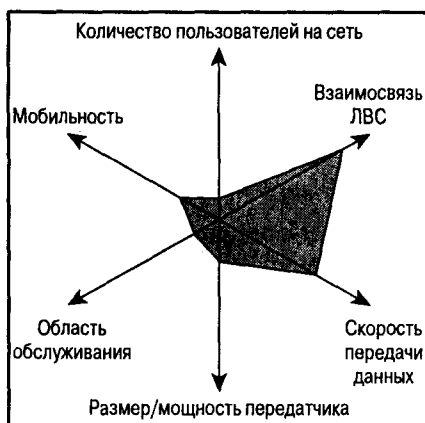
Беспроводные сети должны удовлетворять некоторым требованиям, типичным для всех локальных сетей, в том числе: высокая пропускная способность, возможность охвата небольших расстояний, связность подключенных станций и возможность широковещания. Кроме того, существует набор требований, характерных только для беспроводных локальных сетей. Перечислим важнейшие из них.

- **Производительность.** Протокол управления доступом к среде должен максимально эффективно использовать беспроводную среду для максимизации пропускной способности.
- **Число узлов.** От беспроводных локальных сетей может требоваться поддержка сотен узлов из множества ячеек.
- **Соединение с магистральной локальной сетью.** В большинстве случаев требуется взаимосвязь со станциями магистральной локальной сети. Для беспроводных локальных сетей, имеющих внутреннюю инфраструктуру, это требование легко удовлетворяется посредством использования модулей управления, присоединяемых к локальным сетям обоих типов. Может также понадобиться специальное помещение для мобильных пользователей и организация эпизодических беспроводных сетей.
- **Обслуживаемая область.** Типичная сфера охвата беспроводной локальной сети имеет диаметр 100–300 м.
- **Потребление питания от батарей.** Мобильные сотрудники используют рабочие станции с питанием от батарей, потребление которого не должно быть большим при использовании беспроводных адаптеров. Это делает неприменимым протокол MAC, требующий, чтобы мобильные узлы постоянно следили за точками доступа или часто связывались с основной станцией.
- **Устойчивость передачи и безопасность.** Беспроводные сети, если они разработаны неправильно, могут быть подвержены интерференции и легко прослушиваться. Структура беспроводной локальной сети должна обеспечивать надежную передачу даже в обстановке шума, а также некоторый уровень защиты от прослушивания.
- **Совместная работа в сети.** С ростом популярности беспроводных сетей повысилась вероятность того, что две или более сетей будут работать в одной области или в нескольких областях, допускающих интерференцию разных локальных сетей. Такая интерференция может мешать нормальной работе алгоритма MAC и способствовать несанкционированному доступу к отдельной локальной сети.
- **Работа без лицензии.** Пользователи желали бы приобретать продукты рынка беспроводных локальных сетей и работать с ними на нелицензируемой полосе частот.
- **Переключение/роуминг.** Протокол MAC, используемый в беспроводных локальных сетях, должен позволять мобильным станциям перемещаться из одной ячейки в другую.
- **Динамическая конфигурация.** MAC-адресация и сетевое управление локальной сети должны позволять динамическое и автоматическое добавление, удаление и передислокацию конечных систем, не причиняя неудобств другим пользователям.

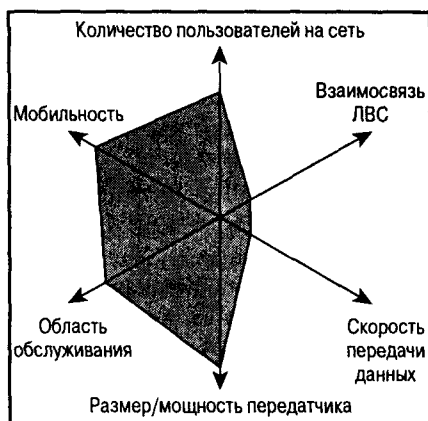
На рис. 13.4 представлены диаграммы Кивиата (Kiviat graphs)<sup>1</sup> для сравнения беспроводных локальных сетей, проводных локальных сетей и мобильных сетей передачи данных.



а) Проводные ЛВС



б) Беспроводные ЛВС



в) Сети мобильных элементов

Рис. 13.4. Диаграммы Кивиата для сетей передачи данных

## Технология беспроводных локальных сетей

Беспроводные локальные сети обычно классифицируются согласно использованной в них технологии передачи. Все современные продукты рынка локальных сетей относятся к одной из следующих категорий.

<sup>1</sup> Диаграмма Кивиата является графическим средством сравнения систем по нескольким параметрам. Переменные располагаются через равные угловые интервалы. На каждой оси ставится одна точка, соответствующая данной системе; соединение точек образует контур, являющийся характеристикой системы.

- **Инфракрасные (infrared — IR) локальные сети.** Отдельная ячейка сети, использующей передачу в инфракрасном диапазоне, ограничена размерами одной комнаты, поскольку инфракрасное излучение не проходит сквозь непрозрачные стены.
- **Локальные сети с расширенным спектром.** Данный тип локальных сетей использует при передаче технологию расширенного спектра. В большинстве случаев эти локальные сети работают на диапазонах ISM (Industrial, Scientific and Medical Radio Frequency Band — радиочастотные диапазоны для промышленного, научного и медицинского применения), поэтому для их использования в США не требуется лицензия FCC (Federal Communications Commission — Федеральная комиссия по средствам связи).
- **Узкополосная СВЧ-передача.** Эти локальные сети работают на СВЧ, но не используют расширенный спектр. Некоторые из этих продуктов работают на частотах, требующих лицензии FCC, другие используют нелицензируемые частоты ISM.

Некоторые ключевые характеристики названных технологий приведены в табл. 13.1.

**Таблица 13.1. Сравнительные характеристики беспроводных локальных сетей**

	Инфракрасное излучение		Расширенный спектр		Радио
	Рассеянное инфракрасное излучение	Инфракрасное излучение с направленным лучом	Перестройка частоты	Прямая последовательность	Узкополосная СВЧ-передача
<b>Скорость передачи данных (Мбит/с)</b>	1–4	1–10	1–3	2–20	10–20
<b>Мобильность</b>	Стационарные/мобильные	Стационарные в пределах линии прямой видимости	Мобильные		Стационарные/мобильные
<b>Диапазон (м)</b>	15–60	25	30–100	30–250	10–40
<b>Возможность обнаружения</b>	Незначительная		Малая		Некоторая
<b>Длина волны/частота</b>	$\lambda$ : 800–900 нм		902–928 МГц 2,4–2,4835 ГГц 5,725–5,85 ГГц		902–928 МГц 5,2–5,775 ГГц 18,825–19,205 ГГц
<b>Схема модуляции</b>	ASK		FSK	QPSK	FS/QPSK
<b>Получаемая мощность</b>	—		< 1 Вт		25 мВт
<b>Схема доступа</b>	CSMA	Token Ring, CSMA	CSMA		ALOHA с резервированием, CSMA
<b>Необходимость лицензии</b>	Не требуется		Не требуется		Только вне частот ISM



## 13.2. ИНФРАКРАСНЫЕ ЛОКАЛЬНЫЕ СЕТИ

Оптическая беспроводная связь в инфракрасной части спектра нашла свое применение в помещениях, где имеются многочисленные удаленные устройства. В последнее время все большее внимание уделяется использованию инфракрасных технологий для создания беспроводных локальных сетей. В данном разделе сравниваются характеристики инфракрасных сетей и радиосетей, после чего приводятся некоторые подробности относительно инфракрасных локальных сетей.

### Плюсы и минусы

В беспроводных локальных сетях используются две технологии: передача в инфракрасном диапазоне и радиопередача в СВЧ-диапазоне с использованием расширенного спектра или узкополосной передачи. Инфракрасная передача предлагает несколько существенных преимуществ. Во-первых, спектр для такой связи обычно неограничен, что дает возможность получать весьма высокие скорости передачи. Кроме того, для инфракрасного диапазона не существует регулирующих правил или стандартов, чего нельзя сказать о некоторых участках СВЧ-диапазона.

Кроме того, инфракрасное излучение имеет некоторые свойства видимого, что делает его привлекательным для определенных конфигураций локальных сетей. Инфракрасное излучение диффузно отражается от светлоокрашенных объектов; таким образом, для покрытия всей комнаты можно использовать отражение от потолка. Инфракрасное излучение не проникает сквозь стены или другие непрозрачные объекты. Это дает два преимущества: во-первых, связь в инфракрасном диапазоне легче защитить от прослушивания, чем связь в СВЧ-диапазоне; во-вторых, в каждой комнате здания может существовать своя инфракрасная конфигурация, и они не будут интерферировать между собой, что позволяет создавать значительные инфракрасные локальные сети.

Еще одним плюсом передачи в инфракрасном диапазоне является относительная простота и дешевизна соответствующего оборудования. При инфракрасной передаче данных используется модуляция интенсивности, так что ИК-приемники должны обнаруживать только амплитуду оптических сигналов, тогда как большинство СВЧ-приемников должны обнаруживать частоту или фазу.

Следует также отметить несколько недостатков рассматриваемой среды. Многие устройства, используемые внутри помещений, дают существенное фоновое излучение в ИК-диапазоне. Это внешнее излучение воспринимается ИК-приемником как шум, значит, требуются передатчики большей мощности. В то же время следует учитывать вопросы чрезмерных затрат мощности и безопасности для зрения.

### Методы передачи

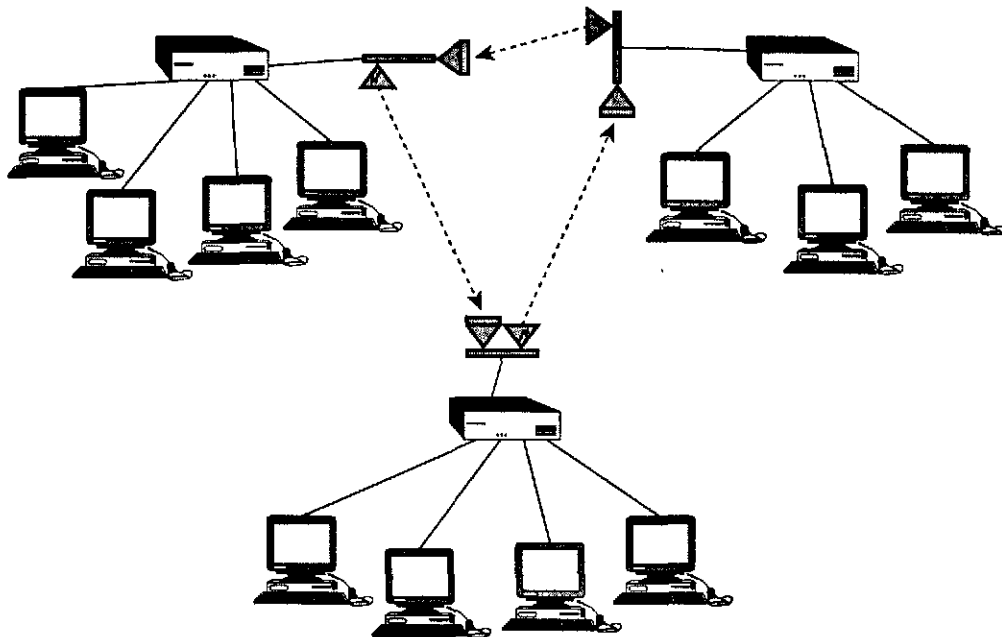
В настоящее время для передачи данных в ИК-диапазоне используются три альтернативных метода: сигнал может быть сфокусированным и направленным (пульта дистанционного управления телевизором); может излучаться равномерно во всех направлениях; может отражаться от светлоокрашенных потолков.

#### Направленный луч

Инфракрасная передача с направленным лучом используется для создания двухточечных каналов связи. В этом режиме радиус связи зависит от излучаемой

мощности и степени фокусировки. Сфокусированный инфракрасный канал передачи данных может измеряться километрами. Такие диапазоны передачи не нужны для создания комнатных беспроводных локальных сетей. В то же время инфракрасный канал связи используется для связи зданий, точнее, мостов или маршрутизаторов, расположенных в зданиях, находящихся в пределах прямой видимости.

При комнатном использовании двухточечных инфракрасных каналов может настраиваться локальная сеть типа Token Ring (рис. 13.5). Набор инфракрасных трансиверов настраивается так, чтобы данные циркулировали между ними как в кольцевой конфигурации. Каждый трансивер поддерживает рабочую станцию или концентратор станций, причем концентратор выполняет функции маршрутизатора.



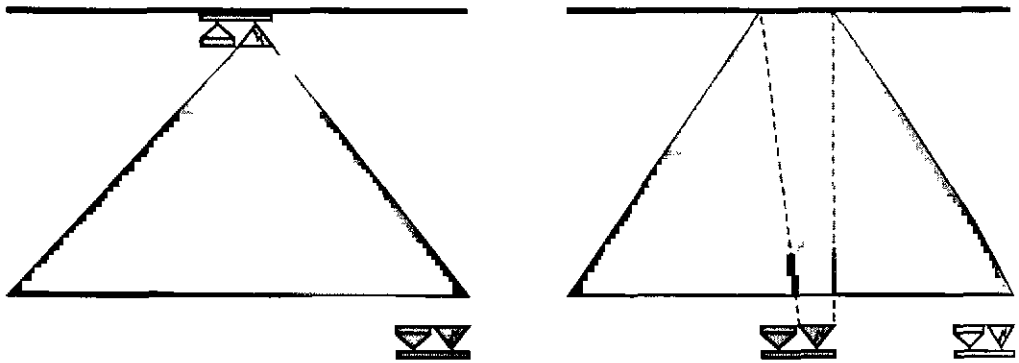
*Рис. 13.5. Локальная сеть Token Ring с использованием двухточечных инфракрасных каналов*

### **Ненаправленная передача**

Ненаправленная конфигурация включает одну базовую станцию, находящуюся в пределах прямой видимости остальных станций локальной сети. Как правило, такая выделенная станция располагается на потолке (рис. 13.6, а) и действует как многопортовый ретранслятор. Связь осуществляется следующим образом. Передатчик базовой станции ретранслирует ненаправленный сигнал, который могут принимать все другие ИК-приемники, находящиеся в охваченном диапазоне, а данные трансиверы передают направленный луч, нацеленный на базовый модуль станции.

### **Отражение**

В такой конфигурации все ИК-передатчики сфокусированы и нацелены на точку в диффузно отражающем потолке (рис. 13.6, б). Инфракрасное излучение поступает на потолок и отражается от него, после чего принимается всеми приемниками области.

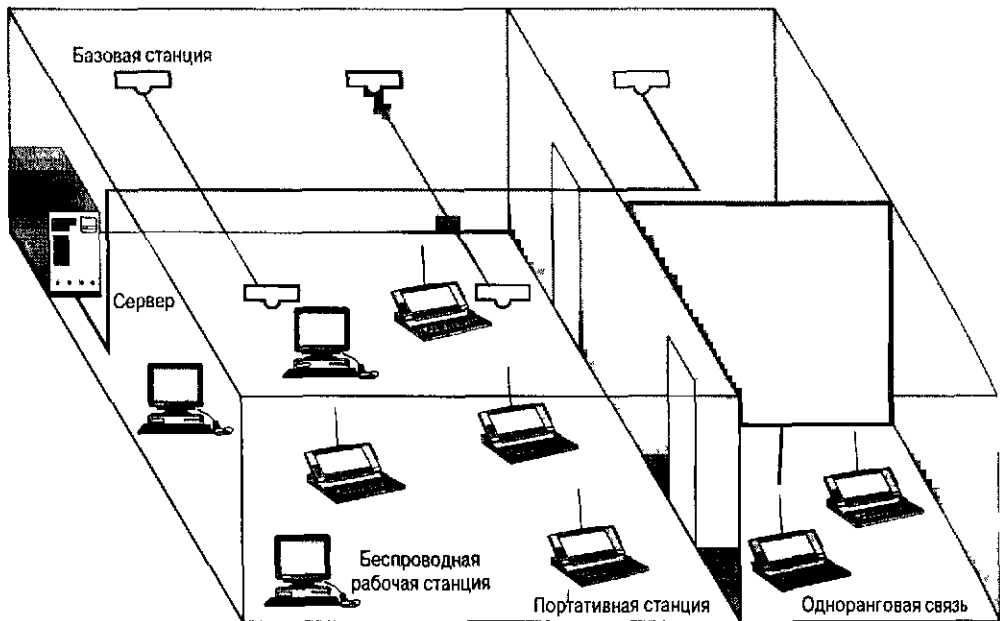


а) Связь в пределах линии прямой видимости

б) Диффузная конфигурация

*Рис. 13.6. Конфигурация ненаправленной инфракрасной локальной сети*

Пример типичной беспроводной локальной сети с инфракрасной передачей приведен на рис. 13.7. Имеется несколько (по одной на комнату) размещенных на потолке базовых станций. Каждая станция обеспечивает связь между несколькими стационарными и мобильными рабочими станциями, находящимися в обслуживаемой комнате. Между собой базовые станции соединяются проводом, они также связаны с сервером, который может выступать как точка доступа к проводной ЛВС или глобальной сети. Кроме того, могут существовать конференц-залы, в которых отсутствуют базовые станции, а связь организуется эпизодическим образом.



*Рис. 13.7. Сети портативных и стационарных беспроводных станций, использующих передачу в инфракрасном диапазоне*

## 13.3. СЕТИ С РАСШИРЕННЫМ СПЕКТРОМ

В настоящее время наиболее популярные беспроводные локальные сети используют технологию расширенного спектра.

### Конфигурация

Беспроводные локальные сети с расширенным спектром используются в многоячеечных конфигурациях, исключая только небольшие офисы (рис. 13.2). Чтобы избежать интерференции между соседними ячейками в них используются различные центральные частоты.

В пределах данной ячейки топология может иметь выделенную станцию-концентратор или быть одноранговой. Первый случай изображен на рис. 13.2. Концентратор обычно располагается на потолке и соединяется с магистральной локальной сетью с целью обеспечения связи со станциями, присоединенными к проводной локальной сети, и станциями, принадлежащими той же локальной сети, но расположенными в других ячейках. Концентратор может также управлять доступом: выступать как точечный координатор IEEE 802.11 или как многопортовый ретранслятор с функциями, подобными ретрансляторам 10 Мбит/с и 100 Мбит/с Ethernet. В последнем случае все станции ячейки передают только концентратору и принимают только от него. Возможен альтернативный вариант, не зависящий от механизма управления доступом: каждая станция использует свою ненаправленную антенну для широковещания, так что эту передачу могут принимать все другие станции ячейки; данная конфигурация соответствует шинной топологии.

Еще одной потенциальной функцией концентратора является автоматическое переключение мобильных станций. В любой момент времени число станций, обслуживаемых данным концентратором, определяется динамически и зависит от пространственной близости. Если концентратор отмечает, что сигнал станции начинает слабеть, он может автоматически перепоручить обслуживание этой станции ближайшему концентратору.

В однородной топологии концентратор отсутствует и для управления доступом используется алгоритм MAC, например CSMA. Лучше всего такая топология подходит эпизодическим сетям.

### Вопросы передачи

Желательной, но не необходимой, характеристикой беспроводной локальной сети является возможность ее использования без процедуры лицензирования. Условия лицензирования отличаются в разных странах, что только усложняет ситуацию. В США Федеральная комиссия по средствам связи (FCC) разрешила использование приложений двух типов в нелицензируемой полосе ISM: системы с расширенным спектром с мощностью до 1 Вт и маломощные системы ( $\leq 0,5$  Вт). Таким образом, для использования беспроводных локальных сетей с расширенным спектром не требуется получения лицензии, и их популярность неуклонно растет.

В США для нелицензируемого использования в системах с расширенным спектром выделены три частотных диапазона: 902–928 МГц (“полоса 915 МГц”), 2,4–2,4835 ГГц (“полоса 2,4 ГГц”) и 5,725–5,825 ГГц (“полоса 5,8 ГГц”). Второй из этих диапазонов также используется в Европе и Японии. Чем выше частота, тем больше потенциальная ширина полосы, так что данные три диапазона пере-

числены в порядке возрастания привлекательности с точки зрения пропускной способности. Кроме того, следует рассматривать возможность интерференции. Существует множество устройств, работающих на частоте порядка 900 МГц, в том числе радиотелефоны, радиомикрофоны и любительские радиостанции. На частоте порядка 2,4 ГГц работает меньшее число устройств; к ним относится микроволновая печь, внешнее излучение которой увеличивается с возрастом печи. В настоящее время полоса 5,8 ГГц используется относительно мало; впрочем, чем выше частота, тем дороже оборудование для ее эксплуатации.

До недавнего времени типичные беспроводные локальные сети с расширенным спектром ограничивались скоростями передачи 1–3 Мбит/с. Как показано в главе 14, новые стандарты дают возможность использовать скорости 5,5 и 11 Мбит/с, а некоторые частные решения — более высокие скорости.

## 13.4. СЕТИ С УЗКОПОЛОСНОЙ СВЧ-ПЕРЕДАЧЕЙ

Термином *узкополосная СВЧ-передача* (narrowband microwave) обозначают использование для передачи сигнала полосы сверхвысоких частот при относительно узкой полосе — широкой ровно настолько, насколько требуется для вмещения сигнала. До недавнего времени все локальные продукты сетей с узкополосной СВЧ-передачей использовали СВЧ-диапазон, требующий получения лицензии. Затем по крайней мере один поставщик произвел продукт, предназначенный для полосы ISM.

### Лицензируемые узкополосные системы

СВЧ, которые используются для передачи речи, данных и видео, требуют получения лицензии, и их использование координируется, что позволяет избежать потенциальной интерференции между системами. В США лицензирование контролирует Федеральная комиссия по средствам связи. Каждая географическая область имеет радиус 28 км, и в ней могут выдаваться до пяти лицензий, причем каждая лицензия выдается на две частоты. Motorola имеет 600 лицензий (1200 частот) в полосе 18 ГГц, которые позволяют охватить все области с населением свыше 30 000 человек.

В узкополосной схеме обычно используется конфигурация, представленная на рис. 13.2. Соседние ячейки используют неперекрывающиеся диапазоны частот в пределах общей полосы 18 ГГц. Поскольку Motorola контролирует распределение своих частот, эта компания может гарантировать, что близко расположенные независимые локальные сети не будут мешать друг другу. Отметим также, что для защиты от подслушивания все передачи шифруются.

Одним из преимуществ лицензируемых узкополосных локальных сетей является гарантированная связь без интерференции. В отличие от использования нелицензируемого спектра, такого, как диапазон ISM, владелец лицензии имеет легальное право на создание канала передачи данных без интерференции. Передачи пользователей локальных сетей, функционирующих в диапазоне ISM, могут повреждаться вследствие интерференции с передачами других пользователей и воспрепятствовать этому законным способом нельзя.

### Нелицензируемые узкополосные системы

В 1995 году компания RadioLAN стала первым поставщиком продуктов для локальных беспроводных сетей с использованием нелицензируемой полосы частот.

ISM, которая может применяться для маломощной ( $\leq 0,5$  Вт) узкополосной передачи. Продукты RadioLAN обеспечивают скорость передачи 10 Мбит/с и работают в полосе 5,8 ГГц. Область действия продуктов ограничена расстоянием 50 м в полуоткрытом офисе и 100 м в открытом офисе.

Продукты RadioLAN предназначены для одноранговой конфигурации, но с одной любопытной особенностью. Взамен стационарного концентратора в сети автоматически выбирается одна динамическая ведущая станция (Dynamic Master), причем выбор основывается на параметрах станции, таких, как расположение, наличие интерференции и интенсивность сигнала. При изменении условий ведущая станция также может меняться. Еще одной особенностью данных сетей является наличие функции динамической ретрансляции, которая позволяет каждой станции выступать как ретранслятор при передаче данных станциям, находящимся вне пределов досягаемости.

### 3.5. РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА И WEB-САЙТЫ

Великолепными обзорными статьями по беспроводным локальным сетям являются [PAHL95] и [BANT94]. В работе [KAHN97] рассмотрены сети, использующие передачу в инфракрасном диапазоне.

BANT94 Bantz D., Bauchot F. Wireless LAN Design Alternatives. — *IEEE Network*, March/April, 1994.

KAHN97 Kahn J., Barry J. Wireless Infrared Communications. — *Proceedings of the IEEE*, February 1997.

PAHL95 Pahlavan K., Probert T., Chase M. Trends in Local Wireless Networks. — *IEEE Communications Magazine*, March 1995.



#### Рекомендуемые Web-сайты

- Wireless LAN Alliance ([www.wlana.org/](http://www.wlana.org/)). Предлагает введение в технологию, в том числе включает обсуждение вопросов реализации и детальное описание конкретных примеров, предоставленных пользователями. Приводятся ссылки на родственные узлы.

### 3.6. ТЕРМИНЫ И ВОПРОСЫ

#### Основные термины

■ беспроводные ЛВС

■ инфракрасный диапазон

■ диаграмма Кивиата

кочевой доступ

организация

эпизодических сетей

расширение локальных

сетей

# ГЛАВА 14

## СТАНДАРТ БЕСПРОВОДНЫХ ЛОКАЛЬНЫХ СЕТЕЙ IEEE 802.11

- 14.1. Архитектура IEEE 802
  - Архитектура протоколов
  - Формат кадра MAC
  - Управление логическим каналом
- 14.2. Архитектура и услуги IEEE 802.11
  - Архитектура IEEE 802.11
  - Услуги IEEE 802.11
- 14.3. Управление доступом к среде в IEEE 802.11
  - Надежная доставка данных
  - Управление доступом
  - Кадр MAC
  - Вопросы безопасности
- 14.4. Физический уровень IEEE 802.11
  - Исходный физический уровень IEEE 802.11
  - IEEE 802.11a
  - IEEE 802.11b
- 14.5. Рекомендуемая литература и Web-сайты
- 14.6. Термины и вопросы

**Н**аиболее популярной спецификацией беспроводных локальных сетей является стандарт, разработанный рабочей группой IEEE 802.11. В начале данной главы рассматривается общая архитектура стандартов IEEE 802, затем речь пойдет об особенностях стандарта IEEE 802.11.

## 14.1. АРХИТЕКТУРА IEEE 802

Архитектуру локальной сети лучше всего описывать через многоуровневую протокольную структуру, упорядочивающую основные функции локальной сети. Данный раздел начинается с описания стандартизированной протокольной архитектуры для локальных сетей, которая включает физический уровень, уровень управления доступом к среде и уровень управления логическим каналом. Затем подробно рассматриваются два последних.

### Архитектура протоколов

Протоколы, определенные специально для передачи по локальным и городским сетям, решают вопросы, связанные с передачей блоков данных по сети. В терминах OSI протоколы высокого уровня (уровень 3 или 4 и выше) независимы от архитектуры сети и применимы в локальных, городских и глобальных сетях. Таким образом, обсуждение протоколов локальных сетей в основном связано с нижними уровнями модели OSI.

На рис. 14.1 сопоставляются протоколы локальных сетей и архитектура OSI (рис. 4.3). Эта архитектура была разработана комитетом IEEE 802 и принята всеми организациями, работающими в области спецификации стандартов локальных сетей. Обычно она называется эталонной моделью IEEE 802 (IEEE 802 reference model)<sup>1</sup>.

Рассмотрим эталонную модель IEEE 802. Нижний уровень этой модели соответствует физическому уровню модели OSI и включает следующие функции:

- кодирование/декодирование сигнала;
- генерирование/удаление начальной комбинации битов (для синхронизации);
- прием/передача битов.

Кроме того, физический уровень модели 802 включает спецификацию среды передачи и топологии (поскольку выбор среды передачи и топологии является важным вопросом в определении структуры локальной сети).

Над физическим уровнем рассматриваются функции, связанные с предоставлением услуг пользователям локальной сети.

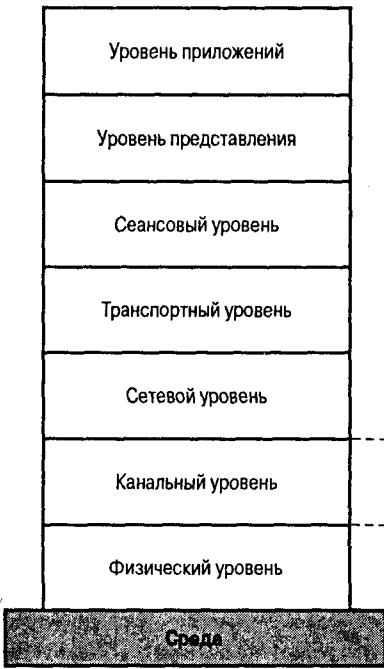
- При передаче — дополнение данных кадра информацией относительно адреса и полями выявления ошибок.
- При приеме — выделение информационной части кадра, распознавание адреса и выявление ошибок.
- Регулирование доступа к среде передачи локальной сети.
- Организация сопряжения с вышестоящими уровнями и реализация управления потоком и защиты от ошибок.

---

<sup>1</sup> Работа комитета IEEE 802 рассмотрена в приложении А.



Эталонная модель OSI



Эталонная модель IEEE 802

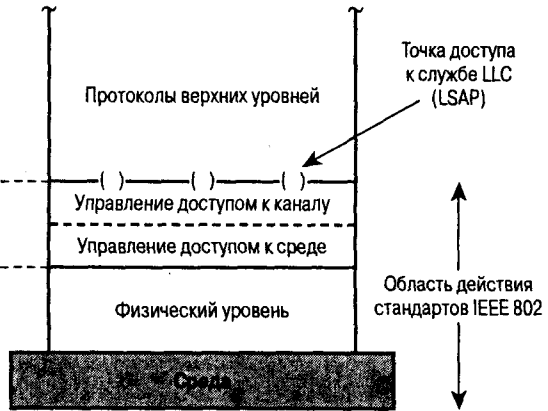


Рис. 14.1. Сопоставление моделей OSI и IEEE 802

Эти функции обычно связываются со вторым уровнем модели OSI. Набор функций в последнем пункте списка выделяется в уровень **управления логическим каналом** (logical link control — LLC). Функции, перечисленные в первых трех пунктах списка, также рассматриваются как отдельный уровень, называющийся **уровнем управления доступом к среде** (medium access control — MAC). Назовем причины такого разделения.

- Логика, необходимая для управления доступом к среде совместного использования, не находится на традиционном уровне 2 управления каналом передачи данных.
- При одинаковом управлении логическим каналом могут быть реализованы разные схемы управления доступом к среде.

На рис. 14.2 (копия приводившегося ранее рис. 11.16) показана связь уровней архитектуры. Высокоуровневые данные спускаются на уровень LLC, на котором к ним в виде заголовка присоединяется управляющая информация, при этом создается *модуль данных протокола LLC*, или PDU LLC. Эта управляющая информация используется в операциях протокола LLC. PDU LCC спускается на уровень MAC, где к нему в начале и конце пакета присоединяется управляющая информация, формируя в итоге *кадр MAC*. Эта дополнительная управляющая информация необходима для работы протокола MAC. Для полноты картины на рисунке над протоколами локальной сети показан также уровень приложений и использование протокола TCP/IP.

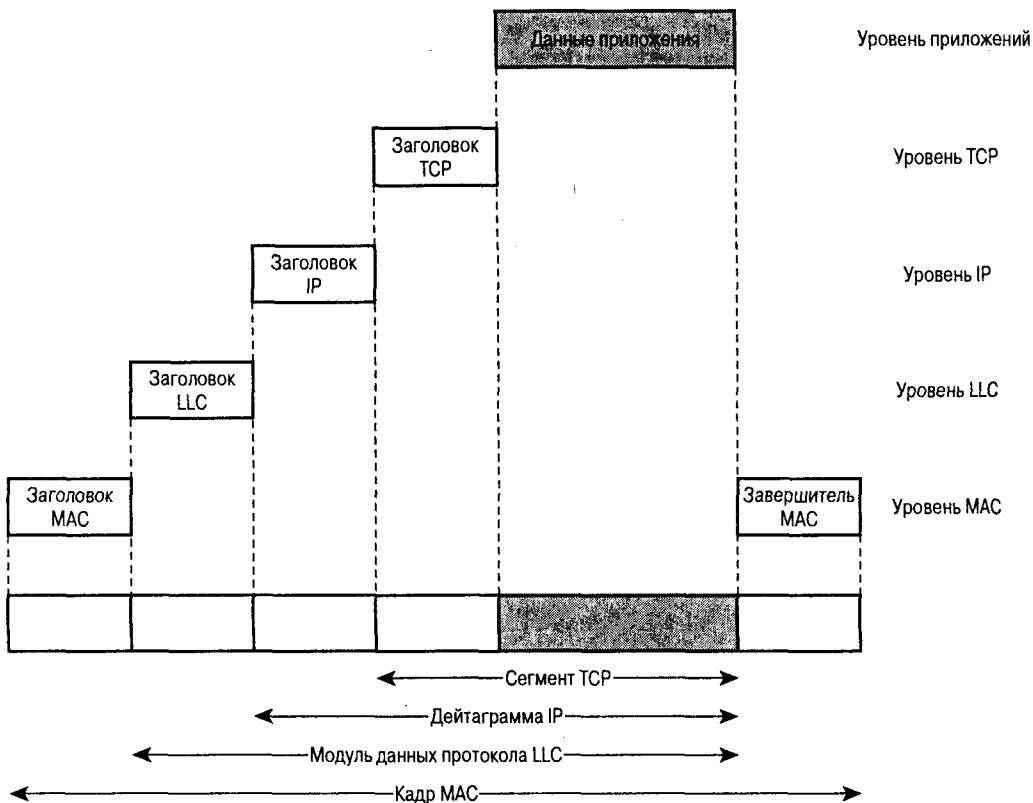


Рис. 14.2. Протоколы IEEE 802 в контексте

## Формат кадра MAC

Уровень MAC получает блок данных от уровня LLC и отвечает за выполнение функций, связанных с доступом к среде, и за передачу данных. Как и другие протокольные уровни, MAC реализует указанные функции, используя на своем уровне модуль данных протокола (protocol data unit — PDU). В этом случае PDU называется MAC-кадром.

Точный формат MAC-кадра несколько отличается для различных используемых протоколов MAC. Вообще все MAC-кадры имеют формат, подобный тому, что показан на рис. 14.3. Перечислим поля этого кадра.

- **Управление MAC.** Это поле содержит всю управляющую информацию, необходимую для функционирования протокола MAC. Например, здесь может указываться уровень приоритета.
- **MAC-адрес получателя.** Целевая физическая точка подключения к локальной сети для этого кадра.
- **MAC-адрес отправителя.** Исходная физическая точка подключения к локальной сети для этого кадра.
- **Данные.** Тело MAC-кадра. Здесь могут находиться данные LLC от вышестоящего уровня или управляющая информация, связанная с работой протокола MAC.



Рис. 14.3. PDU LLC в общем формате MAC-кадра

- **CRC.** Поле циклической проверки четности с избыточностью (cyclic redundancy check — CRC), называемое также полем контрольной последовательности кадра (frame check sequence — FCS). Это код выявления ошибок, как было показано для протокола HDLC и других протоколов управления каналом передачи данных (раздел 8.1).

В большинстве протоколов управления каналом передачи данных объект протокола канала передачи данных отвечает не только за выявление ошибок с использованием CRC, но и за исправление этих ошибок путем повторной передачи поврежденных кадров. В протокольной архитектуре локальной сети эти две функции разделены между уровнями MAC и LLC. Уровень MAC отвечает за выявление ошибок и отклонение кадров с ошибками. Уровень LLC может отслеживать, какие кадры были успешно приняты, и повторно передавать остальные.

## Управление логическим каналом

Уровень LLC для локальных сетей во многих отношениях подобен другим широко используемым канальным уровням. Он связан с передачей PDU канального уровня между двумя станциями, без промежуточных коммутационных узлов. Отличают протокол LLC от других протоколов управления каналом передачи данных следующие две характеристики.

1. Он должен поддерживать множественный доступ, составляющий природу канала с совместно используемой средой (что отличается от многоточечной линии, в которой отсутствует выделенный основной узел).
2. Протокол LLC несколько упрощен за счет передачи уровню MAC некоторых элементов управления доступом.

Адресация в LLC включает задание исходного и целевого пользователей LLC. Обычно пользователь — это вышестоящий протокол или функция управления сетью на станции. Эти адреса пользователей LLC называются точками доступа к службе (service access point — SAP), что согласуется с терминологией OSI для пользователя протокольного уровня.

Рассмотрим услуги, которые LLC предоставляет пользователю вышестоящего уровня, после чего перейдем к самому протоколу LLC.

## Услуги LLC

LLC определяет механизмы задания в среде адресов станций и управления информационным обменом между двумя пользователями. В основе работы формата этих стандартов лежит протокол HDLC. Подключенным устройствам, использующим LLC, предлагаются на выбор три вида услуг.

- **Услуги без подтверждений и без организации соединения.** Эти услуги работают по принципу использования дейтаграмм. Они крайне просты и не включают механизмы управления потоком и защиты от ошибок. Таким образом, доставка данных не гарантируется. Впрочем, в большинстве устройств за надежность отвечает некоторый более высокий программный уровень.
- **Услуги с установкой соединения.** Подобны аналогичным услугам, предлагаемым протоколом HDLC. Между двумя пользователями, обменивающимися данными, устанавливается логическое соединение; кроме того, обеспечивается управление потоком и защита от ошибок.
- **Услуги с подтверждениями без организации соединения.** Это нечто среднее между двумя видами услуг, описанными выше. Дейтаграммы при этом подтверждаются, но логическое соединение не устанавливается.

Обычно поставщик предлагает эти услуги как альтернативные варианты, из которых заказчик может выбрать желаемый, приобретая оборудование. Возможна также ситуация, когда заказчик приобретает оборудование, предлагающее два или все три вида услуг, и выбирает нужное исходя из конкретных обстоятельств.

Услуги без подтверждений и без организации соединения требуют минимальной логики и полезны в двух случаях. Во-первых, часто бывает так, что вышестоящие уровни программного обеспечения гарантируют необходимую надежность и предлагают механизм управления потоком, поэтому целесообразнее избежать ненужного дублирования. Например, TCP может предлагать механизмы, необходимые для гарантирования надежной доставки данных. Во-вторых, бывает, что служебные издержки по установке и поддержанию соединения неоправданы или даже вредны (например, действия по сбору данных, включающие периодическую выборку из источников информации, таких, как сенсоры, или автоматические самотестирующиеся отчеты, поступающие от средств обеспечения безопасности или сетевых компонентов). При организации наблюдения потеря одного из периодически поступающих модулей данных не составит большой проблемы, если следующий отчет поступит в скором времени. Следовательно, в большинстве описанных случаев неподтвержденная служба без организации соединения является предпочтительным вариантом.

Услуги с установкой соединения могут использоваться в очень простых устройствах, таких, как терминальные контроллеры, имеющих малофункциональные высшие программные уровни. В этих случаях предлагаемые услуги будут обеспечивать управление потоком и надежность, что обычно реализуется на более высоких уровнях программного обеспечения связи.

Услуги с подтверждениями без организации соединения также полезны в нескольких случаях. Если используются услуги с установкой соединения, программа, управляющая логическим каналом, должна хранить какую-то таблицу для каждого

активного соединения, чтобы отслеживать его состояние. Если пользователю требуется гарантировать доставку, но при этом существует большое число адресатов данных, то услуги с установкой соединения могут оказаться непрактичными из-за большого числа требуемых таблиц. Пример: управление процессом или автоматизированная промышленная среда, где центральному узлу может потребоваться связаться с большим числом устройств обработки или программируемых контроллеров. Другим применением данной службы является обработка важных и требующих немедленного реагирования сигналов тревоги или аварийного управления. Вследствие важности этих сигналов необходимо подтверждение их приема, чтобы отправитель был уверен, что сигнал дошел по назначению. Пользователь может передать срочный сигнал, не тратя времени на установку соединения.

## Протокол LLC

Основной протокол LLC моделировался после протокола HDLC, поэтому формат и функции этих протоколов очень похожи. Отличия же можно сформулировать следующим образом.

- LLC использует асинхронный сбалансированный режим работы HDLC для поддержки службы LLC, ориентированной на установление соединения; это называется режимом работы 2. Другие режимы HDLC не используются.
- LLC поддерживает неподтвержденную службу без организации соединения с использованием нумерованных PDU; это называется режимом работы 1.
- LLC поддерживает подтвержденную службу без организации соединения путем использования двух новых нумерованных PDU; это называется режимом работы 3.
- LLC допускает уплотнение путем использования точек доступа к службе LLC (LLC service access point — LSAP).

Все три протокола LLC используют одинаковый формат PDU (см. рис. 14.3), состоящий из четырех полей. Поля DSAP и SSAP содержат адреса исходного и целевого пользователей LLC по 7 бит каждый. Один бит DSAP показывает, является ли DSAP индивидуальным или групповым адресом. Один бит SSAP указывает, является ли PDU командой или откликом. Формат управляющего поля LLC идентичен формату аналогичного поля HDLC (см. рис. Г.1, приложение Г) и использует расширенные порядковые номера (состоящие из 7 бит).

В режиме работы 1, поддерживающем неподтвержденную службу без организации соединения, для передачи пользовательских данных служат нумерованные информационные (unnumbered information — UI) PDU. Подтверждение отсутствует, управление потоком не осуществляется, защита от ошибок не обеспечивается. Впрочем, функции выявления ошибок и отклонения кадров имеются на уровне MAC.

Два других типа PDU используются для поддержки функций управления, связанных со всеми тремя режимами работы. Оба типа PDU используются следующим образом. Объект LLC может выдать команду (бит C/R = 0) XID или TEST. Принимающий объект LLC выдает соответствующую команду XID или TEST в ответ. PDU XID используется для обмена информацией двух типов: о поддерживаемых режимах и о размере окна. PDU TEST используется для проведения теста канала передачи данных между двумя объектами LLC методом об-

ратной петли. После получения команды TEST целевой объект LLC должен как можно быстрее ответить PDU TEST.

В режиме работы 2 соединение с использованием канала передачи данных устанавливается между двумя SAP LLC до обмена данными. Протокол режима 2 пытается установить соединение в ответ на запрос пользователя. Объект LLC выдает PDU SABME<sup>2</sup> для запроса логического соединения с другим объектом LLC. Если соединение принято пользователем LLC, указанным точкой DSAP, то целевой объект LLC возвращает PDU с нумерованным подтверждением (unnumbered acknowledgment — UA). С этого момента соединение уникально определяется парой пользовательских SAP. Если целевой пользователь LLC отклоняет запрос на установление соединения, объект LLC возвращает PDU режима разрыва (disconnected mode — DM).

После установления соединения обмен данными, как и в протоколе HDLC, выполняется посредством информационных PDU. Эти PDU включают порядковые номера приема и передачи, необходимые для упорядочения потока и управления им. PDU управления используются, как и в протоколе HDLC, для управления потоком и защиты от ошибок. Любой объект LLC может разорвать логическое соединение, передав PDU разрыва соединения (disconnect — DISC).

При режиме работы 3 каждый переданный PDU подтверждается. В отличие от протокола HDLC, в LLC определен новый нумерованный модуль PDU, информационный PDU с подтверждением без организации соединения (Acknowledged Connectionless — AC). Пользовательские данные передаются в командных модулях PDU AC, и их прием должен подтверждаться с использованием отклика PDU AC. Для защиты от потери PDU используется 1-битовый порядковый номер. Отправитель чередует 0 и 1 в передаваемых командных PDU AC, а приемник отвечает PDU AC с номерами, противоположными соответствующим командным PDU. В каждый момент времени в обращении может находиться не более одного PDU с каждой стороны.

## 14.2. АРХИТЕКТУРА И УСЛУГИ IEEE 802.11

Работу по стандартизации беспроводных локальных сетей в комитете IEEE 802 начала в 1987 году группа IEEE 802.4. Изначально планировалось разработать беспроводную локальную сеть, являющуюся эквивалентом шинного протокола MAC с передачей маркера, функционирующую в диапазоне ISM. После некоторой работы было признано, что шинная топология с передачей маркера является не самой удачной схемой для радиопередачи с точки зрения эффективного использования спектра радиочастот. В 1990 году было принято решение о создании новой рабочей группы IEEE 802.11, задачей которой стало рассмотрение беспроводных локальных сетей и которая получила право разрабатывать спецификации протокола MAC и физической среды. Ключевые термины, фигурирующие в стандартах IEEE 802.11, перечислены в табл. 14.1.

---

<sup>2</sup> SABME означает Set Asynchronous Balanced Mode Extended (установить асинхронный симметричный режим, расширенный). Эта команда используется в протоколе HDLC для выбора асинхронного симметричного режима (ABM) и расширенной (7-битовой) нумерации. Для режима работы 2 режим ABM и 7-битовые порядковые номера являются обязательными.

**Таблица 14.1. Терминология IEEE 802.11**

---

Базовый набор услуг (basic service set — BSS)	Набор станций, которыми управляет одна координационная функция
Координационная функция	Логическая функция, определяющая, когда станция может передавать и получать PDU
Модуль данных протокола MAC (MPDU)	Модуль данных, которым обмениваются два одноранговых объекта MAC, используя услуги физического уровня
Модуль данных службы MAC (MSDU)	Информация, передаваемая единым блоком между пользователями MAC
Распределительная система (distribution system — DS)	Система, которая используется для соединения нескольких базовых наборов услуг и интеграции локальной сети в расширенный набор услуг
Расширенный набор услуг (extended service set — ESS)	Две или большее число базовых наборов услуг, соединенных распределительной системой. Для уровня управления логическим каналом (logical link control — LLC) расширенный набор услуг представляется единой логической локальной сетью
Станция	Любое устройство, физический уровень и MAC-уровень которого соответствуют стандарту IEEE 802.11
Точка доступа (access point — AP)	Любой объект, обладающий функциональными возможностями станции и обеспечивающий доступ соотнесенных с ним станций к распределительной системе посредством беспроводной среды

---

## Архитектура IEEE 802.11

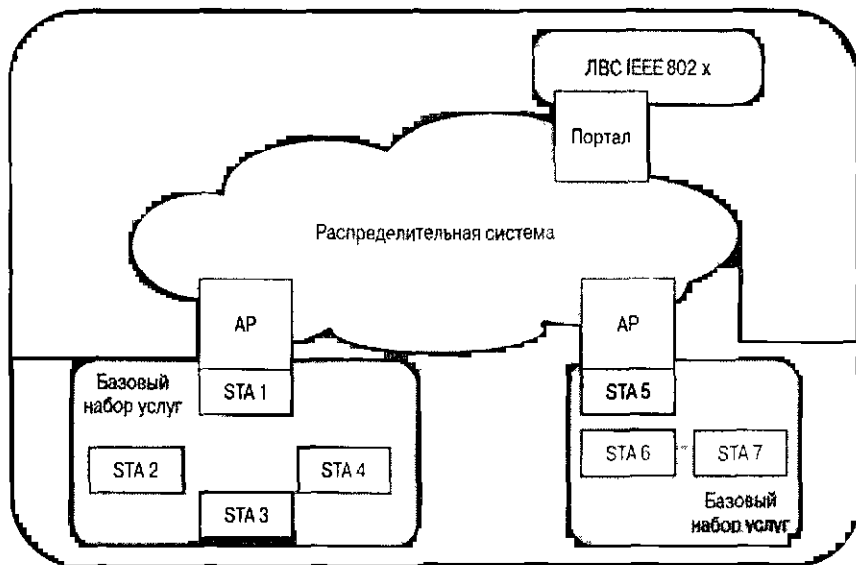
На рис. 14.4 показана модель, разработанная рабочей группой 802.11. Элементарной структурной единицей беспроводной локальной сети является базовый набор услуг (basic service set — BSS), который состоит из нескольких станций, реализующих одинаковый протокол MAC и состоящих за доступ к одной совместно используемой среде. BSS может быть изолирован или соединен с магистральной распределительной системой через точку доступа. Точка доступа в этой ситуации играет роль моста. Протокол MAC может быть полностью распределенным или управляемым центральной координационной функцией, размещенной в точке доступа. Базовый набор услуг обычно соответствует тому, что в литературе называют ячейкой. Распределительная система может быть коммутатором, проводной или беспроводной сетью.

На рис. 14.4 показана простейшая конфигурация, при которой каждая станция принадлежит к одному BSS; т.е. каждая станция находится в пределах досягаемости только станций того же BSS. Диапазоны охвата BSS могут перекрываться (географически), так что одна станция может входить в несколько BSS. Более того, соотношение станции и BSS является динамическим. Станции могут отключаться, выходить из диапазона охвата и входить в этот диапазон.

Расширенный набор услуг (extended service set — ESS) состоит из двух или большего числа базовых наборов услуг, соединенных распределительной системой. Обычно в качестве распределительной системы выступает проводная магистральная локальная сеть. Для уровня управления логическим каналом (logical link control — LLC) расширенный набор услуг представлен единой логической локальной сетью.

На рис. 14.4 показано, что точка доступа (access point — AP) реализуется как часть станции. Точка доступа — это логика в пределах станции, обеспечивающая

доступ к распределительной системе путем предоставления ей услуг; кроме того, точка действует как станция. Для интеграции архитектуры IEEE 802.11 с традиционной проводной локальной сетью используется портал. Логика портала реализуется в устройстве, таком, как мост или маршрутизатор, являющемся частью проводной локальной сети и присоединенном к распределительной системе.



STA — станция

Рис. 14.4. Архитектура IEEE 802.11

## Услуги IEEE 802.11

Группа IEEE 802.11 определила девять услуг, которые должна предлагать беспроводная локальная сеть для обеспечения функциональных возможностей проводной сети. Данные услуги перечислены в табл. 14.2, там также представлены два способа группировки этих услуг.

Таблица 14.2. Услуги IEEE 802.11

Услуга	Поставщик	Объект поддержки
Ассоциация	Распределительная система	Доставка MSDU
Аутентификация	Станция	Доступ к ЛВС и безопасность
Доставка MSDU	Станция	Доставка MSDU
Интеграция	Распределительная система	Доставка MSDU
Конфиденциальность	Станция	Доступ к ЛВС и безопасность
Отмена аутентификации	Станция	Доступ к ЛВС и безопасность
Повторная ассоциация	Распределительная система	Доставка MSDU
Разрыв ассоциации	Распределительная система	Доставка MSDU
Распределение	Распределительная система	Доставка MSDU



1. Поставщиком услуги может быть либо станция, либо распределительная система (DS). Услуги станций реализуются на каждой станции 802.11', в том числе на станциях, являющихся точками доступа (AP). Услуги распределительных систем предлагаются между базовыми наборами услуг (BSS); эти услуги могут реализовываться на точках доступа или на других специализированных устройствах, присоединенных к распределительной системе.
2. Для управления доступом к ЛВС и безопасностью используются три услуги. Еще шесть услуг используются для поддержки доставки модулей данных службы MAC (MSDU) от станции к станции. MSDU — это блок данных, передаваемых пользователем MAC на уровень MAC; обычно это PDU уровня LLC. Если модуль MSDU чересчур большой для передачи одним кадром MAC, его можно фрагментировать и передать в последовательности кадров MAC. Вопросы фрагментации рассмотрены в разделе 14.3.

Приведенные ниже услуги описаны так, как они определены в документах IEEE 802.11, причем выбран такой порядок их рассмотрения, чтобы была максимально понятна работа сети IEEE 802.11. Основная служба, именуемая доставкой MSDU, рассматривалась выше.

### **Распределение сообщений в пределах распределительной системы**

В распределении сообщений в пределах DS задействованы две услуги: распределение и интеграция. **Распределение** — это основная услуга, используемая станциями для обмена MAC-кадрами, когда кадр должен пройти через распределительную систему, чтобы проследовать от станции из одного BSS к станции, находящейся в другом BSS. Предположим, например, что требуется доставить кадр от станции 2 (STA2) к станции 7 (STA7) (рис. 14.4). Кадр передается от STA2 к STA1, которая для данного BSS является точкой доступа. Точка доступа передает кадр распределительной системе, которая должна направить кадр к точке доступа (STA5) целевого BSS. STA5 получает кадр и передает его STA7. Как сообщение передается через распределительную систему, в стандарте IEEE 802.11 не описывается.

Если две станции сообщаются в пределах одного BSS, услуга распределения предоставляется точкой доступа этого BSS.

Услуга **интеграции** позволяет передавать данные между станцией локальной сети IEEE 802.11 и станцией интегральной локальной сети IEEE 802.x. Термином *интегральная* называется проводная локальная сеть, физически соединенная с распределительной системой, причем станции этой сети могут логически соединяться с локальной сетью IEEE 802.11 посредством услуги интеграции. Услуга интеграции разрешает все вопросы, связанные с логикой трансляции адресов и преобразований среды, требуемых для обмена данными.

### **Услуги, связанные с ассоциацией**

Основной задачей уровня MAC является передача модулей MSDU между объектами MAC; выполняет эту задачу распределительная система, для функционирования которой требуется информация о станциях в пределах ESS. Эта информация поставляется услугами, связанными с ассоциацией. Перед тем как распределительная система сможет передавать или принимать информацию станции, требует-

ся установить *ассоциацию*. Данную концепцию мы рассмотрим немного позже, а пока необходимо описать понятие мобильности. На основе мобильности станций в стандарте определены три типа переходов.

- **Без перехода.** Станция либо стационарна, либо перемещается в пределах досягаемости станций, принадлежащих к тому же BSS.
- **Переход BSS.** Переход станции из одного BSS в другой в пределах одного ESS. В этом случае для доставки данных требуется найти новое местоположение станции.
- **Переход ESS.** Перемещение станции из BSS одного ESS в BSS другого ESS. Переход этого типа поддерживается только в том смысле, что станция может двигаться. Сохранность соединений высшего уровня, поддерживаемых сетью 802.11, гарантировать нельзя. Фактически наиболее вероятным следствием подобного перехода является разрыв услуги.

Для доставки сообщений в пределах распределительной системы услуге распределения должно быть известно, где расположена станция-адресат. В частности, распределительная система должна знать, какая станция выступает в роли точки доступа, т.е. кому передавать сообщения, предназначенные станции-адресату. Для удовлетворения этого требования станция должна поддерживать ассоциацию с точкой доступа в пределах текущего BSS. С этим требованием связаны три услуги.

- **Ассоциация.** Установление первоначальной ассоциации между станцией и точкой доступа. Перед тем как станция начнет передавать или получать кадры в беспроводной локальной сети, ее нужно идентифицировать. Для этого станция должна установить ассоциацию с точкой доступа в пределах конкретного BSS. Затем эта точка доступа может передать необходимую информацию о станции остальным точкам доступа данного ESS, что облегчит будущую маршрутизацию и адресную доставку.
- **Повторная ассоциация.** Разрешает передавать установленную ассоциацию между точками доступа, позволяя мобильной станции перемещаться между наборами BSS.
- **Разрыв ассоциации.** Уведомление от станции или точки доступа об аннулировании существующей ассоциации. Станция должна получить это уведомление до выхода из ESS или отключения. В то же время средство управления MAC защищено от станций, которые исчезают без уведомления.

### Услуги доступа и безопасности

Существуют две характеристики проводной локальной сети, отличающие ее от беспроводной.

1. Для передачи по проводной ЛВС станция должна быть физически соединена с этой сетью. С другой стороны, в беспроводной сети передавать может любая станция, находящаяся в пределах радиосвязи других устройств. В определенном смысле в проводной сети имеется некоторая разновидность аутентификации — для соединения с сетью станция должна выполнить кое-какие предопределенные действия.

2. Подобным образом, для приема передачи от станции, принадлежащей проводной локальной сети, принимающая станция должна также быть присоединена к проводной сети. С другой стороны, в беспроводной сети принимать может любая станция, находящаяся в пределах радиосвязи других устройств. Таким образом, проводная ЛВС обеспечивает некоторую конфиденциальность, ограничивая число возможных получателей сообщения станциями, соединенными с сетью.

Поскольку беспроводная локальная сеть не имеет данных характеристик, в стандарте IEEE 802.11 определены следующие услуги.

- **Аутентификация.** Используется для идентификации станций. В проводной ЛВС обычно предполагается, что доступ к физическому соединению дает право на связь с ЛВС. Это предположение не является верным для беспроводных сетей, где связность устройств обуславливается простым наличием антенны, настроенной соответствующим образом. Услуга аутентификации используется станциями для идентификации себя в среде уже общающихся станций. В IEEE 802.11 определены несколько схем аутентификации, имеется также возможность расширения этих схем. Стандарт не навязывает никакой обязательной схемы, поэтому может использоваться как относительно небезопасная схема квитирования, так и весьма надежная схема шифрования с открытым ключом. При этом, правда, требуется взаимно приемлемая, успешная аутентификация, только после которой станция может устанавливать ассоциацию с точкой доступа.
- **Отмена аутентификации.** Эта услуга используется при аннуляции существующей аутентификации.
- **Конфиденциальность.** Предотвращает чтение сообщения теми, кому оно не предназначено. Для обеспечения конфиденциальности стандарт предлагает (но не навязывает) использовать шифрование, согласно алгоритму WEP, который рассмотрен в разделе 14.3.

## 14.3. УПРАВЛЕНИЕ ДОСТУПОМ К СРЕДЕ В IEEE 802.11

Уровень MAC стандарта IEEE 802.11 охватывает три функциональные области: надежную доставку данных, управление доступом и безопасность. Все они подробно рассмотрены ниже.

### Надежная доставка данных

Как и в любой беспроводной сети, физический уровень и MAC-уровень сети IEEE 802.11 довольно ненадежны. Шум, интерференция и другие эффекты распространения приводят к потере большого числа кадров. Даже при использовании кодов с коррекцией ошибок успешно принять удается не все кадры MAC. Решение данной проблемы можно поручить вышестоящему уровню, такому, как TCP. В то же время использовать таймеры повторной передачи высших уровней — это терять порядка нескольких секунд. Следовательно, обработку ошибок эффективнее производить на уровне MAC. С этой целью в стандарте IEEE 802.11

определен протокол обмена кадров. Когда станция получает от другой станции информационный кадр, она возвращает исходной станции кадр подтверждения (acknowledgement — ACK). Данный сеанс считается элементарным актом, который не должен прерываться передачей другой станции. Если станция не получает кадра ACK в течение короткого периода времени (либо по причине повреждения информационного кадра, либо вследствие повреждения подтверждения), источник передает кадр повторно.

Таким образом, базовый механизм передачи в стандарте IEEE 802.11 включает обмен двумя кадрами. Для дальнейшего повышения надежности может использоваться обмен четырьмя кадрами. В этой схеме источник вначале передает запрос на передачу (request to send — RTS). Адресат отвечает кадром “готов к приему” (clear to send — CTS). После приема кадра CTS источник передает кадр данных, а адресат отвечает подтверждением (ACK). Кадр RTS оповещает все станции в пределах области приема исходной станции, что происходит обмен информацией; все станции, принявшие этот кадр, воздерживаются от передачи, чтобы избежать конфликта вследствие одновременной передачи двух кадров. Подобным образом, кадр CTS оповещает все станции, находящиеся в пределах области приема адресата, что происходит обмен информацией. В заключение отметим, что кадры RTS/CTS требуются MAC, но их использование можно заблокировать.

## Управление доступом

Рабочая группа 802.11 рассмотрела два типа предложенных алгоритмов MAC: протоколы распределенного доступа, которые, подобно Ethernet, распределяют решение относительно передачи по всем узлам, используя для этого механизм детектирования несущей; протоколы централизованного доступа, которые включают регулирование передачи централизованным средством принятия решений. Протокол распределенного доступа имеет смысл использовать в эпизодических сетях равноправных рабочих станций, его также целесообразно применять в других беспроводных конфигурациях, трафик в которых носит преимущественно пульсирующий характер. Централизованный протокол доступа естественным образом подходит для конфигураций, в которых несколько беспроводных станций связаны между собой и имеется также определенная базовая станция, подключенная к магистральной проводной локальной сети; особенно полезны такие протоколы, если некоторые данные требуют немедленного реагирования или имеют высокий приоритет.

Конечным результатом работы комитета 802.11 является алгоритм MAC, называемый DFWMAC (distributed foundation wireless MAC — распределенный основной беспроводный протокол MAC), предоставляющий механизм распределенного управления доступом с возможным надстроенным централизованным управлением. На рис. 14.5 показана соответствующая архитектура. Нижний подуровень уровня MAC представляет собой распределенную координационную функцию (distributed coordination function — DCF). DCF использует алгоритм состязания для обеспечения доступа ко всему информационному потоку. Обычный асинхронный информационный поток напрямую использует DCF. Точечная координационная функция (point coordination function — PCF) — это централизованный алгоритм управления доступом к среде, используемый для обеспечения бесконфликтного обслуживания. PCF надстраивается над DCF и использует

функции DCF для обеспечения доступа всем пользователям. Ниже последовательно рассматриваются оба названных подуровня.

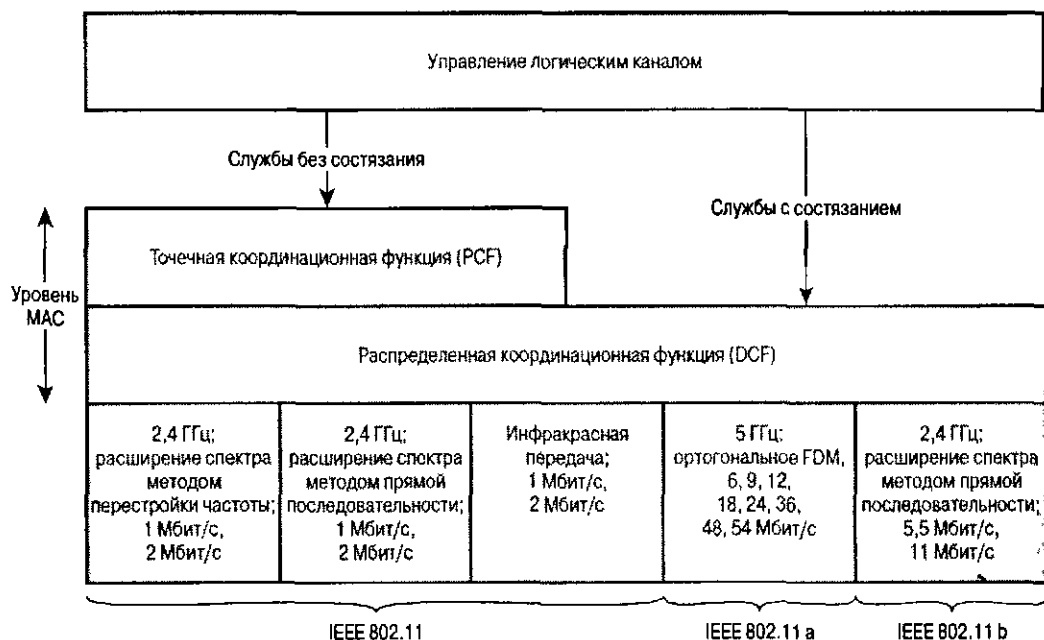


Рис. 14.5. Протокольная архитектура IEEE 802.11

### Распределенная координационная функция

Подуровень DCF использует простой алгоритм CSMA. Если станция имеет готовый к передаче кадр MAC, она прослушивает среду. Если среда свободна, станция может передавать; в противном случае станция должна подождать, пока не завершится текущая передача, а лишь затем передавать. DCF не содержит функций обнаружения конфликтов (например, CSMA/CD), поскольку обнаружение конфликтов в беспроводной среде непрактично. Динамический диапазон сигналов в среде очень велик, так что передающая станция не может эффективно отличить поступающий слабый сигнал от шума и результатов собственной передачи.

Для обеспечения бесперебойной и ровной работы описанного алгоритма подуровень DCF включает набор задержек, которые эквивалентны схеме приоритетов. Рассмотрим вначале отдельную задержку, называемую межкадровым промежутком (interframe space — IFS). На самом деле существует три различных значения IFS, но действие алгоритма более понятно, если вначале проигнорировать эти детали. Использование IFS следующим образом меняет правила доступа алгоритма CSMA (рис. 14.6).

1. Станция, имеющая готовые к передаче кадры, прослушивает среду. Если среда свободна, станция ожидает, будет ли среда свободна в течение времени IFS. Если да, то станция может передавать немедленно.
2. Если среда занята (либо была занята на начало прослушивания, либо стала занятой за время IFS), станция откладывает передачу и продолжает наблюдать за средой, пока не будет завершена текущая передача.

3. Как только будет завершена текущая передача, станция выдерживает еще один промежуток времени IFS. Если в течение этого времени среда остается свободной, станция освобождает среду и снова прослушивает ее. Если среда по-прежнему остается свободной, станция может начинать передачу. Если в течение времени отсрочки среда занята, таймер отсрочки останавливается и запускается после освобождения среды.

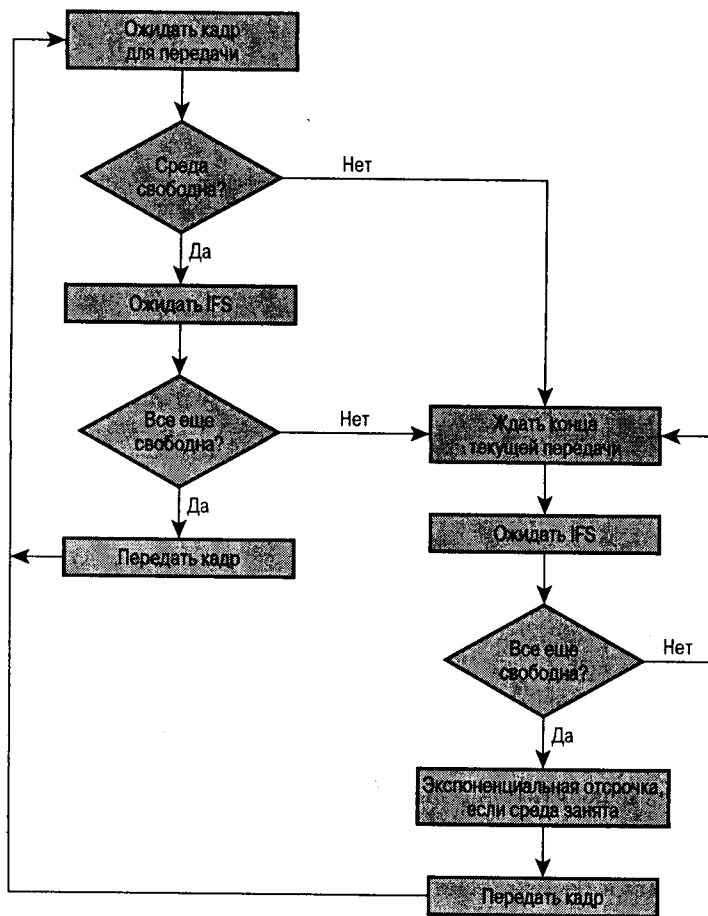


Рис. 14.6. Логика управления доступом к среде стандарта IEEE 802.11

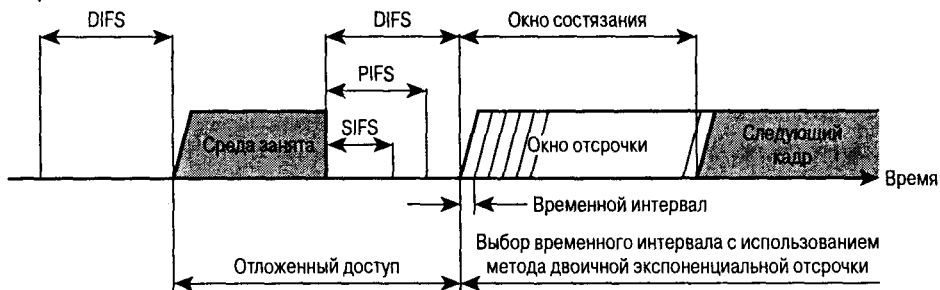
Для обеспечения устойчивости схемы используется метод двоичной экспоненциальной отсрочки. Станция постоянно пытается начать передачу, каждый раз сталкиваясь с конфликтом, но после каждого конфликта среднее время случайной задержки удваивается. Двоичная экспоненциальная отсрочка позволяет справиться со значительной нагрузкой. Постоянные попытки передачи, заканчивающиеся конфликтами, приводят к увеличению времени ожидания, что, в свою очередь, сглаживает нагрузку. Без подобной отсрочки может произойти следующее: две или более станции одновременно пытаются начать передачу, что приводит к конфликту, станции немедленно пытаются передать данные повторно, возникает новый конфликт и т.д.

Для создания схемы, подобной алгоритму доступа с использованием приоритетов, описанная схема несколько уточняется, для чего вводятся три значения IFS.

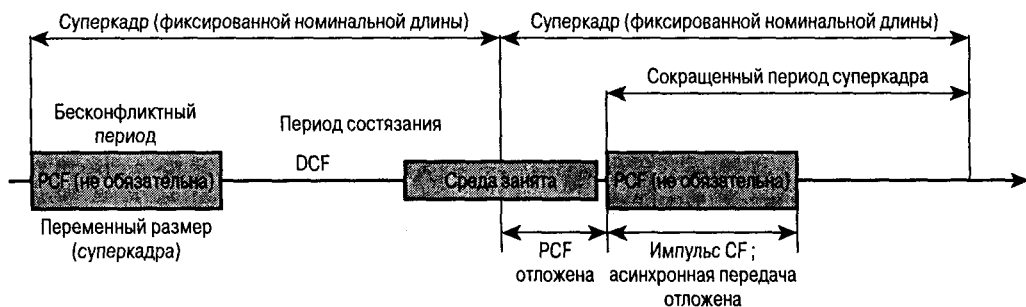
- **SIFS (short IFS — краткий IFS).** Минимальный IFS, используемый для всех немедленных ответных действий, как будет объяснено далее.
- **PIFS (point coordination function IFS — IFS точечной координационной функции).** IFS среднего размера, используемый централизованным средством управления в схеме PCF при организации опроса.
- **DIFS (distributed coordination function IFS — IFS распределенной координационной функции).** Наибольший IFS, используемый как минимальная задержка для асинхронных кадров, состоящих за доступ.

На рис. 14.7, а показано использование описанных временных промежутков. Рассмотрим вначале SIFS. Каждая станция, использующая SIFS для определения возможности передачи, имеет, по сути, наивысший приоритет, поскольку при получении доступа она всегда имеет преимущество перед станцией, ожидающей в течение времени PIFS или DIFS. SIFS используется в следующих ситуациях.

Немедленный доступ,  
если среда свободна  
дольше срока DIFS



а) Основной метод доступа



б) Структура суперкадра PCF

Рис. 14.7. Временная диаграмма управления доступом к среде IEEE 802.11

- **Подтверждение (acknowledgment — ACK).** Когда станция получает кадр, адресованный только ей (т.е. имеем не групповую или широкове-

щательную передачу), она отвечает кадром ACK после ожидания только в течение времени SIFS. Получаем два желательных результата. Во-первых, поскольку не используется обнаружение конфликтов, вероятность их возникновения больше, чем в схеме CSMA/CD, а подтверждение на уровне MAC обеспечивает эффективное восстановление после ошибок. Во-вторых, SIFS может использоваться для обеспечения эффективной доставки модуля данных протокола (protocol data unit — PDU) LLC, которая требует нескольких кадров MAC. В этом случае имеем следующий сценарий. Станция с готовым к передаче многокадровым PDU LLC последовательно передает кадры MAC. Каждый кадр подтверждается получателем по прошествии времени SIFS. Когда станция-источник получает кадр ACK, она немедленно (по прошествии времени SIFS) посылает следующий кадр последовательности. В результате как только станция захватывает контроль над каналом, она будет его удерживать до окончания передачи всех фрагментов PDU LLC.

- “Свободен для передачи” (clear to send — CTS). Станция может гарантировать доставку своего кадра, выпустив небольшой кадр RTS (Request to Send — запрос на передачу). Станция, которой адресован этот кадр, должна немедленно ответить кадром CTS, если она готова к приему. Все остальные станции получают кадр RTS и выжидают, используя среду, пока не увидят соответствующий кадр CTS или пока не истечет predeterminedное время ожидания.
- Ответ на опрос (poll response). Этот момент будет объяснен позднее, при рассмотрении PCF.

Следующим по размеру интервалом IFS является PIFS. Он используется централизованным средством управления при организации упорядоченных опросов и имеет преимущество перед обычным конкурирующим информационным потоком. В то же время кадры, переданные с использованием SIFS, имеют преимущество перед опросом PCF.

И последний интервал, DIFS, используется для организации обычного асинхронного информационного потока.

### **Точечная координационная функция**

PCF — это альтернативный метод доступа, реализованный на один уровень выше DCF. Его работа заключается в опросе, производимом централизованным мастером опроса (точечным координатором). Точечный координатор использует PIFS при отправке запросов. Поскольку время PIFS меньше времени DIFS, точечный координатор может захватывать среду и блокировать асинхронный поток информации на время отправки запросов и получения ответов.

Рассмотрим следующий предельный случай. Беспроводная сеть сконфигурирована так, что несколько станций с чувствительным ко времени информационным потоком контролируются точечным координатором, тогда как оставшийся поток состоит из доступа с использованием CSMA. Точечный координатор может циклически выпускать запросы ко всем станциям, сконфигурированным для упорядоченного опроса. При запуске запроса опрашиваемая станция может ответить, используя SIFS. Если точечный координатор получает ответ, он выпускает другой запрос, используя PIFS. Если в течение predeterminedного времени не получено отклика, координатор выпускает запрос.





делить канал для успешной передачи кадра MAC. В некоторых кадрах управления в этом поле указывается идентификатор ассоциации, или соединения.

- **Адреса.** Число и значение полей адреса зависит от контекста. Возможны следующие типы адреса: источника, назначения, передающей станции, принимающей станции.
- **Управление очередностью.** Содержит 4-битовое подполе номера фрагмента, используемое для фрагментации и повторной сборки, и 12-битовый порядковый номер, используемый для нумерации кадров, передаваемых между данными приемником и передатчиком.
- **Тело кадра.** Содержит модуль MSDU или фрагмент MSDU. В данном случае MSDU — это модуль данных протокола LLC или управляющая информация MAC.
- **Контрольная последовательность кадра.** 32-битовая проверка четности с избыточностью.

Поле управления кадром, показанное на рис. 14.8, б, состоит из следующих полей.

- **Версия протокола.** Версия 802.11, текущая версия — 0.
- **Тип.** Определим тип кадра: контроль, управление или данные.
- **Подтип.** Дальнейшая идентификация функций кадра. Разрешенные сочетания типов и подтипов перечислены в табл. 14.3.

**Таблица 14.3. Разрешенные комбинации типа и подтипа**

Значение типа	Описание типа	Значение подтипа	Описание подтипа
00	Управление	0000	Запрос ассоциации
00	Управление	0001	Ответ на запрос ассоциации
00	Управление	0010	Запрос повторной ассоциации
00	Управление	0011	Ответ на запрос повторной ассоциации
00	Управление	0100	Пробный запрос
00	Управление	0101	Ответ на пробный запрос
00	Управление	1000	Сигнальный кадр
00	Управление	1001	Объявление наличия трафика
00	Управление	1010	Разрыв ассоциации
00	Управление	1011	Аутентификация
00	Управление	1100	Отмена аутентификации
01	Контроль	1010	PS-опрос
01	Контроль	1011	Запрос передачи
01	Контроль	1100	“Готов к передаче”
01	Контроль	1101	Подтверждение
01	Контроль	1110	Без состязания (CF)-конец

Значение типа	Описание типа	Значение подтипа	Описание подтипа
01	Контроль	1111	CF-конец + CF-подтверждение
10	Данные	0000	Данные
10	Данные	0001	Данные + CF-подтверждение
10	Данные	0010	Данные + CF-опрос
10	Данные	0011	Данные + CF-подтверждение + CF-опрос
10	Данные	0100	Нулевая функция (без данных)
10	Данные	0101	Данные + CF-подтверждение
10	Данные	0110	Данные + CF-опрос
10	Данные	0111	Данные + CF-подтверждение + CF-опрос

- **К DS.** Координационная функция MAC присваивает этому биту значение 1, если кадр предназначен распределительной системе.
- **От DS.** Координационная функция MAC присваивает этому биту значение 0, если кадр исходит от распределительной системы.
- **Больше фрагментов.** 1, если за данным фрагментом следует еще несколько.
- **Повтор.** 1, если данный кадр является повторной передачей предыдущего.
- **Управление мощностью.** 1, если передающая станция находится в режиме ожидания.
- **Больше данных.** Указывает, что станция передала не все данные. Каждый блок данных может передаваться как один кадр или как группа фрагментов в нескольких кадрах.
- **WEP.** 1, если реализован алгоритм конфиденциальности проводного эквивалента (wired equivalent privacy — WEP). Протокол WEP используется для обмена ключами шифрования при безопасном обмене данными.
- **Порядок.** 1, если используется услуга строгого упорядочения, указывающая адресату, что кадры должны обрабатываться строго по порядку.

Рассмотрим теперь различные типы кадров MAC.

### Контрольные кадры

Контрольные кадры способствуют надежной доставке информационных кадров. Существует шесть подтипов контрольных кадров.

- **Опрос после выхода из экономичного режима (PS-опрос).** Данный кадр передается любой станцией станции, включающей точку доступа. В кадре запрашивается передача кадра, прибывшего, когда станция находилась в режиме энергосбережения, и в данный момент размещенного в буфере точки доступа.
- **Запрос передачи (RTS).** Данный кадр является первым из четверки, используемой для обеспечения надежной передачи данных (см. начало раздела 14.3). Станция, пославшая это сообщение, предупреждает адресата и остальные станции, способные принять данное сообщение, о своей попытке передать адресату информационный кадр.

- **“Готов к передаче” (CTS).** Второй кадр четырехкадровой схемы. Передается станцией-адресатом станции-источнику и предоставляет право отправки информационного кадра.
- **Подтверждение (ACK).** Подтверждение успешного приема предыдущих данных, кадра управления или кадра PS-опрос.
- **Без состязания(CF)-конец.** Объявляет конец периода без состязания; часть стратегии использования точечной координационной функции.
- **CF-конец + CF-подтверждение.** Подтверждает кадр CF-конец. Данный кадр завершает период без состязания и освобождает станции от ограничений, связанных с этим периодом.

### Информационные кадры

Существует восемь подтипов информационных кадров, собранных в две группы. Первые четыре подтипа определяют кадры, переносящие данные высших уровней от исходной станции к станции-адресату. Перечислим эти кадры.

- **Данные.** Просто информационный кадр. Может использоваться как в период состязания, так и в период без состязания.
- **Данные + CF-подтверждение.** Может передаваться только в период без состязания. Помимо данных в этом кадре имеется подтверждение полученной ранее информации.
- **Данные + CF-опрос.** Используется точечным координатором для доставки данных к мобильной станции и для запроса у мобильной станции информационного кадра, который находится в ее буфере.
- **Данные + CF-подтверждение + CF-опрос.** Объединяет в одном кадре функции двух описанных выше кадров.

Остальные четыре подтипа информационных кадров фактически не переносят данных пользователя. Информационный кадр “нулевая функция” не переносит ни данных, ни запросов, ни подтверждений. Он используется только для передачи точке доступа бита управления питанием в поле управления кадром, указывая, что станция перешла в режим работы с пониженным энергопотреблением. Оставшиеся три кадра (CF-подтверждение, CF-опрос, CF-подтверждение + CF-опрос) имеют те же функции, что и описанные выше подтипы кадров (данные + CF-подтверждение, данные + CF-опрос, данные + CF-подтверждение + CF-опрос), но не несут пользовательских данных.

### Кадры управления

Кадры управления используются для управления связью станций и точек доступа. Возможны следующие подтипы.

- **Запрос ассоциации.** Посылается станцией к точке доступа с целью запроса ассоциации с данным BSS. Кадр включает информацию о возможностях, например, будет ли использоваться шифрование, или способна ли станция отвечать при опросе.
- **Ответ на запрос ассоциации.** Возвращается точкой доступа и указывает, что запрос ассоциации принят.

- **Запрос повторной ассоциации.** Посылается станцией при переходе между BSS, когда требуется установить ассоциацию с точкой доступа в новом BSS. Использование повторной ассоциации, а не просто ассоциации позволяет новой точке доступа договариваться со старой о передаче информационных кадров по новому адресу.
- **Ответ на запрос повторной ассоциации.** Возвращается точкой доступа и указывает, что запрос повторной ассоциации принят.
- **Пробный запрос.** Используется станцией для получения информации от другой станции или точки доступа. Кадр используется для локализации BSS стандарта IEEE 802.11.
- **Ответ на пробный запрос.** Отклик на пробный запрос.
- **Сигнальный кадр.** Передается периодически, позволяет мобильным станциям локализовать и идентифицировать BSS.
- **Объявление наличия трафика.** Посылается мобильной станцией с целью уведомления других (которые могут находиться в режиме пониженного энергопотребления), что в буфере данной станции находятся кадры, адресованные другим.
- **Разрыв ассоциации.** Используется станцией для аннуляции ассоциации.
- **Аутентификация.** Как объясняется ниже, для аутентификации станций используются множественные кадры.
- **Отмена аутентификации.** Передается для прекращения безопасного соединения.

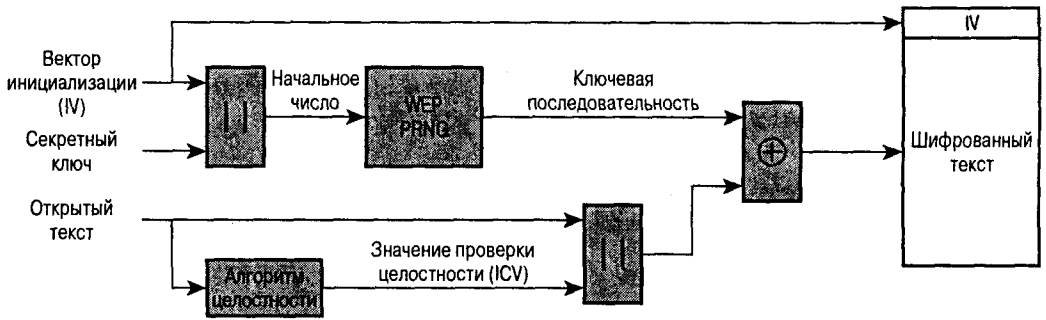
## Вопросы безопасности

Стандарт IEEE 802.11 предлагает и механизм конфиденциальности, и механизм аутентификации.

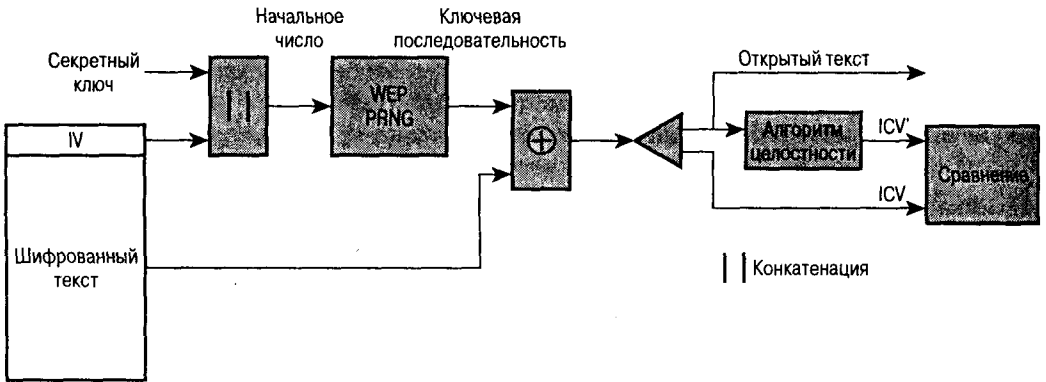
### Алгоритм конфиденциальности проводного эквивалента

В беспроводной локальной сети вопрос прослушивания имеет особую важность — ведь уловить передачу так просто! Для обеспечения современного уровня безопасности стандарт IEEE 802.11 включает схему WEP. Для обеспечения конфиденциальности (а также целостности данных) используется алгоритм, основанный на алгоритме шифрования RC4.

На рис. 14.9 представлен процесс шифрования. Алгоритм обеспечения целостности — это простая 32-битовая последовательность циклической проверки четности с избыточностью (CRC), присоединяемая к концу кадра MAC (рис. 14.8, а). Для процесса шифрования 40-битовый секретный ключ делится между двумя общающимися сторонами. К секретному ключу присоединяется вектор инициализации (IV). Получившийся блок — это начальное число генератора псевдослучайной последовательности (PRNG), определенного в RC4. Генератор создает последовательность битов, длина которой равна длине кадра MAC плюс CRC. Побитовое применение операции исключающего ИЛИ к кадру MAC и псевдослучайной последовательности дает зашифрованный текст. К данному тексту присоединяется вектор инициализации, и результат передается. Вектор инициализации периодически меняется (при каждой новой передаче), следовательно, меняется и псевдослучайная последовательность, что усложняет задачу расшифровки перехваченного текста.



а) Шифрование



б) Дешифрование

Рис. 14.9. Формат кадра MAC IEEE 802.11

После получения сообщения (рис. 14.9, б) приемник извлекает вектор инициализации и присоединяет его к совместно используемому секретному ключу, после чего генерирует ту же псевдослучайную последовательность, что и источник. К полученному таким образом ключу и поступившим данным побитово применяется операция исключающего ИЛИ, результатом которой является исходный текст. Данный алгоритм основан на следующем свойстве исключающего ИЛИ:

$$A \oplus B \oplus B = A.$$

Таким образом, если взять исходный текст, применить к нему и ключевой последовательности операцию исключающего ИЛИ, а затем применить операцию исключающего ИЛИ к результату и той же ключевой последовательности, то в итоге получится исходный текст. В заключение приемник сравнивает поступившую последовательность CRC и CRC, вычисленную по восстановленным данным: если величины совпадают, данные считаются неповрежденными.

### Аутентификация

Стандарт IEEE 802.11 предлагает два типа аутентификации: “открытая система” и “общий ключ”. Аутентификация открытых систем просто позволяет двум сторонам договориться о передаче данных без рассмотрения вопросов безо-

пасности. В этом случае одна станция передает другой управляющий кадр MAC, именуемый кадром аутентификации. В данном кадре указывается, что имеет место аутентификация открытых систем. Другая сторона отвечает собственным кадром аутентификации — и процесс завершен. Таким образом, при аутентификации открытых систем стороны просто обмениваются информацией о себе.

**Аутентификация с общим ключом** требует, чтобы две стороны совместно владели секретным ключом, не доступным третьей стороне. Процедура аутентификации между двумя сторонами, А и В, выглядит следующим образом.

1. А посылает кадр аутентификации, в котором указан тип “общий ключ” и идентификатор станции, определяющий станцию-отправителя.
2. В отвечает кадром аутентификации, который включает 128-октетный *текст запроса*. Текст запроса создается с использованием генератора случайных чисел WEP. Ключ и вектор инициализации, используемые при генерации текста запроса, не важны, поскольку далее в процедуре они не используются.
3. А передает кадр аутентификации, который включает полученный от В текст запроса. Кадр шифруется с использованием схемы WEP.
4. В получает зашифрованный кадр и дешифрует его, используя WEP и секретный ключ, которым владеют А и В. Если дешифрование прошло успешно (совпали CRC), В сравнивает принятый текст запроса с текстом, который был послан на втором этапе процедуры. После этого В передает А сообщение аутентификации, содержащее код состояния (успех или неудача).

## 14.4. ФИЗИЧЕСКИЙ УРОВЕНЬ IEEE 802.11

Спецификация физического уровня стандарта IEEE 802.11 выпускалась в три этапа: первая часть увидела свет в 1997 году, остальные две — в 1999. Первая часть, именуемая просто IEEE 802.11, включала описание уровня MAC и три спецификации физического уровня — две в диапазоне 2,4 ГГц и одну в инфракрасном диапазоне, — работающие при скоростях 1 и 2 Мбит/с. Спецификация IEEE 802a — это полоса 5 ГГц и скорость до 54 Мбит/с; IEEE 802.b — 2,4 ГГц и 5,5 и 11 Мбит/с, соответственно. Все спецификации последовательно изучаются ниже.

### Исходный физический уровень IEEE 802.11

В исходном стандарте 802.11 были определены три физические среды передачи.

- Спектр, расширенный методом прямой последовательности. Полоса ISM 2,4 ГГц, скорость передачи данных 1 или 2 Мбит/с.
- Спектр, расширенный методом скачкообразной перестройки частоты. Полоса ISM 2,4 ГГц, скорость передачи данных 1 или 2 Мбит/с.
- Инфракрасный диапазон (длина волны 850–950 нм). Скорость передачи данных 1 или 2 Мбит/с.

Ключевые элементы спецификаций перечислены в табл. 14.4.

**Таблица 14.4. Спецификации физического уровня IEEE 802.11**

<b>а) Расширение спектра методом прямой последовательности</b>					
<b>Скорость передачи данных</b>	<b>Длина раздробленного кода</b>	<b>Модуляция</b>	<b>Скорость передачи символов</b>	<b>Бит/символ</b>	
1 Мбит/с	11 (последовательность Баркера)	DBPSK	1 млн. символов/с	1	
2 Мбит/с	11 (последовательность Баркера)	DQPSK	1 млн. символов/с	2	
5,5 Мбит/с	8 (ССК)	DBPSK	1,375 млн. символов/с	4	
11 Мбит/с	8 (ССК)	DQPSK	1,375 млн. символов/с	8	
<b>б). Расширение спектра методом скачкообразной перестройки частоты</b>					
<b>Скорость передачи данных</b>	<b>Модуляция</b>	<b>Скорость передачи символов</b>	<b>Бит/символ</b>		
1 Мбит/с	Двухуровневая GFSK	1 млн. символов/с	1		
2 Мбит/с	Четырехуровневая GFSK	1 млн. символов/с	2		
<b>в). Передача в инфракрасном диапазоне</b>					
<b>Скорость передачи данных</b>	<b>Модуляция</b>	<b>Скорость передачи символов</b>	<b>Бит/символ</b>		
1 Мбит/с	16-PPM	4 млн. символов/с	0,25		
2 Мбит/с	4-PPM	4 млн. символов/с	0,5		
<b>г) Ортогональное частотное уплотнение (OFDM)</b>					
<b>Скорость передачи данных</b>	<b>Модуляция</b>	<b>Степень кодирования</b>	<b>Кодированных битов на поднесущую</b>	<b>Битов кода на символ OFDM</b>	<b>Битов данных на символ OFDM</b>
6 Мбит/с	BPSK	1/2	1	48	24
9 Мбит/с	BPSK	3/4	1	48	36
12 Мбит/с	QPSK	1/2	2	6	48
18 Мбит/с	QPSK	3/4	2	96	72
24 Мбит/с	16-QAM	1/2	4	192	96
36 Мбит/с	16-QAM	3/4	4	192	144
49 Мбит/с	64-QAM	2/3	6	288	192
54 Мбит/с	16-QAM	3/4	6	288	216

**Расширение спектра методом прямой последовательности (DSSS)**

В системе DSSS может использоваться до семи каналов со скоростью передачи в каждом 1 или 2 Мбит/с. Количество доступных каналов зависит от ширины полосы, выделяемой конкретным государственным органом регулирования. В Европе, например, доступно до 13 каналов, а в Японии — только один. Ширина полосы каждого канала равна 5 МГц, схема кодирования: DBPSK — для скорости 1 Мбит/с и DQPSK — для скорости 2 Мбит/с.



Как отмечалось в главе 7, система DSSS использует раздробленный код, или псевдослучайную последовательность, для расширения скорости передачи данных, а следовательно, полосы сигнала. В стандарте IEEE 802.11 определено использование последовательности Баркера.

Последовательность Баркера (Barker sequence) — это последовательность  $\{s(t)\}$  элементов  $+1$  и  $-1$ . Длина последовательности равна  $n$ , а для автокорреляционного значения  $R(\tau)$  справедливо следующее:  $|R(\tau)| \leq 1$  для всех  $|\tau| \leq (n-1)$ . Более того, данное свойство справедливо для таких преобразований (и их сочетаний):

$$s(t) \rightarrow -s(t) \quad s(t) \rightarrow (-1)^l s(t) \quad \text{и} \quad s(t) \rightarrow -s(n-1-t).$$

В настоящее время известны такие последовательности Баркера:

$n = 2$	++
$n = 3$	++-
$n = 4$	+++-
$n = 5$	+++ - +
$n = 7$	+++ - - + -
$n = 11$	+ - + + - + + - - -
$n = 13$	+++ + - - + + - - +

В 802.11 используется последовательность Баркера  $n = 11$ . Таким образом, каждая двоичная единица отображается в последовательность  $\{+ - + + - + + - - -\}$ , а каждый двоичный нуль — в последовательность  $\{- + - - + - - - + +\}$ .

Важной особенностью последовательностей Баркера является их устойчивость к интерференции и эффектам многолучевого распространения.

### Расширение спектра методом скачкообразной перестройки частоты (FHSS)

В главе 7 указывалось, что система FHSS использует множественные каналы, причем перестройка с одного канала на другой выполняется на основе псевдослучайной последовательности. В схеме IEEE 802.11 используются каналы шириной 1 МГц. Число каналов колеблется от 23 (Япония) до 70 (США).

Параметры схемы FHSS стандартом не задаются. В США, например, минимальная скорость перестройки частоты составляет 2,5 раза в секунду. Минимальное расстояние перехода (по частоте) в Северной Америке и большей части Европы составляет 6 МГц, в Японии — 5 МГц.

В качестве схемы модуляции для скорости 1 Мбит/с используется двухуровневая гауссова частотная манипуляция (GFSK). Двоичные нуль и единица кодируются как отклонение от текущей несущей частоты. Для скорости 2 Мбит/с используется четырехуровневая GFSK, в которой четыре различных отклонения от несущей представляют четыре 2-битовые комбинации нулей и единиц.

### Передача в инфракрасном диапазоне

Схема передачи в инфракрасном диапазоне, как она определена в IEEE 802.11, является ненаправленной (рис. 13.6). Связь возможна на расстоянии до 20 м. В качестве схемы модуляции для скорости 1 Мбит/с выбрана фазоимпульсная модуляция (pulse position modulation — PPM): каждая группа, состоящая из 4 бит данных, отображается в один из 16 символов PPM; каждый символ

представляет собой строку, включающую 16 бит, причем каждая такая строка состоит из 15 нулей и одной двоичной единицы. При скорости 2 Мбит/с каждая группа, состоящая из 2 бит, отображается в одну из четырех 4-битовых последовательностей, каждая из которых содержит три нуля и одну двоичную единицу. При передаче используется схема модуляции интенсивности, когда присутствие сигнала обозначает двоичную единицу, а отсутствие — двоичный нуль.

## IEEE 802.11a

Спецификация IEEE 802.11a использует полосу 5 ГГц. В отличие от спецификации 2,4 ГГц, здесь применяется не расширенный спектр, а ортогональное частотное уплотнение (OFDM). Напомним (см. раздел 11.2), что при OFDM, также именуемом модуляцией со многими несущими, используется несколько несущих сигналов на различных частотах, посредством каждого из которых осуществляется передача части битов. Данная схема подобна частотному уплотнению (FDM), отличие состоит лишь в том, что все подканалы выделены одному источнику.

Стандарт IEEE 802.11a определяет такие скорости передачи данных: 6, 9, 12, 18, 24, 36, 48 и 54 Мбит/с. Имеется до 52 поднесущих, которые модулируются с использованием схем BPSK, QPSK, 16-QAM или 64-QAM, в зависимости от требуемой скорости передачи. Расстояние между поднесущими составляет 0,3125 МГц. Для прямой коррекции ошибок используются сверточные код со степенью кодирования 1/2, 2/3 или 3/4.

## IEEE 802.11b

IEEE 802.11b — это расширение спецификации IEEE 802.11 DSSS, допускающее скорости передачи данных 5,5 и 11 Мбит/с. Скорость передачи раздробленного сигнала равна 11 МГц, т.е. такая же, как в исходной схеме DSSS, следовательно, обе схемы требуют одинаковой полосы. Для получения более высокой скорости при неизменной полосе и скорости передачи раздробленного сигнала используется манипуляция дополнительным кодом (complementary code keying — CCK).

Модуляция CCK является достаточно сложной и здесь рассматриваться не будет. Пример схемы CCK приводится на рис. 14.10 для скорости передачи 11 Мбит/с. Входные данные рассматриваются как 8-битовые блоки со скоростью 1,375 МГц ( $8 \text{ бит/символ} \times 1,375 \text{ МГц} = 11 \text{ Мбит/с}$ ). Шесть битов отображаются в одну из 64 кодовых последовательностей, при этом применяется матрица Уолша  $8 \times 8$  (рис. 7.17). Результат плюс оставшиеся два бита подаются на вход модулятора QPSK.

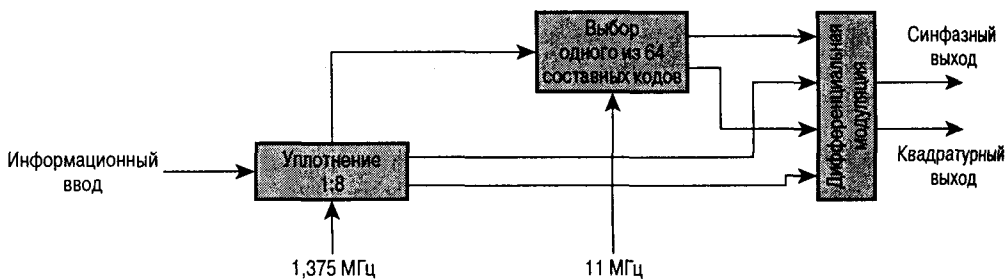


Рис. 14.10. Схема модуляции CCK для скорости 11 Мбит/с

## 14.5. РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА И WEB-САЙТЫ

Великолепной технической трактовкой IEEE 802.11 является работа [OHAR99]. Стандарты IEEE 802.11 также рассмотрены в [GEIE99], причем в этой работе подробно разбираются многочисленные практические ситуации. Хорошей обзорной статьей по стандартам 802.11 является [CROW97]. Ни в одной из двух последних работ не рассматриваются стандарты IEEE 802.11a и IEEE 802.11b. О стандарте IEEE 802.11a речь идет в работе [GEIE01].

CROW97 Crow B. et al. IEEE 802.11 Wireless Local Area Networks. — *IEEE Communications Magazine*, September 1997.

GEIE99 Geier J. *Wireless LANs*. — New York: Macmillan Technical Publishing, 1999.

GEIE01 Geier J. Enabling Fast Wireless Networks with OFDM. — *Communications System Design*, February 2001. ([www.csdmag.com](http://www.csdmag.com))

OHAR99 Ohara B., Petrick A. *IEEE 802.11 Handbook: A Designer's Companion*. — New York: IEEE Press, 1999.



### Рекомендуемые Web-сайты

- The IEEE 802.11 Wireless LAN Working Group ([grouper.ieee.org/groups/802/11/](http://grouper.ieee.org/groups/802/11/)). Документы рабочей группы плюс архивы дискуссий.
- Wireless Ethernet Compatibility Alliance ([www.wi-fi.org/](http://www.wi-fi.org/)). Промышленная группа, выступающая за способность к взаимодействию продуктов IEEE 802.11 друг с другом и с Ethernet.

## 14.6. ТЕРМИНЫ И ВОПРОСЫ

### Основные термины

алгоритм	манипуляция	расширенный набор
конфиденциальности	дополнительным	услуг (ESS)
проводного	кодом (ССК)	точечная
эквивалента (WEP)	модуль данных	координационная
аутентификация	протокола MAC	функция (PCF)
открытых систем	(MPDU)	точка доступа (AP)
аутентификация с	модуль данных службы	управление доступом к
общим ключом	MAC (MSDU)	среде (MAC)
базовый набор услуг	последовательность	управление логическим
(BSS)	Баркера	каналом (LLC)
двоичная	распределенная	
экспоненциальная	координационная	
отсрочка	функция (DCF)	
координационная	распределительная	
функция	система (DS)	

## Вопросы

1. Перечислите и кратко охарактеризуйте уровни протоколов IEEE 802.
2. Чем отличаются MAC-адрес и LLC-адрес?
3. Перечислите и кратко охарактеризуйте услуги LLC.
4. В чем различие между точкой доступа и порталом?
5. Является ли распределительная система беспроводной сетью?
6. Перечислите и кратко охарактеризуйте услуги IEEE 802.11.
7. Как связаны понятия ассоциации и мобильности?
8. С какими характеристиками беспроводной локальной сети связаны уникальные вопросы безопасности, не возникающие в проводной сети?
9. Какая аутентификация безопаснее и почему: аутентификация открытых систем или аутентификация с общим ключом?

# ГЛАВА 15

## BLUETOOTH

### 15.1. Обзор

Области применения Bluetooth

Стандарты Bluetooth

Архитектура протоколов

Модели использования

Пикосети и рассеянные сети

### 15.2. Радиоспецификация

### 15.3. Узкополосная спецификация

Перестройка частоты

Физические каналы

Пакеты

Исправление ошибок

Логические каналы

Управление каналом

Аудио Bluetooth

Безопасность Bluetooth

### 15.4. Спецификация администратора связей

### 15.5. Протокол управления логическим каналом и адаптации (L2CAP)

Каналы L2CAP

Пакеты L2CAP

Сигнальные команды

Качество обслуживания

### 15.6. Рекомендуемая литература и Web-сайты

### 15.7. Термины и вопросы

**В**luetooth<sup>1</sup> — это внедренное в микрочип невыключающееся радиоустройство ближнего действия. Эта технология была представлена шведским производителем мобильных средств связи Ericsson в 1994 году как средство, позволяющее портативным компьютерам совершать звонки по мобильным телефонам. С тех пор несколько тысяч компаний работают над тем, чтобы технология Bluetooth стала стандартом множества маломощных, действующих на близком расстоянии беспроводных устройств. Промышленные наблюдатели считают, что к 2005 году чипами Bluetooth будут оборудованы миллиарды аппаратов (*Business Week*, 18 September 2000).

Стандарты Bluetooth публикуются промышленным консорциумом Bluetooth SIG (Special Interest Group — специальная группа). Данная глава посвящена обзору этой технологии.

## 15.1. ОБЗОР

Цель Bluetooth — унифицировать возможности ближней радиосвязи. В диапазоне 2,4 ГГц (общедоступные нелицензируемые частоты для маломощных устройств) два аппарата Bluetooth, находящиеся на расстоянии до 10 м, могут совместно использовать пропускную способность до 720 Кбит/с. Bluetooth предназначена для поддержки многих приложений (полный список достаточно объемный и продолжает пополняться: передача данных, аудио, графики, видео и т.д.). Например, чип Bluetooth может внедряться в такие аудиоустройства, как наушники, беспроводные и обычные телефоны, домашние стереопроигрыватели и цифровые MP3-плееры. С помощью Bluetooth потребители могут делать следующее:

- звонить с беспроводного головного телефона, удаленно связанного с телефоном ячейки сотовой связи;
- соединять компьютеры с периферией (принтеры, клавиатуры, мыши), не используя кабель;
- подключать без проводов MP3-плееры к другим машинам с целью загрузки музыкальных файлов;
- организовывать домашние сети, чтобы пользователь-домосед мог, не двигаясь с места, наблюдать за кондиционером, микроволновой печью и тем, как его дети используют Internet;
- звонить домой для того, чтобы включить или выключить бытовой прибор, устройство аварийной сигнализации или просто понаблюдать, что происходит в отсутствие хозяина.

## Области применения Bluetooth

Технология Bluetooth предназначена для работы в среде со многими пользователями. В маленькую сеть, называемую пикосетью (piconet), могут объединяться

---

<sup>1</sup>Название “Bluetooth” (“голубой зуб”) произошло от прозвища датского короля Гарольда Блаатанда (Harald Blaaland), жившего в X веке нашей эры. В историю он вошел как человек, который принес христианство в Скандинавию и объединил разрозненные земли Дании и Норвегии. Логотип Bluetooth, которым помечаются устройства, поддерживающие данную технологию, представляет собой руническую запись инициалов короля.

до восьми устройств. Десять таких пикосетей могут сосуществовать в одном радиодиапазоне Bluetooth. Для обеспечения безопасности каждый канал связи кодируется и защищается от подслушивания и интерференции.

Bluetooth предусматривает поддержку трех основных областей применения с использованием беспроводной связи ближнего действия.

- **Точки доступа для ввода данных** (в том числе — посредством голоса). Bluetooth способствует передаче в реальном времени данных и речи, обеспечивая удобную беспроводную связь портативных и стационарных аппаратов связи.
- **Замена кабеля.** При наличии Bluetooth отпадает необходимость в многочисленных кабельных проводках, которые сопровождают практически все устройства связи. Соединение Bluetooth устанавливается мгновенно, причем связываемые устройства не обязательно должны находиться в пределах прямой видимости. Радиус охвата — порядка 10 м, а если использовать усилитель, эту величину можно довести до 100 м.
- **Организация эпизодических сетей.** Устройство, оснащенное чипом Bluetooth, может мгновенно устанавливать связь с другим устройством, находящимся в пределах области охвата.

Несколько примеров того, как можно использовать технологию Bluetooth, приведены в табл. 15.1.

**Таблица 15.1. Сценарии пользователя Bluetooth [НААР98]**

---

#### **Телефон “три в одном”**

Если вы в офисе — телефон работает как интерком (не нужно платить за услуги телефонии), дома это беспроводный телефон (оплачивается как стационарное устройство), а если вы перемещаетесь, его можно использовать как обычный мобильный телефон (это уже сотовая связь)

#### **Мост Internet**

Портативный ПК можно связать с Internet откуда угодно либо посредством мобильного телефона (беспроводное соединение), либо с помощью кабеля (PSTN, ISDN, JIBC, xDSL)

#### **Интерактивная конференция**

На встречах и конференциях информацию можно мгновенно распространять между всеми участниками. Кроме того, проектором можно управлять, и не имея проводного соединения

#### **Удаленный головной телефон**

Соедините головной телефон с вашим мобильным ПК или любым проводным соединением — и ваши руки свободны для более важных дел в офисе или в движущемся автомобиле

#### **Громкоговоритель портативного ПК**

Соедините беспроводный головной микрофон с вашим портативным ПК — и используйте его как микрофон в офисе, машине или дома

#### **Почта из портфеля**

Получите доступ к электронной почте, не вынимая портативный ПК из портфеля. Как только ПК получает почту, вам об этом сообщает ваш мобильный телефон. Используя тот же мобильный телефон, входящую почту можно просмотреть и прочитать сообщения

#### **Задержанные сообщения**

Если вы летите в самолете, электронную почту можно занести в ПК. Как только вы приземлитесь и сможете пользоваться мобильным телефоном, сообщения будут посланы немедленно

#### **Автоматическая синхронизация**

Автоматически синхронизируйте ваш настольный компьютер, портативный ПК, ноутбук и мобильный телефон. Как только вы войдете в офис, список адресов и календарь настольного компьютера будут автоматически обновлены согласно файлам вашего ноутбука (или наоборот)

---

**Мгновенная цифровая открытка**

Подключите (без проводов!) камеру к мобильному телефону или к любому выходу проводной связи. Введите комментарий с мобильного телефона, ноутбука или портативного ПК — и готовую открытку можно отправлять немедленно

**Беспроводный настольный компьютер**

Вашему компьютеру не нужны провода, чтобы соединиться с принтером, сканером, клавиатурой, мышью или локальной сетью

**Стандарты Bluetooth**

Стандарты Bluetooth — это фундаментальный труд: более 1500 страниц, содержащих технологическую (внутреннюю) спецификацию Bluetooth и спецификацию профиля. **Внутренние спецификации** описывают детали разнообразных уровней протокольной архитектуры Bluetooth (от радиointерфейса до управления каналом связи). Рассматриваются и родственные темы, такие, как возможность взаимодействия со сходными технологиями, требования к тестированию и определения разнообразных таймеров Bluetooth и связанных с ними значений.

В **спецификациях профиля** описано использование технологии Bluetooth для поддержки различных приложений. В каждой спецификации рассматривается применение технологии, определенной во внутренней спецификации, для реализации конкретной модели использования. В спецификации профиля указывается, какие аспекты внутренних спецификаций Bluetooth являются обязательными, необязательными и неприменимыми. Цель спецификации профиля — определить стандарт возможности взаимодействия, чтобы продукты разных производителей, заявленные как поддерживающие данную модель использования, могли действительно работать вместе. Грубо говоря, спецификации профиля относятся к одной из двух категорий: замена кабеля или беспроводное аудио. Профили замены кабеля дают близко расположенным устройствам удобное средство организации логического соединения и обмена данными. Например, если два аппарата попадают в диапазон взаимной досягаемости, они могут автоматически согласовать общий профиль. Возможно, в связи с этим пользователь будет предупрежден об установлении общего профиля или же произойдет автоматический обмен информацией. Профили беспроводного аудио занимаются созданием голосовых соединений ближнего действия.

На рис. 15.1 (взят из [WILS00]) изображена схема, по которой разработчик, интересующийся конкретным приложением, может найти нужный ему документ среди множества существующих. Список начинается с обзора некоторых необходимых внутренних спецификаций и профиля общего доступа. Этот профиль является одним из многих, составляющих основу других профилей, и не задает каких-либо самодостаточных функциональных возможностей. Профиль общего доступа указывает, как узкополосная архитектура Bluetooth (определенная во внутренних спецификациях) должна использоваться между устройствами, которые реализуют один или несколько профилей. После базового набора документов список на рисунке разделяется на два столбца, в одном из которых перечислены документы, относящиеся к замещению кабеля, а в другом — к беспроводному аудио.



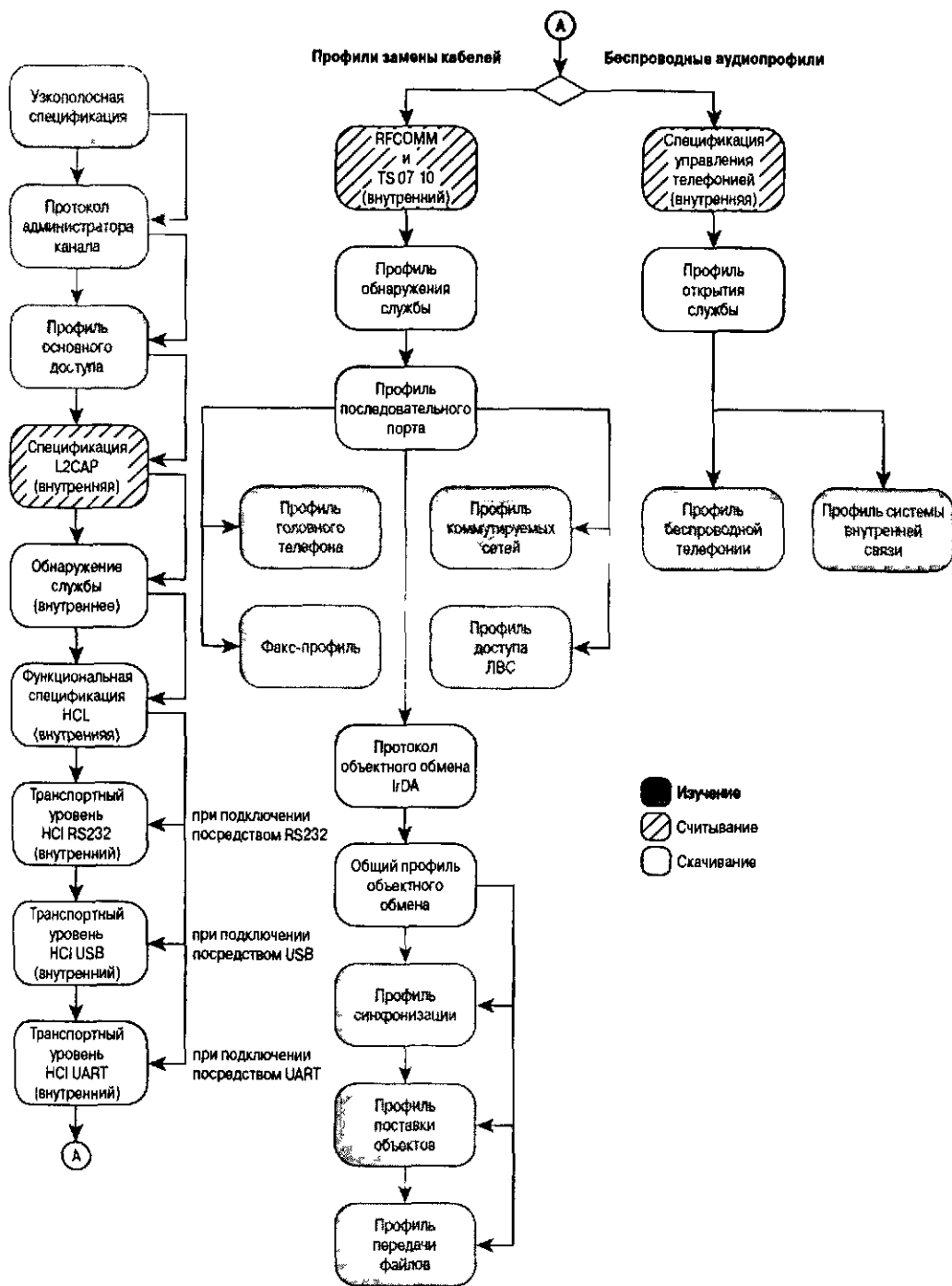
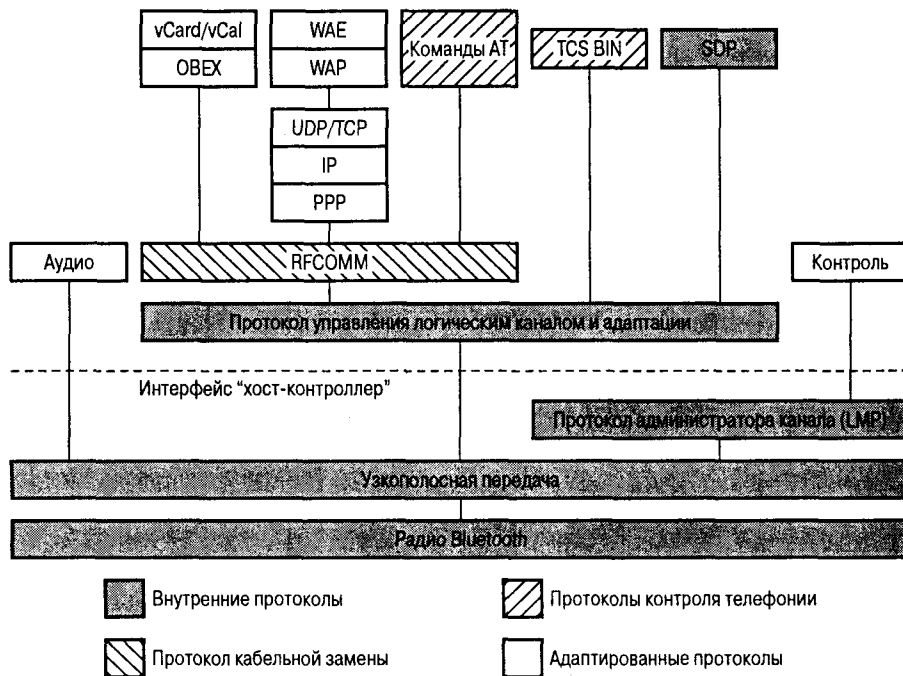


Рис. 15.1. Документация Bluetooth

## Архитектура протоколов

Bluetooth определяется как многоуровневая протокольная архитектура (рис. 15.2), состоящая из внутренних протоколов, протоколов замены кабеля и управления телефонией и адаптированных протоколов.



AT — сигнальная последовательность (префикс модема)  
 IP — протокол Internet  
 OBEX — протокол объектного обмена  
 PPP — протокол двухточечного соединения  
 RFCOMM — связь на радиочастотах  
 SDP — протокол обнаружения службы  
 TCP — протокол управления передачей

TCS BIN — спецификация управления телефонией (бинарная)  
 UDP — протокол пользовательских дейтаграмм  
 vCal — виртуальный календарь  
 vCard — виртуальная карта  
 WAE — среда беспроводных приложений  
 WAP — протокол беспроводных приложений

Рис. 15.2. Стек протоколов Bluetooth

**Внутренние протоколы** формируют пятиуровневый стек, состоящий из следующих элементов.

- **Радио.** Устанавливает детали радиоинтерфейса, включая используемые частоты, схему перестройки частоты, тип модуляции и мощность передачи.
- **Узкополосная передача.** На этом уровне решаются вопросы установления соединения в пределах пикосети, адресации, формата пакета и управления мощностью.
- **Протокол администратора канала связи (Link Manager Protocol — LMP).** Отвечает за установление канала связи между устройствами Bluetooth и его текущее администрирование, куда включается решение вопросов безопасности, таких, как аутентификация и шифрование, плюс контроль и согласование размеров узкополосных пакетов.

- **Протокол управления логическим каналом и адаптации (Logical Link Control and Adaptation Protocol — L2CAP).** Адаптирует протоколы высших уровней к уровню узкополосной передачи. L2CAP предлагает как службы без установления соединения, так и службы на основе соединения.
- **Протокол обнаружения службы (Service Discovery Protocol — SDP).** Для установления связи между двумя устройствами Bluetooth или большим числом устройств может запрашиваться информация об устройстве, службах и характеристиках служб.

В спецификацию Bluetooth включен еще один протокол: RFCOMM — **протокол замещения кабеля.** RFCOMM предоставляет виртуальный последовательный порт, задача которого — сделать замену кабельных технологий максимально незаметной для пользователей и программ. Последовательные порты — это один из наиболее общих типов интерфейсов связи, которые используются вычислительными устройствами и устройствами связи. Следовательно, RFCOMM позволяет замещать кабели последовательного порта с минимальной модификацией существующих устройств. RFCOMM обеспечивает транспортировку двоичных данных и эмулирует сигналы управления EIA-232 над узкополосным уровнем Bluetooth. EIA-232 (ранее известный как RS-232) — это широко используемый стандарт интерфейса последовательного порта.

В Bluetooth определен **протокол управления телефонией.** Протокол TCS BIN (telephony control specification — binary — спецификация управления телефонией — бинарная) — это протокол с битовой структурой, который определяет передачу сигналов управления вызовами с целью установления сеансов передачи речи и данных между устройствами Bluetooth. Кроме того, он определяет процедуры управления мобильностью для управления группами устройств Bluetooth TCS.

**Адаптированный протокол** определяется в спецификациях, выпускаемых другими организациями по стандартизации, и вводится в общую архитектуру Bluetooth. Стратегия Bluetooth заключается в создании только необходимых протоколов при максимально возможном использовании имеющихся стандартов. В число адаптированных протоколов входят следующие.

- **PPP.** Протокол двухточечного соединения — это стандартный протокол Internet для перемещения дейтаграммы IP по двухточечному каналу связи.
- **TCP/UDP/IP.** Фундаментальные протоколы из набора TCP/IP (описаны в главе 4).
- **OBEX.** Протокол объектного обмена — это протокол сеансового уровня, разработанный Ассоциацией передачи данных в инфракрасном диапазоне (Infrared Data Association — IrDA) для обмена объектами. Функции OBEX подобны функциям HTTP, но выполняются проще. Кроме того, данный протокол дает модель представления объектов и операций. Примерами форматов содержимого, передающегося посредством OBEX, являются vCard и vCalendar, которые предлагают формат электронных визитных карточек и персонального календаря/делового расписания, соответственно.
- **WAE/WAP.** Bluetooth вмещает в своей архитектуре среду беспроводных приложений и протокол беспроводных приложений (описаны в главе 12).

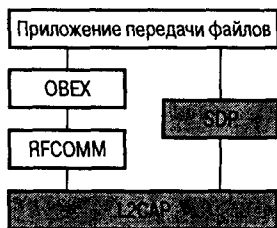
## Модели использования

В документах по профилям Bluetooth определены несколько моделей использования. По сути, модель использования — это набор протоколов, которые реализуют конкретное приложение на основе Bluetooth. Каждый профиль определяет протоколы и свойства протоколов, поддерживающие конкретную модель использования. На рис. 15.3 (взят из [МЕТТ99]) представлены модели использования с наивысшим приоритетом. Перечислим эти модели.

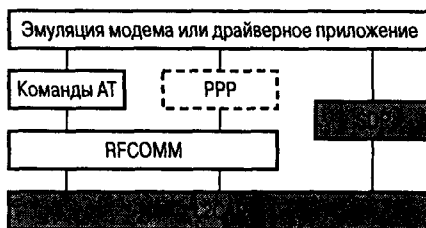
- **Передача файлов.** Модель поддерживает передачу каталогов, файлов, документов, изображений и потоковую информацию. Данная модель использования также содержит возможность просмотра папки с удаленного устройства.
- **Мост Internet.** Используя данную модель, ПК связывается без проводов с мобильным телефоном или беспроводным модемом для удаленного телефонного доступа к сети или факсу. Относительно удаленного доступа следует сказать, что управление мобильным телефоном или модемом осуществляется с помощью команд AT, а для передачи данных используется другой стек протоколов (например, PPP над RFCOMM). При факсимильной же связи программное обеспечение факса работает непосредственно над уровнем RFCOMM.
- **Доступ к локальной сети.** Данная модель использования позволяет устройствам пикосети получить доступ к локальной сети. После соединения работа устройства та же, что и при проводном подключении.
- **Синхронизация.** Данная модель обеспечивает синхронизацию содержащейся на устройствах персональной информации, такой, как записи в телефонной книге, календаре, сообщения и заметки. Здесь следует упомянуть IrMC (Ir mobile communications — мобильная связь в инфракрасном диапазоне), протокол IrDA, который позволяет передавать между устройствами обновленную персональную информацию (по схеме клиент/сервер).
- **Телефон “три в одном”.** Телефонные трубки, которые реализуют данную модель использования, могут работать как беспроводный телефон, подсоединенный к голосовой базовой станции, как интерком, связанный с другими телефонами, и как сотовый телефон.
- **Головной телефон.** Головной телефон может использоваться как устройство аудиоввода/вывода удаленного устройства.

## Пикосети и рассеянные сети

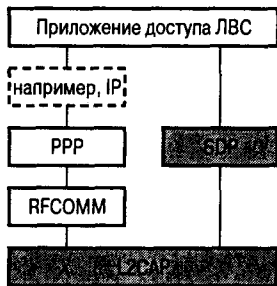
Ранее уже упоминалось, что основные элементы организации сетей Bluetooth — это пикосети, состоящие из ведущего устройства (master) и из 1–7 активных подчиненных устройств (slave). Радиоустройство, обозначенное как ведущее, определяет канал (последовательность перестройки частоты) и фазу (т.е. когда передавать), которые будут использоваться всеми устройствами данной пикосети. Эти параметры определяются ведущим устройством на основе собственного адреса. Подчиненное устройство может общаться только с ведущим, причем лишь тогда, когда это разрешает ведущее устройство. Любое устройство одной пикосети может также входить в другую пикосеть в качестве как подчиненного, так и ведущего (рис. 15.4). Данная схема с перекрытием называется **рассеянной сетью** (scatternet). На рис. 15.5 (на основе иллюстрации из [HARR00a]) архитектура пикосети/рассеянной сети сравнивается с другими формами беспроводных сетей.



а) Передача файлов



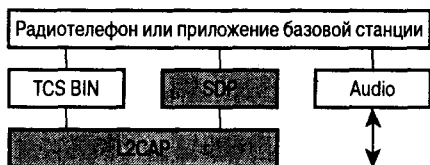
б) Коммутируемые сети



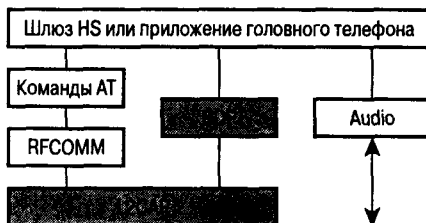
в) Доступ LAN



г) Синхронизация



д) Радиотелефон и система внутренней связи



е) Головной телефон

Рис. 15.3. Модели использования Bluetooth

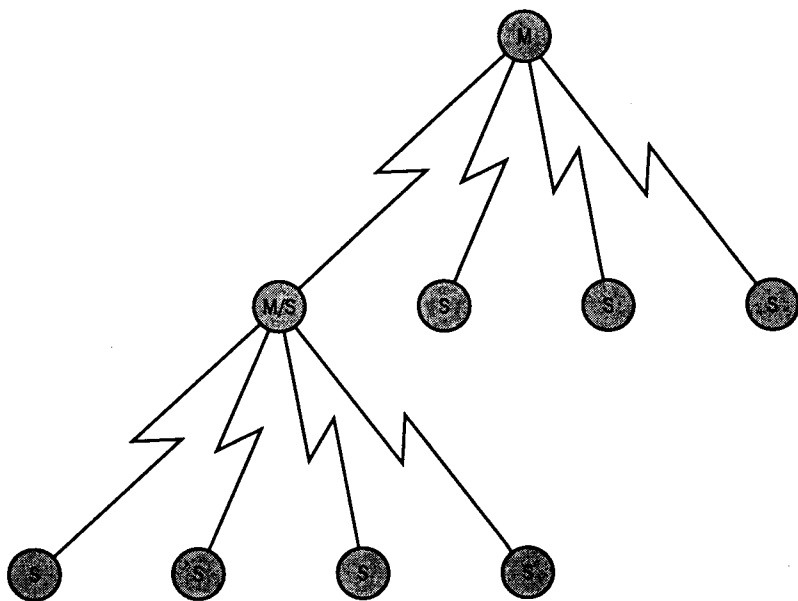
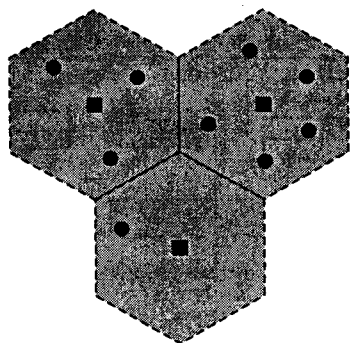
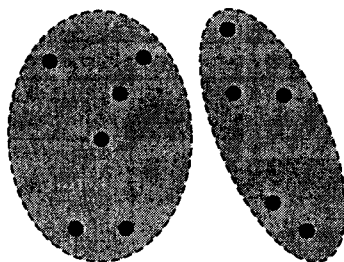


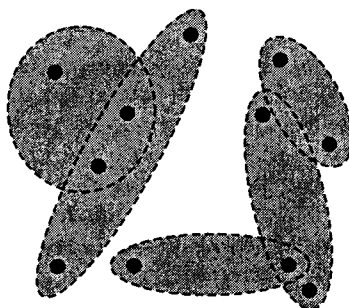
Рис. 15.4. Ведущие (M) и подчиненные (S) устройства



а) Сотовая система (квадраты представляют стационарные базовые станции)



б) Обычные эпизодические системы



в) Рассеянные сети

Рис. 15.5. Конфигурации беспроводной сети

Преимущество схемы пикосеть/рассеянная сеть состоит в том, что она позволяет многим аппаратам находиться в одной физической области и эффективно использовать общую полосу. В системе Bluetooth задействована схема перестройки частоты с расстоянием между несущими 1 МГц. Как правило, при общей полосе 80 МГц может быть до 80 различных частот. Если передача ведется на одной частоте, то каждый отдельный канал соответствует полосе 1 МГц. При перестройке частоты логический канал определяется последовательностью изменения частоты. В любой момент времени доступна полоса 1 МГц, а максимальное число устройств, совместно использующих полосу, — 8. В общей полосе 80 МГц могут одновременно функционировать различные логические каналы (с разными последовательностями перестройки). Если устройства из различных пикосетей, работающие в различных логических каналах, одновременно перестроятся на одну частоту, произойдет конфликт. С увеличением числа пикосетей в области и, соответственно, числа конфликтов падает производительность. Подытожим: в рассеянной сети совместно используются физическая область и суммарная полоса, а в пикосети — логический канал и возможность передачи данных.

## 15.2. РАДИОСПЕЦИФИКАЦИЯ

Радиоспецификация Bluetooth — это небольшой документ, в котором представлены основные детали радиопередачи устройств Bluetooth. Некоторые ключевые параметры приведены в табл. 15.2.

Таблица 15.2. Параметры Bluetooth

Топология	До 7 одновременных каналов в логической звездообразной сети
Модуляция	GFSK
Максимальная скорость передачи данных	1 Мбит/с
Ширина полосы радиочастот	220 кГц (по уровню -3 дБ), 1 МГц (по уровню -20 дБ)
Полоса радиочастот	2,4 ГГц, диапазон ISM
Число радионесущих	23/79
Расстояние между несущими	1 МГц
Мощность передачи	0,1 Вт
Схема доступа к пикосети	FH-TDD-TDMA
Скорость изменения частоты	1600 раз в секунду
Схема доступа к рассеянной сети	FH-CDMA

Одним из аспектов радиоспецификации является определение трех классов передатчиков на основе выходной мощности.

- **Класс 1.** Минимальная мощность выходного сигнала — 1 мВт (0 дБмВт); для передачи на максимальное расстояние — 100 мВт (+20 дБмВт). В этом классе контроль мощности является обязательным, мощность должна находиться в диапазоне 4–20 дБмВт. Данный режим обеспечивает наибольшее расстояние связи.

- **Класс 2.** Минимальная мощность выходного сигнала — 0,25 мВт (–6 дБмВт); для передачи на максимальное расстояние — 2,4 мВт (+4 дБмВт). Контроль мощности — по желанию.
- **Класс 3.** Наименьшая мощность. Номинальный выход — 1 мВт.

Bluetooth использует полосу 2,4 ГГц в диапазоне ISM (полоса частот для промышленного, научного и медицинского использования). В большинстве стран этой полосы достаточно для определения 79 физических каналов шириной 1 МГц (табл. 15.3). Чтобы аппараты не излучали больше мощности, чем требуется, используется контроль мощности. Алгоритм контроля мощности реализуется посредством протокола администратора канала связи между ведущим и подчиненными устройствами в пикосети.

**Таблица 15.3. Международное распределение частот Bluetooth**

Зона	Диапазон, ГГц	Радиоканалы
США, большая часть стран Европы и большинство других стран	2,4–2,4835	$f = (2,402 + n)$ МГц, $n = 0, \dots, 78$
Япония	2,471–2,497	$f = (2,473 + n)$ МГц, $n = 0, \dots, 22$
Испания	2,445–2,475	$f = (2,449 + n)$ МГц, $n = 0, \dots, 22$
Франция	2,4465–2,4835	$f = (2,454 + n)$ МГц, $n = 0, \dots, 22$

Для модуляции в устройствах Bluetooth выбрана гауссова частотная манипуляция, причем двоичная единица представляется положительным отклонением от центральной частоты, а двоичный нуль — отрицательным. Минимальное отклонение — 115 кГц.

## 15.3. УЗКОПОЛОСНАЯ СПЕЦИФИКАЦИЯ

Один из наиболее сложных документов Bluetooth — узкополосная спецификация, ключевые элементы которой рассмотрены ниже.

### Перестройка частоты

Перестройка частоты (Frequency Hopping — FH) в Bluetooth используется с двумя целями.

1. Обеспечить сопротивление интерференции и эффектам многолучевого пространства.
2. Предоставить некоторую схему множественного доступа соседствующим устройствам из разных пикосетей.

Опишем работу схемы FH. Общая полоса делится на 79 (почти во всех странах) физических каналов, каждый занимает полосу в 1 МГц. Перестройка частоты производится путем перехода из одного физического канала в другой, причем последовательность изменения частоты определяется псевдослучайной последовательностью. Все устройства отдельной пикосети используют одинаковые последовательности изменения частоты; далее такую по-



следовательность будем называть FH-каналом<sup>2</sup>. Скорость изменения частоты — 1600 раз в секунду, так что каждый физический канал занят в течение 0,625 мс. Каждый интервал в 0,625 мс называется слотом, все слоты нумеруются последовательно.

Радиоустройства Bluetooth сообщаются с использованием дуплекса с временным разделением (time division duplex — TDD). Напомним (см. главу 11), что TDD — это схема передачи по каналу, в которой в любой момент времени данные передаются только в одну сторону, а направления передачи чередуются. Поскольку среду пикосети совместно используют более двух устройств, как схема доступа используется TDMA. Таким образом, схему доступа к пикосети можно обозначить как FH-TDD-TDMA, а ее общий вид приведен на рис. 15.6. Здесь  $k$  обозначает номер слота, а  $f(k)$  — физический канал, выбранный на это время.

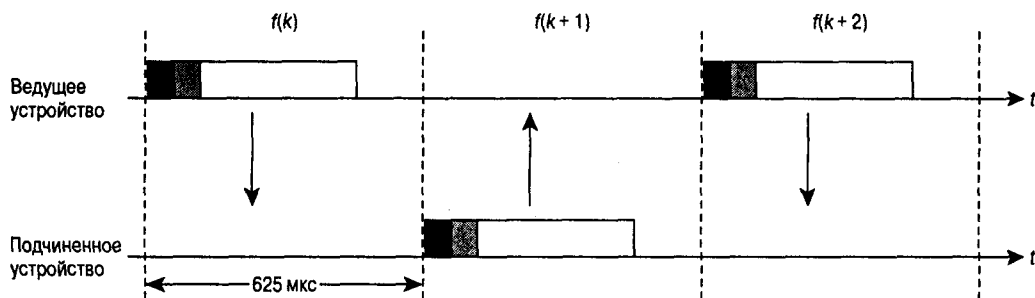


Рис. 15.6. Дуплекс с временным разделением (TDD) с перестройкой частоты (FH)

Передача пакета начинается в начале слота. Длина пакета может равняться 1, 3 или 5 слотам. Передача многослотового пакета ведется на одной частоте, без перестройки (рис. 15.7). В течение следующего слота передатчик переходит на частоту, определенную для данного периода времени последовательностью изменения частоты (таким образом, несколько частот последовательности пропускаются).

Использование TDD предотвращает перекрестные помехи между операциями приема и передачи в радиотрансивере, что просто необходимо, если каждое устройство должно иметь один чип Bluetooth. Поскольку передача и прием происходят в разных слотах, в них используются разные частоты.

Последовательность перестройки частоты определяется ведущим устройством пикосети и является функцией его Bluetooth-адреса. При этом для генерации этой последовательности используется довольно сложная математическая операция, включающая перестановку и выполнение операций исключающего ИЛИ.

Поскольку разные пикосети в одной области будут использовать разные ведущие устройства, отличаться будут и последовательности перестройки частоты. Таким образом, большую часть времени передачи двух устройств разных пикосетей в одной области будут идти по разным каналам. В некоторые моменты времени передачи в двух пикосетях будут вестись по одному физическому каналу, что приведет к конфликту и потере данных. Впрочем, такое случается довольно редко, и для страховки достаточно использовать схемы прямого исправления ошибок и выявления ошибок/запроса ARQ. Итак, некая форма множественного

<sup>2</sup>В документах Bluetooth такой термин не используется, здесь он введен только для того, чтобы облегчить обсуждение.

доступа с кодовым разделением каналов (code division multiple access — CDMA) успешно применяется к аппаратам разных пикосетей в одной рассеянной сети; именуется эта схема FH-CDMA.

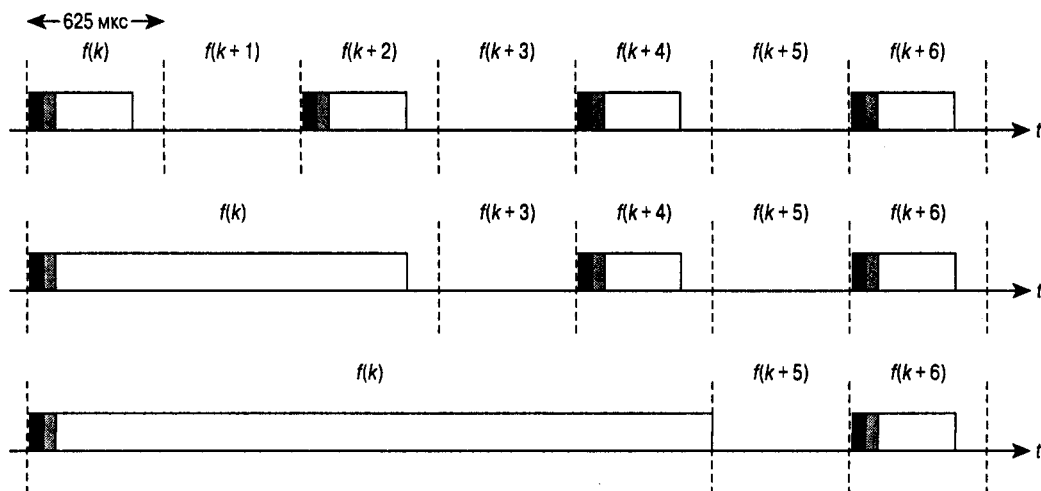


Рис. 15.7. Пример многословтовых пакетов

## Физические каналы

Между ведущим и подчиненным устройствами могут устанавливаться каналы связи двух типов.

- **Ориентированный на установление соединения синхронный канал (Synchronous Connection Oriented — SCO).** Выделяет фиксированную полосу двухточечному соединению, в котором задействованы ведущее устройство и одно из подчиненных устройств. Ведущее устройство поддерживает канал SCO путем использования через постоянные интервалы зарезервированных слотов. Основная единица резервирования — два последовательных слота (по одному в каждом направлении передачи). Ведущее устройство может одновременно поддерживать до трех каналов SCO, а подчиненное устройство — два или три канала SCO. Пакеты SCO никогда не передаются повторно.
- **Асинхронный канал без установления соединения (Asynchronous Connectionless — ACL).** Многоточечный канал между ведущим и всеми подчиненными устройствами пикосети. В слотах, не зарезервированных для каналов SCO, ведущее устройство может обмениваться пакетами с любым подчиненным устройством, в том числе и с устройствами, с которыми уже установлены каналы SCO. Может существовать только единственный канал ACL. Вследствие различных сбоев большинство пакетов ACL передаются повторно.

В основном каналы SCO используются для обмена срочными данными, требующими гарантированной скорости передачи, но не требующих гарантированной доставки. Например, в нескольких профилях Bluetooth используются оцифрованные аудиоданные, которым присуща допустимость потери данных. Гарантированная скорость передачи данных достигается посредством резервирования для передачи определенного числа слотов.

Каналы ACL обеспечивают соединения по принципу коммутации пакетов. Резервирование полосы невозможно, и доставка гарантируется посредством использования схем обнаружения ошибок и повторной передачи. Подчиненному устройству разрешается возвращать пакет ACL в слоте передачи от подчиненного устройства ведущему тогда и только тогда, когда он был передан в предыдущем слоте передачи от ведущего устройства подчиненному. Для каналов ACL определены 1-слотовые, 3-слотовые и 5-слотовые пакеты. Данные могут посылаться незащищенными (хотя на более высоких уровнях может использоваться схема ARQ) или защищенными кодом прямого исправления ошибок со степенью кодирования 2/3. Максимальная скорость передачи данных достигается при использовании 5-слотовых незащищенных пакетов и асимметричного распределения пропускной способности, это скорость 721 Кбит/с в прямом направлении и 57,6 Кбит/с — в обратном. Все существующие возможности приведены в табл. 15.4.

Таблица 15.4. Доступные скорости передачи данных по каналу ACL (Кбит/с)

Тип	Симметричная схема	Асимметричная схема	
DM1	108,8	108,8	108,8
DH1	172,8	172,8	172,8
DM3	256,0	384,0	54,4
DH3	384,0	576,0	86,4
DM5	286,7	477,8	36,3
DH5	432,6	721,0	57,6

DMx — x слотов, кодируется по схеме FEC

DHx — x слотов, незащищенная передача

## Пакеты

Формат всех пакетов Bluetooth показан на рис. 15.8. Он состоит из трех полей.

- **Код доступа.** Используется для временной синхронизации, компенсации сдвига, опроса и запроса.
- **Заголовок.** Используется для определения типа пакета и переноса управляющей информации протокола.
- **Полезная нагрузка.** При наличии этого поля в нем содержатся речь или данные пользователя и (в большинстве случаев) заголовок полезной нагрузки.

### Код доступа

Существует три типа кодов доступа.

- **Код доступа к каналу (channel access code — SAC).** Определяет пикосеть (уникален для пикосети).
- **Код доступа к устройству (device access code — DAC).** Используется для избирательного доступа и передачи последующих откликов.
- **Код доступа к опросу (inquiry access code — IAC).** Используется при опросах.

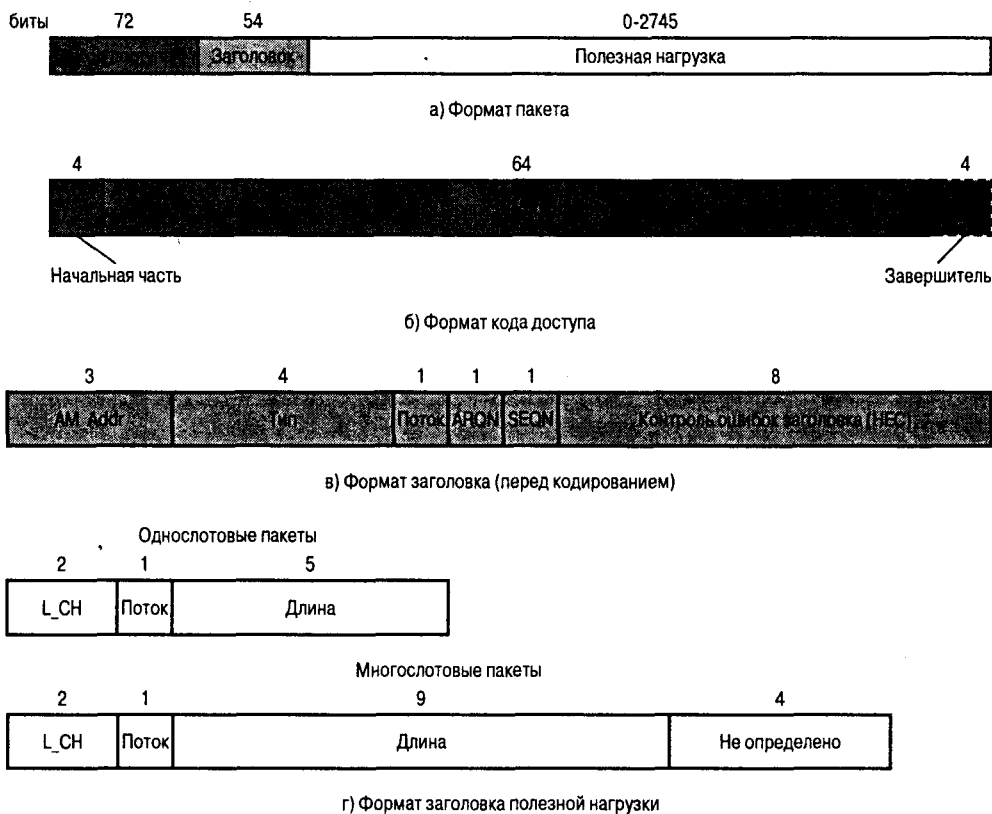


Рис. 15.8. Узкополосные форматы Bluetooth

В код доступа входят начальная комбинация битов, синхронизирующее слово и завершитель. **Начальная комбинация битов** (preamble) используется для компенсации постоянной составляющей сигнала передачи и состоит из последовательности 0101, если младший (крайний левый) бит в синхронизирующем слове — 0, и последовательности 1010, если младший бит в синхронизирующем слове — 1. Подобным образом, **завершитель** (trailer) равен 0101, если старший (крайний справа) бит в синхронизирующем слове — 1, и 1010, если старший бит в синхронизирующем слове — 0.

64-битовое **синхронизирующее слово** состоит из трех компонентов (рис. 15.9), и его стоит рассмотреть подробнее. Каждому устройству Bluetooth присваивается глобально уникальный 48-битовый адрес. 24 младших бита называются нижней частью адреса (lower address part — LAP) и используются для формирования синхронизирующего слова. Для кода SAC используется LAP ведущего устройства; для кода DAC — LAP устройства, к которому идет обращение. Существует два разных кода IAC. Общий код IAC (General IAC — GIAC) — это сообщение общего опроса, которое используется для выявления всех устройств Bluetooth в радиусе действия; для этого имеется специальное зарезервированное значение LAP. Выделенный IAC (Dedicated IAC — DIAC) принадлежит выделенной группе устройств Bluetooth, которые имеют общие характеристики, и в этом случае используется адрес LAP, ранее определенный для этих характеристик.

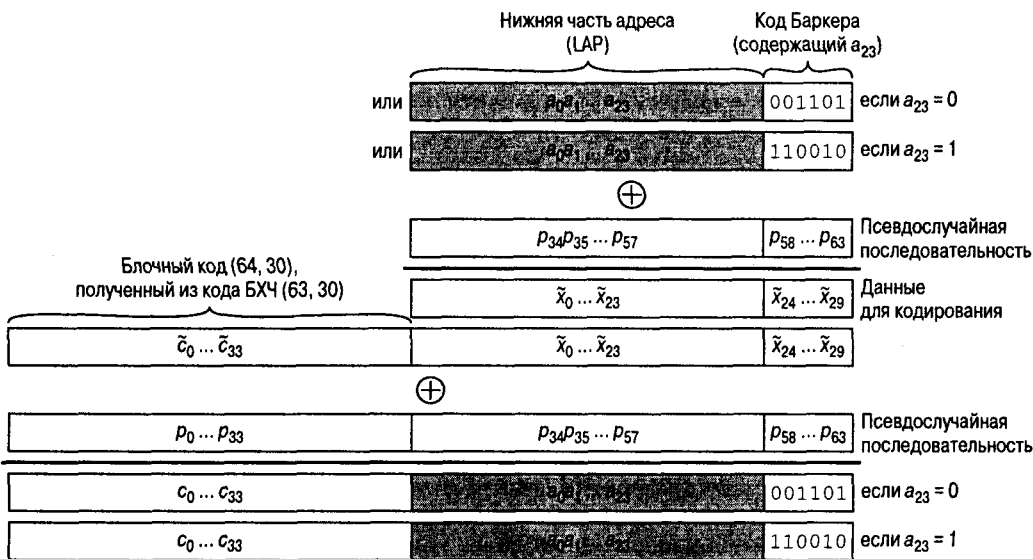


Рис. 15.9. Структура синхронизирующего слова

Используя соответствующий адрес LAP, синхронизирующее слово создается следующим образом.

1. К LAP добавляется 6 бит 001101, если старший бит LAP равен 0, и 110010, если старший бит — 1. В результате получается 7-битовая последовательность Баркера<sup>3</sup>. Цель введения последовательности Баркера — улучшить автокорреляционные свойства синхронизирующего слова.
2. Генерируется 64-битовая шумоподобная последовательность,  $p_0, p_1, \dots, p_{63}$ . Последовательность определяется уравнением  $P(X) = 1 + X + X^3$ , и ее можно реализовать с помощью линейного регистра сдвига с обратной связью. Начальное значение для генерации псевдослучайной последовательности — 100000<sup>4</sup>.
3. К последовательности  $p_{34}, p_{35}, \dots, p_{63}$  и 30-битовой последовательности, синтезированной на этапе 1, применяется операция исключающего ИЛИ, что позволяет скремблировать информацию, удаляя нежелательные регулярные последовательности.
4. Для скремблированного блока информации генерируется 34-битовый код с коррекцией ошибок, он помещается перед блоком, в результате чего формируется 64-битовое кодовое слово. Таким образом, имеем код (64, 30). Для генерации данного кода следует начать с кода БХЧ (63, 30)<sup>5</sup>. Потом определяется порождающий многочлен  $g(X) = (1 + X)g'(X)$ , где  $g'(X)$  — это порождающий многочлен кода БХЧ (63, 30). В результате имеем желаемый 34-битовый код.
5. К последовательностям  $p_0, p_1, \dots, p_{63}$  и 64-битовой последовательности, синтезированной на этапе 4, применяется операция исключающего ИЛИ. На этом шаге дешифруется информационная часть кодового слова, так что пе-

<sup>3</sup>Более подробно об этом см. в разделе 14.4.

<sup>4</sup>О генерации псевдослучайных последовательностей рассказывается в разделе 7.5.

<sup>5</sup>Подробнее о кодах БХЧ см. в разделе 8.2.

редаются исходные часть LAP и последовательность Баркера. На данном этапе также скремблируется блочный код.

Скремблирование информационной части кодового слова на этапе 3 происходит с целью усиления устойчивости блочного кода к ошибкам. Последующее дешифрование позволяет получателю легко восстанавливать LAP. С точки зрения спецификации, дешифрование 34-битового кода с коррекцией ошибок устраняет цикличность исходного кода, благодаря чему улучшаются спектральные характеристики передачи и автокорреляционные свойства.

### Заголовок пакета

Формат заголовка всех пакетов Bluetooth показан на рис. 15.8, *в*. Он состоит из трех полей.

- **AM\_ADDR.** Напомним, что пикосеть может содержать не более семи активных подчиненных устройств. 3-битовое поле AM\_Addr содержит адрес “активного режима” (временный адрес, присвоенный данному подчиненному устройству в данной пикосети) одного из подчиненных устройств. Передача от ведущего к подчиненному устройству содержит данный адрес подчиненного устройства, передача от подчиненного устройства — адрес этого подчиненного устройства. Значение 0 зарезервировано для передачи от ведущего устройства всем подчиненным устройствам в пикосети.
- **Тип.** Определяет тип пакета (табл. 15.5). Для контрольных пакетов зарезервированы четыре кода, общие для каналов SCO и ACL. Остальные типы пакетов используются для передачи информации пользователя. Пакеты HV1, HV2, HV3 в каналах SCO применяются для передачи речи со скоростью 64 Кбит/с. Отличаются эти пакеты степенью защиты от ошибок, что, в свою очередь, определяет, насколько часто должны посылаться пакеты для поддержания скорости передачи данных 64 Кбит/с. Пакет DV переносит и речь, и данные. Для каналов ACL определены 6 разных пакетов. Эти пакеты плюс пакет DM1 переносят пользовательские данные с разной степенью защиты от ошибок и разной скоростью передачи информации (табл. 15.4). Существует еще один тип пакета, общий для обоих физических каналов; он состоит только из кода доступа, а его размер равен 68 бит (не включая завершитель). Этот пакет обозначается ID и используется в процедурах опроса и доступа.
- **Поток.** Предоставляет 1-битовый механизм управления потоком только для трафика ACL. При приеме пакета с нулевым значением данного поля станция должна временно остановить передачу пакетов ACL по этому каналу. При приеме пакета со значением 1 передача может возобновляться.
- **ARQN.** Обеспечивает 1-битовый механизм подтверждения для трафика ACL, защищенного CRC (табл. 15.5). После успешного приема возвращается подтверждение ACK (ARQN = 1); в противном случае возвращается отрицательное подтверждение NAK (ARQN = 0). При отсутствии ответа предполагается отрицательное подтверждение, а если получено (или предполагается) отрицательное подтверждение, соответствующий пакет передается повторно.

Таблица 15.5. Типы пакетов Bluetooth

Код типа	Физический канал	Название	Число слотов	Описание
0000	Общий	NULL	1	Нет полезной нагрузки. Используется для возвращения источнику информации о канале в зависимости от успеха предыдущей передачи (ARQN) или о состоянии буфера приемника (Поток). Не подтверждается
0001	Общий	POLL	1	Нет полезной нагрузки. Используется ведущим устройством для опроса подчиненного устройства. Подтверждается
0010	Общий	FHS	1	Специальный управляющий пакет для определения адреса устройства и времени отправителя. Используется в ответе ведущего устройства на запрос, на опрос и при синхронизации перестройки частоты. Защита: FEC 2/3
0011	Общий	DM1	1	Поддерживает управляющие сообщения, может также переносить данные пользователя. Защита: 16-битовая CRC, FEC 2/3
0101	SCO	HV1	1	Переносится 10 байт информации; обычно используется для передачи речи со скоростью 64 Кбит/с. Защита: FEC 1/3
0110	SCO	HV2	1	Переносится 20 байт информации; обычно используется для передачи речи со скоростью 64 Кбит/с. Защита: FEC 2/3
0111	SCO	HV3	1	Переносится 30 байт информации; обычно используется для передачи речи со скоростью 64 Кбит/с. Не защищается FEC
1000	SCO	DV	1	Данные (150 бит) и речь (50 бит). Поле данных защищается схемой FEC 2/3
0100	ACL	DH1	1	Переносится 28 байт информации плюс 16-битовое поле CRC. Не защищается FEC. Обычно используется для передачи высокоскоростных данных
1001	ACL	AUX1	1	Переносится 30 байт информации без защиты CRC или FEC. Обычно используется для передачи высокоскоростных данных
1010	ACL	DM3	3	Переносится 123 байт информации плюс 16-битовое поле CRC. Защита: FEC 2/3
1011	ACL	DH3	3	Переносится 185 байт информации плюс 16-битовое поле CRC. Не защищается FEC
1110	ACL	DM5	5	Переносится 226 байт информации плюс 16-битовое поле CRC. Защита: FEC 2/3
1111	ACL	DH5	5	Переносится 341 байт информации плюс 16-битовое поле CRC. Не защищается FEC

- **SEQN**. Предоставляет 1-битовую схему последовательной нумерации. Передаваемые пакеты поочередно помечаются 1 или 0. Это нужно адресату для фильтрации повторных передач; если повторная передача вызвана потерей подтверждения АСК, адресат получает один пакет дважды.

- **Контроль ошибок в заголовке (Header Error Check — HEC).** 8-битовый код выявления ошибок. Используется для защиты заголовка пакета.

### Формат полезной нагрузки

Для некоторых типов пакетов узкополосная спецификация определяет формат поля полезной нагрузки. Если ею является речь, заголовок не определен. В то же время заголовок определяется для всех пакетов ACL и данных в пакете DV SCO. Формат полезной нагрузки для данных состоит из трех полей.

- **Заголовок полезной нагрузки.** 8-битовый заголовок определяется для пакетов из одного слота, и 16-битовый заголовок — для многослотовых пакетов.
- **Тело полезной нагрузки.** Содержит пользовательскую информацию.
- **CRC.** 16-битовый код проверки четности с избыточностью, который используется для защиты всех полей полезной нагрузки, исключая поля в пакетах AUX1.

Заголовок полезной нагрузки (если он есть) состоит из трех полей (рис. 15.8, г).

- **L\_CN.** Определяет логический канал (описаны ниже). Возможны следующие опции: сообщение LMP (11); нефрагментированное сообщение L2CAP или начало фрагментированного сообщения L2CAP (10); продолжение фрагментированного сообщения L2CAP (01); другое (00).
- **Поток.** Используется для управления потоком на уровне L2CAP. Аналогичен механизму включения/выключения, который предлагает поле потока в заголовке пакета трафика ACL.
- **Длина.** Число байтов данных в полезной нагрузке, исключая заголовок полезной нагрузки и CRC.

## Исправление ошибок

На узкополосном уровне Bluetooth использует три схемы исправления ошибок:

- FEC (forward error correction — прямое исправление ошибок) со степенью кодирования 1/3;
- FEC со степенью кодирования 2/3;
- ARQ (automatic repeat request — автоматический запрос повторной передачи).

Эти схемы исправления ошибок предназначены для удовлетворения противоречивых требований. Схемы исправления ошибок должно быть достаточно для организации связи по потенциально надежному беспроводному каналу, но при этом она должна быть простой и рациональной.

Схема FEC 1/3 используется в 18-битовом заголовке пакета и поле передачи речи в пакете HV1. Данная схема — это просто передача трех копий каждого бита. Используется мажоритарная логика: каких битов в принятой тройке больше, такая величина и считается переданной.

Схема FEC 1/2 используется во всех пакетах DM, в поле данных пакета DV, в пакете FHS и в пакете HV2. Кодер реализует разновидность кода Хэмминга с параметрами (15, 10). Этот код в каждом кодовом слове может исправлять все 1-битовые ошибки и обнаруживать все 2-битовые ошибки.



Схема ARQ используется с пакетами DM и DH, с полем данных пакета DV и подобна схемам ARQ, которые применяются в протоколах управления каналами данных (см. раздел 8.4). Напомним, что схемы ARQ имеют следующие элементы.

- **Выявление ошибок.** С помощью кода обнаружения ошибок CRC, дополняющегося кодом FEC, адресат выявляет ошибки и отбрасывает искаженные пакеты.
- **Положительное подтверждение (ACK).** Адресат возвращает положительное подтверждение по всем успешно полученным, безошибочным пакетам.
- **Повторная передача после истечения времени ожидания.** По истечении предопределенного промежутка времени источник повторно передает пакет, который не был подтвержден.
- **Отрицательное подтверждение (NAK) и повторная передача.** Адресат возвращает отрицательное подтверждение по пакету, в котором обнаружена ошибка, после чего источник передает этот пакет повторно.

В Bluetooth применяется так называемая схема *быстрого ARQ*, которая использует тот факт, что ведущее и подчиненное устройства связываются через разные временные слоты (рис. 15.10). Когда станция получает пакет, она (на основе 16-битового кода CRC) определяет, имеются ли ошибки. Если да, биты ARQN в заголовке присваивается значение 0 (NAK), если нет — значение 1 (ACK). В случае приема станцией сообщения NAK она повторно передает тот же пакет, что и в предыдущем слоте, с тем же значением бита SEQN в заголовке пакета. При описанном подходе о неудачной передаче отправитель извещается в следующем слоте и может попытаться повторить ее. Использование 1-битовых порядковых номеров и немедленная повторная передача пакета минимизирует служебные издержки и максимизирует быстроту реагирования.

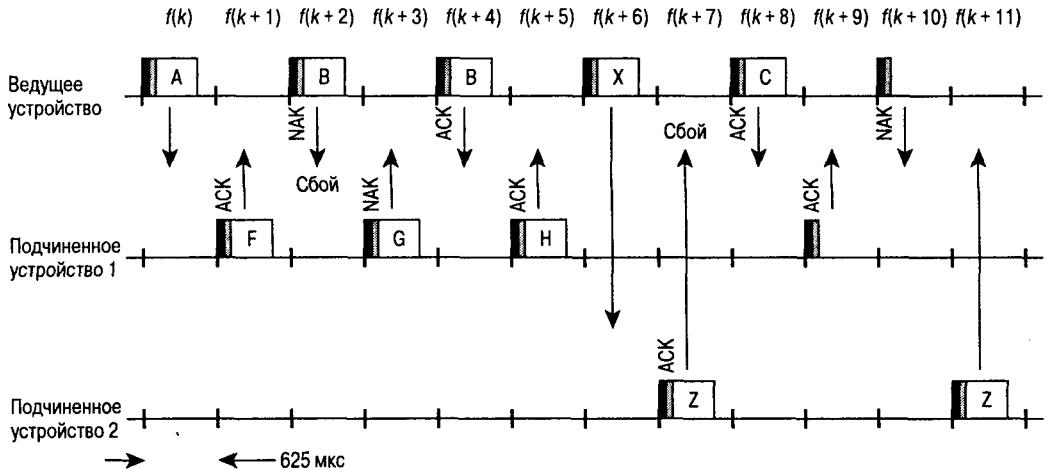
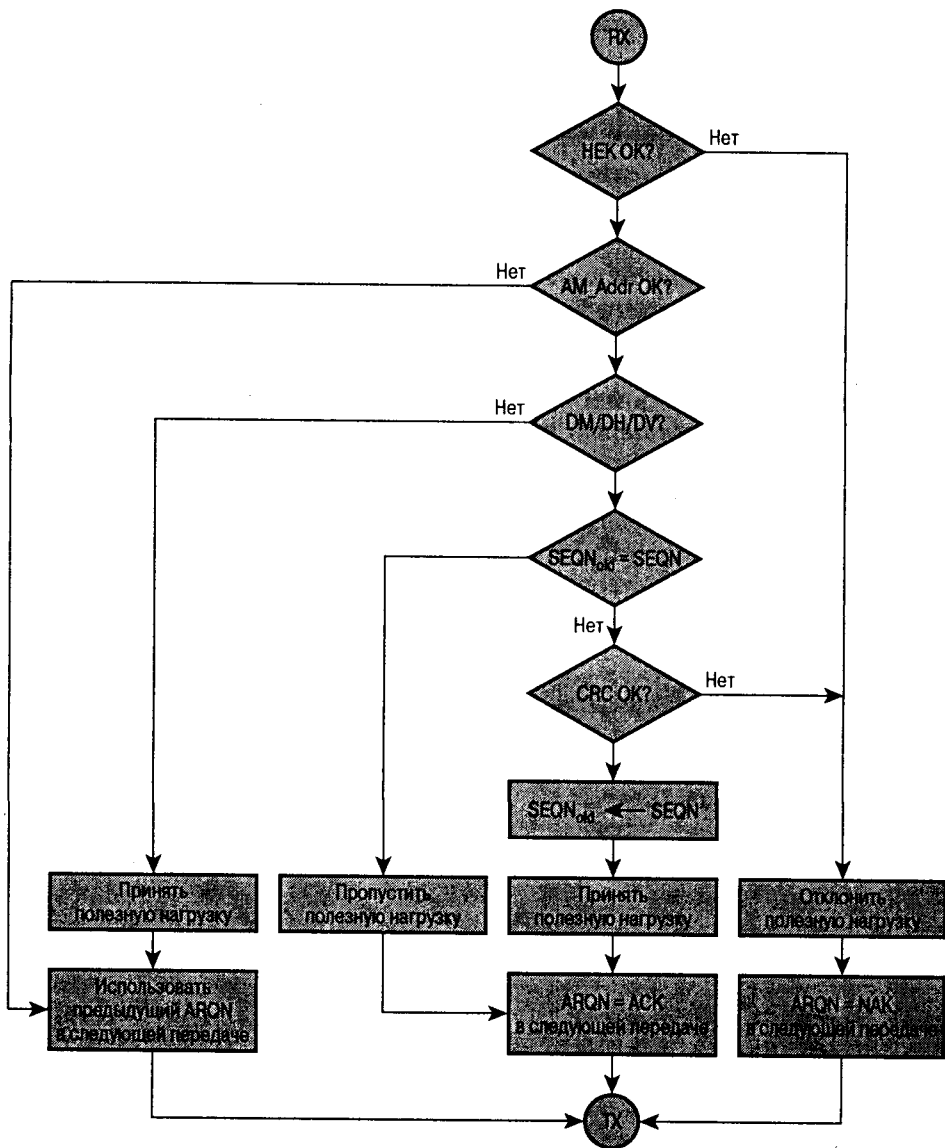


Рис. 15.10. Пример операции повторной передачи

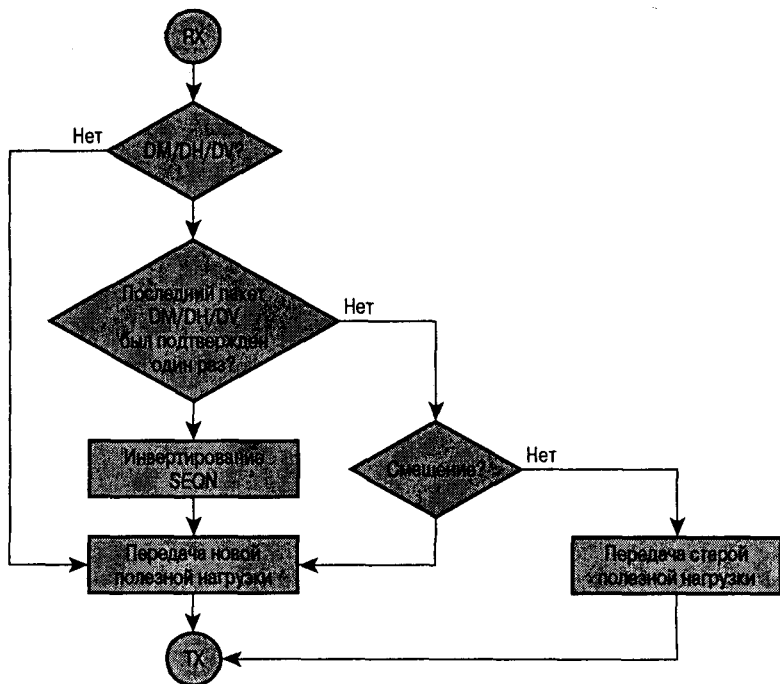
На рис. 15.11 механизм ARQ показан более подробно. После приема пакета (рис. 15.11, а) устройство, используя поле НЕС, проверяет, не поврежден ли заголовок. Если он поврежден, пакет отклоняется, а полю ARQN в следующем

слоте присваивается значение NAK. Далее устройство сопоставляет адреса и проверяет, используется ли в пакетах указанного типа механизм ARQ. После того как обе проверки будут успешно пройдены, устройство проверяет, не совпадает ли новый порядковый номер (SEQN) с предыдущим. Если они различны, устройство проверяет CRC. При отсутствии ошибок в следующем исходящем пакете значение поля ARQN устанавливается равным ACK, при сбое на любом этапе проверок значение поля устанавливается равным NAK.



а) Протокол получения для определения бита ARQN

Рис. 15.11. Схема ARQ в Bluetooth



б) Фильтрация повторной передачи для пакетов с CRC

Рис. 15.11. Схема ARQ в Bluetooth (продолжение)

Со стороны передатчика (рис. 15.11, б) следующий передаваемый пакет DM, DH или DV определяется значением предшествующего поля ARQN. Если принимаются поля со значениями ACK, устройство передает новую полезную нагрузку, в которой периодически чередуются значения SEQN = 0 и 1. При получении сообщения NAK или отсутствии подтверждения устройство будет передавать старую полезную нагрузку до тех пор, пока не будет получено сообщение ACK или пока число неудачных попыток повторной передачи не превысит некоторую пороговую величину. В последнем случае из буфера передачи удаляется старая полезная нагрузка и начинает передаваться новая.

## Логические каналы

В Bluetooth определено пять типов логических каналов передачи данных, предназначенных для переноса различных типов трафика.

- **Управление каналом (Link Control — LC).** Используется для управления потоком пакетов по интерфейсу канала. Канал LC отображается в заголовок пакета и переносит низкоуровневую управляющую информацию, относящуюся к схеме ARQ, управлению потоком и определению характеристик полезной нагрузки. Канал LC переносится во всех пакетах, кроме пакета ID, у которого нет заголовка.
- **Администратор канала (Link Manager — LM).** Используется для передачи информации по управлению каналом между станциями. Этот логический

канал поддерживает трафик LMP и может передаваться либо по каналу SCO, либо по каналу ACL.

- **Пользовательский асинхронный канал (User Asynchronous — UA).** Переносит асинхронные данные пользователя. Этот канал обычно переносится по каналу ACL, но может переноситься в пакетах DV по каналу SCO.
- **Пользовательский изохронный канал (User Isochronous — UI).** Переносит изохронные данные пользователя<sup>6</sup>. Этот канал обычно переносится по каналу ACL, но может также переноситься в пакетах DV по каналу SCO. На узкополосном уровне канал UI трактуется так же, как канал UA. Синхронизация для поддержки изохронности обеспечивается на высшем уровне.
- **Пользовательский синхронный канал (User synchronous — US).** Переносит синхронные данные пользователя по каналу SCO.

## Управление каналом

Работу пикосети можно понять исходя из состояний работы в процессе установления и эксплуатации канала (рис. 15.12). Существуют два основных состояния.

- **Холостое.** Состояние по умолчанию. Состояние с низким энергопотреблением, в котором работают только собственные часы устройства.
- **Соединение.** Устройство присоединяется к пикосети в качестве ведущего или подчиненного.

Кроме того, существуют семь промежуточных подсостояний, которые используются для введения в пикосеть новых подчиненных устройств. Для смены состояния используются либо команды от администратора Bluetooth, либо внутренние сигналы в контроллере канала. Перечислим возможные подсостояния.

- **Запрос.** Устройство выпустило запрос. Используется ведущим устройством для активизации подчиненного устройства и связи с ним. Ведущее устройство отправляет сообщение запроса путем передачи кода доступа к подчиненному устройству (device access code — DAC) в различных каналах, на которые может перестраиваться частота.
- **Поиск запроса.** Устройство ожидает запрос с его собственным кодом DAC.
- **Ответ ведущего устройства.** Устройство, выступающее как ведущее, получает от подчиненного устройства ответ на запрос. Теперь устройство может переходить в состояние соединения или возвращаться в состояние запроса для обращения к другим подчиненным.
- **Ответ подчиненного устройства.** Устройство, выступающее как подчиненное, отвечает на запрос ведущего. Если установление соединения прошло успешно, устройство входит в состояние соединения; в противном случае оно возвращается в состояние поиска запроса.

---

<sup>6</sup>Изохронными называются блоки данных, которые повторяются с известной периодичностью.

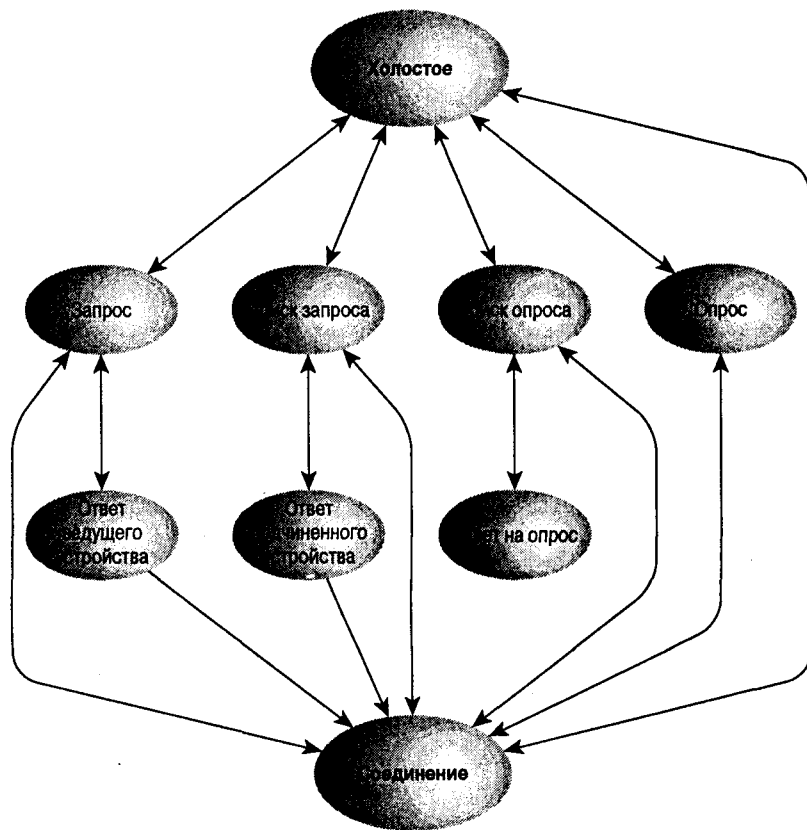


Рис. 15.12. Диаграмма переходов между состояниями Bluetooth

- **Опрос.** Устройство пытается определить, что находится в пределах его досягаемости.
- **Поиск опроса.** Устройство ожидает опрос.
- **Ответ на опрос.** Устройство, которое инициировало опрос, получает ответ на него.

### Процедура опроса

Первое, что должно сделать потенциальное ведущее устройство для организации пикосети, — определить, какие еще устройства находятся в пределах его досягаемости, там, где предполагается развертывание пикосети. Для этого устройство начинает процедуру опроса, призванную активизировать пользователя или приложение на устройстве. Процедура начинается с передачи потенциальным ведущим устройством пакета ID с кодом доступа к опросу (inquiry access code — IAC), который является общим для всех устройств Bluetooth. Напомним, что пакет ID не имеет ни заголовка, ни полезной нагрузки.

Из 79 радионесущих 32 считаются активизирующими (wake-up carriers). Ведущее устройство ретранслирует код IAC на всех 32 таких несущих по очереди. Само устройство при этом находится в состоянии опроса (рис. 15.12). Между

тем, устройства в холостом состоянии периодически переходят в состояние поиска опроса с целью поиска сообщений IAC на активизирующих радионесущих. Как только устройство получает это сообщение, оно переходит в состояние ответа на опрос и возвращает пакет FHS (табл. 15.5), который содержит адрес этого устройства и информацию по синхронизации, необходимую ведущему устройству для инициации подключения. Ведущее устройство не отвечает на пакет FHS и может остаться в состоянии опроса, пока не убедится, что найдены все устройства в пределах его досягаемости.

Как только устройство ответило на опрос, оно переходит в состояние поиска запроса и ожидает от ведущего устройства запрос на установление соединения. В то же время если в фазе ответа на опрос случится конфликт (два или более устройств одновременно ответят на опрос), не будет получен ни один запрос, устройство будет вынуждено вернуться в состояние поиска опроса и попытаться получить другие сообщения опроса и запроса<sup>7</sup>.

### Процедура запроса

Как только ведущее устройство нашло устройства в пределах своего диапазона, оно может установить соединение с каждым из них, формируя, таким образом, пикосеть. Для запроса каждого устройства ведущее устройство использует адрес данного устройства (BD\_ADDR), на основе которого вычисляет последовательность изменения частоты в передаче-запросе. Запрос осуществляется посредством ID-пакета, в котором указан код доступа (DAC) к требуемому устройству. Напомним, что DAC — это нижняя часть адреса подчиненного устройства. В ответ подчиненное устройство возвращает ведущему тот же пакет ID DAC на тех же частотах (последовательность их изменения именуется последовательностью режима запроса), что использовались ведущим устройством. После этого ведущее устройство в следующем слоте передает собственный пакет FHS, который содержит его адрес устройства и время. Далее подчиненное устройство еще раз посылает ответный пакет ID DAC, таким образом подтверждая получение пакета FHS. В этот момент подчиненное устройство переходит из состояния ответа подчиненного устройства в состояние соединения и начинает использовать последовательность перестройки частоты режима соединения, определенную ведущим устройством в пакете FHS. Тем временем ведущее устройство может запрашивать другие устройства, пока не соединится со всеми желаемыми подчиненными; после этого ведущее устройство переходит в состояние соединения.

### Состояние соединения

Для каждого подчиненного устройства состояние соединения начинается с отправки ведущим устройством пакета Poll (опрос), после которого подчиненное устройство синхронизируется с временем ведущего и начинает использовать ту же последовательность изменения частоты. Подчиненное устройство может ответить пакетом любого типа.

Как только подчиненное устройство перешло в состояние соединения, возможны четыре режима работы.

---

<sup>7</sup>Диаграмма состояний на рис. 15.12 взята из узкополосной спецификации, и на ней не показан переход из состояния ответа на опрос в состояние поиска запроса, но показан переход из состояния ответа на опрос в состояние поиска опроса. Впрочем, в тексте спецификации приведено то же описание, что и в данной книге.

- **Активный.** Подчиненное устройство активно участвует в работе пикосети, ожидая, передавая и принимая пакеты. Ведущее устройство периодически передает подчиненному устройству пакеты, на основе которых поддерживается синхронизация.
- **Подслушивание.** Подчиненное устройство прослушивает не все слоты приема, а только те, в которых могут находиться сообщения, предназначенные данному устройству. В течение других слотов подчиненное устройство может работать в режиме пониженного энергопотребления. Для передачи подчиненному устройству, находящемуся в состоянии подслушивания, ведущее устройство выделяет меньше слотов, чем обычно.
- **Удержание.** В этом состоянии устройство не поддерживает пакеты ACL и переходит в режим пониженного энергопотребления. Подчиненное устройство может по-прежнему участвовать в обменах по каналам SCO. В периоды неактивности подчиненное устройство может прослушивать сеть в состоянии пониженного энергопотребления или участвовать в работе другой пикосети.
- **Парковка.** Если от подчиненного устройства не требуется активного участия в работе пикосети, но оно должно оставаться ее частью, устройство может переходить в режим парковки — состояние пониженного энергопотребления с крайне малой активностью. Устройство теряет активный адрес (AM\_ADDR) и получает адрес элемента в режиме парковки (PM\_ADDR). Если в пикосети используется режим парковки, она может состоять более чем из семи подчиненных устройств.

## Аудио

В узкополосной спецификации указано, что могут использоваться две схемы шифрования речи: импульсно-кодовая модуляция (Pulse Code Modulation — PCM) или дельта-модуляция с непрерывной переменной огибающей (continuously variable slope delta — CVSD). Схему выбирают администраторы двух связанных устройств, исходя из конкретного приложения.

Модуляция PCM была рассмотрена в разделе 6.4. CVSD — это разновидность дельта-модуляции (delta modulation — DM), которая также рассматривалась в указанном разделе. Напомним, что при дельта-модуляции, аналоговые входные данные аппроксимируются ступенчатой функцией, значение которой изменяется на один уровень квантования ( $\delta$ ) на каждом интервале дискретизации ( $T_s$ ). Таким образом, для представления выхода дельта-модуляции достаточно одного бита на выборку. По сути, поток таких битов — это аппроксимация производной аналогового сигнала, а не его амплитуды: значение 1 генерируется, если ступенчатая функция увеличивает свое значение на следующем интервале дискретизации; в противном случае генерируется значение 0.

Ранее уже отмечалось, что в схемах DM существует два типа ошибок: шум квантования, имеющий место при очень медленном изменении сигнала, и шум перегрузки по крутизне — при очень быстром изменении сигнала (рис. 6.18). Схема CVSD предназначена для минимизации ошибок обоих типов путем применения разных уровней квантования: малых, когда сигнал меняется медленно, и больших, когда сигнал меняется быстро (рис. 15.13; основан на иллюстрации из [HAAR98]). Крутизна сигнала определяется на основе изучения  $K$  предшествую-

щих битов выходных данных. Получающаяся в результате схема устойчивее к ошибкам, чем PCM, и более устойчивая к ошибкам квантования и перегрузки по крутизне, чем DM.

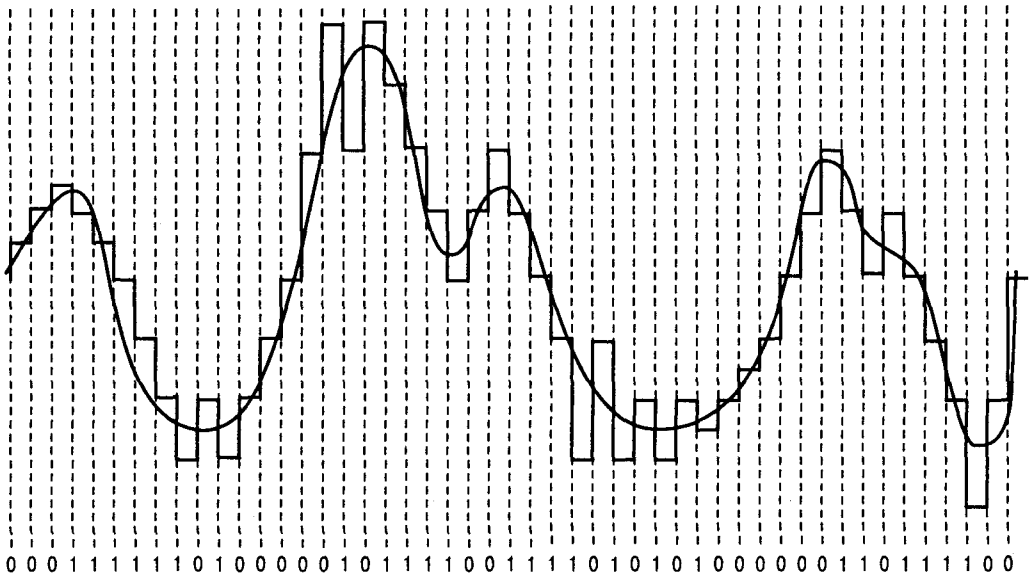


Рис. 15.13. Пример дельта-модуляции с непрерывной переменной огибающей

На рис. 15.14 показано кодирование и декодирование CVSD (сравните с рис. 6.19). Как и при дельта-модуляции, двоичный выход преобразуется в ступенчатую функцию, которая пытается максимально близко повторить оригинал. При кодировании происходит следующее. Вход — это выборки PCM, поступающие со скоростью 64 Кбит/с. В каждый момент дискретизации вход PCM сравнивается с предшествующим значением аппроксимирующей ступенчатой функции, которое представляется как  $\hat{x}(k-1)$ . Выход компаратора  $b(k)$  определяется следующим образом:

$$b(k) = \begin{cases} 1 & x(k) - \hat{x}(k) \geq 0 \\ -1 & x(k) - \hat{x}(k) < 0 \end{cases}$$

При передаче эти числа представляются битом знака (отрицательные числа отображаются в двоичную 1; положительные — в двоичный 0). Затем выход  $b(k)$  используется для определения амплитуды на следующем интервале ступенчатой функции,  $\delta(k)$ .

$$\delta(k) = \begin{cases} \min[\delta_{\min} + \delta(k-1), \delta_{\max}] & \text{если по крайней мере } J \text{ из } K \text{ последних битов равны} \\ \max[\beta \times \delta(k-1), \delta_{\max}] & \text{в противном случае} \end{cases}$$



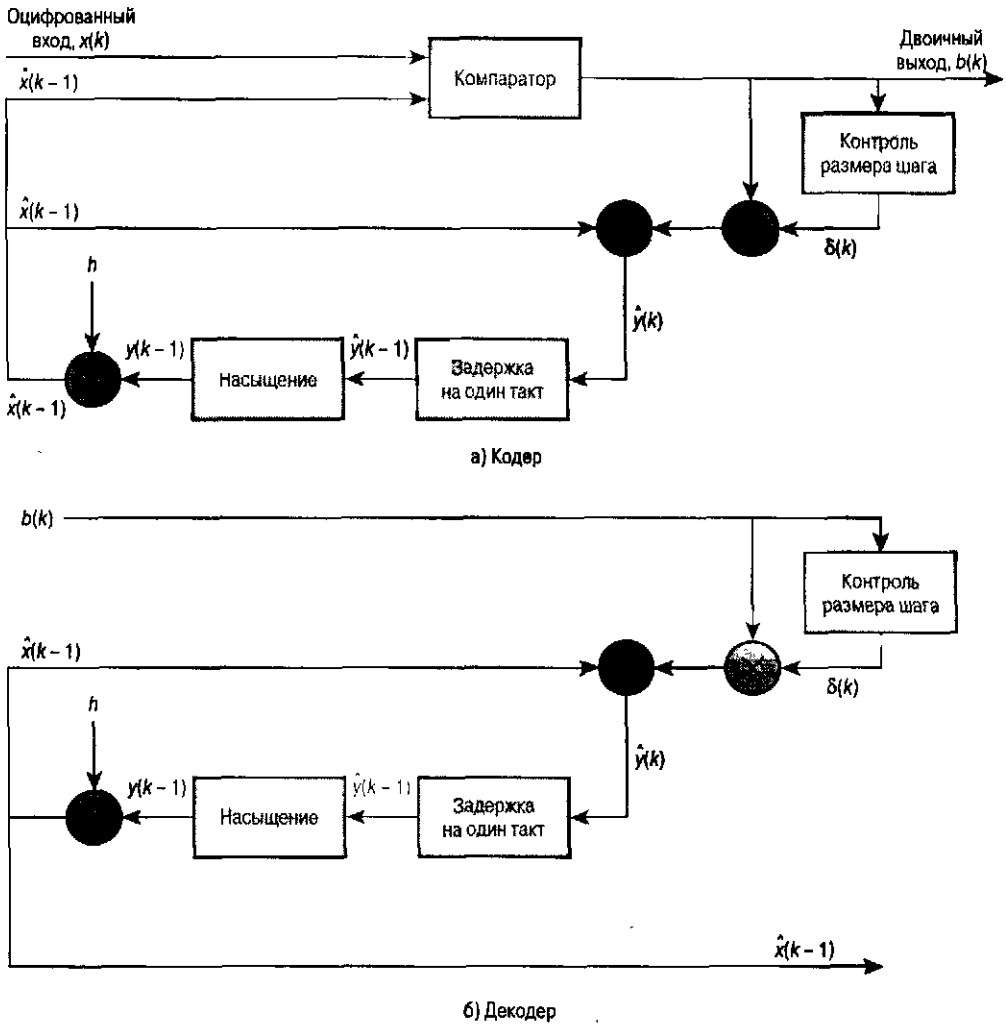


Рис. 15.14. Дельта-модуляция с непрерывной переменной огибающей

В табл. 15.6 показаны значения параметров по умолчанию. Из данного выше определения следует, что если сигнал меняется быстро (по крайней мере  $J$  из последних  $K$  шагов выполнены в одном направлении), то величина изменения шага ( $\delta(k)$ ) линейно увеличивается на постоянную величину  $\delta_{\min}$ , вплоть до некоторой максимальной величины  $\delta_{\max}$ . С другой стороны, если сигнал изменяется медленно, то величина изменения шага постепенно (с коэффициентом затухания  $\beta$ ) уменьшается вплоть до минимального значения  $\delta_{\min}$ . Знак изменения шага определяется значением выходных данных  $b(t)$ .

**Таблица 15.6. Значения параметров CVSD**

Параметр	Значение
$H$	$1 - \frac{1}{32} = 0,96875$
$\beta$	$1 - \frac{1}{1024} \approx 0,999$
$J$	4
$K$	4
$\delta_{\min}$	10
$\delta_{\max}$	1280
$Y_{\min}$	$-2^{15}$ или $-2^{15} + 1$
$Y_{\max}$	$2^{15} - 1$

После выполнения указанных вычислений изменение шага добавляется к последнему значению ступенчатой функции; результат обозначается  $\hat{y}(k)$ . Затем это значение задерживается на один интервал дискретизации, что записывается как  $\hat{y}(k-1)$ . После этого применяется *функция насыщения*, которая определяется следующим образом:

$$y(k) = \begin{cases} \min[\hat{y}(k), y_{\max}] & \hat{y}(k) \geq 0 \\ \max[\hat{y}(k), y_{\min}] & \hat{y}(k) < 0 \end{cases}$$

где  $y_{\min}$  и  $y_{\max}$  — значения отрицательного и положительного насыщения кодера, ограничивающие диапазон ступенчатой функции.

В завершение  $y(k-1)$  умножается на коэффициент затухания  $h$ , в результате чего получается оценка сигнала  $\hat{x}(k-1)$ .

## Безопасность

В узкополосной спецификации Bluetooth определены следующие средства организации безопасного канала между любыми двумя устройствами Bluetooth:

- аутентификация;
- шифрование (конфиденциальность);
- управление использованием ключей.

В алгоритмах безопасности фигурируют четыре параметра.

- **Адрес модуля.** Общеизвестный 48-битовый адрес устройства.
- **Секретный ключ аутентификации.** Секретный 128-битовый ключ.
- **Секретный частный ключ.** Секретный ключ длиной от 4 до 128 бит.
- **Случайное число.** 128-битовое случайное число, сгенерированное на устройстве Bluetooth.

Два названных секретных ключа создаются и используются только устройством Bluetooth; разглашаться они не должны.

Цель аутентификации — проверить, действительно ли устройство Bluetooth, с которым планируется обмен данными, является тем, за кого себя выдает. На рис. 15.15 показана процедура аутентификации модулем А модуля В. Аутентификация заключается в проверке того, что оба устройства используют один predetermined общий ключ аутентификации. Вначале А генерирует случайное число  $AU\_RAND_A$  и передает это значение В. Обе стороны используют алгоритм аутентификации  $E_1$  для генерации 32-битового подписанного ответа SRES. На вход  $E_1$  подается значение  $AU\_RAND_A$ , 48-битовый адрес устройства В и общий секретный ключ, а на выходе получается 128-битовое значение; 32 бит выхода формируют SRES. Алгоритм  $E_1$  (основанный на алгоритме шифрования SAFER) генерирует код аутентификации сообщения (message authentication code — MAC), представляющий собой хэш-код входных данных, основанный на секретном ключе. После того как В сгенерирует значение SRES, это значение возвращается А. А сравнивает поступившее значение SRES с локально сгенерированным. Если они совпадают — модуль В прошел аутентификацию.

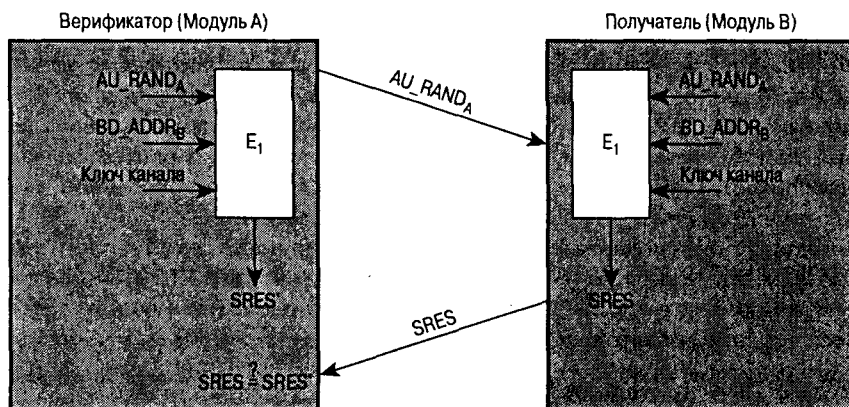


Рис. 15.15. Аутентификация с запросом и ответом

На рис. 15.15 показана аутентификация только в одном направлении. Для взаимной аутентификации описанную выше процедуру должен инициировать модуль В.

Пользовательская информация может защищаться путем шифрования полезной нагрузки пакета; код доступа и заголовок пакета никогда не шифруются. Шифрование производится по алгоритму, известному как  $E_0$ . Соответствующий процесс показан на рис. 15.16. Если шифрование данных возможно, ведущее устройство посылает подчиненному случайное число RAND. Для каждой передачи пакета генерируется новый ключ шифрования. Вначале генерируется разовый ключ полезной нагрузки: сложным способом комбинируются число RAND, адрес ведущего устройства, текущее время и общий секретный ключ. Этот ключ полезной нагрузки используется как вход алгоритма  $E_0$ , в результате получается непрерывный поток битов (на рис. 15.16 обозначен символом z). Алгоритм  $E_0$  реализуется на линейном регистре сдвига с обратной связью (linear feedback shift register — LFSR), для инициализации которого используется ключ полезной нагрузки. При шифровании (и дешифровании) к полезной нагрузке и выходу  $E_0$

применяется операция побитового исключающего ИЛИ. Поскольку для каждого сеанса шифрования текущее время будет разным, безопасность увеличивается еще и за счет постоянного изменения ключей шифрования.

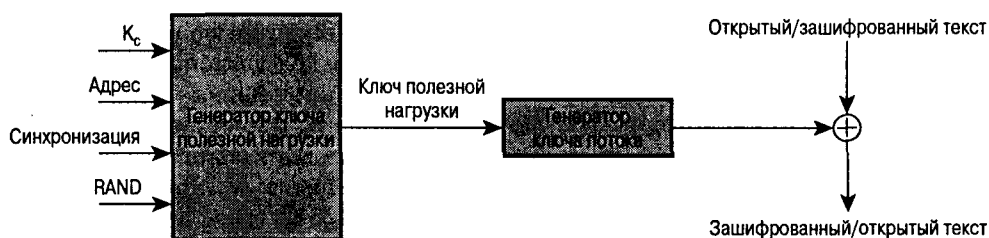


Рис. 15.16. Поточное шифрование Bluetooth

## 15.4. СПЕЦИФИКАЦИЯ АДМИНИСТРАТОРА СВЯЗЕЙ

Протокол LMP (link manager protocol — протокол администратора связей) отвечает за различные аспекты радиоканала между ведущим и подчиненным устройствами. Протокол включает обмен сообщениями в форме PDU (protocol data unit — модуль данных протокола) LMP между объектами LMP ведущего и подчиненного устройств. Сообщения всегда посылаются как однословные пакеты с 1-байтовым заголовком полезной нагрузки, идентифицирующим тип сообщения и тело полезной нагрузки, в котором содержится дополнительная информация сообщения.

Процедуры, определенные для LMP, группируются в 24 функциональных области и каждая включает обмен одним или несколькими сообщениями. Эти области перечислены в табл. 15.7, указаны также соответствующие PDU<sup>8</sup>. Ниже приводится краткий обзор каждой области.

Таблица 15.7. PDU протокола LMP

Функция	PDU
Общий ответ	accepted, not_accepted
<b>Службы безопасности</b>	
Аутентификация	au_rand, sres
Образование пар	in_rand, au_rand, sres, comb_key, unit_key
Изменение ключа канала	comb_key
Изменение текущего ключа канала	temp_rand, temp_key, use_semi_permanent_key
Шифрование	encryption_mode_req, encryption_key_size_req, start_encryption_req, stop_encryption_req
<b>Время/синхронизация</b>	
Запрос расхождения во времени	clkoffset_req, clkoffset_res

<sup>8</sup>В спецификации LMP название каждого PDU начинается с LMP\_. Например, первый PDU в табл. 15.7 называется LMP\_accepted. Для краткости префикс LMP\_ опущен.

Функция	PDU
Информация по смещению слотов	Slot_offset
Запрос информации о точности отсчета времени	timing_accuracy_req, timing_accuracy_res
<b>Возможности станции</b>	
Версия LMP	version_req, version_res
Поддерживаемые возможности	features_req, features_res
<b>Управление режимами</b>	
Переключение роли “ведущий/подчиненный”	Swich_req
Запрос имени	name_req, name_res
Отсоединение	detach
Режим удержания	hold, hold_req
Режим подслушивания	sniff, sniff_req, unsniff_req
Режим парковки	park_req, park, set_broadcast_window, modify_beacon, unpark_PM_ADDR_req, unpark_BD_ADDR_req
Контроль мощности	incr_power_req, decr_power_req, max_power, min_power
Переход между пакетами DM и DH на основе качества канала	auto_rate, preferred_rate
Качество обслуживания	quality_of_service, quality_of_service_req
Каналы SCO	SCO_link_req, remove_SCO_link_req
Управление многослотовыми пакетами	max_slot, max_slot_req
Схема избирательного запроса	page_mode_req, page_scan_mode_req
Надзор за каналом	supervision_timeout

Для ответа на PDU во многих процедурах используются два сообщения **общего ответа**. Модуль `accepted` содержит идентификатор принятого сообщения. Модуль `not_accepted` содержит идентификатор непринятого сообщения и причину отказа.

LMP поддерживает различные службы безопасности с механизмами управления аутентификацией, шифрованием и распределением ключей. В число данных служб входят следующие.

- **Аутентификация.** Аутентификация определена в узкополосной спецификации, но включает обмен двумя PDU LMP: содержащим случайное число и содержащим подписанный ответ (рис. 15.15).
- **Образование пар.** Данная служба позволяет взаимно аутентифицированным пользователям автоматически устанавливать ключ шифрования канала. На первом этапе ключ инициализации генерируется обеими сторонами и используется в процедуре аутентификации для подтверждения того, что обе стороны имеют одинаковый ключ. Ключ инициализации генерируется из общего личного идентификационного номера (`personal identification number` — PIN), введен-

ного в оба устройства. Затем обе стороны обмениваются сообщениями и определяют, каким будет ключ канала для дальнейшего шифрования: уже подтвержденным секретным ключом или комбинационным ключом, который вычисляется на основе ключа канала ведущего устройства.

- **Изменение ключа канала.** Если два устройства соединены в пару и используют комбинационный ключ, то этот ключ можно менять. Одна сторона генерирует новый ключ и посылает его другой стороне, применив операцию исключающего ИЛИ к новому и старому ключам канала. Другая сторона ключ принимает или отклоняет.
- **Изменение текущего ключа канала.** Текущий ключ канала можно изменить временно. При создании временного (на один сеанс) ключа фигурируют случайные числа и операции исключающего ИЛИ.
- **Шифрование.** LMP не задействован непосредственно в шифровании канала, но предлагает службы для управления процессом шифрования. В процессе могут оговариваться несколько параметров, в том числе рабочий режим шифрования (шифрование отсутствует, только двухточечное шифрование, двухточечное и широкополосное шифрование), размер ключа и случайный начальный ключ для начала нового сеанса шифрования. Протокол LMP также фигурирует, когда начинается и прекращается использование шифрования.

LMP предлагает механизмы синхронизации часов различных участников пикосети.

- **Запрос расхождения во времени.** Когда подчиненное устройство получает пакет FHS, оно определяет разницу между его собственным временем и временем ведущего устройства, которое указано в полезной нагрузке пакета FHS. Величина расхождения во времени постоянно обновляется с каждым пакетом, полученным от ведущего устройства. Ведущее устройство может запросить значение расхождения во времени в любой момент соединения. Запомнив эту величину, ведущее устройство может предсказать, в каком радиоканале данное устройство отзовется на сообщение поиска запроса, после того как покинет пикосеть. Это может использоваться для уменьшения времени запроса данного устройства, когда оно вновь войдет в данную пикосеть.
- **Информация по смещению слотов.** Иницилирующее устройство может передать сообщение, которое характеризует отличия в отсчете времени (промежутки между границами слотов) в двух соседних пикосетях.
- **Запрос информации о точности отсчета времени.** Используется устройством, чтобы определить, насколько точно работает подсистема синхронизации другого устройства. В число параметров, представляющих интерес, входят накапливающееся расхождение во времени двух устройств и случайное смещение счета времени.

LMP включает два PDU, которые используются для обмена информацией о сообщающихся устройствах.

- **Версия LMP.** Позволяет каждому объекту LMP определить версию LMP, реализованную на других объектах. На настоящий момент существует единственная версия.

- **Поддерживаемые возможности.** Радио и контроллер канала Bluetooth могут поддерживать только некоторые типы пакетов, описанные в узкополосной спецификации и радиоспецификации. Для обмена этой информацией используются PDU LMP\_features\_req и LMP\_features\_res. LMP поддерживает следующее:
  - 3-слововые пакеты;
  - 5-слововые пакеты;
  - шифрование;
  - смещение слотов;
  - точность синхронизации;
  - переключение;
  - режим удержания;
  - режим подслушивания;
  - режим парковки;
  - RSSI;
  - изменение скорости передачи данных на основе качества канала;
  - канал SCO;
  - пакеты HV2;
  - пакеты HV3;
  - логарифмическую  $\mu$ -характеристику;
  - логарифмическую A-характеристику;
  - CVSD;
  - схему избирательного запроса;
  - контроль мощности.

Устройство Bluetooth имеет несколько состояний и режимов, в которых оно может находиться. Для управления режимами протокол LMP предлагает следующие PDU.

- **Переключение роли “ведущий/подчиненный”.** Позволяет подчиненному устройству становиться ведущим устройством пикосети. Это требуется, например, когда устройством, производящим запрос, должно быть ведущее устройство. Если новому устройству требуется произвести запрос, оно может использовать указанную службу.
- **Запрос имени.** Позволяет устройству запросить текстовое имя другого устройства.
- **Отсоединение.** Позволяет устройству разорвать соединение. Данный PDU может передаваться ведущим или подчиненным устройством.
- **Режим удержания.** Переводит канал между ведущим и подчиненным устройствами в режиме удержания на заданный период времени.

- **Режим подслушивания.** Чтобы перейти в режим подслушивания, ведущее и подчиненное устройства согласовывают интервал подслушивания  $T$  и смещение подслушивания  $D$ , которое определяет время подслушивания. Смещение определяет время первого слота подслушивания; после этого слоты будут прослушиваться с периодом  $T$ .
- **Режим парковки.** Переводит подчиненное устройство в режим парковки.
- **Контроль мощности.** Используется устройством для того, чтобы указать другому устройству увеличить или уменьшить мощность передачи.
- **Переход между пакетами DM и DH на основе качества канала.** Устройство можно настроить так, чтобы оно постоянно использовало пакеты DM, постоянно использовало пакеты DH или автоматически выбирало тип пакетов, соответствующий качеству канала. Эта служба позволяет явное переключение между тремя указанными вариантами. Различие между пакетами DM и DH состоит в том, что первые защищаются кодом FEC 2/3, а вторые не защищены кодом FEC.
- **Качество обслуживания (QoS).** Качество обслуживания Bluetooth определяется двумя параметрами. *Интервал опроса*, максимальное время между передачами от ведущего устройства данному подчиненному, используется для распределения пропускной способности и управления ожиданием. Второй параметр — это *число повторений для широковещательных пакетов* (number of repetitions for broadcast packets — NBC). Широковещательные пакеты не подтверждаются, таким образом, чем большее число раз они будут транслироваться повторно, тем выше вероятность их приема.
- **Каналы SCO.** Используются для создания канала SCO.
- **Управление многослововыми пакетами.** Оговаривает максимальное число слотов в пакете. Значение по умолчанию — 1; возможны также значения 3 и 5.
- **Схема избирательного запроса.** Управляет типом схемы избирательного запроса устройств в пикосети. Существует схема по умолчанию, реализация которой является обязательной. Могут определяться дополнительные схемы.
- **Надзор за каналом.** Определяет максимальное время, по прошествии которого устройство может объявлять о сбое в работе канала.

## 15.5. ПРОТОКОЛ УПРАВЛЕНИЯ ЛОГИЧЕСКИМ КАНАЛОМ И АДАПТАЦИИ (L2CAP)

Подобно протоколу управления логическим каналом (logical link control — LLC) в спецификации IEEE 802 L2CAP предоставляет объектам среды совместного использования протокол канального уровня. Как и LLC, L2CAP предлагает некоторые услуги и полагается на нижестоящий уровень (в данном случае — узкополосный) в вопросах управления потоком и защиты от ошибок.

L2CAP использует каналы ACL; каналы SCO он не поддерживает. С помощью каналов ACL протокол L2CAP предоставляет протоколам высших уровней две альтернативные услуги.



- **Служба без установления соединения.** Одна из услуг, основанных на надежных дейтаграммах.
- **Служба режима соединения.** Эта услуга подобна услуге предлагаемой протоколом HDLC. Между двумя пользователями, обменивающимися данными, устанавливается логическое соединение и обеспечивается управление потоком и защита от ошибок.

## Каналы L2CAP

L2CAP предлагает три типа логических каналов.

- **Без установления соединения.** Поддерживает услуги без установления соединения. Каждый канал является однонаправленным и обычно используется для широковещания от ведущего устройства.
- **На основе соединения.** Поддерживает услуги на основе соединений. Каждый канал является двунаправленным (полнодуплексным), причем в обоих направлениях задается определенное качество обслуживания (quality of service — QoS).
- **Канал передачи сигналов.** Предусмотрен для обмена сигнальными сообщениями между объектами L2CAP.

Пример использования логических каналов L2CAP приведен на рис. 15.17. С каждым логическим каналом соотнесен идентификатор канала (channel identifier — CID). Для каналов на основе соединения уникальное значение CID присваивается обоим концам канала, таким образом каждое соединение идентифицируется и сопоставляется с пользователем L2CAP на соответствующей стороне канала. Для идентификации каналов без установления соединения используется значение CID = 2, а каналов передачи сигналов — CID = 1. Таким образом, между ведущим и любым подчиненным устройством существует только один канал без установления соединения и один канал передачи сигналов, а каналов на основе соединения может быть несколько.

## Пакеты L2CAP

На рис. 15.18 показан формат пакетов L2CAP. Для услуг без установления соединения пакет содержит следующие поля.

- **Длина.** Размер (в байтах) полезной информационной нагрузки плюс поле PSN.
- **Идентификатор канала.** 2 — канал без установления соединения.
- **Мультиплексор протоколов/услуг (protocol/service multiplexer — PSM).** Определяет пользователя на высшем уровне, которому предназначена полезная нагрузка данного пакета.
- **Полезная информационная нагрузка.** Данные пользователя более высокого уровня. Размер данного поля может доходить до 65533 ( $2^{16} - 3$ ) байт.

Пакеты на основе соединения имеют такой же формат, но без поля PSM. Это поле не требуется, поскольку получатель данных на вышестоящем уровне определяется из значения CID. Поле полезной нагрузки информации может иметь размер до 65535 ( $2^{16} - 1$ ) байт.



Пакеты сигнальных команд имеют такой же формат заголовка, что и пакеты на основе соединения. В данном случае значение CID = 1 означает, что по этому каналу передается сигнальная информация. Полезная нагрузка сигнального пакета — это одна или большее число команд L2CAP, каждая из которых состоит из четырех полей.

- **Код.** Идентифицирует тип команды.
- **Идентификатор.** Используется для сопоставления запроса с ответом. При ответе реагирующее устройство использует то же значение этого поля, что было установлено запрашивающим устройством. Для каждой новой исходящей команды должен использоваться другой идентификатор.
- **Длина.** Размер (в байтах) поля данных этой команды.
- **Данные.** Если необходимо, команду могут сопровождать дополнительные данные.

## Сигнальные команды

Существует одиннадцать сигнальных команд, разбитых на пять категорий (табл. 15.9). Для отмены любой команды может передаваться команда отклонения команды. Причиной отмены может быть, например, неприемлемый идентификатор CID или превышение длины.

Таблица 15.8. Коды сигнальных команд L2CAP

Код	Описание	Параметры
0x01	Отмена команды	Причина
0x02	Запрос соединения	PSM, CID источника
0x03	Ответ на запрос соединения	CID адресата, CID источника, результат, статус
0x04	Запрос конфигурации	CID адресата, флаги, опции
0x05	Ответ на запрос конфигурации	CID источника, флаги, результат, опции
0x06	Запрос разрыва соединения	CID адресата, CID источника
0x07	Ответ на запрос разрыва соединения	CID адресата, CID источника
0x08	Эхо-запрос	Данные (не обязательно)
0x09	Эхо-ответ	Данные (не обязательно)
0x0A	Информационный запрос	Тип информации
0x0B	Ответ на информационный запрос	Тип информации, результат, данные (не обязательно)

Команды соединения используются для установления нового логического канала. Команда запросов содержит значение PSM, определяющее пользователя L2CAP для данного соединения. На настоящий момент разрешены три значения: для протокола обнаружения услуги, RFCOMM и протокола управления телефонией. Другие значения PSM присваиваются динамически и зависят от реализации. Команда запроса также содержит значение CID, которое присвоится источником этому соединению. Команда ответа включает CID источника и адресата, последний присваивается кана-

ду отвечающей стороной. Параметр результата определяет результат запроса (успешный, в стадии рассмотрения, отклоненный), и если запрос находится в стадии рассмотрения (например, незаконченная аутентификация, незаконченная авторизация), в поле статуса указывается текущее состояние запроса.

Для установления исходного логического канала между двумя объектами L2CAP и пересмотра параметров этого канала могут передаваться команды конфигурации. Каждый параметр в команде-запросе связан с входящим или с исходящим трафиком. Команда запроса содержит поле флагов; в настоящее время определен только один флаг, указывающий, будут ли переданы дополнительные команды конфигурации. Поле опций содержит список параметров и их значений, подлежащих обсуждению. Каждый параметр определяется тремя полями.

- **Тип (1 байт).** Семь младших битов данного байта определяют опцию. Если старший бит равен 0, опция является обязательной, и если распознать ее не удалось, адресат должен отклонить конфигурационный запрос. Если старший бит равен 1, опция необязательна и может игнорироваться адресатом.
- **Длина (1 байт).** Размер полезной нагрузки, в качестве которой выступают опции. Длина 0 указывает на отсутствие полезной нагрузки.
- **Полезная нагрузка — опции.** Дальнейшая информация об опции.

Согласовываться могут следующие параметры.

- **Максимальный модуль передачи (maximum transmission unit — MTU).** Наибольшая полезная нагрузка пакета L2CAP (в байтах), которую автор запроса может принять в данном канале. Вследствие асимметричности MTU отправитель запроса должен указать, какой MTU он готов принять по обратному каналу (если это значение отличается от значения, принятого по умолчанию). Реализации L2CAP должны поддерживать MTU минимального размера — 48 байт. Значение по умолчанию — 672 байт. Оно не подлежит согласованию, это просто информация о том, MTU какого размера может принимать отправитель запроса.
- **Время ожидания отказа.** Напомним, что в схеме ARQ определен отказ от полезной нагрузки, если некоторое число попыток ее передачи завершилось сбоем. Время ожидания отказа — это параметр, определяющий промежуток времени, в течение которого источник будет предпринимать попытки передачи, прежде чем отказаться от пакета L2CAP.
- **Качество обслуживания (quality of service — QoS).** Определяет спецификацию потока трафика в канале для трафика локального устройства (исходящего трафика). Этот параметр описывается ниже.

В двух последних случаях возможно согласование параметров: сторона-ответчик может принять параметры QoS и время ожидания отказа или предложить согласовать их. Сторона-источник может в ответ принять или отклонить измененные параметры.

Команда ответа на запрос конфигурации также включает поле флагов, имеющее такое же значение, как в команде-запросе. Поле результата в команде ответа указывает, был переданный запрос принят или отклонен. Поле опций содержит тот же список параметров, что имеется в соответствующей команде запроса. Если запрос был успешен, эти параметры включают возвратные значения всех безразличных параметров (см. последующее обсуждение QoS). Если запрос

был неудачным, в ответ должны включаться отклоненные параметры, значения которых меняются на приемлемые для станции-ответчика.

**Команды разрыва соединения** используются для разрыва логического канала.

**Эхо-команды** — это требование запроса от удаленного объекта L2CAP. Такие команды обычно используются для тестирования канала или передачи информации о поставщике с использованием дополнительного поля данных.

**Информационные команды** — это запросы информации о реализации от удаленного объекта L2CAP.

## Качество обслуживания

Параметр QoS в L2CAP определяет на основе RFC 1363<sup>9</sup> спецификацию потока трафика. По сути, спецификация потока — это набор параметров, отражающих уровень производительности, к которому будет стремиться передатчик.

Если данная опция включена в запрос конфигурации, она описывает поток трафика, исходящий от устройства, которое послало запрос, по направлению к запрашиваемому устройству. Если данная опция включена в положительный запрос конфигурации, она описывает соглашения по потоку входящего трафика, как их представляет отвечающее устройство. Если данная опция включена в отрицательный запрос конфигурации, она описывает предпочтительный поток входящего трафика, как его представляет отвечающее устройство.

Спецификация потока состоит из следующих параметров:

- тип услуги;
- скорость маркера (байты в секунду);
- размер сегмента в маркерах (байты);
- максимальная ширина полосы (байты в секунду);
- ожидание (микросекунды);
- разброс задержек (микросекунды).

Параметр тип услуги показывает уровень услуги для данного потока. При значении 0 по данному каналу трафик распространяться не будет; 1 — услуга “без обязательств”, устройство будет передавать данные с максимально возможной скоростью, но без гарантий производительности; 2 — гарантированная услуга, отправитель будет передавать данные, которые соответствуют остальным параметрам QoS.

Параметры скорости маркера и размера сегмента в маркерах определяют схему маркерного сегмента, которая часто используется в спецификациях QoS. Преимуществом этой схемы является емкое описание максимального и среднего трафика, который может ожидать получатель, и удобная реализация стратегии управления трафиком.

Спецификация трафика по схеме маркерного сегмента включает два параметра: скорость пополнения маркера ( $R$ ) и размер сегмента ( $B$ ). Маркерная скорость  $R$  задает среднюю непрерывную скорость передачи данных; т.е. за относительно долгий период времени среднее значение скорости передачи данных этого потока равно  $R$ . Размер сегмента  $B$  определяет величину разрешенного кратко-

---

<sup>9</sup>A Proposed Flow Specification, RFC 1363, September 1992.

временного превышения скорости  $R$ . Точное условие звучит так: в течение любого периода времени  $T$  объем посланных данных не может превышать  $RT + B$ .

Данная схема показана на рис. 15.19, там же приведены иллюстрации используемых терминов. *Сегмент* представляет счетчик, который в любое время отображает допустимое количество байтов данных, которые могут послаться. Сегмент заполняется маркерами байтов со скоростью  $R$  (т.е. значение счетчика увеличивается  $R$  раз в секунду) вплоть до его полного заполнения (максимальное значение счетчика). Данные поступают от пользователя L2CAP в форме пакетов, которые для передачи выстраиваются в очередь. Пакет может передаваться, только если есть достаточное количество маркеров байтов для маркировки всего пакета. Если маркеров достаточно, пакет передается и из сегмента уходит соответствующее число маркеров. Если маркеров недостаточно, то спецификация потока запрещает передачу такого пакета. В стандартах не указывается, что происходит с такими пакетами дальше; обычно пакет просто находится в очереди на передачу до тех пор, пока не будет накоплено нужное число маркеров.

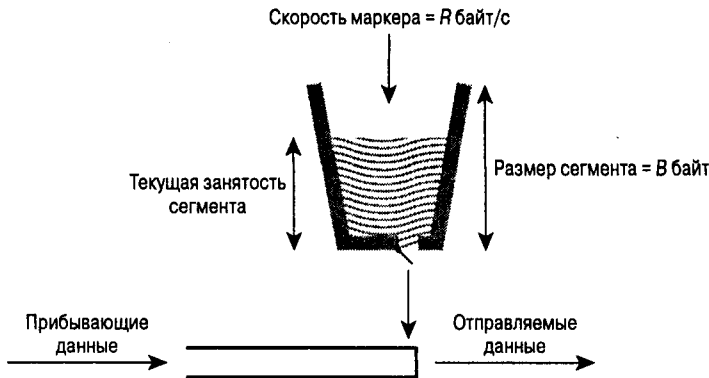


Рис. 15.19. Схема маркерного сегмента

В течение достаточно большого промежутка времени скорость передачи данных, которую разрешает маркерный сегмент, равна  $R$ . Однако в период отсутствия данных или при относительно медленной скорости их поступления емкость сегмента увеличивается, чтобы в дальнейшем принять на  $B$  байтов больше, чем определяется средней скоростью. Таким образом,  $B$  обозначает допустимую степень неравномерности потока данных.

Если оба параметра протокола L2CAP имеют значение 0, то для данного приложения маркерная схема не требуется и использоваться не будет. Если оба параметра равны 1, это значение называется безразличным. Для услуг “без обязательств” безразличное значение указывает, что запрашивающей стороне требуется настолько большой маркер или сегмент, насколько может гарантировать отвечающая сторона, а для гарантированной услуги оно означает, что на момент запроса скорость передачи данных будет максимальной или максимальным будет размер сегмента.

**Максимальная ширина полосы** (измеряется в байтах в секунду) определяет, насколько быстро может происходить пересылка пакетов между приложениями. Некоторые промежуточные системы используют эту информацию для более эффективного распределения ресурсов. Предположим, что маркерный сегмент полон, а в поток поступает серия пакетов, размер которых равен размеру сегмента. Если не ввести огра-

ничивающий механизм, такое событие может надолго “засорить” сегмент. Максимальная скорость передачи — это и есть такой ограничитель, который определяет, с какой частотой последовательные пакеты могут вводиться в сеть.

**Ожидание** — это максимально допустимая задержка (выражается в микросекундах) между отправкой бита и началом его передачи по радиоканалу.

**Разброс задержек** — это разность (в микросекундах) возможных максимальной и минимальной задержек пакета. Указанное значение используется приложениями для вычисления объема буфера, необходимого принимающей стороне для восстановления схемы передачи данных. Если принимающему приложению требуется, чтобы данные поступали так же, как передавались, ему, возможно, придется временно занести поступающие данные в буфер, чтобы приемник смог подстроиться под схему передачи. В качестве примера можно привести ситуацию, когда приложение передает данные, такие, как выборки речи, которые генерируются и проигрываются через регулярные промежутки времени. В зависимости от размера буфера, который может предоставить принимающий хост, будет меняться и разброс задержек, допустимый при передаче отдельных пакетов потока.

## 15.6. РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА И WEB-САЙТЫ

Обзоры Bluetooth содержатся в [HAAR00a] и [HAAR00b]. Определенный интерес представляют также два других обзора (состоят из нескольких частей), приведенных в [WILS00] и [RODB00]. В работах [BRAY01] и [MILL01] изложены технические детали, причем в первой — более подробно.

**BRAY01** Bray J., Sturman C. *Bluetooth: Connect Without Cables*. — Upper Saddle River, NJ: Prentice Hall, 1992.

**HAAR00a** Haartsen J. The Bluetooth Radio System. — *IEEE Personal Communications*, February 2000.

**HAAR00b** Haartsen J., Mattisson S. Bluetooth — A New Low-Power Radio Interface Providing Short-Range Connectivity. — In: *Proceedings of the IEEE*, October 2000.

**MILL01** Miller B., Bisdikian C. *Bluetooth Revealed*. — Upper Saddle River, NJ: Prentice Hall, 2001.

**RODB00** Rodbell M. Bluetooth: Wireless Local Access, Baseband and RF Interfaces, and Link Management. — *Communications System Design*, March, April, May 2000 ([www.csdmag.com](http://www.csdmag.com)).

**WILS00** Wilson J., Kronz J. Inside Bluetooth: Part I and Part II. — *Dr. Dobb's Journal*, March, April 2000.



### Рекомендуемые Web-сайты

- **Bluetooth SIG** ([www.bluetooth.com/](http://www.bluetooth.com/)). Содержит все стандарты, множество других документов и новостей, а также информацию о продуктах и производителях Bluetooth.
- **Infotooth** ([www.palowireless.com/infotooth/](http://www.palowireless.com/infotooth/)). Отличный дополнительный источник информации о Bluetooth.

## 15.7. ТЕРМИНЫ И ВОПРОСЫ

### Основные термины

Bluetooth	пикосеть	протокол управления
RFCOMM	протокол замены кабеля	логическим каналом
администратор канала (LM)	протокол обнаружения службы	и адаптации (L2CAP)
базовая спецификация		рассеянная сеть
модель использования		спецификация потока
		спецификация профиля

### Вопросы

1. Какие прикладные области поддерживаются технологией Bluetooth?
2. Какая разница между базовой спецификацией и спецификацией профиля?
3. Что такое “модель использования”?
4. Как связаны ведущее и подчиненное устройства пикосети?
5. Каким образом можно объединить перестройку частоты и дуплекс с временным разделением?
6. Чем отличаются схемы FH-CDMA и DS-CDMA?
7. Перечислите и кратко определите типы каналов, которые могут устанавливаться между ведущим и подчиненным устройствами.
8. Какие схемы исправления ошибок используются в узкополосной спецификации Bluetooth?
9. Перечислите и кратко определите узкополосные логические каналы Bluetooth.
10. Какие службы безопасности предлагает Bluetooth?
11. Перечислите и кратко опишите логические каналы L2CAP.
12. Что такое “спецификация потока”?



**Приложение А. Стандарты и организации  
стандартизации**

**Приложение Б. Анализ трафика**

**Приложение В. Анализ Фурье**

**Приложение Г. Протоколы управления каналом  
передачи данных**

ПРИЛОЖЕНИЕ **A**

**СТАНДАРТЫ  
И ОРГАНИЗАЦИИ  
СТАНДАРТИЗАЦИИ**

**В** данной книге часто упоминается важная концепция стандарта. В этом приложении приводятся некоторые базовые сведения относительно природы и значимости стандартов, а также рассматриваются основные организации, занимающиеся разработкой стандартов в области связи и организации сетей.

## **A.1. ВАЖНОСТЬ СТАНДАРТОВ**

В телекоммуникационной индустрии для регулирования физических, электрических и процедурных характеристик средств связи существуют определенные стандарты. В компьютерной индустрии этого правила ранее не было. Если производители средств связи давно поняли, что их оборудование должно взаимодействовать с оборудованием других поставщиков, то производители компьютеров традиционно пытались добиться монопольного положения среди своих пользователей. Быстрое распространение компьютеров и распределенной обработки данных показало несостоятельность этой позиции. Компьютеры различных производителей должны взаимодействовать друг с другом и, при том, что стандарты протоколов постоянно развиваются, пользователи уже не согласны приобретать специализированные программы для преобразования протоколов. В результате стандартизация в настоящее время распространилась на все технологические области, рассматриваемые в книге.

Процесс стандартизации обладает как преимуществами, так и недостатками. Перечислим наиболее яркие из них. Принципиальными преимуществами стандартизации являются следующие.

- Стандарт обеспечивает существование большого рынка для конкретного оборудования или программного обеспечения. Это способствует организации серийного производства и, в некоторых случаях, использованию технологий интеграции высокого (large-scale-integration — LSI) или очень высокого (very-large-scale-integration — VLSI) уровня, что приводит к уменьшению стоимости производства.
- Стандарт позволяет взаимодействовать изделиям различных производителей, что дает покупателю большую гибкость в выборе и использовании оборудования.

Принципиальные недостатки стандартизации.

- Стандарт имеет склонность к “замораживанию” технологии. Пока стандарт пройдет стадии разработки, проверки, переделки и опубликования, возможно появление более эффективных технологий.
- Один и тот же предмет описывается несколькими стандартами. Само по себе это является недостатком не стандартов, а существующего положения дел в отрасли. К счастью, за последние годы различные организации, занимающиеся проблемами стандартизации, стали более тесно сотрудничать. Несмотря на это, еще существуют области, в которых имеется несколько противоречивых стандартов.

## **А.2. СТАНДАРТЫ И РЕГУЛИРОВАНИЕ**

Читателю полезно будет различать три концепции.

- Рекомендательный стандарт.
- Регулятивный стандарт.
- Регулятивное использование рекомендательного стандарта.

Рекомендательные стандарты разрабатываются организацией по стандартизации, подобной описываемым ниже. Эти стандарты являются необязательными в том смысле, что существующего стандарта не заставляют придерживаться. Т.е. производители добровольно реализуют продукт, который соответствует стандарту, если они считают, что это выгодно для них; нет никаких официальных норм, которым нужно следовать. Стандарты являются добровольными и в том смысле, что они разрабатываются добровольцами, которым организация, управляющая процессом разработки стандарта, ничего не платит. Эти добровольцы — сотрудники заинтересованных сторон, таких, как производители и правительственные органы.

Рекомендательные стандарты соблюдаются потому, что обычно они разрабатываются на основе широкого консенсуса, и потому, что существует пользовательский спрос на продукты, на основе которого производители и инициировали разработку стандарта.

Регулятивные стандарты, наоборот, разрабатываются государственным регулятивным органом для удовлетворения некоторых общественных целей — экономических, охраны здоровья и государственной безопасности. Данные стандарты — это норма, которой должно следовать так, как она изложена в официальном документе. В качестве знакомых примеров регулятивных стандартов можно назвать правила противопожарной безопасности и санитарные нормы. В то же время регулятивные стандарты могут относиться к разнообразным продуктам, в том числе связанным с компьютерами и связью. Например, Федеральная комиссия по средствам связи США определяет разрешенное электромагнитное излучение.

Относительно новое, или по крайней мере недавно получившее широкое распространение, явление — регулятивное использование рекомендательного стандарта. Типичный пример: регулирование, при котором государство приобретает продукт, соответствующий некоторому набору рекомендательных стандартов. Данный подход имеет следующие преимущества.

- Уменьшает работу по нормотворчеству в государственных органах.
- Способствует сотрудничеству государственных органов и организаций по стандартизации с целью выработки стандартов широкого применения.
- Снижает разнообразие стандартов, которым должны следовать поставщики.

## **А.3. МЕЖДУНАРОДНЫЙ ТЕЛЕКОММУНИКАЦИОННЫЙ СОЮЗ**

Международный телекоммуникационный союз (International Telecommunications Union — ITU) является специализированным учреждением ООН. Следовательно, членами ITU являются правительства стран. Представительство США располо-

жено в Государственном департаменте США. В уставе союза ITU указано, что он “отвечает за изучение технических, эксплуатационных и тарифных вопросов и выпуск Рекомендаций (Recommendations) по ним с целью стандартизации телекоммуникаций в мировом масштабе”. Основной целью ITU является стандартизация до требуемой степени методов и операций в сфере телекоммуникаций для достижения сквозной совместимости международной телекоммуникационной связи, вне зависимости от страны источника и адресата.

## **Сектор радиосвязи**

Сектор радиосвязи союза ITU (ITU Radiocommunication Sector — ITU-R) был создан 1 марта 1993 года на основе существовавших ранее организаций CCIR и IFRB (созданных в 1927 и 1947 годах, соответственно). Сектор ITU-R отвечает за стандартизацию в области радиосвязи. Перечислим основные виды деятельности сектора ITU-R.

- Разработка проектов Рекомендаций ITU-R для технических характеристик и операционных процедур радиокommunikационных служб и систем.
- Создание Руководств по управлению спектром и совместной работе радиокommunikационных служб и систем.

Сектор делится на следующие группы:

- SG 1 — управление спектром;
- SG 3 — распространение радиосигналов;
- SG 4 — услуги стационарных спутников;
- SG 6 — широкоэвещательные услуги (наземные и спутниковые);
- SG 7 — научные цели;
- SG 8 — спутниковые услуги для мобильной, любительской и т.п. радиосвязи;
- SG 9 — стационарные услуги;
- SC — специальный комитет по регуляторным/процедурным вопросам;
- CPM — комитет подготовки конференций.

## **Сектор стандартизации телекоммуникаций**

Сектор ITU-T был создан 1 марта 1993 года в результате преобразования союза ITU. Он заменил собой Международный консультативный комитет по телеграфии и телефонии, МККТТ (International Telegraph and Telephone Consultative Committee — CCITT), имевший практически те же устав и цели, что и новый сектор ITU-T. Сектор ITU-T занимается телекоммуникационными стандартами, изучая технические, операционные и тарифные вопросы и выпуская по ним Рекомендации с целью стандартизации телекоммуникаций в мировом масштабе.

Структура сектора ITU-T включает 14 комиссий, подготавливающих Рекомендации по следующим вопросам.

2. Сетевые и сервисные операции.
3. Принципы тарифов и учета.

4. Сетевое управление телекоммуникациями и эксплуатация сети.
5. Защита окружающей среды от электромагнитного воздействия.
6. Внешние предприятия.
7. Сети передачи данных и связь открытых систем.
8. Характеристики телематических систем.
9. Теле- и радиопередача.
10. Языки и общие подходы к программному обеспечению для телекоммуникационных систем.
11. Требования к сигналам и сигнальные протоколы.
12. Производительность сквозной передачи сетей и терминалов.
13. Общие вопросы сетей.
15. Транспортные сети, системы и оборудование.
16. Мультимедийные службы и системы.

## Схема работы

Работа в ITU-T проводится в форме четырехлетних циклов. Каждые четыре года организуется Всемирная конференция по стандартизации телекоммуникаций (World Telecommunications Standardization Conference), на которой утверждается рабочая программа на следующие четыре года, имеющая форму вопросов, формулируемых на основе запросов, сделанных членами различных комиссий, и предоставляемых на рассмотрение этим комиссиям. На конференции определяются эти вопросы, пересматриваются сферы деятельности комиссий, создаются новые или упраздняются прежние комиссии, а также распределяются вопросы между существующими комиссиями.

На основе сформулированных вопросов каждая комиссия готовит проект Рекомендаций, который может быть представлен на утверждение следующей (через четыре года) конференции. Правда, все чаще Рекомендации утверждаются сразу после их разработки, не дожидаясь конца четырехлетнего периода изучения. Эта ускоренная процедура была принята в 1988 году. Таким образом, в 1988 году в последний раз большая серия документов была опубликована одновременно в виде перечня Рекомендаций.

## A.4. СТАНДАТЫ INTERNET И ОБЩЕСТВО INTERNET

Многие из протоколов, входящих в набор протоколов TCP/IP, были стандартизованы или еще находятся в процессе стандартизации. По общему согласию за разработку и публикацию этих стандартов отвечает организация, известная как Общество Internet (Internet Society). Эта организация представляет собой сообщество профессионалов, наблюдающих за многочисленными специальными комиссиями и советами, которые участвуют в развитии и стандартизации Internet.

В данном разделе кратко описана методика разработки стандартов для набора протоколов TCP/IP.

## Организации Internet и документы RFC

Общество Internet является координационным комитетом по проектированию, конструированию и управлению Internet. В сферу ответственности этой организации входят собственно Internet, а также стандартизация протоколов, используемых конечными системами Internet с целью взаимодействия. Фактически за работы по созданию и публикации стандартов отвечают следующие три организации, входящие в Общество Internet.

- **Совет по архитектуре Internet (Internet Architecture Board — IAB).** Отвечает за определение общей архитектуры Internet, обеспечивая руководство и управление группой IETF.
- **Проблемная группа проектирования Internet (Internet Engineering Task Force — IETF).** Основная организация, занимающаяся вопросами проектирования и развития протоколов.
- **Управляющий комитет группы IETF (Internet Engineering Steering Group — IESG).** Осуществляет техническое руководство действиями группы IETF и процессами стандартизации Internet.

Фактическая разработка новых стандартов и протоколов для Internet осуществляется рабочими группами, которые организует группа IETF. Членство в рабочей группе является добровольным; принять участие в ее работе может любая заинтересованная сторона. В процессе разработки спецификации рабочая группа создает черновую версию документа, доступную в сети под названием “Internet Draft” (проект Internet) и находящуюся в архиве “Internet Drafts” группы IETF. Документ может оставаться на стадии проекта до шести месяцев, при этом заинтересованные стороны могут его пересматривать и комментировать. В течение этого времени комитет IESG может одобрить публикацию черновой версии как документа RFC (Request for Comment). Если в указанный период времени черновая версия документа не перешла в разряд документа RFC, то она удаляется из архива. Впоследствии рабочая группа может опубликовать переработанную версию того же проекта.

Именно группа IETF отвечает за публикацию документов RFC в случае их одобрения комитетом IESG. Сами документы RFC являются рабочими материалами для научно-исследовательского сообщества Internet. По существу, документ этой серии может затрагивать любую тему, связанную с системами передачи данных, и может содержать что угодно: от отчета о собрании до спецификации стандарта.

### Процесс стандартизации

Решение о том, какой из документов RFC становится стандартом Internet, принимается комитетом IESG при учете рекомендации группы IETF. Чтобы спецификация стала стандартом, она должна удовлетворять следующим требованиям:

- быть непротиворечивой и понятной;
- быть технически компетентной;
- иметь различные независимые и способные ко взаимодействию реализации со значительным опытом применения;

- обладать значительной общественной поддержкой;
- приносить очевидную пользу в некоторых или во всех областях Internet.

Основным различием между данными критериями и критериями, которые используют организации ISO и ITU-T при создании международных стандартов, является наличие у стандарта опыта применения.

В левой части рис А.1 показана последовательность этапов, называемая *путем стандарта*, которую должна пройти спецификация для превращения в стандарт; сам этот процесс описывается документом RFC 2026. Каждый следующий этап включает все более значительный объем проверки и тестирования. На каждом этапе группа IETF должна предоставлять рекомендацию для продвижения протокола, а комитет IESG должен ее одобрять. Процесс начинается, когда комитет IESG разрешает публикацию проекта Internet как документа RFC со статусом предложенного стандарта (Proposed Standard).

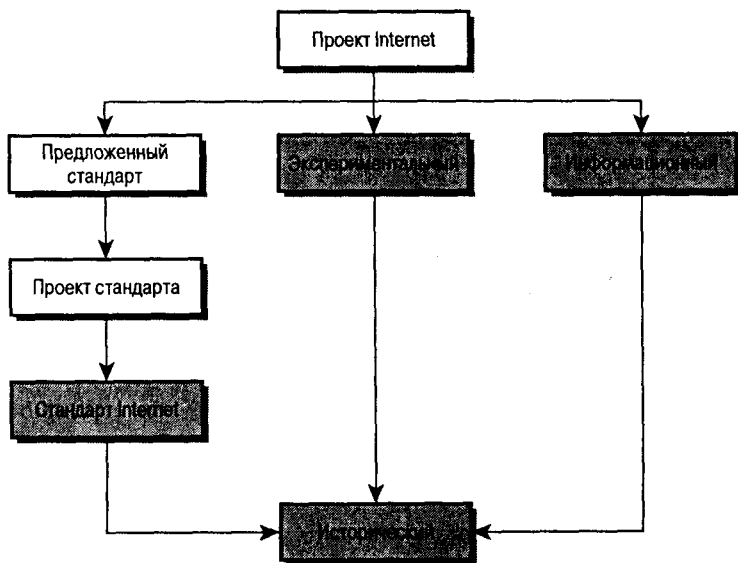


Рис. А.1. Процесс публикации документа RFC

Белые прямоугольники на рисунке представляют временные состояния, в которых документ должен находиться минимальное приемлемое время. Впрочем, документ может иметь статус предложенного стандарта в течение по меньшей мере шести месяцев, а статус проекта стандарта (Draft Standard) — по меньшей мере четырех месяцев, чтобы оставалось время для пересмотра и комментирования. Серыми прямоугольниками обозначены долговременные состояния, в которых документ может находиться годы.

Чтобы спецификация получила статус проекта стандарта, должно существовать не менее двух независимых и способных ко взаимодействию реализаций, проверенных на практике.

После того как спецификация была реализована и проработала достаточное время, ей может быть присвоен статус стандарта Internet (Internet Standard). На этом этапе спецификации присваивается номер STD и номер RFC.

И, наконец, когда стандарт устаревает, ему присваивается статус исторического.



## Документы, не проходящие путем стандарта

Протокол или другая спецификация, признанная не готовой к стандартизации, может быть опубликована как экспериментальный документ RFC (Experimental RFC). После доработки спецификация может быть повторно представлена на рассмотрение. Если в общем спецификация непротиворечива, способствует решению важных проблем, считается понятной, получила необходимое число рецензий и достаточно удовлетворяет интересам компьютерного сообщества, чтобы считаться полезной, то документ RFC может получить статус предложенного стандарта.

И последнее замечание: информация общего порядка Internet-сообщества публикуется в так называемых информационных спецификациях (Informational Specification).

## **A.5. СТАНДАРТЫ IEEE 802**

Ключом к развитию рынка локальных сетей является наличие дешевого интерфейса. Цена подключения оборудования к локальной сети должна быть во много раз меньше цены собственно оборудования. Этим требованием, а также сложностью логики локальной сети обусловлено решение, основанное на использовании чипов и VLSI (Very Large Scale Integration — сверхбольшая интегральная схема, СБИС). В то же время производители чипов не расположены к фиксации необходимых ресурсов, пока не возникнет обширный рынок. Широко принятые стандарты локальных сетей гарантируют этот объемный рынок и также позволяют общаться устройствам от различных производителей. Такова подоплека создания комитета IEEE 802.

Комитет выпустил набор принятых в 1985 году Национальным институтом стандартизации США (American National Standards Institute — ANSI) национальных стандартов. Впоследствии эти стандарты были пересмотрены, и в 1987 году Международная организация по стандартизации (International Organization for Standardization — ISO) выпустила их повторно уже как международные стандарты. Эти стандарты были обозначены как ISO 8802. С тех пор комитет IEEE 802 продолжает перерабатывать и расширять стандарты, которые в конечном счете принимает и ISO.

Комитет быстро пришел к двум выводам. Во-первых, задача организации связи в локальной сети достаточно сложна, и поэтому требуется разбить ее на более управляемые подзадачи. Это привело к созданию эталонной модели IEEE 802, представленной в главе 14 (см. рис. 14.1).

Во-вторых, ни один технический подход не удовлетворит всех требований. Второй вывод был неохотно сделан, когда стало ясно, что ни один стандарт не удовлетворит всех участников комитета. Необходима была поддержка различных топологий, методов доступа и сред передачи. Реакция комитета была следующей: решение стандартизировать все серьезные предложения и отказ от попыток свести их воедино.

## Модель реализации локальных сетей

Помимо разработки эталонной модели протоколов для структурирования работы комитета также вводится модель реализации, которая должна направлять разработку некоторых стандартов. Эти две модели сравниваются на рис. А.2. Отличие

заклучено в том, что модель реализации подчеркивает сложность стандартов физического уровня. Данному уровню управления доступом к среде (MAC) соответствуют несколько альтернативных физических сред. В большинстве случаев имеется независимая от среды спецификация, в которой рассмотрены вопросы кодирования сигнала, синхронизации и другие темы, общие для всех сред. Кроме того, существует зависимая от среды спецификация, в которой описана электромеханическая спецификация данной среды. Эти две части могут реализовываться независимо или с промежуточным интерфейсом.

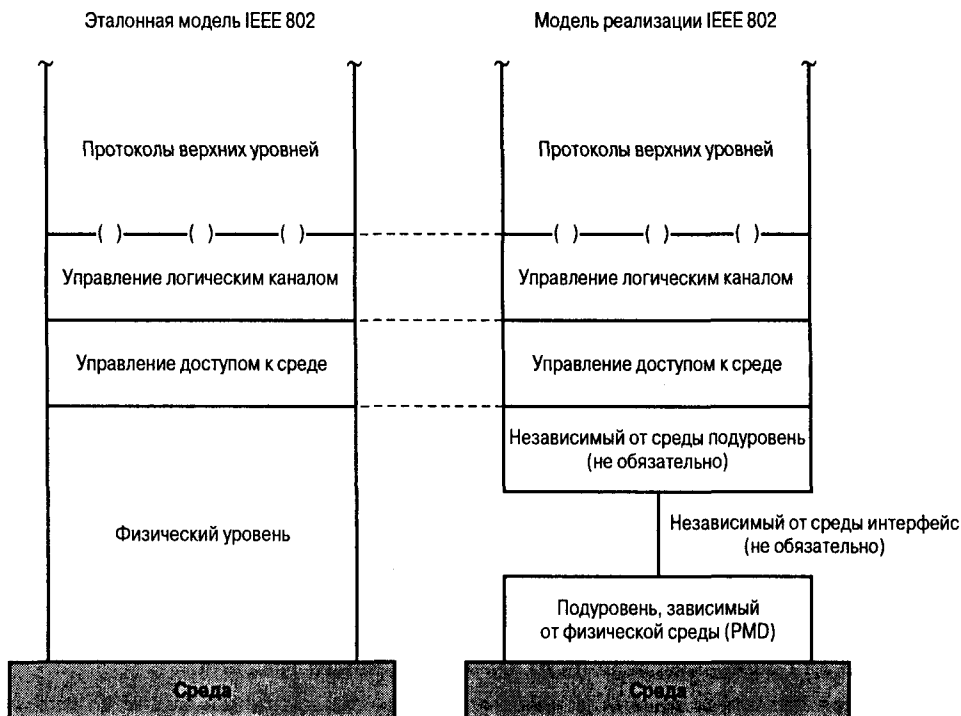


Рис. А.2. Эталонная модель и модель реализации IEEE 802

## Структура IEEE 802

Текущее состояние стандартизации характеризуется работой, которую выполняют различные подкомитеты IEEE 802. Перечислим эти группы.

- **802.1 Higher Layer LAN Protocols Working Group.** Разработка стандартов и рекомендаций по следующим темам: архитектура локальных и городских сетей IEEE 802, соединение локальных, городских и глобальных сетей IEEE 802, общее управление сетями, уровни протоколов, находящиеся над MAC и LLC.
- **802.2 Logical Link Control Working Group.** В настоящее время неактивна.
- **802.3 Ethernet Working Group.** Разработка стандартов для сетей на основе технологии CSMA/CD (Ethernet).

- **802.4 Token Bus Working Group.** Разработка стандартов для шинных топологий с передачей маркера. В настоящее время неактивна.
- **802.5 Token Ring Working Group.** Разработка стандартов для кольцевых топологий с передачей маркера.
- **802.6 Metropolitan Area Network Working Group.** Разработка стандартов для городских сетей. В настоящее время неактивна.
- **802.7 Broadband TAG.** Консультативная группа по разработке рекомендательных стандартов IEEE Std 802.7–1989, IEEE Recommended Practice for Broadband Local Area Networks. Группа неактивна, работа по всем текущим проектам завершена. Преемник — группа 802.14.
- **802.8 Fiber Optic TAG.** Разработка стандартов по оптоволоконным технологиям.
- **802.9 Isochronous LAN Working Group.** Разработка стандартов для изохронных ЛВС.
- **802.10 Security Working Group.** Разработка совместимых стандартов безопасности для всех сетей IEEE 802.
- **802.11 Wireless LAN Working Group.** Разработка стандартов для беспроводных локальных сетей.
- **802.12 Demand Priority Working Group.** Разработка стандартов для локальных сетей с запросом приоритета (100VG-AnyLAN).
- **802.13.** Не используется.
- **802.14 Cable Modem Working Group.** Стандарты для передачи данных по традиционным кабельным телевизионным сетям. Эталонная архитектура задает собой смешанную оптоволоконную/кабельную структуру, охватывающую площадь до 80 км от головного узла. Основной сетевой протокол разрабатывается в расчете на трафик LLC IEEE 802.2 (представитель — Ethernet). В то же время высказывается предположение, что сеть должна поддерживать технологию ATM для передачи различного мультимедийного трафика.
- **802.15 Wireless Personal Area Networks Working Group.** Разработка стандартов для малоразмерных беспроводных сетей.
- **802.16 Broadband Wireless Access Study Group.** Разработка стандартов для широкополосного беспроводного доступа.
- **802.17 Resilient Packet Ring Working Group.** Стандарт определит протокол сети с очень высокой скоростью передачи, оптимизированной для пакетной передачи в оптоволоконных сетях с кольцевой топологией.

ПРИЛОЖЕНИЕ **Б**

**АНАЛИЗ ТРАФИКА**

**В** разделе 10.1 было введено понятие анализа трафика. В данном приложении приводится более подробная информация об этом предмете<sup>1</sup>.

## **Б.1. ОСНОВНЫЕ ПОНЯТИЯ**

Понятия регулирования трафика были разработаны при проектировании телефонных коммутаторов и телефонных сетей с коммутацией каналов, но эту концепцию можно с таким же успехом применить к сотовым сетям. Рассмотрим ячейку, которая может одновременно обслуживать  $N$  пользователей (пропускная способность —  $N$  каналов) и имеет  $L$  потенциальных абонентов ( $L$  мобильных устройств). Если  $L < N$ , система называется *неблокируемой*; в любой момент времени могут обрабатываться все звонки. Если  $L > N$ , система называется *блокируемой*; попытавшись позвонить, абонент может обнаружить, что вся имеющаяся пропускная способность уже используется, и данный абонент блокируется. В блокируемых системах нас интересуют следующие основные вопросы.

1. Какова степень блокирования; т.е. чему равна вероятность того, что запрос соединения будет блокирован? Можно рассмотреть альтернативный вопрос: какая пропускная способность ( $N$ ) требуется для получения определенного верхнего предела вероятности блокирования?
2. Если заблокированный вызов поставлен в очередь ожидания услуги, чему равна средняя задержка ее предоставления? Альтернативный вопрос: какая пропускная способность требуется для получения определенной средней задержки?

Для определения загрузки системы введем два параметра:

- $\lambda$  — средняя частота звонков (запросов соединения) в единицу времени;
- $h$  — среднее время разговора при удачном звонке.

Основной мерой трафика является его **интенсивность**, которая выражается в безразмерных единицах — эрлангах:

$$A = \lambda h.$$

При моделировании трафика параметр  $A$  — мера информационного обмена в часы наибольшей нагрузки (см. раздел 10.1) — считается заданным. Сама модель должна дать ответы на вопросы, подобные тем, что приведены в начале раздела. Природу этой модели определяют два ключевых фактора:

- способ обработки заблокированных звонков;
- число источников трафика.

Возможны два способа обработки заблокированных вызовов. Во-первых, вызов может помещаться в очередь вызовов, ожидающих свободного канала; этот подход называется **задержкой неудачных вызовов** (lost call delayed —

---

<sup>1</sup> В данном приложении использованы понятия из области анализа систем массового обслуживания. Дополнительный материал по этому вопросу можно найти на [WilliamStallings.com/StudentSupport.html](http://WilliamStallings.com/StudentSupport.html).

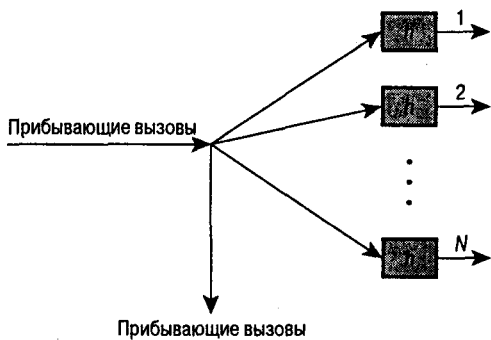
LCD), хотя фактически вызов не является неудачным, он просто отложен. Во-вторых, заблокированный вызов может отклоняться, и в этом случае снова имеем две возможных линии поведения пользователя. Если пользователь перезванивает повторно через некоторый случайный промежуток времени, это называется **очисткой неудачного вызова** (lost call cleared — LCC). Если пользователь постоянно предпринимает попытки дозвониться, это называется **удержанием неудачного вызова** (lost call held — LCH). Для каждого варианта выведены формулы, описывающие производительность системы. Для сотовых систем обычно используется модель LCC, которая, как правило, более точно описывает систему.

Вторым ключевым элементом модели трафика является предположение о том, является число пользователей конечным или бесконечным. Отличия показаны на рис. Б.1. В модели с **бесконечным числом источников** предполагается фиксированное среднее время поступления вызовов. При **конечном числе источников** частота поступления вызовов будет зависеть от числа источников, уже задействованных в разговорах. В частности, если общее число потенциальных пользователей равно  $L$  и каждый из них совершает звонки со средней частотой  $\lambda/L$ , то при абсолютно свободной ячейке частота поступления вызовов равна  $\lambda$ . Если же в момент времени  $t$  обслуживается  $K$  пользователей, мгновенная частота поступления вызовов в этот момент равна  $\lambda(L - K)/L$ . Отметим, что модели с бесконечным числом источников аналитически проще, а применять их стоит, если число источников по крайней мере в 5–10 раз превышает пропускную способность системы.

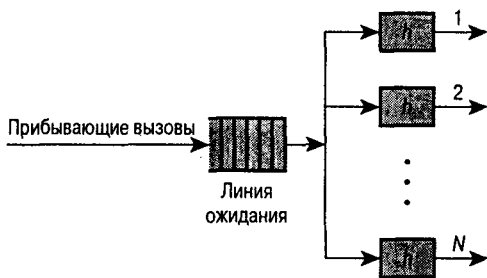
## **Б.2. МНОГОСЕРВЕРНЫЕ МОДЕЛИ**

Теперь мы готовы рассмотреть использование различных моделей трафика для расчета систем. Для каждой из четырех моделей на рис. Б.1 выведены выражения для величин, представляющих практический интерес. Важнейшие соотношения приведены в табл. Б.1. При их выводе были сделаны следующие предположения:

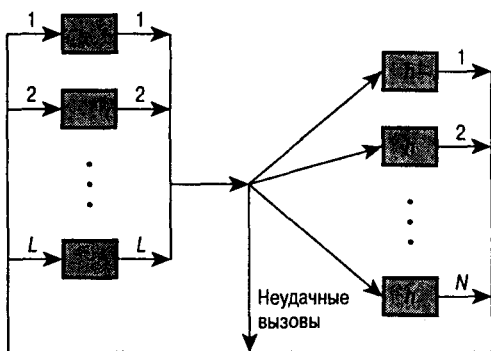
- пуассоновское распределение моментов поступления звонков;
- экспоненциальное распределение времени разговора (не требуется для модели с бесконечным числом источников и очисткой неудачных вызовов);
- равная интенсивность трафика для всех источников;
- звонки обслуживаются в порядке поступления (важно при вычислении задержки).



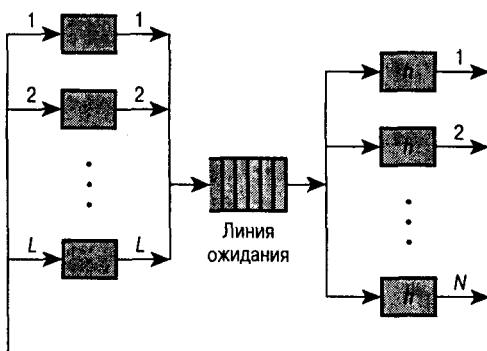
а) Бесконечное число источников, очистка неудачных вызовов



б) Бесконечное число источников, задержка неудачных вызовов



в) Конечное число источников, очистка неудачных вызовов



г) Конечное число источников, задержка неудачных вызовов

Рис. Б.1. Многосерверные модели систем обслуживания

Видно, что даже при сделанных предположениях формулы содержат большие суммы. На раннем этапе развития данной теории основная работа аналитика заключалась в упрощении предположений до такой степени, чтобы уравнения имели красивое аналитическое решение. Были получены определенные результаты, которые в табличной форме приводятся в разных работах (например, [MART94]). К сожалению, имеется тенденция к использованию этих таблиц (а соответственно, и предположений) там, где данные предположения являются заведомо неверными. Однако при современном уровне развития компьютерных технологий данная проблема уже не является существенной. В любом случае табличные результаты полезны для быстрой, грубой оценки.

Некоторые параметры в табл. Б.1 стоит прокомментировать. Для систем с очисткой неудачных звонков  $P$  — вероятность очистки, или потери, запроса вызова. Данный параметр — это отношение числа звонков, которые не могут быть обслужены, к общему числу звонков; в телефонии данную величину также называют **уровнем обслуживания** (grade of service — GOS). Для систем с задержкой неудачных вызовов поступающий вызов задерживается, а не отклоняется. В этом случае  $P(> 0)$  — вероятность того, что при поступлении вызова система используется полностью, а следовательно, вызов задерживается.  $P(> 1)$  — вероят-

ность того, что любой запрос вызова задерживается на время, превышающее  $t$ , а  $P_2(> t)$  — вероятность того, что задержанный вызов задерживается на время, превышающее  $t$ .

**Таблица Б.1. Формулы трафика**

а) Бесконечное число источников, очистка неудачных вызовов

$$P = \frac{\frac{A^N}{N!}}{\sum_{x=0}^N \frac{A^x}{x!}}$$

б) Бесконечное число источников, задержка неудачных вызовов

$$P(> 0) = \frac{\frac{A^N}{N!} \frac{N}{N-A}}{\sum_{x=0}^{N-1} \frac{A^x}{x!} + \frac{A^N}{N!} \frac{N}{N-A}}$$

$$P(> t) = P(> 0)e^{-(N-A)t/h}$$

$$D_1 = P(> 0) \frac{h}{N-A}$$

$$D_2 = \frac{h}{N-A}$$

$$P_2(> t) = e^{-(N-A)t/h}$$

в) Конечное число источников, очистка неудачных вызовов

$$P = \frac{\binom{L-1}{N} M^N}{\sum_{x=0}^N \binom{L-1}{x} M^x},$$

где

$$\binom{L-1}{x} = \frac{(L-1)!}{x!(L-1-x)!}$$

$$M = \frac{A}{L-A(1-P)}$$

г) Бесконечное число источников, задержка неудачных вызовов

$$P(> 0) = \frac{\sum_{x=N}^L \frac{L!}{N!} \frac{M^x}{(L-x)! N^{x-N}}}{\sum_{x=0}^{N-1} \binom{L}{x} M^x + \sum_{x=N}^L \frac{L!}{N!} \frac{M^x}{(L-x)! N^{x-N}}},$$

где

$$M = \frac{A}{L+1-A(1-P)} \approx \frac{A}{L+1-A}$$

$A$  — поступающий трафик, эрланги  
 $N$  — число серверов  
 $L$  — число источников  
 $h$  — среднее время разговора  
 $P$  — вероятность неудачного вызова (блокирования, задержки)

$P(> 0)$  — вероятность того, что задержка  $> 0$   
 $P(> t)$  — вероятность того, что задержка  $> t$   
 $P_2(> t)$  — вероятность того, что задержка  $> t$  для задержанных звонков  
 $D_1$  — средняя задержка для всех звонков  
 $D_2$  — средняя задержка для задержанных звонков

## Бесконечное число источников, очистка неудачных вызовов

Данный случай рассмотрен в разделе 10.1. В табл. 10.3 указана пропускная способность таких систем в зависимости от числа обслуживающих серверов и вероятности блокирования вызова, а на рис. Б.2 для различного числа серверов изображена вероятность неудачного вызова как функции поступающей нагрузки.



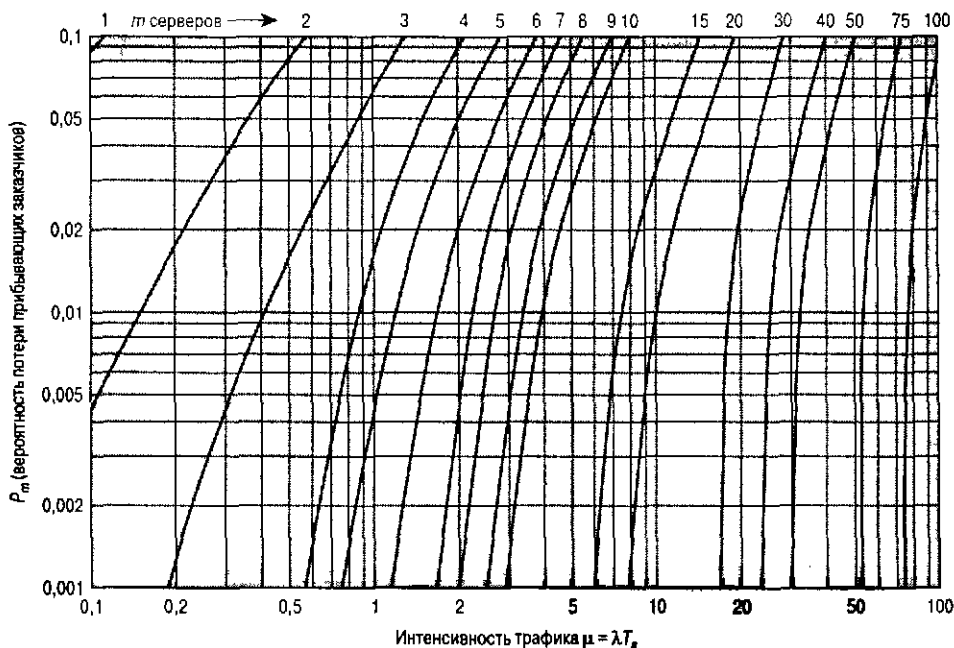


Рис. Б.2. Вероятность неудачного вызова для системы с бесконечным числом источников и очисткой неудачных вызовов

## Конечное число источников, очистка неудачных вызовов

В реальной системе число источников конечно, так что следует применять табл. Б.1, в. Здесь формулы сложнее и, из-за присутствия дополнительного параметра  $L$ , таблицы более громоздки, чем в предыдущем случае. Следует отметить, правда, что во многих случаях достаточно предположения о бесконечном числе источников.

Чтобы оценить порядок отличия систем с конечным и бесконечным числом источников, сравним две системы. Одна из них — система с бесконечным числом источников и частотой звонков  $\lambda$ ; вторая — система с конечным числом источников, частота звонков на источник —  $\lambda/L$ . Имеем

$$\begin{aligned}
 A_{\infty} &= \lambda h, \\
 A_L &= \frac{\lambda}{L}(L - K)h = \\
 &= \lambda h \left(1 - \frac{K}{L}\right) \geq \lambda h \left(1 - \frac{N}{L}\right).
 \end{aligned}$$

Здесь

$A_{\infty}$  — поступающий трафик, бесконечное число источников;

$A_L$  — поступающий трафик,  $L$  источников;

$K$  — число источников, в данный момент занятых в разговорах.

Неравенство, приведенное выше, отражает тот факт, что число источников, занятых в разговорах, не может превышать общую пропускную способность системы,  $N$ .

Имеем

$$A_{\infty} > A_L > A_{\infty} \left(1 - \frac{N}{L}\right).$$

Таким образом, для  $L \gg N$   $A_{\infty}$  является хорошей аппроксимацией  $A_L$ .

Этот вывод согласуется с данными табл. Б.2, в которой показано, что когда число источников становится больше пропускной способности системы, эффективная пропускная способность системы приближается к величине, рассчитанной для бесконечного числа источников. Отметим, что при снижении вероятности неудачного звонка аппроксимация становится менее точной. Изучая табл. Б.2, можно также прийти к заключению, что вероятность неудачного звонка в системе с конечным числом источников всегда меньше, чем в подобной системе с бесконечным числом источников. Таким образом, модель с бесконечным числом источников — это оценка с запасом.

**Таблица Б.2. Очистка неудачных вызовов, конечное и бесконечное число источников (число серверов — 10)**

Число источников	Поступающая нагрузка		
	$p = 0,005$	$p = 0,01$	$p = 0,05$
$\infty$	2,8	4,46	6,22
50	4,16	4,64	6,4
25	4,6	5,1	6,81
20	4,81	5,32	7,06
15	5,27	5,78	7,41
12	5,97	6,47	8,02
10	10	10	10

## Задержка неудачных вызовов

Если неудачные запросы можно ставить в очередь ожидания услуги, то при проектировании системы важнее могут быть характеристики задержки, а не блокирования. Некоторые формулы для варианта задержки неудачных вызовов в системе с бесконечным числом источников показаны в табл. Б.1, б. Проектирование системы может производиться с учетом вероятности любой задержки, вероятности задержки, превышающей заданную величину, или средней вероятности задержки.

Для поступающих звонков все три названных выше параметра выражаются как сложные функции  $A$  и  $N$  и вычисляются только с помощью компьютера или аппроксимирующих таблиц. К счастью, для задержанных звонков средняя задержка и задержка, превышающая заданную величину, вычисляются достаточно просто и представляют некоторый интерес.

**Пример.** Чему равна нагрузка, которую может обработать коммутатор с пропускной способностью 1500, если среднее время разговора равно 1000 с, а доля задержанных звонков, ожидающих более 1 мин, не превышает 10%? Итак, при  $N = 1500$ ,  $h = 1000$ ,  $t = 60$ ,  $P_2(> 60) = 0,1$  имеем

$$\ln(0,1) = -(1500 - A) \times (0,06).$$

Отсюда  $A = 1462$ . Таким образом, при заданных требованиях возможная нагрузка системы довольно близка к общей пропускной способности.

**Пример.** В данном примере использован рис. Б.3, на котором представлена вероятность задержки для бесконечного числа источников и задержки неудачных вызовов. Рассмотрим коммутатор с пропускной способностью 20 и систему, в которой среднее время разговора равно 5 мин. Если нагрузка системы равна 16 эрлангов, чему равна доля всех звонков которые будут задержаны более чем на 2,5 мин? Прежде всего нужно вычислить долю звонков, которые вообще будут задержаны. Из рис. Б.3 для параметров  $A = 16$  и  $N = 20$  имеем  $P(> 0) \approx 0,25$ . Далее,

$$\begin{aligned} P(> 150) &= P(> 0) \times e^{-(N-A)/h} = \\ &= 0,25 \times e^{-(20-16)(150/300)} = 0,03. \end{aligned}$$

Средняя задержка задержанных вызовов равна  $h/(N - A) = 75$  с. Средняя задержка для всех вызовов равна  $P(> 0) \times 75 = 18,75$  с.

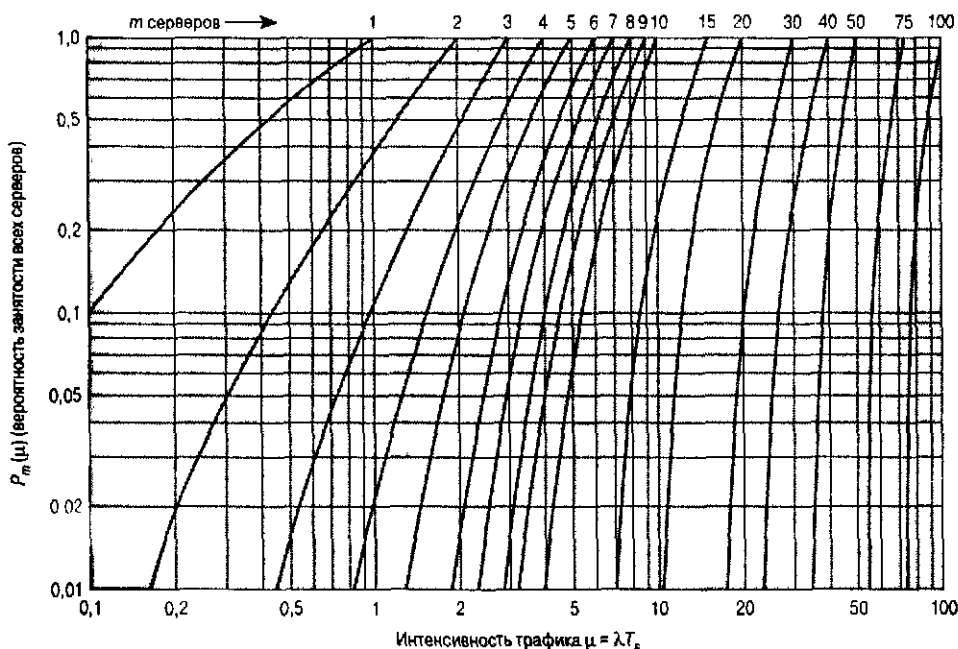


Рис. Б.3. Вероятность задержки для систем с бесконечным числом источников и задержкой неудачных вызовов

Для систем с конечным числом источников и задержкой неудачных вызовов формулы приведены в табл. Б.1, г. Данный случай является аналитически наиболее сложным, для первой аппроксимации стоит использовать модель с бесконечным числом источников.

## Резюме

Уравнения из табл. Б.1 могут использоваться по-разному: для вычислений с использованием таблиц или для компьютерных расчетов. Для системы с бесконечным числом источников и очисткой неудачных вызовов имеется три переменных:  $P$ ,  $N$  и  $A$  (которую можно определить из  $\lambda$  и  $h$ ). Зная любые две переменные, можно вычислить третью. Приведем примеры соответствующих задач.

- Имея значение  $P$ , измеренное для определенного  $N$ , нужно вычислить, насколько потребуется увеличить  $N$  для снижения  $P$  до заданного уровня.
- Зная  $\lambda$ ,  $h$  и желаемое значение  $P$ , нужно вычислить требуемую пропускную способность ( $N$ ).

В системах с бесконечным числом источников и задержкой неудачных вызовов появляется дополнительный параметр — задержка. Как и ранее, зная две из трех переменных  $P(> 0)$ ,  $N$  и  $A$ , третью можно вычислить. Отметим, что для определения величины  $D$  необходимо знать  $h$ .

Те же рассуждения можно привести и для систем с конечным числом источников (и задержкой или очисткой неудачных вызовов). В таких системах используется фиксированное значение  $L$ , а постановка задач может быть такой же, как и ранее.

## Б.3. РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

Удачным практическим руководством по теории обслуживания, которая лежит в основе вопросов, рассмотренных в данном приложении, является [TANN95]; здесь предложены подробные рекомендации по применению анализа систем массового обслуживания и приведено множество примеров. К книге также прилагается компакт-диск с богатой библиотекой подпрограмм на языке Pascal для вычисления характеристик многих практических ситуаций. Другим заслуживающим внимания руководством по данному предмету является [GUNT00]. Удачный обзор анализа трафика предложен в [MART94]. Здесь же содержатся графики и таблицы, которые можно использовать для выполнения быстрого анализа.

GUNT00 Gunther N. *The Practical Performance Analyst*. — Lincoln, NE: Authors Choice Press, 2000.

MART94 Martine R. *Basic Traffic Analysis*. — Englewood Cliffs, NJ: Prentice-Hall, 1994.

TANN95 Tanner M. *Practical Queueing Analysis*. — New York: McGraw-Hill, 1995.

ПРИЛОЖЕНИЕ **В**

**АНАЛИЗ ФУРЬЕ**

**В** этом приложении дан обзор ключевых понятий анализа Фурье.

## В.1. ПРЕДСТАВЛЕНИЕ ПЕРИОДИЧЕСКОГО СИГНАЛА В ВИДЕ РЯДА ФУРЬЕ

Определение частотной природы многих сигналов — это поразительно простая задача... если только есть хорошая таблица интегралов. Начнем с периодических сигналов. Любой периодический сигнал может быть представлен в виде суммы синусоид, называемой рядом Фурье<sup>1</sup>:

$$x(t) = \frac{A_0}{2} + \sum_{n=1}^{\infty} [A_n \cos(2\pi n f_0 t) + B_n \sin(2\pi n f_0 t)],$$

где  $f_0$  — величина, обратная периоду сигнала ( $f_0 = 1/T$ ), называемая *собственной частотой*, или *собственной гармоникой*. Частоты, кратные  $f_0$ , называются *гармониками*. Следовательно, периодический сигнал с периодом  $T$  состоит из собственной частоты  $f_0 = 1/T$  и частот, кратных ей. Если член  $A_0 \neq 0$ , то сигнал  $x(t)$  содержит *постоянную составляющую*.

Значения коэффициентов, входящих в приведенное выше выражение, вычисляются следующим образом:

$$A_0 = \frac{2}{T} \int_0^T x(t) dt,$$

$$A_n = \frac{2}{T} \int_0^T x(t) \cos(2\pi n f_0 t) dt,$$

$$B_n = \frac{2}{T} \int_0^T x(t) \sin(2\pi n f_0 t) dt.$$

Такое представление, называемое *синусно-косинусным*, является наиболее простым для вычисления, но неудобно тем, что на каждой частоте присутствуют два компонента. Существует другое, более выразительное представление, называемое *амплитудно-фазовым*, которое имеет вид

$$x(t) = \frac{C_0}{2} + \sum_{n=1}^{\infty} C_n \cos(2\pi n f_0 t + \theta_n).$$

С предыдущим представлением оно связано следующим образом:

$$C_0 = A_0,$$

---

<sup>1</sup> При записи рядов Фурье и преобразований Фурье математики обычно используют переменную  $\omega_0$ , имеющую размерность радиан в секунду и  $\omega_0 = 2\pi f_0$ . В физике и технике удобнее использовать запись через  $f_0$ , поскольку при этом выражения становятся проще, кроме того, если частота будет в герцах, а не в радианах в секунду, выражения будут интуитивно восприниматься проще.

$$C_n = \sqrt{A_n^2 + B_n^2},$$

$$\theta_n = \text{tg}^{-1}\left(\frac{-B_n}{A_n}\right).$$

Примеры рядов Фурье для периодических сигналов приведены на рис. В.1.

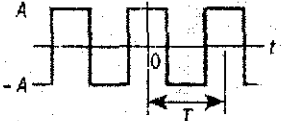
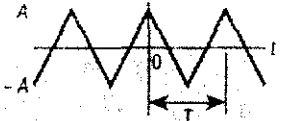
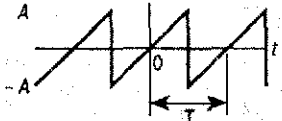
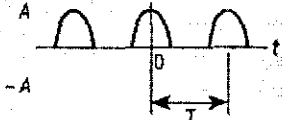
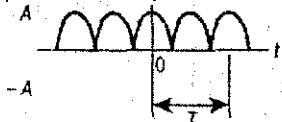
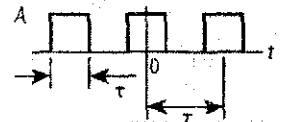
Сигнал	Ряд Фурье
<p>Меандр</p> 	$\left(\frac{4A}{\pi}\right) \times \left[\cos(2\pi f_1 t) - \frac{1}{3} \cos(2\pi(3f_1)t) + \frac{1}{5} \cos(2\pi(5f_1)t) - \frac{1}{7} \cos(2\pi(7f_1)t) + \dots\right]$
<p>Треугольный импульс</p> 	$C_0 = 0$ $C_n = 0 \quad \text{для четных } n$ $C_n = 8A/(n\pi)^2 \quad \text{для нечетных } n$
<p>Пилообразный импульс</p> 	$A_0 = 0$ $A_n = 0 \quad \text{для четных } n$ $B_n = -(-1)^n \times (2A/n\pi)$
<p>Однополупериодный выпрямленный косинус</p> 	$C_0 = 2A/\pi$ $C_n = 0 \quad \text{для нечетных } n$ $C_n = (2A/\pi) \times (-1)^{(1+n/2)} \times \{2/(n^2 - 1)\} \quad \text{для четных } n$
<p>Двухполупериодный выпрямленный косинус</p> 	$C_0 = 4A/\pi$ $C_n = (4A/\pi) \times (-1)^n \times \{1/(4n^2 - 1)\}$
<p>Серия импульсов</p> 	$C_n = (2A\tau/T) \times (\sin(n\pi\tau/T)/(n\pi\tau/T))$

Рис. В.1. Некоторые распространенные формы периодических сигналов и соответствующие им ряды Фурье

## В.2. РАЗЛОЖЕНИЕ В РЯД ФУРЬЕ АПЕРИОДИЧЕСКОГО СИГНАЛА

Мы видели, что спектр периодического сигнала состоит из дискретных частотных составляющих, находящихся на собственной частоте и ее гармониках. Этот спектр можно определить с помощью преобразования Фурье: между сигналом  $x(t)$  и его спектром  $X(f)$  существует следующая связь:

$$x(t) = \int_{-\infty}^{\infty} X(f)e^{2\pi i f t} df,$$

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-2\pi i f t} dt,$$

где  $i = \sqrt{-1}$ . Наличие в формуле мнимых чисел — вопрос удобства. Мнимые компоненты имеют физическую интерпретацию и связаны с фазой сигнала; впрочем, данная тема не входит в число вопросов, рассматриваемых в этой книге.

Некоторые примеры сигналов и их Фурье-образов приведены на рис. В.2.

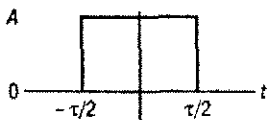
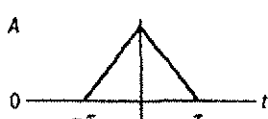
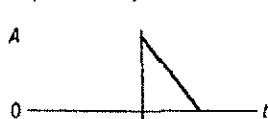

Сигнал $x(t)$	Фурье-образ $X(f)$
<p>Прямоугольный импульс</p> 	$A\tau \frac{\sin(\pi f \tau)}{\pi f \tau}$
<p>Треугольный импульс</p> 	$A\tau \left( \frac{\sin(\pi f \tau)}{\pi f \tau} \right)^2$
<p>Пилообразный импульс</p> 	$(A/2\pi f) \times \{ [\sin(\pi f \tau)/\pi f \tau] \exp(-\pi f \tau) - 1 \}$
<p>Косинусоидальный импульс</p> 	$\frac{2A\tau}{\pi} \times \frac{\cos(\pi f \tau)}{1 - (2f\tau)^2}$

Рис. В.2. Некоторые распространенные аperiodические сигналы и их Фурье-образы



## Спектральная плотность мощности и ширина полосы

Абсолютная ширина полосы любого ограниченного во времени сигнала бесконечна. На деле же основная мощность сигнала сосредоточена в некотором конечном диапазоне частот, и эффективная ширина полосы определяется именно той частью спектра, на которую приходится большая часть мощности сигнала. Чтобы строго определить эту концепцию, нам требуется ввести спектральную плотность мощности (power spectral density — PSD). Эта величина, по сути, описывает распределение мощности сигнала как функцию частоты, показывая, таким образом, какая мощность приходится на различные диапазоны частот.

Рассмотрим вначале мощность во временном интервале. Как правило, функция  $x(t)$  задает сигнал в единицах напряжения или силы тока. В любом случае мгновенная мощность сигнала пропорциональна  $|x(t)|^2$ . Среднюю мощность ограниченного во времени сигнала мы определим как

$$P = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} |x(t)|^2 dt .$$

Для периодического сигнала средняя мощность одного периода равна

$$P = \frac{1}{T} \int_0^T |x(t)|^2 dt .$$

Желательно знать, как распределение мощности зависит от частоты. Для периодических сигналов это распределение легко выражается через коэффициенты показательного ряда Фурье. Спектральная плотность мощности  $S(f)$  удовлетворяет следующему равенству:

$$S(f) = \sum_{n=-\infty}^{\infty} |C_n|^2 \delta(f - nf_0) .$$

Здесь  $f_0$  — величина, обратная периоду сигнала ( $f_0 = 1/T$ ),  $C_n$  — коэффициент в амплитудно-фазовом представлении ряда Фурье, а  $\delta(t)$  — единичный импульс, или дельта-функция, определяемая следующим образом:

$$\delta(t) = \begin{cases} 0, & \text{если } t \neq 0 \\ \infty, & \text{если } t = 0 \end{cases}$$
$$\int_{-\infty}^{\infty} \delta(t) dt = 1 .$$

Для аперидических функций спектральную плотность мощности  $S(f)$  определить труднее. На практике для ее определения делают следующее: задают “период”  $T_0$  и неограниченно увеличивают его.

Для функции  $S(f)$ , принимающей непрерывные значения, мощность, заключенная в полосе частот  $f_1 < f < f_2$ , равна

$$P = 2 \int_{f_1}^{f_2} S(f) df .$$

Для периодического сигнала мощность, приходящаяся на первые  $j$  гармоник, равна

$$P = C_0^2 + \frac{1}{2} \sum_{n=1}^j C_n^2 .$$

После того как мы определили все нужные понятия, мы можем дать определение ширины полосы половинной мощности, которая является наиболее общей характеристикой полосы. Ширина полосы половинной мощности представляет собой интервал между частотами, на которых  $S(f)$  принимает значения, равные половине максимального значения мощности, или, эквивалентное определение, на 3 дБ ниже максимального значения.

### **В.3. РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА**

Весьма доступной работой по рядам Фурье и преобразованиям Фурье является книга [JAME95]. Другая интересная работа — [HSU84]. Для изучения рядов и преобразований Фурье можно порекомендовать работу [КАММ00].

HSU Hsu H. *Applied Fourier Analysis*. — New York: Harcourt Brace, 1984.

JAME95 James J. *A Student's Guide to Fourier Transforms*. — Cambridge, England: Cambridge University Press, 1995.

КАММ00 Kammler D. *A First Course in Fourier Analysis*. — Upper Saddle River, NJ: Prentice Hall, 2000.

ПРИЛОЖЕНИЕ **Г**

**ПРОТОКОЛЫ  
УПРАВЛЕНИЯ  
КАНАЛОМ  
ПЕРЕДАЧИ ДАННЫХ**

**В** данной книге неоднократно говорилось об использовании протоколов управления каналом передачи данных как элемента архитектуры протоколов. Ключевой механизм большинства таких протоколов — защита от ошибок с применением автоматического запроса повторной передачи (ARQ), рассмотренного в разделе 8.4. В данном приложении речь идет о протоколе HDLC, являющемся представителем протоколов управления каналом передачи данных и основой, на которой развились многие такие протоколы.

## Г.1. HDLC

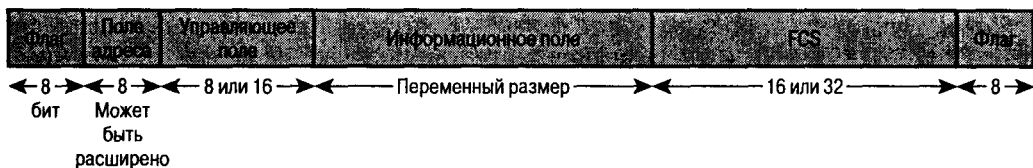
Рассмотрим важнейший протокол управления каналом связи — протокол HDLC (High-level Data Link Control — высокоуровневый протокол управления каналом) (стандарты ISO 3009 и ISO 4335). Это не единственный распространенный протокол, но он является основой многих других важных протоколов управления каналом связи, использующих те же механизмы, что применены в протоколе HDLC, или им подобные.

### Структура кадра

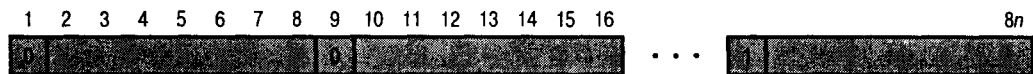
Пожалуй, лучший способ понять протокол HDLC — рассмотреть структуру его кадров. Функционирование протокола — это обмен информацией двух типов между двумя связанными станциями. Во-первых, HDLC принимает пользовательские данные от некоторого вышестоящего уровня и доставляет эти данные указанной стороне. Кроме того, HDLC принимает пользовательские данные и доставляет их вышестоящему уровню со своей стороны канала передачи данных. Во-вторых, два модуля HDLC должны обмениваться управляющей информацией с целью обеспечения управления потоком, защиты от ошибок и реализации других функций управления. Для этого информация форматируется блоками, называемыми кадрами. Кадр — это предопределенная структура, в которой каждому типу информации отведено свое место.

На рис. Г.3 показана структура кадра протокола HDLC. Сам кадр содержит следующие поля.

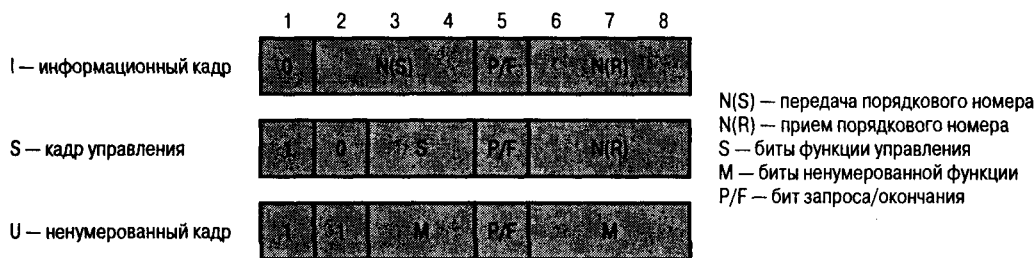
- **Флаг.** Используется для синхронизации. Находится в начале и конце кадра и всегда содержит последовательность битов 01111110.
- **Адрес.** Указывает подчиненную станцию — приемник или передатчик кадра и является необходимым в многоточечных каналах связи, где главный узел может посылать данные нескольким подчиненным и несколько подчиненных могут посылать данные главному. Размер поля адреса, согласно предварительной договоренности, обычно составляет 8 бит, может также использоваться расширенный формат (рис. Г.3, б).
- **Управляющее поле.** Указывает цель и функции кадра. Описывается ниже в этом разделе.
- **Информационное поле.** Содержит передаваемые пользовательские данные.
- **Контрольная последовательность кадра (FCS).** Содержит 16- или 32-битовую проверку четности с избыточностью, используется для обнаружения ошибок.



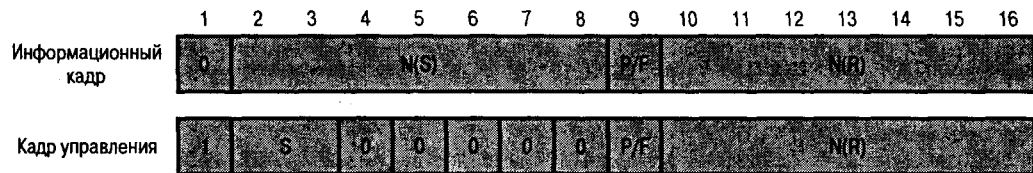
а) Формат кадра



б) Расширенное поле адреса



в) Формат управляющего поля из 8 бит



г) Формат управляющего поля из 16 бит

Рис. Г.3. Примеры операций протокола HDLC

В протоколе HDLC определены три типа кадров, каждый из которых имеет собственный формат управляющего поля. *Информационные кадры* (information frames, I-кадры) переносят данные, которые необходимо передать пользователю. Кроме того, в информационном кадре передается информация управления потоком и защиты от ошибок. *Кадры управления* (supervisory frames, S-кадры) обеспечивают реализацию других механизмов управления потоком данных и защиты от ошибок. *Ненумерованные кадры* (unnumbered frames, U-кадры) обеспечивают дополнительные функции управления каналом связи.

Первые один-два бита управляющего поля служат для определения типа кадра. Остальные разряды организованы в подполя, показанные на рис. Г.3, в и г. Отметим, что в основном управляющем поле S- и I-кадров используются трехбитовые порядковые номера. При соответствующей команде выбора режима в S- и I-кадрах может использоваться расширенное управляющее поле, в котором применяются семибитовые порядковые номера.

Все форматы управляющих полей содержат бит запроса/окончания (poll/final bit, P/F bit). Использование этого бита зависит от контекста. В управляющих кадрах, как правило, он называется Р-битом, а его значение устанавливается равным 1 для запроса передачи ответного кадра от однорангового объекта протокола HDLC. В ответных кадрах этот бит называется F-битом, а его значение устанавливается равным 1 и указывает на то, что данный кадр передается в ответ на команду запроса.

## Функционирование

Функционирование протокола HDLC заключается в обмене I-кадрами, S-кадрами и U-кадрами между двумя станциями. Различные команды и отклики, определенные для этих типов кадров, перечислены в табл. Г.1. Ниже, при описании работы протокола HDLC, мы рассмотрим эти три типа кадров.

**Таблица Г.1. Команды и отклики протокола HDLC**

Название	Команда/отклик	Описание
Информационный кадр (I)	K/O	Обмен пользовательскими данными
Кадр управления (S)		
Готов к приему (RR)	K/O	Положительное подтверждение; готовность к приему I-кадра
Не готов к приему (RNR)	K/O	Положительное подтверждение; неготовность к приему I-кадра
Отказ (REJ)	K/O	Отрицательное подтверждение; возврат
Выборочный отказ (SREJ)	K/O	Отрицательное подтверждение; выборочный отказ
Ненумерованный кадр (U)		
Выбор обычного/расширенного режима отклика (SNRM/SNRME)	K	Выбор режима; расширенный — семибитовые порядковые номера
Выбор обычного/расширенного асинхронного режима (SARM/SARME)	K	Выбор режима; расширенный — семибитовые порядковые номера
Выбор обычного/расширенного асинхронного сбалансированного режима (SABM/SABME)	K	Выбор режима; расширенный — семибитовые порядковые номера
Установка режима инициализации (SIM)	K	Инициализация функций управления каналом связи на станции-адресате
Разрыв соединения (DISC)	K	Разрыв логического соединения
Ненумерованное подтверждение (UA)	O	Подтверждение приема одной из команд выбора режима
Режим разрыва соединения (DM)	O	Станция-адресат находится в режиме разрыва соединения
Запрос на разрыв соединения (RD)	O	Запрос на команду DISC
Запрос на установление режима инициализации (RIM)	O	Необходима инициализация; запрос на команду SIM

Название	Команда/отклик	Описание
Ненумерованная информация (UI)	K/O	Используется для обмена управляющей информацией
Ненумерованный опрос (UP)	K	Используется для запроса управляющей информации
Сброс (RSET)	K	Используется для восстановления; обновляет значения N(R) и N(S)
Идентификация обмена (XID)	K/O	Используется для запроса/отчета о состоянии
Тестирование (TEST)	K/O	Обмен идентичными информационными полями с целью тестирования
Отклонение кадра (FRMR)	O	Отчет о получении неприемлемого кадра

Работа протокола HDLC происходит в три этапа. Вначале одна из сторон инициализирует канал передачи данных, чтобы обеспечить упорядоченный обмен кадрами. Во время этой фазы обговариваются все параметры, которые будут использоваться. После инициализации две стороны обмениваются данными и управляющей информацией с целью осуществления управления потоком и защиты от ошибок. И, наконец, на последнем этапе одна из сторон сообщает о прекращении работы.

### Инициализация

Инициализацию может запросить любая из сторон, передав одну из шести команд выбора режима. Эти команды необходимы для выполнения следующих действий.

1. Сообщение другой стороне о необходимости инициализации.
2. Определение режима передачи данных: одна из сторон является ведущей, и контролируется весь обмен информацией; две стороны являются равноправными и совместно управляют обменом.
3. Определение размера используемого порядкового номера (3 или 7 бит).

При принятии другой стороной этого запроса ее модуль HDLC передает исходной стороне кадр ненумерованного подтверждения (UA). При отклонении запроса передается кадр режима разрыва (DM).

### Передача данных

После запроса и принятия инициализации устанавливается логическое соединение. Обе стороны могут начинать передавать пользовательские данные, помещенные в I-кадры, начиная с порядкового номера 0. Поля N(S) и N(R) I-кадра — это порядковые номера, поддерживающие управление потоком и защиту от ошибок. При передаче модулем HDLC I-кадров последние последовательно нумеруются по модулю 8 или 128, в зависимости от размера порядковых номеров (соответственно 3 или 7 бит), после чего эти номера помещаются в поле N(S). Поле N(R) — это подтверждение приема I-кадров; оно позволяет модулю HDLC указывать номер следующего ожидаемого I-кадра.

S-кадры также используются для управления потоком и защиты от ошибок. Кадр готовности к приему (RR) указывает следующий ожидаемый I-кадр, подтверждая таким образом получение предыдущего. Кадр RR используется при отсутствии обратного потока пользовательских данных (I-кадров), которым могло бы передаваться подтверждение приема. Наличие в I-кадре команды RNR (Не готов к приему) также подтверждает получение предыдущего I-кадра, но, в отличие от команды RR, просит у однорангового объекта временно приостановить передачу I-кадров. Как только объект, передавший команду RNR, снова будет готов к приему, он пошлет команду RR. Команда REJ (отказ) инициирует возвратный механизм ARQ. Эта команда показывает, что прием последнего I-кадра был отклонен и что требуется повторная передача всех I-кадров, начиная с кадра с номером N(R). При использовании выборочного отказа (SREJ) запрашивается повторная передача не всех, а единственного кадра.

### Разрыв соединения

Инициировать разрыв соединения любой модуль HDLC может либо по собственной инициативе (при наличии некоторого сбоя), либо в ответ на запрос со стороны пользователя вышестоящего уровня. Собственно разрыв соединения происходит в ответ на передачу модулем HDLC кадра разрыва соединения (DISC). Удаленный объект должен принять разрыв, для чего он отвечает кадром UA.

### Примеры операций

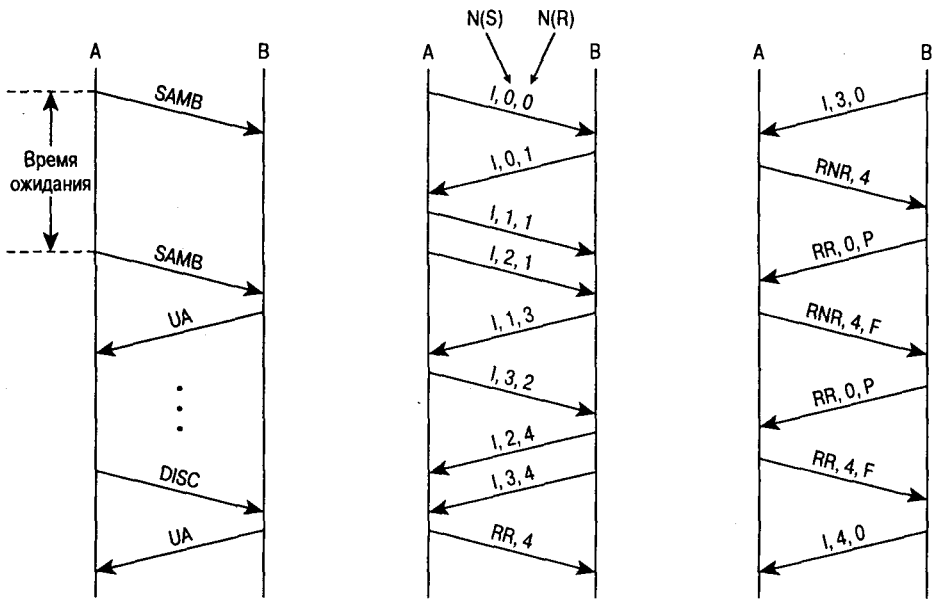
Для лучшего понимания операций протокола HDLC рассмотрим рис. Г.4. Каждая стрелка на представленных диаграммах подписана, указаны названия кадров, значение P/F-бита и, где это необходимо, значения полей N(R) и N(S). Если указана буква P или F, то значение соответствующего бита равно 1, в противном случае значение равно 0.

На рис. Г.4, а показаны кадры, участвующие в установке и разрыве соединения. Объект протокола HDLC одной стороны передает другой стороне команду SABM<sup>1</sup> и запускает таймер. Противоположная сторона получает команду SABM, возвращает отклик UA, устанавливает значения ее переменных и счетчиков и останавливает таймер. С этого момента логическое соединение считается активным и обе стороны могут начинать передачу кадров. Если по истечении времени, на которое установлен таймер, не будет получено подтверждение приема команды SABM, то исходная сторона передает эту команду повторно, как показано на рисунке. Этот процесс будет повторяться до тех пор, пока не будет получен отклик UA или DM либо пока попытка установки соединения не будет предпринята заданное число раз. В последнем случае управляющий объект получит сообщение о неудачной попытке установления соединения, после чего потребуется вмешательство со стороны вышестоящего уровня. Процедура разрыва соединения показана на том же рис. Г.4, а. Одна сторона передает команду DISC, а другая отвечает откликом UA.

---

<sup>1</sup> SABM (Set Asynchronous Balanced Mode) — выбор асинхронного сбалансированного режима. Команда SABM — это запрос начала обмена. ABM — название режима передачи, что для нас в данном случае не важно.

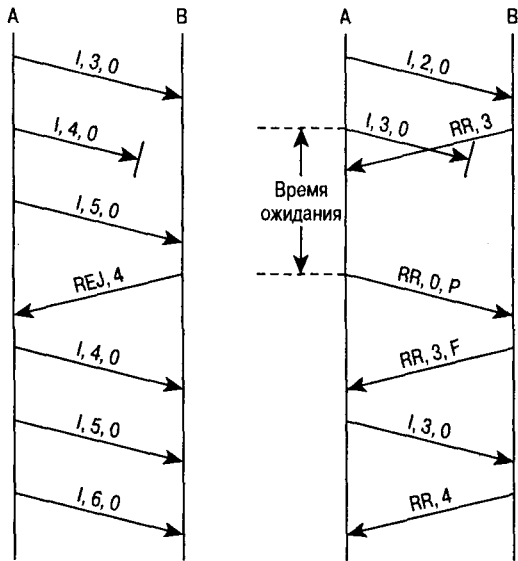




а) Установление и разрыв соединения

б) Двусторонний обмен данными

в) Состояние "занято"



г) Восстановление после отказа

д) Восстановление после истечения времени ожидания

Рис. Г.4. Примеры операций протокола HDLC

На рис. Г.4, б приведен пример дуплексного обмена I-кадрами. Если объект посылает подряд несколько I-кадров, при этом не получая какой-либо информации, то порядковые номера приема просто повторяются (например,  $I,1,1$ ;  $I,2,1$  в направлении от А к В). Если объект получает несколько I-кадров подряд, не передавая при этом какой-либо информации, то порядковый номер приема в следующем передаваемом кадре должен отражать общую ситуацию (например,  $I,1,3$  в направлении от В к А). Отметим, что помимо I-кадров в информационном обмене могут быть задействованы кадры управления.

На рис. Г.4, в показаны операции, касающиеся состояния “занято”. Такое состояние возникает, когда объект HDLC не может обработать I-кадры со скоростью их поступления или когда пользователь не может принять данные со скоростью их поступления в I-кадры. В любом случае приемный буфер объекта переполняется и объект должен остановить поступающий поток I-кадров, используя для этого команду RNR. В нашем примере объект А передает команду RNR, которая требует от объекта В остановить передачу I-кадров. Как правило, станция, получившая команду RNR, через некоторое время начинает периодически запрашивать станцию, находящуюся в состоянии “занято”, посылая для этого команду RR с заданным Р-битом. После снятия состояния “занято” объект А возвращает отклик RR и передача кадров от станции В может возобновиться.

На рис. Г.4, г показан пример восстановления после ошибки с использованием команды REJ. В этом примере станция А передает I-кадры с номерами 3, 4 и 5. Вследствие ошибки кадр 4 теряется. После того как станция В получает пятый I-кадр, она отклоняет его (поскольку он пришел не вовремя) и посылает команду REJ с полем  $N(R) = 4$ . Станция А получает эту команду и повторно передает все кадры, начиная с 4. После завершения повторной передачи станция А может продолжать передавать следующие кадры.

На рис. Г.4, д показан пример восстановления после ошибки по истечении времени ожидания. В этом примере станция А передает I-кадр с номером 3 как последний из набора I-кадров. При передаче в кадр вкрадывается ошибка. Станция В эту ошибку обнаруживает и отклоняет кадр. Поскольку при обнаружении ошибки в кадре “подозрительными” являются все биты этого кадра, то приемник не может отреагировать на него. Следовательно, станция В не может послать команду REJ, поскольку она не может определить, был ли отклоненный кадр I-кадром. В это же время после передачи кадра станция А запускает таймер, установленный на время, достаточное для получения отклика. По истечении заданного времени станция А инициирует действия по восстановлению, которые обычно заключаются в запросе станции В с помощью команды RR с установленным Р-битом, чтобы определить состояние этой станции. Поскольку данный запрос требует отклика, то объект А получает кадр, содержащий поле  $N(R)$ , и может продолжать свои действия. В нашем случае отклик указывает на потерю кадра 3, который станция А передает повторно.

Приведенные примеры не охватывают всех случаев, но они должны помочь читателю получить представление о работе протокола HDLC.

# СЛОВАРЬ ТЕРМИНОВ

Некоторые определения данного словаря взяты из *Нового национального стандартного словаря электрических и электронных терминов* (США) — стандарта IEEE 100-1992, 1993 г. Эти определения помечены звездочкой.

**Абонентская линия.** Линия связи, обычно витая пара, между отдельным абонентом и ближайшим узлом коммутации общедоступной сети связи.

**Автоматический запрос повторной передачи.** Функция, которая при обнаружении ошибки в передаче автоматически выдает запрос на повторную передачу.

**Амплитудная манипуляция.** Модуляция, при которой два двоичных значения представляются двумя различными амплитудами несущей частоты.

**Амплитудная модуляция\*.** Форма модуляции, при которой амплитуда несущего сигнала изменяется в соответствии с некоторой характеристикой модулирующего сигнала.

**Амплитудное замирание\*.** Замирание, при котором все частотные составляющие принятого сигнала флуктуируют одновременно и пропорционально.

**Аналоговая передача.** Передача аналоговых сигналов безотносительно к их содержанию. В ходе передачи сигнал может подвергаться усилению, но данные из него не извлекаются.

**Аналоговые данные\*.** Данные, представленные физической величиной, которая считается изменяющейся непрерывно и значение которой прямо пропорционально данным или некоторой функции данных.

**Аналоговый сигнал.** Непрерывно изменяющиеся электромагнитные колебания, которые могут распространяться в различных средах.

**Антенна.** Часть передающей или принимающей системы, разработанная для излучения или получения электромагнитных волн.

**Апериодический.** Функция или сигнал, не являющиеся периодическими.

**Асинхронный режим передачи (АТМ).** Способ передачи пакетов, при котором используются пакеты фиксированного размера, называемые ячейками. АТМ является интерфейсом передачи данных для В-ISDN. В отличие от X.25, АТМ не имеет механизмов управления потоком данных и защиты от ошибок.

**Атмосферное поглощение.** Потеря энергии при передаче радиоволн вследствие рассеяния в атмосфере.

**Беспроводная передача.** Передача через воздух, вакуум или воду посредством антенн.

**Бит четности\*.** Контрольный бит, добавляемый к массиву двоичных цифр, с той целью, чтобы сумма всех двоичных цифр, включая этот бит, всегда была четной либо чтобы она всегда была нечетной.

**Блочный код коррекции ошибок.** Схема, в которой  $k$ -битовый блок данных отображается в  $n$ -битовый ( $n > k$ ), именуемый кодовым словом. В процессе отображения используется кодер FEC.

**Виртуальный канал.** Режим пакетной коммутации, в котором в начале передачи между двумя станциями устанавливается соединение (виртуальный канал). Все пакеты передаются по одному и тому же маршруту, без необходимости указывать в них полный адрес, и прибывают в исходном порядке.

**Восходящий канал.** Канал связи от наземной станции к спутнику.

**Представление во временной области.** Описание характеристик функции или сигнала как функции времени.

**Геосинхронная орбита.** Круговая орбита в плоскости экватора. Для планеты с земной массой и радиусом данная орбита расположена в 66 454 км над поверхностью.

**Геостационарный спутник.** Геосинхронный спутник с нулевым склонением, так что спутник кажется зависшим над определенной точкой земной поверхности.

**Дейтаграмма\*.** При пакетной коммутации — пакет, независимый от других пакетов и содержащий достаточную информацию для его пересылки от терминального оборудования (DTE) отправителя до DTE получателя без необходимости устанавливать соединение между устройствами DTE и сетью.

**Децибел.** Мера сравнения силы двух сигналов. Число децибелов равно десятикратному десятичному логарифму отношения мощностей сигналов или двадцатикратному десятичному логарифму отношения напряжений сигналов.

**Диаграмма направленности.** Графическое представление характеристик излучения антенны как функции пространственных координат. Как правило, расстояние от антенны до любой точки диаграммы направленности пропорционально мощности, излученной антенной в этом направлении.

**Диполь\*.** Представитель класса антенн, диаграмма направленности которых близка к диаграмме элементарного электрического диполя.

**Дифракция\*.** Отклонение части луча, определяемое волновой природой излучения, которое происходит при прохождении луча через края непрозрачных препятствий.

- Длина волны\***. Расстояние между двумя точками периодической волны, имеющими одинаковую фазу.
- Доплеровский сдвиг\***. Эффективное изменение частоты принятого сигнала вследствие относительной скорости передатчика и приемника.
- Дуплекс с временным разделением**. Схема передачи, в которой в любой момент времени передача данных ведется только в одном направлении; направление передачи периодически меняется.
- Заголовок**. Системная управляющая информация, которая предшествует данным пользователя.
- Замирание\***. Изменение во времени мощности принятого сигнала, вызванное изменениями в среде или такте передачи.
- Затухание**. Уменьшение значения тока, напряжения или мощности сигнала при передаче.
- Земная радиоволна\***. Радиоволна, распространяющаяся над землей и обычно подвергающаяся влиянию со стороны земли и тропосферы. *Примечания:* 1. Земная радиоволна включает все компоненты радиоволн над поверхностью земли за исключением ионосферных и тропосферных волн. 2. Земная радиоволна преломляется вследствие изменения в тропосфере диэлектрической константы.
- Изохронность\***. Временная характеристика события или сигнала, повторяющегося через известные промежутки времени.
- Импульсная помеха**. Шумовой импульс с большой амплитудой и малой длительностью.
- Импульсно-кодовая модуляция**. Метод оцифровки сигнала, при котором величины отсчетов сигнала измеряются по отношению к некоторому заданному значению и посредством кодирования преобразуются в цифровой сигнал.
- Инкапсуляция**. Добавление объектом протокола к данным, полученным от пользователя протокола, управляющей информации.
- Интермодуляционный шум**. Помехи, возникающие вследствие нелинейной комбинации сигналов разных частот.
- Инфракрасный сигнал**. Электромагнитные волны с частотой от  $3 \times 10^{11}$  до  $4 \times 10^{14}$  Гц.
- Ионосфера**. Часть внешней атмосферы Земли, в которой на передачу радиоволн влияет ионизация, вызванная солнечной радиацией. Ионосфера находится на расстоянии 50–400 км от земной поверхности.
- Ионосферная волна\***. Радиоволна, распространяющаяся под углом к ионосфере и отражающаяся от нее.
- Канал\***. Отдельный тракт передачи электрических сигналов. *Примечание:* Слово *тракт* используется в широком смысле, включая обособление по принципу частотного или временного разделения. Под *каналом* может подразумеваться одно- или двухсторонний тракт.

- Код с коррекцией ошибок\***. Код, в котором каждое выражение соответствует определенным правилам построения, так что отклонение от данной структуры автоматически обнаруживается, и возможна автоматическая коррекция некоторых или всех ошибок.
- Код расширения**. Последовательность цифр, используемая для расширения полосы частот в системе расширенного спектра. Другие названия: расширяющая последовательность, раздробленный код.
- Код с обнаружением ошибок**. Код, в котором каждое выражение соответствует определенным правилам построения, поэтому если в выражении происходит какая-то ошибка, оно уже не будет соответствовать правилам построения, что позволяет обнаружить наличие ошибки.
- Коммутация каналов**. Метод связи, при котором между двумя устройствами устанавливается заранее определенный тракт через один или более промежуточных узлов коммутации. В отличие от пакетной коммутации, цифровые данные пересылаются как непрерывный поток битов. Пропускная способность гарантирована, а задержка практически определяется временем распространения сигнала. Коммутация каналов применяется в телефонных сетях.
- Контрольная последовательность кадра**. Код с обнаружением ошибок, вставленный в виде поля в блок передаваемых данных. Этот код предназначен для обнаружения ошибок при приеме данных.
- Контрольная сумма**. Код с обнаружением ошибок на основе операции суммирования, выполняемой над битами, которые требуется проверить.
- Коэффициент направленного действия (антенны)\***. Отношение мощности, излученной в данном направлении, к мощности излучения изотропной антенны такой же номинальной мощности.
- Локальная сеть**. Сеть связи, которая обеспечивает взаимосвязь разнообразных устройств работы с данными в пределах маленькой области.
- Максимальная амплитуда**. Максимальное значение напряжения сигнала; обычно измеряется в вольтах.
- Миллиметровый диапазон**. Неточный термин. Как правило, системами миллиметрового диапазона называют системы, работающие в области между 10 ГГц (длина волны = 30 мм) и 300 ГГц (длина волны = 1 мм).
- Многолучевое распространение**. Эффект распространения, в результате которого сигнал прибывает на приемник по двум трактам или большему числу трактов.
- Модуль данных протокола (PDU)\***. Набор данных, определенный в протоколе данного уровня и включающий управляющую информацию протокола этого уровня и, возможно, пользовательские данные этого уровня.
- Модуляция\***. Процесс или результат процесса изменения некоторых характеристик сигнала, называемого несущим сигналом, в соответствии с сигналом сообщения.
- Направляемая среда**. Среда передачи, в которой волны направляются вдоль твердой среды, такой, как медная витая пара, медный коаксиальный кабель или оптоволокно.

- Ненаправляемая среда.** Среда передачи, такая, как воздух, вакуум или вода, в которой для передачи и приема используются антенны.
- Несущая частота.** Непрерывная частота, которая может модулироваться или на которую может накладываться другой сигнал (содержащий информацию).
- Нисходящий канал.** Канал связи от спутника к наземной станции.
- Обратный канал.** В беспроводной или сотовой сети канал связи от мобильного устройства к базовой станции.
- Объединенная сеть (internet).** Совокупность сетей с коммутацией пакетов, соединенных посредством маршрутизаторов.
- Отражение.** Явление, происходящее при поступлении электромагнитного сигнала на поверхность, большую по сравнению с длиной волны сигнала; угол падения равен углу отражения.
- Пакет ошибок.** Пакет ошибок длиной  $B$  — это непрерывная последовательность  $B$  ошибочных битов.
- Пакет.** Группа битов, содержащая данные и управляющую информацию. Обычно пакетом называется модуль данных протокола сетевого уровня (уровня 3 OSI).
- Пакетная коммутация.** Метод передачи сообщений через сеть связи, при котором длинные сообщения разбиваются на короткие пакеты. Затем эти пакеты передаются как отдельные сообщения.
- Передающая среда.** Физический тракт между передатчиками и приемниками системы связи.
- Перекрестные помехи\*.** Нежелательная энергия, появляющаяся в одном тракте вследствие процессов, проходящих в других трактах.
- Период.** Абсолютная величина минимального интервала, по истечении которого повторяются прежние характеристики периодического сигнала.
- Периодический сигнал.** Сигнал  $f(t)$ , удовлетворяющий условию  $f(t) = f(t + nk)$  для всех целых  $n$  при постоянном  $k$ .
- Полуволновой диполь\*.** Проводная антенна, состоящая из двух прямых коллинеарных проводников равной длины, разделенных небольшой щелью, на которую подается сигнал; длина каждого проводника приблизительно равна четверти длины волны.
- Потери в свободном пространстве.** Потеря энергии сигнала, вызванная его рассеянием в пространстве.
- Поячеечная передача.** Механизм пакетной коммутации, применяемый для пакетов фиксированного размера, называемых ячейками. На технологии поячеечной передачи базируется АТМ.
- Преломление\*.** Отклонение луча от прямолинейного распространения при переходе между двумя различными средами или в среде, коэффициент преломления которой является непрерывной функцией координат.
- Пропускная способность канала\*.** Максимально возможная скорость передачи информации через канал, определяемая ограничениями канала.

- Протокол Internet.** Межсетевой протокол, обеспечивающий связь без установки соединения через несколько сетей с пакетной коммутацией.
- Протокол.** Набор правил, согласно которым работают функциональные модули при осуществлении связи.
- Прямое исправление ошибок.** Выполняемые приемником процедуры, которые заключаются в коррекции ошибок на основании информации, содержащейся в принятом сигнале.
- Прямой канал.** В беспроводной или сотовой сети канал связи от базовой станции к мобильному устройству.
- Псевдослучайная последовательность.** Последовательность чисел, выглядящих случайными со статистической точки зрения.
- Рассеяние\*.** Порождение волн с иным направлением, частотой или поляризацией при падении радиоволны на препятствие.
- Расстояние Хэмминга\*.** Число разрядов, в которых отличаются два двоичных числа равной длины.
- Расширение спектра методом прямой последовательности.** Разновидность расширения спектра, когда каждый бит исходного сигнала представляется несколькими битами передаваемого сигнала, для чего используется код расширения.
- Расширение спектра путем скачкообразной перестройки частоты.** Разновидность расширения спектра, при которой сигнал передается на разных радиочастотах якобы в случайной последовательности, переходя с одной частоты на другую через фиксированные промежутки времени.
- Расширенный спектр.** Метод распространения информации, содержащейся в сигнале, по расширенной полосе частот с использованием кода расширения.
- СВЧ.** Электромагнитные волны в диапазоне частот приблизительно от 1 до 40 ГГц.
- Сеансовый уровень.** Уровень 5 модели OSI. Управляет логическим соединением (сеансом связи) двух общающихся процессов или приложений.
- Селективное замирание\*.** Замирание, по-разному сказывающееся на различных спектральных компонентах радиосигнала.
- Сетевой уровень.** Уровень 3 в модели OSI. Отвечает за пересылку данных по сети связи.
- Синхронное уплотнение с временным разделением.** Метод уплотнения с временным разделением, при котором временные интервалы в общей линии передачи выделяются каналам ввода-вывода на неизменной заранее установленной основе.
- Собственная частота\*.** Наименьший частотный компонент в ряде Фурье для периодической величины.
- Сотовая сеть.** Беспроводная сеть связи, в которой станции общаются через стационарные антенны, расположенные в форме гексагональной структуры.
- Спектр.** Понятие, означающее абсолютный диапазон частот. Например, инфракрасный спектр простирается примерно от  $3 \times 10^{11}$  до  $4 \times 10^{14}$  Гц.
- Тепловой шум.** Статистически однородный шум, обусловленный температурой передающей среды.



- Точка доступа к службе.** Средство идентификации пользователем услуг объекта протокола. Объект протокола выделяет для использования объектам высшего уровня одну или более точек доступа.
- Транспортный уровень.** Уровень 4 модели OSI. Обеспечивает надежный, прозрачный перенос данных между конечными точками.
- Тропосфера\*.** Часть атмосферы, в которой температура уменьшается с высотой, формируются облака и действуют конвекционные потоки. Тропосфера начинается от поверхности Земли и заканчивается приблизительно на уровне от 6 км (полюса) до 18 км (экватор).
- Угловая модуляция\*.** Модуляция, при которой изменяется угол синусоидального несущего сигнала. Частными формами угловой модуляции являются фазовая и частотная модуляция.
- Уплотнение с временным разделением.** Разделение средства передачи между двумя или более каналами путем поочередного предоставления средства нескольким разным информационным каналам.
- Уплотнение с частотным разделением.** Разделение средства передачи между двумя или более каналами путем разделения полосы частот, в которой работает это средство, на более узкие полосы, каждая из которых образует отдельный канал.
- Уплотнение.** При передаче данных — функция, позволяющая двум или более источникам данных совместно использовать одну передающую среду таким образом, что каждый источник данных имеет собственный канал.
- Управление доступом к среде (MAC).** В широкоэмитательных сетях — метод, позволяющий определить, какое устройство обращалось к передающей среде в любой момент времени.
- Управление потоком данных.** Функция, выполняемая принимающим объектом с целью ограничения количества данных или скорости передачи данных, отправляемых передающим объектом.
- Уровень ошибок битов.** Вероятность того, что переданный бит будет принят с ошибкой.
- Уровень представления\*.** Уровень 6 модели OSI. Определяет общий синтаксис для представления данных и для преобразования данных приложения в соответствии с общим синтаксисом и в обратную сторону.
- Уровень приложений.** Уровень 7 модели OSI. Этот уровень определяет интерфейс системы с пользователем.
- Фаза.** Для периодического сигнала  $f(t)$  — дробная часть  $t/P$  периода  $P$ , на которую  $t$  сдвинута относительно произвольного начала координат. Началом координат обычно считается момент предыдущего перехода через нуль в направлении от отрицательных значений к положительным.
- Фазовая манипуляция.** Модуляция, при которой цифровые данные представляются сдвигом фазы несущего сигнала.
- Фазовая модуляция.** Модуляция, при которой изменяемой характеристикой является фазовый угол несущего сигнала.

- Физический уровень.** Уровень 1 модели OSI. Охватывает электрические, механические и временные аспекты передачи сигнала в среде.
- Хэш-функция.** Функция, отображающая блок данных или сообщение переменной длины на значение фиксированной длины, называемое хэш-кодом. Если эта функция засекречена, ее можно использовать для аутентификации данных или сообщения. Она также называется профилем сообщения.
- Циклическая проверка четности с избыточностью (CRC).** Код обнаружения ошибок, вычисляемый как остаток от деления проверяемых битов на заданное двоичное число.
- Цифровой сигнал.** Дискретный или прерывистый сигнал, например импульсы напряжения.
- Цифровые данные.** Данные, состоящие из последовательности дискретных элементов.
- Частота.** Количество колебаний сигнала в секунду; измеряется в герцах.
- Частотная манипуляция.** Модуляция, при которой два двоичных значения представляются двумя разными частотами в окрестности несущей частоты.
- Частотная модуляция.** Модуляция, при которой изменяемой характеристикой является частота переменного тока.
- Представление в частотной области.** Описание характеристик функции или сигнала через их частотные составляющие.
- Ширина полосы\*.** Разность между граничными частотами непрерывного спектра частот.
- Широкополосные системы.** В сфере передачи данных так обычно называются системы, предлагающие пользователям скорости передачи от 2 Мбит/с до сотен Мбит/с.
- Шум.** Ложные сигналы, которые объединяются с сигналом, предназначенным для передачи и приема, и искажают его.
- Электрический диполь\*.** Элементарный излучатель, состоящий из пары равных, осциллирующих в противофазе электрических зарядов на бесконечно малом расстоянии друг от друга.

# ЛИТЕРАТУРА

## Аббревиатуры

ACM (Association for Computing Machinery) —  
Ассоциация по вычислительной технике

IEEE (Institute of Electrical and Electronics Engineers) —  
Институт инженеров по электротехнике и электронике

ADAM91 Adamek, J. *Foundations of Coding*. — New York: Wiley, 1991.

AGRA99 Agrawal P., Sreenan C. Get Wireless: A Mobile Technology Spectrum. — In: *IT Pro*, August 1999.

ANDE95 Anderson J., Rappaport T., Yoshida S. Propagation Measurements and Models for Wireless Communications Channels. — *IEEE Communications Magazine*, January 1995.

ASH90 Ash R. *Information Theory*. — New York: Dover, 1990.

BANT94 Bantz D., Bauchot F. Wireless LAN Design Alternatives. — *IEEE Network*, March/April, 1994.

BELL00 Bellamy J. *Digital Telephony*. — New York: Wiley, 2000. // Беллами Дж. *Цифровая телефония*. — М: Радио и связь, 1986.

BERL87 Berlekamp E., Peile R., Pope S. The Application of Error Control to Communications. — *IEEE Communications Magazine*, April 1987.

BERR96 Berrou C., Glavieux A. Near Optimum Error Correcting Codes and Decoding: Turbo Codes. — *IEEE Transactions on Communications*, October 1996.

BERT92 Bertsekas D., Gallager R. *Data Networks*. — Englewood Cliffs, NJ: Prentice Hall, 1992. // Бертсекас Д., Галлагер Р. *Сети передачи данных*. — М: Мир, 1998.

BERT94 Bertoni H., Honcharenko W., Maciel L., Xia H. UHF Propagation Prediction for Wireless Personal Communications. — In: *Proceedings of the IEEE*, September 1994.

- BERT00** Bertoni H. *Radio Propagation for Modern Wireless Systems*. — Upper Saddle River, NJ: Prentice Hall, 2000.
- BHAR83** Bhargava V. Forward Error Correction Schemes for Digital Communications. — *IEEE Communications Magazine*, January 1983.
- BLAC99a** Black U. *Second-generation Mobile and Wireless Networks*. — Upper Saddle River, NJ: Prentice Hall, 1999.
- BLAC99b** Black U. *ATM Vol. I. Foundation for Broadband Networks*. — Upper Saddle River, NJ: Prentice Hall, 1992.
- BOLC01** Bolcskei H. et al. Fixed Broadband Wireless Access: State of the Art, Challenges, and Future Directions. — *IEEE Communications Magazine*, January 2001.
- BRAY01** Bray J., Sturman C. *Bluetooth: Connect Without Cables*. — Upper Saddle River, NJ: Prentice Hall, 2001.
- BUCH99** Bushanan B. *Handbook of Data and Communication Networks*. — Boston: Kluwer Academic Publishers, 1999.
- CARN99a** Carne E. *Telecommunications Primer: Data, Voice, and Video Communications*. — Upper Saddle River, NJ: Prentice Hall, 1999.
- CARN99b** Carne E. *Telecommunications Topics: Applications of Functions and Probabilities in Electronic Communications*. — Upper Saddle River, NJ: Prentice Hall, 1999.
- CISC00** Cisco Systems, Inc. *Overcoming Multipath in Non-Line-of-Sight High-Speed Microwave Communication Links*. White Paper, 2000 ([www.cisco.com](http://www.cisco.com)).
- COTT00** Cotterell R., Langhammer M., Mauer V. Turbo Decoding for Comm Apps. — *Communications Systems Design*, August 2000.
- COUC01** Couch L. *Digital and Analog Communication Systems*. — Upper Saddle River, NJ: Prentice Hall, 2001.
- CROW97** Crow B. et al. IEEE 802.11 Wireless Local Area Networks. — *IEEE Communications Magazine*, September 1997.
- DALK96** Dalke R., Hufford G., Ketchum R. *Radio Propagation Considerations for Local Multipoint Distribution Systems*. National Telecommunications and Information Administration Publication PB97116511, August 1996.
- DAUM82** Daumer W. Subjective Evaluation of Several Efficient Speech Coders. — *IEEE Transactions on Communications*, April 1982.
- DINA98** Dinan E., Jabbari B. Spreading Codes for Direct Sequence CDMA and Wideband CDMA Cellular Networks. — *IEEE Communications Magazine*, September 1998.
- DIX094** Dixon R. *Spread Spectrum Systems with Commercial Applications*. — New York: Wiley, 1994.
- ECON99** *The World in Your Pocket*. — *The Economist*, October 1999.
- EDEL82** Edelson B., Marsten R., Morgan W. Greater Message Capacity for Satellites. — *IEEE Spectrum*, March 1982.

- ELBE99** Elbert B. *Introduction to Satellite Communication*. — Boston: Artech House, 1999.
- ENGE00** Englemann R. The Origins of Radio. — *IEEE Potentials*, October/November 2000.
- EVAN98** Evans J. New Satellite for Personal Communications. — *Scientific American*, April 1998.
- EVER94** Everitt D. Traffic Engineering of the Radio Interface for Cellular Mobile Networks. — In: *Proceedings of the IEEE*, September 1994.
- FREE96** Freeman R. *Telecommunication System Engineering*. — New York: Wiley, 1996.
- FREE97** Freeman R. *Radio System Design for Telecommunications*. — New York: Wiley, 1997.
- FREE98a** Freeman R. *Telecommunications Transmission Handbook*. — New York: Wiley, 1998.
- FREE98b** Freeman R. Bits, Symbols, Bauds, and Bandwidth. — *IEEE Communications Magazine*, April 1998.
- FREE99** Freeman R. *Fundamentals of Telecommunications*. — New York: Wiley, 1999.
- FUNG98** Fung P. A Primer on MMDS Technology. — *Communication Systems Design*, April 1998. (csdmag.com).
- GARG99** Garg V., Wilkes J. *Principles and Applications of GSM*. — Upper Saddle River, NJ: Prentice Hall, 1999.
- GEIE99** Geier J. *Wireless LANs*. — New York: Macmillan Technical Publishing, 1999.
- GEIE01** Geier J. Enabling Fast Wireless Networks with OFDM. — *Communications System Design*, February 2001 (www.csdmag.com).
- GIBS93** Gibson J. *Principles of Digital and Analog Communications*. — New York: Macmillan, 1993.
- GIBS97** Gibson J., ed. *The Communications Handbook*. — Boca Raton, FL: CRC Press, 1997.
- GLOV98** Glover I., Grant P. *Digital Communications*. — Upper Saddle River, NJ: Prentice Hall, 1998.
- GOOD97** Goodman D. *Wireless Personal Communications Systems*. — Reading, MA: Addison-Wesley, 1997.
- GUNT00** Gunther N. *The Practical Performance Analyst*. — Lincoln, NE: Authors Choice Press, 2000.
- HAAR98** Haartsen J. Bluetooth: The Universal Radio Interface for ad hoc, Wireless Connectivity. — *Ericsson Review*, No. 3, 1998 (www.ericsson.se).
- HAAR00a** Haartsen J. The Bluetooth Radio System. — *IEEE Personal Communications*, February 2000.
- HAAR00b** Haartsen J., Mattisson S. Bluetooth — A New Low-Power Radio Interface Providing Short-Range Connectivity. — In: *Proceedings of the IEEE*, October 2000.

- HAAS00** Haas Z. *Wireless and Mobile Networks*. — In: [TERP00].
- HATA80** Hata M. Empirical Formula for Propagation Loss in Land Mobile Radio Services. — *IEEE Transactions on Vehicular Technology*, March 1980.
- HAYK94** Haykin S. *Communication Systems*. — New York: Wiley, 1995.
- HEIJ00** Heijden M., Taylor M., eds. *Understanding WAP: Wireless Applications, Devices, and Services*. — Boston: Artech House, 2000.
- HSU84** Hsu H. *Applied Fourier Analysis*. — New York: Harcourt Brace, 1984.
- INGL97** Inglis A., Luther A. *Satellite Technology: An Introduction*. — Boston: Focal Press, 1997.
- JAIN90** Jain R. Covolutional Codes Improve Bit-Error Rate in Digital Systems. — In: *EDN*, August 20, 1990.
- JAME95** James J. *A Student's Guide to Fourier Transforms*. — Cambridge, England: Cambridge University Press, 1995.
- JAYA84** Jayant N., Noll P. *Digital Coding of Waveforms*. — Englewood Cliffs, NJ: Prentice Hall, 1984.
- JONE93** Jones E. *Digital Transmission*. — New York: McGraw-Hill, 1993.
- KAHN97** Kahn J., Barry J. Wireless Infrared Communications. — In: *Proceedings of the IEEE*, February 1997.
- KAMM00** Kammler D. *A First Course in Fourier Analysis*. — Upper Saddle River, NJ: Prentice Hall, 2000.
- KELL00** Keller T., Keller H. Adaptive Multicarrier Modulation: A Convenient Framework for Time-Frequency Processing in Wireless Communication. — In: *Processings of the IEEE*, May 2000.
- KNUT98** Knuth D. *The Art of Computing Programming. Vol. 2. Seminumerical Algorithms*. — Reading, MA: Addison-Wesley, 1998 // Кнут Д. Э. *Искусство программирования. Т. 2. Получисленные методы*, 3-е изд. — М: Издательский дом "Вильямс", 2000.
- KROO86** Kroon P, Deprettere E. Regular Pulse Excitation — A Novel Approach to Effective Multipulse Coding of Speech. — *IEEE Transactions on Acoustics, Speech and Signal Processing*, No. 5, 1986.
- LATH98** Lathi B. *Modem Digital and Analog Communication Systems*. — New York: Oxford University Press, 1998.
- LEBO98** Lebow I. *Understanding Digital Transmission and Recording*. — New York: IEEE Press, 1998.
- MACW76** Macwilliams F., Sloane N. Pseudo-Random Sequences and Arrays. — In: *Proceedings of the IEEE*, December 1976 (имеется также в [TANT98]).
- MANN99** Mann S. The Wireless Application Protocol. — *Dr. Dobb's Journal*, October 1999.
- MARK99** Marks R. The IEEE 802.16 Working Group on Broadband Wireless. — *IEEE Network*, March/April 1999.

- MART94** Martine R. *Basic Traffic Analysis*. — Upper Saddle River, NJ: Prentice Hall, 1994.
- MCDY99** McDysan D., Spohn D. *ATM: Theory and Applications*. — New York: McGraw-Hill, 1999.
- METT99** Mettala R. et al. *Bluetooth Protocol Architecture Version 1.0*. Bluetooth Whitepaper 1.C.120/1.0, 25 August 1999 ([www.bluetooth.com](http://www.bluetooth.com)).
- MILL01** Miller B., Bisdikian C. *Bluetooth Revealed*. — Upper Saddle River, NJ: Prentice Hall, 2001.
- MOSH89** Moshos G. *Data Communications: Principles and Problems*. — New York: West Publishing Co., 1989.
- MURH98** Murhammer M. et al. *TCP/IP: Tutorial and Technical Overview*. — Upper Saddle River, NJ: Prentice Hall, 1998.
- NORD00** Nordbotten A. LMDS Systems and Their Application. — *IEEE Communications Magazine*, June 2000.
- OHAR99** Ohara B., Petrick A. *IEEE 802.11 Handbook: A Designer's Companion*. — New York: IEEE Press, 1999.
- OJAN98** Ojanpera T., Prasad G. An Overview of Air Interface Multiple Access for IMT-2000/UMTS. — *IEEE Communications Magazine*, September 1998.
- OKUM68** Okumura T. et al. Field Strength and Its Variability in VHF and UHF Land Mobile Radio Service. — In: *Rev. Elec. Communication Lab*. 1968.
- ORLI98** Orlik P., Rappaport S. Traffic Performance and Mobility Modeling of Cellular Communications with Mixed Platforms and Highly Variable Mobilities. — In: *Proceedings of the IEEE*, July 1998.
- ORTI00** Ortiz S. Broadband Fixed Wireless Travels the Last Mile. — *Computer*, July 2000.
- PAHL95** Pahlavan K., Probert T., Chase M. Trends in Local Wireless Networks. — *IEEE Communications Magazine*, March 1995.
- PAPA97** Papazian P., Hufford G., Aschatz R., Hoffman R. Study of the Local Multipoint Distribution Service Radio Channel. — *IEEE Transactions on Broadcasting*, June 1997.
- PARK88** Park S., Miller K. Random Number Generators: Good Ones are Hard to Find. — *Communications of the ACM*, October 1998.
- PASS00** Passani L. Creating WAP Services. — *Dr. Dobb's Journal*, July 2000.
- PEAR92** Pearson J. *Basic Communication Theory*. — Englewood Cliffs, NJ: Prentice Hall, 1992.
- PERK97** Perkins C. Mobile IP. — *IEEE Communications Magazine*, May 1997.
- PERK98** Perkins C. Mobile Networking Through Mobile IP. — *IEEE Internal Computing*, January-February 1998.
- PETE61** Peterson W., Brown D. Cyclic Codes for Error Detection. — In: *Proceedings of the IEEE*, January 1961.

- PETE95** Peterson R., Ziemer R., Borth D. *Introduction to Spread Spectrum Communications*. — Englewood Cliffs, NJ: Prentice Hall, 1995.
- PHIL98** Phillips J., Namee G. *Personal Wireless Communications with DECT and PWT*. — Boston: Artech House, 1998.
- PICH97** Pichna R., Wang Q. *Power Control*. — In: [GIBS97].
- PICK82** Pickholtz R., Schilling D., Milstein L. Theory of Spread Spectrum Communications — A Tutorial. — *IEEE Transactions on Communications*, May 1982 (имеется также в [TANT98]).
- POLL96** Pollini G. Trends in Handover Design. — *IEEE Communications Magazine*, March 1996.
- POLL98** Polle I. *Your Guide to Propagation*. Potters Bar, Herts., U.K.: Radio Society of Great Britain, 1998.
- PRAS98** Prasad R., Ojanpera T. An Overview of CDMA Evolution: Toward Wideband CDMA. — *IEEE Communications Surveys*, Fourth Quarter 1998 (<http://www.comsoc.org/>).
- PRAS00** Prasad R., Mohr W., Konhauser W., eds. *Third Generation Mobile Communication Systems*. — Boston: Artech House, 2000.
- PROA01** Proakis J. *Digital Communications*. — New York: McGraw-Hill, 2001.
- RABI95** Rabiner L. Toward Vision 2001: Voice and Audio Processing Considerations. — *AT&T Technical Journal*, March/April 1995.
- RAHN93** Rahnema M. Overview of the GSM System and Protocol Architecture. — *IEEE Communications Magazine*, April 1993.
- RAMA88** Ramabadrans T., Gaitonde S. A Tutorial on CRC Computations. — *IEEE Micro*, August 1988.
- RAPP97** Rappaport T., Rias M., Kapoor V. *Propagation Models*. — In: [GIBS97].
- RODB00** Rodbell M. Bluetooth: Wireless Local Access, Baseband and RF Interfaces, and Link Management. — *Communications System Design*, March, April, May 2000 ([www.csdmag.com](http://www.csdmag.com)).
- SCHI00** Shiller J. *Mobile Communications*. — Reading, MA: Addison-Wesley, 2000.  
// Й. Шиллер, *Мобильные коммуникации*. — М: Издательский дом "Вильямс", 2002.
- SKLA93** Sklar B. Defining, Designing, and Evaluating Digital Communication Systems. — *IEEE Communications Magazine*, November 1993.
- SKLA97a** Sklar B. Rayleigh Fading Channels in Mobile Digital Communication systems. — *IEEE Communications Magazine*, July 1997.
- SKLA97b** Sklar B. A Primer on Turbo Code Concepts. — *IEEE Communications Magazine*, December 1997.
- SKLA01** Sklar B. *Digital Communications: Fundamentals and Applications*. Upper Saddle River, NJ; Prentice Hall, 2001. // Склар Б. *Цифровая связь: теоретические основы и практическое применение*. — М: Издательский дом "Вильямс", 2003.



- SOLO98** Solomon J. *Mobile IP: The Internet Unplugged*. — Upper Saddle River, NJ; Prentice Hall PTR, 1998.
- SPOH97** Spohn D. *Data Network Design*. — New York: McGraw-Hill, 1994.
- SPRA91** Spragins J., Hammond J., Pawlikowski K. *Telecommunications Protocols and Design*. — Reading, MA: Addison-Wesley, 1991.
- STAL99** Stallings W. *Cryptography and Network Security: Principles and Practice*, 2nd Edition. — Upper Saddle River, NJ: Prentice Hall, 1999. // Столлингс В. *Криптография и защита сетей: принципы и практика*, 2-е изд. — М.: Издательский дом “Вильямс”, 2001.
- STAL00** Stallings, W. *Data and Computer Communications*, 6th ed. — Upper Saddle River, NJ: Prentice Hall, 2000 // Столлингс В. *Компьютерные системы передачи данных*, 6-е изд. — М.: Издательский дом “Вильямс”, 2002.
- TANN95** Tanner M. *Practical Queueing Analysis*. — New York: McGraw-Hill, 1995.
- TANT98** Tantaratana S., Ahmed K., eds. *Wireless Applications of Spread Spectrum Systems: Selected Readings*. — Piscataway, NJ: IEEE Press, 1998.
- TERP00** Terplan K., Morreale P. Eds. *The Telecommunications Handbook*. — Boca Raton, FL: CRC Press, 2000.
- THUR00** Thurwachter C. *Data and Telecommunications: Systems and Applications*. — Upper Saddle River, NJ: Prentice Hall, 2000.
- VUCE00** Vucetic B., Yuan J. *Turbo Codes: Principles and Applications*. — Boston: Kluwer Academic Publishers, 2000.
- WAPF98** WAP Forum. *Wireless Application Environment Overview*. — WAP Forum document, April 1998.
- WEBB00** Webb W. *Introduction to Wireless Local Loop: Broadband and Narrowband Systems*. — Boston: Artech House, 2000.
- WILS00** Wilson J., Kronz J. *Inside Bluetooth: Part I and Part II*. — *Dr. Dobb's Journal*, March, April 2000.
- XION00** Xiong F. *Digital Modulation Techniques*. — Boston: Artech House, 2000.
- ZENG00** Zeng M., Annamalai A., Bhargava V. *Harmonization of Global Third-Generation Mobile Systems*. — *IEEE Communications Magazine*, December 2000.

# СОКРАЩЕНИЯ

2G	Second Generation (второе поколение)
3G	Third Generation (третье поколение)
AM	Amplitude Modulation (амплитудная модуляция)
AMPS	Advances Mobile Phone System (усовершенствованная система мобильной телефонной связи)
ARQ	Automatic Repeat Request (автоматический запрос повторной передачи)
ASK	Amplitude-Shift Keying (амплитудная манипуляция)
ATM	Asynchronous Transfer Mode (асинхронный режим передачи)
BER	Bit Error Rate (уровень ошибок битов)
bps	Bits Per Second (биты в секунду, бит/с)
CDMA	Code Division Multiple Access (множественный доступ с кодовым разделением каналов)
CRC	Cyclic Redundancy Check (циклическая проверка четности с избыточностью)
DAMA	Demand Assignment Multiple Access (множественный доступ с распределением по запросу)
dBm	decibel-milliWatt (децибел-милливатт, дБмВт)
dBW	decibel-Watt (децибел-ватт, дБВт)
DECT	Digital Enhanced Cordless Telecommunications (цифровые расширенные беспроводные телекоммуникации)
DM	Delta Modulation (дельта-модуляция)
DSSS	Direct Sequence Spread Spectrum (расширение спектра методом прямой последовательности)
ECC	Error Correction Code (код с коррекцией ошибок)
EHF	Extremely High Frequency (крайне высокая частота, КВЧ)
ELF	Extremely Low Frequency (крайне низкая частота, КНЧ)
FAMA	Fixed Assignment Multiple Access (множественный доступ с фиксированным распределением)
FCC	Federal Communications Commission (Федеральная комиссия по средствам связи США)
FCS	Frame Check Sequence (контрольная последовательность кадра)
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 (расширение спектра методом скачкообразной перестройки частоты)
FM	Frequency Modulation (частотная модуляция)
FSK	Frequency-Shift Keying (частотная манипуляция)
GEOS	Geostationary (геостационарный); Geosynchronous (геосинхронный)

GSM	Global System for Mobile Communications (глобальная система мобильной связи)
GW	Ground Wave (земная волна)
HDLC	High-Level Data Link Control (высокоуровневый протокол управления каналом)
HF	High Frequency (высокая частота, ВЧ)
IAB	Internet Architecture Board (Совет по архитектуре Internet)
ICMP	Internet Control Message Protocol (протокол управляющих сообщений Internet)
IEEE	Institute of Electrical and Electronics Engineers (Институт инженеров по электротехнике и радиоэлектронике, ИИЭР)
IETF	Internet Engineering Task Force (Проблемная группа проектирования Internet)
IP	Internet Protocol (протокол Internet)
ISO	International Organization for Standardization (Международная организация по стандартизации)
ITU	International Telecommunication Union (Международный телекоммуникационный союз)
ITU-T	ITU Telecommunication Standardization Sector (Сектор стандартизации телекоммуникаций Международного телекоммуникационного союза)
LAN	Local Area Network (локальная сеть)
LEO	Low Earth Orbit (низкая околоземная орбита)
LF	Low Frequency (низкая частота, НЧ)
LLC	Logical Link Control (управление логическим каналом)
LMDS	Local Multipoint Distribution Service (локальная многоточечная распределительная служба)
LOS	Line of Sight (линия прямой видимости)
LSFR	Linear Feedback Shift Register (линейный регистр сдвига с обратной связью)
MAC	Medium Access Control (управление доступом к среде)
MAN	Metropolitan-Area Network (городская сеть)
MEO	Medium Earth Orbit (средняя околоземная орбита)
MF	Medium Frequency (средняя частота, СЧ)
MMDS	Multichannel Multipoint Distribution Service (многоканальная многоточечная распределительная служба)
MOS	Mean Opinion Score (усредненная оценка разборчивости речи)
MSS	Mobile Service Satellite (спутник службы мобильной связи)
OFDM	Orthogonal Frequency Division Multiplexing (ортогональное уплотнение с частотным разделением)
OSI	Open Systems Interconnection (взаимодействие открытых систем)
PCM	Pulse Code Modulation (импульсно-кодовая модуляция)
PDU	Protocol Data Unit (модуль данных протокола)
PM	Phase Modulation (фазовая модуляция)
PN	Pseudonoise (псевдошум)
PRN	Pseudorandom Number (псевдослучайное число)
PSK	Phase-Shift Keying (фазовая манипуляция)
PWT	Personal Wireless Telecommunications (персональные беспроводные телекоммуникации)

QAM	Quadrature Amplitude Modulation (квадратурная амплитудная модуляция)
QoS	Quality of Service (качество обслуживания)
RF	Radio Frequency (радиочастота)
SAP	Service Access Point (точка доступа к службе)
SDU	Service Data Unit (модуль данных службы)
SHF	Super High Frequency (сверхвысокая частота, СВЧ)
SS	Spread Spectrum (расширенный спектр)
SW	Sky Wave (ионосферная волна); Short Wave (короткая волна)
TCP	Transmission Control Protocol (протокол управления передачей)
TDD	Time Division Duplex (дуплекс с временным разделением)
TDM	Time Division Multiplexing (уплотнение с временным разделением)
TDMA	Time Division Multiple Access (множественный доступ с временным разделением)
UDP	User Datagram Protocol (протокол пользовательских дейтаграмм)
UHF	Ultra High Frequency (ультравысокая частота, УВЧ)
VF	Voice Frequency (частота речевого диапазона)
VHF	Very High Frequency (очень высокая частота, ОВЧ)
VLf	Very Low Frequency (очень низкая частота, ОНЧ)
WAN	Wide Area Network (глобальная сеть)
WAP	Wireless Application Protocol (протокол беспроводных приложений)
WLL	Wireless Local Loop (беспроводная абонентская линия)
XOR	Exclusive OR (исключающее ИЛИ)

# ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

## A

AMPS, 24; 347; 349  
    распределение частот, 347  
ATM, 84; 85; 89  
    услуги, 92

## B

BER, 170; 244  
Bluetooth, 535-578

## D

DECT, 387-393

## F

FDM, 60  
FDMA, 312  
FTP, 105

## G

GSM, 356-363

## I

IETF, 25  
internet, 110  
Internet Draft, 586  
IP, 99-119  
    дейтаграмма, 104  
    заголовок, 119  
IPv6, 122  
ITU, 24; 583  
ITU-R, 584  
ITU-T, 584

## L

L2CAP, 541  
LMP, 540

## M

Mobile IP, 435-449  
    терминология, 439  
MOS, 400

## O

OSI, 100  
OSI (взаимодействие открытых систем), 106  
    сравнение с TCP/IP, 108  
    стандартизация на основе архитектуры, 110  
    уровни, 106

## R

RFCOMM, 541

## S

SDP, 541  
SMTP, 105

## T

TCP, 104-124  
    заголовок, 104; 105; 124  
    управление потоком, 123  
TCP/IP, 101; 105  
    архитектура протоколов, 101  
    действие, 102  
    межсетевой уровень, 102  
    транспортный уровень, 102

уровень доступа к сети, 102  
уровень приложений, 102  
физический уровень, 101  
TCS BIN, 541  
TDM, 60  
TELNET, 106

## U

UDP, 126

## V

VCC, 85  
VPC, 85

## W

WAP, 451

## A

Абонент, 75  
Абонентская линия, 75  
Автоматический запрос повторной  
передачи, 278  
Адресация, 110  
Амплитудная манипуляция, 172;  
173  
Амплитудная модуляция, 184  
Анализ Фурье, 602–604  
Аналоговая передача сигналов, 168  
Антенна, 132  
    диаграмма направленности,  
    132; 135  
    диполь, 133  
    изотропная, 132  
    Маркони, 133  
    направленная, 132  
    параболическая отражающая,  
    134  
    эффективная площадь, 136  
Антенна  
    вибратор Герца, 132  
    диаграмма приема, 133  
    коэффициент усиления, 136  
    ширина луча, 133  
Архитектура протоколов, 100

## Б

Безопасность  
    в Bluetooth, 564  
    в Mobile IP, 445  
    в беспроводных сетях, 495

## В

Верхняя боковая полоса, 185  
Виртуальный канал, 81  
Виртуальный тракт, 85  
Выравнивание, 159

## Г

Глобальные сети, 70  
Городские сети, 72

## Д

Данные  
    аналоговые, 44  
    цифровые, 44  
Двухполосная связь с передачей  
несущей, 184  
Дейтаграмма, 80  
Дельта-модуляция, 194  
Диаграмма  
    решетчатая, 272  
    состояний, 271  
Дифракция, 153  
Документы RFC, 586  
Документы, не проходящие путем  
стандарта, 588  
Дуплекс  
    с временным разделением  
    (TDD), 387

## З

Замирание, 153  
Затухание, 144  
Защита от ошибок, 243; 283  
    ARQ, 278  
    CRC, 246  
    FEC, 253  
    проверка четности, 245

Защитная полоса, 60  
Зоны Френеля, 406

## И

Избыточность кода, 257  
Импульсно-кодовая модуляция, 191

## К

Кадр, 61  
Канал, 60  
Канал связи  
    восходящий, 296  
    нисходящий, 296  
Квадратурная амплитудная  
    модуляция, 183  
Квадратурная фазовая  
    манипуляция, 176  
Код  
    RSC, 278  
    БХЧ, 263; 267  
    Рида-Соломона, 263; 268  
    сверточный, 270  
    циклический, 262  
Код Уолша, 234  
Кодирование, 243  
    сигналов, 167  
Коды  
    блочные, 253  
    ортогональные, 233  
Коммутатор, 76  
Коммутация  
    методы, 72  
    каналов, 74; 77  
    пакетов, 77; 80  
Коэффициент ошибок битов, 170

## Л

Локальные сети, 71

## М

Магистраль, 76  
Маркер, 575  
Маршрутизатор, 111; 112

Международная организация по  
    стандартам (ISO), 106  
Микроячейка, 329  
Многоуровневая фазовая  
    манипуляция, 178  
Многочастотная манипуляция, 174  
Множественный доступ  
    CDMA, 366  
    DAMA, 313  
    FAMA, 313  
    с временным разделением  
        (TDMA), 317  
    с кодовым разделением (CDMA),  
        218; 365  
    с частотным разделением  
        (FDMA), 310; 312  
Мобильность, 436  
Модуляция  
    ADPCM, 395  
    CVSD, 561  
Мост, 111

## Н

Нижняя боковая полоса, 185

## О

Общество Internet, 585  
Однополосная связь, 186  
Организации Internet, 586  
Отражение, 153

## П

Передача  
    анализ Фурье, 601  
    аналоговая, 48  
    затухание, 144  
    цифровая, 49  
    терминология, 170  
    частотное представление сигнала  
Передающая среда, 53  
Переключение, 329  
Пикосеть, 536  
Помехоустойчивость, 171  
Порт, 103  
Последовательность Касами, 233

Голда, 231  
Постоянная составляющая, 602  
Предложенный стандарт, 587  
Преломление, 141  
Проблемная группа проектирования  
  Internet (IETF), 586  
Проект Internet, 586  
Проект стандарта, 587  
Пропускная способность, 49  
  безошибочная, 52  
Протокол  
  FTP, 105  
  IP, 99–119  
  IGMP, 542  
  L2CAP, 570  
  LMP, 566  
  TCP, 104–124  
Протоколы  
  архитектура, 100  
  основные элементы, 101  
Процесс стандартизации, 586  
Псевдослучайные  
  последовательности, 223

## Р

Радиосигналы  
  преломление, 141  
  распространение, 137  
Разнесение, 159  
Рассеяние, 153  
Расширение  
  множественное, 235  
Расширение спектра  
  прямая последовательность  
  (DSSS), 213  
  скачкообразная перестройка  
  частоты (FHSS), 207

## С

Сегментация, 112  
Сети передачи данных, 70  
  глобальные сети, 70  
  городские сети, 72  
  локальные сети, 71  
  пикосеть, 542  
  рассеянная, 542

Сигнал, 36; 44  
  аналоговый, 45  
  собственная частота, 41  
  цифровой, 45  
  ширина полосы, 41  
Синхронизация, 101; 171  
Слот, 62  
Собственная гармоника, 602  
Собственная частота, 602  
Совет по архитектуре Internet (IAB),  
  586  
Сотовая связь, 326  
Спектр сигнала, 171  
Спектральная плотность мощности,  
  605  
Спутники  
  геостационарные, 299  
  низкоорбитальные, 302  
  орбиты, 297  
  среднеорбитальные, 304  
  угол возвышения, 297  
Среда, 53  
Стандарт  
  802.11b, 25  
  Bluetooth, 538  
  HiperLAN, 25  
  IEEE 802, 506  
  IEEE 802.11, 505  
  IEEE 802.16, 415  
  процесс стандартизации, 587  
Станция, 72  
  наземная, 296  
Степень кодирования, 257  
Схема кодирования, 170

## Т

Теорема  
  о дискретном представлении, 202  
Тепловой шум, 147  
Точки доступа к службе (SAP), 110  
Транспондер, 56; 296  
Трафик  
  интенсивность, 343  
  регулирование, 342  
Турбокодирование, 276



## У

Угловая модуляция, 187

Узел, 72

Уплотнение, 59

ортогональное, с частотным  
разделением (OFDM), 410

с временным разделением (TDM),  
60

с частотным разделением (FDM),  
60; 311

Управление потоком, 279

Управляющий комитет группы IETF  
(IESG), 586

## Ф

Фазовая манипуляция, 172; 175

Фазовая модуляция, 187

Формула Шеннона (для пропускной  
способности канала), 50

Форум ATM, 96

Форум WAP, 451

Фрагментация, 112

## Ц

Цифровая передача сигналов, 168

## Ч

Частота

собственная, 602

Частота битовых ошибок, 170

Частотная манипуляция, 172; 173

Частотная модуляция, 187

## Ш

Шум, 147

импульсные помехи, 149

интермодуляционный, 148

перекрестные помехи, 149

## Я

Ячейка, 85; 89