

**СОВРЕМЕННЫЕ
ТЕЛЕКОММУНИКАЦИОННЫЕ
СИСТЕМЫ**

Учебно-методический комплекс

2011

УДК 681.326(075)
ББК 32.973.202
075

Авторы-составители: преподаватель кафедры инженерной физики УО «ВГУ им. П.М. Машерова» **Е.А. Краснобаев**; студент V курса физического факультета УО «ВГУ им. П.М. Машерова» **Р.В. Парфинович**

Рецензенты:

заведующий кафедрой общей физики и астрономии, кандидат педагогических наук,
доцент *И.В. Галузо*; директор НУП «Информационно-обучающие технологии»,
кандидат технических наук *В.А. Ключников*

Учебно-методический комплекс по дисциплине специализации «Современные телекоммуникационные системы» предназначен для студентов специальности «Физика» (производственная деятельность) и составлен в соответствии с учебной программой указанного курса для специализации 1-31 04-01-02-16 «Компьютерное моделирование физических процессов».

УДК 681.326(075)
ББК 32.973.202

© УО «ВГУ им. П.М. Машерова», 2011

СОДЕРЖАНИЕ

Введение	7
Структурное содержание курса	8
Раздел I. Теоретические основы	9
Лекция 1. Элементы теории передачи информации	9
1.1. Информация, сообщение, сигнал.....	9
1.2. Преобразование непрерывных сообщений в цифровую форму.....	13
1.3. Обобщенная структурная схема системы связи. Канал связи.	14
1.4. Информационные характеристики источника дискретных сообщений и канала связи	18
1.5. Скорость передачи информации по дискретному каналу. Пропускная способность	24
1.6. Согласование физических характеристик сигнала и канала связи	26
Лекция 2. Цифровые методы обработки и передачи звуковых сообщений	27
2.1. Особенности представления звуковых сообщений в цифровой форме.....	27
2.2. Классификация методов и показатели качества компрессии сигналов при передаче звука	30
2.3. Кодеры формы	31
2.4. Вокодеры	35
2.5. Гибридные кодеры	39
Лекция 3. Цифровые методы передачи видеоизображений	40
3.1. Основные характеристики цифрового видео	40
3.2. Свойства системы зрения человека	42
3.3. Межкадровая и внутрикадровая избыточность изображения	42
3.4. Алгоритмы сжатия цифровых сигналов при передаче видеоизображений.....	43
3.5. Примеры форматов цифрового видео	45

Лекция 4. Телекоммуникационные каналы	48
4.1. Свойства физических каналов	48
4.2. Системы многоканальной связи	65
Лекция 5. Сети ЭВМ	72
5.1. Назначение локальных вычислительных сетей	72
5.2. Классификация ЛВС	73
5.3. Базовая модель взаимодействия открытых систем OSI (Open System Interconnection).....	73
5.4. Топологии вычислительной сети.....	76
5.5. ЛВС стандарта Ethernet.....	79
5.6. ЛВС стандарта Token Ring	91
Лекция 6. Глобальная информационная сеть Интернет	101
6.1. Предпосылки возникновения Интернет.....	101
6.2. Адресация в сети Интернет	103
6.3. Стек протоколов TCP/IP	108
6.4. Маршрутизация сообщений	111
6.5. Развитие стека TCP/IP: протокол IPv.6	112
Лекция 7. Системы подвижной связи	116
7.1. Назначение и классификация систем подвижной связи	116
7.2. Ведомственные системы подвижной связи	116
7.3. Системы персонального радиовызова.....	118
7.4. Системы сотовой подвижной связи.....	119
Раздел II. Тестовые задания	126
Тест к лекции 1. Элементы теории передачи информации	126
Тест к лекции 2. Цифровые методы обработки и передачи звуковых сообщений.....	127
Тест к лекции 3. Цифровые методы передачи видеоизображений	128
Тест к лекции 4. Телекоммуникационные каналы.....	129
Тест к лекции 5. Сети ЭВМ.....	130
Тест к лекции 6. Глобальная информационная сеть Интернет.....	131
Тест к лекции 7. Системы подвижной связи	132

Раздел III. Лабораторные работы.....	133
1. Расчет конфигурации сети Ethernet	133
1.1. Критерии корректности конфигурации	133
1.2. Методика расчета времени двойного оборота и уменьшения межкадрового интервала	134
1.3. Пример расчета конфигурации сети	136
1.4. Задание на лабораторную работу	137
1.5. Справочные данные IEEE	143
1.6. Контрольные вопросы	145
2. Изучение структуры IP-адреса	145
2.1. Типы адресов стека TCP/IP.....	145
2.2. Классы IP-адресов.....	146
2.3. Особые IP-адреса	147
2.4. Использование масок в IP-адресации	149
2.5. Задание на лабораторную работу	150
2.6. Контрольные вопросы.....	151
3. Взаимодействие прикладных программ с помощью транспортного протокола TCP	151
3.1. Транспортный протокол TCP.....	151
3.2. Транспортный протокол UDP.....	153
3.3. Порты, мультиплексирование и демультиплексирование	154
3.4. Логические соединения.....	156
3.5. Программирование обмена данными на основе транспортных протоколов.....	157
3.6. Пример реализации простейшего клиент- серверного приложения на основе сокетов.....	160
3.7. Задание на лабораторную работу	161
3.8. Справочные данные.....	162
3.9. Контрольные вопросы.....	163
4. Взаимодействие прикладных программ с помощью протокола электронной почты SMTP	163
4.1. Модель протокола, команды и коды ответов SMTP	163
4.2. Кодировка сообщений.....	165

4.3. Процесс передачи сообщений	166
4.4. Пример последовательности команд почтовой транзакции	167
4.5. Задание на лабораторную работу	169
4.6. Справочные данные	170
4.7. Контрольные вопросы	172
5. Взаимодействие прикладных программ с помощью протокола электронной почты POP3	173
5.1. Модель протокола POP3, его назначение и стадии POP3-сессии	173
5.2. Формат сообщений	174
5.3. Процесс получения сообщений. Команды и ответы протокола POP3.....	175
5.4. Задание на лабораторную работу	176
5.5. Справочные данные.....	177
5.6. Контрольные вопросы.....	179
6. Взаимодействие прикладных программ с помощью протокола передачи данных FTP	179
6.1. Назначение и модели работы протокола FTP.....	179
6.2. Особенности управления процессом обмена данными	182
6.3. Команды и ответы протокола FTP.....	183
6.4. Задание на лабораторную работу	184
6.5. Справочные данные.....	184
6.6. Контрольные вопросы.....	187
Список используемой литературы	188

ВВЕДЕНИЕ

Цель дисциплины – изучение теоретических основ современных телекоммуникационных систем, изучение построения сетей, адресации, маршрутизации и их практического применения.

Задачи дисциплины:

- изучить понятия, математические основы, особенности современных телекоммуникационных систем;
- ознакомиться с цифровыми методами обработки и передачи видеоизображений и звуковых сообщений;
- изучить принципы построения сетей по стандарту Ethernet;
- изучить принципы адресации в сетях TCP/IP и приобрести практические навыки применения и назначения IP-адресов;
- ознакомиться с назначением, классификацией, топологией вычислительных сетей.

В результате изучения дисциплины студент должен:

знать:

- базовые понятия, математические основы и методы построения современных телекоммуникационных систем;
- методы расчета конфигурации сети;
- особенности структуры IP адресов;
- взаимодействие прикладных программ с помощью протоколов TCP/IP и SMTP;

уметь:

- применять для выполнения лабораторных работ полученные теоретические знания.

Для успешного усвоения дисциплины необходимы знания по высшей математике, основам радиоэлектроники и программированию.

СТРУКТУРНОЕ СОДЕРЖАНИЕ КУРСА

№ п/п	<i>Вид занятия</i>	<i>Тема занятия и изучаемые вопросы</i>	<i>Количество часов</i>
1.	Лекция 1	Элементы теории передачи информации	4
2.	Лекция 2	Цифровые методы обработки и передачи звуковых сообщений	4
3.	Лекция 3	Цифровые методы передачи видеоизображений	4
4.	Лекция 4	Телекоммуникационные каналы	4
5.	Лекция 5	Сети ЭВМ	4
6.	Лекция 6	Глобальная информационная сеть Интернет	4
7.	Лекция 7	Системы подвижной связи	4
8.	Лабораторная работа 1	Расчет конфигурации сети Ethernet	4
9.	Лабораторная работа 2	Изучение структуры IP-адреса	4
10.	Лабораторная работа 3	Взаимодействие прикладных программ с помощью транспортного протокола TCP	4
11.	Лабораторная работа 4	Взаимодействие прикладных программ с помощью протокола электронной почты SMTP	4
12.	Лабораторная работа 5	Взаимодействие прикладных программ с помощью протокола электронной почты POP3	4
13.	Лабораторная работа 6	Взаимодействие прикладных программ с помощью протокола передачи данных FTP	4
	Всего		52

Раздел I. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ

Лекция 1. Элементы теории передачи информации

1.1. Информация, сообщение, сигнал

Понятие «информация» имеет много различных аспектов, и в связи с этим существует и несколько различных подходов к ее определению и оценкам (количественным, качественным и др.). Например, в философии принято считать, что информация есть отражение реального мира.

Исходя из специфики задач теории информации, академик А.А. Харкевич определил информацию как совокупность сведений о каком-либо событии, явлении, предмете и т.д., являющихся объектом хранения, передачи и преобразования. Для выполнения указанных действий используют условные символы (знаки) – буквы, математические знаки, рисунки, жесты, слова и т.п., позволяющие выразить информацию в необходимой форме. Совокупность знаков, которые используются для хранения, передачи и преобразования информации, называют *сообщением*. Так, при телеграфной передаче информация представляется в виде букв и цифр. Соответственно сообщением является текст телеграммы, представляющий последовательность этих знаков. В телефонных системах сообщением является речь – определенный набор звуков, отображающих не только содержание, но и интонацию, тембр, ритм и иные свойства речи. В различных технических системах информация представляется в двоичной форме, т.е. только двумя условными символами, например, 1 и 0. Соответственно сообщением служит последовательность конечного числа двоичных символов. При передаче движущихся изображений в телевизионных системах сообщение представляет собой изменение во времени яркости элементов изображения.

Различают дискретные и непрерывные сообщения. *Дискретные сообщения* формируются в результате последовательной выдачи источником сообщений отдельных знаков. Множество различных знаков называют *алфавитом источника сообщений*, а их количество – *объемом алфавита*. *Непрерывные сообщения* не делимы на элементы. Они описываются функциями времени, принимающими непрерывное множество значений. Типичными примерами непрерывных сообщений могут служить речь, телевизионное изображение.

Передача сообщений (а следовательно, и информации) на расстояние осуществляется с помощью какого-либо материаль-

ного носителя (бумаги, магнитной ленты и т.п.) или физического процесса (звуковых или электромагнитных волн, тока и т.п.). Физический процесс, посредством которого сообщение передается на расстояние, называется *сигналом*.

В качестве сигнала можно использовать любой физический процесс, изменяющийся в соответствии с переносимым сообщением. В современных системах управления и связи чаще всего используют электрические сигналы. Физической величиной, определяющей такой сигнал, является ток или напряжение. Сигналы формируются путем изменения тех или иных параметров физического носителя по закону передаваемых сообщений. Процесс изменения параметров носителя принято называть *модуляцией*.

Сообщения могут быть функциями времени, например, речь при передаче телефонных разговоров, температура или давление воздуха при передаче телеметрических данных, спектакль при передаче по телевидению и т.п. В других случаях сообщение не является функцией времени (например, текст телеграммы, неподвижное изображение и т.д.). Сигнал является функцией времени, даже если сообщение таковым не является.

По своей природе сигналы могут быть электрическими, световыми, звуковыми и т.п. В системах телекоммуникаций (ТК) используются электрические и световые сигналы. Электрические сигналы получили наиболее широкое применение в системах ТК. Это связано с тем, что обработка электрических сигналов осуществляются гораздо проще с технической точки зрения, чем обработка сигналов другой физической природы. Поэтому при передаче сообщения неэлектрической природы предварительно преобразуются в электрические колебания с помощью преобразователей: микрофонов, передающих телевизионных трубок, датчиков температуры, давления и т.п. Эти электрические колебания обычно называют *первичными сигналами*. Наряду с электрическими сигналами в системах ТК могут использоваться световые сигналы, передача которых на расстояние по волоконно-оптическим линиям связи может оказаться предпочтительнее, чем передача электрических сигналов.

Различают следующие виды сигналов:

- непрерывные по уровню и по времени (рис. 1.1, а);
- непрерывные по уровню и дискретные по времени (рис. 1.1, б);
- дискретные (квантованные) по уровню и непрерывные по времени (рис. 1.1, в);
- дискретные по уровню и по времени (рис. 1.1, г).

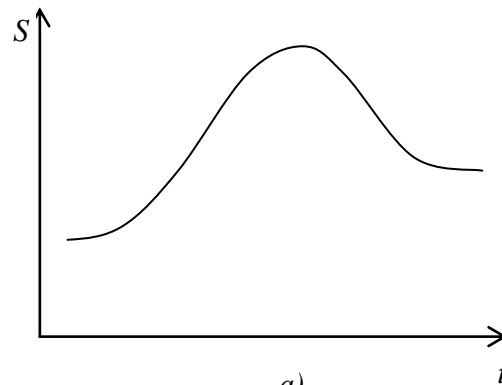
Сигналы первого вида (рис. 1.1, а), называемые непрерывными, задаются на конечном или бесконечном временном интервале и могут принимать любые значения в некотором диапазоне. Примером таких сигналов, называемых *аналоговыми*, являются сигналы на выходах микрофона, датчиков температуры, давления, положения и т.п.

Сигналы второго вида (рис. 1.1, б) задаются в определенные дискретные моменты времени и могут принимать любые значения из некоторого диапазона. Их можно получить из непрерывных сигналов путем взятия отсчетов в определенные моменты. Это преобразование называется *дискретизацией* во времени. *Шаг дискретизации* T_d (промежуток времени между двумя соседними отсчетами) может быть как постоянным, так и переменным. Обычно его значение выбирают, исходя из допустимой погрешности при восстановлении непрерывного сигнала по конечному числу его отсчетов. Устройство, осуществляющее формирование дискретных отсчетов сигналов, называется *дискретизатором*.

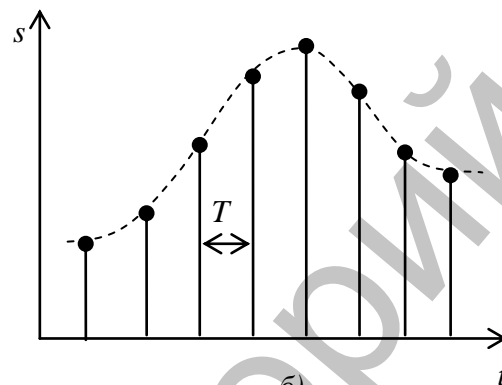
Сигналы третьего вида (рис. 1.1, в), называемые *квантованными по уровню*, задаются на некотором временном интервале и характеризуются тем, что принимают только вполне определенные дискретные значения. Их можно получить из непрерывных сигналов, применяя к ним операцию квантования по уровню. В результате этой операции непрерывный сигнал заменяется ступенчатой функцией. *Шаг квантования* Δ_a (расстояние между двумя соседними разрешенными уровнями) может быть как постоянным, так и переменным. Его обычно выбирают из условия обеспечения требуемой точности восстановления непрерывного сигнала из квантованного. Устройство, которое выполняет указанную операцию, носит название *квантователя*.

Сигналы четвертого вида (рис. 1.1, г), называемые *дискретными*, задаются в определенные дискретные моменты и принимают определенные дискретные значения. Их можно получить, например, из непрерывных сигналов, осуществляя операции дискретизации по времени и квантования по уровню. Такие сигналы легко представить в цифровой форме, т.е. в виде чисел с конечным числом разрядов. По этой причине их часто называют *цифровыми*. Аналогичная классификация возможна и для сообщений. Сообщения, подлежащие передаче, являются или случайной величиной, или случайной функцией. Детерминированные (заранее известные) сообщения не содержат информации, и нет смысла их передавать. Соответственно сигнал также следует рассматривать как случайный процесс. Детерминированные сигналы

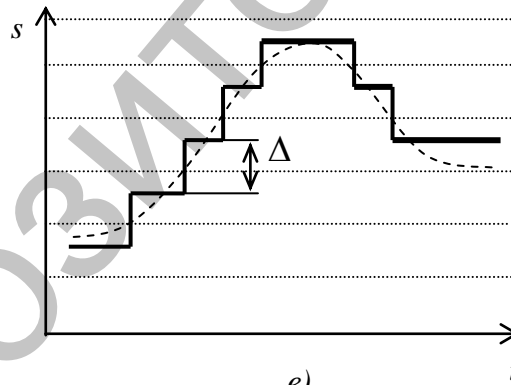
не несут информацию. В технике связи они используются для изучения свойств различных радиотехнических цепей.



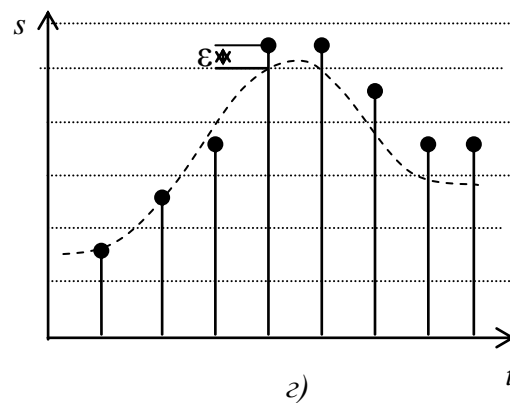
a)



б)



в)



г)

Рис. 1.1.

1.2. Преобразование непрерывных сообщений в цифровую форму

Для сообщения с ограниченным спектром дискретизация осуществляется на основе теоремы Котельникова, в соответствии с которой любую непрерывную функцию со спектром $0 \leq F \leq F_{\max}$ можно однозначно определить последовательностью ее мгновенных значений, взятых через интервалы:

$$T_{\text{Д}} \leq \frac{1}{2F_{\max}}. \quad (1.1)$$

Восстановление непрерывной функции производится в соответствии с выражением:

$$\tilde{U}(t) = \sum_{i=-\infty}^{\infty} U(iT_{\text{Д}}) \frac{\text{Sin}2\pi(t-iT_{\text{Д}})}{2\pi F_{\max}(t-iT_{\text{Д}})}, \quad (1.2)$$

называемым рядом Котельникова.

Условие ограничения спектра может выполняться не для всех сигналов. Но на практике вследствие ограниченности полосы пропускания канала, спектр сигнала можно считать ограниченным некоторой частотой F_{\max} . Обычно она определяется на основе частотного критерия. Спектр сигнала ограничивается областью частот от 0 до F_{\max} , в которой сосредоточена большая часть энергии сигнала (80–95%). Такое ограничение, естественно, приводит к некоторому искажению сигнала.

Восстановление непрерывного сигнала по дискретным отсчетам также сопровождается погрешностью. Основные причины:

1. Вместо δ -импульсов при формировании отсчетов используются импульсы конечной длительности.

2. Ограничение спектра сигнала (если оно есть).

3. Пренебрежение вкладом бесконечного числа функций отсчетов, соответствующих выборкам за пределами интервала T , так как обычно восстановление непрерывного сигнала проводится по конечному числу членов ряда Котельникова.

4. Для восстановления непрерывной функции последовательность дискретных значений необходимо подать на вход так называемого «идеального» фильтра нижних частот (ФНЧ). Поскольку «идеальный» ФНЧ является абстракцией и на практике используется реальный ФНЧ с характеристиками, отличающимися от характеристик «идеального» ФНЧ, то это также является причиной искажений при восстановлении непрерывного сигнала по дискретным отсчетам.

Таким образом, следует иметь в виду, что теорема Котельникова выражает предельные соотношения для идеализированных условий (ограниченность спектра и бесконечное время наблюдения). Однако в большинстве практических случаев ее применение для реальных сигналов позволяет получить сравнительно небольшую погрешность, приемлемую в инженерной практике.

1.3. Обобщенная структурная схема системы связи. Канал связи

Системой связи (телекоммуникационной системой) называют совокупность технических средств, предназначенных для передачи информации, включая источник информации и получателя информации. Структурная схема простейшей системы связи показана на рис. 1.2. Рассмотрим назначение основных ее элементов.

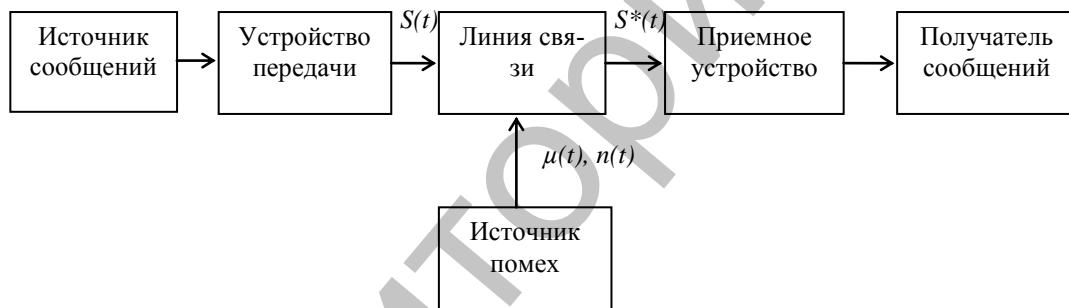


Рис. 1.2.

Источником сообщений может быть человек или различного рода устройства (датчики, ЭВМ и т.п.). Источник сообщений осуществляет выбор сообщений из ансамбля сообщений. Если сообщение на выходе источника имеет неэлектрическую природу, то для его передачи в системе связи оно преобразуется в первичный электрический сигнал. Указанная операция производится в первичном преобразователе. В телефонии, например, в качестве первичного преобразователя применяется микрофон, превращающий перепады звукового давления в пропорционально изменяющийся электрический ток. В телевидении функции первичного преобразователя выполняет телевизионная камера.

Первичные сигналы обычно являются низкочастотными и не предназначены для передачи на расстояние. Для передачи на большие расстояния используют специальные электромагнитные колебания высокой частоты, называемые переносчиками, которые могут эффективно распространяться по *линии связи*. В пере-

дающем устройстве первичный сигнал превращается во вторичный (высокочастотный) сигнал $S(t)$. В качестве переносчика могут использоваться электромагнитные колебания, имеющие гармоническую $u(t) = A_0 \cdot \sin(\omega_0 t - \varphi_0)$ или импульсную форму.

Сами переносчики не содержат информации о передаваемом сообщении. Для того, чтобы заложить в них эту информацию, применяют операцию модуляции, которая заключается в изменении одного или нескольких параметров переносчика по закону передаваемого сообщения. Например, в гармоническом переносчике можно изменять амплитуду, частоту или фазу колебания. При этом возможны три вида модуляции: амплитудная (АМ), фазовая (ФМ) и частотная (ЧМ), когда модулированные параметры могут быть представлены в следующем виде

$$\begin{aligned} A(t) &= A_0 + \Delta A \cdot a(t), \\ \omega(t) &= \omega_0 + \Delta \omega \cdot a(t), \\ \varphi(t) &= \varphi_0 + \Delta \varphi \cdot a(t), \end{aligned}$$

где $a(t)$ – закон изменения передаваемого сообщения (полагаем, что $-1 \leq a(t) \leq 1$),

ΔA , $\Delta \omega$, $\Delta \varphi$ – максимальные изменения соответственно амплитуды, частоты и фазы.

В импульсном переносчике можно изменять амплитуду, временное положение импульсов относительно выбранного начала отсчета, их длительность, период следования, параметры формы импульса и т.п.

Устройство, осуществляющее изменение одного или нескольких параметров переносчика, называется модулятором.

Линия связи – это среда, используемая для передачи сигналов. Линии связи могут быть проводные и беспроводные, например, радиолинии. В радиолиниях средой служит часть пространства, в котором распространяются электромагнитные волны от передатчика к приемнику.

Источник помех. В реальной системе сигнал передается при наличии помех, под которыми понимают любые случайные воздействия, накладывающиеся на сигнал и затрудняющие его прием. Поэтому сигнал $S^*(t)$, поступающий на вход приемного устройства (рис. 1.2), в общем случае отличается от сигнала $S(t)$, который был на выходе радиопередающего устройства. В некоторых случаях действие помех $n(t)$ можно описать соотношением

$$S^*(t) = S(t) + n(t),$$

где $n(t)$ не зависит от $S(t)$.

Помеха, удовлетворяющая такому условию, называется аддитивной. Если соотношение, связывающее сигналы на выходе радиопередатчика и приемника, имеет вид

$$S^*(t) = \mu(t) \cdot S(t),$$

где $\mu(t)$ – некоторая случайная функция, то помеха называется мультипликативной.

В реальных линиях связи действуют как аддитивная, так и мультипликативная помехи, поэтому

$$S^*(t) = \mu(t) \cdot S(t) + n(t).$$

В зависимости от характера изменения во времени различают флуктуационные, импульсные (сосредоточенные во времени) и узкополосные (сосредоточенные по частоте) помехи. Флуктуационная помеха порождается различного рода флуктуациями, т.е. случайного рода отклонениями физических величин от их средних значений. Флуктуационная помеха может быть обусловлена дискретной природой носителей заряда в электронных приборах, тепловым движением носителей заряда и некоторыми другими причинами.

Импульсная помеха представляет собой случайную последовательность импульсов, следующих столь редко, что реакция приемника на текущий импульс успевает затухнуть к моменту появления очередного импульса. Типичным примером такой помехи могут служить атмосферная помеха, различного рода индустриальные помехи и т.д.

Узкополосная помеха, как следует из ее названия, сосредоточена в сравнительно узкой полосе частот, существенно меньшей по сравнению с полосой частот сигнала.

Приемное устройство обрабатывает принятое колебание $S^*(t) = \mu(t) \cdot S(t) + n(t)$ и восстанавливает по нему сообщение (первичный сигнал) $b^*(t)$, которое с некоторой погрешностью отражает переданное сообщение $b(t)$. Другими словами, приемник должен на основе анализа колебания $S^*(t)$ определить, какое из возможных сообщений передавалось.

Получателем сообщения может быть человек, для которого оно предназначено, или различного рода устройства (автомат, ЭВМ, магнитофон и т.п.).

Совокупность технических средств передачи информации, включающая среду распространения и обеспечивающая передачу сигнала от некоторой точки А системы до точки В (рис. 1.3), называется *каналом*. Точки А и В могут быть выбраны произвольно в зависимости от решаемых задач. Так, в одних случаях канал

может состоять только из линии связи, в других – из модулятора, линии связи и демодулятора и т.д. Часть системы связи, расположенная до точки А, является источником сигнала для этого канала. Если сигнал, поступающий на вход канала и снимаемый на его выходе, является дискретным по состояниям, то такой канал называется дискретным. Если входные и выходные сигналы канала непрерывные, то канал называется непрерывным. Встречаются также дискретно-непрерывные и непрерывно-дискретные каналы, на вход которых поступают дискретные сигналы, а на выходе снимаются непрерывные, и наоборот.

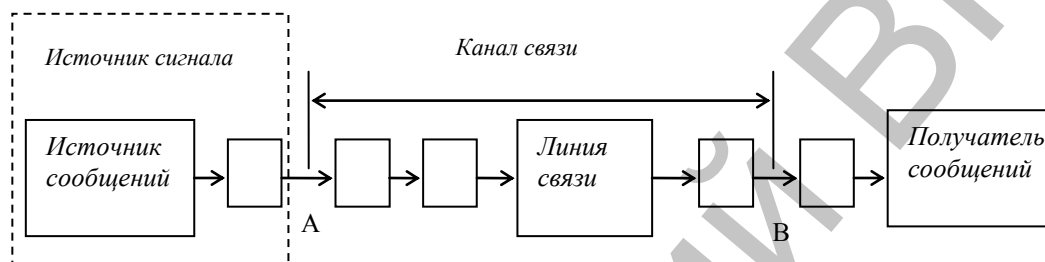


Рис. 1.3.

Любая телекоммуникационная система характеризуется рядом показателей. Рассмотрим наиболее существенные из них с точки зрения передачи информации.

Достоверность передачи информации характеризует степень соответствия принятых сообщений переданным. Она зависит от параметров самой системы, степени ее технического совершенства и условий работы. Последние определяются типом и состоянием линий связи, видом и интенсивностью помех и т.д. Для различных телекоммуникационных систем критерии соответствия принятого сигнала переданному могут существенно отличаться. При передаче дискретных сообщений действие помех проявляется в том, что вместо переданного символа принимается другой. В этом случае достоверность передачи сообщений целесообразно характеризовать или вероятностью правильного приема символа $p_{ПР}$, или вероятностью ошибки $p_{ОШ} = 1 - p_{ПР}$.

При передаче непрерывных сообщений отличие принятого сообщения $b^*(t)$ от переданного $b(t)$ носит также непрерывный характер.

Для оценки достоверности передачи сообщений в данном случае обычно используют средний квадрат ошибки

$$\overline{\varepsilon^2} = \overline{[b^*(t) - b(t)]^2}$$

или относительный средний квадрат ошибки

$$\delta^2 = \overline{\varepsilon^2} / P_b = P_\varepsilon / P_b,$$

где усреднение производится по всем реализациям сообщений $b(t)$ и их оценкам $b^*(t)$, $P_b = \frac{1}{T_c} \int_0^{T_c} b^2(t) dt$ – средняя мощность сообщения $b(t)$, T_c – его длительность, P_ε – мощность помехи на выходе приемника.

Возможны и другие показатели достоверности, как, например, показатель максимальной абсолютной ошибки $\varepsilon_{\max} = \max |\varepsilon(t)|$.

Под *помехоустойчивостью* понимают способность системы противостоять вредному действию помех на передачу сообщений. Количественно помехоустойчивость телекоммуникационных систем можно характеризовать вероятностью ошибки $p_{OШ}$ при заданном отношении средних мощностей сигнала и помехи в полосе частот занимаемой сигналом, или требуемым отношением средних мощностей сигнала и помехи на входе приемника системы, при котором обеспечивается заданная вероятность ошибки $p_{OШ}$. Еще одна важная характеристика – скорость передачи информации – будет введена ниже.

1.4. Информационные характеристики источника дискретных сообщений и канала связи

Количественная оценка информации. Для сравнения различных систем телекоммуникаций необходимо ввести количественную меру, позволяющую оценивать объем информации, содержащейся в сообщении.

Строгие методы количественного определения информации были предложены К. Шенноном в 1948 г. и привели к построению теории информации, являющейся основой теории связи, информатики и ряда смежных отраслей науки и техники.

$$U = \left(\begin{array}{l} u_1; u_2; \dots; u_N; \\ P(u_1); P(u_2); \dots; P(u_N) \end{array} \right).$$

Рассмотрим основные идеи этой теории применительно к дискретному источнику сообщений, который в каждый момент времени случайным образом может принимать одно из конечного множества возможных состояний. Каждому состоянию источника сообщений ставится в соответствие условное обозначение в виде знака (в частности буквы) из алфавита данного источника $u_1, u_2, u_3, \dots, u_N$. Одни состояния выбираются источником сообще-

ний чаще, другие реже. Поэтому наряду с множеством состояний целесообразно задать вероятность их появления:

$$U = \left(\begin{array}{c} u_1; u_2; \dots; u_N \\ p_1; p_2; \dots; p_N \end{array} \right),$$

$$\text{причем } \sum_{i=1}^N p_i = 1.$$

Совокупность состояний и вероятностей их получения называется ансамблем U .

Перед тем как ввести определение количества информации, сформулируем условия, которым должна удовлетворять эта величина:

1) она должна быть аддитивной величиной, т.е. если рассматривать два последовательных события u_i и u_k , происходящих независимо друг от друга, как одно укрупненное, то количество информации в таком событии должно равняться сумме количества информации в каждом из них

$$I(u_i, u_k) = I(u_i) + I(u_k); \quad (1.3)$$

2) количество информации в сообщении о достоверном событии ($p=1$) равно нулю (такое сообщение ничего не добавляет к нашим знаниям);

3) данная величина должна быть неотрицательной;

4) количество информации не должно зависеть от качественного содержания сообщения, в частности, от степени его важности для получателя, эмоциональной окраски и т.д.

Итак, для определения количества информации в некотором сообщении u_i из ансамбля U необходимо основываться только на таком параметре, который характеризует в самом общем виде это сообщение. Таким параметром, очевидно, является вероятность p_i появления данного сообщения на выходе источника.

Дальнейшее уточнение искомого определения не составит труда, если принять во внимание первые два из перечисленных выше условий. Пусть u_i и u_k – два независимых события. Вероятность того, что оба этих сообщения появятся на выходе источника одно за другим,

$$P(u_i, u_k) = P(u_i) \cdot P(u_k), \quad (1.4)$$

а количество информации в этих сообщениях должно удовлетворять условию (1.3). Следовательно, необходимо найти функцию, обладающую свойством, что при перемножении двух аргументов значения функции складываются. Единственная такая функция – это логарифмическая функция $I(u) = k \log P(u)$, где k – постоян-

ный коэффициент. Заметим, что при таком определении количества информации выполняется и второе требование – при $P(u)=1$, $I(u)=0$. Основание логарифма не имеет принципиального значения и определяет только масштаб функции. Так как информационная техника широко использует элементы, имеющие два устойчивых состояния, то обычно основание логарифма выбирают равным 2. В дальнейшем обозначение \log , если основание не оговаривается особо, будет означать двоичный логарифм. Чтобы количество информации $I(u)$ было неотрицательной величиной, выбирают $k = -1$. Поэтому $I(u) = -\log P(u)$.

Если источник передает последовательность зависимых между собой сообщений, то получение предшествующих сообщений может изменить вероятность последующего, а следовательно, и количество информации в нем. Оно должно определяться по условной вероятности передачи данного сообщения u_k при известных предшествовавших сообщениях u_{k-1}, u_{k-2}, \dots :

$$I(u_k | u_{k-1}, u_{k-2}, \dots) = -\log P(u_k | u_{k-1}, u_{k-2}, \dots). \quad (1.5)$$

Введенное выше определение характеризует количество информации, содержащееся в одном сообщении из ансамбля U . При этом $I(u)$ является случайной величиной, зависящей от того, какое состояние источника в действительности реализуется. Для характеристики всего ансамбля (или источника) используется математическое ожидание количества информации, называемое энтропией и показывающее, какое количество информации в среднем содержится в одном сообщении данного источника

$$H(U) = M \{-\log P(u)\}. \quad (1.6)$$

Для источника независимых сообщений выражение (1.6) можно представить в виде

$$H(U) = -\sum P(u_i) \log P(u_i), \quad i = 1, \dots, N. \quad (1.7)$$

Пример 1.

Рассмотрим случай, когда алфавит состоит из двух знаков, появляющихся на выходе источника сообщений независимо друг от друга. Обозначим $P(u_1) = P$. Соответственно $P(u_2) = 1 - P$. Тогда на основании (1.7) имеем

$$H(U) = -P \log P - (1 - P) \log(1 - P).$$

Если события u_1 и u_2 являются равновероятными, то $P = 1/2$. Подставив это значение в данное уравнение, получим $H(U) = 1$. В общем виде зависимость $H(U)$ от P показана на рис. 1.4.

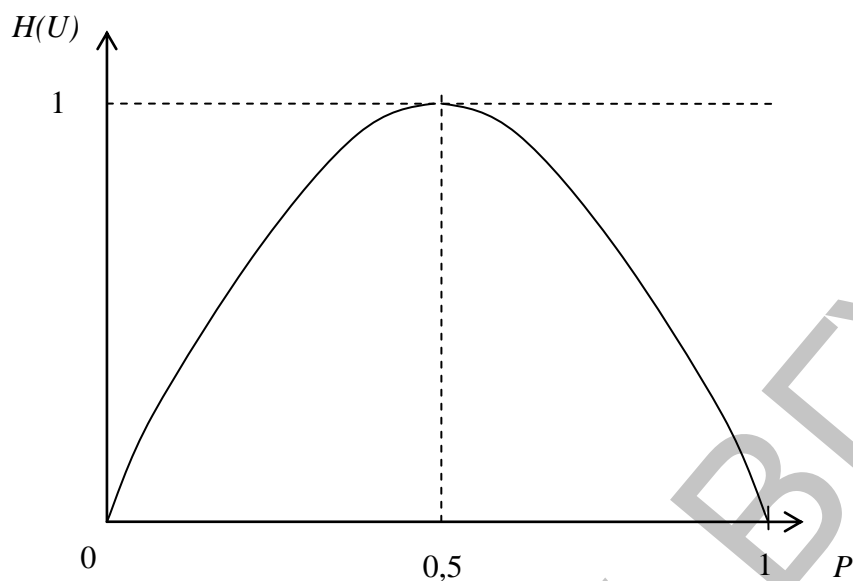


Рис. 1.4.

Количество информации, содержащееся в среднем в одном символе алфавита, состоящего из двух знаков, которые появляются на выходе источника сообщений независимо друг от друга и равновероятно, получило название «1 бит информации» или просто 1 бит.

Случай, когда алфавит источника сообщений состоит всего из двух знаков, широко распространен на практике. В качестве примера можно привести цифровые системы, использующие алфавит, состоящий из знаков 0 и 1.

Чем больше энтропия источника, тем больше степень неожиданности передаваемых им сообщений в среднем, т.е. тем более неопределенным является ожидаемое сообщение. Поэтому энтропию часто называют мерой неопределенности сообщений. При этом имеется в виду неопределенность, существующая до того, как сообщение передано. После приема сообщения, если оно передано верно, всякая неопределенность устраняется. Это позволяет трактовать количество информации как меру уменьшения неопределенности.

Свойства энтропии источника дискретных сообщений

Перечислим основные свойства энтропии:

энтропия неотрицательна;

энтропия равна нулю для вырожденного ансамбля, когда одно сообщение передается с вероятностью 1, а остальные имеют нулевую вероятность;

энтропия аддитивна (смысл этого свойства рассмотрен выше);

если ансамбль содержит N различных сообщений, то

$$H(U) \leq \log N, \quad (1.8)$$

причем равенство имеет место только тогда, когда все сообщения передаются равновероятно и независимо ($P(u_i) = 1/N; i = 1, \dots, N$).

Для доказательства последнего свойства воспользуемся неравенством

$$\ln w \leq (w - 1). \quad (1.9)$$

Рассмотрим разность функций, стоящих в левой и правой частях неравенства (1.8). Путем несложных преобразований ее можно представить в виде:

$$\begin{aligned} H(U) - \log N &= \sum P_i \log(1/P_i) - \sum P_i \log N = \\ &= \sum P_i \log(1/NP_i), \end{aligned} \quad (1.10)$$

причем суммирование ведется по $i = 1, \dots, N$.

Чтобы воспользоваться неравенством (1.9), перейдем в последнем выражении к натуральному основанию логарифма:

$$H(U) - \log N = \sum P_i \log(1/NP_i) = \log e \sum P_i \ln(1/NP_i). \quad (1.11)$$

Используя (1.9), преобразуем выражение (1.11). Получим:

$$\begin{aligned} H(U) - \log N &= \log e \sum P_i \ln(1/NP_i) \leq \log e \sum P_i [1/NP_i - 1] = \\ &= \log e (\sum 1/N - \sum P_i) = 0. \end{aligned} \quad (1.12)$$

Знак равенства в (1.8) имеет место тогда, когда в неравенстве (1.9) $w = 1$. Так как в нашем случае $w = 1/NP_i$, то условие $w = 1$ эквивалентно $P_i = 1/N; i = 1, \dots, N$. То есть, $H(U) = \log N$ в случае, когда все элементы алфавита появляются на выходе источника равновероятно. Свойство (1.8) доказано.

Анализируя свойство (1.8), можно прийти к выводу, что чем больше объем алфавита источника дискретных сообщений, тем большее количество информации содержится в среднем в одном символе этого алфавита (при условии, что все элементы алфавита появляются на выходе источника равновероятно). Казалось бы, рассмотренный выше алфавит $\{0,1\}$, который широко применяется на практике, следует безоговорочно заменить алфавитами с $N > 2$. Однако следует иметь в виду, что в реальных условиях различение сигналов, которые могут принимать количество значений $N > 2$, осуществить на фоне помех гораздо сложнее, чем в случае, когда сигналы могут принимать только два значения – 0 и 1.

Пример 2.

В теории информации доказывается, что энтропия источника зависимых сообщений всегда меньше энтропии источника независимых сообщений при том же объеме алфавита и тех же безусловных вероятностях сообщений.

Пусть, например, источник выдает последовательность букв из алфавита объемом $N=32$. Если буквы выбираются равновероятно и независимо друг от друга, то энтропия источника

$$H(U) = \log 32 = 5 \text{ [бит]}.$$

Однако смысловое содержание такой последовательности букв вряд ли удовлетворит получателя сообщения. Если буквы передаются не хаотически, а составляют связный текст, например, на русском языке, то они оказываются неравновероятными и, главное, зависимыми (так, после гласных не может появиться «ь»; мала вероятность появления 3 гласных или согласных подряд и т.д.). В качестве примера приведем относительные частоты использования некоторых букв русского алфавита в текстах (в порядке убывания):

$$\begin{aligned} \text{«о»} &- 0,090; \\ \text{«е»}, \text{«ё»} &- 0,072; \\ \text{«а»}, \text{«и»} &- 0,062; \\ \text{«н»}, \text{«т»} &- 0,053; \\ &\dots\dots\dots \\ \text{«ц»} &- 0,004; \\ \text{«щ»}, \text{«э»} &- 0,003; \\ \text{«ф»} &- 0,002. \end{aligned}$$

Как видно из этих данных, различие в частоте появления букв в текстах достигает 45 раз!

Если рассматривать ансамбль текстов русской художественной прозы, то энтропия оказывается меньше 1,5 бит на букву. Еще меньше, около 1 бит на букву, энтропия ансамбля поэтических произведений, так как в них имеются дополнительные вероятностные связи, обусловленные ритмом и рифмами.

Избыточность сообщений. Рассмотрим ансамбль U , состоящий из N различных символов: u_1, u_2, \dots, u_N . Энтропия такого дискретного источника достигает максимального значения $H_{\max}(U) = \log N$, если символы статистически независимы и равновероятны. На практике может оказаться так, что символы, образующие алфавит, нельзя рассматривать как независимые и равновероятные, поэтому для такого источника $H(U) < H_{\max}(U)$. Предположим, на выходе такого источника появилось сообщение, состоящее из n символов. Количество информации, содержащееся в нем

$$I = nH(U). \quad (1.13)$$

При использовании алфавита с максимальной энтропией для передачи такого же количества информации потребовалось бы меньшее число символов

$$I = n_{\min} H_{\max}(U). \quad (1.14)$$

Приравнивая (1.13) и (1.14), находим

$$n_{\min} = n \cdot \frac{H(U)}{H_{\max}(U)} = \mu \cdot n, \quad (1.15)$$

где $\mu = H(U)/H_{\max}(U) < 1$ – коэффициент, характеризующий допустимую степень сжатия сообщений.

Величина $x = 1 - \mu$ называется избыточностью источника. Последствия от наличия избыточности неоднозначны. С одной стороны, избыточные сообщения требуют дополнительных затрат на передачу (например, увеличивается длительность передачи). С другой стороны, наличие избыточности способствует повышению помехоустойчивости сообщений, подчиняющихся априорно известным условиям (ограничениям), т.к. можно обнаружить и исправить ошибки, приводящие к нарушению этих ограничений.

Для сокращения избыточности на практике применяется *кодирование* источника дискретных сообщений, заключающееся в преобразовании исходного дискретного сообщения по определенному правилу в последовательность кодовых символов, удовлетворяющую требованиям равновероятности и статистической независимости.

Устройство, осуществляющее указанную операцию, называется *кодером* источника. В случае его использования, оно размещается в структурной схеме, приведенной на рис. 1.2, между источником сообщений и передающим устройством. Соответственно, на приемной стороне на выходе приемного устройства необходимо добавить устройство, которое будет осуществлять обратную операцию перед тем, как сообщение поступит получателю. Такое устройство называется *декодером*.

1.5. Скорость передачи информации по дискретному каналу. Пропускная способность

Наряду с введенными в разделе 1.3 такими характеристиками, как достоверность и помехоустойчивость, при оценке эффективности систем связи используется ряд других характеристик, рассмотренных ниже.

Технической скоростью V_T называется число элементарных сигналов (символов), передаваемых по каналу в единицу времени. Она зависит от свойств линии связи и быстродействия аппаратуры канала. С учетом возможных различий в длительностях символов

$$V_T = 1/\tau_{cp},$$

τ_{cp} – средняя длительность символа.

Единицей измерения технической скорости служит Бод. 1 Бод – скорость, при которой за одну секунду передается один символ.

Скорость передачи сигналов по дискретному каналу устанавливается с учетом ширины F_K полосы пропускания непрерывного канала. Ширина спектра сигнала $S^*(t)$ на выходе непрерывного канала не может превышать F_K . Учитывая, что длительность сигнала и ширина его спектра связаны обратной пропорциональной зависимостью, получим $(V_T)_{\max} \sim F_K$.

Информационная скорость, или скорость передачи информации, определяется средним количеством информации относительно заданного сообщения, которое передается по каналу в единицу времени. Она зависит как от характеристик данного канала связи, таких, как объем алфавита используемых символов, техническая скорость их передачи, статистические свойства помех в линии, так и от вероятностей поступающих на вход символов и их статистической взаимосвязи.

При известной скорости V_T скорость передачи информации $W(V_T, U)$ относительно некоторого сообщения по каналу задается соотношением

$$W(V_T, U) = V_T \cdot H(U), \quad (1.16)$$

где $H(U)$ – среднее количество информации, переносимое одним символом.

Для практических применений телекоммуникационных систем важно выяснить, до какого предела и каким путем можно повысить скорость передачи информации по конкретному каналу. Предельные возможности канала по передаче сообщений характеризуются его пропускной способностью.

Пропускная способность канала C_D равна максимальной скорости передачи информации по данному каналу, которой можно достигнуть при самых совершенных способах передачи и приема:

$$C_D = \max W(V_T, U) = \max \{V_T \cdot H(U)\}. \quad (1.17)$$

Пропускная способность канала, как и скорость передачи информации, измеряется числом двоичных единиц информации в секунду – бит/сек.

Пропускная способность канала является характеристикой его самого и не зависит от статистики сигнала.

Если передаваемые символы независимы, то для дискретного канала без помех соотношение (1.17) с учетом (1.8) может быть представлено в виде

$$C_D = V_T \cdot \log . \quad (1.18)$$

Каналы без помех можно рассматривать как некую идеализацию реальных каналов. Если мешающим действием помехи пренебречь нельзя, то можно показать, что для дискретных каналов с помехой пропускная способность будет тем меньше, чем больше энтропия источника помехи.

1.6. *Согласование физических характеристик сигнала и канала связи*

Канал можно охарактеризовать тремя основными параметрами:

- 1) Временем T_K , в течение которого он представлен для передачи сигнала.
- 2) Шириной полосы пропускания F_K .
- 3) Максимально допустимым превышением сигнала над помехой D_K .

Произведение указанных параметров называется объемом канала (рис. 1.5):

$$V_K = T_K F_K D_K.$$

Аналогично объему канала вводится понятие объем сигнала:

$$V_C = T_C F_C D_C,$$

где T_C – длительность сигнала, F_C – ширина его спектра, D_C – превышение сигнала над помехой.

V_C и V_K можно представить в трехмерном пространстве.

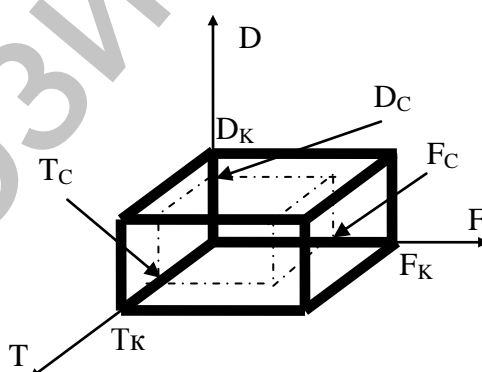


Рис. 1.5.

Необходимым условием принципиальной возможности неискаженной передачи сигнала по данному каналу является условие:

$$V_C \leq V_K. \quad (1.20)$$

Для того, чтобы выполнялись достаточные условия передачи $T_C \leq T_K$; $F_C \leq F_K$; $D_C \leq D_K$, могут потребоваться преобразо-

вания. Если канал имеет полосу пропускания меньше, чем ширина спектра подлежащего передаче сигнала, ширину спектра можно уменьшить в требуемое количество раз за счет такого же увеличения длительности сигнала. Объем сигнала при этом останется неизменным. Практически такое преобразование можно осуществить, например, путем записи сигнала на магнитную ленту с высокой скоростью и последующего воспроизведения со скоростью, при которой ширина его спектра равна полосе пропускания сигнала. При низком допустимом уровне превышения сигнала в канале D_K преобразование заключается в уменьшении уровня превышения D_C с одновременным увеличением его длительности. Возможны и другие преобразования.

С целью наилучшего соответствия характеристик сигнала и канала связи в радиопередающем устройстве (рис. 1.2) обычно применяется *канальное кодирование*.

Устройство, осуществляющее указанную операцию, называется *кодером* канала. Соответственно на приемной стороне необходимо добавить *декодер*, который будет осуществлять обратную операцию.

Лекция 2. Цифровые методы обработки и передачи звуковых сообщений

При реализации современных систем телекоммуникаций предпочтение отдают цифровым методам обработки и передачи сигналов. Цифровые системы по сравнению с аналоговыми имеют ряд существенных преимуществ при обработке, запоминании и передаче сигналов. Представление сообщений в цифровой форме обеспечивает высокую помехоустойчивость, возможность более полного использования пропускной способности каналов, стабильность параметров передачи и гибкость при построении телекоммуникационных сетей. Особое значение приобретает применение цифровых методов при передаче звуковых сообщений. В данном разделе рассмотрены методы кодирования источника звуковых сообщений. Устройство, осуществляющее преобразование первичного речевого сигнала в цифровую форму, называют *речевым кодером*.

2.1. Особенности представления звуковых сообщений в цифровой форме

Человеческий голос порождает первичный аналоговый сигнал, который занимает полосу частот примерно от 50 до 10000 Гц. Пред-

ставление этого сигнала в цифровой форме осуществляется путем дискретизации во времени и квантования по уровням (рис. 2.1) и сопровождается неустранимой ошибкой, называемой шумом квантования. Шум квантования – один из факторов, определяющих верность передачи непрерывных сообщений по дискретному каналу (вторым фактором являются помехи в канале передачи, накладывающиеся на полезный сигнал и приводящие к ошибочному приему).

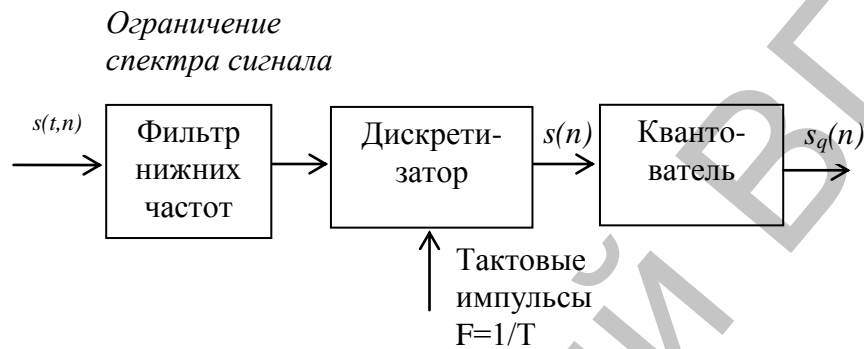


Рис. 2.1.

При квантовании возникает ошибка квантования $\varepsilon(n)$, равная разнице между квантованным $s_q(n)$ и истинными значениями сигнала:

$$\varepsilon(n) = s_q(n) - s(n), \quad (2.1)$$

где n – порядковый номер отсчета.

Искажения, вносимые квантователем, оценивают значением среднеквадратичной ошибки (СКО):

$$D = \sqrt{\frac{1}{M} \sum_{n=1}^M [s_q(n) - s(n)]^2}. \quad (2.2)$$

Если значение квантуемого сигнала не выходит за пределы рабочего диапазона квантователя, то ошибка

$$|\varepsilon(n)| \leq \frac{\Delta}{2}, \quad (2.3)$$

где Δ – шаг квантования.

Как следует из выражения (2.3) для снижения ошибки (шума) квантования необходимо снижать шаг квантования и соответственно увеличивать число уровней квантования. Если полный размах непрерывного сигнала равен $2S_{\max}$, то число уровней квантования

$$L_{KB} = 2S_{\max} / \Delta + 1. \quad (2.4)$$

Наиболее очевидный подход заключается в использовании квантователя с постоянным шагом квантования Δ . В этом случае для высококачественной передачи звуковых сообщений с малой

ошибкой квантования, как показывает практика, требуется, чтобы $L_{KB} \geq 4000$. При цифровом кодировании такого сигнала с помощью двоичных символов на каждый дискретный отсчет потребуется не менее $n = 12$ разрядов, поскольку $L_{KB} = 2^n$.

Оценим скорость цифрового потока в телекоммуникационном канале при передаче звуковых сообщений.

Для передачи речи в аналоговой телефонии в 60-х годах XX столетия была выбрана полоса частот 0,3–3,4 кГц. Решающими в выборе такой полосы были экономические соображения и нехватка телефонных каналов. Несмотря на определенное ухудшение восприятия ряда звуков (например, шипящих, существенная часть энергии которых сосредоточена в верхней части речевого спектра), такое ограничение незначительно повлияло на разборчивость речи.

При представлении речевых сигналов в цифровой форме верхнюю частоту в спектре дискретизируемого сигнала выбирают равной 4 кГц. Согласно теореме Котельникова при $F_{\max} = 4$ кГц, период дискретизации составляет $T_D = 1/(2F_{\max}) = 125$ мкс. При этом частота дискретизации $F_D = 1/T_D = 8000$ Гц. Скорость цифрового потока соответственно равна

$$W = F_D n. \quad (2.5)$$

При передаче речевых сообщений, использующей 12-разрядное кодирование отсчетов, скорость цифрового потока, поступающего на вход телекоммуникационного канала, составит 96 кбит/сек. Еще более высокие требования будут предъявляться к пропускной способности канала при передаче высококачественных звуковых сообщений, например, музыки. Известно, что для высококачественного воспроизведения музыки на компакт-дисках частота дискретизации составляет $F_D = 44,1$ кГц при 16-разрядном кодировании отсчетов. Подставляя эти значения в (2.5), определим скорость цифрового потока, которая составит:

$$W = 44100 \text{ отсчетов/сек} \cdot 16 \text{ бит/отсчет} = 705,6 \text{ кбит/сек.}$$

При использовании 2-х стереофонических каналов скорость цифрового потока превысит 1400 кбит/сек.

Необходимость эффективного использования телекоммуникационных каналов явилась причиной разработки специальных технических решений, позволяющих уменьшить скорость цифрового потока при передаче речевых сообщений. Процедуру преобразования речевых сигналов, при которой уменьшается скорость цифрового потока, назвали *компрессией* (сжатием). Практический эффект такого уменьшения скорости очевиден – появляется воз-

возможность обслужить большее количество абонентов на телекоммуникационном канале с заданной пропускной способностью или осуществлять передачу речевых сообщений по низкоскоростным каналам, по которым передача сигналов в некомпьюрированном виде была бы невозможна.

2.2. Классификация методов и показатели качества компрессии сигналов при передаче звука

Речевые кодеры можно разделить на 3 основные группы: кодеры формы, вокодеры и гибридные кодеры.

В кодерах формы обработке подвергается каждый отсчет дискретизированной последовательности. Кодеры данного типа обеспечивают сохранение и передачу формы исходного аналогового сигнала. При этом, как правило, достигается достаточно высокое качество восстановленного сигнала, поскольку основным источником искажений формы выходного сигнала является квантование. Однако скорость цифрового потока на выходе кодера формы остается все-таки достаточно высокой. Так при передаче речевых сообщений кодеры формы формируют цифровой поток со скоростью от 24 до 64 кбит/с.

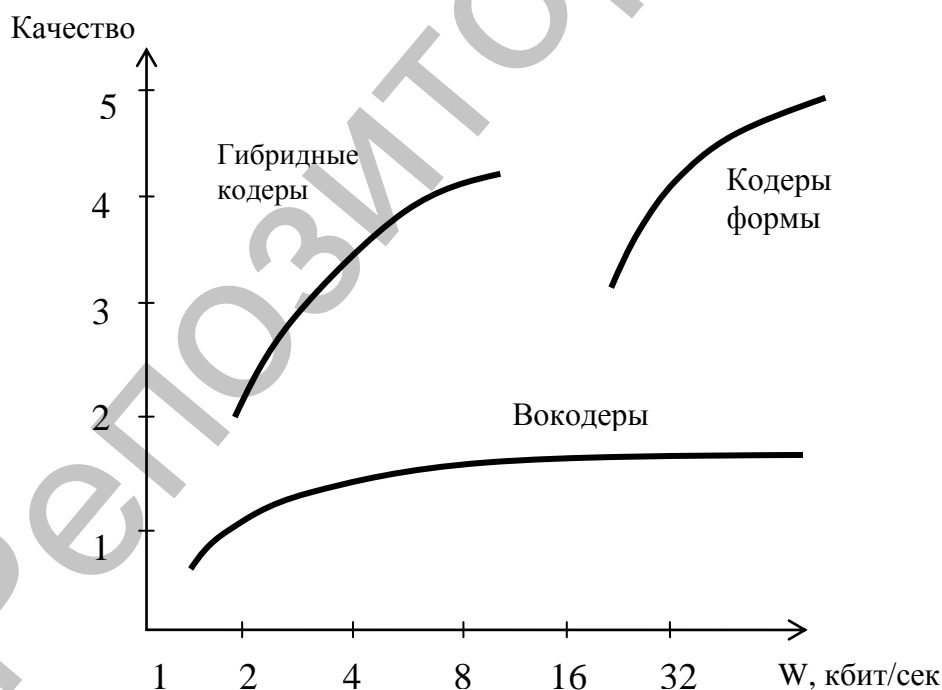


Рис. 2.2.

Работа вокодеров (от английских слов VOice – голос и CODER – кодировщик) основана на моделировании речевых сигналов с учетом их характерных особенностей. Это позволяет снизить

скорость передачи до 0,5–16 кбит/с. Однако до середины 80-х годов XX столетия качество сигналов при вокодерном кодировании было плохим, и это ограничивало их практическое использование. Современные вокодеры обеспечивают качество, ненамного уступающее принятому в телефонной сети общего пользования, и их широко применяют, в частности, в системах подвижной радиосвязи.

В гибридных кодерах используется метод, объединяющий преимущества кодеров формы и вокодеров.

Качество кодирования и восстановления речевых сигналов измеряется часто по пятибалльной шкале MOS (mean opinion score – средняя субъективная оценка). Поскольку человек как получатель информации является ключевым элементом любой телекоммуникационной системы, качество сигнала оценивается по его субъективному восприятию речи. Оценка по шкале MOS определяется путем обработки оценок, даваемых группами слушателей нескольким речевым сигналам, воспроизводимым различными громкоговорятелями. Каждый слушатель выносит оценку каждого сигнала: 1 – плохо, 2 – слабо, 3 – разборчиво, 4 – хорошо, 5 – отлично. Затем результаты усредняются. Соотношение качества и скорости для рассматриваемых методов приведено на рис. 2.2 [2].

2.3. Кодеры формы

Нелинейное кодирование. Для уменьшения сравнительно большого количества уровней квантования, которое вытекает из соотношений, полученных в разделе 2.1 в предположении об использовании квантователя с равномерным шагом квантования, следует учесть особенности работы слухового аппарата человека.

Человеческое ухо воспринимает звук нелинейно: наиболее заметными оказываются искажения при слабом уровне звука, в то время как при большом уровне звука чувствительность к искажениям звукового сигнала снижается. Принимая во внимание указанные особенности, можно уменьшить количество уровней квантования и, соответственно, скорость цифрового потока в телекоммуникационном канале, применив квантование с неравномерным шагом. Суть такого подхода состоит в изменении шага квантования пропорционально уровню входного сигнала. При этом малые уровни сигнала квантуются с меньшей ошибкой, чем большие. Закон изменения шага квантования определяют из условия, чтобы отношение сигнал–шум сохранялось постоянным при изменении уровня сигнала.

Условно неравномерное квантование можно представить как последовательное соединение устройства компрессии вход-

ного сигнала и равномерного квантователя (рис. 2.3). При приеме нелинейные искажения сигнала, вносимые компрессором, устраняют экспандером, нелинейным устройством с амплитудной характеристикой, обратной характеристике компрессора.

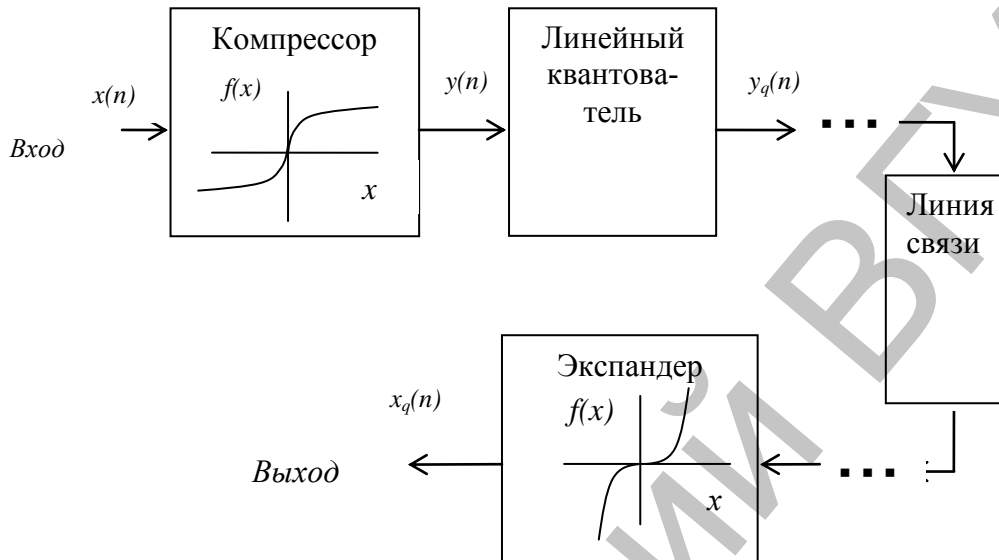


Рис. 2.3.

Для обозначения процессов компрессии и экспандирования для краткости пользуются одним термином – компандирование, а совокупность 2-х устройств – компрессора и экспандера называют компандером.

При передаче речи используют два типа компандирования: по μ -закону и по A -закону.

Первый метод используют в США и Японии. При μ -законе сигнал в компрессоре преобразуется следующим образом:

$$y_{\mu} = \frac{\text{sign}(x)}{\ln(1 + \mu)} \ln \left(1 + \mu \left| \frac{x}{x_{\max}} \right| \right), \quad (2.6)$$

где x – сигнал на входе компрессора,

x_{\max} – его максимальное значение,

μ – константа (обычно $\mu=255$).

A -закон используется в Европе. В этом случае компрессор преобразует сигнал следующим образом:

$$y_A = \begin{cases} \frac{A}{1 + \ln A} \left(\frac{x}{x_{\max}} \right), & \left| \frac{x}{x_{\max}} \right| \leq \frac{1}{A} \\ \frac{\text{sign}(x)}{1 + \ln A} \left(1 + \ln A \left| \frac{x}{x_{\max}} \right| \right), & \frac{1}{A} \leq \left| \frac{x}{x_{\max}} \right| \leq 1. \end{cases}$$

Наиболее часто используют значение параметра $A=87.6$.

Применение рассмотренных методов компадирования позволяет в одном и том же заданном диапазоне изменения речевого сигнала вместо 12-разрядных двоичных чисел использовать восьмиразрядные двоичные числа. Таким образом, скорость цифрового потока при передаче речевого сигнала уменьшится с 96 до 64 кбит/сек. Указанный способ компрессии речи закреплен в международной рекомендации G.711.

Учитывая, что сжатие и последующее восстановление к первоначальному виду непрерывных по величине отсчетов звуковых сигналов может сопровождаться появлением погрешности из-за отклонений характеристик компрессора и экспандера от расчетных значений, вместо рассмотренного выше на практике обычно применяется другой способ нелинейного кодирования (рис.2.4). После равномерного квантования при числе уровней $L = 2^{12}$ и предварительного кодирования производится цифровая компрессия, в результате чего длина кодовой комбинации уменьшается до $n=8$ разрядов. Результатом преобразования является двоичная последовательность со скоростью 64 кбит/с.

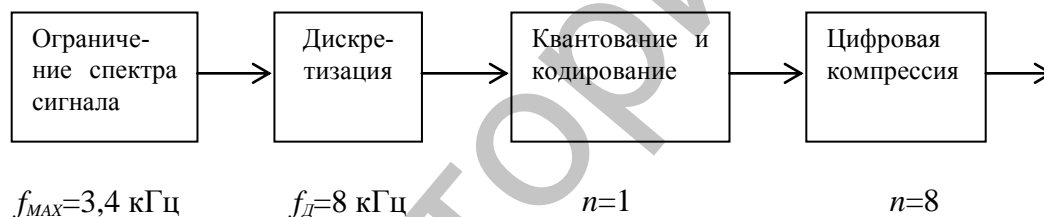


Рис. 2.4.

Дифференциальная импульсно-кодовая модуляция (ДИКМ). Обычно между двумя соседними отсчетами сигнала существует определенная взаимосвязь, которую в радиотехнике обозначают термином корреляция. Это справедливо для всех сигналов за исключением так называемого белого шума, отсчеты которого некоррелированы. Степень корреляции между отсчетами возрастает с ростом частоты дискретизации. Наличие корреляции указывает на наличие избыточности в сигнале.

Учитывая корреляцию между отсчетами, можно сжать сигнал по сравнению с обычной ИКМ. Самый распространенный метод кодирования, основанный на учете корреляции между отсчетами – ДИКМ-кодирование.

При ДИКМ кодируют и передают по каналу не сам отсчет (как в ИКМ), а разность (или ошибку) между текущим отсчетом и предварительной оценкой (предсказанным значением) этого отсчета, полученной из анализа предыдущих отсчетов

$$\varepsilon(n) = s(n) - s_{\text{ПРЕДСК}}(n). \quad (2.7)$$

Чем точнее осуществляется предсказание очередного отсчета, тем меньше по величине разностный сигнал и, следовательно, тем меньшее количество разрядов потребуется для его кодирования в цифровом виде. В качестве сигнала предсказания можно использовать либо предыдущий отсчет $s(n-1)$, либо M предшествующих отсчетов, что позволяет повысить точность предсказания:

$$s_{\text{ПРЕДСК}} = \sum_{i=1}^M c_i \cdot s(n-i), \quad (2.8)$$

где c_i – коэффициенты.

Типовой вариант реализации метода ДИКМ приведен на рис. 2.5.

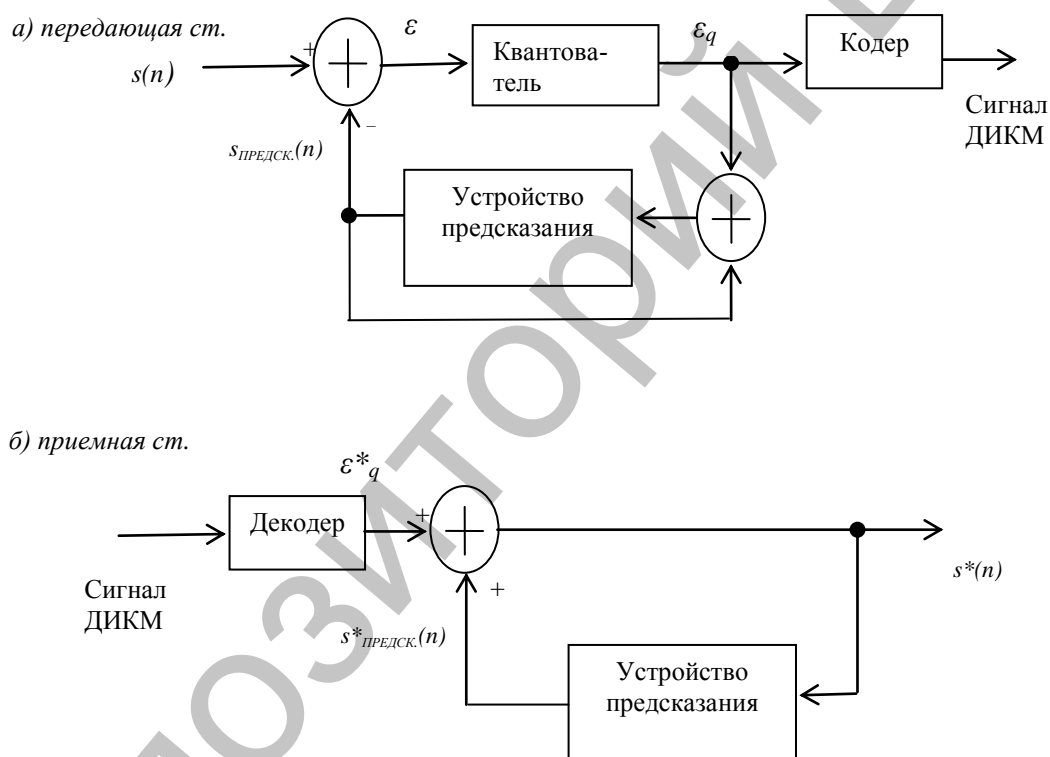


Рис. 2.5.

На приемной стороне (рис. 2.5,б) восстановление исходного сообщения осуществляется на основе соотношения

$$s^*(n) = \epsilon^*(n) + s^*_{\text{ПРЕДСК}}(n).$$

Эффективность метода ДИКМ может быть повышена путем перехода к адаптивной дифференциальной импульсно-кодовая модуляции (АДИКМ). При этом производится автоматическое регулирование величины шага квантования сигнала ошибки предсказания, а также автоматическая подстройка коэффициентов c_i в (2.8) в соответствии с изменением текущего спектра

передаваемого сообщения. Для этого как в передающее, так и в приемное устройства вводятся дополнительные цепи автоматической регулировки усиления и подстройки параметров передатчика на основе статистического оценивания параметров передаваемого сообщения. Этот алгоритм дает практически такое же качество воспроизведения речи, как и ИКМ, однако для передачи информации при его использовании требуется полоса всего в 32 кбит/с. Алгоритм АДИКМ был принят в качестве международного стандарта G.726 в 1984 г.

2.4. Вокодеры

Принципы вокодерного кодирования. Вокодеры предназначены для кодирования исключительно речевого сигнала. При их построении максимально учитывают особенности образования речи и ее восприятия человеком. Форма восстановленного сигнала при этом может радикально отличаться от формы исходного сигнала. В качестве примера на рис. 2.6 приведены осциллограммы двух сигналов, внешне весьма различных. Тем не менее, при их воспроизведении человек не заметит разницы. Дело в том, что спектральный состав обоих сигналов одинаков: они являются суммой синусоиды и ее третьей гармоники. Различны лишь значения начальной фазы третьей гармоники. Органы слуха человека не реагируют на фазовые соотношения.

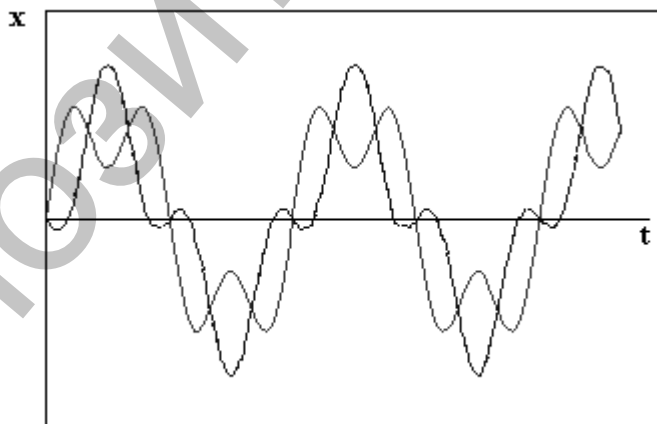


Рис. 2.6.

Задача ИКМ, ДИКМ, АДИКМ и других аналогичных им методов — максимально точно передать информацию о форме сигнала. Именно поэтому эти методы кодирования называют кодированием формы. Задача вокодерной обработки другая — обеспечить, чтобы восстановленный сигнал звучал как можно более сходно с исходным сигналом.

Принципиальное отличие вокодерного кодирования от кодирования формы состоит в том, что по каналу связи передают не сам сигнал, а параметры модели его образования. На приемном конце восстановленный сигнал синтезируют.

Существует большое число идей построения вокодеров. Например, в канальных или полосовых вокодерах спектр речи делят на 7–20 полос (каналов) аналоговыми или цифровыми полосовыми фильтрами. Большее число каналов дает большую натуральность и разборчивость. С каждого полосового фильтра сигнал поступает на детектор и фильтр низких частот. На приемный конец раз в 20 мс передают информацию об уровне сигнала в каждом канале. Синтезатор речи представляет собой набор синусоидальных генераторов и регулируемых аттенюаторов, устанавливающих требуемые соотношения между амплитудами колебаний разных частот. Передача информации об уровне сигнала в каждом канале возможна в аналоговом или цифровом виде.

В фонемных вокодерах используют тот факт, что речь передается ограниченным числом слогов – фонем. Например, русский язык использует 42 фонемы. Выполняя фонемный анализ речи, можно периодически (например один раз в 20 мс) передавать на приемный конец номер соответствующей фонемы, закодированный 6 битами, а также информацию об уровне сигнала (еще 6 бит). Таким образом, скорость цифрового потока составит $(6+6)/20=0,6$ кбит/с. На приемном конце синтезатор воспроизводит соответствующую фонему, извлекая ее из памяти. Известны и другие принципы вокодерного кодирования. Хотя первые вокодеры были предложены в 1930-е годы, до начала 1980-х годов качество восстанавливаемой речи было крайне низким. Область применения вокодеров ограничивалась линиями командной связи, речевого управления и говорящими автоматами информационно-справочных служб. При этом достигалась низкая скорость передачи (порядка 0,6–4 кбит/с).

Прогресс вокодеров в 80-е и 90-е годы XX в. непосредственно связан с новыми возможностями цифровой обработки сигналов и микропроцессорной техники. С другой стороны, он явился ответом на потребности быстро развивающегося рынка массовых цифровых систем подвижной радиосвязи, в частности сотовых систем.

Вокодеры используют достаточно сложные алгоритмы обработки речевых сигналов и по этой причине выполняются на основе цифровых сигнальных процессоров (ЦСП). Производительность ЦСП обычно оценивают в миллионах операций в секунду. Вокодеры, использующие ЦСП, способные выполнять 15 млн

операций в секунду, относятся к низкопроизводительным, если указанный параметр превышает 30 млн операций в секунду, то такие вокодеры считаются высокопроизводительными.

Некоторые особенности процесса речеобразования, учитываемые в вокодерах. Рассмотрим особенности процесса речеобразования. При разговоре грудная клетка сжимается и расширяется, поток воздуха проходит из легких через трахею и гортань в полости глотки, рта и носа. Голосовой тракт простирается от голосовой щели (отверстия между голосовыми складками в гортани) до губ. В процессе речеобразования его форма меняется.

Если произносятся звонкие звуки (гласные, носовые, звонкие согласные), голосовые складки в гортани смыкаются и размыкаются с той или иной частотой, которая называется частотой основного тона. Получается последовательность импульсов воздушного потока, которые возбуждают полости голосового тракта. Говоря, человек меняет геометрические размеры этих полостей, соответственно меняются и их резонансные частоты, которые называют формантами. Звонкие звуки называются также вокализованными.

Частота основного тона обычно находится в интервале от 50 до 400 Гц. На рис. 2.7 приведены временная зависимость и спектр, соответствующие гласному звуку «и». Хорошо виден периодический характер сигнала; в спектре ярко выражены основной тон и форманты.

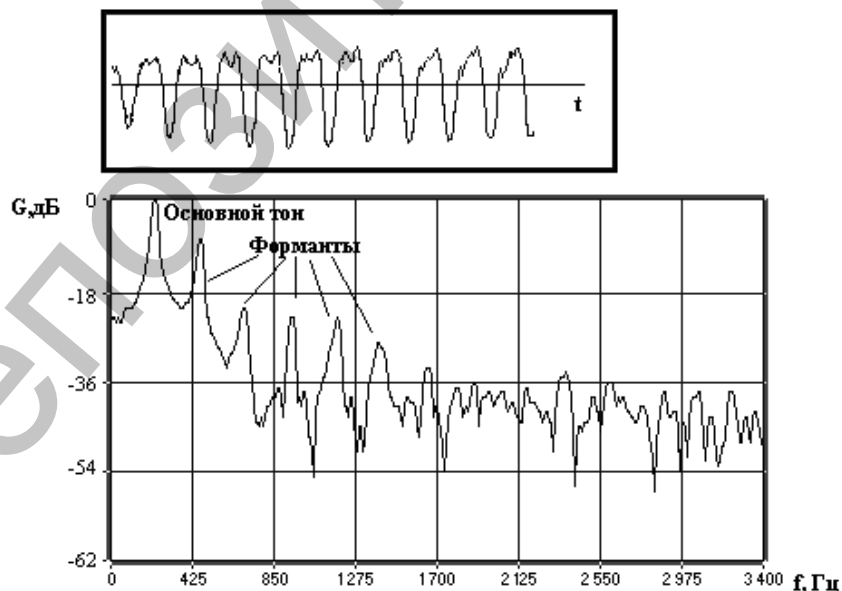


Рис. 2.7.

При произнесении глухих (невокализованных) звуков голосовые складки расслаблены. Проходя по суженному голосовому тракту, воздух создает турбулентный поток. Полости рта и носа возбуждаются при этом шумоподобным сигналом. На рис. 2.8

показаны временная зависимость и спектр, соответствующие глухому согласному звуку «с».

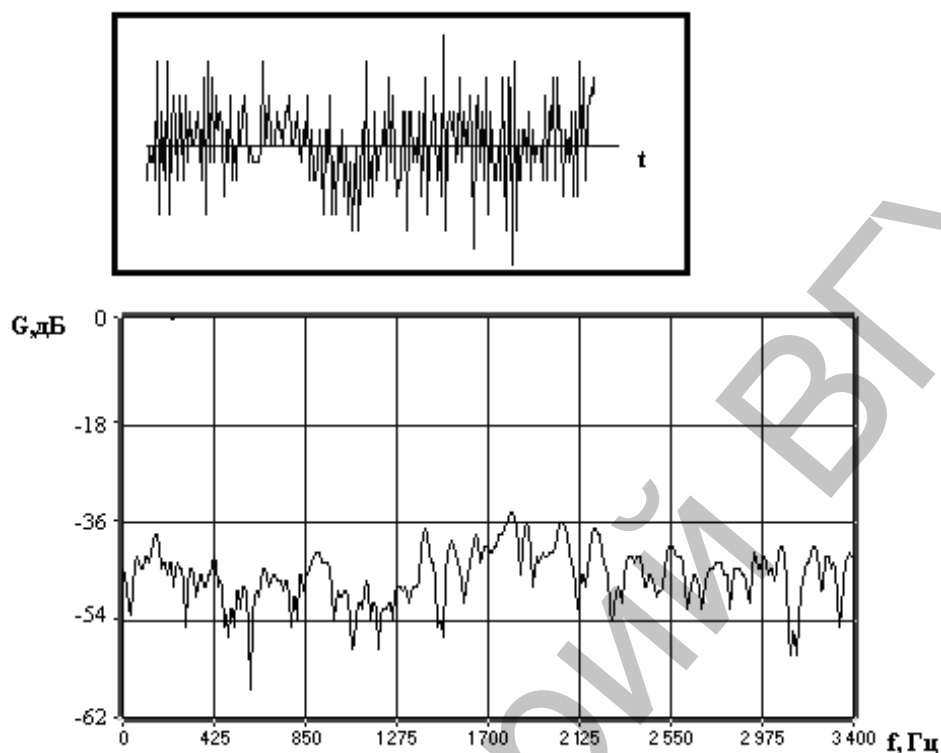


Рис. 2.8.

Сигнал не содержит периодических составляющих и подобен шуму; в спектре отсутствуют форманты и основной тон.

Взрывные (смычные) звуки получаются путем кратковременного выхлопа – полного перекрытия речевого тракта, нагнетания давления и внезапного открытия тракта. Взрывные звуки бывают звонкие (б, д, г) и глухие (п, т, к), то есть могут образовываться с участием голосовых складок и без них.

Органы речи обладают инерционностью: на интервале 20–30 мс параметры речи можно считать постоянными.

Метод линейного кодирования с предсказанием. Многие из методов вокодерного кодирования берут свое начало от изобретенного довольно давно метода LPC (Linear Predictive Coding). В качестве входного сигнала в LPC используется та же последовательность цифровых значений амплитуды, однако этот метод применяется не к отдельным цифровым значениям, а к определенным их блокам. Для каждого такого блока значений вычисляются его характерные параметры: частота, амплитуда и ряд других. Именно эти значения и передаются по сети. При таком подходе к кодированию речи, во-первых, возрастают требования к вычислительным мощностям специализированных процессоров, используемых для обработки сигнала, а во-вторых, увеличивается задержка при передаче, поскольку кодирование применяется не к

отдельным значениям, а к некоторому их набору, который перед началом преобразования следует накопить в определенном буфере. Подчеркнем, что задержка в передаче речи при использовании этого метода связана не только с необходимостью обработки цифрового сигнала (эту задержку можно уменьшать, увеличивая мощность процессора), а непосредственно следует из характера метода сжатия. Этот метод позволяет, вообще говоря, достигать очень больших степеней сжатия, которым соответствует полоса пропускания 2,4 или 4,8 кбит/с, однако качество звука здесь сильно страдает. Поэтому в коммерческих приложениях он не используется, а применяется в основном для ведения служебных переговоров.

2.5. Гибридные кодеры

Более сложные методы сжатия речи основаны на применении LPC в сочетании с элементами кодирования формы сигнала. В этих алгоритмах используется кодирование с обратной связью, когда при передаче сигнала осуществляется оптимизация кода. Закодировав сигнал, процессор пытается восстановить его форму и сличает результат с исходным сигналом, после чего начинает варьировать параметры кодировки, добиваясь наилучшего совпадения. Достигнув такого совпадения, аппаратура передает полученный код по линиям связи; на противоположном конце происходит восстановление звукового сигнала. Ясно, что для использования такого метода требуются еще более серьезные вычислительные мощности.

Одной из наиболее распространенных разновидностей описанного метода кодирования является метод LD-CELP (Low-Delay Code-Excited Linear Prediction). Этот метод позволяет достичь удовлетворительного качества воспроизведения при пропускной способности 16 кбит/с; он был стандартизован Международным союзом электросвязи (International Telecommunications Union – ITU) в 1992 году как алгоритм кодирования речи G.728 [3]. Алгоритм применяется к последовательности цифр, получаемых в результате аналого-цифрового преобразования голосового сигнала с 16-разрядным разрешением.

Пять последовательных цифровых значений кодируются одним 10-битовным блоком – это и дает те самые 16 кбит/с. Для применения этого метода требуются очень большие вычислительные мощности, в частности, для прямолинейной реализации G.728 необходим процессор с быстродействием 44 млн операций в секунду.

В марте 1995 года ИТУ принял новый стандарт G.723, который предполагается использовать при сжатии речи для организации видеоконференций по телефонным сетям. Этот стандарт является частью более общего стандарта H.324, описывающего подход к организации таких видеоконференций, при этом целью является обеспечение видеоконференций с использованием обычных модемов. Основой G.723 является метод сжатия речи MP-MLQ (Multipulse Maximum Likelihood Quantization). Он позволяет добиться весьма существенного сжатия речи при сохранении достаточно высокого качества звучания.

В основе метода лежит описанная выше процедура оптимизации; с помощью различных усовершенствований можно сжимать речь до уровня 4, 8; 6, 4; 7, 2 и 8,0 кбит/с. Структура алгоритма позволяет на основе программного обеспечения изменять степень сжатия голоса в ходе передачи. Вносимая кодированием задержка не превышает 20 мс.

Как показали испытания, проведенные ведущими американскими и европейскими телекоммуникационными компаниями, качество голоса, получаемое при сжатии методом MP-MLQ до уровня 6,4 кбит/с, не ниже того, что дает ADPCM при сжатии до 32 кбит/с.

Лекция 3. Цифровые методы передачи видеоизображений

3.1. Основные характеристики цифрового видео

Цифровое видео характеризуется четырьмя основными факторами [3]:

- частотой кадров;
- экранным разрешением;
- глубиной цвета;
- качеством изображения.

Частота кадров. Вследствие инерционности процесса восприятия изображения человеческий глаз интерпретирует последовательность быстро сменяющихся друг друга кадров не как отдельные изображения, а как непрерывно протекающий процесс. Это происходит при смене кадров с частотой 16–25 Гц. С учетом данной особенности стандартная скорость воспроизведения видеосигнала телевизионных системах – 25 кадров/с (в США и Японии – 30 кадров/с). Но даже при частоте 25 кадров/с может появляться эффект мелькания изображения. Для гарантированного устранения данного недостатка в телевизионных системах каждый

кадр разбивается на 2 полукадра, каждый из которых прорисовывается либо по четным, либо по нечетным строкам изображения. Эти полукадры передаются поочередно с частотой 50 Гц каждый. И хотя на самом деле частота смены кадров полного изображения составляет 25 кадров/с, глаз воспринимает изображение при чересстрочной развертке как смену кадров 50 раз в секунду.

Монитор компьютера для формирования изображения на экране использует метод прогрессивного сканирования, при котором строки кадра формируются последовательно, сверху вниз, а полный кадр прорисовывается 30 раз каждую секунду или даже более часто. В этом заключается основное отличие между компьютерным и телевизионным методом формирования видеосигнала.

Глубина цвета. Этот показатель является комплексным и определяет количество цветов, одновременно отображаемых на экране. Компьютеры обрабатывают цвет в RGB-формате (красный-зеленый-синий), в то время как видео использует и другие методы. Одна из наиболее распространенных моделей цветности для видеоформатов – YUV. Здесь Y обозначает яркостный сигнал, а сигналы цветности определяются по следующим формулам:

$$U=B-Y,$$

$$V=R-Y.$$

Разностные сигналы U и V формируют вместе с сигналом Y полный видеосигнал.

Каждая из моделей RGB и YUV может быть представлена разными уровнями глубины цвета (максимального количества цветов).

Для цветовой модели RGB обычно характерны следующие режимы глубины цвета: 8 бит/пиксел (256 цветов), 16 бит/пиксел (65535 цветов) и 24 бит/пиксел (16,7 млн цветов).

Экранное разрешение. Данным термином обозначают количество точек, из которых состоит изображение на экране. В современных мониторах может применяться разрешение 1024x768 пикселей, 1440x900 (широкий формат) пикселей и выше.

Телевизионный стандарт NTSC предусматривает разрешение 768 на 484. Стандарт PAL, распространенный в Европе, имеет несколько большее разрешение – 768 на 576 точек.

Качество изображения. Последняя, и наиболее важная характеристика – это качество видеоизображения. Требования к качеству зависят от конкретной задачи. Иногда достаточно, чтобы картинка была размером в четверть экрана с палитрой из 256-ти цветов (8 бит), при скорости воспроизведения 15 кадров/с. В других случаях требуется полноэкранное видео (768 на 576) с палитрой в 16,7 млн цветов (24 бит) и полной кадровой разверткой (25 или 30 кадров/с).

3.2. Свойства системы зрения человека

Очень часто окончательную оценку изображения делает человек. Исследование системы зрения человека показывает, что она обладает нелинейной характеристикой, а ее отклик не является абсолютно верным. Рассмотрим указанные особенности подробнее.

Одной из характеристик системы зрения человека является способность восприятия яркости света. Эксперименты по определению восприятия людьми минимально различимых градаций яркости света, поступающего от калиброванного источника, показали, что яркость света воспринимается глазом нелинейно [1]. Если начертить график зависимости величины этой минимально различимой градации яркости от эталонной яркости, то при изменении яркости в пределах нескольких порядков этот график имеет логарифмический характер. Такие субъективные экспериментальные результаты согласуются с объективными данными, полученными в экспериментах на животных, в которых установлено, что светочувствительные клетки сетчатки и оптический нерв возбуждаются с частотой, пропорциональной логарифму интенсивности подводимого к ним света. По вполне понятным причинам подобные объективные измерения на людях не проводились.

Другой отличительной особенностью системы зрения человека является ее пространственно-частотный отклик. Точная форма частотной характеристики глаза исследовалась с помощью ряда психовизуальных экспериментов. Было показано, что глаз подавляет низкие и ослабляет высокие пространственные частоты. В определенном смысле пространственно-частотный отклик имеет полосовой характер.

Наконец, для системы зрения человека характерна способность к насыщению, т.е. к ограничению отклика при очень больших или очень малых интенсивностях наблюдаемого светового потока. Рассмотренные особенности были учтены при разработке методов сокращения избыточности изображений, рассматриваемых ниже.

3.3. Межкадровая и внутрикадровая избыточность изображения

Избыточность изображения проявляется в высокой степени взаимной статистической прогнозируемости элементов изображения. В радиотехнике такую связь сигналов характеризуют по-

нятием *корреляция*. Конечной целью операции сжатия видеoinформации является устранение этой статистической прогнозируемости, т.е. необходимо в максимально возможной степени уменьшить коррелированность отсчетов, полученных при дискретизации видеосигналов.

Типичное изображение содержит очень много избыточной информации. Межкадровая избыточность изображений связана с необходимостью передавать изображение с достаточно высокой частотой. При этом изменение либо целого изображения, либо отдельных его участков от кадра к кадру может быть или небольшим, или даже отсутствовать. Такой эффект имеет место, например, когда в телевидении передается статическая заставка экрана. Поэтому одним из способов уменьшения избыточности передаваемых сигналов изображения является передача не абсолютных значений сигналов, соответствующих определенным элементам изображений в различных кадрах, а их изменения от кадру к кадру.

Внутрикадровая избыточность обусловлена высокой степенью однородности изображения на малых участках изображения в пределах одного кадра. Размер этого участка можно оценить, вычисляя коэффициент корреляции между яркостями точек изображения, находящихся на различном расстоянии от точки, выбранной в качестве опорной. Расстояние, при котором коэффициент корреляции становится меньше некоторой заданной величины (обычно 5–10% от максимального значения), и есть искомый размер. Анализ показал, что для большинства изображений размер участка, в пределах которого проявляется взаимосвязь яркостей точек изображения, составляет 16x16 точек [1].

3.4. Алгоритмы сжатия цифровых сигналов при передаче видеоизображений

Для того чтобы оценить, насколько актуальной является проблема компрессии цифровых сигналов при передаче видеоизображений, определим скорость цифрового потока, которая получится при преобразовании изображения размером 800x600 пикселей, следующего с частотой 25 кадров/секунду (напомним, что это минимальная частота, которая требуется для устранения эффекта мелькания при смене кадров), при глубине цвета 24 бит/пиксел:

$$W=800 \times 600 \times 25 \times 24=288 \text{ Мбит/с.}$$

Еще сравнительно недавно в середине 90-х годов XX столетия не существовало телекоммуникационных технологий, с помо-

щью которых такой цифровой поток в несжатом виде можно было бы передать по линиям связи. Другая проблема заключается в том, что при сохранении несжатых сигналов видеоизображений в компьютере объем его дискового пространства может оказаться сравнительно быстро исчерпанным. Ну и конечно же рассмотренный выше пример показал, что проблема компрессии цифровых сигналов при передаче по телекоммуникационным линиям связи сигналов видеоизображений является намного более актуальной, чем при передаче аудиосигналов.

Эффективность процесса компрессии сигналов видеоизображений оценивают с помощью коэффициента сжатия. Коэффициент сжатия – это цифровое выражение соотношения между объемом сжатого и исходного видеоматериала. Для примера коэффициент 200:1 означает, что если принять объем полученного после компрессии ролика за единицу, то исходный оригинал занимал объем в 200 раз больший. Обычно, чем выше коэффициент сжатия, тем хуже качество видео. Но многое, конечно, зависит от используемого алгоритма.

При определении необходимой степени сжатия сигналов видеоизображений следует исходить из разумной достаточности. При этом необходимо учитывать, как четыре характеристики (частота кадра, экранное разрешение, глубина цвета и качество изображения) влияют на объем и качество видео. Очень важно иметь представление, какую цену придется заплатить за качественное изображение. Чем больше глубина цвета, выше разрешение и лучше качество, тем более высокая производительность компьютера потребуется, не говоря уж о громадных объемах дискового пространства, необходимого под цифровое видео. Учитывая эти характеристики, можно выбрать оптимальный коэффициент сжатия. Надо отметить, что в профессиональном видео действует простое правило – чем ниже коэффициент сжатия, тем лучше.

Различают сжатие в режиме реального времени, симметричное или асимметричное, с потерей качества или без потери, сжатие видеопотока или покадровое сжатие.

Сжатие в режиме реального времени. Термин real-time (реальное время) имеет много толкований. Применительно к сжатию данных используется его прямое значение, т.е. работа в реальном времени. Многие системы оцифровывают видео и одновременно сжимают его, иногда параллельно совершая и обратный процесс декомпрессии и воспроизведения. Для качественного выполнения этих операций требуются очень высокопроизводительные специальные процессоры, поэтому некоторые платы вво-

да/вывода видео для персональных компьютеров не способны оперировать с полнометражным видео и часто пропускают кадры.

Симметричное и асимметричное сжатие. Этот признак классификации связан с соотношением способов сжатия и декомпрессии видео. Симметричное сжатие предполагает возможность проиграть видеофрагмент с разрешением, например, 640 на 480 пикселей при скорости в 30 кадров/с, если оцифровка и запись его выполнялась с теми же параметрами. Асимметричное сжатие – это процесс обработки одной секунды видео за значительно большее время. Степень асимметричности сжатия обычно задается в виде отношения. Так цифры 150:1 означают, что сжатие одной минуты видео занимает примерно 150 минут реального времени. Асимметричное сжатие обычно более удобно и эффективно для достижения качественного видео и оптимизации скорости его воспроизведения. К сожалению, при этом кодирование полнометражного ролика может занять слишком много времени, вот почему подобный процесс выполняют специализированные компании, куда отсылают исходный материал на кодирование (что увеличивает материальные и временные расходы на проект).

Сжатие с потерей или без потери качества. Чем выше коэффициент сжатия, тем больше уменьшается качество видео. Почти все методы сжатия явно приводят к потере качества. Даже если это не заметно на глаз, всегда есть разница между исходным и сжатым материалом.

Сжатие видеопотока или покадровое сжатие. Покадровый метод подразумевает сжатие и хранение каждого видеокadra как отдельного изображения. Сжатие видеопотока основано на следующей идее: не смотря на то, что изображение все время претерпевает изменения, задний план в большинстве видеосцен остается постоянным – отличный повод для соответствующей обработки и сжатия изображения. Создается исходный кадр, а каждый следующий сравнивается с предыдущим и последующим изображениями, а фиксируется лишь разница между ними. Этот метод позволяет существенно повысить коэффициент сжатия, практически сохранив при этом исходное качество.

3.5. Примеры форматов цифрового видео

AVI (Audio Video Interleave). Разработанный фирмой Microsoft метод сжатия, записи и воспроизведения движущихся изображений (Live Video) и звука на компьютере с использованием только программных средств. Файлы, созданные с использованием этого метода, имеют расширение AVI.

AVI может иметь или не иметь звуковые дорожки. При создании AVI файлов, включающих звуковое сопровождение, важным является правильная синхронизация звука с видеоизображением. Для этого используется технология чередования видеокадров и звука, которой, собственно, и определяется аббревиатура AVI (Audio Video Interleaved). Разные по типу видео- и аудиоданные записываются в один файл на диске следующим образом: все информационные потоки разбиваются на множество равных частей (chunks) и затем записываются в один файл друг за другом по очереди. Например, сначала записывается заголовок; затем – 1-я часть видео; затем – 1-я часть звука; затем – 2-я часть видео; затем – 2-я часть звука и т.д.

Motion-JPEG. Стандарт компрессии JPEG был разработан объединенной группой экспертов по фотографии (JPEG – Joint Photographic Expert Group) международной организации стандартов (ISO). Как ясно уже из названия, схема компрессии была разработана для неподвижных изображений. Так как телевидение, в сущности, и есть последовательность неподвижных изображений, то JPEG-кодирование может применяться и для компрессии видеоизображений. Иногда этот стандарт называют «динамический» JPEG.

В основе схемы компрессии JPEG лежит дискретное косинусное преобразование (DCT). К преимуществам JPEG относится тот факт, что каждый кадр сжимается независимо от остальных и для восстановления исходного изображения не нужно задействовать информацию из соседних кадров. Такое построение сжатых данных позволяет осуществлять произвольный доступ, коммутацию и монтаж видеофрагментов проще, чем при использовании других методов кодирования. Недостатком данного формата является сравнительно невысокое значение коэффициента компрессии, а также высокие требования к производительности процессора, от которого требуется декодировать каждый кадр скомпрессированного изображения за 1/25 долю секунды. Различные варианты Motion-JPEG позволяют получить значение коэффициента компрессии от 5:1 до 100:1, однако следует подчеркнуть, что уже при значении коэффициента компрессии 20:1 качество изображения в большинстве случаев становится неудовлетворительным.

MPEG. В январе 1992 года группа экспертов в области движущихся изображений MPEG (Motion Picture Experts Group) представила первую часть стандарта для сжатия цифрового видео и звука – MPEG phase 1, или просто MPEG-1 (ISO 11172). Стандарт определяет методы компрессии и воспроизведения видео- и аудиоданных. Комитет MPEG также определил ряд других форма-

тов для сжатого видео- и аудиоматериала. Форматы MPEG различаются по качеству результатов и скорости передачи данных:

MPEG-1: оригинальный формат для хранения и воспроизведения видео- и аудиоданных на мультимедиа носителях данных (компакт-дисках). Потенциально поддерживает телевизионное качество видео. Однако, при скорости передачи данных в диапазоне 150–255 Кбайт/сек. качество сопоставимо с видеозаписью VHS (разрешение 352 x 228 (PAL) или 320 x 240 (NTSC) при частоте 25 или 30 кадров в секунду соответственно).

MPEG-2: более новый стандарт (утвержден в ноябре 1994 г.). Разработан как дополнение к стандарту MPEG-1. Поддерживает передачу высококачественного видео по высокоскоростным цифровым каналам. Интенсивность потока данных от до 2 до 15 Мбит/сек. Разрешение 720 x 480 и 1280 x 720, частота 60 кадров в секунду со звуковыми данными CD-качества. Подходит для всех стандартов телевидения и даже систем телевидения высокой точности (High Definition Television). Используется при записи DVD дисков.

MPEG-4: предназначен для передачи видео- и аудиоданных по низкоскоростным линиям. Этот формат рассчитан для применения в системах видеотелефонии, мультимедийной электронной почте, электронных информационных изданиях и т.п. Базируется на формате файлов QuickTime. MPEG-4 версия 1 одобрен в октябре 1998 г. Стандарт ориентирован на разрешение 174 x 144 пиксела при 10 кадрах в секунду и позволяет передавать данные со скоростью от 4800 до 64000 бит/сек.

MPEG-1 и MPEG-2 признаны международными стандартами для сжатия видео.

Технология MPEG использует поточное сжатие видео, при котором обрабатывается не каждый кадр по отдельности (как это происходит при сжатии видео с помощью алгоритмов Motion-JPEG), а анализируется динамика изменений видеофрагментов и устраняются избыточные данные. Поскольку в большинстве моментов фон изображения остается достаточно стабильным, а действие происходит только на переднем плане, алгоритм MPEG начинает сжатие с создания исходного (ключевого) кадра. Играя роль опорных при восстановлении остальных изображений, они размещаются последовательно через каждые 10–15 кадров. Только некоторые фрагменты изображений, которые находятся между ними, претерпевают изменения, и именно эта разница сохраняется при сжатии. Таким образом, MPEG-последовательность содержит три типа кадров:

Intro frames (кадры типа «I») – опорные кадры, которые компрессируются без обращения к другим кадрам, кадры данного типа имеют самый большой размер;

Predicted frames (кадры типа «Р») – кадры данного типа компрессируются на основе обращения к предшествующему кадру типа «I» или предыдущему кадру типа «Р» с целью предсказания динамики изменений видеофрагментов и формирования разностного сигнала;

Bi-directional interpolated frames (кадры типа «В» – двунаправленные кадры), сжатие которых осуществляется при обращении к одному предшествующему и одному последующему кадру типа Р, имеют минимальный размер из рассматриваемых типов кадров.

Последовательность кадров, которая передается по телекоммуникационному каналу, в соответствии с рекомендациями стандарта MPEG, имеет вид:

I B B P B B P B B P B B P B B I ...

Отдельные изображения состоят из структурных единиц – макроблоков, соответствующих участку изображения размером 16X16 пикселей. Компьютер анализирует изображения и ищет идентичные или похожие макроблоки, сравнивая базовые и последующие кадры. В результате сохраняется только данные о различиях между кадрами, называемые вектором смещения (vector movement code) .

Макроблоки, которые не претерпевают изменений или претерпевают сравнительно небольшие изменения, при формировании разностного сигнала практически не вносят вклад, так что количество данных для реального сжатия и хранения существенно снижаются. В результате при использовании MPEG-технологии можно достигнуть рабочего коэффициента более чем 200:1, хотя это приводит к некоторой потере качества.

Лекция 4. Телекоммуникационные каналы

4.1. Свойства физических каналов

Было введено понятие телекоммуникационного канала (канала связи) как совокупности технических устройств и линии связи, необходимых для передачи сигнала между пунктами связи. Линию связи, представляющую собой физическую среду, в которой распространяется сигнал, называют физическим каналом. Классификация физических каналов приведена на рис. 4.1.

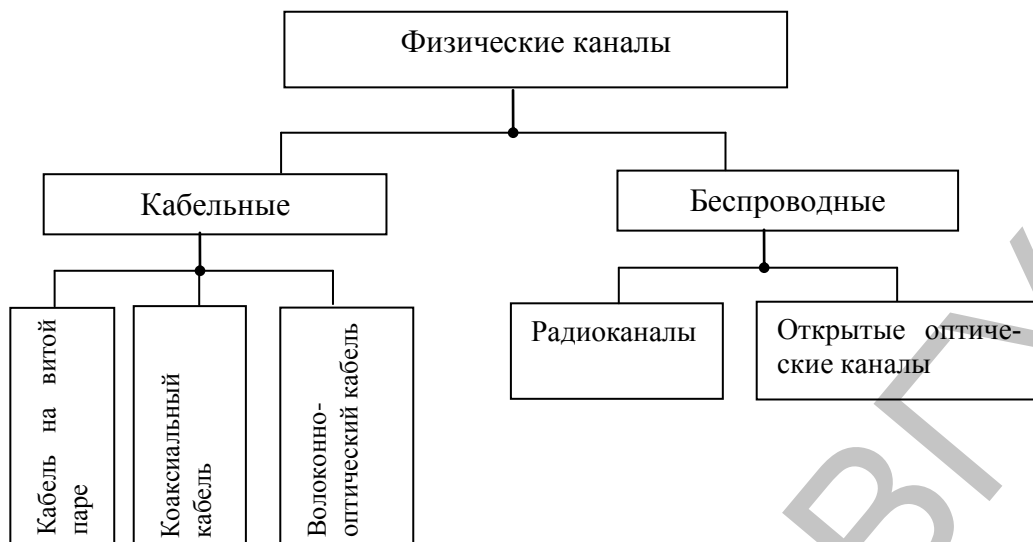


Рис. 4.1.

Кабели на витой паре. Витая пара представляет собой два изолированных медных провода, скрученных вместе. Скрутка проводов позволяет уменьшить индуктивность проводов, приводящую к ограничению технической скорости. Кроме того, скрутка проводов способствует уменьшению электрических помех, наводимых соседними парами (перекрестные помехи), а также внешними источниками, например, двигателями, реле и т.п. Все провода в кабеле обычно имеют один и тот же диаметр, хотя в специальных кабелях встречаются провода разных диаметров. В таких кабелях провода с большим диаметром часто используются для передачи высокочастотного сигнала, поскольку потери сигнала снижаются при увеличении диаметра проводника.

Кабель на витой паре характеризуется исключительной простотой монтажа и ремонта повреждений. К недостаткам такого кабеля относятся низкий уровень защищенности от электромагнитных помех, сравнительно большой уровень собственных излучений, способствующих затуханию сигнала, а также возможности «прослушивания» передаваемых сигналов. Для уменьшения недостатков на практике (в том числе и в локальных вычислительных сетях) иногда используют экранированный кабель (Shielded Twisted Pair, STP) – несколько витых пар, помещенных в защитную оболочку, которая заземляется. Неэкранированный кабель на витой паре (Unshielded Twisted Pair, UTP) также широко используется в качестве среды передачи в локальных вычислительных сетях. Кроме того, он применяется в телефонных сетях.

Типичные значения диаметра проводов – 0,4; 0,5; 0,63 и 0,9 мм (это диаметр самого провода). Провода покрываются изоляционным материалом, поэтому общий диаметр каждого прово-

да с изоляцией больше и обычно в два раза превышает диаметр самого провода. До 60-х гг. XX века применялась бумажная изоляция. Затем в качестве изоляционного материала стали использовать поливинилхлорид (ПВХ) или полиэтилен. Тип и толщина слоя изоляционного материала определяют емкость между проводами в паре. ПВХ обычно используется для внутренних кабелей, а полиэтилен – для внешних. Следует иметь в виду, что при горении ПВХ выделяет хлорный газ и другие вредные для здоровья людей вещества. Это привело к необходимости разработки специального изоляционного материала, который не выделяет вредных веществ при горении, а кабель, в котором он используется, получил название «кабель с малодымной изоляцией».

Провода в кабеле имеют определенный цвет изоляции, и поэтому пользователю легко выбрать провода, которые он собирается подключать. Провода с бумажной изоляцией маркировались цветными полосками через определенные интервалы вдоль провода. Современное пластмассовое покрытие провода кодируется цветом в соответствии с установленным стандартом. Каждый провод имеет два цвета: основной и вторичный, который используется либо при нанесении продольной полосы вдоль провода, либо для полос вокруг провода, наносимых через определенные интервалы.

Телекоммуникационные кабели могут содержать от 2 до 3000 витых пар. Число пар в кабелях внутренней проводки не превышает 200.

Входное полное сопротивление (импеданс) витой пары можно смоделировать двумя последовательно включенными сопротивлениями, параллельно одному из которых включена емкость (рис. 4.2).



Рис. 4.2.

На рис. 4.2а представлена эквивалентная схема замещения полного входного сопротивления телефонной линии для диапазона звуковых частот. На рис. 4.2б показана высокочастотная эквивалентная схема замещения полного входного сопротивления телефонной линии для частот выше звукового диапазона.

В диапазоне звуковых частот входное полное сопротивление витой пары имеет емкостный характер, но с увеличением частоты

реактивная составляющая постепенно уменьшается и входное полное сопротивление можно считать активным. Так, например, на частоте 100 кГц входное полное сопротивление витой пары имеет активный характер и равно 100 Ом. Зависимость входного полного сопротивления витой пары с проводом диаметром 0,5 мм от частоты показана на рис. 4.3. Учитывая, что в диапазоне звуковых частот входное полное сопротивление витой пары изменяется в несколько раз от значения 1500 Ом на частоте 300 Гц до значения приблизительно 250 Ом на частоте 3,4 кГц, при расчетах трансформаторов и других элементов линейного тракта в звуковом диапазоне частот его принимают равным постоянному значению 600 Ом, которое имеет место только на фиксированной частоте 1200 Гц. Это сделано для того, чтобы упростить процесс расчетов и измерений характеристик линий на витой паре, а также иметь опорную точку отсчета для сравнения этих характеристик.

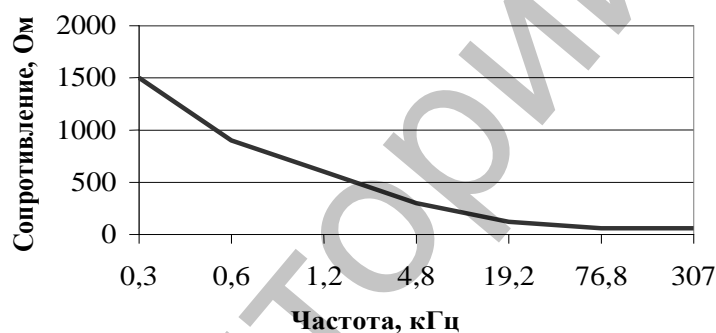


Рис. 4.3.

Затухание сигнала в кабеле на витой паре сильно зависит от частоты передаваемого сигнала. Затухание возрастает пропорционально частоте, но в определенных частотных полосах затухание остается практически постоянным (рис. 4.4).

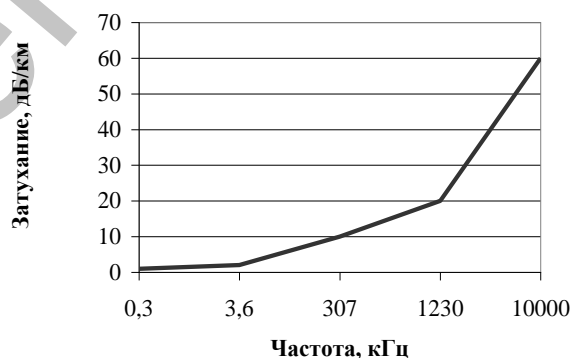


Рис. 4.4.

Характеристики кабелей на витой паре регламентируются в стандарте EIA/TIA 568, который был разработан усилиями нескольких организаций: Американского национального института стандартов ANSI, Ассоциации электронной промышленности EIA, Ассоциации телекоммуникационной промышленности TIA и лаборатории Underwriters Labs (UL). Этот стандарт включает 5 категорий неэкранированных кабелей на витой паре [4]:

- *категория 1*

традиционный телефонный кабель, по которому можно передавать только речь, передача данных не допускается;

- *категория 2*

кабель, предназначенный для передачи данных со скоростью 4 Мбит/с, состоит из 4-х витых пар;

- *категория 3*

кабель, предназначенный для передачи данных со скоростью 10 Мбит/с;

- *категория 4*

кабель, предназначенный для передачи данных со скоростью 16 Мбит/с;

- *категория 5*

кабель, предназначенный для передачи данных со скоростью 100 Мбит/с.

Максимальная длина сегмента кабеля типа UTP в локальных вычислительных сетях составляет 100 м. Затухание сигнала на частоте 10 МГц составляет около 1,0–3,0 дБ/км, а задержка сигнала не превышает 8–12 нс/м [5].

Экранированный кабель на витой паре также нашел применение в компьютерных сетях. Он применяется в ЛВС, реализованных на основе технологии Token Ring. Экранированный кабель делится по своим характеристикам на 9 типов: Type 1, Type 2, ..., Type 9. Электрические параметры кабеля Type 1 примерно соответствуют параметрам кабеля UTP категории 5. Однако ввиду того, что один из важных параметров кабеля – волновое сопротивление – у этих двух типов кабелей имеет заметно отличающиеся значения (у UTP категории 5 – 100 Ом, у STP типа 1 – 150 Ом), они не являются взаимозаменяемыми, так как это скажется на работе сетевых адаптеров – специальных устройств, с помощью которых компьютеры подключаются к линии связи.

Коаксиальные кабели. Коаксиальный кабель известен многим как подводный кабель телевизионной антенны. Этот кабель состоит из центрального проводника, изолированного твердым нейлоном или полиэтиленом (рис. 4.5). Изолирующий слой затем покрывается провололочной оплеткой или фольгой, которые вы-

полняют функции экрана, защищающего центральный проводник от внешних помех. Экран покрывают изолирующей оболочкой из ПВХ.

Стоимость коаксиального кабеля в несколько раз выше стоимости кабелей на витой паре, а монтировать его гораздо сложнее. Но наличие экрана существенно увеличивает помехозащищенность и снижает собственное излучение. Несанкционированное подключение к коаксиальному кабелю сложнее, чем в витой паре, но все равно возможно. В режиме модуляции высокочастотного сигнала пропускная способность коаксиального кабеля может достигать 500 Мбит/с, в немодулированном режиме она обычно лежит в пределах 50–100 Мбит/с [5].

Допустимая длина линии связи – несколько километров. Затухание сигнала на частоте 10 МГц составляет порядка 0,1–1,0 дБ/км. Задержка распространения сигнала – от 4 до 5 нс/м [5].



Рис. 4.5.

Входное полное сопротивление коаксиального кабеля для подсоединения радиоаппаратуры обычно равно 50 Ом. Кабель для телевизионных фидеров и в устройствах связи имеет входное полное сопротивление 75 Ом.

В локальных вычислительных сетях (ЛВС), реализованных на основе технологии Ethernet, получили распространение 2 типа коаксиальных кабелей, которые получили названия соответственно «тонкий Ethernet» и «толстый Ethernet». Характеристики этих кабелей приведены в табл. 4.1.

Затухание сигнала в коаксиальном кабеле меньше, чем у кабеля на витой паре и на частоте 10 МГц составляет 0,1–1,0 дБ/км.

Волоконно-оптические кабели. Волоконно-оптические кабели предназначены для передачи на расстояние оптических сигналов. Основным элементом волоконно-оптических кабелей являются световоды – тонкие нити из высокопрозрачных материалов. Оптическим сигналом служит модулированное оптическое (световое) излучение лазера или светодиода, переносимое по световоду в виде совокупности собственных оптических волн (мод) этих линий передачи.

Таблица 4.1

Наименование кабеля	Диаметр центрального проводника, мм	Наружный диаметр кабеля, мм	Максимальная длина сегмента кабеля, м	Примечание
тонкий Ethernet	0,89	6	185	
толстый Ethernet	2,17	12	500	из-за высокой жесткости трудно монтировать

Для описания процесса распространения оптических волн в материальных средах пользуются волновым и лучевым методами. Первый метод основан на решении уравнений Максвелла и позволяет получить точное решение электродинамической задачи. Однако во многих случаях решение уравнений Максвелла происходит крайне сложно. В этой связи на практике широкое распространение получили лучевые методы (методы геометрической оптики). Законы геометрической оптики используются для описания и анализа процесса передачи оптических сигналов, когда длина волны излучения λ значительно меньше размеров диаметра сердцевины оптического волокна (ОВ), т.е. $\lambda < 2a$, где a – радиус сердцевины.

В геометрической оптике световые волны изображаются световыми лучами, которые в однородной среде распространяются прямолинейно. При падении на границу раздела двух сред с разными значениями показателей преломления световой луч изменяет свое направление, и в общем случае появляются преломленный и отраженный лучи.

Углы, которые образуют падающий, отраженный и преломленный лучи с нормалью к границе раздела сред, восстановленной в точке падения, называют соответственно углами падения $\theta_{пад}$, отражения $\theta_{отр}$ и преломления $\theta_{плм}$. Угол падения равен углу отражения: $\theta_{пад} = \theta_{отр}$. Углы падения и преломления согласно закону Снеллиуса связаны соотношением [7]:

$$n_2 \sin \theta_{пад} = n_1 \sin \theta_{плм}, \quad (4.1)$$

где n_2 – показатель преломления среды, в которой распространяется падающий луч; n_1 – показатель преломления среды, в которой распространяется преломленный луч.

В общем случае показатель преломления n является характеристикой среды, показывающей отношение скорости распространения света в вакууме к скорости в рассматриваемой среде ($n = c/v$). Среда с большим и меньшим абсолютными показателями преломления называется соответственно оптически более плотной и оптически менее плотной. При переходе из оптически более плотной в оптически менее плотную среду, как видно из (4.1), изменяется $\theta_{над}$, а следовательно – $\theta_{нлм}$. В соответствии с этим при определенном значении $\theta_{над}$ преломленный световой луч будет распространяться по поверхности раздела сред, при этом $\theta_{нлм} = 90^\circ$. Угол падения, при котором происходит указанное явление, называется критическим или предельным и определяется выражением:

$$\theta_{кр} = \arcsin(n_1 / n_2). \quad (4.2)$$

При $\theta_{над} > \theta_{кр}$ световой луч полностью отражается от границы раздела сред и возвращается в среду падения. Это явление называется полным внутренним отражением. Для распространения световых лучей в оптически более плотной среде n_2 , без проникновения в менее плотную n_1 , необходимо соблюдать условие $\theta_{над} > \theta_{кр}$.

Процесс распространения световых лучей в оптически более плотной среде, окруженной менее плотной, показан на рис. 4.6. Траектория, показанная сплошной линией, соответствует световому лучу, который падает на границу раздела под углом $\theta_{над} > \theta_{кр}$, отражается от нее и возвращается в область более плотной среды, где распространяется зигзагообразно. Такие лучи называются направляемыми, и их траектории полностью расположены внутри среды распространения. Поскольку вся мощность направляющих лучей практически полностью возвращается в область более плотной среды, они могут распространяться на значительные расстояния. Траектория, показанная штриховой линией, соответствует лучу, который падает на границу раздела под углом $\theta_{над} < \theta_{кр}$ и испытывает не только отражение, но и, проникая в менее плотную среду, преломление. Такие лучи называются лучами излучения. В этом случае лучи, падающие на границу раздела под углом $\theta_{над} < \theta_{кр}$, при каждом отражении теряют часть своей мощности и поэтому претерпевают значительное затухание в процессе распространения.

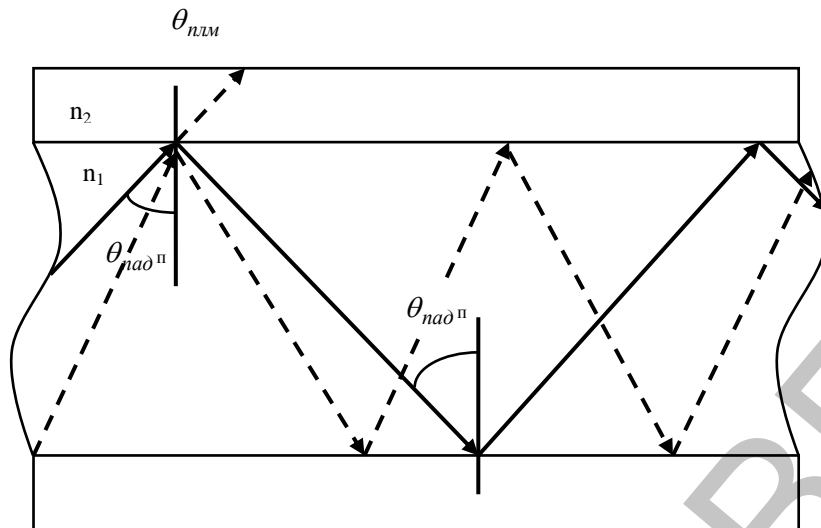


Рис. 4.6.

На рисунке изображен процесс распространения световых лучей в оптически более плотной среде, окруженной менее плотной, где:

- лучи сердцевины (направляемые лучи);
- лучи оболочки.

Оптические волокна, у которых показатель преломления на границе раздела сердцевины и оболочки изменяется по радиусу скачком, называются *ступенчатыми*. Оптические волокна, у которых показатель преломления сердцевины изменяется по радиусу плавно, называются *градиентными*.

В зависимости от параметров ОВ уравнения Максвелла могут иметь одно или несколько решений, каждому из которых соответствует определенный тип волны, называемый модой. Тип волны характеризует сложность структуры поля в ОВ числом максимумов и минимумов в поперечном направлении. Если в ОВ существует один тип волн, то оно называется *одномодовым*. При наличии нескольких типов волн ОВ называется *многомодовым*.

Важнейшим параметром ОВ являются потери, приводящие к ослаблению сигнала. Они определяют дальность передачи по оптическому кабелю и его эффективность. Ослабление световодных трактов волоконно-оптических кабелей α обусловлено *собственными потерями* в волоконных световодах (α_C) и *дополнительными потерями*, обусловленными деформацией и изгибами световодов при наложении покрытий и защитной оболочки при изготовлении кабеля (α_K).

Собственные потери волоконных световодов α_C состоят, в первую очередь, из потерь *поглощения* (α_{II}) и потерь *рассеяния* (α_P). Потери на поглощение существенно зависят от чистоты

материала и при наличии *посторонних примесей* ($\alpha_{\text{ПР}}$) могут быть значительными. Потери $\alpha_{\text{ПР}}$ возникают за счет посторонних примесей, таких, как гидроксильные группы, наличие ионов металла и других включений. Наличие этих примесей приводит к появлению резонансных всплесков поглощения на определенных длинах волн в кварце: 0,95; 1,24; 1,39 мкм, поэтому рабочие длины волн в световодах на базе кварца можно выбирать только в окнах прозрачности – 0,85; 1,3; 1,55 мкм. Кроме этого в настоящее время для изменения величины показателя преломления оптического волокна в чистый кварц сознательно вносятся *легирующие примеси (добавки)*, естественно увеличивающие суммарные потери. Поэтому при выборе легирующих примесей учитывают это обстоятельство.

С учетом изложенного можем записать:

$$\alpha_c = \alpha_{\text{П}} + \alpha_{\text{Р}} + \alpha_{\text{ПР}}. \quad (4.3)$$

Зависимость α_c от длины волны имеет 3 минимума, называемых «окнами прозрачности», соответственно в диапазонах волн 0,8; 0,9; 1,2; 1,3; 1,5; 1,6 мкм (рис. 4.7). При этом в диапазонах с увеличением длины волны ослабление существенно уменьшается. Так, при $\lambda = 0,85$ мкм $\alpha = 5$ дБ/км, при $\lambda = 1,3$ мкм $\alpha = 1$ дБ/км и при $\lambda = 1,55$ мкм $\alpha = 0,5$ дБ/км.

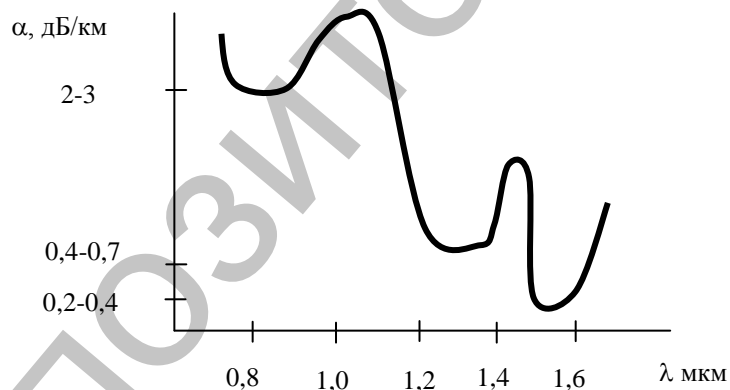


Рис. 4.7.

В связи с этим последнее время активно ведутся работы по освоению диапазона волн 1,5–1,6 мкм. В этом случае длину регенерационного участка удастся довести до 100 км и исключить из ОК металлические жилы для дистанционного питания линейных регенераторов.

В настоящее время проводятся работы по созданию оптических световодов для длинноволновой инфракрасной области на основе материалов, отличных от кварца. В световодах из поликристалла бромистого и бромидистого таллия на длинах волн 4–5 мкм получено затухание $\alpha = 0,01 \text{--} 0,005$ дБ/км.

Дополнительные потери в оптических кабелях (α_K) обусловлены деформацией оптических волокон в процессе изготовления кабеля, скруткой, изгибами волокон и технологическими неоднородностями в процессе изготовления волокна.

Наряду с ослаблением важнейшим параметром волоконно-оптических систем передачи является полоса частот ΔF , пропускаемая световодом. Она определяет объем информации, который можно передавать по оптическому кабелю (ОК). Ограничение ΔF применительно к цифровым системам передачи обусловлено тем, что импульс на приеме приходит размытым, искаженным вследствие различия скоростей распространения в световоде отдельных его частотных составляющих. Данное явление носит название дисперсии. Количественно дисперсия характеризуется разностью τ максимального и минимального времени распространения сигнала в ОВ□.

Полоса частот связана с дисперсией τ соотношением $\Delta F \cong 0,44/\tau$ [6]. Пропускная способность ОК существенно зависит от типа и свойств волоконных световодов (одномодовых или многомодовых, градиентных или ступенчатых), а также от типа излучателя (лазер или светоизлучающий диод).

Дисперсия возникает по двум причинам: некогерентность источников излучения и существование большого количества мод. Дисперсия, вызванная первой причиной, называется *хроматической (частотной)* τ_{xp} , она состоит из двух составляющих – *материальной* τ_m и *волноводной (внутримодовой)* τ_v дисперсией. Причина хроматической дисперсии – некогерентность источника излучения. Волноводная дисперсия связана с зависимостью коэффициента распространения от длины волны. Материальная дисперсия обусловлена зависимостью показателя преломления от длины волны.

Дисперсия, вызванная второй причиной, называется *модовой (межмодовой)* τ_{mod} . Она обусловлена наличием большого количества мод, каждая из которых распространяется со своей скоростью.

Сравнивая дисперсионные характеристики различных световодов, можно отметить, что лучшими данными обладают одномодовые световоды. Хорошие данные также у градиентных световодов с параболическим изменением показателя преломления. Наиболее сильно проявляется дисперсия у многомодовых ступенчатых световодов.

Частотная полоса пропускания существующих конструкций ОК колеблется в широких пределах и составляет 30–10000 МГц□км,

она различна для различных типов световодов. Для градиентных световодов с лазерным источником излучения частотная полоса составляет 100–250 МГц*км. В многомодовых ступенчатых световодах она сужается до 30 МГц*км. Наивысшей пропускной способностью обладают одномодовые световоды. У них полоса пропускания достигает 0,5–10 ГГц*км.

Явление дисперсии приводит как к ограничению пропускной способности ОК, так и к снижению дальности передачи по ним, т.к. чем длиннее линия, тем больше проявляется дисперсия и тем больше уширение импульсов.

Диаметры сердцевины и отражающей оболочки одномодового волокна (ОМВ) в соответствии с рекомендациями международного союза электросвязи (ITU) составляют соответственно 8–10 и 125 мкм. Для многомодового волокна (ММВ) эти параметры равны 50 и 125 мкм соответственно. Сразу после вытяжки ОВ на него наносится первичное защитное покрытие толщиной 5–10 мкм из материала, обладающего несколько большим коэффициентом преломления, чем светоотражающая оболочка ОВ, и более высокими потерями на поглощение для подавления распространения нежелательных мод. В отдельных случаях для защиты от механических воздействий на ОВ наносится вторичное защитное покрытие толщиной 200–300 мкм из материала с высокими механическими характеристиками и влагостойкостью. Иногда ОВ размещается в специальной второпластовой трубке (рис. 4.8), обеспечивающей механическую развязку ОВ от других элементов конструкции кабеля при его сгибе. Для предотвращения механической нагрузки на ОВ при прокладке, подвеске или в других аналогичных случаях используются металлические или синтетические упрочняющие стержни (рис. 4.8). Для защиты от воздействий внешней среды оптический кабель защищается наружной оболочкой. Структура некоторых типов волоконно-оптических кабелей (без наружной оболочки) показаны на рис. 4.8.



1 – оптическое волокно; 2 – сердечник; 3 – пластмассовая трубка.

Рис. 4.8.

Радиоканалы. Полоса частот, используемых в радиосвязи, составляет 3 Гц... 3000 ГГц, что соответствует длинам волн от 10^8 м до 0,1 мм. Классификация диапазонов частот и соответствующих им длин волн приведена в табл. 4.2.

Таблица 4.2

Диапазон частот	Диапазон волн	Название частот	Название волн
3...30 Гц	100000...10000 км	Крайне низкие (КНЧ)	
30...300 Гц	10000...1000 км	Сверхнизкие (СНЧ)	
300... 3000 Гц	1000...100 км	Инфранизкие (ИНЧ)	
3...30 кГц	100...10 км	Очень низкие (ОНЧ)	Мириаметровые
30...300 кГц	10...1 км	Низкие (НЧ)	Километровые
300...3000 кГц	1000...100 м	Средние (СЧ)	Гектометровые
3...30 МГц	100...10 м	Высокие (ВЧ)	Декаметровые
30...300 МГц	10...1 м	Очень высокие (ОВЧ)	Метровые
300...3000 МГц	100...10 см	Ультравысокие (УВЧ)	Дециметровые '
3...30 ГГц	10...1 см	Сверхвысокие (СВЧ)	Сантиметровые
30...300 ГГц	1...10 мм	Крайне высокие (КВЧ)	Миллиметровые
300...3000 ГГц	1...0,1 мм	Гипервысокие (ГВЧ)	Децимиллиметровые

По способу распространения радиоволн различают каналы с *открытым* и с *закрытым* распространением. В каналах с закрытым распространением электромагнитная энергия распространяется по направляющим линиям (кабельные, проводные, волноводные СВЧ тракты и др.). Для них характерны малый уровень помех и постоянство параметров сигнала, что позволяет передавать информацию с высокой скоростью и достоверностью.

Рассмотрим кратко особенности использования радиоволн различных диапазонов в каналах с открытым распространением [9]. В диапазонах КНЧ–НЧ на небольших расстояниях поле в месте приема создается за счет дифракционного огибания волнами выпуклой поверхности Земли. На больших расстояниях радиоволны распространяются в своеобразном сферическом волноводе, внутренняя стенка которого образуется поверхностью Земли, а внешняя – ионосферой. Такой механизм распространения позволяет принимать сигналы в любой точке Земли, причем параметры принятых сигналов отличаются достаточно высокой стабильностью. Особенностью этих диапазонов является также способность волн проникать в толщу Земли и воды. При этом

глубина проникновения возрастает с уменьшением частоты. Однако следует иметь в виду, что антенные системы в указанных диапазонах могут иметь достаточно большие размеры. Это обусловлено тем, что для эффективного излучения радиоволн размеры антенны должны быть соизмеримы с длиной волны. Если этому требованию удовлетворить невозможно (например, в диапазонах КНЧ и СНЧ), то тогда приходится увеличивать мощность радиопередающих устройств.

В распространении волн диапазона ВЧ принимает участие ионосфера. Однако если волны длиннее 1 км отражаются от нижнего ее слоя практически зеркально, то декаметровые волны достаточно глубоко проникают в ионосферу, что приводит к эффекту *многолучевости*, когда в точку приема приходит одновременно несколько сигналов с разными временами запаздывания. Многолучевость может носить *дисперсионный* или дискретный характер. Дисперсия (рассеяние) сигнала определяется отражением радиоволн от некоторого объема ионосферы, а дискретная многолучевость – отражением от разных слоев ионосферы. Так как глубина проникновения в ионосферу зависит от длины волны, то для передачи информации между двумя пунктами можно указать оптимальную рабочую частоту (ОРЧ), на которой связь будет наиболее надежной (максимум мощности принимаемого сигнала, минимум эффекта многолучевости). Значения ОРЧ рассчитывают для определенных трасс и времени связи. Для этого составляют долговременные и кратковременные прогнозы по данным мировой сети станций ионосферного зондирования. Декаметровые волны широко применяются для глобальной связи и радиовещания. С их помощью можно передавать информацию сравнительно большого объема в пределах всего земного шара при ограниченной мощности передатчика и небольших по размеру антеннах. До появления спутниковых систем связи этот диапазон был единственным пригодным для организации связи между двумя любыми пунктами на Земле без промежуточной ретрансляции. Однако эффект глобального распространения коротких волн имеет и свою отрицательную сторону – в точке приема могут появиться сильные помехи от дальних радиостанций.

Гектометровые волны днем распространяются как земные, а ночью – как ионосферные. Дальность распространения земной волны над сушей не превышает 500 км, а над морем – 1000 км. Диапазон СЧ широко используется в радиовещании, связи и радионавигации.

Волны диапазона частот от 30 МГц и выше слабо дифрагируют и поэтому распространяются в пределах прямой видимости. Не-

которого увеличения дальности можно достичь, применив поднятые антенны, а для организации связи на расстояния, превышающие прямую видимость, – ретрансляцию сигналов. Системы с ретрансляцией сигналов называются *радиорелейными линиями*. Одним из основных достоинств высокочастотных диапазонов является большой частотный ресурс, что позволяет создавать радиосистемы передачи информации с высокой скоростью передачи и радиосети с большим числом одновременно работающих радиостанций.

Стремление увеличить дальность радиолинии в этих диапазонах без промежуточной ретрансляции нашло свое решение в радиосистемах, использующих рассеяние радиоволн на неоднородностях тропосферы, ионосферы и метеорных следах. Однако такие системы по качеству передачи информации не могут конкурировать с радиорелейными линиями того же диапазона, поэтому их имеет смысл применять тогда, когда строительство радиорелейных линий по тем или иным причинам невозможно.

Стремление увеличить ширину полосы частот канала, а также повысить пространственную селекцию сигналов за счет использования остронаправленных антенн при их ограниченных размерах привело к освоению диапазона миллиметровых и децимиллиметровых волн. Главной особенностью их с точки зрения распространения является сильное поглощение в дожде и тумане, что ограничивает их применение в наземных системах большой дальности. Однако в космических и спутниковых системах они весьма перспективны.

Новую эру в освоении высокочастотной области радиодиапазона для средств связи открыл запуск искусственных спутников Земли (ИСЗ). Обычно ИСЗ находятся на высоте от 500 до 40 000 км от поверхности Земли и поэтому обеспечивают радиосвязь между земными станциями, удаленными на расстояния до 10–17 тыс. км. Линия спутниковой связи состоит из двух конечных земных станций и одного или нескольких спутников-ретрансляторов, обращающихся вокруг Земли по заданным орбитам.

Выбор рабочих частот для линии радиосвязи через ИСЗ определяется следующими факторами: условиями распространения и поглощения радиоволн, уровнем внешних помех, принимаемых антенной, техническими средствами (коэффициент шума приемного устройства, ширина лепестка диаграммы направленности антенны, точность ориентации и т.п.), взаимными помехами между системами связи через ИСЗ и другими службами, работающими в смежных или совмещенных диапазонах частот. Ограничение диапазона частот снизу определяется экранирующим действием ионосферы, а сверху – поглощением в тропосфере. Эти два фактора предопределили диапазон рабочих частот 30 МГц –

40 ГГц. В настоящее время наибольшее использование находит диапазон 1–12 ГГц.

Открытые оптические каналы. В системах, в которых используются открытые оптические каналы (ООК), передача оптических сигналов осуществляется непосредственно через атмосферу. ООК могут использоваться при объединении компьютеров в локальные вычислительные сети, при объединении различных ЛВС, в телеметрических системах, системах безопасности и т.п.

Выпускаемая в настоящее время аппаратура позволяет обеспечить передачу оптических сигналов на расстояние до 3-х километров. Однако в зависимости от погодных условий это расстояние может уменьшаться. Передача осуществляется в дуплексном синхронном режиме со скоростью 2048 кбит/с. К приемно-передающей станции подводятся информационный кабель от компьютерной сети и напряжение питания.

Оптическая система, устанавливаемая на каждом конце линии связи, состоит из двух станций, содержащих излучающий лазерный блок, приемный оптический блок и соответствующее электронное оборудование. В передающем блоке обычно используется полупроводниковый лазер, например, GaAlAs, инфракрасное излучение которого формируется внешней оптической системой в луч с очень малой угловой расходимостью. Излучение лазера модулируется передаваемым цифровым потоком управляющего интерфейса. После прохождения излучения через атмосферу, линзовая оптическая система противоположной станции фокусирует принятое излучение на фотоприемник. Электронные системы станции усиливают принятый фотоприемником сигнал и осуществляют синхронную регенерацию исходного цифрового потока. В оптическом приемно-передающем блоке используется линзовая система с большой апертурой для обеспечения приема и передачи информации в сложных погодных условиях. Оптическая линия связи хорошо защищена от несанкционированного доступа к передаваемым данным. Поскольку сигнал невидим и хорошо сфокусирован, попытка доступа к данным невозможна без нарушения связи, что немедленно обнаруживается пользователями.

Высококачественная оптическая система и электронное оборудование устанавливаются в водонепроницаемом корпусе, благодаря этому неблагоприятные погодные условия не оказывают влияния на работоспособность системы связи. Каждый приемно-передающий блок оптической атмосферной системы связи устанавливается на универсальном юстировочном устройстве и комплектуется оптическим прицелом для визуальной настройки станций друг на друга (рис. 4.9).



Рис. 4.9.

Для надежной работы во всепогодных условиях оптический приемо-передающий блок снабжают системой антиобледенения, исключающей образование конденсата влаги и ледяной корки на оптических компонентах системы при любых изменениях состояния атмосферы.

Ниже в качестве примера в табл. 4.3 приведены характеристики оптической атмосферной системы связи SBAL-2/3, выпускаемой фирмой Granch (г. Новосибирск).

Таблица 4.3

Скорость передачи данных	2048 кбит/с
Режим передачи	Синхронный, полный дуплекс
Рабочая дистанция (вероятность доступа 85% в погодных условиях средней полосы России)	3000 м
Рекомендуемая дистанция (вероятность доступа 95% в погодных условиях средней полосы России)	1500 м
Расходимость оптического излучения	2 мрад
Оптический излучатель	Ga Al As лазер
Длина волны оптического излучения	0,85 мкм
Мощность излучения	20 мВт
Чувствительность	-45 dBm
Расстояние от приемо-передающего блока до интерфейсного устройства	1000 м
Размеры приемо-передающего блока	360 x 318 x 362 мм
Масса	12 кг
Потребляемая мощность / с системой антиобледенения	10/50 Вт
Напряжение питания	220 В

Благодаря низкой мощности выходного сигнала (менее 100 мВт) не требуется специального технического лицензирования для безопасной работы и управления.

При ошибке передачи данных BER $<10^{-9}$ и скорости 2000 кбит/с, сильный снегопад и густой туман существенно сокращают максимальную дистанцию устойчивой связи (Примечание: BER – от англ. Bit Error Rate – коэффициент ошибок, определяемый как отношение числа неверно принятых битов (0 вместо 1 и наоборот) к полному числу переданных битов при передаче по каналу связи. В современных сетях связи характерные значения BER – $1E-9$ и лучше). Как показывают наблюдения, и в сильный снегопад, и в густой туман на расстоянии 3 км связь сохраняется, но, поскольку относительное количество испорченных пакетов очень велико, скорость передачи данных существенно уменьшается.

4.2. Системы многоканальной связи

Учитывая относительно высокую стоимость линейных сооружений, особое значение приобретает обеспечение их эффективного использования. Одним из путей решения указанной проблемы является создание систем многоканальной связи (СМС), позволяющих осуществить передачу по общей линии связи большого количества сигналов.

Для того, чтобы на приемной стороне можно было различить поступающие сигналы, они должны обладать определенными свойствами: не перекрываться либо по частоте, либо во времени. Существует еще один способ различения сигналов – по форме. С учетом изложенного, СМС подразделяют на системы с частотным разделением каналов, временным разделением каналов, а также системы с разделением сигналов по форме.

Системы с частотным разделением каналов. Системы с частотным разделением каналов (ЧРК) нашли широкое применение в различных областях связи. Рассмотрим особенности их реализации в телефонии. Как известно, полоса частот канала тональной частоты (ТЧ) составляет 300...3400 Гц. Это означает, что для передачи по общей линии связи сигналов источников сообщений I_1, I_2, \dots, I_N (рис. 4.10), занимающих одну и ту же полосу частот, их необходимо предварительно преобразовать, в противном случае их невозможно будет различить на приемной стороне. Такое преобразование осуществляется в канальных модуляторах $КМ_1, КМ_2, \dots, КМ_N$, на один вход которых поступают сигналы источников сообщений I_1, I_2, \dots, I_N , а на второй вход – гармонические сигналы, формируемые генератором канальных

сигналов (ГКС). Последние имеют одинаковую амплитуду и различающиеся между собой частоты: f_1, f_2, \dots, f_N .

В канальных модуляторах $КМ_1, КМ_2, \dots, КМ_N$ сигналы источников сообщений $И_1, И_2, \dots, И_N$ модулируют гармонические сигналы, поступающие от ГКС. В результате такого преобразования спектры исходных сигналов источников сообщений переносятся в более высокочастотную часть и располагаются на оси частот упорядоченно возле соответствующей частоты: f_1, f_2, \dots, f_N (рис. 4.11). За счет рационального выбора опорных частот f_1, f_2, \dots, f_N можно выполнить условие, чтобы сигналы, поступающие на выход канальных модуляторов, не перекрывались по частоте. Для надежного исключения перекрытия по частоте вводятся небольшие защитные интервалы Δf_s (рис. 4.11).

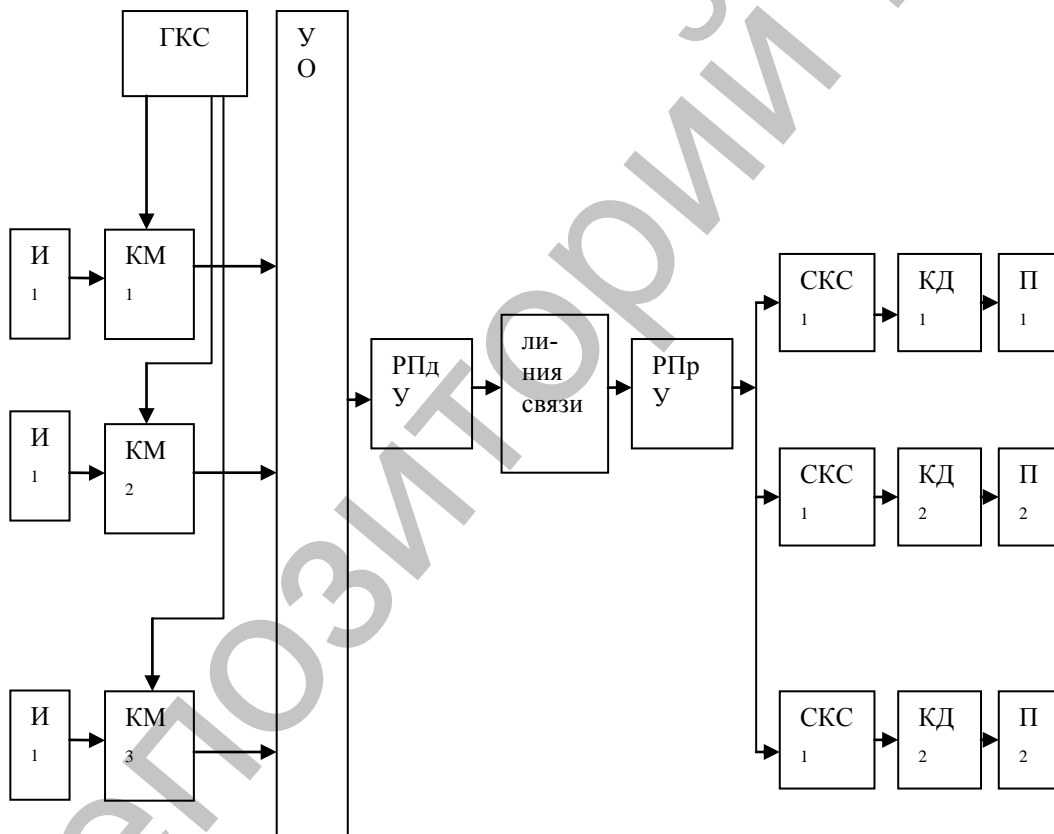


Рис. 4.10.

С выхода канальных модуляторов сигналы поступают на вход устройства объединения (УО), на выходе которого формируется групповой сигнал, представляющий собой линейную суперпозицию канальных сигналов. Групповой сигнал поступает на вход радиопередающего устройства (РПДУ) и далее – в линию связи и в радиоприемное устройство (РПРУ). Более детально рассмотрим преобразования, которым подвергается групповой сиг-

нал, поступающий на выход РПрУ. Для того, чтобы выделить из группового сигнала соответствующий каналный сигнал, используются селекторы каналных сигналов (СКС). Каждый СКС должен обладать свойством пропускать частотные составляющие спектра «своего» каналного сигнала, лежащие в полосе частот $(f_{нi}) \dots (f_{ви})$, где $i=1, 2, \dots, N$ (рис. 4.11), и в максимально возможной степени ослаблять частотные составляющие остальных каналных сигналов, лежащие вне указанной полосы частот. Для осуществления этой функции в качестве СКС используют полосовые фильтры.

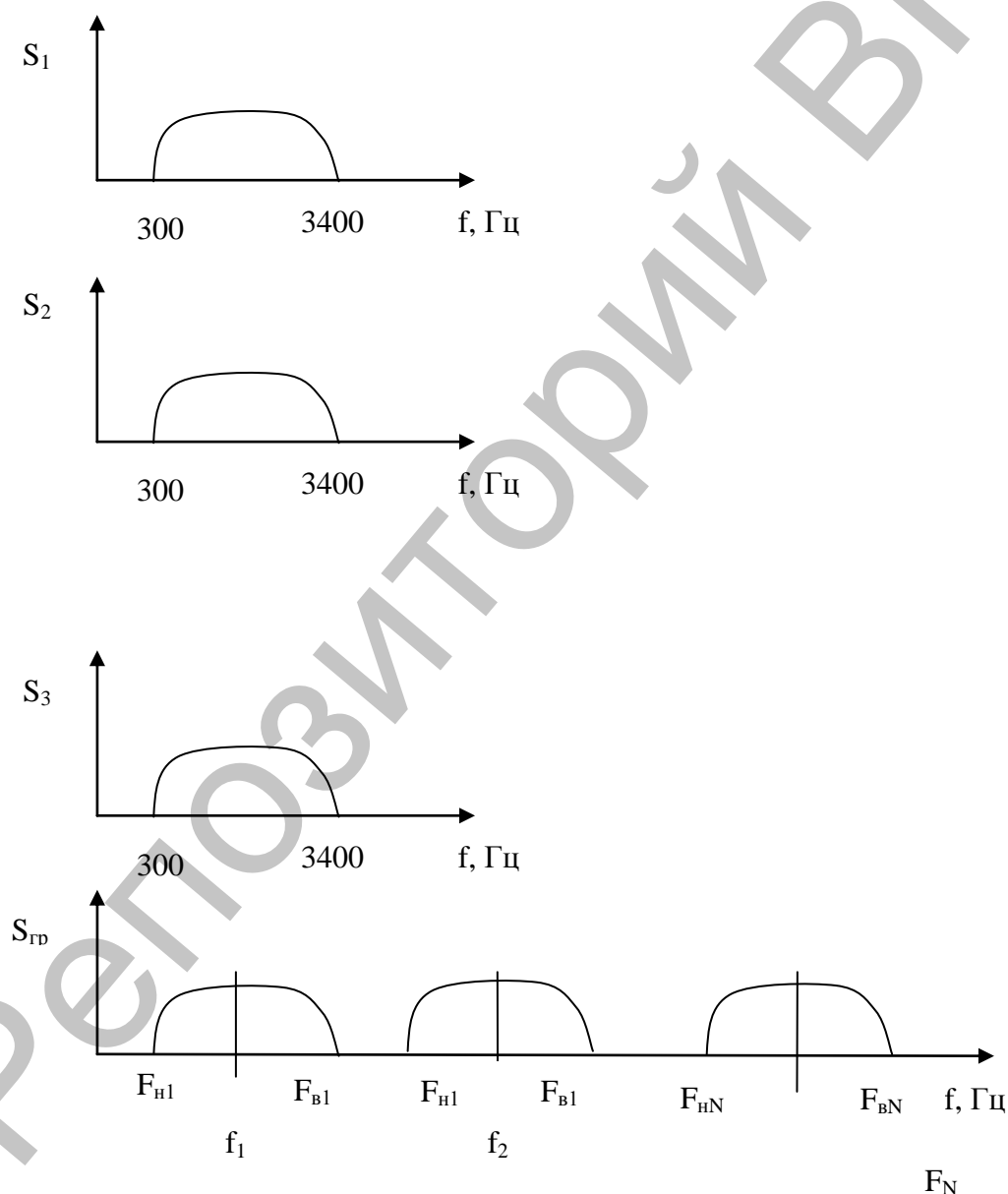


Рис. 4.11.

Необходимо помнить, что сигналы, поступающие на выходы селекторов каналных сигналов, еще не является теми сигналами,

лами, которые могут быть переданы соответствующему получателю ($\Pi_1, \Pi_2, \dots, \Pi_N$), поскольку их спектры были смещены в более высокочастотную область в канальных модуляторах. Поэтому перед тем, как сигналы поступят получателям, их спектры необходимо вернуть в исходную полосу частот (300–3400 Гц). Эта операция выполняется в канальных детекторах $КД_1, КД_2, \dots, КД_N$.

Таким образом, рассмотренные выше принципы позволяют реализовать передачу сравнительно большого количества сигналов, спектры которых перекрываются по частоте, по одной линии связи.

В соответствии с рассмотренными принципами была реализована аппаратура систем многоканальной связи с ЧРК, которая разрабатывалась с учетом рекомендаций Международного консультативного комитета по телеграфии и телефонии (МККТТ). В соответствии с рекомендациями каналообразующая аппаратура классифицируется по нескольким группам (табл. 4.4) [8].

Таблица 4.4

Наименование группы	Количество объединяемых групп предыдущего уровня	Количество объединяемых каналов ТЧ	Занимаемая полоса частот, кГц
Первичная	–	12	60–108
Вторичная	5	60	312–552
Третичная	5	300	812–2204
Четверичная	3	900	8516–12388

Хотя технически несложно распространить мультиплексирование и на более высокие частоты (в некоторых европейских странах используется 10800-канальная система, в которой в полосе частот 4–60 МГц передается двенадцать 900-канальных четверичных групп каналов), развитие цифровых методов передачи сделало экономически невыгодным дальнейшее развитие систем многоканальной связи с ЧРК.

Системы с временным разделением каналов. Принцип *временного разделения каналов* (ВРК) состоит в том, что групповой тракт предоставляется поочередно для передачи сигналов каждого канала многоканальной системы.

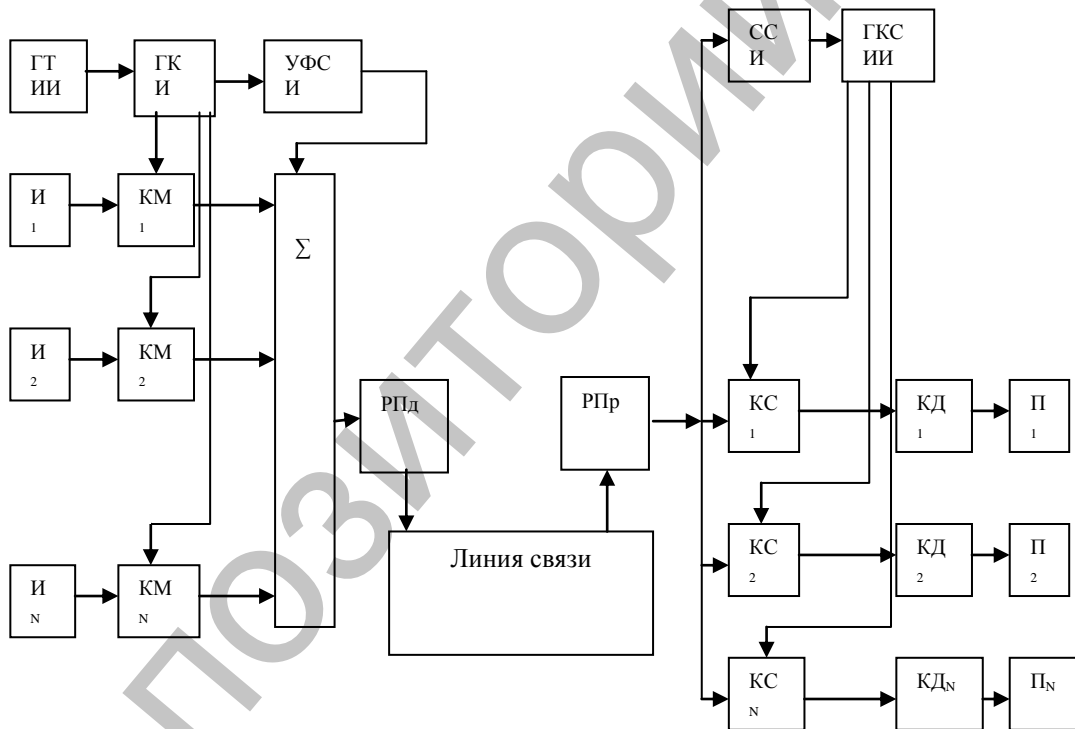
В зарубежных источниках для обозначения принципа временного разделения каналов используется термин Time Division Multiply Access (TDMA).

При передаче используется дискретизация во времени (импульсная модуляция). Сначала передается импульс 1-го канала, затем следующего канала и т.д. до последнего канала за номером N , после чего опять передается импульс 1-го канала и процесс

повторяется периодически. На приеме устанавливается аналогичный коммутатор, который поочередно подключает групповой тракт к соответствующим приемникам. В определенный короткий промежуток времени к групповой линии связи оказывается подключена только одна пара приемник/передатчик.

Это означает, что для нормальной работы многоканальной системы с ВРК необходима синхронная и синфазная работа коммутаторов на приемной и передающей сторонах. Для этого один из каналов занимают под передачу специальных импульсов синхронизации.

Интервал времени между ближайшими импульсами группового сигнала T_K называется канальным интервалом или тайм-слотом (Time Slot). Промежуток времени между соседними импульсами одного индивидуального сигнала называется циклом передачи $T_{Ц}$. От соотношения $T_{Ц}$ и T_K зависит число импульсов, которое можно разместить в цикле, т.е. число временных каналов.



4.12. Структурная схема системы с ВРК.

ГТИ – генератор тактовых импульсов.

ГКИ – генератор канальных импульсов, задающих момент формирования дискретного отсчета.

$И_1...И_N$ – источники сообщений.

$КМ_1...КМ_N$ – канальные модуляторы, формирующие отсчеты в моменты времени, задаваемые ГКИ.

УФСИ – устройство формирования синхроимпульсов, необходимых для синхронизации приемной аппаратуры.

Σ – устройство формирования группового сигнала.

РПд – радиопередающее устройство.

РПр – радиоприемное устройство.

ССИ – селектор синхроимпульсов, обеспечивающий выделение (селекцию сигналов синхронизации).

ГКСИ – генератор канальных синхроимпульсов.

$КС_1 \dots КС_N$ – канальные селекторы, выделяющие из группового сигнала отсчеты данного канала.

$КД_1 \dots КД_N$ – канальные демодуляторы, восстанавливающие непрерывные сигналы на основе дискретных отсчетов.

$П_1 \dots П_N$ – получатели сообщений.

Как уже отмечалось, в системах с ВРК для правильной обработки и распределения сигналов по каналам на приемной стороне необходимо обеспечить синхронизацию. Различают 3 вида синхронизации: тактовую, цикловую и сверхцикловую. Тактовая синхронизация обеспечивает на приемной стороне правильное «считывание» из группового сигнала каждого разряда передаваемого канального сигнала. Для обеспечения тактовой синхронизации используют так называемые самосинхронизирующиеся коды, которые позволяют сохранить синхронизацию тактовых передатчика и приемника даже при передаче длинных последовательностей «1» и «0». Примерами самосинхронизирующихся кодов являются манчестерский код, применяемый в локальных вычислительных сетях, реализованных по технологии Ethernet или Token Ring, или код HDB-3, используемый в цифровых системах передачи сообщений первичного уровня (см. ниже).

Цикловая синхронизация обеспечивает на приемной стороне правильное определение каждого канального интервала. Она реализуется путем передачи сигнала синхронизации в позиции нулевого канального интервала. Пятнадцать циклов объединяются в сверхцикл. Для синхронизации цифровых потоков в сверхциклах в групповой сигнал вводятся сигналы сверхцикловой синхронизации.

Системы многоканальной связи с временным разделением каналов также как и системы многоканальной связи с частотным разделением каналов строятся по *иерархическому принципу*. Применительно к цифровым системам этот принцип заключается в том, что на каждом более высоком уровне иерархии происходит объединение в групповой сигнал нескольких выходных сигналов предыдущего уровня иерархии.

Цифровая система передачи, соответствующая первой ступени иерархии, называется первичной; в этой ЦСП осуществляется прямое преобразование относительно небольшого числа первичных сигналов в первичный цифровой поток. Системы передачи второй ступени иерархии объединяют определенное число первичных потоков во вторичный цифровой поток и т.д.

В рекомендациях МСЭ-Т представлено два типа иерархий ЦСП: плезиохронная цифровая иерархия (ПЦИ) и синхронная цифровая иерархия (СЦИ). Появившаяся исторически первой плезиохронная цифровая иерархия имеет *европейскую, североамериканскую и японскую* разновидности (табл. 4.5).

Используя данные, приведенные в табл. 4.5, нетрудно посчитать количество каналов, образующих групповой сигнал на каждом уровне иерархии. Например, для европейской системы количество каналов на разных уровнях будет равно:

- 1-ый уровень: 30;
- 2-ый уровень: $4 \times 30 = 120$;
- 3-ый уровень: $4 \times 120 = 480$;
- 4-ый уровень: $4 \times 480 = 1920$;
- 5-ый уровень: $4 \times 1920 = 7680$.

Таблица 4.5

Уровень иерархии	Европа		Северная Америка		Япония	
	Скорость, Мбит/с	Количество объединяемых каналов предыдущего уровня	Скорость, Мбит/с	Количество объединяемых каналов предыдущего уровня	Скорость, Мбит/с	Количество объединяемых каналов предыдущего уровня
0	0,064	–	0,064	–	0,064	–
1	2,048	30	1,544	24	1,544	24
2	8,448	4	6,312	4	6,312	4
3	34,368	4	44,736	7	32,064	5
4	139,264	4			97,728	3

Для цифровых потоков ПЦИ применяют соответствующие обозначения. Для североамериканской и японской ПЦИ применяется обозначение Т (иногда DS), для европейской ПЦИ – Е. Цифровые потоки первого уровня обозначаются соответственно Т-1 и Е-1, второго – Т-2 и Е-2 и т.д. К использованию в сетях связи РФ принята европейская ПЦИ.

Лекция 5. Сети ЭВМ

В настоящее время в мире существует свыше 100 миллионов компьютеров, большая часть из которых объединена в различные информационно-вычислительные сети от малых локальных сетей в офисах и домах до глобальных сетей типа Internet. Настоящий раздел посвящен рассмотрению принципов построения и основных характеристик локальных вычислительных сетей.

5.1. Назначение локальных вычислительных сетей

Локальная вычислительная сеть (ЛВС) – это сеть, объединяющая два или более компьютеров, с целью совместного использования программных и/или аппаратных ресурсов: принтеров, файлов, приложений, баз данных и т.п.

Понятие локальная вычислительная сеть – (англ. LAN – Local Area Network) относится к географически ограниченным (территориально или производственно) аппаратно-программным реализациям. В первых ЛВС объединялись от нескольких компьютеров до нескольких десятков компьютеров, расположенных на сравнительно небольшом удалении друг от друга: в одном офисе, на одном этаже здания, в одном здании.

Рассмотрим преимущества, получаемые при объединении персональных компьютеров в ЛВС.

Совместное использование аппаратных ресурсов. Совместное использование аппаратных ресурсов позволяет использовать периферийные устройства, например, лазерные принтеры, со всех компьютеров, входящих в состав данной ЛВС. При этом повышается загрузка устройств и отпадает необходимость их дублирования на каждом компьютере.

Совместное использование ресурсов процессора. Объединение компьютеров в ЛВС позволяет обеспечить доступ каждого из компьютеров к использованию вычислительных мощностей одного или нескольких более производительных компьютеров, входящих в состав этой же сети. Так же же, как и в предыдущем случае, это позволяет повысить загрузку и улучшить использование дорогостоящих устройств, а также исключить необходимость дублирования таких устройств на каждом рабочем месте.

Совместное использование данных. Совместное использование данных предоставляет возможность доступа и управления базами данных с периферийных рабочих мест, нуждающихся в информации.

Совместное использование программных средств. Объединение компьютеров в ЛВС позволяет использовать совместно не

только аппаратные ресурсы, но и программные средства. Это преимущество становится особенно ощутимым в случае, когда та или иная программа требует при работе повышенных ресурсов компьютера (оперативной памяти, памяти накопителя на жестком диске, быстродействия процессора и т.п.).

5.2. Классификация ЛВС

ЛВС принято делить на одноранговые сети и сети на основе выделенного сервера. В одноранговых сетях все пользователи компьютеров обладают одинаковыми правами. Право доступа других участников (пользователей) ЛВС к ресурсам того или иного компьютера определяет его владелец. Одноранговые сети способны эффективно функционировать, когда число компьютеров в сети не превышает нескольких десятков. При дальнейшем росте количества компьютеров в сети целесообразно использовать сети на основе выделенного сервера. Работа таких сетей строится на основе использования специального программного обеспечения – сервера, устанавливаемого на одном из компьютеров и предназначенного для обслуживания запросов, поступающих от других компьютеров сети – клиентов.

Руководит работой сети на основе выделенного сервера администратор, который регистрирует пользователей сети, назначает им права доступа к сетевым ресурсам, отвечает за проведение грамотной политики безопасности сети, а также решает многие другие вопросы, необходимые для нормального функционирования сети. Количество компьютеров в сети на основе выделенного сервера может составлять от нескольких десятков до нескольких сотен и больше.

На практике при решении вопроса о доступе к различным ресурсам ЛВС на основе выделенного сервера очень часто реализуется комбинированный подход, при котором общие сетевые ресурсы находятся в ведении администратора, а ресурсами конкретного компьютера, как и в случае одноранговых сетей, распоряжается пользователь этого компьютера.

5.3. Базовая модель взаимодействия открытых систем OSI (Open System Interconnection)

В общем случае передача сообщений в любых сетях реализуется на основе модели взаимодействия открытых систем (Open System Interconnection, OSI). Эта модель была разработана в 1984 г. Международной организацией по стандартизации (англ. ISO – International Standards Organization). В соответствии с принятой тер-

минологией открытыми системами называют любые устройства, принимающие участие в обмене сообщениями и реализованные на основе общепринятых международных стандартов. Соблюдение указанных стандартов является необходимым условием для того, чтобы эти устройства могли «понимать» друг друга.

Эталонная модель OSI делит общую проблему передачи сообщений между компьютерами через среду сети на семь менее крупных составных частей. Такая структуризация позволяет упростить решение каждой из составных частей и, следовательно, упростить до определенного уровня проблему передачи сообщений. Каждая из семи составных частей проблемы решается с помощью одного из уровней модели (рис. 5.1).

Прикладной уровень
Уровень представления
Сеансовый уровень
Транспортный уровень
Сетевой уровень
Канальный уровень
Физический уровень

Рис. 5.1. Модель OSI.

Два самых низших уровня модели OSI реализуются аппаратным и программным обеспечением; остальные пять высших уровней, как правило, реализуются программным обеспечением. Кроме того, первые 3 уровня являются сетезависимыми. То есть их программная и аппаратная (для уровней 1 и 2) реализация зависит от конкретных характеристик сети. Остальные 4 уровня являются сетезависимыми.

Модель OSI описывает, каким образом сообщение продвигается путь через физический канал сети от одной прикладной программы (например, программы моделирования радиоэлектронных устройств) до другой прикладной программы, находящейся в другом компьютере. Т.к. сообщение, которое должно быть отослано, проходит вниз через уровни системы, по мере этого продвижения оно утрачивает привычную для человека форму представления и все больше приобретает ту форму, которую понимают компьютеры, а точнее протоколы, определяющие правила обработки сообщений на передающей и приемной сторонах.

Уровень 1 расположен в основании модели и выполняет наиболее простые функции. Как правило, по мере увеличения номера уровня усложняются его функции.

После того, как стали понятными основные особенности принципа деления на уровни модели OSI, можно приступить к

обсуждению каждого отдельного уровня и его функций. Каждый уровень имеет заранее заданный набор функций, которые он должен выполнить для того, чтобы связь могла состояться.

Физический уровень. В функции физического уровня входит согласование таких характеристик, как типы используемых сигналов и скорость их передачи, уровни напряжений сигналов, типы разъемов и другие аналогичные характеристики.

Канальный уровень. Канальный уровень решает вопросы:

- формирования кадров, включающих блоки данных пользователя, которые требуется передать по сети, и необходимую для их передачи служебную информацию,
- проверки правильности передачи кадров,
- передачи подтверждений на правильно переданные кадры,
- отправки запросов на повторную передачу кадров, искаженных в процессе доставки,
- управления доступом к среде передачи данных и ряд других важных вопросов.

Сетевой уровень. Сетевой уровень обеспечивает возможность передачи сообщений между двумя конечными системами, подключенными к разным сетям. Т.к. две конечные системы, желающие организовать связь, может разделять значительное расстояние и множество сетей, протоколы маршрутизации, находящиеся на сетевом уровне, отвечают за выбор оптимального маршрута через последовательность соединенных между собой сетей.

Транспортный уровень. В протоколах сетевого уровня, отвечающих за доставку сообщений между конечными системами, обменивающимися сообщениями, не предусмотрено никаких механизмов, обеспечивающих надежную доставку. Эту функцию берут на себя протоколы транспортного уровня. Для обеспечения надежной доставки сообщений, транспортный уровень предусматривает механизмы для установления, поддержания и упорядоченного завершения действия виртуальных каналов, систем обнаружения и устранения неисправностей транспортировки и управления информационным потоком (с целью предотвращения переполнения системы данными из другой системы).

Сеансовый уровень. Как указывает его название, сеансовый уровень устанавливает, управляет и завершает сеансы взаимодействия между прикладными задачами. Сеансы состоят из диалога между двумя или более объектами представления. Сеансовый уровень синхронизирует диалог между объектами вышележащего уровня и управляет обменом информацией между ними.

Уровень представления. Данный уровень отвечает за то, чтобы сообщения, приходящие с прикладного уровня одной ко-

нечной системы (компьютера), были читаемыми для прикладного уровня другой конечной системы (компьютера) при условии, что они могут иметь различные операционные системы, аппаратную платформу и т.д. При необходимости протоколы данного уровня осуществляют трансляцию между множеством форматов представления сообщений путем использования общего метаязыка.

Прикладной уровень. Протоколы прикладного уровня отвечают за реализацию функций, относящихся к общесетевым приложениям: передачу файлов (FTP), передачу почтовых сообщений (POP-3, SMTP и др.), передачу сообщений в формате разметки гипертекста (HTTP) и другие.

5.4. Топологии вычислительной сети

Топология «Звезда» (рис. 5.2). Концепция топологии сети в виде звезды пришла из области больших ЭВМ, в которой головная машина получает и обрабатывает все данные с периферийных устройств как активный узел обработки данных. Вся информация между двумя периферийными рабочими местами проходит через центральный узел вычислительной сети – концентратор (hub).

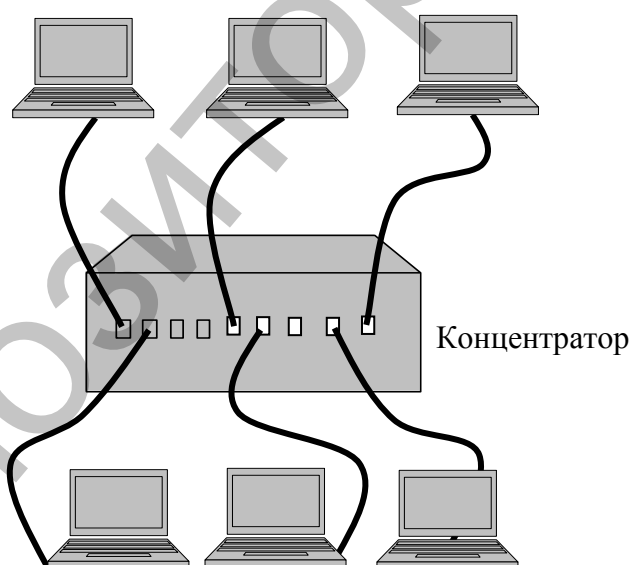


Рис. 5.2. Топология «звезда».

Пропускная способность сети определяется вычислительной мощностью узла. Кабельное соединение довольно простое, так как каждая рабочая станция связана с узлом. Затраты на прокладку кабелей высокие, особенно когда центральный узел географически расположен не в центре топологии.

При расширении вычислительных сетей не могут быть использованы ранее выполненные кабельные связи: к новому рабо-

чему месту необходимо прокладывать отдельный кабель из центра сети.

Производительность вычислительной сети в первую очередь зависит от характеристик концентратора. Он может быть узким местом вычислительной сети. В случае выхода из строя центрального узла нарушается работа всей сети.

Топология «Кольцо» (рис. 5.3). При кольцевой топологии сети рабочие станции связаны одна с другой последовательно, т.е. рабочая станция 1 с рабочей станцией 2, рабочая станция 3 с рабочей станцией 4 и т.д. Последняя рабочая станция связана с первой. Маршрут, по которому осуществляется передача сообщений, замыкается в кольцо.

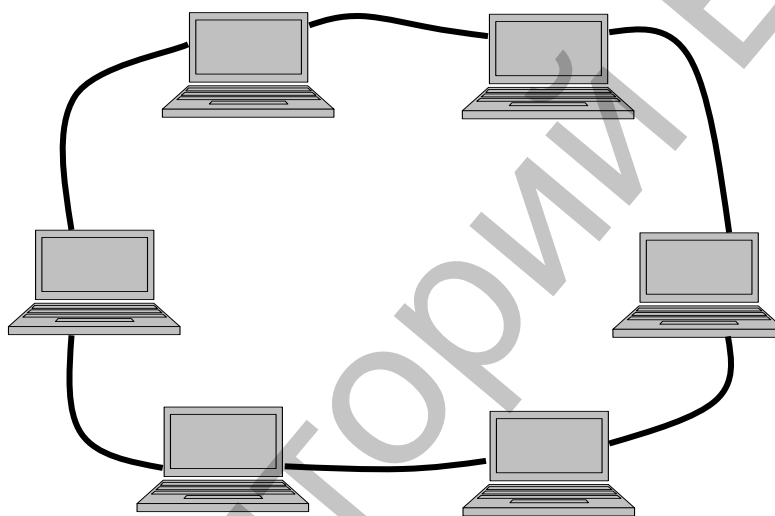


Рис. 5.3. Кольцевая топология.

Сообщения циркулируют регулярно по кругу. Основная проблема при кольцевой топологии заключается в том, что каждая рабочая станция должна активно участвовать в пересылке информации, и в случае выхода из строя хотя бы одной из них, если не предпринять специальных мер, вся сеть может оказаться неработоспособной.

Топология «Общая шина» (рис. 5.4). При шинной топологии среда передачи информации представляется в форме коммуникационного пути, доступного для всех рабочих станций, к которому они все должны быть подключены. Все рабочие станции могут непосредственно вступать в контакт с любой рабочей станцией, имеющейся в сети.

Рабочие станции в любое время, без прерывания работы всей вычислительной сети, могут быть подключены к ней или отключены. Функционирование вычислительной сети не зависит от состояния отдельной рабочей станции.

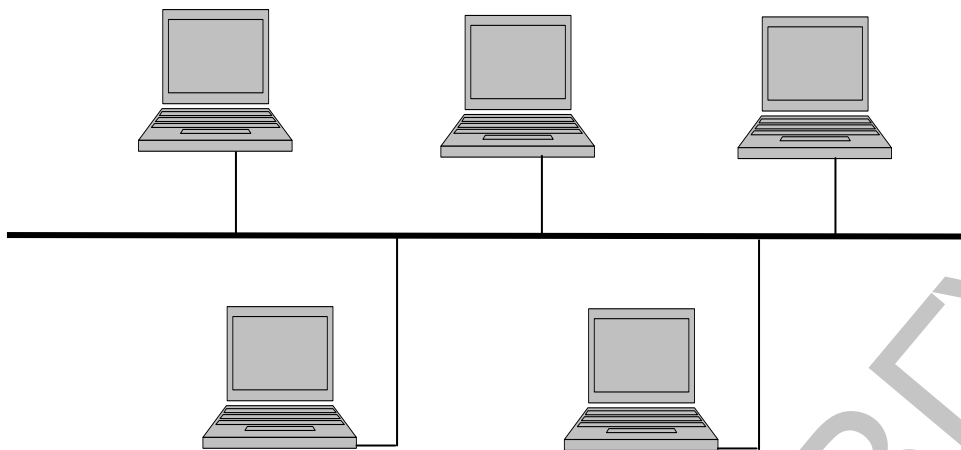


Рис. 5.4. Топология «Общая шина».

Иерархическая топология (рис. 5.5). Наряду с известными топологиями вычислительных сетей: «кольцо», «звезда» и «шина», на практике применяется и комбинированная, например, иерархическая (древовидная) топология. Вычислительные сети с древовидной структурой применяются там, где непосредственное применение базовых сетевых структур («кольцо», «звезда», «шина»), в чистом виде реализовать затруднительно. Для подключения большого числа рабочих станций применяют концентраторы. Концентраторы могут быть пассивными или активными. Пассивный концентратор обычно используют как разветвитель. Он не осуществляет усиления сигнала. При этом сигнал, поступивший на один из портов (входов/выходов), передается одновременно на все остальные порты. Активный концентратор в отличие от пассивного осуществляет восстановление первоначального уровня сигнала и его формы. Эта процедура носит название регенерации сигнала.

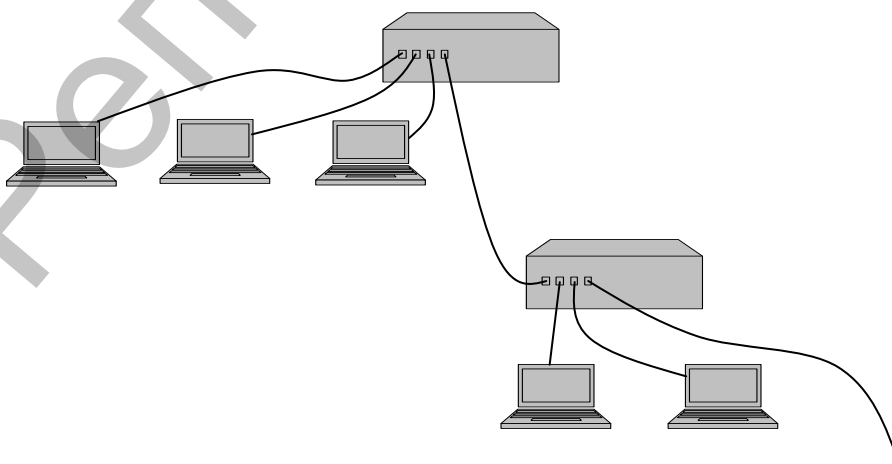


Рис. 5.5. Иерархическая топология.

В зависимости от фирмы производителя и модели концентратора, его вид может изменяться. Подключение к портам концентратора осуществляется с помощью 8-контактного разъема RJ-45, используемого с кабелем «витая пара».

Для того, чтобы компьютер мог передавать и принимать сообщения из сети, он должен иметь специальную плату – сетевой адаптер (синоним – сетевая карта). Эта плата устанавливается в слот расширения материнской платы компьютера.

Если в качестве среды передачи ЛВС используется коаксиальный кабель, то для его подключения к сетевой карте применяются тройниковые соединители (T-connectors).

5.5. ЛВС стандарта Ethernet

Сети стандарта Ethernet используют топологию «общая шина». В сетях Ethernet используется метод доступа к среде передачи данных, называемый *методом коллективного доступа с опознаванием несущей и обнаружением коллизий (carrier-sense-multiple-access with collision detection, CSMA/CD)*. Этот метод используется исключительно в сетях с общей шиной (к которым относятся и радиосети, породившие этот метод). Все компьютеры такой сети имеют непосредственный доступ к общей шине, поэтому она может быть использована для передачи данных между любыми двумя узлами сети. Простота схемы подключения – это один из факторов, определивших успех стандарта Ethernet. Говорят, что кабель, к которому подключены все станции, работает в режиме *коллективного доступа (multiple-access, MA)*.

Все данные, передаваемые по сети, помещаются в кадры определенной структуры и снабжаются уникальным адресом станции назначения. Затем кадр передается по кабелю. Все станции, подключенные к кабелю, могут распознать факт передачи кадра, и та станция, которая узнает собственный адрес в заголовках кадра, записывает его содержимое в свой внутренний буфер для дальнейшей обработки. При описанном подходе возможна ситуация, когда две или более станции одновременно пытаются передать кадр данных по общему кабелю (рис. 5.6). Для уменьшения вероятности этой ситуации непосредственно перед отправкой кадра передающая станция слушает кабель (то есть принимает и анализирует распространяющиеся в нем электрические сигналы), чтобы обнаружить, не передается ли уже по кабелю кадр данных от другой станции. Если *опознается несущая (carrier-sense, CS)*, то станция откладывает передачу своего кадра до окончания чужой передачи, и только потом пытается вновь его

передать. Но даже при таком алгоритме две станции практически одновременно (с учетом конечного времени распространения сигнала в кабеле) могут решить, что по шине в данный момент времени нет передачи, и начать одновременно передавать свои кадры. Говорят, что при этом происходит *коллизия*, так как сигналы обоих кадров накладываются друг на друга в общей кабеле, что приводит к их искажению. При этом ни один из сигналов, участвующих в коллизии, не будет принят правильно.

Чтобы корректно обработать коллизию, все станции одновременно наблюдают за возникающими на кабеле сигналами. Если передаваемые и наблюдаемые сигналы отличаются, то фиксируется *обнаружение коллизии (collision detection, CD)*. Для увеличения вероятности немедленного обнаружения коллизии всеми станциями сети, ситуация коллизии усиливается посылкой в сеть станциями, начавшими передачу своих кадров, специальной последовательности битов, называемой *jam-последовательностью*.

После обнаружения коллизии передающая станция обязана прекратить передачу и ожидать в течение короткого случайного интервала времени, а затем может снова сделать попытку передачи кадра.

Из описания метода доступа видно, что он носит случайный характер, и вероятность успешного получения в свое распоряжение общей среды зависит от загруженности сети, то есть от интенсивности возникновения в станциях потребности передачи кадров. При разработке этого метода предполагалось, что скорость передачи данных в 10 Мбит/с очень высока по сравнению с потребностями компьютеров во взаимном обмене данными, поэтому загрузка сети будет всегда небольшой. Это предположение остается часто справедливым и по сей день, однако уже появились приложения, работающие в реальном масштабе времени с мультимедийной информацией, для которых требуются гораздо более высокие скорости передачи данных. Поэтому наряду с классическим Ethernet'ом растет потребность и в новых высокоскоростных технологиях.

Метод CSMA/CD определяет основные временные и логические соотношения, гарантирующие корректную работу всех станций в сети.

Между двумя последовательно передаваемыми по общей шине кадрами информации должна выдерживаться пауза в 9.6 мкс; эта пауза нужна для приведения в исходное состояние сетевых адаптеров узлов, а также для предотвращения монопольного захвата среды передачи данных одной станцией.

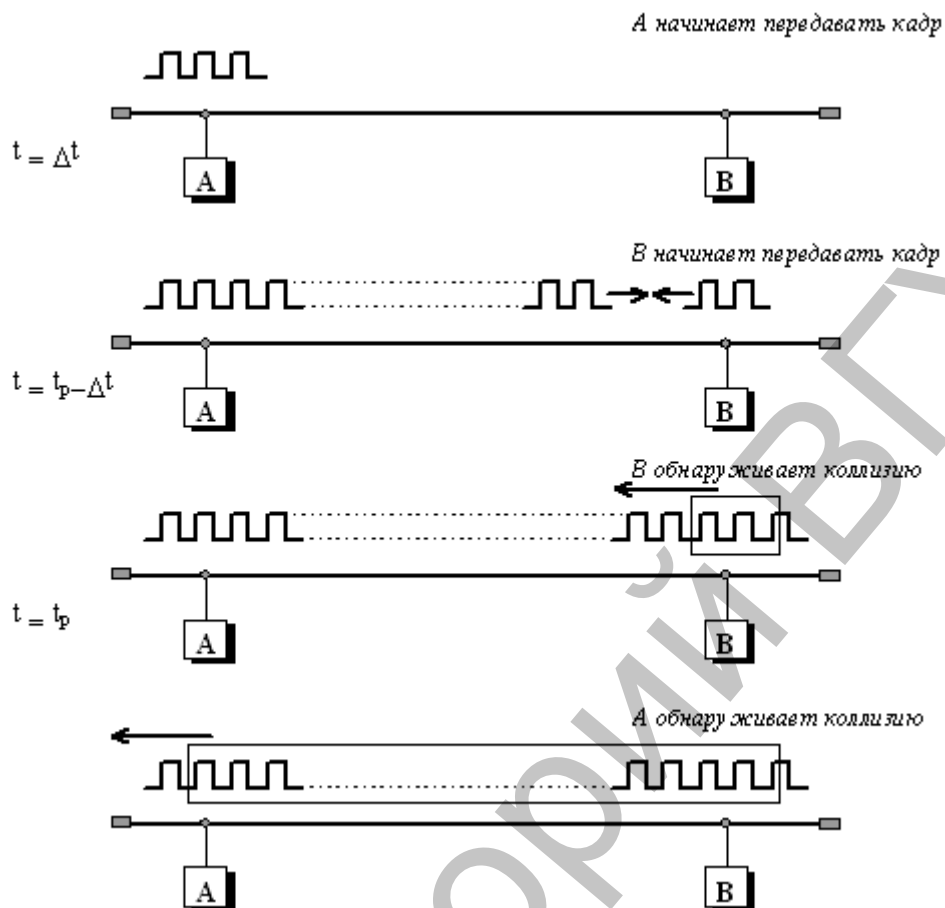


Рис. 5.6. Схема возникновения коллизии в методе случайного доступа CSMA/CD (t_p – задержка распространения сигнала между станциями A и B).

При обнаружении коллизии (условия ее обнаружения зависят от применяемой физической среды) станция выдает в среду специальную 32-х битную последовательность (jam-последовательность), усиливающую явление коллизии для более надежного распознавания ее всеми узлами сети.

После обнаружения коллизии каждый узел, который передавал кадр и столкнулся с коллизией, после некоторой задержки пытается повторно передать свой кадр. Узел делает максимально 16 попыток передачи этого кадра информации, после чего отказывается от его передачи. Величина задержки выбирается как равномерно распределенное случайное число из интервала, длина которого экспоненциально увеличивается с каждой попыткой. Такой алгоритм выбора величины задержки снижает вероятность коллизий и уменьшает интенсивность выдачи кадров в сеть при ее высокой загрузке.

Четкое распознавание коллизий всеми станциями сети является необходимым условием корректной работы сети Ethernet.

Если какая-либо передающая станция не распознает коллизию и решит, что кадр данных ею передан верно, то этот кадр данных будет утерян, так как информация кадра исказится из-за наложения сигналов при коллизии, он будет отбракован принимающей станцией (скорее всего из-за несовпадения контрольной суммы). Конечно, скорее всего, искаженная информация будет повторно передана каким-либо протоколом верхнего уровня, например, транспортным или прикладным, работающим с установлением соединения и нумерацией своих сообщений. Но повторная передача сообщения протоколами верхних уровней произойдет через гораздо более длительный интервал времени (десятки секунд) по сравнению с микросекундными интервалами, которыми оперирует протокол Ethernet. Поэтому, если коллизии не будут надежно распознаваться узлами сети Ethernet, то это приведет к заметному снижению полезной пропускной способности данной сети.

Все параметры протокола Ethernet подобраны таким образом, чтобы при нормальной работе узлов сети коллизии всегда четко распознавались. Именно для этого минимальная длина поля данных кадра должна быть не менее 46 байт (что вместе со служебными полями дает минимальную длину кадра в 72 байта или 576 бит). Длина кабельной системы выбирается таким образом, чтобы за время передачи кадра минимальной длины сигнал коллизии успел бы распространиться до самого дальнего узла сети. Поэтому для скорости передачи данных 10 Мбит/с, используемой в стандартах Ethernet, максимальное расстояние между двумя любыми узлами сети не должно превышать 2500 метров.

С увеличением скорости передачи кадров, что имеет место в новых стандартах, базирующихся на том же методе доступа CSMA/CD, например, Fast Ethernet, в котором передача сообщений осуществляется со скоростью 100 Мбит/с, максимальная длина сети уменьшается пропорционально увеличению скорости передачи. В стандарте Fast Ethernet она составляет 210 м, а в «гигабитном» Ethernet (1000 Мбит/с) ограничена 25 метрами.

Независимо от реализации физической среды, в классическом Ethernet (10 Мбит/с) должны быть выполнены два ограничения, связанные с методом доступа:

максимальное расстояние между двумя любыми узлами не должно превышать 2500 м,

в сети не должно быть более 1024 узлов.

Кроме того, каждый вариант физической среды добавляет к этим ограничениям свои ограничения, которые также должны выполняться.

Значения основных параметров процедуры передачи кадра стандарта «Ethernet» приведено в табл. 5.1.

Таблица 5.1

Скорость передачи информации	10 Мб/с
Межкадровый интервал	9.6 мкс
Максимальное число попыток передачи	16
Максимальная кратность возрастания диапазона паузы	10
Длина jam-последовательности	32 бита
Максимальная длина кадра (без преамбулы)	1518 байтов
Минимальная длина кадра (без преамбулы)	64 байта (512 бит)
Длина преамбулы	64 бита

Стандарты технологии Ethernet.

Стандарт 10Base-5. Стандарт 10Base-5 соответствует экспериментальной сети Ethernet фирмы Хегох и может считаться классическим Ethernet'ом. Он использует в качестве среды передачи данных коаксиальный кабель с диаметром центрального медного провода 2,17 мм и внешним диаметром около 10 мм (толстый Ethernet).

Кабель используется как моноканал для всех станций. Сегмент кабеля имеет максимальную длину 500 м (без повторителей) и должен иметь на концах согласующие терминаторы сопротивлением 50 Ом, поглощающие распространяющиеся по кабелю сигналы и препятствующие возникновению отраженных сигналов.

Различные компоненты сети, выполненной на толстом коаксиале, показаны на рис. 5.7.

Станция должна подключаться к кабелю при помощи приемопередатчика – трансивера. Трансивер устанавливается непосредственно на кабеле и питается от сетевого адаптера компьютера.

Трансивер соединяется с сетевым адаптером интерфейсным кабелем AUI (Attachment Unit Interface) длиной до 50 м, состоящим из 4 витых пар (адаптер должен иметь разъем AUI). Допускается подключение к одному сегменту не более 100 трансиверов, причем расстояние между подключениями трансиверов не должно быть меньше 2,5 м.

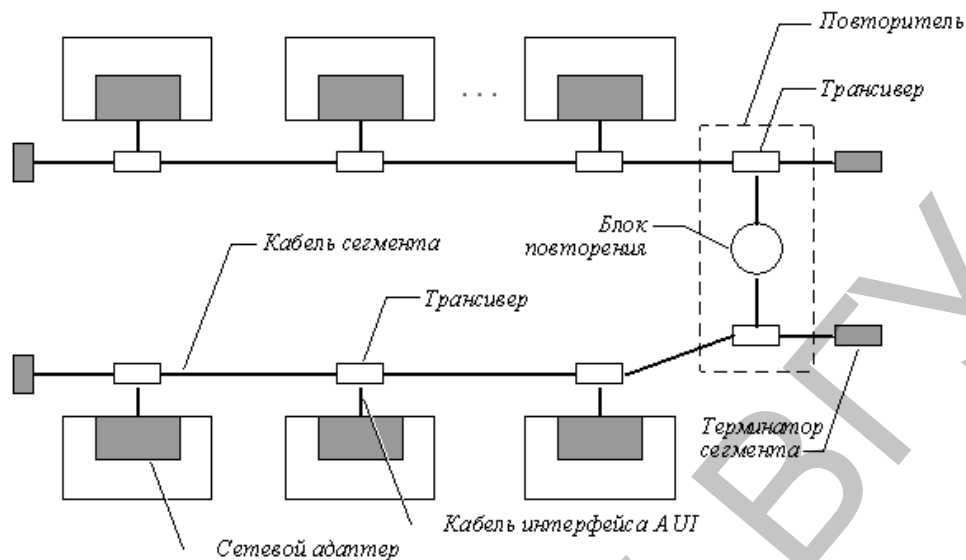


Рис. 5.7. Компоненты физического уровня сети стандарта 10Base-5, состоящей из двух сегментов.

Трансивер – это часть сетевого адаптера, которая выполняет следующие функции:

- прием и передача данных с кабеля на кабель,
- определение коллизий на кабеле,
- электрическая развязка между кабелем и остальной частью адаптера,
- защита кабеля от некорректной работы адаптера.

Последнюю функцию часто называют *контролем болтливости (jabber control)*. При возникновении неисправностей в адаптере может возникнуть ситуация, когда на кабель будет непрерывно выдаваться последовательность случайных сигналов. Так как кабель – это общая среда для всех станций, то работа сети будет заблокирована одним неисправным адаптером. Чтобы этого не случилось, на выходе передатчика ставится схема, которая проверяет количество битов, переданных в пакете. Если максимальная длина пакета превышает, то эта схема просто отсоединяет выход передатчика от кабеля.

Упрощенная структурная схема трансивера показана на рис. 5.8. Детектор коллизий определяет наличие коллизии в коаксиальном кабеле по повышенному уровню постоянной составляющей сигналов. Если постоянная составляющая превышает определенный порог, то значит, на кабель работает более чем один передатчик.

К *достоинствам* стандарта 10Base-5 относятся:

- хорошая защищенность кабеля от внешних воздействий,
- сравнительно большое расстояние между узлами,

- возможность простого перемещения рабочей станции в пределах длины кабеля AUI.

К недостаткам следует отнести:

- высокую стоимость кабеля,
- сложность его прокладки из-за большой жесткости,
- наличие специального инструмента для заделки кабеля,
- при повреждении кабеля или плохом соединении происходит останов работы всей сети,
- необходимо заранее предусмотреть подводку кабеля ко всем возможным местам установки компьютеров.

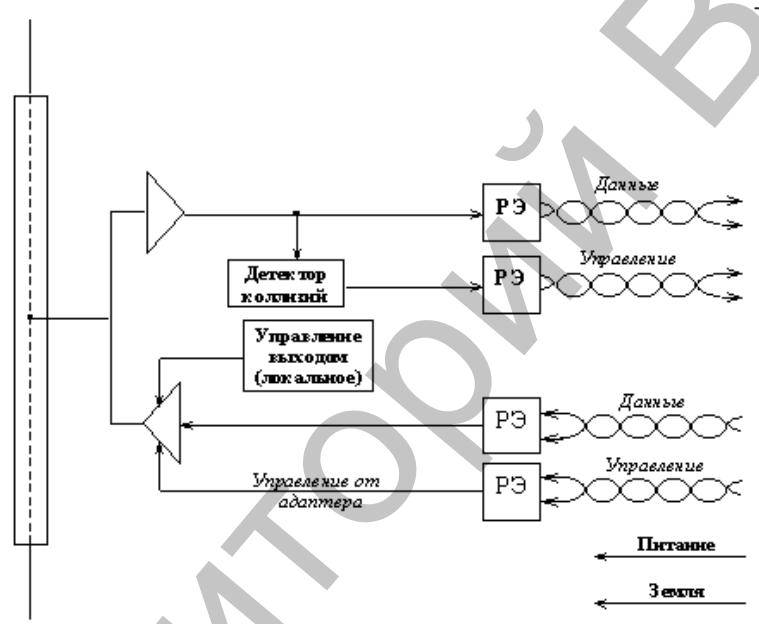


Рис. 5.8. Структурная схема трансивера.

Стандарт 10Base-2. Стандарт 10Base-2 использует в качестве передающей среды *коаксиальный кабель* с диаметром центрального медного провода 0,89 мм и внешним диаметром около 5 мм (тонкий Ethernet, волновое сопротивление кабеля 50 Ом). Максимальная длина сегмента без повторителей составляет 185 м, сегмент должен иметь на концах согласующие терминаторы 50 Ом.

Станции подключаются к кабелю с помощью T-коннектора, который представляет собой тройник, один отвод которого соединяется с сетевым адаптером, а два других – с двумя концами разрыва кабеля. Максимальное количество станций, подключаемых к одному сегменту, – 30. Минимальное расстояние между станциями – 1 м.

Этот стандарт очень близок к стандарту 10Base-5. Но трансиверы в нем объединены с сетевыми адаптерами за счет того, что более гибкий тонкий коаксиальный кабель может быть под-

веден непосредственно к выходному разъему платы сетевого адаптера, установленной в шасси компьютера. Кабель в данном случае «висит» на сетевом адаптере, что затрудняет физическое перемещение компьютеров.

Топология сегмента сети стандарта 10Base-2 показана на рис. 5.9.

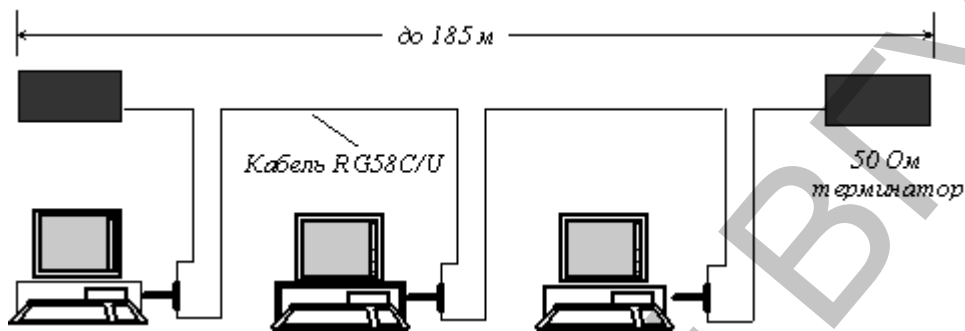


Рис. 5.9. Сеть стандарта 10Base-2.

Реализация этого стандарта на практике приводит к наиболее простому решению для кабельной сети, так как для соединения компьютеров требуются только сетевые адаптеры и Т-коннекторы. Однако этот вид кабельных соединений наиболее сильно подвержен авариям и сбоям: кабель восприимчив к помехам, в моноканале имеется большое количество механических соединений (каждый Т-коннектор дает три механических соединения, два из которых имеют жизненно важное значение для всей сети), пользователи имеют доступ к разъемам и могут нарушить целостность моноканала. Кроме того, эстетика и эргономичность этого решения оставляют желать лучшего, так как от каждой станции через Т-коннектор отходят два довольно заметных провода, которые под столом часто образуют моток кабеля – запас, необходимый на случай даже небольшого перемещения рабочего места.

Общим недостатком стандартов 10Base-5 и 10Base-2 является отсутствие оперативной информации о состоянии моноканала. Повреждение кабеля обнаруживается сразу же (сеть перестает работать), но для поиска отказавшего отрезка кабеля необходим специальный прибор – кабельный тестер.

Стандарт 10Base-T. Стандарт принят в 1991 году как дополнение к существующему набору стандартов Ethernet и имеет обозначение 802.3i. Данный стандарт предусматривает использование в качестве среды двойной неэкранированной витой пары (Unshielded Twisted Pair, UTP). Соединения станций осуществляются по топологии «точка–точка» со специальным устройством – многопортовым повторителем с помощью двух витых пар. Одна витая пара используется для передачи данных от станции к по-

вторителю (выход Tx сетевого адаптера), а другая – для передачи данных от повторителя станции (вход Rx сетевого адаптера). На рис. 5.10 показан пример трехпортового повторителя.

Многопортовые повторители в данном случае обычно называются концентраторами (англоязычные термины – hub или concentrator). Концентратор осуществляет функции повторителя сигналов на всех отрезках витых пар, подключенных к его портам, так что образуется единая среда передачи данных – моноканал (шина). Повторитель обнаруживает коллизии в сегменте в случае одновременной передачи сигналов по нескольким своим Rx входам и посылает jam-последовательность на все свои Tx выходы. Стандарт определяет битовую скорость передачи данных 10 Мб/с и максимальное расстояние отрезка витой пары между двумя непосредственно связанными узлами (станциями и концентраторами) не более 100 м при использовании витой пары качества не ниже категории 3.

Возможно иерархическое соединение концентраторов в дерево. Для обеспечения синхронизации станций при реализации процедур доступа CSMA/CD и надежного распознавания станциями коллизий в стандарте определено максимально число концентраторов между любыми двумя станциями сети. Общее количество станций в сети 10Base-T не должно превышать 1024.

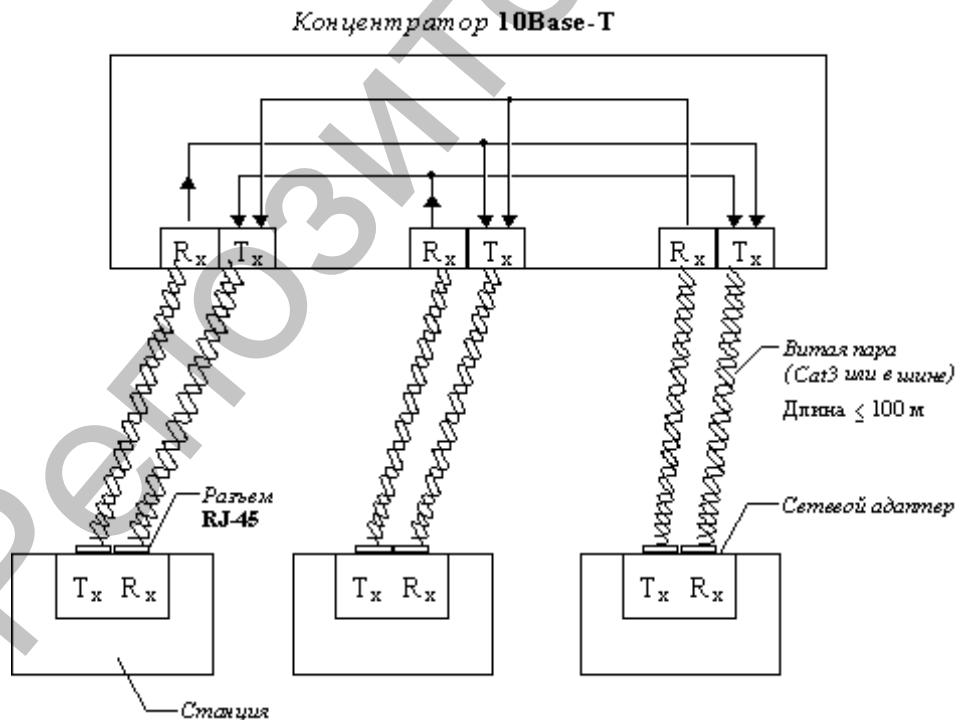


Рис. 5.10. Сеть 10Base-T – один домен коллизий
Tx – передатчик, Rx – приемник.

Сети, построенные на основе стандарта 10Base-T, обладают по сравнению с коаксиальными вариантами Ethernet'a многими преимуществами. Эти преимущества связаны с разделением общего физического кабеля на отдельные кабельные отрезки, подключенные к центральному коммуникационному устройству. И хотя логически эти отрезки по-прежнему образуют общий домен коллизий, их физическое разделение позволяет контролировать их состояние и отключать в случае обрыва, короткого замыкания или неисправности сетевого адаптера на индивидуальной основе. Это обстоятельство существенно облегчает эксплуатацию больших сетей Ethernet, так как концентратор обычно автоматически выполняет такие функции, уведомляя при этом администратора сети о возникшей проблеме.

Стандарт 10Base-F. Стандарт 10Base-F использует в качестве среды передачи данных оптоволокно. Функционально сеть стандарта 10Base-F состоит из тех же элементов, что и сеть стандарта 10Base-T, – сетевых адаптеров, многопортового повторителя и отрезков кабеля, соединяющих адаптер с портом повторителя. Как и при использовании витой пары, для соединения адаптера с повторителем используется два оптоволоконна – одно соединяет выход Tx адаптера со входом Rx повторителя, а другое – вход Rx адаптера с выходом Tx повторителя.

Стандарт FOIRL (Fiber Optic Inter-Repeater Link) – это первый стандарт комитета 802.3 для использования оптоволоконна в сетях Ethernet. Он гарантирует длину оптоволоконной связи между повторителями до 1 км при общей длине сети не более 2500 м. Максимальное число повторителей – 4.

Стандарт 10Base-FL предназначен для соединения конечных узлов с концентратором и работает с сегментами оптоволоконна длиной не более 2000 м при общей длине сети не более 2500 м. Максимальное число повторителей – 4.

Стандарт 10Base-FB предназначен для магистрального соединения повторителей. Он позволяет иметь в сети до 5 повторителей при максимальной длине одного сегмента 2000 м и максимальной длине сети 2740 м. Повторители, соединенные по стандарту 10Base-FB, постоянно обмениваются специальными последовательностями сигналов, отличающимися от сигналов кадров данных, для обнаружения отказов своих портов. Поэтому концентраторы стандарта 10Base-FB могут поддерживать резервные связи, переходя на резервный порт при обнаружении отказа основного с помощью тестовых специальных сигналов. Концентраторы этого стандарта передают как данные, так и сигналы проста линии синхронно, поэтому биты синхронизации кадра не

нужны и не передаются. Стандарт 10Base-FB поэтому называют также *синхронный Ethernet*.

Стандарты 10Base-FL и 10Base-FB не совместимы между собой.

Правило 4-х повторителей. При описании топологии сети стандарта 10Base-5 приводились ограничения на длину одного непрерывного отрезка коаксиального кабеля, используемого в качестве общей шины передачи данных для всех станций сети. Отрезок кабеля, завершающийся на обоих концах терминаторами и имеющий общую длину не более 500 м, называется физическим сегментом сети. Однако при расчете окна коллизий общая максимальная длина сети 10Base-5 считалась равной 2500 м. Противоречия здесь нет, так как стандарт 10Base-5 (впрочем, как и остальные стандарты физического уровня Ethernet) допускает соединение нескольких сегментов коаксиального кабеля с помощью повторителей, которые обеспечивают увеличение общей длины сети.

Повторитель соединяет два сегмента коаксиального кабеля и выполняет функции регенерации электрической формы сигналов и их синхронизации (*retiming*). Повторитель прозрачен для станций, он обязан передавать кадры без искажений, модификации, потери или дублирования. Имеются ограничения на максимально допустимые величины дополнительных задержек распространения битов нормального кадра через повторитель, а также битов jam-последовательности, которую повторитель обязан передать на все подключенные к нему сегменты при обнаружении коллизии на одном из них. Воспроизведение коллизии на всех подключенных к повторителю сегментах – одна из его основных функций. Говорят, что сегменты, соединенные повторителями, образуют один *домен коллизий (collision domain)*.

Повторитель состоит из трансиверов, подключаемых к коаксиальным сегментам, а также блока повторения, выполняющего основные функции повторителя.

На рис. 5.7 показан пример сети, состоящей из двух сегментов, соединенных одним повторителем.

В общем случае стандарт 10Base-5 допускает использование *до 4-х повторителей*, соединяющих в этом случае *5 сегментов* длиной до 500 метров каждый, если используемые повторители удовлетворяют ограничениям на допустимые величины задержек сигналов. При этом общая длина сети будет составлять 2500 м, и такая конфигурация гарантирует правильное обнаружение коллизии крайними станциями сети. Только 3 сегмента из 5 могут быть нагруженными, то есть сегментами с подключенными к ним трансиверами конечных станций.

Правила 4-х повторителей и максимальной длины каждого из сегментов легко использовать на практике для определения корректности конфигурации сети. Однако эти правила применимы только тогда, когда все соединяемые сегменты представляют собой одну физическую среду, то есть в нашем случае толстый коаксиальный кабель, а все повторители также удовлетворяют требованиям физического стандарта 10Base-5. Аналогичные простые правила существуют и для сетей, все сегменты которых удовлетворяют требованиям другого физического стандарта, например, 10Base-T или 10Base-F. Однако для смешанных случаев, когда в одной сети Ethernet присутствуют сегменты различных физических стандартов, правила, основанные только на количестве повторителей и максимальных длинных сегментов, становятся более запутанными. Поэтому более надежно рассчитывать время полного оборота сигнала по смешанной сети с учетом задержек в каждом типе сегментов и в каждом типе повторителей и сравнивать его с максимально допустимым временем, которое для любых сетей Ethernet с битовой скоростью 10 Мб/с не должно превышать 575 битовых интервалов (количество битовых интервалов в пакете минимальной длины с учетом преамбулы).

Формат кадра Ethernet. Форматы нескольких вариантов технологии Ethernet показаны на рис. 5.11.

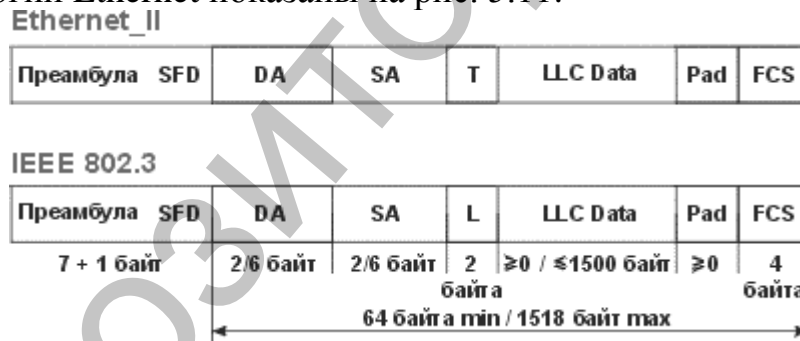


Рис. 5.11. Формат кадра Ethernet.

Поля имеют следующие назначения:

Преамбула: 7 байт, каждый из которых представляет чередование единиц и нулей 10101010. Преамбула позволяет установить битовую (тактовую) синхронизацию на приемной стороне.

Ограничитель начала кадра (SFD, start frame delimiter): 1 байт, последовательность 10101011, указывает, что далее последуют информационные поля кадра. Этот байт можно отнести к преамбуле.

Адрес назначения (DA, destination address): 6 байт, указывает адрес станции, для которой (которых) предназначен этот кадр. Это может быть единственный физический адрес (unicast), групповой адрес (multicast) или широковещательный адрес (broadcast).

Адрес отправителя (SA, source address): 6 байт, указывает адрес станции, которая посылает кадр.

Поле типа или длины кадра (T or L, type or length): 2 байта. Существуют два базовых формата кадра Ethernet (в английской терминологии raw formats – сырые форматы) – Ethernet_II и IEEE 802.3 (рис. 5.11), причем различное назначение у них имеет именно рассматриваемое поле. Для кадра Ethernet_II в этом поле содержится информация о типе кадра. Для кадра IEEE 802.3 в этом поле содержится выраженный в байтах размер следующего поля – поля данных (LLC Data). Если эта цифра приводит к общей длине кадра меньше 64 байт, то за полем LLC Data добавляется поле Pad (заполнение). Для протокола более высокого уровня не возникает путаницы с определением типа кадра, так как для кадра IEEE 802.3 значение этого поля не может быть больше 1500 (0x05DC). Поэтому в одной сети могут свободно сосуществовать оба формата кадров, более того, один сетевой адаптер может взаимодействовать с обоими типами посредством стека протоколов.

Данные (LLC Data): поле данных.

Дополнительное поле (Pad) – заполняется только в том случае, когда поле данных невелико, с целью увеличения размера поля данных до 46 байт. Ограничение снизу на минимальную длину кадра необходимо для правильного разрешения коллизий.

Контрольная последовательность кадра (FCS, frame check sequence): 4-байтовое поле, в котором указывается контрольная сумма, вычисленная с использованием циклического избыточного кода по полям кадра за исключением преамбулы, SDF и FCS.

5.6. ЛВС стандарта Token Ring

Сети стандарта Token Ring, так же, как и сети Ethernet, используют разделяемую среду передачи данных, которая состоит из отрезков кабеля, соединяющих все станции сети в кольцо. Кольцо рассматривается как общий разделяемый ресурс, и для доступа к нему используется не случайный алгоритм, как в сетях Ethernet, а детерминированный, основанный на передаче станциями права на использование кольца в определенном порядке. Право на использование кольца передается с помощью кадра специального формата, называемого маркером (token).

Стандарт Token Ring был принят комитетом 802.5 в 1985 году. В это же время компания IBM приняла стандарт Token Ring в качестве своей основной сетевой технологии. В настоящее время именно компания IBM является основным законодателем мо-

ды технологии Token Ring, производя около 60% сетевых адаптеров этой технологии.

Сети Token Ring работают с двумя значениями скоростями – 4 Мб/с и 16 Мб/с. Первая скорость определена в стандарте 802.5, а вторая является новым стандартом де-факто, появившимся в результате развития технологии Token Ring. Смещение станций, работающих на различных скоростях, в одном кольце не допускается.

Сети Token Ring, работающие со скоростью 16 Мб/с, имеют и некоторые усовершенствования в алгоритме доступа по сравнению со стандартом 4 Мб/с.

Маркерный метод доступа к среде передачи данных. В сетях с *маркерным методом* доступа право на доступ к среде передается циклически от станции к станции по логическому кольцу. Кольцо образуется отрезками кабеля, соединяющими соседние станции. Таким образом, каждая станция связана со своей предшествующей и последующей станцией и может непосредственно обмениваться данными только с ними. Для обеспечения доступа станций к физической среде по кольцу циркулирует кадр специального формата и назначения – *маркер* (англ. – *token*).

Получив маркер, станция анализирует его, при необходимости модифицирует и при отсутствии у нее данных для передачи обеспечивает его продвижение к следующей станции. Станция, которая имеет данные для передачи, при получении маркера изымает его из кольца, что дает ей право доступа к физической среде и передачи своих данных. Затем эта станция выдает в кольцо кадр данных установленного формата последовательно по битам. Переданные данные проходят по кольцу всегда в одном направлении от одной станции к другой.

При поступлении кадра данных к одной или нескольким станциям эти станции копируют для себя этот кадр и вставляют в этот кадр подтверждение приема. Станция, выдавшая кадр данных в кольцо, при обратном его получении с подтверждением приема изымает этот кадр из кольца и выдает новый маркер для обеспечения возможности другим станциям сети передавать данные.

На рис. 5.12 описанный алгоритм доступа к среде иллюстрируется временной диаграммой. Здесь показана передача пакета А в кольце, состоящем из 6 станций, от станции 1 к станции 3.

Время удержания одной станцией маркера ограничивается *тайм-аутом удержания маркера*, по истечении которого станция обязана передать маркер далее по кольцу.

В сетях Token Ring 16 Мб/с используется также несколько другой алгоритм доступа к кольцу, называемый алгоритмом *раннего освобождения маркера* (*Early Token Release*). В соответст-

вии с ним станция передает маркер следующей станции сразу же после окончания передачи последнего бита кадра, не дожидаясь возвращения по кольцу этого кадра с битом подтверждения приема. В этом случае пропускная способность кольца используется более эффективно и приближается к 80% от номинальной.

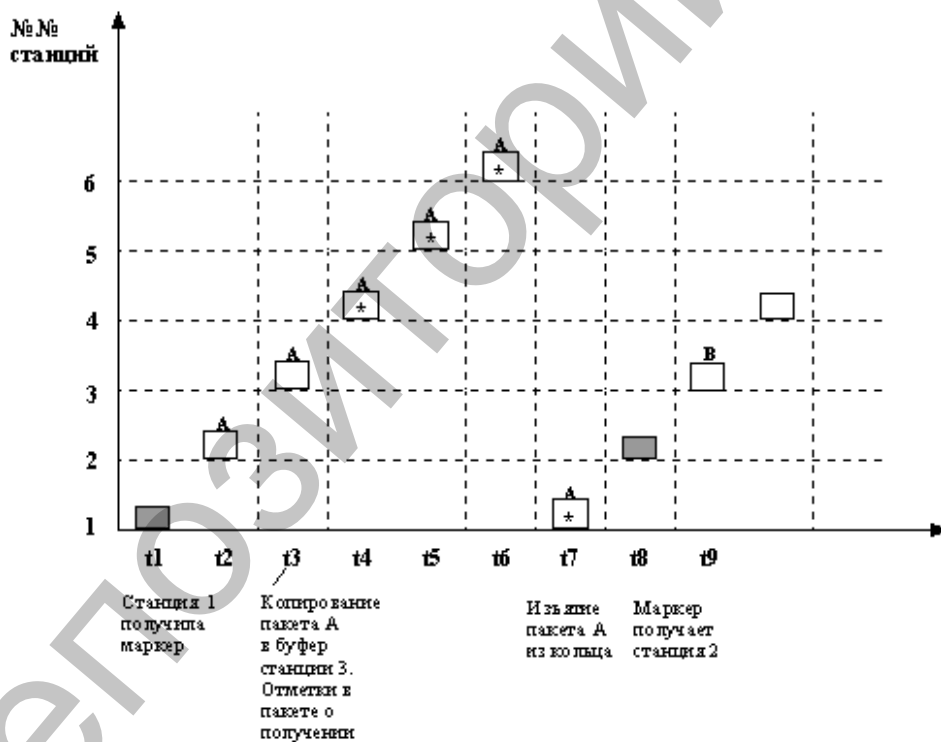
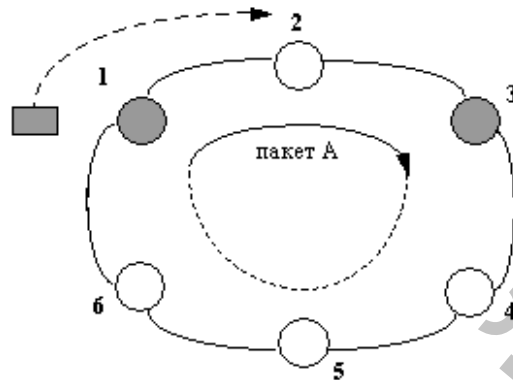


Рис. 5.12. Принцип маркерного доступа.

Для различных видов сообщений передаваемым данным могут назначаться различные *приоритеты*.

Каждая станция имеет механизмы обнаружения и устранения неисправностей сети, возникающих в результате ошибок передачи или переходных явлений (например, при подключении и отключении станции).

Не все станции в кольце равны. Одна из станций обозначается как *активный монитор*, что означает дополнительную ответственность по управлению кольцом. Активный монитор осуществляет управление тайм-аутом в кольце, порождает новые маркеры (если необходимо), чтобы сохранить рабочее состояние, и генерирует диагностические кадры при определенных обстоятельствах. Активный монитор выбирается, когда кольцо инициализируется, и в этом качестве может выступить любая станция сети. Если монитор отказал по какой-либо причине, существует механизм, с помощью которого другие станции (резервные мониторы) могут договориться, какая из них будет новым активным монитором.

Форматы кадров Token Ring. В Token Ring существует три различных формата кадров:

- маркер;
- кадр данных;
- прерывающая последовательность.

Маркер. Кадр маркера состоит из трех полей, каждое длиной в один байт.

Поле начального ограничителя появляется в начале маркера, а также в начале любого кадра, проходящего по сети. Поле состоит из уникальной серии электрических импульсов, которые отличаются от тех импульсов, которыми кодируются единицы и нули в байтах данных. Поэтому начальный ограничитель нельзя спутать ни с какой битовой последовательностью.

Поле контроля доступа. Разделяется на четыре элемента данных: PPP T M RRR, где PPP – биты приоритета, T – бит маркера, M – бит монитора, RRR – резервные биты.

Каждый кадр или маркер имеет приоритет, устанавливаемый битами приоритета (значение от 0 до 7, 7 – наивысший приоритет). Станция может воспользоваться маркером, если только она получила маркер с приоритетом, меньшим, чем ее собственный, или равным ему. Сетевой адаптер станции, если ему не удалось захватить маркер, помещает свой приоритет в резервные биты маркера, но только в том случае, если записанный в резервных битах приоритет ниже его собственного. Эта станция будет иметь преимущественный доступ при последующем поступлении к ней маркера.

Схема использования приоритетного метода захвата маркера показана на рис. 5.13. Сначала монитор помещает в поле текущего приоритета P максимальное значение приоритета, а поле резервного приоритета R обнуляется (маркер 7110). Маркер проходит по кольцу, в котором станции имеют текущие приоритеты 3, 6 и 4. Так как эти значения меньше, чем 7, то захватить маркер станции не могут, но они записывают свое значение приоритета в

поле резервного приоритета, если их приоритет выше его текущего значения. В результате маркер возвращается к монитору со значением резервного приоритета $R = 6$. Монитор переписывает это значение в поле P , а значение резервного приоритета обнуляет, и снова отправляет маркер по кольцу. При этом обороте его захватывает станция с приоритетом 6 – наивысшим приоритетом в кольце в данный момент времени.

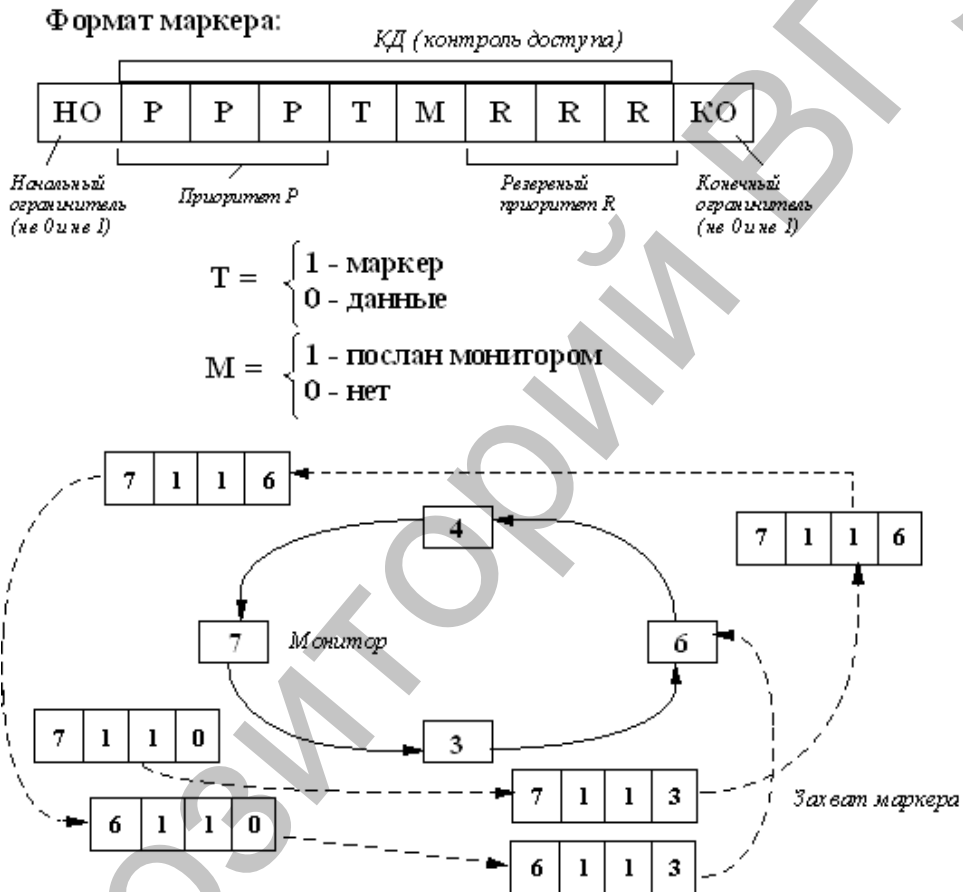


Рис. 5.13. Приоритеты в кольце Token Ring.

Бит маркера имеет значение 0 для маркера и 1 для кадра.

Бит монитора устанавливается в 1 активным монитором и в 0 любой другой станцией, передающей маркер или кадр. Если активный монитор видит маркер или кадр, содержащий бит монитора в 1, то активный монитор знает, что этот кадр или маркер уже однажды обошел кольцо и не был обработан станциями. Если это кадр, то он удаляется из кольца. Если это маркер, то активный монитор переписывает приоритет из резервных битов полученного маркера в поле приоритета. Поэтому при следующем проходе маркера по кольцу его захватит станция, имеющая наивысший приоритет.

Поле конечного ограничителя – последнее поле маркера. Так же, как и поле начального ограничителя, это поле содержит уникальную серию электрических импульсов, которые нельзя спутать с данными. Кроме отметки конца маркера это поле также содержит два подполя: бит промежуточного кадра и бит ошибки. Эти поля относятся больше к кадру данных, структура которого рассмотрена ниже.

Кадр данных. Кадр данных состоит из нескольких групп полей:

- последовательность начала кадра;
- адрес получателя;
- адрес отправителя;
- поле данных;
- последовательность контроля кадра;
- последовательность конца кадра.

Кадр данных может переносить либо данные для управления кольцом, либо пользовательские данные.

Адрес получателя (либо 2, либо 6 байтов). Первый бит определяет групповой или индивидуальный адрес как для 2-байтовых, так и для 6-байтовых адресов. Второй бит в 6-байтовых адресах говорит, назначен адрес локально или глобально.

Адрес отправителя имеет тот же размер и формат, что и адрес получателя.

Поле данных кадра может содержать служебные данные, используемые для управления работой сети, или данные пользователя, предназначенных для (или получаемых от) протокола более высокого уровня. Это поле не имеет определенной максимальной длины, хотя существуют практические ограничения на его размер, основанные на временных требованиях к тому, как долго некоторая станция может управлять кольцом.

Последовательность контроля кадра – используется для обнаружения ошибок, состоит из четырех байтов остатка циклически избыточной контрольной суммы, вычисляемой по алгоритму CRC-32, осуществляющему циклическое суммирование по модулю 32.

Последовательность конца кадра состоит из двух полей: конечный ограничитель и статус кадра.

Конечный ограничитель в кадре данных имеет дополнительное значение по сравнению с маркером. Кроме уникальной последовательности электрических импульсов он содержит два однобитовых поля: бит промежуточного кадра и бит обнаружения ошибки. Бит промежуточного кадра устанавливается в 1, ес-

ли этот кадр является частью многокадровой передачи, или в 0 для последнего или единственного кадра. Бит обнаружения ошибки первоначально установлен в 0; каждая станция, через которую передается кадр, проверяет его на ошибки (по коду CRC) и устанавливает бит обнаружения ошибки в 1, если она выявлена. Очередная станция, которая видит уже установленный бит обнаружения ошибки, должна просто передать кадр. Исходная станция заметит, что возникла ошибка, и повторит передачу кадра.

Статус кадра имеет длину 1 байт и содержит 4 резервных бита и два подполя: бит распознавания адреса и бит копирования кадра. Так как это поле не сопровождается вычисляемой суммой CRC, то используемые биты дублируются в байте. Когда кадр создается, передающая станция устанавливает бит распознавания адреса в 0; получающая станция устанавливает бит в 1, чтобы сообщить, что она опознала адрес получателя. Бит копирования кадра также вначале установлен в 0, но устанавливается в 1 получающей станцией (станцией назначения), когда она копирует содержимое кадра в собственную память (другими словами, когда она реально получает данные). Данные копируются (и бит устанавливается), если только кадр получен без ошибок. Если кадр возвращается с обоими установленными битами, исходная станция знает, что произошло успешное получение. Если бит распознавания адреса не установлен во время получения кадра, это означает, что станция назначения больше не присутствует в сети (возможно, вследствие неполадок). Возможна другая ситуация, когда адрес получателя опознается, но бит копирования кадра не установлен. Это говорит исходной станции, что кадр был искажен во время передачи (бит обнаружения ошибки в конечном ограничителе также будет установлен). Если оба бита опознавания адреса и копирования кадра установлены, и бит обнаружения ошибки также установлен, то исходная станция знает, что ошибка случилась после того, как этот кадр был корректно получен.

Практическая реализация сетей Token Ring. Стандарт Token Ring фирмы IBM предусматривает построение связей в сети как с помощью непосредственного соединения станций друг с другом, так и образование кольца с помощью концентраторов (называемых MAU – Media Attachment Unit или MSAU – Multi-Station Access Unit).

На рис. 5.14 показаны основные аппаратные элементы сети Token Ring и способы их соединения.

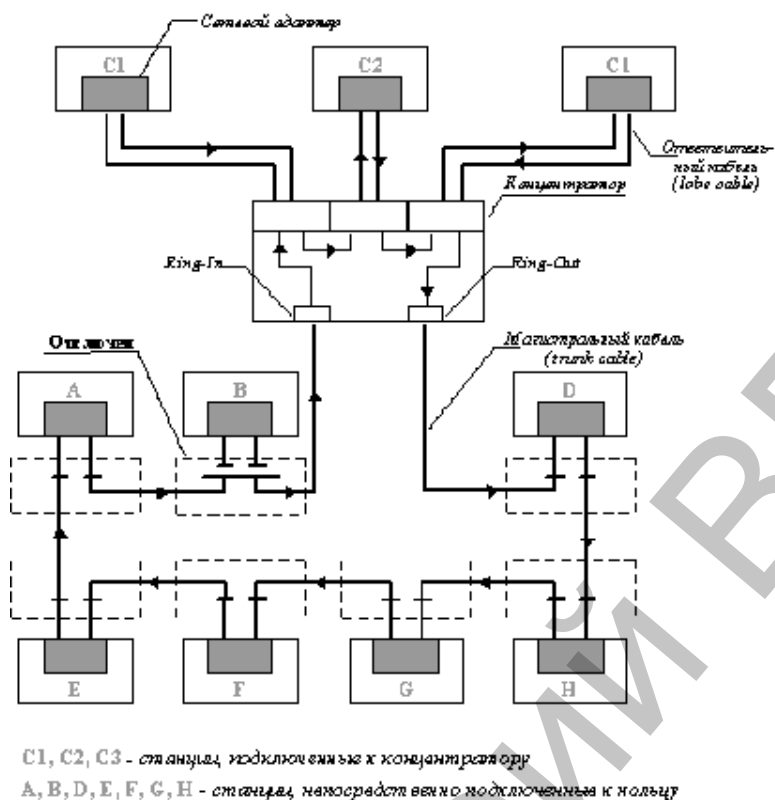


Рис. 5.14. Конфигурация кольца Token Ring.

В приведенной конфигурации показаны станции двух типов. Станции C1, C2 и C3 – это станции, подключаемые к кольцу через концентратор. Обычно такими станциями являются компьютеры с установленными в них сетевыми адаптерами. Станции этого типа соединяются с концентратором ответвительным кабелем (lobe cable), который обычно является экранированной витой парой (Shielded Twisted Pair, STP), соответствующей стандартному типу кабеля из кабельной системы IBM (Type 1, 2, 6, 8, 9).

Максимальная длина ответвительного кабеля зависит от типа концентратора, типа кабеля и скорости передачи данных. Обычно для скорости 16 Мб/с максимальная длина кабеля Type 1 может достигать 200 м, а для скорости 4 Мб/с – 600 м. Концентраторы Token Ring делятся на активные и пассивные. Пассивные концентраторы обеспечивают только соединения портов внутри концентратора в кольцо, активные выполняют и функции повторителя, обеспечивая ресинхронизацию сигналов и исправление их амплитуды и формы. Естественно, что активные концентраторы поддерживают большие расстояния до станции, чем пассивные.

Остальные станции сети соединены в кольцо непосредственными связями. Такие связи называются магистральными (trunk cable). Обычно связи такого рода используются для соединения концентраторов друг с другом для образования общего

кольца. Порты концентраторов, предназначенные для такого соединения, называются портами Ring-In и Ring-Out.

Для предотвращения влияния отказавшей или отключенной станции на работу кольца станции подключаются к магистрали кольца через специальные устройства, называемые устройствами подключения к магистрали (Trunk Coupling Unit, TCU). В функции такого устройства входит образование обходного пути, исключающего заход магистрали в MAC-узел станции при ее отключении или отказе. Обычно для этих целей в TCU используются реле, которые подпитываются постоянным током во время нормальной работы. При пропадании тока подпитки контакты реле переключаются и образуют обходной путь, исключая станцию.

При подключении станции в кольцо через концентратор, устройства TCU встраивают в порты концентратора.

Максимальное количество станций в одном кольце – 250.

Кроме экранированной витой пары существуют сетевые адаптеры и концентраторы Token Ring, поддерживающие неэкранированную витую пару и оптоволокно.

Кадр Token Ring. В технологии Token Ring каждый кадр начинается с «последовательности начала кадра», которая содержит следующие поля:

Начальный ограничитель, такой же, как и для маркера;

Управление доступом, также совпадает для кадров и для маркеров;

Поле управления кадром, в котором передается служебная информация, указывающая на то, что передается в кадре: данные пользователя или управляющая информация. В последнем случае также указывается тип управляющей информации.

Адрес получателя (либо 2, либо 6 байтов). Первый бит определяет групповой или индивидуальный адрес как для 2-байтовых, так и для 6-байтовых адресов. Вторым битом в 6-байтовых адресах говорит, назначен адрес локально или глобально.

Адрес отправителя имеет тот же размер и формат, что и адрес получателя.

Поле данных кадра может содержать данные одного из описанных управляющих кадров MAC или запись пользовательских данных, предназначенных для (или получаемых от) протокола более высокого уровня, такого, как IPX или NetBIOS. Это поле не имеет определенной максимальной длины, хотя существуют практические ограничения на его размер, основанные на временных требованиях к тому, как долго некоторая станция может управлять кольцом.

Последовательность контроля кадра – используется для обнаружения ошибок, состоит из четырех байтов остатка циклически избыточной контрольной суммы, вычисляемой по алгоритму CRC-32, осуществляющему циклическое суммирование по модулю 32.

Последовательность конца кадра состоит из двух полей: конечный ограничитель и статус кадра.

Конечный ограничитель в кадре данных имеет дополнительное значение по сравнению с маркером. Кроме уникальной последовательности электрических импульсов он содержит два однобитовых поля: бит промежуточного кадра и бит обнаружения ошибки. Бит промежуточного кадра устанавливается в 1, если этот кадр является частью многокадровой передачи, или в 0 для последнего или единственного кадра. Бит обнаружения ошибки первоначально установлен в 0; каждая станция, через которую передается кадр, проверяет его на ошибки (по коду CRC) и устанавливает бит обнаружения ошибки в 1, если она выявлена. Очередная станция, которая видит уже установленный бит обнаружения ошибки, должна просто передать кадр. Исходная станция заметит, что возникла ошибка, и повторит передачу кадра.

Статус кадра имеет длину 1 байт и содержит 4 резервных бита и два подполя: бит распознавания адреса и бит копирования кадра. Так как это поле не сопровождается вычисляемой суммой CRC, то используемые биты дублируются в байте. Когда кадр создается, передающая станция устанавливает бит распознавания адреса в 0; получающая станция устанавливает бит в 1, чтобы сообщить, что она опознала адрес получателя. Бит копирования кадра также вначале установлен в 0, но устанавливается в 1 получающей станцией (станцией назначения), когда она копирует содержимое кадра в собственную память (другими словами, когда она реально получает данные). Данные копируются (и бит устанавливается), если только кадр получен без ошибок. Если кадр возвращается с обоими установленными битами, исходная станция знает, что произошло успешное получение. Если бит распознавания адреса не установлен во время получения кадра, это означает, что станция назначения больше не присутствует в сети (возможно, вследствие неполадок). Возможна другая ситуация, когда адрес получателя опознается, но бит копирования кадра не установлен. Это говорит исходной станции, что кадр был искажен во время передачи (бит обнаружения ошибки в конечном ограничителе также будет установлен). Если оба бита опознавания адреса и копирования кадра установлены и бит обнаружения ошибки также установлен то исходная станция знает, что ошибка случилась после того, как этот кадр был корректно получен.

Лекция 6. Глобальная информационная сеть Интернет

6.1. Предпосылки возникновения Интернет

Передача данных между локальными сетями компьютеров стала неотъемлемой частью современных информационных технологий. Локальные сети, разбросанные по всему миру, могут содержать данные, представляющие интерес для пользователей других локальных сетей, удаленных от первых на значительные расстояния. Эти данные могут относиться к различным сферам человеческой деятельности: экономике, охране окружающей среды, добыче полезных ископаемых и т.д. В научном мире передача данных между локальными сетями компьютеров позволяет ученым посылать программы и данные на удаленные суперкомпьютеры для обработки, получать результаты и обмениваться научной информацией с коллегами.

К сожалению, большинство локальных сетей создаются независимо друг от друга. Пользователи выбирают аппаратную технологию, подходящую для их коммуникационных проблем. Некоторым пользователям нужна высокоскоростная сеть, соединяющая их машины, но такие сети не могут быть расширены на большие расстояния, их эффективность снижается при увеличении количества компьютеров свыше некоторого оптимального значения. Другим подходит более медленная сеть, но она должна соединять достаточно большое количество компьютеров и т.д. Может оказаться так, что локальные сети используют одну и ту же технологию, но они разнесены на значительные расстояния, что не позволяет осуществить обмен между данными сетями на основе общей технологии. Кроме того, применяемые на практике технологии локальных сетей ориентированы на конкретные сетевые топологии и не обеспечивают доставки сообщений в сетях с произвольной топологией.

Во всех рассмотренных выше примерах для организации обмена данными между сетями нужна новая технология в отличие от тех, что используются в локальных сетях. Эта технология должна сделать возможным взаимное соединение большого числа разделенных физических сетей и заставить их работать как одно единое целое. Она получила название межсетевого обмена (internetworking), поскольку приспособливает друг к другу различные аппаратные технологии, лежащие в основе физических сетей, с помощью добавления как физических соединений сетей, так и нового набора соглашений. Технология межсетевого обмена скрывает детали сетевого оборудования и позволяет компьютерам взаимодействовать вне зависимости от типа их физических соединений.

Новая технология межсетевого обмена является примером модели Взаимодействия Открытых Систем. Они называются открытыми потому, что в отличие от узкоспециализированных коммуникационных систем, продаваемых тем или иным производителем, ее спецификации доступны всем. Поэтому любой пользователь может создать программное обеспечение, необходимое для взаимодействия в объединенной сети. Более того, вся эта технология была разработана для того, чтобы упростить взаимодействие между машинами с различными аппаратными архитектурами, чтобы использовать почти любое оборудование сети с коммутацией пакетов и чтобы позволить взаимодействие различных операционных систем.

Межсетевая технология появилась в результате проведения разработок, финансировавшихся Агентством перспективных исследований Министерства обороны США DARPA (Defense Advanced Research Projects Agency). Технология DARPA включает набор сетевых стандартов, описывающих детально процесс взаимодействия компьютеров, а также ряд соглашений при взаимодействии сетей и маршрутизации трафика. Официально называемый Стеком межсетевых протоколов TCP/IP (а в обыденной речи – TCP/IP по именам двух основных стандартов), набор сетевых стандартов может использоваться для взаимодействия компьютеров, принадлежащих неограниченному числу сетей.

Технология DARPA стала базовой технологией для огромного количества сетей, которые в настоящее время связывают большинство исследовательских центров и институтов, промышленных компаний, университетов и учебных заведений, других организаций. Для обозначения получившейся распределенной сетевой среды стали использовать термины Интернет DARPA, Интернет TCP/IP или просто Интернет.

Способы объединения сетей. Как соединяются сети между собой для создания объединенной сети? Физически две сети могут соединяться только с помощью компьютера, присоединенного к каждой из них. Компьютеры, соединяющие две сети и передающие пакеты из одной в другую, называются межсетевыми шлюзами (gateway) или межсетевыми маршрутизаторами (router).

Рассмотрим пример, показанный на рис. 6.1. Две физические сети соединены со шлюзом G. Шлюз принимает пакеты из сети 1, предназначенные компьютерам в другой сети, и передает их в сеть 2. Аналогично G должен принимать пакеты из сети 2, которые предназначены для компьютеров в сети 1, и передавать их туда.

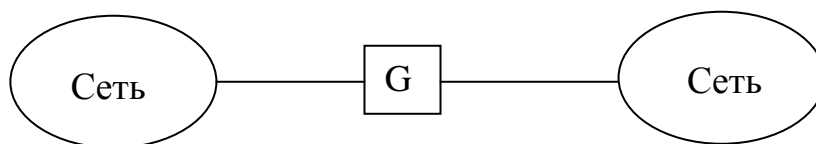


Рис. 6.1. Схема соединения 2-х сетей.

Когда соединения физических сетей становятся более сложными, шлюзам нужно знать о топологии Интернета за пределами сетей, к которым они присоединены. На рис. 6.2 в качестве примера изображены три сети, соединенные между собой двумя шлюзами.

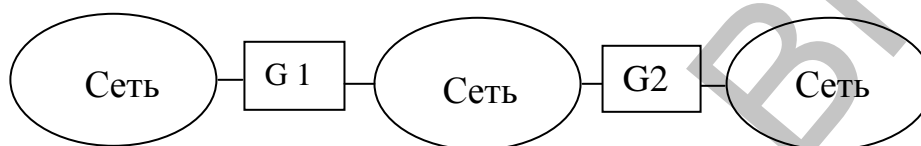


Рис. 6.2. Схема соединения 3-х сетей с помощью двух шлюзов.

В этом примере шлюз G1 должен перемещать из сети 1 в сеть 2 все пакеты, предназначенные для машин либо в сети 2, либо в сети 3. По мере того, как количество физических сетей увеличивается, задача шлюза по принятию решений о том, куда посылать пакеты, становится более сложной. Может сложиться впечатление, что шлюзы, которые знают, как направить пакеты к их получателю, являются большими машинами, имеющими достаточное количество основной или внешней памяти для хранения информации о каждой машине в Интернете. Тем не менее, шлюзы, используемые в сети Интернет, обычно являются миникомпьютерами; они часто имеют небольшую дисковую память или не имеют ее вообще, а также имеют ограниченную оперативную память. Причина использования маленьких межсетевых шлюзов заключена в следующем: шлюзы маршрутизируют пакеты, основываясь на информации о сети получателя, а не на информации о компьютере-получателе. Если маршрутизация сообщений или данных осуществляется между сетями, то количество информации, которую нужно хранить шлюзу, пропорционально количеству сетей в Интернете, а не числу машин в нем.

6.2. Адресация в сети Интернет

Для того, чтобы осуществлять обмен сообщениями, каждый компьютер или маршрутизатор в сети Интернет должен иметь уникальный адрес. Адрес устройства, подключенного к сети Интернет, называется IP-адресом и имеет размер 32 бита (4 байта). В качестве примера рассмотрим произвольный IP-адрес 10000000000010100000001000011110. Именно с такими адресами,

представленными в виде 32 разрядного двоичного числа, оперируют любые устройства (компьютеры и маршрутизаторы), осуществляющие обмен сообщениями в сети Интернет. Двоичная форма представления IP-адреса удобна для компьютеров, но очень неудобна для людей, которым пришлось бы читать, запоминать и вводить с клавиатуры такие числа. Поэтому на практике используют IP-адрес, представленный в более компактной форме, в которой каждый из 4 байт двоичного числа записывается в виде десятичных чисел, разделенных точкой: 128.10.2.30. Такой формат представления IP-адреса называется «десятичное с точкой».

Классы IP-адресов. IP-адрес назначается администратором во время конфигурирования компьютеров или маршрутизаторов. IP-адрес состоит из двух частей: номера сети и номера узла. Номер сети назначается по рекомендации специального подразделения Интернет-Сетевого информационного центра (Network Information Center, NIC). Обычно провайдеры услуг Интернет получают наборы (диапазоны) адресов у подразделений NIC, а затем распределяют их между своими абонентами.

Все IP-адреса делятся на 5 классов, которые приведены в табл. 6.1.

Таблица 6.1

	биты адреса 0	7 8	15 16	23 24	31
Класс А	0	номер сети	номер узла (хоста)		
Класс В	10	номер сети		номер узла (хоста)	
Класс С	110	номер сети			номер узла (хоста)
Класс D	1110	Адреса для обращения к группам компьютеров (multicast)			
Класс E	11110	Зарезервировано			

Адрес состоит из двух логических частей – номера сети и номера узла в сети. Какая часть адреса относится к номеру сети, а какая к номеру узла, определяется значениями первых битов адреса:

если адрес в двоичном представлении начинается с 0, то сеть относят к классу А, и номер сети занимает один байт, остальные 3 байта интерпретируются как номер узла в сети. Сети класса А имеют номера в диапазоне от 1 до 126;

если первые два бита адреса в двоичном представлении равны 10, то сеть относится к классу В. В адресах этого класса номер сети занимает 16 битов (с 0 по 15), остальные 16 битов интерпретируются как номер узла в сети;

если адрес начинается с последовательности 110, то это сеть класса С. Для обозначения номера сети отводится 24 бита, остальные 8 битов интерпретируются как номер узла в сети;

если адрес начинается с последовательности 1110, то он является адресом класса D и обозначает особый адрес, который используется для групповой передачи. Если в пакете в качестве адреса назначения указан адрес класса D, то такой пакет должны получить все узлы, которым присвоен данный адрес;

если адрес начинается с последовательности 11110, то он относится к классу E. Адреса этого класса зарезервированы.

В табл. 6.2 представлены дополнительные сведения об адресах различных классов.

Таблица 6.2

Класс адреса	Диапазон возможных значений 1-го байта адреса	Возможное количество сетей	Возможное количество узлов
A	1–126	126	16.777,214
B	128–191	16,382	65,534
C	192–223	2.097,150	254
D	224–239	–	2^{28}
E	240–247	–	2^{27}

Сети класса A имеют номера в диапазоне от 1 до 126. Номер 0 не используется, а номер 127 зарезервирован для специальных целей. Существует несколько соглашений об особой интерпретации IP-адресов:

- если IP-адрес состоит только из двоичных нулей (00000.....00000), то он обозначает адрес того узла, который сгенерировал этот пакет;
- если в поле номера сети стоят 0, то по умолчанию считается, что этот узел принадлежит той же самой сети, что и узел, который отправил пакет;
- если все двоичные разряды IP-адреса равны 1 (111111.....111111), то пакет с таким адресом назначения должен рассылаться всем узлам, находящимся в той же сети, что и источник этого пакета. Такая рассылка называется ограниченным широковещательным сообщением (limited broadcast);
- если в поле номера узла адреса назначения стоят двоичные единицы (111...111), то пакет, имеющий такой адрес, рассылается всем узлам сети с заданным номером. Такая рассылка называется широковещательным сообщением (broadcast);
- адрес 127.0.0.1 зарезервирован для организации обратной связи при тестировании работы программного обеспечения узла без реальной отправки пакета по сети. Этот адрес имеет название loopback.

Уже упоминавшаяся форма группового IP-адреса – multicast – означает, что данный пакет должен быть доставлен сразу нескольким узлам, которые образуют группу с номером, указанным в поле адреса. Узлы сами идентифицируют себя, то есть определяют, к какой из групп они относятся. Один и тот же узел может входить в несколько групп. Такие сообщения в отличие от широковещательных называются мультивещательными. Групповой адрес не делится на поля номера сети и узла и обрабатывается маршрутизатором особым образом.

IP-адрес назначается независимо от локального адреса узла в конкретной сети, в которую входит данный узел. Один и тот же узел локальной сети может одновременно входить в несколько IP-сетей. В этом случае узел должен иметь несколько IP-адресов, по числу сетевых связей. Таким образом следует иметь в виду, что IP-адрес характеризует не отдельный компьютер или маршрутизатор, а одно сетевое соединение.

Система доменных имен Интернет. Хотя IP-адрес, представленный в формате «десятичное с точкой», гораздо легче запомнить по сравнению со случаем, когда используется двоичный формат, для многих людей и это оказывается достаточно трудной проблемой. По этой причине разработчики Интернет реализовали так называемую систему имен доменов Интернет (Domain Name System, DNS). Система имен доменов позволяет обращаться к сетевым компьютерам не только по их IP-адресам, но и по индивидуальным именам узлов, представленным в символьной форме. Например, для того, чтобы связаться с WWW-сервером ВГУ, можно использовать либо его IP-адрес (например) 195.209.230.198, либо его имя *www.vsu.by*.

Имя *www.vsu.by* состоит из трех элементов: *www*, *vsu* и *by*, отделенных друг от друга точками. Эти элементы описывают сферу деятельности, принадлежности или выполняемых функций определенного компьютера и носят название доменов. Домен *www* означает то, что данный компьютер является узлом WWW, т.е. на нем работает WWW-сервер. Домен *vsu* означает организацию, которой принадлежит компьютер – Витебский государственный университет им. П.М. Машерова (в английском варианте – Vitebsk State University). Домен *by* означает страну (Беларусь), к которой относится организация *vsu*, и называется доменом верхнего уровня.

Преобразование доменных имен в IP-адреса осуществляется специальными программами-серверами DNS (серверами имен). Если для сервера DNS отводится отдельный компьютер, то тогда сервером DNS можно назвать весь компьютер целиком.

Поскольку ни один сервер DNS не может содержать необходимой информации об IP-адресах всех компьютеров, подключенных к сети Интернет, то аналогичное программное обеспечение сервера DNS работает на тысячах компьютерах в сети Интернет, организованных в иерархическую структуру. Поэтому DNS можно рассматривать как распределенную базу данных, в которой данные о конкретных компьютерах находятся на различных серверах DNS. При этом серверы DNS могут посылать друг другу запросы и получать ответы на них. Например, когда какой-нибудь прикладной программе нужно соединиться с удаленным компьютером, она первым делом соединяется с ближайшим DNS-сервером, IP-адрес которого при конфигурировании компьютера вводится администратором. Если данные о запрошенном соответствии доменного имени и IP-адреса хранятся в базе данного DNS-сервера, то он сразу посылает ответ клиенту, если же нет – то он посылает запрос DNS-серверу другого домена, который, в свою очередь, может сам обработать запрос либо передать его другому DNS-серверу. Все DNS-серверы соединены иерархически, в соответствии с иерархией доменов сети Интернет. Исходный запрос будет передаваться по цепочке DNS-серверов, пока не будет найдено нужное соответствие имени и IP-адреса. Этот процесс ускоряется из-за того, что серверы имен постоянно кэшируют информацию, предоставляемую по запросам. Для повышения надежности своей работы компьютеры могут использовать в своей работе IP-адреса нескольких DNS-серверов.

База данных DNS имеет структуру дерева, называемого доменным пространством имен, в котором каждый домен (узел дерева) имеет имя и может содержать поддомены. Имя домена идентифицирует его положение в этой базе данных по отношению к родительскому домену, причем точки в имени отделяют части, соответствующие узлам домена. Корень базы данных DNS управляется центром Internet Network Information Center. Домены верхнего уровня назначаются для каждой страны, а также для определенных классов организаций. Имена этих доменов должны следовать международному стандарту ISO 3166. Для обозначения стран используются трехбуквенные и двухбуквенные аббревиатуры, а для различных классов организаций используются следующие аббревиатуры:

- com – коммерческие организации (например, microsoft.com);
- edu – образовательные (например, mit.edu);
- mil – военные организации (только для организаций США);
- int – международные организации
- gov – правительственные организации США (например, whitehouse.gov);

org – некоммерческие организации (например, fidonet.org);
net – организации, поддерживающие сети (например, nsf.net).

Каждый домен DNS администрируется отдельной организацией, которая обычно разбивает свой домен на поддомены и передает функции администрирования этих поддоменов другим организациям. Каждый домен имеет уникальное имя, а каждый из поддоменов имеет уникальное имя внутри своего домена. Имя домена может содержать до 63 символов. Каждый хост в сети Internet однозначно определяется своим *полным доменным именем* (*fully qualified domain name, FQDN*), которое включает имена всех доменов по направлению от хоста к корню.

6.3. Стек протоколов TCP/IP

На концептуальном уровне программное обеспечение TCP/IP организовано в виде 5 уровней, показанных на рис. 6.3. Следует обратить внимание на то, что, хотя программное обеспечение TCP/IP также организовано в виде многоуровневой структуры, тем не менее эта модель отличается от известной 7-уровневой модели взаимодействия открытых систем. Дело в том, что стек TCP/IP был разработан до появления модели взаимодействия открытых систем (Open Systems Interconnection, OSI). Основные протоколы стека TCP/IP – Transmission Control Protocol (TCP) и Internet Protocol (IP), давшие название всему семейству протоколов межсетевого обмена, появились в 1982 г., в то время как модель взаимодействия открытых систем была принята международной организацией по стандартизации (ISO) в 1984 г.

На рис. 6.3 указаны основные протоколы, участвующие в передаче данных между компьютерами в сети Интернет, а также форма, в которой передаются данные между уровнями.

Два самых нижних уровня на рис. 6.3 в стеке протоколов TCP/IP не регламентируются и соответствуют физическому и каналному уровням модели OSI. Они поддерживают все популярные стандарты физического и канального уровня: для локальных сетей это Ethernet, Token Ring, FDDI, Fast Ethernet, 100VG-AnyLAN, для глобальных сетей – протоколы соединений «точка-точка» SLIP и PPP, протоколы территориальных сетей с коммутацией пакетов X.25, Frame Relay. Разработана также специальная спецификация, определяющая использование технологии ATM в качестве транспорта канального уровня. Обычно при появлении новой технологии локальных или глобальных сетей она быстро включается в стек TCP/IP за счет разработки соответствующего RFC, определяющего метод инкапсуляции пакетов IP в ее кадры.

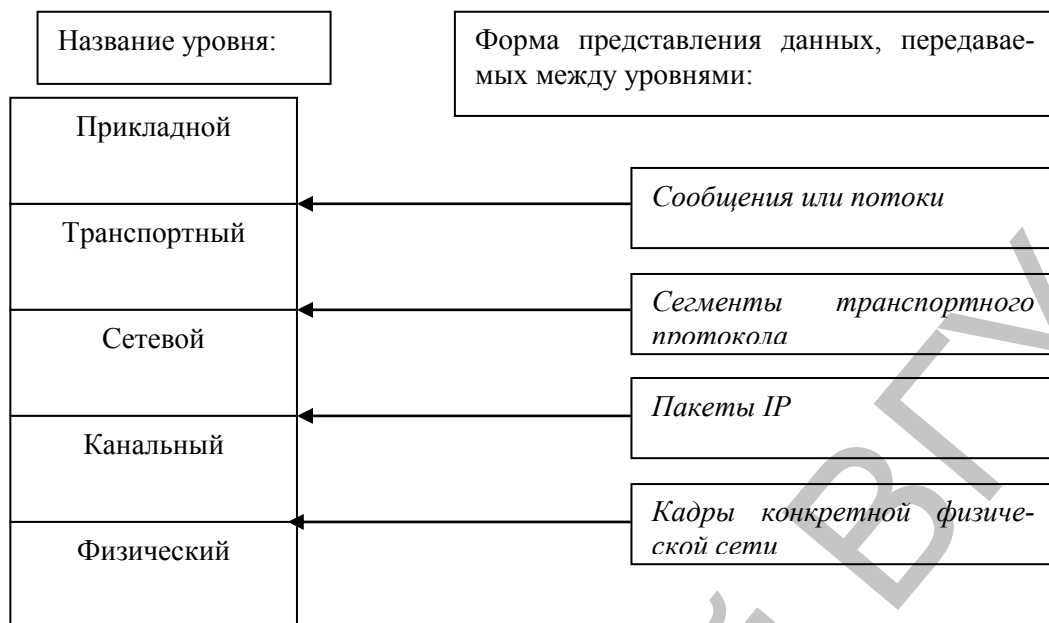


Рис. 6.3. Пятиуровневая модель реализации стека протоколов TCP/IP.

Третий уровень – это уровень межсетевого взаимодействия, который занимается передачей пакетов с использованием различных транспортных технологий локальных сетей, территориальных сетей, линий специальной связи и т.п. В качестве основного протокола сетевого уровня (в терминах модели OSI) в стеке используется протокол IP, который изначально проектировался как протокол передачи пакетов в составных сетях, состоящих из большого количества локальных сетей, объединенных как локальными, так и глобальными связями. Поэтому протокол IP хорошо работает в сетях со сложной топологией. Протокол IP является ненадежным протоколом, то есть он не гарантирует доставку пакетов до узла назначения.

К уровню межсетевого взаимодействия относятся и все протоколы, связанные с составлением и модификацией таблиц маршрутизации, такие, как протоколы сбора маршрутной информации RIP (Routing Internet Protocol) и OSPF (Open Shortest Path First), а также протокол межсетевых управляющих сообщений ICMP (Internet Control Message Protocol). Последний протокол предназначен для обмена информацией об ошибках между маршрутизаторами сети и узлом – источником пакета. С помощью специальных пакетов ICMP сообщает о невозможности доставки пакета, о превышении времени жизни или продолжительности сборки пакета из фрагментов, об аномальных величинах параметров, об изменении маршрута пересылки и типа обслуживания, о состоянии системы и т.п.

На четвертом уровне функционируют протокол управления передачей TCP (Transmission Control Protocol) и протокол дейта-

грамм пользователя UDP (User Datagram Protocol). Протокол TCP обеспечивает надежную передачу сообщений между удаленными прикладными процессами за счет образования виртуальных соединений. В отличие от TCP протокол UDP является ненадежным протоколом и не гарантирует доставку пакетов по назначению. Однако для ряда приложений, таких, например, как компьютерная видеоконференцсвязь, такой протокол подходит больше, чем высоконадежный протокол TCP.

Верхний уровень называется прикладным. За долгие годы использования в сетях различных стран и организаций стек TCP/IP накопил большое количество протоколов и сервисов прикладного уровня. К ним относятся такие широко используемые протоколы, как протокол пересылки файлов FTP, протокол эмуляции терминала telnet, почтовый протокол SMTP, используемый в электронной почте сети Интернет, гипертекстовые сервисы доступа к удаленной информации, такие, как WWW и многие другие. Остановимся несколько подробнее на некоторых из них.

Протокол пересылки файлов FTP (File Transfer Protocol) реализует удаленный доступ к файлу. Для того чтобы обеспечить надежную передачу, FTP использует в качестве транспорта протокол с установлением соединений – TCP. Кроме пересылки файлов протокол FTP предлагает и другие услуги. Так, пользователю предоставляется возможность интерактивной работы с удаленной машиной, например, он может распечатать содержимое ее каталогов. Наконец, FTP выполняет аутентификацию пользователей. Прежде чем получить доступ к файлу, в соответствии с протоколом пользователи должны сообщить свое имя и пароль. Для доступа к публичным каталогам FTP-архивов сети Интернет парольная аутентификация не требуется, и ее обходят за счет использования для такого доступа предопределенного имени пользователя Anonymous. В стеке TCP/IP протокол FTP предлагает наиболее широкий набор услуг для работы с файлами, однако он является и самым сложным для программирования. Приложения, которым не требуются все возможности FTP, могут использовать другой, более экономичный протокол – простейший протокол пересылки файлов TFTP (Trivial File Transfer Protocol). Этот протокол реализует только передачу файлов, причем в качестве транспорта используется более простой, чем TCP, протокол без установления соединения – UDP. Протокол TFTP целесообразно использовать для не требующих высокой надежности доставки приложений.

Протокол виртуального терминала telnet обеспечивает передачу потока байтов между процессами, а также между процессом и терминалом. Наиболее часто этот протокол используется

для эмуляции терминала удаленного компьютера. При использовании сервиса telnet пользователь фактически управляет удаленным компьютером так же, как и локальный пользователь, поэтому такой вид доступа требует хорошей защиты. Поэтому серверы telnet всегда используют как минимум аутентификацию по паролю, а иногда и более мощные средства защиты.

Протокол SNMP (Simple Network Management Protocol) используется для организации сетевого управления. Изначально протокол SNMP был разработан для удаленного контроля и управления маршрутизаторами Интернет, которые традиционно часто называют также шлюзами. С ростом популярности протокол SNMP стали применять и для управления любым коммуникационным оборудованием – концентраторами, мостами, сетевыми адаптерами и т.д., и т.п. Проблема управления в протоколе SNMP разделяется на две задачи. Первая задача связана с передачей информации. Протоколы передачи управляющей информации определяют процедуру взаимодействия SNMP-агента, работающего в управляемом оборудовании, и SNMP-монитора, работающего на компьютере администратора, который часто называют также консолью управления. Протоколы передачи определяют форматы сообщений, которыми обмениваются агенты и монитор. Вторая задача связана с контролируемыми переменными, характеризующими состояние управляемого устройства. Стандарты регламентируют, какие данные должны сохраняться и накапливаться в устройствах, имена этих данных и синтаксис этих имен. В стандарте SNMP определена спецификация информационной базы данных управления сетью. Эта спецификация, известная как база данных MIB (Management Information Base), определяет те элементы данных, которые управляемое устройство должно сохранять, и допустимые операции над ними.

6.4. Маршрутизация сообщений

Многосвязность базовой сети передачи данных порождает проблему маршрутизации. Пакет вводится в сеть, имея в своем заголовке адрес пункта назначения. В принципе, этот адрес в первом же узле коммутации может быть преобразован в полный маршрут к узлу назначения (маршрутизация от источника). Однако, более часто решение о дальнейшем маршруте пакета принимается каждым УК, через который проходит пакет. Для принятия такого решения УК должен содержать таблицу маршрутизации, указывающую для каждого адресата номер выходной линии, в которую следует отправить пакет. Например, в табл. 6.3 пред-

ставлены основной и альтернативный маршруты из УК1 к другим УК сети, топология которой изображена на рис. 6.4.

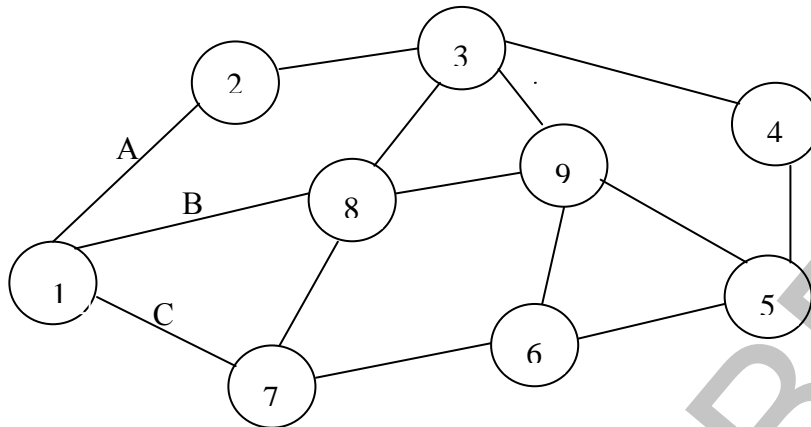


Рис. 6.4. Топология многосвязной сети.

Таблица 6.3

Таблица маршрутизации в УК1

Адрес назначения	Выходная линия	Альтернативная выходная линия
УК2	А	В
УК3	А	В
УК4	В	А
УК5	В	С
УК6	С	В
УК7	С	В
УК8	В	С
УК9	В	А

В небольших сетях на каждом узле коммутации можно поддерживать статический набор альтернативных маршрутов, учитывающих большинство наиболее вероятных отказов. Однако более гибким является адаптивный метод, который в случае нарушения работоспособности какой-либо линии передачи, или УК, корректирует таблицу маршрутизации узла. Для этого УК должен располагать информацией о топологии сети (для вычисления маршрута) и регулярно получать информацию о ее состоянии (работоспособности линий связи и уровне загрузки других УК). Эти задачи решаются специализированными алгоритмами маршрутизации.

6.5. Развитие стека TCP/IP: протокол IPv6

Технология стека TCP/IP сложилась в основном в конце 1970-х годов и с тех пор основные принципы работы базовых протоколов, таких, как IP, TCP, UDP и ICMP, практически не из-

менились. Однако сам компьютерный мир за эти годы значительно изменился, поэтому долго назревавшие усовершенствования в технологии стека TCP/IP сейчас стали необходимостью.

Основные обстоятельства, из-за которых требуется модификация базовых протоколов стека TCP/IP, перечислены ниже.

Повышение производительности компьютеров и коммуникационного оборудования. За время существования стека производительность компьютеров возросла на два порядка, объемы оперативной памяти выросли более чем в 30 раз, пропускная способность магистрали Internet в Соединенных Штатах выросла в 800 раз.

Появление новых приложений. Коммерческий бум вокруг Internet и использование ее технологий при создании intranet привели к появлению в сетях TCP/IP, ранее использовавшихся в основном в научных целях, большого количества приложений нового типа, работающих с мультимедийной информацией. Эти приложения чувствительны к задержкам передачи пакетов, так как такие задержки приводят к искажению передаваемых в реальном времени речевых сообщений и видеоизображений. Особенностью мультимедийных приложений является также передача очень больших объемов информации.

Некоторые технологии вычислительных сетей, например, frame relay и АТМ, уже имеют в своем арсенале механизмы для резервирования полосы пропускания для определенных приложений. Однако эти технологии еще не скоро вытеснят традиционные технологии локальных сетей, не поддерживающие мультимедийные приложения (например, Ethernet). Следовательно, необходимо компенсировать такой недостаток средствами сетевого уровня, то есть средствами протокола IP.

Бурное расширение сети Internet. В начале 1990-х годов сеть Internet расширялась очень быстро, новый узел появлялся в ней каждые 30 секунд, но 1995-й год стал переломным – перспективы коммерческого использования Internet стали отчетливыми и сделали ее развитие просто бурным. Первым следствием такого развития стало почти полное истощение адресного пространства Internet, определяемого полем адреса IP в четыре байта.

Новые стратегии администрирования. Расширение Internet связано с его проникновением в новые страны и новые отрасли промышленности. При этом в сети появляются новые органы администрирования, которые начинают использовать новые методы администрирования. Эти методы требуют появления новых средств в базовых протоколах стека TCP/IP.

Сообщество Internet уже несколько лет работает над разработкой новой спецификации для базового протокола стека – протокола IP. Выработано уже достаточно много предложений, от простых, предусматривающих только расширения адресного пространства IP, до очень сложных, приводящих к существенному увеличению стоимости реализации IP в высокопроизводительных (и так недешевых) маршрутизаторах.

Основным предложением по модернизации протокола IP является предложение, разработанное группой IETF. Сейчас принято называть ее предложение версией 6 – IPv6, а все остальные предложения группируются под названием *IP Next Generation, IPng*.

В предложении IETF протокол IPv6 оставляет основные принципы IPv4 неизменными. К ним относятся дейтаграммный метод работы, фрагментация пакетов, разрешение отправителю задавать максимальное число хопов для своих пакетов. Однако, в деталях реализации протокола IPv6 имеются существенные отличия от IPv4. Эти отличия коротко можно описать следующим образом.

Использование более длинных адресов. Новый размер адреса – наиболее заметное отличие IPv6 от IPv4. Версия 6 использует 128-битные адреса.

Гибкий формат заголовка. Вместо заголовка с фиксированными полями фиксированного размера (за исключением поля Резерв), IPv6 использует базовый заголовок фиксированного формата плюс набор необязательных заголовков различного формата.

Поддержка резервирования пропускной способности. В IPv6 механизм резервирования пропускной способности заменяет механизм классов сервиса версии IPv4.

Поддержка расширяемости протокола. Это одно из наиболее значительных изменений в подходе к построению протокола – от полностью детализированного описания протокола к протоколу, который разрешает поддержку дополнительных функций.

Адресация в IPv6. Адреса назначения и источника в IPv6 имеют длину 128 бит или 16 байт. Версия 6 обобщает специальные типы адресов версии 4 в следующих типах адресов:

Unicast – индивидуальный адрес. Определяет отдельный узел – компьютер или порт маршрутизатора. Пакет должен быть доставлен узлу по кратчайшему маршруту.

Cluster – адрес кластера. Обозначает группу узлов, которые имеют общий адресный префикс (например, присоединенных к одной физической сети). Пакет должен быть маршрутизирован

группе узлов по кратчайшему пути, а затем доставлен только одному из членов группы (например, ближайшему узлу).

Multicast – адрес набора узлов, возможно в различных физических сетях. Копии пакета должны быть доставлены каждому узлу набора, используя аппаратные возможности групповой или широковещательной доставки, если это возможно.

Как и в версии IPv4, адреса в версии IPv6 делятся на классы, в зависимости от значения нескольких старших бит адреса. Большая часть классов зарезервирована для будущего применения. Наиболее интересным для практического использования является класс, предназначенный для провайдеров услуг Internet, названный *Provider-Assigned Unicast*. Адрес этого класса имеет следующую структуру:

010	Идентификатор провайдера	Идентификатор абонента	Идентификатор подсети	Идентификатор узла
-----	--------------------------	------------------------	-----------------------	--------------------

Каждому провайдеру услуг Internet назначается уникальный идентификатор, которым помечаются все поддерживаемые им сети. Далее провайдер назначает своим абонентам уникальные идентификаторы и использует оба идентификатора при назначении блока адресов абонента. Абонент сам назначает уникальные идентификаторы своим подсетям и узлам этих сетей. Абонент может использовать технику подсетей, применяемую в версии IPv4, для дальнейшего деления поля идентификатора подсети на более мелкие поля.

Описанная схема приближает схему адресации IPv6 к схемам, используемым в территориальных сетях, таких, как телефонные сети или сети X.25. Иерархия адресных полей позволит магистральным маршрутизаторам работать только со старшими частями адреса, оставляя обработку менее значимых полей маршрутизаторам абонентов. Под поле идентификатора узла требуется выделения не менее 6 байт, для того чтобы можно было использовать в IP-адресах MAC-адреса локальных сетей непосредственно.

Для обеспечения совместимости со схемой адресации версии IPv4, в версии IPv6 имеется класс адресов, имеющих 0000 0000 в старших битах адреса. Младшие 4 байта адреса этого класса должны содержать адрес IPv4. Маршрутизаторы, поддерживающие обе версии адресов, должны обеспечивать трансляцию при передаче пакета из сети, поддерживающей адресацию IPv4, в сеть, поддерживающую адресацию IPv6, и наоборот.

Лекция 7. Системы подвижной связи

7.1. Назначение и классификация систем подвижной связи

Системы подвижной связи (СПС) предназначены для удовлетворения потребности абонентов в услугах связи, обеспечивая при этом им возможность свободного перемещения по определенной территории, называемой зоной обслуживания. Классификация СПС приведена на рис. 7.1.

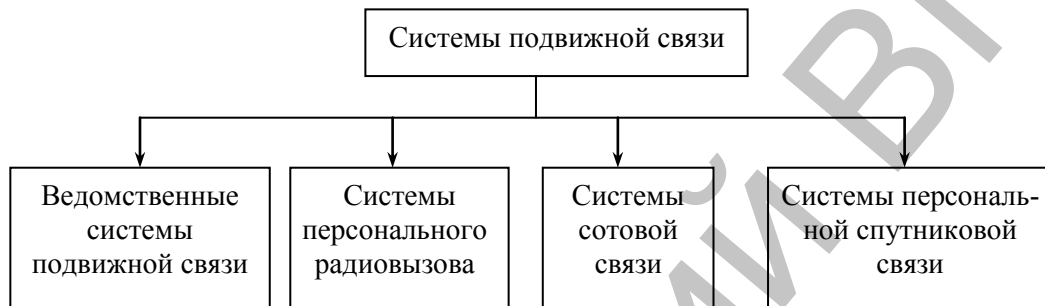


Рис. 7.1.

7.2. Ведомственные системы подвижной связи

Исторически сложилось так, что первыми в Советском Союзе стали использоваться ведомственные (профессиональные) системы подвижной связи. Это было связано с тем, что в условиях достаточно жестких ограничений на радиосвязь возможность ее применения для связи подвижными абонентами предоставлялась, как правило, государственным организациям (органам внутренних дел, пожарной охране, скорой помощи, такси и т.п.). Для них были разработаны и используются до настоящего времени такие комплексы оборудования радиосвязи, как «Лен», «Колос», «Гранит», «Виляя» и др. [11]. В последние годы развитие этих и им подобных систем было направлено на расширение функциональных возможностей, видов услуг, улучшение качественных характеристик и обеспечение конфиденциальности связи. Результатом предпринятых усилий явились такие комплексы, как «Маяк» и «Сапфир» [11]. Созданы отечественные цифровые и аналого-цифровые радиостанции «Альфа», «Риф» и др. с автоматическим поиском свободного канала связи, цифровой маскировкой и шифрованием передаваемой информации. Поскольку процедура поиска свободного канала называется «транкинг» (от англ. – trunking), ведомственные системы подвижной связи в литературе часто стали называть «транкинговыми», хотя это является некорректным, т.к. поиск

свободного канала используется и во многих других системах подвижной связи, например, в системах сотовой связи.

В ведомственных системах подвижной радиосвязи достигается наиболее полное использование выделенного диапазона частот по сравнению со всеми существовавшими ранее типами систем связи. Различают ведомственные системы с последовательным (сканирующим) поиском свободного канала связи и системы с выделенным каналом управления. Основным недостатком сканирующих систем является значительное время установления связи, поэтому их использование наиболее эффективно при небольшом количестве каналов (до 10). Более распространены системы с выделенным каналом управления, использующие аналоговые стандарты MPT1327, MPT1317, MPT1343, MPT1347. Кроме того, известны и аналоговые системы с совмещенным каналом управления, в которых сигналы управления передаются в полосе частот, расположенной ниже спектра речевого сигнала (полоса частот от 0 до 15 Гц). Такие системы были разработаны в США; им было присвоено обозначение LTR.

Системы подвижной связи обеспечивают своих абонентов качественной связью не только в пределах какого-либо отдельно взятого региона (города, области и т.п.), но и в глобальном масштабе (страна, континент). Такой режим работы называется *роумингом* (от англ. *roam* – скитаться, блуждать). Для организации роуминга необходимо, чтобы системы были одного стандарта или имели специальное оборудование, позволяющее абонентам систем разных стандартов связываться друг с другом.

По принципу организации связи различают три вида роуминга:

- ручной – простой обмен одного средства связи на другое;
- полуавтоматический, когда абоненту необходимо зарегистрироваться у местного оператора;
- автоматический, т.е. предоставление абоненту возможности выйти на связь «в любое время в любом месте».

Общей тенденцией развития ведомственных систем подвижной радиосвязи является переход от аналоговых стандартов к единым международным цифровым стандартам, обеспечивающим конфиденциальность и повышенное качество связи, более эффективное использование частотного диапазона, роуминг для всех абонентов и возможность передачи данных с высокой скоростью.

Первая цифровая транкинговая система EDACS (*Enhanced Digital Access Communications System*) была разработана и внедрена в скандинавских странах для обслуживания полиции, служб безопасности и при охране границ. Для этих же служб с

1997 г. стали применяться системы общеевропейского стандарта TETRA (*Trans European Trunked Radio*) – ETS 300.392, ETS 300.394, разработанные Европейским институтом стандартов связи (ETSI). Системы этого нового стандарта обеспечивают передачу речевых сообщений в цифровой форме и используют частотно-временное разделение каналов. Этим стандартом предусматривается опознавание абонента и организация прямой связи между абонентами без участия базовых станций. Передача данных происходит со скоростью до 28,8 кбит/с.

Основными требованиями, предъявляемыми абонентами и операторами к профессиональным системам подвижной связи, являются:

- обеспечение связи в заданной зоне обслуживания независимо от местоположения подвижных абонентов;
- возможность взаимодействия отдельных групп абонентов и организации циркулярной связи;
- оперативность управления связью, в том числе на различных уровнях;
- обеспечение связи через центры управления;
- возможность приоритетного установления каналов связи;
- низкие энергетические затраты подвижной станции;
- конфиденциальность разговоров.

7.3. Системы персонального радиовызова

Системы персонального радиовызова, или *пейджинговые* системы (от английского слова page – выкликать), обеспечивают одностороннюю (в последние годы – и двухстороннюю) передачу информации своим абонентам в пределах обслуживаемой зоны. Сети этой подвижной связи в России создаются на основе систем и средств, соответствующих международным стандартам, прежде всего принятым большинством стран Европейского сообщества.

Необходимость разработки и использования систем персонального радиовызова обусловлена тем, что до недавнего времени в различных отраслях производства, на транспорте и в сфере обслуживания между работниками, деятельность которых сопряжена с пребыванием на каких-либо объектах или с передвижением по городу, могла осуществляться только радиотелефонная связь. Сложность реализации такой связи определялась ограниченностью и занятостью диапазона радиочастот, громоздкостью и высокой стоимостью аппаратуры. Использование же систем персонального радиовызова позволяет избежать указанных трудностей и осуществить избирательный вызов по узкополосному

каналу любого из абонентов, свободно передвигающихся в пределах города и его окрестностей. При вызове, о котором сигнализирует миниатюрный абонентский приемник – пейджер, извещаемый абонент использует ближайший телефон для переговоров или получает необходимые инструкции в виде текстовой информации на дисплее своего пейджера. Применение систем персонального радиовызова в значительной мере сокращает потерю времени на поиски требуемого абонента,

Наряду с системами персонального радиовызова городского типа разработаны системы государственных и континентальных масштабов, использующие спутники.

Благодаря сравнительно низкой стоимости услуг телекоммуникационных операторов системы персонального радиовызова, рационально сочетающиеся с телефонной сетью, стали доступными для значительного числа абонентов. Еще сравнительно недавно они завоевали широкое признание во многих странах. Однако в последние годы в связи со значительным снижением цен на услуги операторов сотовой подвижной связи количество пользователей систем персонального радиовызова стало сокращаться.

7.4. Системы сотовой подвижной связи

Среди современных телекоммуникационных средств наиболее стремительно развиваются системы сотовой радиотелефонной связи. Их внедрение позволило решить проблему экономичного использования выделенной полосы радиочастот путем многократного использования одних и тех же частот. В отличие от ведомственных систем, которые создавались (и создаются) в интересах сравнительно небольшого числа абонентов, принадлежащих какой-нибудь организации, сотовые системы подвижной связи предназначены для массового обслуживания абонентов.

Повторное использование частот. Первая коммерческая система радиотелефонной связи была организована в 1946 году в г. Сент-Луис (США) [11]. За годы, прошедшие с тех пор, системы радиотелефонной связи претерпели существенные изменения. Уменьшение габаритов устройств, а также снижение стоимости услуг данного типа связи привели к небывалому росту спроса на них среди населения. Однако широкому распространению систем радиотелефонной связи препятствовал дефицит частотных ресурсов: увеличение количества фиксированных частот, на которых осуществлялась передача, в определенном ограниченном частотном диапазоне приводило к возникновению взаимных помех между радиотелефонами с близкими по частоте рабочими каналами.

Удачное решение, позволившее решить указанную проблему, было предложено исследовательским центром Bell Laboratories американской компании AT&T [11]. Оно состояло в том, чтобы разбить всю обслуживаемую территорию на небольшие участки, называемые сотами. Нужно отметить, что для того, чтобы обеспечить равномерное покрытие какой-нибудь зоны обслуживания без перекрытий соседних ячеек или без пропусков территории между ячейками, где обслуживание клиентов будет отсутствовать, подходят ячейки трех типов: треугольные, квадратные и шестиугольные (рис. 7.2). На этапе разработки систем подвижной связи выбор пал на шестиугольные ячейки, поскольку диаграммы направленности антенн в первых системах сотовой связи имели форму круга, а из перечисленных выше 3-х геометрических фигур именно шестиугольник больше всего соответствует этой форме.

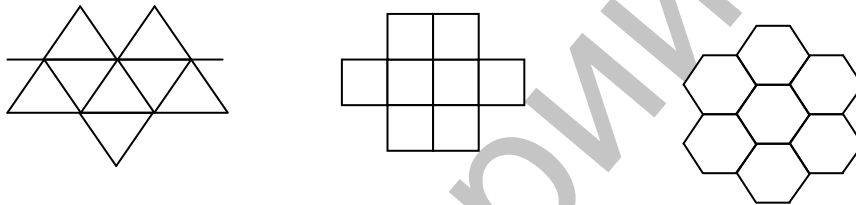


Рис. 7.2. Возможные способы разбиения зоны обслуживания на ячейки.

Каждая сота должна обслуживаться передатчиком с ограниченным радиусом действия и фиксированной частотой. Это позволяет без взаимных помех многократно использовать ту же самую частоту в других ячейках (сотах), удаленных на определенное расстояние D , называемое защитным интервалом (рис. 7.3).

В данном примере для организации связи в зоне обслуживания использовано семь несущих частот F_1-F_7 . В общем случае их число может равняться C . Если на каждой несущей частоте сформирован набор из m каналов с шириной полосы частот каждого канала F_k , то общая ширина полосы частот F_c , занимаемая системой сотовой связи

$$F_c = F_k \cdot m \cdot C. \quad (7.1)$$

Количество абонентов N , которым может предоставить услуги данная система сотовой связи, составляет:

$$N = m \cdot n, \quad (7.2)$$

где n – количество ячеек, используемых в зоне обслуживания системы подвижной связи.

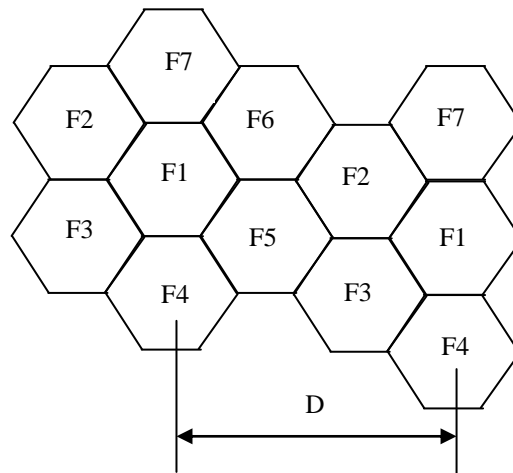


Рис. 7.3. Повторное использование частот в несмежных сотах.

Из (7.1) и (7.2) следует, что при ограниченной полосе частот, занимаемой системой сотовой подвижной связи, количество обслуживаемых абонентов может возрасти за счет увеличения количества ячеек в зоне обслуживания.

Состав системы сотовой подвижной связи. Каждая из сот обслуживается многоканальным приемопередатчиком, называемым базовой станцией (БС). Число каналов базовой станции обычно кратно 8, например, 8, 16, 32... Один из каналов является управляющим (*control channel*). В некоторых ситуациях он может называться также каналом вызова (*calling channel*). На этом канале происходит непосредственное установление соединения при вызове подвижного абонента сети, а сам разговор начинается только после того, как будет найден свободный в данный момент канал и произойдет переключение на него. Любой из каналов сотовой связи представляет собой пару частот для дуплексной связи, т.е. частоты базовой и подвижной станций разнесены. Это делается для того, чтобы улучшить фильтрацию сигналов и исключить взаимное влияние передатчика на приемник одного и того же устройства при их одновременной работе.

Все базовые станции соединены с центром коммутации подвижной связи (коммутатором) по выделенным проводным или радиорелейным каналам связи. *Центр коммутации MSC* – это автоматическая телефонная станция системы сотовой связи, обеспечивающая все функции управления сетью. Она организует их эстафетную передачу, в процессе которой достигается непрерывность связи при перемещении подвижной станции из соты в соту и переключение рабочих каналов в соте при появлении помех или неисправностей, производит соединение подвижного абонента с абонентом телефонной сети общего пользования и др.

Алгоритмы функционирования систем сотовой связи. Несмотря на разнообразие стандартов сотовой связи, алгоритмы их функционирования, независимо от имеющихся особенностей, в основном сходны. Для абонента практически нет никакой разницы, в каком стандарте осуществляется связь. Если ему нужно позвонить, то он просто нажимает соответствующую клавишу на своем радиотелефоне (подвижной абонентской станции), что соответствует снятию трубки обычного телефона. Когда же подвижная абонентская станция находится в режиме ожидания (состояние «трубка положена» обычного телефона), ее приемное устройство постоянно сканирует (просматривает) либо все каналы системы, либо только управляющие.

Вызов абонента. Для вызова соответствующего абонента всеми базовыми станциями сотовой системы связи по управляющим каналам передается сигнал вызова. Радиотелефон вызываемого абонента при получении этого сигнала отвечает по одному из свободных каналов управления. Базовые станции, принявшие ответный сигнал, передают информацию о его параметрах в центр коммутации, который, в свою очередь, переключает разговор на ту базовую станцию, где зафиксирован максимальный уровень сигнала сотового радиотелефона вызываемого абонента. Радиотелефон вызываемого абонента занимает один из свободных каналов в этой ячейке.

Эстафетная передача. По мере удаления абонента от базовой станции уровень сигнала уменьшается, что ведет к ухудшению качества связи. Поддержание необходимого качества связи достигается путем автоматического переключения абонента на обслуживание базовой станцией соседней ячейки. Это происходит следующим образом. Для контроля таких ситуаций каждая базовая станция снабжена специальным приемником, периодически измеряющим уровень сигнала сотового телефона разговаривающего абонента и сравнивающим его с допустимым пределом. Когда уровень сигнала становится меньше этого предела, то информация об этом автоматически передается в центр коммутации по служебному каналу связи. Центр коммутации выдает команду об измерении уровня сигнала сотового радиотелефона абонента на ближайшие к нему базовые станции. После получения информации от базовых станций об уровне этого сигнала центр коммутации переключает радиотелефон на ту из них, где уровень сигнала оказался наибольшим. Это происходит сравнительно быстро, поэтому абонент может не заметить этих переключений.

Аналогичные действия предпринимаются при снижении качества связи из-за влияния помех, в связи с ухудшением усло-

вий распространения радиоволн или при возникновении неисправностей коммутационного оборудования. В такой ситуации разговор переключается на свободный канал базовой станции, в зоне действия которой находится в это время абонент.

Роуминг. Одна из важных услуг сети сотовой связи – предоставление возможности использования одного и того же радиотелефона при поездке в другой город, область или даже страну, причем сотовая сеть позволяет не только самому абоненту звонить из другого города или страны, но и получать звонки. Для организации роуминга применительно к сотовым системам необходимо выполнение трех условий:

- наличие в требуемых регионах сотовых систем стандарта, совместимого со стандартом компании, у которой был приобретен радиотелефон;
- наличие соответствующих организационных и экономических соглашений о роуминговом обслуживании абонентов между различными операторами сотовой связи;
- наличие каналов связи, по которым будут передаваться переадресуемые звонки.

Аналоговые системы сотовой подвижной связи. Аналоговые системы сотовой подвижной связи принадлежат к первому поколению сотовых систем и используют для передачи информации аналоговые сигналы с частотной модуляцией. В этих системах сотовой подвижной связи, так же, как и в любых других радиосистемах, использующих аналоговые сигналы, применяется частотное разделение каналов.

Аналоговые системы обеспечивают:

- организацию связи между подвижными абонентами и абонентами телефонной сети общего пользования;
- регистрацию стоимости разговора;
- возможность автоматического поиска подвижного абонента (в системах стандарта NMT-450 и NMT-900).

В целом для всех аналоговых систем характерны общие недостатки:

- низкое качество связи;
- несовместимость стандартов;
- невозможность взаимодействия с сетью ISDN и сетями пакетной передачи данных;
- возможность прослушивания разговоров;
- отсутствие методов борьбы с замираниями сигналов, обусловленными влиянием окружающего ландшафта и зданий или вследствие передвижения абонентов.

Сравнительные характеристики систем сотовой связи основных аналоговых стандартов представлены в табл. 7.1.

Таблица 7.1

Характеристика	Наименование стандарта						
	NMT-450	NMT-900	TACS (ETACS)	AMPS	C-450	HCMTS	Radio-com-200
Страна-разработчик	Скандинавские страны	Скандинавские страны	Англия	США	Германия	Япония	Франция
Год ввода в эксплуатацию	1981	1986	1985/1987	1983	1985	1979	1985
Полоса частот, МГц: базовая станция	463–467,5	935–960	935–950 (917–933)	870–890	461,3–465,74	870–885	424,8–427,9
подвижная станция	453–457,5	890–915	890–905 (872–888)	825–845	451,3–455,74	925–940	418,8–421,9
Общее число дуплексных каналов	180	1000 (1999)	600 (640)	666	222	600 (1200)	256
Максимальная мощность БС, Вт	50	100	100 (100)	100	100	50/25	70
Номинальная мощность передатчика ПС, Вт	0,15–15	0,1–6	2–20 (2–20)	3	15	1–5	11
Типичный радиус соты, км	1–40	0,5–20	2–20	2–20	2–30	5–10	20
Время переключения каналов на границе сот, мс	1250	270	290	250	300	800	–

Из приведенных в табл. 7.1 стандартов аналоговых систем в России используются два: NMT-450 и AMPS, причем стандарт NMT-450 утвержден в качестве федерального. Учитывая отмеченные выше недостатки аналоговых систем, в настоящее время во всем мире наблюдается снижение роста количества их пользователей. Кроме того, в стадии проработки находится проект по замене сетей стандарта NMT-450 на сети цифрового стандарта, работающие в близком диапазоне частот (GSM-400 и CDMA-450).

Цифровые системы сотовой подвижной связи. Цифровые системы сотовой подвижной связи представляют собой системы второго поколения. По сравнению с аналоговыми системами они

предоставляют абонентам большой набор услуг и обеспечивают повышенное качество связи, а также взаимодействие с цифровыми сетями с интеграцией служб (ISDN) и пакетной передачи данных (PDN). Среди этих систем широкое распространение получили те, которые базируются на стандартах GSM (DCS 1800), D-AMPS (ADC), JDC, CDMA. Сравнительные характеристики стандартов представлены в табл. 7.2.

Таблица 7.2

Характеристика	GSM (DCS1800)	D-AMPS (ADC)	JDC	CDMA
Метод доступа	TDMA	TDMA	TDMA	CDMA
Количество речевых каналов	8	3	3	32
Рабочий диапазон частот, МГц	935–960; 890–915	824–840 869–894	810–826; 940–956; 1429–1441; 1447–1489; 1501–1513	824–840 869–894
Разнос каналов, кГц	200	30	25	1250
Экв. полоса частот на один разговорный канал, кГц	25	10	8,3	–
Вид модуляции	0,3 GMSK	n/4 DQPSK	n/4 DQPSK	QPSK
Скорость передачи информации, Кбит/с	270	48	42	

Раздел II. ТЕСТОВЫЕ ЗАДАНИЯ

Тест к лекции 1. Элементы теории передачи информации

1. Совокупность знаков, которые используются для хранения, передачи и преобразования информации, называют:
 - a) *сообщение;*
 - b) *сигнал;*
 - c) *код.*
2. Различают следующие виды сообщений:
 - a) *продолжительные;*
 - b) *дискретные;*
 - c) *непрерывные;*
 - d) *непродолжительные.*
3. Множество различных знаков называют:
 - a) *объемом алфавита;*
 - b) *алфавитом источника сообщений;*
 - c) *алфавитом.*
4. Сопоставьте термины:
 - a) *непрерывные по уровню и по времени;* 1) аналоговые;
 - b) *дискретные по уровню и непрерывные по времени;* 2) цифровые;
 - c) *дискретные по уровню и по времени.* 3) квантованные по уровню.
5. Системой связи (телекоммуникационной системой) называют:
 - a) *совокупность технических средств, предназначенных для передачи информации, включая источник информации и получателя информации;*
 - b) *совокупность технических средств, предназначенных для передачи, обработки и хранения информации, включая источник информации и получателя информации;*
 - c) *совокупность технических средств, предназначенных для передачи информации, не включая источник информации и получателя информации.*
6. Среда, используемая для передачи сигналов, – это:
...
7. В зависимости от характера изменения помех во времени различают:
 - a) *флуктуационные;*
 - b) *узкополосные;*
 - c) *широкополосные;*

d) импульсные;

e) дискретные.

8. Совокупность технических средств передачи информации, включающая среду распространения и обеспечивающая передачу сигнала – это:

...

9. Число элементарных сигналов (символов), передаваемых по каналу в единицу времени – это:

...

10. Выберите неверное утверждение:

a) пропускная способность канала равна максимальной скорости передачи информации по данному каналу;

b) пропускная способность канала измеряется числом двоичных единиц информации в секунд – бит/сек;

c) скорость передачи информации измеряется числом двоичных единиц информации в секунду – бит/сек;

d) пропускная способность канала не является характеристикой его самого, а зависит от статистики сигнала.

Тест к лекции 2. Цифровые методы обработки и передачи звуковых сообщений

1. Устройство, осуществляющее преобразование первичного речевого сигнала в цифровую форму – это:

...

2. Процедура преобразования речевых сигналов, при которой уменьшается скорость цифрового потока – это:

...

3. Речевые кодеры подразделяются на группы:

a) кодеры формы;

b) вокодеры;

c) кодеры звука;

d) гибридные кодеры;

e) видеокодеры.

4. Нелинейное устройство с амплитудной характеристикой, обратной характеристике компрессора – это:

...

5. Человеческий голос порождает первичный аналоговый сигнал, который занимает полосу частот:

a) 10...10000 Гц;

b) 50...20000 Гц;

c) 100...10000 Гц;

d) 50...10000 Гц.

6. Какого метода кодирования не существует:

- a) ДИКМ;
- b) ИКМ;
- c) ДИКМК;
- d) АДИКМ.

Тест к лекции 3. Цифровые методы передачи видеоизображений

1. Что не является основной характеристикой цифрового видео:

- a) глубина цвета;
- b) качество изображения;
- c) скорость развертки;
- d) разрешение;

2. Стандартная скорость воспроизведения видеосигнала в телевизионных системах:

- a) 15 кадров/с;
- b) 24 кадра/с;
- c) 25 кадров/с;
- d) 30 кадров/с.

3. Какой модели цветности для видеоизображений не существует:

- a) RGB;
- b) YUO;
- c) YUV.

4. Избыточность изображения проявляется в высокой степени взаимной статистической прогнозируемости элементов изображения. В радиотехнике такую связь сигналов характеризуют понятием:

...

5. Различают следующие виды сжатия (выберите верные ответы):

- a) сжатие в режиме реального времени;
- b) симметричное и асимметричное сжатие;
- c) сплошное и дискретное сжатие;
- d) сжатие видеопотока или пок кадровое сжатие;
- e) сжатие с частичной потерей изображения.

6. Какого формата цифрового видео не существует:

- a) Motion-JPEG;
- b) MPEG-6;
- c) Audio Video Interleave.

7. Какой тип кадров не содержит MPEG-последовательность:

- a) Intro frames (кадры типа «I»);
- b) Predicted frames (кадры типа «P»);
- c) Si-directional interpolated frames (кадры типа «S»).

Тест к лекции 4. Телекоммуникационные каналы

1. Выберите верное утверждение:
 - a) *затухание сигнала в кабеле на витой паре сильно зависит от частоты передаваемого сигнала;*
 - b) *затухание сигнала в кабеле на витой паре не зависит от частоты передаваемого сигнала;*
 - c) *затухание сигнала в кабеле на витой паре обратно пропорционально частоте передаваемого сигнала.*
2. Характеристики кабелей на витой паре регламентируются в стандарте EIA/TIA 568. Сколько категорий неэкранированных кабелей на витой паре включает в себя этот стандарт?
...
3. Каким соотношением связаны углы падения и преломления согласно закону Снеллиуса (n_1 -показатель преломления среды преломленного луча, n_2 – падающего):
 - a) $n_1 \sin \theta_{пад} = n_2 \sin \theta_{плм}$;
 - b) $n_2 \sin \theta_{пад} = n_1 \sin \theta_{плм}$;
 - c) $n_2 \sin \theta_{пад} = n_1 \cos \theta_{плм}$.
4. Полоса частот, используемых в радиосвязи, составляет:
 - a) *2...2000 МГц;*
 - b) *5...5000 ГГц;*
 - c) *3...3000 ГГц.*
5. Принцип временного разделения каналов (ВРК) состоит в том, что:
 - a) *групповой тракт предоставляется поочередно для передачи сигналов каждого канала многоканальной системы;*
 - b) *групповой тракт предоставляется одновременно для передачи всех многоканальной системы.*
6. Какой принцип лежит в основе построения систем многоканальной связи с временным разделением каналов и системы многоканальной связи с частотным разделением каналов?
...
7. Появившаяся исторически первой плезиохронная цифровая иерархия имеет разновидности:
 - a) *Европейскую;*
 - b) *Азиатскую;*
 - c) *Японскую;*
 - d) *Южно-американскую;*
 - e) *Северо-американскую.*

Тест к главе 5. Сети ЭВМ

1. Сеть, объединяющая два или более компьютеров, с целью совместного использования программных и/или аппаратных ресурсов – это:
...
2. ЛВС подразделяются на:
 - a) одноранговые;
 - b) многоранговые;
 - c) сети на основе выделенного сервера;
 - d) администрируемые;
 - e) не администрируемые.
3. Какие из перечисленных уровней эталонной модели OSI являются сетезависимыми:
 - a) физический;
 - b) сетевой;
 - c) сеансовый;
 - d) канальный;
 - e) прикладной.
4. Какие из перечисленных уровней эталонной модели OSI являются сетезависимыми:
 - a) прикладной;
 - b) уровень представления;
 - c) канальный;
 - d) транспортный;
 - e) сеансовый;
 - f) физический.
5. Отметьте существующие топологии вычислительной сети:
 - a) «Звезда»;
 - b) «Круг»;
 - c) «Общая шина»;
 - d) иерархическая;
 - e) последовательная.
6. Какую топология используют сети стандарта Ethernet:
 - a) «Круг»;
 - b) «Иерархическая»;
 - c) «Общая шина»;
 - d) «Кольцо».
7. Если передаваемые и наблюдаемые сигналы отличаются, то фиксируется:
...
8. Стандарты технологии Ethernet:
 - a) 10Base-3;
 - b) 10Base-5;

- c) 10Base-2;
- d) 10Base-H;
- e) 10Base-T.

9. Сегменты, соединенные повторителями, образуют один (продолжите фразу)

...

10. Как называется метод, в котором право на доступ к среде передается циклически от станции к станции по логическому кольцу?

...

11. В Token Ring существует три различных формата кадров: маркер; кадр данных и

...

Тест к лекции 6. Глобальная информационная сеть Интернет

1. Адрес устройства, подключенного к сети Интернет, называется IP-адресом и имеет размер:

- a) 8 бита;
- b) 16 бита;
- c) 24 бита;
- d) 32 бита.

2. На какое кол-во классов делятся все IP-адреса:

...

3. Что является доменом верхнего уровня из трех элементов имени: *www.vsu.by*?

- a) *www*;
- b) *vsu*;
- c) *by*.

4. Сопоставьте различные классы организаций с соответствующими им аббревиатурами:

- | | |
|-----------------|--------------------------------------|
| a) <i>com</i> ; | 1) коммерческие организации; |
| b) <i>edu</i> ; | 2) международные организации; |
| c) <i>int</i> ; | 3) образовательные; |
| d) <i>org</i> ; | 4) организации, поддерживающие сети; |
| e) <i>net</i> . | 5) некоммерческие организации. |

5. Сопоставьте названия протоколов стека TCP/IP с их назначением:

- | | |
|--------------------|-------------------------------------|
| a) <i>FTP</i> ; | 1) пересылка файлов; |
| b) <i>Telnet</i> ; | 2) почтовый протокол; |
| c) <i>SMTP</i> ; | 3) эмуляция терминала; |
| d) <i>SNMP</i> . | 4) организации сетевого управления. |

6. Укажите различия протоколов IPv6 и IPv4:
- a) *фрагментация пакетов;*
 - b) *дейтаграммный метод работы;*
 - c) *использование более длинных адресов;*
 - d) *поддержка расширяемости протокола.*

Тест к лекции 7. Системы подвижной связи

1. На что НЕ подразделяются СПС:
- a) *ведомственные СПС;*
 - b) *системы сотовой связи;*
 - c) *системы персональной спутниковой связи;*
 - d) *системы глобальной спутниковой связи.*
2. Как называется процедура поиска свободного канала?
...
3. Как называется режим работы, при котором системы подвижной связи обеспечивают своих абонентов качественной связью не только в пределах какого-либо отдельно взятого региона, но и в глобальном масштабе?
...
4. Какого вида роуминга не существует:
- a) *машинный;*
 - b) *ручной;*
 - c) *автоматический.*
5. Как по-другому можно назвать системы персонального радиовызова?
...
6. В каком году была организована первая коммерческая система радиотелефонной связи:
- a) *1945 г. ;*
 - b) *1946 г. ;*
 - c) *1955 г. ;*
 - d) *1960 г.*
7. Что обеспечивают аналоговые системы:
- a) *регистрацию стоимости вызова;*
 - b) *высокое качество связи;*
 - c) *возможность автоматического поиска подвижного абонента;*
 - d) *совместимость стандартов.*

Раздел III. Лабораторные работы

1. Расчет конфигурации сети Ethernet

Цель работы: изучение принципов построения сетей по стандарту Ethernet и приобретение практических навыков оценки корректности их конфигурации.

Необходимое оборудование: калькулятор.

1.1. Критерии корректности конфигурации

Соблюдение многочисленных ограничений, установленных для различных стандартов физического уровня сетей Ethernet, гарантирует корректную работу сети (естественно, при исправном состоянии всех элементов физического уровня). Основные характеристики и ограничения технологии Ethernet приведены в табл. 1.1 и 1.2.

Наиболее часто приходится проверять ограничения, связанные с длиной отдельного сегмента кабеля, а также количеством повторителей и общей длиной сети.

Правила «5-4-3» для коаксиальных сетей и «4 хабов» для сетей на основе витой пары и оптоволокна не только дают гарантии работоспособности сети, но и оставляют большой «запас прочности» сети. Например, если посчитать время двойного оборота в сети, состоящей из 4 повторителей 10Base-5 и 5 сегментов максимальной длины 500 м, то окажется, что оно составляет 537 битовых интервала. А так как время передачи кадра минимальной длины (вместе с преамбулой), составляющей 72 байт, равно 575 битовым интервалам, то видно, что разработчики стандарта Ethernet оставили 38 битовых интервала в качестве запаса для обеспечения надежности. Тем не менее, в документах комитета IEEE 802.3 утверждается, что и 4 дополнительных битовых интервала создают достаточный запас надежности.

Комитет IEEE 802.3 приводит исходные данные о задержках (табл. 1.3 и 1.4), вносимых повторителями и различными средами передачи данных, для тех специалистов, которые хотят самостоятельно рассчитывать максимальное количество повторителей и максимальную общую длину сети, не довольствуясь теми значениями, которые приведены в правилах «5-4-3» и «4 хабов». Особенно такие расчеты полезны для сетей, состоящих из смешанных кабельных систем, например, коаксиала и оптоволокна, на которые правила о количестве повторителей не рассчитаны. При этом максимальная длина каждого отдельного физического

сегмента должна строго соответствовать стандарту, то есть 500 м для «толстого» коаксиала, 100 м для витой пары и т.д.

Чтобы сеть Ethernet, состоящая из сегментов различной физической природы, работала корректно, необходимо выполнение четырех основных условий:

- количество станций в сети – не более 1024;
- максимальная длина каждого физического сегмента – не более величины, определенной в соответствующем стандарте физического уровня;
- время двойного оборота сигнала (Path Delay Value, PDV) между двумя самыми удаленными друг от друга станциями сети – не более 575 битовых интервала;
- сокращение межкадрового интервала (Path Variability Value, PVV) при прохождении последовательности кадров через все повторители – не больше, чем 49 битовых интервала (т.к. при отправке кадров конечные узлы обеспечивают начальное межкадровое расстояние в 96 битовых интервала, то после прохождения повторителя оно должно быть не меньше, чем $96 - 49 = 47$ битовых интервала).

Соблюдение этих требований обеспечивает корректность работы сети даже в случаях, когда нарушаются простые правила конфигурирования, определяющие максимальное количество повторителей и общую длину сети в 2500 м.

1.2. Методика расчета времени двойного оборота и уменьшения межкадрового интервала

Для упрощения расчетов обычно используются справочные данные IEEE, содержащие значения задержек распространения сигналов в повторителях, приемопередатчиках и различных физических средах (табл. 1.3). Битовый интервал обозначен как bt.

Комитет 802.3 старался максимально упростить выполнение расчетов, поэтому данные, приведенные в таблице, включают сразу несколько этапов прохождения сигнала. Например, задержки, вносимые повторителем, состоят из задержки входного трансивера, задержки блока повторения и задержки выходного трансивера. Тем не менее, в таблице все эти задержки представлены одной величиной, названной базой сегмента.

Чтобы не нужно было два раза складывать задержки, вносимые кабелем, в таблице даются удвоенные величины задержек для каждого типа кабеля.

В таблице используются также такие понятия, как левый сегмент, правый сегмент и промежуточный сегмент. Поясним

эти термины на примере сети, приведенной на рис. 1.1. Левым сегментом называется сегмент, в котором начинается путь сигнала от выхода передатчика конечного узла. На рис. 1.1 это сегмент 1. Затем сигнал проходит через промежуточные сегменты 2–5 и доходит до приемника наиболее удаленного узла наиболее удаленного сегмента 6, который называется правым. Именно здесь в худшем случае происходит столкновение кадров и возникает коллизия.

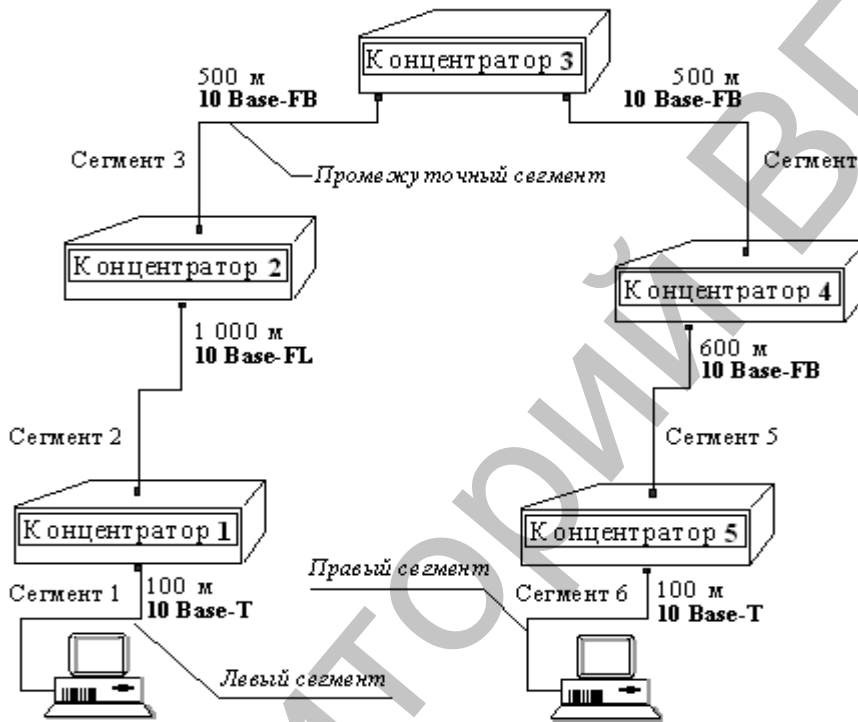


Рис. 1.1. Пример сети Ethernet, состоящей из сегментов различных физических стандартов.

С каждым сегментом связана постоянная задержка, названная базой, которая зависит только от типа сегмента и от положения сегмента на пути сигнала (левый, промежуточный или правый). База правого сегмента, в котором возникает коллизия, намного превышает базу левого и промежуточных сегментов.

Кроме этого, с каждым сегментом связана задержка распространения сигнала вдоль кабеля сегмента, которая зависит от длины сегмента и вычисляется путем умножения времени распространения сигнала по одному метру кабеля (в битовых интервалах) на длину кабеля в метрах.

Расчет PDV заключается в вычислении задержек, вносимых каждым отрезком кабеля (приведенная в таблице задержка сигнала на 1 м кабеля умножается на длину сегмента), а затем суммировании этих задержек с базами левого, промежуточных и правого сегментов. Общее значение PDV не должно превышать 575.

Так как левый и правый сегменты имеют разные величины базовой задержки, то в случае различных типов сегментов на удаленных краях сети необходимо выполнить расчеты дважды: один раз принять в качестве левого сегмента сегмент одного типа, а во второй – сегмент другого типа. Результатом можно считать максимальное значение PDV.

Чтобы признать конфигурацию сети корректной, нужно рассчитать также уменьшение межкадрового интервала повторителями, то есть величину PVV.

Для расчета PVV также можно воспользоваться значениями максимальных величин уменьшения межкадрового интервала при прохождении повторителей различных физических сред, рекомендованными IEEE и приведенными в табл. 1.4.

1.3. Пример расчета конфигурации сети

В примере крайние сегменты сети принадлежат к одному типу – стандарту 10Base-T, поэтому двойной расчет не требуется.

Приведенная на рис. 1.1 сеть в соответствии с правилом «4 хабов» не является корректной – в сети между узлами сегментов 1 и 6 имеются 5 хабов, хотя не все сегменты являются сегментами 10Base-FB. Кроме того, общая длина сети равна 2800 м, что нарушает правило 2500 м. Рассчитаем значение PDV.

Левый сегмент 1:

$$15,3 \text{ (база)} + 100 \cdot 0,113 = 26,6.$$

Промежуточный сегмент 2:

$$33,5 + 1000 \cdot 0,1 = 133,5.$$

Промежуточный сегмент 3:

$$24 + 500 \cdot 0,1 = 74,0.$$

Промежуточный сегмент 4:

$$24 + 500 \cdot 0,1 = 74,0.$$

Промежуточный сегмент 5:

$$24 + 600 \cdot 0,1 = 84,0.$$

Правый сегмент 6:

$$165 + 100 \cdot 0,113 = 176,3.$$

Сумма всех составляющих дает значение PDV, равное 568,4.

Так как значение PDV меньше максимально допустимой величины 575, то эта сеть проходит по критерию времени двойного оборота сигнала несмотря на то, что ее общая длина превышает 2500 м, а количество повторителей больше 4.

Рассчитаем значение PVV.

Левый сегмент 1 10Base-T: сокращение в 10,5 bt.

Промежуточный сегмент 2 10Base-FL: 8.

Промежуточный сегмент 3 10Base-FB: 2.

Промежуточный сегмент 4 10Base-FB: 2.

Промежуточный сегмент 5 10Base-FB: 2.

Сумма этих величин дает значение PVV, равное 24,5, что меньше предельного значения в 49 битовых интервала.

В результате сеть соответствует стандартам Ethernet по всем параметрам.

1.4. Задание на лабораторную работу

1. Ознакомиться с теоретическим материалом.

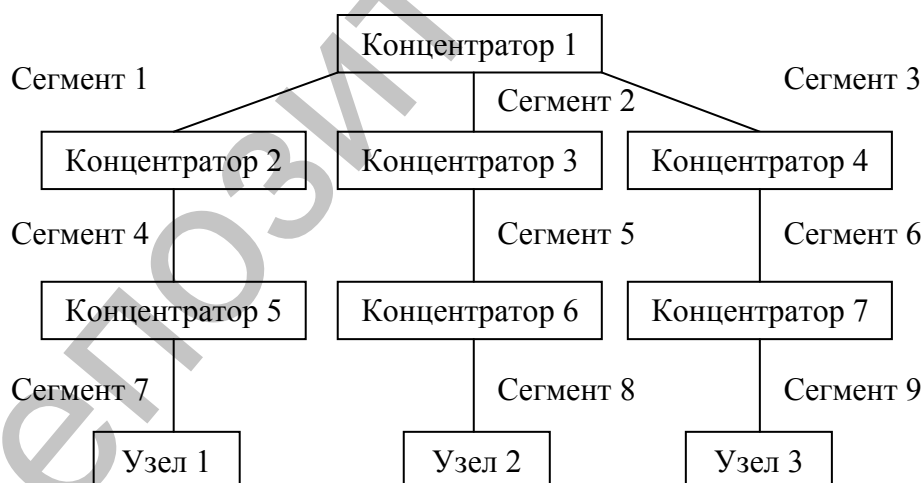
2. Произвести оценку конфигурации сети в соответствии с вариантом:

- по физическим ограничениям: на длину сегмента, на длину сети, правило «4 хаба» («5 хабов» для 10Base-FB);
- по времени двойного оборота сигнала в сети;
- по уменьшению межкадрового интервала.

3. По результатам расчетов сделать вывод о корректности конфигурации сети Ethernet.

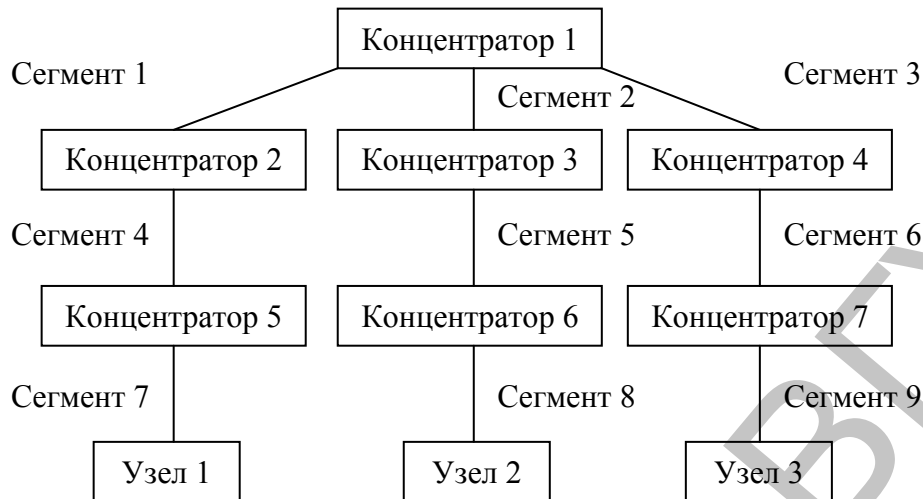
4. По результатам работы оформить отчет. Содержание отчета: исходные данные, расчеты указанных параметров, выводы.

Вариант 1



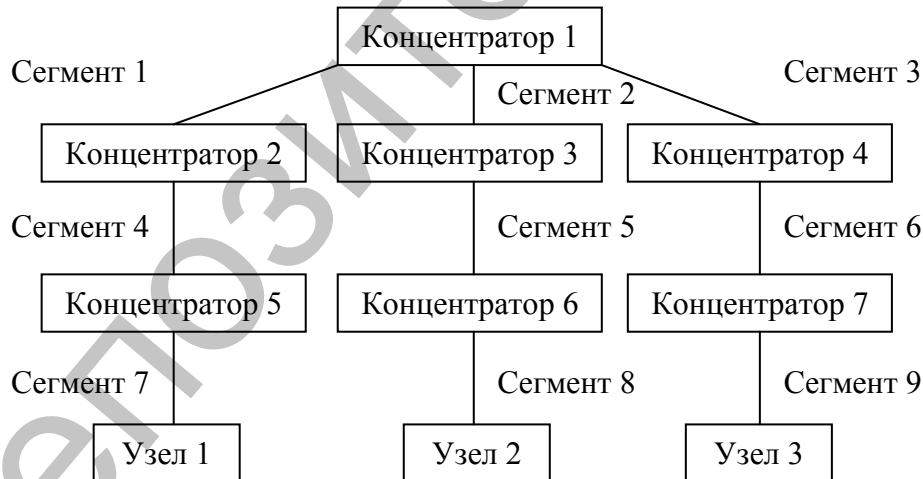
	10Base-FB	10Base-FL	10Base-T	Длина, м
Сегмент 1	+			500
Сегмент 2	+			300
Сегмент 3	+			400
Сегмент 4		+		1000
Сегмент 5		+		300
Сегмент 6		+		400
Сегмент 7			+	100
Сегмент 8			+	50
Сегмент 9			+	100

Вариант 2



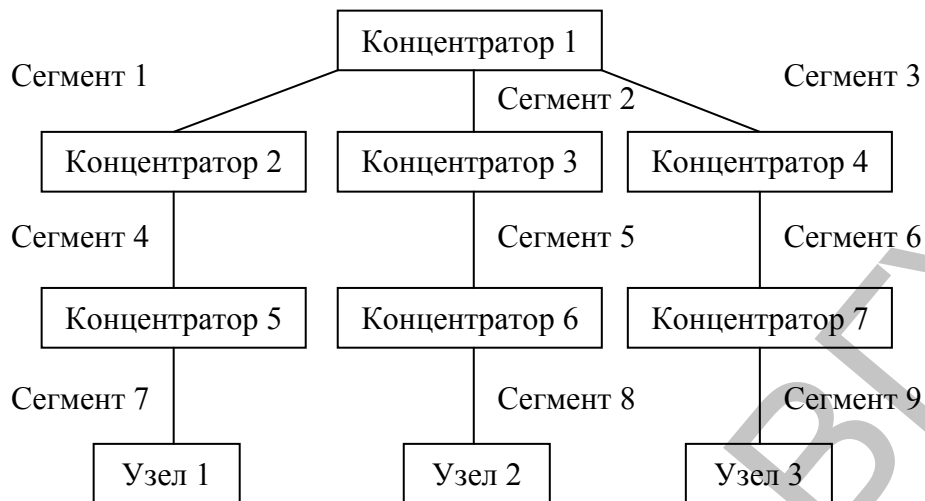
	10Base-FB	10Base-FL	10Base-T	Длина, м
Сегмент 1		+		700
Сегмент 2	+			400
Сегмент 3	+			400
Сегмент 4		+		700
Сегмент 5		+		200
Сегмент 6	+			500
Сегмент 7			+	80
Сегмент 8			+	100
Сегмент 9			+	80

Вариант 3



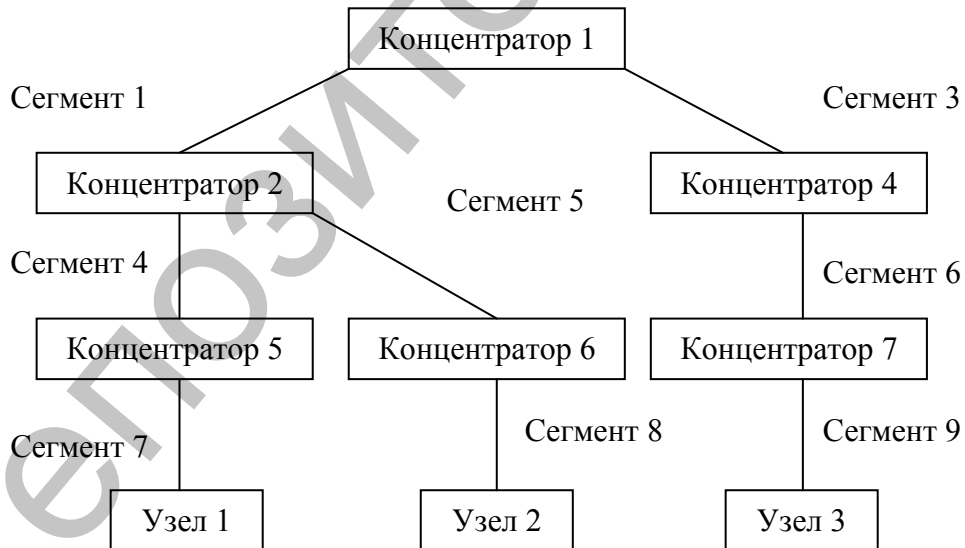
	10Base-FB	10Base-FL	10Base-T	Длина, м
Сегмент 1	+			1000
Сегмент 2		+		200
Сегмент 3		+		200
Сегмент 4		+		400
Сегмент 5	+			300
Сегмент 6		+		200
Сегмент 7			+	100
Сегмент 8			+	100
Сегмент 9			+	40

Вариант 4



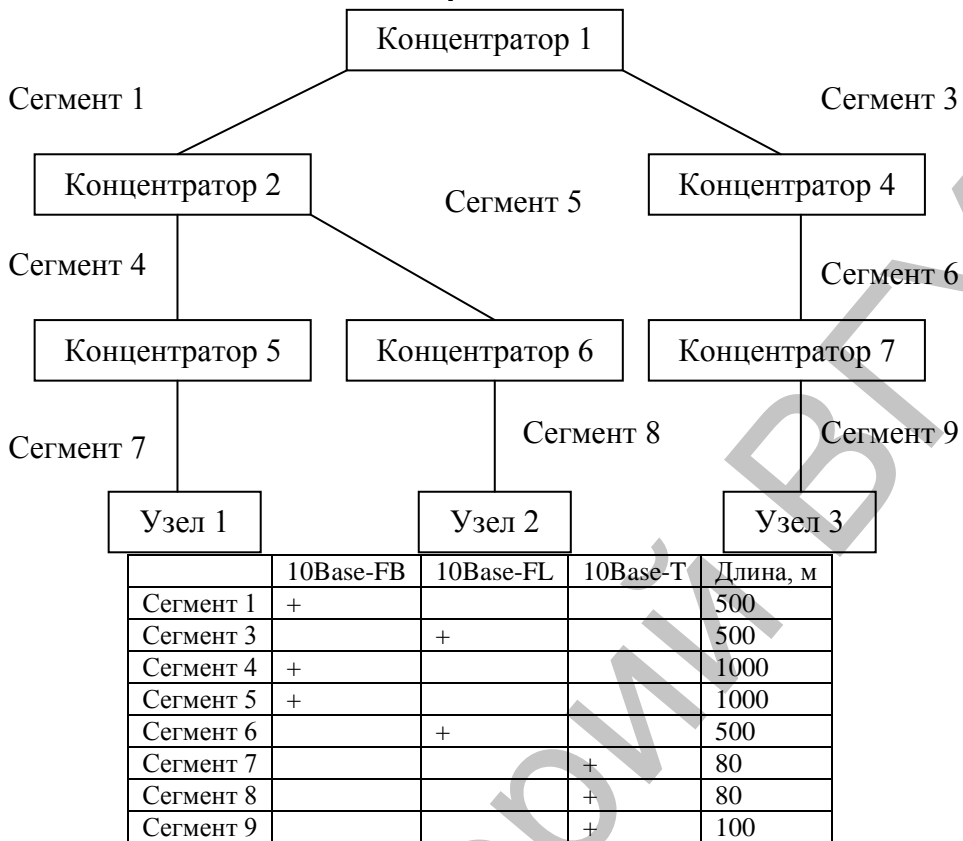
	10Base-FB	10Base-FL	10Base-T	Длина, м
Сегмент 1		+		600
Сегмент 2		+		400
Сегмент 3		+		200
Сегмент 4	+			800
Сегмент 5	+			500
Сегмент 6	+			800
Сегмент 7			+	50
Сегмент 8			+	100
Сегмент 9			+	50

Вариант 5

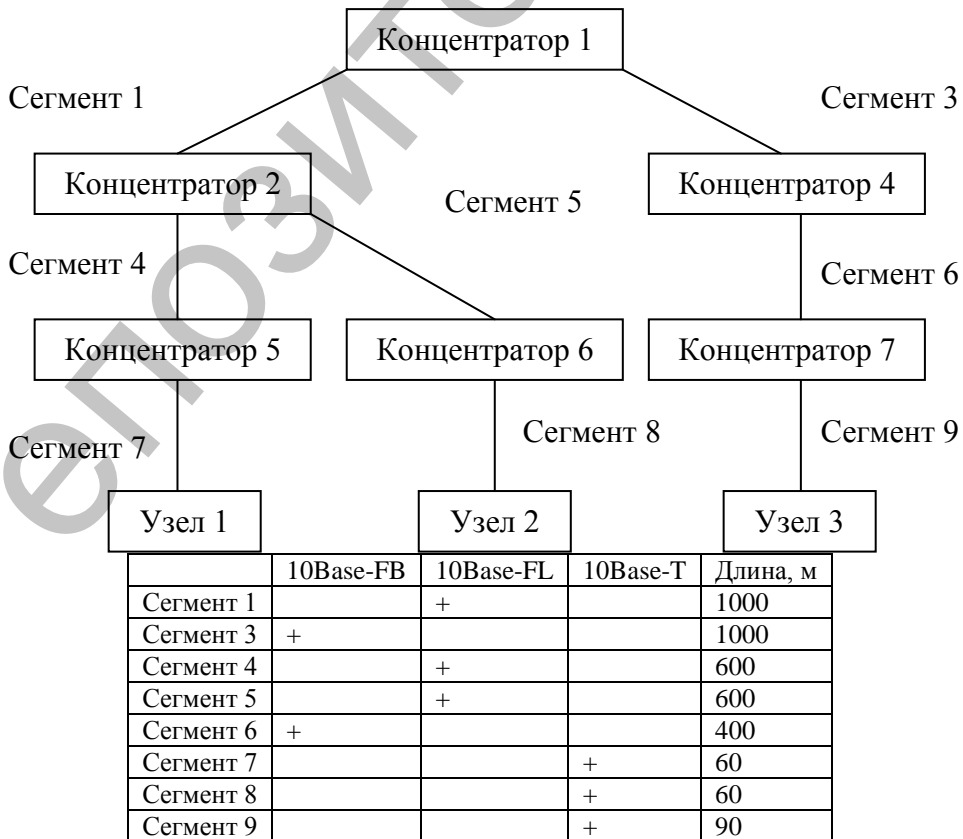


	10Base-FB	10Base-FL	10Base-T	Длина, м
Сегмент 1	+			400
Сегмент 3	+			500
Сегмент 4		+		1100
Сегмент 5		+		1100
Сегмент 6		+		600
Сегмент 7			+	100
Сегмент 8			+	100
Сегмент 9			+	100

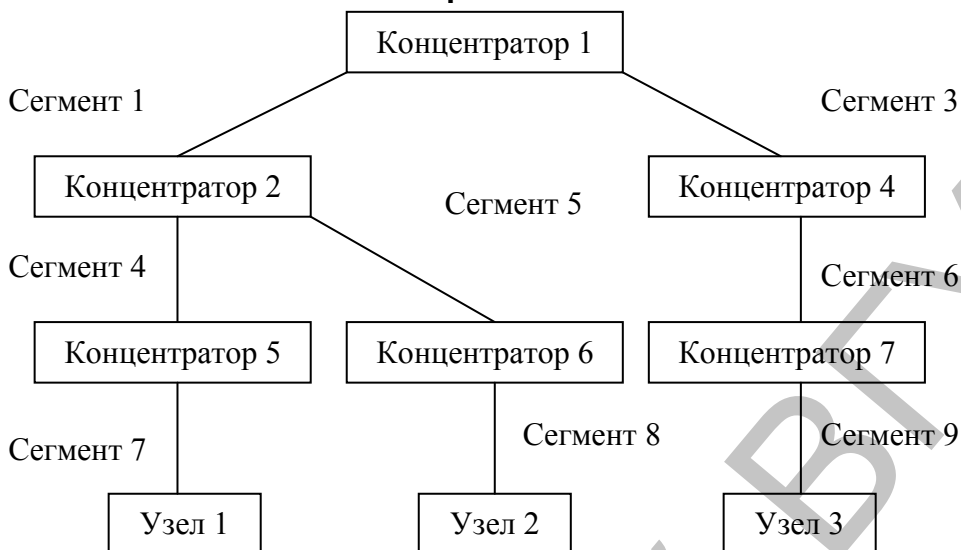
Вариант 6



Вариант 7

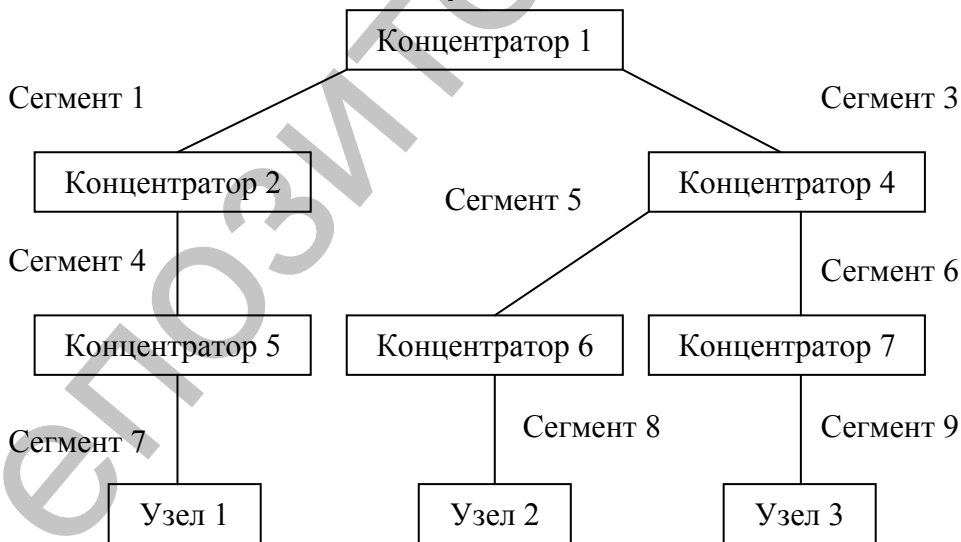


Вариант 8



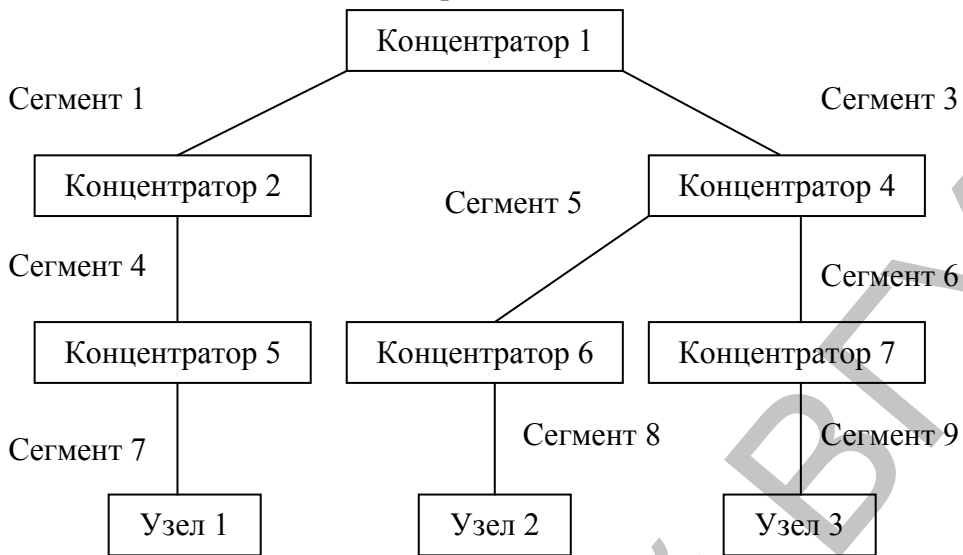
	10Base-FB	10Base-FL	10Base-T	Длина, м
Сегмент 1		+		900
Сегмент 3		+		900
Сегмент 4	+			700
Сегмент 5	+			700
Сегмент 6	+			500
Сегмент 7			+	70
Сегмент 8			+	70
Сегмент 9			+	100

Вариант 9



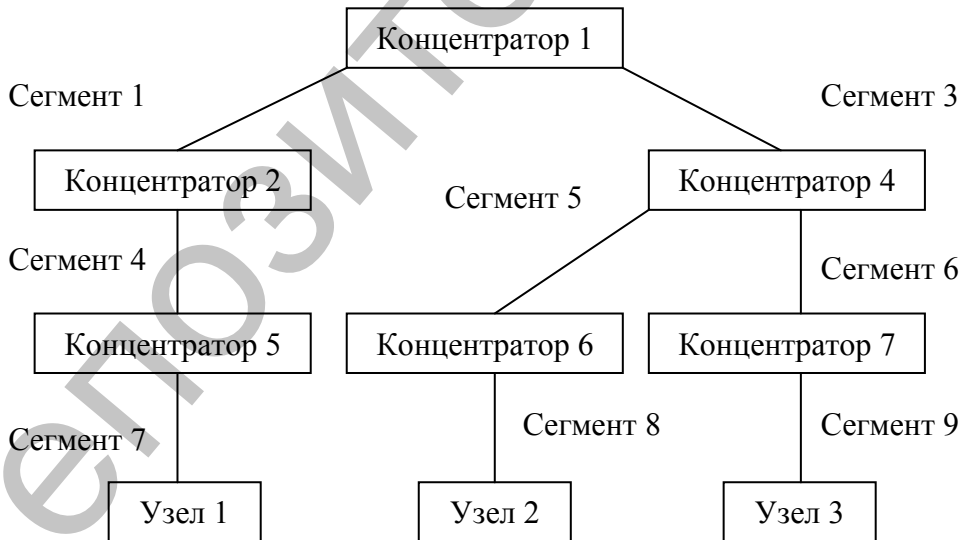
	10Base-FB	10Base-FL	10Base-T	Длина, м
Сегмент 1	+			400
Сегмент 3	+			500
Сегмент 4		+		1100
Сегмент 5		+		1100
Сегмент 6		+		600
Сегмент 7			+	100
Сегмент 8			+	100
Сегмент 9			+	100

Вариант 10

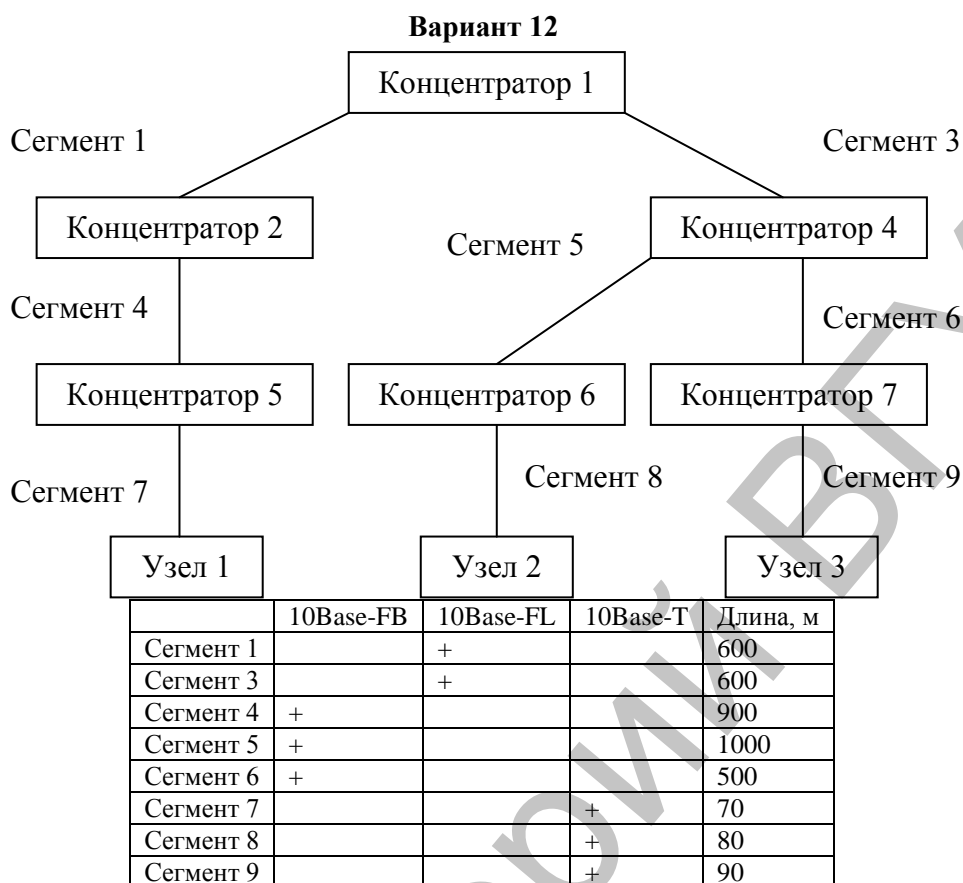


	10Base-FB	10Base-FL	10Base-T	Длина, м
Сегмент 1	+			500
Сегмент 3		+		500
Сегмент 4	+			1000
Сегмент 5	+			1000
Сегмент 6		+		500
Сегмент 7			+	80
Сегмент 8			+	80
Сегмент 9			+	100

Вариант 11



	10Base-FB	10Base-FL	10Base-T	Длина, м
Сегмент 1		+		1000
Сегмент 3	+			1000
Сегмент 4		+		600
Сегмент 5		+		600
Сегмент 6	+			400
Сегмент 7			+	60
Сегмент 8			+	60
Сегмент 9			+	90



1.5. Справочные данные IEEE

Таблица 1.1

Общие ограничения для всех стандартов Ethernet

Характеристика	Значение
Номинальная пропускная способность	10 Мбит/с
Максимальное число станций в сети	1024
Максимальное расстояние между узлами в сети	2500 м (в 10Base-FB – 2750 м)
Максимальное число коаксиальных сегментов в сети	5

Таблица 1.2

Параметры спецификаций физического уровня для стандарта Ethernet

Параметр	10Base-5	10Base-2	10Base-T	10Base-F
Кабель	Толстый коаксиальный кабель RG-8 или RG-11	Тонкий коаксиальный кабель RG-58	Неэкранированная витая пара категорий 3,4,5	Многомодовый волоконно-оптический кабель

Максимальная длина сегмента, М	500	185	100	2000
Максимальное расстояние между узлами сети (при использовании повторителей), м	2500	925	500	2500 (2740 для 10Base-FB)
Параметр	10Base-5	10Base-2	10Base-T	10Base-F
Максимальное число станций в Сегменте	100	30	1024	1024
Максимальное число повторителей между любыми станциями сети	4	4	4	4 (5 для 10Base-FB)

Таблица 1.3

Данные для расчета значения PDV

Тип сегмента	База левого сегмента, bt	База промежуточного сегмента, bt	База правого сегмента, bt	Задержка среды на 1 м, bt	Максимальная длина сегмента, м
10Base-5	11,8	46,5	169,5	0,0866	500
10Base-2	11,8	46,5	169,5	0,1026	185
10Base-T	15,3	42,0	165,0	0,113	100
10Base-FB	–	24,0	–	0,1	2000
10Base-FL	12,3	33,5	156,5	0,1	2000
FOIRL	7,8	29,0	152,0	0,1	1000
AUI (>2 м)	0	0	0	0,1026	2+48

Таблица 1.4

Уменьшение межкадрового интервала повторителями

Тип сегмента	Передающий сегмент, bt	Промежуточный сегмент, bt
10Base-5 или 10Base-2	16	11
10Base-FB	–	2
10Base-FL	10,5	8
10Base-T	10,5	8

1.6. Контрольные вопросы

1. Поясните механизм доступа к разделяемой среде в технологии Ethernet.
2. В каких случаях возможна оценка корректности конфигурации по физическим ограничениям?
3. Сформулируйте условие надежного распознавания коллизий.
4. С какой целью вводится ограничение на уменьшение межкадрового интервала?
5. В каком случае и почему для самого длинного пути проводятся два расчета?

2. Изучение структуры IP-адреса

Цель работы: изучение принципов адресации в сетях TCP/IP и приобретение практических навыков применения и назначения IP-адресов с использованием масок.

Необходимое оборудование: калькулятор.

2.1. Типы адресов стека TCP/IP

В стеке TCP/IP используются три типа адресов: локальные (называемые также аппаратными), IP-адреса и символьные доменные имена.

В терминологии TCP/IP под локальным адресом понимается такой тип адреса, который используется средствами базовой технологии для доставки данных в пределах подсети, являющейся элементом составной интерсети. В разных подсетях допустимы разные сетевые технологии, разные стеки протоколов, поэтому при создании стека TCP/IP предполагалось наличие разных типов локальных адресов. Если подсеть интерсети является локальной сетью, то локальный адрес – это MAC-адрес. Однако протокол IP может работать и над протоколами более высокого уровня, например, над протоколом IPX или X.25. В этом случае локальными адресами для протокола IP соответственно будут адреса IPX и X.25. Компьютер в локальной сети может иметь несколько локальных адресов даже при одном сетевом адаптере. Некоторые сетевые устройства не имеют локальных адресов. Например, к таким устройствам относятся глобальные порты маршрутизаторов, предназначенные для соединений типа «точка-точка».

IP-адреса представляют собой основной тип адресов, на основании которых сетевой уровень передает пакеты между се-

тиями. Эти адреса состоят из 4 байт, например, 109.26.17.100. IP-адрес назначается администратором во время конфигурирования компьютеров и маршрутизаторов. IP-адрес состоит из двух частей: номера сети и номера узла. Номер сети может быть выбран администратором произвольно либо назначен по рекомендации специального подразделения Internet (Internet Network Information Center, InterNIC), если сеть должна работать как составная часть Internet. Номер узла в протоколе IP назначается независимо от локального адреса узла. Маршрутизатор по определению входит сразу в несколько сетей. Поэтому каждый порт маршрутизатора имеет собственный IP-адрес. Конечный узел также может входить в несколько IP-сетей. В этом случае компьютер должен иметь несколько IP-адресов, по числу сетевых связей. Таким образом, IP-адрес характеризует не отдельный компьютер или маршрутизатор, а одно сетевое соединение.

Символьные доменные имена. Символьные имена в IP-сетях называются доменными и строятся по иерархическому признаку. Между доменным именем и IP-адресом узла нет никакого алгоритмического соответствия, поэтому необходимо использовать какие-то дополнительные таблицы или службы, чтобы узел сети однозначно определялся как по доменному имени, так и по IP-адресу. В сетях TCP/IP используется специальная распределенная служба Domain Name System (DNS), которая устанавливает это соответствие на основании создаваемых администраторами сети таблиц соответствия. Поэтому доменные имена называют также DNS-именами.

2.2. Классы IP-адресов

IP-адрес имеет длину 4 байта и обычно записывается в виде четырех чисел, представляющих значения каждого байта в десятичной форме и разделенных точками, например:

128.10.2.30 – традиционная десятичная форма представления адреса;

10000000 00001010 00000010 00011110 – двоичная форма представления этого же адреса.

Адрес состоит из двух логических частей – номера сети и номера узла в сети. Какая часть адреса относится к номеру сети, а какая – к номеру узла, определяется значениями первых бит адреса. Значения этих бит являются также признаками того, к какому классу относится тот или иной IP-адрес. На рис. 2.1 показана структура IP-адресов различных классов.

Если адрес начинается с 0, то сеть относят к классу А и номер сети занимает один байт, остальные 3 байта интерпретируются как номер узла в сети. Сети класса А имеют номера в диапазоне от 1 до 126. (Номер 0 не используется, а номер 127 зарезервирован для специальных целей, о чем будет сказано ниже.) Количество узлов в сетях класса А может достигать 2^{24} , то есть 16 777 216 узлов.

Класс А. 0ссссссс уууууууу уууууууу уууууууу
Класс В. 10сссссс сссссссс уууууууу уууууууу
Класс С. 110сссссс сссссссс сссссссс уууууууу
Класс D. 1110аааа аааааааа аааааааа аааааааа
Класс Е. 11110ззз зззззззз зззззззз зззззззз

Рис. 2.1. Классы IP-адресов (с – бит, входящий в номер сети; у – бит, входящий в номер узла; а – бит, входящий в адрес группы multicast; з – бит, входящий в зарезервированный адрес).

Если первые два бита адреса равны 10, то сеть относится к классу В. В сетях класса В под номер сети и под номер узла отводится по 16 бит. Таким образом, сеть класса В является сетью средних размеров с максимальным числом узлов 2^{16} , что составляет 65 536 узлов.

Если адрес начинается с последовательности 110, то это сеть класса С. В этом случае под номер сети отводится 24 бита, а под номер узла – 8 бит. Сети этого класса наиболее распространены, число узлов в них ограничено 2^8 , то есть 256 узлами.

Если адрес начинается с последовательности 1110, то он является адресом класса D и обозначает особый, групповой адрес – multicast. Если в пакете в качестве адреса назначения указан адрес класса D, то такой пакет должны получить все узлы, которым присвоен данный адрес.

Если адрес начинается с последовательности 11110, то это значит, что данный адрес относится к классу Е. Адреса этого класса зарезервированы для будущих применений.

2.3. *Особые IP-адреса*

В протоколе IP существует несколько соглашений об особой интерпретации IP-адресов.

Если весь IP-адрес состоит только из двоичных нулей, то он обозначает адрес того узла, который сгенерировал этот пакет; этот режим используется только в некоторых сообщениях ICMP.

Если в поле номера сети стоят только нули, то по умолчанию узел назначения принадлежит той же самой сети, что и узел, который отправил пакет.

Если все двоичные разряды IP-адреса равны 1, то пакет с таким адресом назначения должен рассылаться всем узлам, находящимся в той же сети, что и источник этого пакета. Такая рассылка называется ограниченным широковещательным сообщением (limited broadcast).

Если в поле номера узла назначения стоят только единицы, то пакет с таким адресом рассылается всем узлам сети с заданным номером сети. Например, пакет с адресом 192.190.21.255 доставляется всем узлам сети 192.190.21.0. Такая рассылка называется широковещательным сообщением (broadcast).

При адресации необходимо учитывать те ограничения, которые вносятся особым назначением некоторых IP-адресов. Так, ни номер сети, ни номер узла не может состоять только из одних двоичных единиц или только из одних двоичных нулей. Отсюда следует, что максимальное количество узлов, приведенное для сетей каждого класса, на практике должно быть уменьшено на 2. Например, в сетях класса C под номер узла отводится 8 бит, которые позволяют задавать 256 номеров: от 0 до 255. Однако на практике максимальное число узлов в сети класса C не может превышать 254, так как адреса 0 и 255 имеют специальное назначение. Из этих же соображений следует, что конечный узел не может иметь адрес типа 98.255.255.255, поскольку номер узла в этом адресе класса A состоит из одних двоичных единиц.

Особый смысл имеет IP-адрес, первый октет которого равен 127. Он используется для тестирования программ и взаимодействия процессов в пределах одной машины. Когда программа посылает данные по IP-адресу 127.0.0.1, то образуется как бы «петля». Данные не передаются по сети, а возвращаются модулям верхнего уровня как только что принятые. Поэтому в IP-сети запрещается присваивать машинам IP-адреса, начинающиеся со 127. Этот адрес имеет название loopback. В протоколе IP нет понятия широковещательности в том смысле, в котором оно используется в протоколах канального уровня локальных сетей, когда данные должны быть доставлены абсолютно всем узлам. Как ограниченный широковещательный IP-адрес, так и широковещательный IP-адрес имеют пределы распространения в интерсети – они ограничены либо сетью, к которой принадлежит узел-источник пакета, либо сетью, номер которой указан в адресе назначения.

Уже упоминавшаяся форма группового IP-адреса – multicast – означает, что данный пакет должен быть доставлен

сразу несколькими узлами, которые образуют группу с номером, указанным в поле адреса. Узлы сами идентифицируют себя, то есть определяют, к какой из групп они относятся. Один и тот же узел может входить в несколько групп. Члены какой-либо группы multicast не обязательно должны принадлежать одной сети. Групповой адрес не делится на поля номера сети и узла и обрабатывается маршрутизатором особым образом.

Групповая адресация предназначена для экономичного распространения в Internet или большой корпоративной сети аудио- или видеопрограмм, предназначенных сразу большой аудитории слушателей или зрителей. Если такие средства найдут широкое применение, то Internet сможет создать серьезную конкуренцию радио и телевидению.

2.4. Использование масок в IP-адресации

Важным элементом разбиения адресного пространства Internet являются подсети. Подсеть – это подмножество сети, не пересекающееся с другими подсетями. Это означает, что сеть организации может быть разбита на фрагменты, каждый из которых будет составлять подсеть. Реально каждая подсеть соответствует физической локальной сети (например, сегменту Ethernet). Подсети используются для того, чтобы обойти ограничения физических сетей на число узлов в них и максимальную длину кабеля в сегменте сети. Например, сегмент тонкого Ethernet имеет максимальную длину 185 м и может включать до 32 узлов. Самая маленькая сеть класса C может состоять из 254 узлов. Для того чтобы достичь этого значения, необходимо объединить несколько физических сегментов сети. Сделать это можно либо с помощью физических устройств (например, повторителей), либо при помощи машин-шлюзов. В первом случае разбиение на подсети не требуется, так как логически сеть выглядит как одно целое. При использовании шлюза сеть разбивается на подсети.

Разбиение сети на подсети использует ту часть IP-адреса, которая закреплена за номерами компьютеров. Администратор сети может замаскировать часть IP-адреса и использовать ее для назначения номеров подсетей. Фактически способ разбиения адреса на две части теперь будет применяться к адресу компьютера из IP-адреса сети, в которой организуется разбиение на подсети.

Маска подсети – это четыре байта, которые накладываются на IP-адрес для получения номера подсети. Например, маска 255.255.255.0 позволяет разбить сеть класса B на 254 подсети по 254 узла в каждой. Подсети не только решают, но и создают ряд

проблем. Например, происходит потеря адресов, но уже не по причине физических ограничений, а по причине принципа построения адресов подсети. Так, выделение трех битов на адрес подсети приводит к образованию не восьми, а только шести подсетей, так как номера 0 и 7 нельзя использовать в силу специального значения IP-адресов, состоящих из нулей или из единиц.

Для стандартных классов сетей маски имеют следующие значения:

класс А – 11111111.00000000.00000000.00000000 (255.0.0.0);

класс В – 11111111.11111111.00000000.00000000 (255.255.0.0);

класс С – 11111111.11111111.11111111.00000000 (255.255.255.0).

Снабжая каждый IP-адрес маской, можно отказаться от понятий классов адресов и сделать систему адресации более гибкой. Например, адрес 185.23.44.206 попадает в диапазон 128–191, то есть адрес относится к классу В. Следовательно, номером сети являются первые два байта, дополненные двумя нулевыми байтами – 185.23.0.0, а номером узла – 0.0.44.206. Если этот адрес ассоциировать с маской 255.255.255.0, то номером подсети будет 185.23.44.0, а не 185.23.0.0, как это определено системой классов.

В масках количество единиц в последовательности, определяющей границу номера сети, не обязательно должно быть кратным 8, чтобы повторять деление адреса на байты. Пусть, например, для IP-адреса 129.64.134.5 указана маска 255.255.128.0, то есть в двоичном виде:

IP-адрес 129.64.134.5 – 10000001.01000000.10000110.00000101;

Маска 255.255.128.0 – 11111111.11111111.10000000.00000000.

Если использовать для определения границы номера сети маску, то 17 последовательных единиц в маске, «наложенные» на IP-адрес, определяют в качестве номера сети в двоичном выражении число:

10000001.01000000.10000000.00000000 или в десятичной форме записи – номер сети 129.64.128.0, а номер узла 0.0.6.5.

2.5. Задание на лабораторную работу

1. Ознакомиться с теоретическим материалом.

2. Разработать приложение, которое по заданному классу (А, В или С), количеству подсетей N и максимальному количеству компьютеров M в подсети определяет маску для разбиения на подсети и список возможных IP-адресов подсетей. Если разбиение на подсети невозможно, приложение должно выдавать соответствующее сообщение об ошибке.

2.6. Контрольные вопросы

1. Назовите типы адресов, используемые в стеке TCP/IP. Охарактеризуйте их назначение и применяемые схемы адресации.
2. Назовите и охарактеризуйте классы IP-адресов.
3. Для каких целей используются договоренности об особых адресах?
4. Для каких целей при назначении адресов используются маски?
5. Опишите вид маски и принцип ее использования.
6. Опишите методику выбора маски по заданным в лабораторной работе параметрам.

3. Взаимодействие прикладных программ с помощью транспортного протокола TCP

Цель работы: изучение принципов организации обмена сообщениями между сетевыми приложениями по протоколу TCP и приобретение практических навыков создания клиент-серверных приложений на основе компонентов TClientSocket и TServerSocket.

Необходимое оборудование: IBM PC-совместимый компьютер, подключенный к локальной сети, с установленным программным обеспечением Borland Delphi версии 5.0 и выше.

3.1. Транспортный протокол TCP

TCP (Transmission Control Protocol) – это один из самых широко распространенных протоколов транспортного уровня. Главная функция TCP заключается в доставке сообщений без потерь, чего не может гарантировать протокол более низкого уровня IP (Internet Protocol). Для доставки сообщений предварительно устанавливается соединение между процессом-отправителем и процессом-получателем. Данное соединение осуществляет надежную доставку пакетов. Протокол TCP производит повторную передачу искаженного или утерянного пакета.

Выделение всех необходимых для надежной доставки сообщений функций в отдельный уровень освобождает разработчиков прикладных программ и утилит от решения задач управления потоком дейтаграмм. Протокол обеспечивает сквозную передачу данных от отправителя к получателю. Поскольку TCP ориентирован на установление соединения, то адресат, получивший дейтаграмму, должен уведомить отправителя об этом. Под-

разумеается, что между отправителем и получателем устанавливается виртуальный канал, где они обмениваются сообщениями, часть из которых является подтверждениями о получении данных либо кодами ошибок. Виртуальный канал на самом деле может подразумевать несколько реальных физических каналов передачи данных, поскольку сообщение может проходить через один или несколько шлюзов.

Когда некоторое приложение (процесс) прикладного уровня отправляет сообщение другому приложению с помощью ТСП, предполагается, что сообщение является потоком, т.е. представляет собой поток байтов, передаваемых асинхронно. ТСП получает поток байтов и собирает его в сегменты, добавляя заголовки в начало сегментов. Длина сегмента обычно определяется протоколом или выбирается администратором системы.

Процесс обмена данными начинается с передачи запроса на установление соединения от машины-отправителя к машине-получателю. В запросе содержится специальное целое число, называемое номером сокета (socket). В ответ получатель посылает номер своего сокета. Номера сокетов отправителя и получателя однозначно определяют соединение (конечно, соединение также не возможно без указания IP-адресов отправителя и получателя, но эта задача решается протоколами более низкого уровня – IP).

После установления соединения ТСП начинает передавать сегменты сообщения. На более низком IP-уровне отправителя сегменты разбиваются на одну или несколько дейтаграмм. Пройдя через сеть, дейтаграммы поступают к получателю, где IP-уровень снова собирает из них сегменты и передает их ТСП. ТСП собирает все сегменты в сообщении. От ТСП сообщение поступает к процессу-получателю, где обрабатывается протоколом прикладного уровня.

ТСП на машине-получателе собирает целое сообщение из сегментов, руководствуясь порядковыми номерами сегментов, которые записаны в их заголовке. Если какой-то сегмент сообщения потерян или поврежден (что проверяется по контрольной сумме в заголовке сегмента), то отправителю посылается сообщение, содержащее номер ошибочного сегмента. В этом случае отправитель повторно передает сегмент. Если сегмент успешно принят, то получатель посылает отправителю подтверждение-квитанцию.

В ТСП применяется средство ограничения потока данных, называемое скользящим окном. Оно представляет собой фрагмент сообщения, который адресат готов принять. При установлении соединения отправителю сообщается размер окна (размер

окна кратен размеру сегмента). После того, как отправитель передал количество байтов, соответствующее размеру окна, он должен ждать квитанции. Как только будет получена квитанция на переданные сегменты, окно сдвигается вправо на соответствующее число байтов, и новые сегменты могут быть переданы. Отправитель может передать без получения квитанций в сеть максимально столько сегментов, сколько их укладывается в скользящем окне. В процессе обмена данными получатель может присылать квитанции, в которых будет указан новый размер скользящего окна.

Важную роль в протоколе TCP играют таймеры. Сегмент считается потерянным, если квитанция на него не поступила в течение заданного времени ожидания. При этом производится повторная передача сегмента. При получении квитанции таймер останавливается. Если получатель обнаруживает несколько правильных копий одного и того же сегмента, то все лишние копии просто отбрасываются и отправителю передается только одна квитанция.

3.2. Транспортный протокол UDP

Протокол UDP (User Datagram Protocol) является более простым транспортным протоколом, чем протокол TCP. Он предоставляет прикладным процессам услуги транспортного уровня, которые мало чем отличаются от услуг более низкого уровня, предоставляемых протоколом IP. Протокол UDP обеспечивает доставку дейтаграмм, но не требует подтверждения их получения. Поэтому он не требует установления соединения между передающим и принимающим процессами.

Протокол UDP используется в тех случаях, когда требуется передать данные без установления соединения. Такая связь в принципе не надежна, так как отправителю не сообщается, правильно ли принято его сообщение и получено ли оно вообще. Для проверки возникновения ошибок может использоваться контрольная сумма пакета, но ошибки никак не обрабатываются – они либо игнорируются, либо их обработка выполняется уже на более высоком, прикладном уровне.

Данные, отправляемые прикладным процессом через UDP, достигают места назначения как единое целое, не дробясь на части. Например, если процесс-отправитель передал пять сообщений через порт, то и процесс-получатель должен считать из порта пять сообщений. Размер каждого записанного сообщения должен совпадать с размером каждого прочитанного.

Протокол UDP используется тогда, когда требуется простейший механизм передачи данных. Тогда контроль ошибок либо не выполняется (например, в прикладном протоколе TFTP – Trivial File Transfer Protocol – простейший протокол передачи файлов), либо выполняется на прикладном уровне (например, в управляющем протоколе SNMP – Simple Network Management Protocol или в файловой системе NFS – Network File System).

3.3. Порты, мультиплексирование и демультимплексирование

Пакеты, поступающие на транспортный уровень, организуются операционной системой в виде множества очередей к точкам входа различных прикладных процессов. В терминологии TCP/IP такие системные очереди называются портами. Порт однозначно определяет приложение в пределах компьютера. Порты приложений не стоит путать с портами (интерфейсами) оборудования.

Существуют два способа присвоения порта приложению – централизованный и локальный. За каждым из этих способов закреплен свой диапазон номеров портов: для централизованного – от 0 до 1023, для локального – от 1023 до 65 535.

Если процессы представляют собой популярные общедоступные службы, такие, как FTP, telnet, HTTP, TFTP, DNS и т.п., то за ними закрепляются стандартные присвоенные (assigned) номера, иногда также называемые хорошо известными (well-known) номерами. Централизованное присвоение службам номеров портов выполняется организацией Internet Assigned Numbers Authority (IANA). Эти номера закрепляются и публикуются в стандартах Internet (RFC 1700). Примеры хорошо известных портов приведены в табл. 3.1.

Для тех служб, которые еще не стали столь распространенными, чтобы закреплять за ними стандартные номера, номера портов выделяются локально. На каждом компьютере операционная система ведет список занятых и свободных номеров портов. При поступлении запроса от приложения, выполняемого на данном компьютере, операционная система выделяет ему первый свободный номер. Такие номера называют динамическими (dynamic). В дальнейшем все сетевые приложения должны адресоваться к данному приложению с указанием назначенного ему номера порта. После того как приложение завершит работу, выделенный ему локальный номер порта возвращается в список свободных и может быть назначен другому приложению.

Таблица 3.1

**Примеры номеров портов наиболее популярных служб сети
Internet**

Номер Порта	Служба сети	Описание
0		Зарезервирован
7	echo	Эхо-ответ на входящие сообщения
9	discard	Сброс (поглощение) всех входящих сообщений
11	users	Активные пользователи
13	daytime	Отклик, содержащий время дня
19	chargen	Генератор символов
20	ftp data	Передача данных по протоколу FTP
21	ftp	Передача управляющих команд по протоколу FTP
23	telnet	Порт подключения по протоколу TELNET
25	smtp	Протокол передачи почтовых сообщений SMTP
37	time	Отклик, содержащий время
42	name	Сервер имен
53	domain	Сервер имен доменов
67	boots	Протокол удаленной загрузки сервера
68	bootc	Протокол удаленной загрузки клиента
69	tftp	Упрощенный протокол передачи файлов TFTP
79	finger	Протокол получения информации о пользователях FINGER
80	http	Протокол передачи гипертекста HTTP
109	pop2	Протокол почтового ящика POP2
110	pop3	Протокол почтового ящика POP3
111	Rpc	Протокол удаленного вызова процедур RPC
156	sqlserv	Служба SQL
161	snmp	Управляющий протокол SNMP

Все сказанное выше о портах в равной степени относится к обоим протоколам транспортного уровня. Нет никакой зависимости между назначением номеров для приложений, использующих протокол TCP, и приложений, работающих с протоколом UDP. Приложения, которые передают данные на уровень IP, используя протокол UDP, получают номера, называемые портами UDP. Аналогично приложениям, обращающимся к протоколу TCP, выделяются порты TCP. В том и другом случаях это могут быть как назначенные, так и динамические номера. Диапазоны чисел, из которых выделяются номера портов TCP и UDP, совпадают.

Протоколы TCP и UDP ведут для каждого номера порта две очереди: очередь пакетов, поступающих в данный порт из сети, и очередь пакетов, отправляемых данным портом в сеть. Процедура приема данных протоколом TCP (или UDP), поступающих от нескольких прикладных служб, называется мультиплексированием. Обратная процедура распределения протоколом TCP (или UDP) поступающих от сетевого уровня пакетов между набором высокоуровневых служб, идентифицированных номерами портов, называется демультимплексированием.

3.4. Логические соединения

Для надежной передачи данных между двумя прикладными процессами предусматривается установление логического соединения. Номер порта в совокупности с номером сети и номером конечного узла однозначно определяют прикладной процесс в сети. Этот набор идентифицирующих параметров (IP-адрес, номер порта) имеет название сокет (socket).

Каждый взаимодействующий процесс идентифицируется сокетом – парой (IP-адрес интерфейса, номер порта), а каждое соединение – парой сокетов взаимодействующих процессов. Каждый процесс одновременно может участвовать в нескольких соединениях.

Так, например, если (IP1, n1), (IP2, n2), (IP3, n3) – сокеты трех разных процессов, то возможно образование следующих соединений:

- соединение 1 – {(IP1, n1), (IP2, n2)};
- соединение 2 – {(IP1, n1), (IP3, n3)};
- соединение 3 – {(IP2, n2), (IP3, n3)}.

Каждая такая пара однозначно идентифицирует соединение. Сутью же понятия «соединение» является договоренность о параметрах, характеризующих процедуру обмена данными между двумя процессами. В протоколе TCP каждая сторона соединения посылает противоположной стороне следующие параметры:

- максимальный размер сегмента, который она готова принимать;
- максимальный объем данных (возможно несколько сегментов), которые она разрешает другой стороне передавать в свою сторону, даже если та еще не получила подтверждения на предыдущую порцию данных;
- начальный порядковый номер байта, с которого она начинает отсчет потока данных в рамках данного соединения.

После того как в результате переговорного процесса модулей ТСП с двух сторон соединения параметры процедуры обмена определены, одни из них остаются постоянными в течение всего сеанса связи, а некоторые адаптивно изменяются.

Когда устанавливается несколько соединений, то может случиться, что несколько машин пошлют запросы на соединение, в которых указаны одинаковые порты источники и получатели. Однако путаницы с соединениями не возникает, потому что IP-адреса у всех машин разные, следовательно, каждое соединение будет однозначно определено своим сокетом.

3.5. Программирование обмена данными на основе транспортных протоколов

ТСП должен взаимодействовать не только с протоколами нижележащего уровня, но и с протоколами и приложениями прикладного уровня. Связь с прикладным уровнем осуществляется с помощью набора сервисных примитивов. Сервисные примитивы определены в стандарте протокола, а для прикладных программ они доступны в форме библиотек работы с сокетами. При установлении соединения каждая из сторон выполняет некоторые операции, называемые открытием соединения. Открытие может быть пассивным или активным. Как правило, одна из сторон производит активное открытие соединения, а другая – пассивное, тогда соединение устанавливается. Оба режима подчиняются четким правилам. Пассивное соединение еще иногда называют серверным, а активное – клиентским.

При активном соединении процесс прикладного уровня передает программному обеспечению ТСП на той же ЭВМ сервисный примитив запроса на установление соединения с номером сокета, после чего ТСП отправляет получателю запрос на установление соединения, затем ждет ответа. После установления соединения активный процесс (клиент) может инициировать прием или передачу данных.

При пассивном соединении прикладная программа переводит программное обеспечение ТСП в режим ожидания запроса на соединение от удаленной системы. Когда поступает запрос, программное обеспечение ТСП осуществляет установку соединения, после чего пассивный процесс (сервер) готов принимать и передавать данные.

Программный интерфейс сокетов изначально был разработан для операционной системы (ОС) UNIX. Библиотека функций, поддерживающих этот интерфейс, входит в ядро всех ОС типа

UNIX и Linux. Однако принципы работы с этим программным интерфейсом применимы к большинству ОС, поддерживающих TCP/IP (например, в семействе ОС и оболочек типа Windows программный интерфейс сокетов реализован в динамической библиотеке Winsock.dll).

Для протокола TCP пассивное (на стороне сервера) соединение с сокетом приводит к выполнению следующих функций:

- создание сокета и установление его типа (в ОС типа UNIX функция `socket`);
- настройка сокета на конкретное соединение (указывает адрес и номер порта – в ОС типа UNIX – функция `bind`);
- создание очередей клиентов (в ОС типа UNIX – функция `listen`);
- ожидание входящего запроса на соединение с сокетом (в ОС типа UNIX – функция `accept`);
- прием и передача данных от клиента (в ОС типа UNIX – функции `read`, `write`, `send`, `recv` и их модификации);
- закрытие соединения с клиентом (в ОС типа UNIX – функция `close`).

Получив входящий запрос на соединение, сервер должен решать как бы две задачи одновременно: обслуживать уже установленное с клиентом соединение в соответствии с прикладным протоколом (принимать и отдавать данные клиенту) и ожидать поступления новых запросов на соединение от других клиентов. Обычно в развитых ОС (к ним относятся все современные ОС) эта проблема решается за счет возможностей параллельного выполнения нескольких процессов. Сервер может породить новый процесс (или новую цепочку выполнения – `thread`), который и должен будет заняться обслуживанием уже установленного соединения, а основной процесс сервера может закрыть текущее соединение и вновь вернуться к ожиданию запросов на соединение от других клиентов. В ОС типа UNIX создание нового процесса решается с помощью функции `fork`, при этом за вновь созданным процессом сохраняются все соединения, сделанные в основном процессе.

Для протокола TCP активное (на стороне клиента) соединение с сокетом приводит к выполнению следующих функций:

- создание сокета и установление его типа (в ОС типа UNIX функция `socket`);
- установление соединения с сервером (указывает адрес и номер порта – в ОС типа UNIX – функция `connect`);

- прием и передача данных (в ОС типа UNIX – функции read, write, send, recv и их модификации);
- закрытие соединения с сервером (в ОС типа UNIX – функция close).

Клиент, как правило, не требует для своей работы параллельного выполнения нескольких процессов.

В среде программирования Borland Delphi существуют специальные классы, которые позволяют выполнять те же действия, что и библиотека сокетов в ОС UNIX. Они взаимодействуют с библиотекой Winsock.dll на основе специальных технологий ОС (ActiveX технологии и COM-объекты). В среде Borland Delphi версии 3.0 для целей клиентского и серверного соединений служит класс объектов TTCP; а в среде Borland Delphi версии 5.0 и выше для клиентского соединения существует класс объектов TClientSocket, а для серверного – TServerSocket. Естественно, пользователь может на основе базовых классов разрабатывать свои собственные классы, которые будут поддерживать соединения по определенным им самим прикладным протоколам.

Для того чтобы создать сокет, достаточно создать экземпляр объекта выбранного класса (TTCP – в среде Borland Delphi версии 3.0 на страничке компонент «Internet», TClientSocket или TServerSocket – в среде Borland Delphi версии 5.0 и выше также на страничке компонент «Internet»). Это можно выполнить при проектировании приложения в среде разработки или же средствами языка программирования при выполнении приложения. Чтобы специфицировать (настроить) сокет, необходимо созданному экземпляру объекта присвоить нужные значения в указанные свойства (properties) – как правило, это свойства с именами вида «Port» и «Host» (имена и состав свойств зависят от версии среды разработки). Это тоже можно сделать как в режиме проектирования приложения, так и командами присвоения свойств объекта в тексте программы. После этого сокет инициализирован и с ним можно работать.

Для работы сокета клиента необходимо открыть сокет (процедура Open устанавливает в Thru свойство Active), затем использовать процедуры установления соединения, передачи и приема данных, а в конце работы закрыть сокет (процедура Close). При удалении экземпляра объекта автоматически прекратит существование и связанный с ним сокет. Краткое описание основных свойств, событий и методов объектов приведено в пункте 3.8 «Справочные данные».

По установлению соединения наступит событие, которое программист должен соответствующим образом обработать. В

среде Borland Delphi версии 3.0 программист сам должен создавать потоки выполнения для обслуживания соединения (для выполнения потоков служат экземпляры класса объектов TThread), а в среде Borland Delphi версии 5.0 и выше это можно сделать в автоматическом режиме.

Более подробную информацию о функциях программного интерфейса с сокетами можно получить в справке соответствующей среды разработки или специализированной литературе.

3.6. Пример реализации простейшего клиент-серверного приложения на основе сокетов

Запустить Delphi. На пустой форме (возможен вариант совмещения функций клиента и сервера в одном приложении) разместить два компонента с закладки Internet: TClientSocket и TServerSocket (в 6 и 7 версии Delphi они могут отсутствовать; для установки надо выполнить шаги: Component – Install Packages – Add – dclsockets60.bpl или dclsockets70.bpl).

Для клиентской части программы понадобятся два компонента TEdit (один для ввода IP-адреса или DNS-имени сервера, другой – для ввода передаваемой информации) и одна кнопка «Послать» для соединения с сервером и отправки ему сообщения.

Для серверной части понадобится один компонент TMemo для выдачи служебной информации о подключениях и отображения принимаемых сообщений).

В свойствах TServerSocket: установить Port в 4000, затем – Active в true. В обработчике Button1Click написать код для соединения с сервером:

```
ClientSocket1.Host:=Edit1.Text;  
ClientSocket1.Port:=4000;  
ClientSocket1.Active:=true;
```

Теперь, при нажатии на Button1 произойдет соединение с сервером, адрес которого был указан в Edit1. При установлении соединения произойдет событие ClientSocket1.OnConnect, которое, например, надо обработать так, чтобы на сервер передавалась строка из Edit2 и происходило отсоединение:

```
Socket.SendText(Edit2.Text)  
ClientSocket1.Active:=False;
```

При подключении клиента к ServerSocket1 произойдет событие ServerSocket1.OnClientConnect, в котором нужно вывести следующую информацию:

```
Memo1.Lines.Add ('Client connected from: '+Socket.RemoteHost).
```


При попытке передать клиентом на сервер сообщение на сервере произойдет событие `ServerSocket1.OnClientRead`, в котором текст считывается и выводится в `Memo1`:

```
ReceivedString:=""; {локальная переменная типа string}
while Socket.ReceiveLength>0 do
  ReceivedString:=ReceivedString+Socket.ReceiveText;
Memo1.Lines.Add('Received string: '+ReceivedString);
```

При отсоединении клиента произойдет событие `OnClientDisconnect`, в котором мы добавим в `Memo1` соответствующее сообщение: `Memo1.Lines.Add('Client '+Socket.RemoteHost+' disconnected.');`

Запустите проект и протестируйте его. В качестве IP-адреса можно использовать адреса соседних компьютеров, на которых запущены аналогичные программы или адрес 127.0.0.1 для тестирования на локальном компьютере.

Примечание: Оправка сообщений с сервера клиентам осуществляется с использованием тех же методов и событий соответственно на стороне сервера и клиентов. Отличие заключается в необходимости точного указания номера соединения в свойстве `Connections[Index: Integer]: TcustomWinSocket`. Например:

```
ServerSocket1.Socket.Connections[i].SendText("Test");
```

Здесь $i = (0..N-1)$ – номер соединения, N – общее количество соединений.

3.7. Задание на лабораторную работу

1. Ознакомиться с теоретическим и справочным материалом.
2. Реализовать приведенное в пункте 2 простейшее приложение и ознакомиться с работой компонентов `Socket`.
3. Разработать приложение сервера и приложение клиента, позволяющее общаться между клиентами в режиме `online` (один сервер обслуживает несколько клиентов).

Приложение сервера постоянно опрашивает входящие соединения, хранит историю сообщений определенной глубины (например, 20), список зарегистрированных пользователей и другую информацию, необходимую для решения задачи.

Приложение клиента должно обеспечивать: регистрацию нового пользователя и его авторизацию по имени пользователя и паролю на сервере, передачу сообщений на сервер и прием сообщений от сервера, переданных на сервер другими клиентами.

3.8. Справочные данные

Основные свойства компонента ServerSocket:

Active – инициализирует соединение (открывает для прослушивания socket);

Name – имя экземпляра класса для использования в программе;

Port – номер порта, на котором ведется прослушивание входящих соединений.

Основные события компонента ServerSocket:

OnAccept – зафиксировано входящее соединение;

OnClientConnect – установлено соединение с клиентом;

OnClientDisconnect – разорвано соединение с клиентом;

OnClientError – ошибка при работе с клиентом;

OnClientRead – чтение данных, получаемых от клиента, только в этом событии разрешается принимать данные от клиента;

OnClientWrite – запись данных, передаваемых клиенту, только в этом событии разрешается передавать данные клиенту;

OnListen – сервер успешно проинициализирован для прослушивания socket.

Основные свойства компонента ClientSocket:

Active – инициализирует соединение;

Host – IP-адрес сервера, с которым устанавливается соединение;

Name – имя экземпляра класса для использования в программе;

Port – номер порта, на котором производится соединение с сервером.

Основные события компонента ClientSocket:

OnConnect – установлено соединение с сервером;

OnDisconnect – разорвано соединение с сервером; OnError – ошибка при работе с сервером;

OnRead – чтение данных, получаемых от сервера, только в этом событии разрешается принимать данные от сервера;

OnWrite – запись данных, передаваемых серверу, только в этом событии разрешается передавать данные серверу.

Основные методы класса CustomWinSocket:

ReceiveText – принимает данные в виде строки символов;

SendText – отправляет данные в виде строки символов.

3.9. Контрольные вопросы

1. Назначение протокола TCP и принцип его работы.
2. Назначение протокола UDP и принцип его работы.
3. Назовите отличия протоколов TCP и UDP.
4. С какой целью используются порты?
5. В чем заключается сущность мультиплексирования и демultipлексирования?
6. Поясните сущность понятия «логическое соединение». Каким образом оно определяется в сети?
7. Поясните сущность активного режима работы сетевого приложения.
8. Поясните сущность пассивного режима работы сетевого приложения.

4. Взаимодействие прикладных программ с помощью протокола электронной почты SMTP

Цель работы: изучение принципов организации взаимодействия прикладных программ с помощью протокола электронной почты SMTP и приобретение практических навыков создания клиентских почтовых приложений, использующих протокол SMTP.

Необходимое оборудование: IBM PC-совместимый компьютер, подключенный к глобальной сети Internet, с установленным программным обеспечением Borland Delphi версии 5.0 и выше.

4.1. Модель протокола, команды и коды ответов SMTP

Для передачи сообщений по TCP-соединению большинство почтовых агентов пользуются протоколом SMTP (Simple Mail Transfer Protocol – простой протокол электронной почты).

SMTP принят в качестве стандартного метода передачи электронной почты в сети Internet. В качестве транспортного протокола SMTP использует TCP, соединение устанавливается через порт с номером 25. Для обслуживания этого соединения используется специальная программа, которая именуется почтовым сервером. Для формирования сообщения и установления соединения используется почтовая программа (утилита) пользователя.

Главной целью протокола SMTP является надежная и эффективная доставка электронных почтовых сообщений. Для реализации протокола требуется только надежный канал связи. Средой для SMTP может служить отдельная локальная сеть, система

сетей или же вся всемирная сеть Internet. Если между отправителем и получателем письма имеется непосредственная связь, адрес пользователя имеет вид *имя_пользователя@адрес_ЭВМ*. Когда получатель находится на ЭВМ, которая не поддерживает соединение по протоколу SMTP, и передача происходит через промежуточный сервер, то адрес получателя письма может иметь иной вид, например:

имя_пользователя%имя_сервера@адрес_ЭВМ.

Адреса ЭВМ в сети Internet представляют собой имя домена и преобразуются в IP-адреса согласно протоколу DNS.

Протокол SMTP базируется на следующей модели коммуникаций: в ответ на запрос пользователя почтовая программа-отправитель устанавливает двустороннюю связь с программой-приемником (почтовым сервером). Получателем может быть окончательный или промежуточный адресат. Если необходимо, почтовый сервер может установить соединение с другим сервером и передать сообщение дальше. SMTP-команды генерируются отправителем и посылаются получателю. На каждую команду должен быть получен отклик.

Перечень команд протокола SMTP, определенный спецификацией RFC 821, приведен в табл. 4.1. Это внутренние команды протокола. Если пользователь использует для работы с электронной почтой некоторую утилиту, то эти команды ему недоступны. Они представляют интерес только при программировании взаимодействия программ на основе этого протокола.

В соответствии со спецификацией, помеченные крестиком команды обязаны присутствовать в любой реализации SMTP. Остальные команды SMTP могут быть реализованы дополнительно. Каждая SMTP-команда должна заканчиваться либо пробелом (если у нее есть аргумент), либо комбинацией CRLF (Carriage-Return, Line-Feed – возврат каретки, перевод строки; коды ASCII – 13 и 10 соответственно). В описании команд употребляется слово <данные>, а не <сообщение>. Этим подчеркивается, что, кроме текста, SMTP позволяет передавать и двоичную информацию, например, графические или звуковые файлы.

Команды состоят из ключевых слов, за которыми следует один или более параметров. Ключевое слово состоит из 4-х символов и отделено от аргумента одним или несколькими пробелами. Каждая командная строка заканчивается символами CRLF. Синтаксис команд протокола SMTP приведен в пункте 4.6 «Справочные данные».

В спецификации SMTP требуется, чтобы сервер отвечал на каждую команду SMTP-клиента. Сервер отвечает трехзначной

комбинацией цифр, называемой кодом ответа. Вместе с кодом ответа, как правило, передается одна или несколько строк текстовой информации.

Примечание: Несколько строк текста, как правило, сопровождают только команды EXPN и HELP. В спецификации SMTP, однако, ответ на любую команду может состоять из нескольких строк текста.

Каждая цифра в коде ответа имеет определенный смысл. Первая цифра означает, было ли выполнение команды успешным (2), неуспешным (5) или еще не закончилось (3). Как указано в приложении E документа RFC 821, простой SMTP-клиент может анализировать только первую цифру в ответе сервера и на основании ее продолжать свои действия. Вторая и третья цифры кода ответа разъясняют значение первой. В лабораторной работе допускается не анализировать коды ответов. Можно использовать готовые коды, определенные RFC 821 и приведенные в пункте 4.6 «Справочные данные».

Через систему электронной почты передаются сообщения, которые должны иметь строго определенный формат. Любое почтовое сообщение можно разделить на три части: «конверт», заголовки и собственно текст. «Конверт» используется почтовым сервером, он содержит две команды – MAIL и RCPT. Заголовок используется почтовой программой пользователя. Он может содержать несколько специальных полей: From, To, Date, Subject и др. Каждое из этих полей содержит имя, за которым после двоеточия идет его значения. При желании, пользователь может ознакомиться с содержимым всех полей заголовка письма.

4.2. Кодировка сообщений

Текст сообщения должен передаваться в виде 7-разрядных символов ASCII. Конец сообщения представляет собой строку, содержащую только символы точки «.» и перевода строки. Если по каким-то причинам такая строка непосредственно встречается внутри текста сообщения, то передающая сторона автоматически дублирует точку, чтобы принимающая сторона не приняла ее за конец сообщения (на принимающей стороне дублированный символ точки в начале строки заменяется одной точкой). С 7-битной кодировкой сообщений связана проблема пересылки писем, написанных на национальных языках, в т.ч. и на русском, поскольку для русского алфавита требуется 8-битная кодировка. Большинство современных почтовых серверов поддерживают не только 7-битные, но и 8-битные кодировки, однако для совмес-

тимости со старыми почтовыми системами может использоваться специальная схема преобразования 8-битных кодов в 7-битные (естественно, при этом увеличивается объем сообщения, поскольку один 8-битный символ заменяется на специальную последовательность 7-битных).

Более сложная проблема заключается в том, что нет единого стандарта на кодировку русского алфавита, и поэтому на сторонах отправителя и получателя могут использоваться различные кодировки. Для того чтобы сообщение, посланное отправителем в одной кодировке, было понятно адресату, который применяет другую кодировку, почтовый сервер использует специальные таблицы смены кодировок. С помощью них текст сообщения, посланного в одной кодировке, преобразуется в другую кодировку. Чтобы сервер знал, какую кодировку использовать, информация об исходной кодировке включается почтовой системой отправителя в текст или в заголовок сообщения (в зависимости от используемой системы). Одно сообщение в процессе своего прохождения через почтовые сервера может претерпевать несколько перекодировок. В случае, когда у сервера нет информации об используемой кодировке или тип кодировки указан неверно (непонятен серверу), текст сообщения может исказиться.

4.3. Процесс передачи сообщений

Передача сообщения по протоколу SMTP происходит следующим образом: после установления соединения стороны обмениваются кодами аутентификации (с помощью команд HELO), затем одна из них посылает команду MAIL, в которой указан адрес отправителя и сведения о письме. Если получатель готов к приему сообщения, он посылает положительный отклик. Далее отправитель посылает одну или несколько команд RCPT, в которых идентифицирует адресатов сообщения. Если получатель может принять сообщение для указанного адресата, то он снова выдает положительный отклик на каждую команду. После этого передается команда DATA, за которой следует само письмо (сколько бы ни было адресатов, само письмо передается только один раз). В конце сеанса отправитель дает команду QUIT.

Конечно, маловероятно, что при приеме почтовым сервером сообщения, адресованному некоторому пользователю ЭВМ, этот пользователь установит с сервером непосредственную связь по протоколу SMTP. Обычно сервер размещает сообщение в почтовом ящике пользователя. Почтовый ящик – это некоторое промежуточное хранилище электронных сообщений, позволяющее

почтовому серверу осуществить их передачу адресату не сразу, а в тот момент, когда он установит связь с сервером. Обычно сообщения хранятся в почтовом ящике не вечно – по истечении некоторого времени почтовый сервер удаляет их оттуда. Политика обслуживания почтовых ящиков регламентируется администратором системы.

Для того чтобы получить сообщение из своего почтового ящика, почтовая программа пользователя соединяется с сервером уже не по протоколу SMTP, а по специальному почтовому протоколу получения сообщений. Такой протокол позволяет работать с почтовым ящиком: забирать сообщения, удалять сообщения, сортировать их и выполнять другие операции. Самым популярным в настоящее время протоколом такого рода является протокол POP3 (Post Office Protocol v.3). Он предусматривает соединение с почтовым сервером на основе транспортного протокола TCP через порт 110. Другой, более сложный, известный протокол – IMAP4.

Во многих версиях ОС UNIX протокол SMTP, а также другие почтовые протоколы, реализованы в простейших утилитах типа sendmail или mmdf. Первая из них может служить как почтовым клиентом, так и почтовым сервером. Обычно она работает в фоновом режиме и ждет поступления сообщений от пользователя или извне. В ОС типа Windows NT ту же роль играет утилита, именуемая mdaemon (mail daemon).

Для непосредственного взаимодействия с пользователем существует простейшая программа, именуемая mail или mailx. Однако в современных ОС она используется редко, поскольку существует большое количество прикладных программ с развитым интерфейсом, которые предоставляют пользователю возможности работы с электронной почтой. Примерами таких программ могут служить широко известные пакеты Outlook Express (производство фирмы Microsoft) и Netscape Communicator (производство фирмы Netscape). Внешний интерфейс этих программных продуктов чрезвычайно прост, так что использовать их может даже неквалифицированный пользователь. Все эти программы формируют сообщения в требуемом формате и передают их утилитами ОС, работающим в фоновом режиме, для отправки. К ним же они обращаются и для приема сообщений.

4.4. Пример последовательности команд почтовой транзакции

Каждой строке присвоен номер и обозначено, кому они принадлежат – передатчику (C) или приемнику (S). Текст справа

от двоеточия содержит действительно передаваемые данные. Трехзначные цифровые комбинации в начале передаваемых строк обозначают коды ответа. Ответ SMTP похож на сообщения-подтверждения о доставке, поскольку появляется лишь в том случае, когда приемник получил данные.

1. S: 220 mail.ru ESMTP Wed, 07 Sep 2005 09:21:58 +0400
2. C: Helo XName
3. S: 250 mx1.mail.ru Hello XName [62.76.36.11]
4. C: mail from: sidorov@mail.ru
5. S: 250 OK
6. C: rcpt to: ivanov@yandex.ru
7. S: 250 Accepted
8. C: rcpt to: petrov@mail.ru
9. S: 250 Accepted
10. C: data
11. S: 354 Enter message, ending with "." on a line by itself
12. C: From: sidorov@mail.ru
13. C: To: ivanov@yandex.ru
14. C: Subject: Тема
15. C: Текст сообщения
16. C: простой
17. C: .
18. S: 250 OK id=1ECsNi-000ATp-00
19. C: quit 20. S: 221 mx1.mail.ru closing connection

Как видно из строки 1, когда SMTP-клиент устанавливает TCP-соединение с портом протокола 25, SMTP-сервер отвечает кодом 220. Это означает, что соединение успешно установлено.

После того как почтовые агенты компьютеров установили соединение и обменялись приветствиями, первой командой, согласно спецификации, должна быть команда HELO. Как указано в строке 2, SMTP-клиент передает HELO, указывая имя своего компьютера в качестве аргумента (имя может быть вымышленным).

В ответ на HELO приемник выдает код 250, сообщая передатчику о том, что команда принята и обработана.

После установления TCP-соединения и идентификации (при помощи HELO) SMTP-клиент приступает к почтовой транзакции. Для начала он выполняет одну из следующих команд: MAIL, SEND, SOML или SAML. В нашем примере использована команда MAIL.

После того как сервер выдал код ответа 250 (строка 5), согласившись обработать сообщение от sidorov@mail.ru, необходимо указать получателя сообщения. Это делается при помощи команды RCPT. Команда RCPT имеет аргумент – имя получателя.

На одну команду приходится только одно имя, поэтому, если получателей несколько, команда RCPT выдается несколько раз. В нашем примере команды RCPT выполняются в строках 6 и 8. Выдав команду RCPT, клиент ожидает получить ответ с кодом 250.

После того как посланы все команды RCPT, клиент начинает передачу данных при помощи команды DATA. В строке 10 показано, как клиент (передатчик) высылает команду DATA, в строке 11 – как сервер отвечает кодом 354. Этот код означает, что передача данных разрешена и должна заканчиваться комбинацией CRLF-точка-CRLF (новой строкой, содержащей только точку).

После того как получен код 354, клиент может начать передачу данных.

Сервер, в свою очередь, помещает принятые данные в очереди входящих сообщений. Сервер не высылает никаких ответов до тех пор, пока не получит комбинацию CRLF-точка-CRLF от клиента, означающую конец передачи данных. Как показано в строках 17 и 18, в ответ на полученную комбинацию CRLF-точка-CRLF сервер выдает код 250, который означает успешное окончание операции.

Для того чтобы закончить почтовую транзакцию, клиент, по правилам SMTP, обязан послать команду QUIT. Сервер, в свою очередь, отвечает кодом 221. Этот код подтверждает клиенту, что соединение будет закрыто, после чего соединение действительно закрывается.

В любой момент во время транзакции клиент может использовать команды NOOP, HELP, EXPN и VRFY. В ответ на каждую команду сервер высылает клиенту определенную информацию. В зависимости от ответа, клиент может предпринять определенные действия, однако в спецификации SMTP это не оговаривается.

4.5. Задание на лабораторную работу

1. Ознакомиться с теоретическим и справочным материалом.
2. Получить у преподавателя доменное имя или IP-адрес SMTP-сервера, через который возможна отправка сообщений из учебных аудиторий.
3. Разработать приложение почтового клиента, позволяющее выполнять отставку текстовых сообщений по протоколу SMTP указанным адресатам.

Приложение использует для обмена с сервером транспортный протокол TCP/IP. Сообщение должно включать поля From,

To, Subject. Пользователь должен иметь возможность заполнить необходимые поля и дать команду на отправку сообщения. Процесс обмена командами и ответами должен происходить в автоматическом режиме и отображаться на экране в формате, приведенном в примере.

В случае возвращения сервером кода ответа, отличающегося от требуемого, необходимо вывести сообщение об этом и разорвать соединение.

4.6. Справочные данные

Таблица 4.1

Команды протокола SMTP

Команда	Обязательна	Описание
HELO	X	Идентифицирует модуль-передатчик для модуля-приемника (hello).
MAIL	X	Начинает почтовую транзакцию, которая завершается передачей данных в один или несколько почтовых ящиков (mail).
RCPT	X	Идентифицирует получателя почтового сообщения (recipient).
DATA		Строки, следующие за этой командой, рассматриваются получателем как данные почтового сообщения. В случае SMTP почтовое сообщение заканчивается комбинацией символов: CRLF-точка-CRLF.
RSET		Прерывает текущую почтовую транзакцию (reset).
NOOP		Требует от получателя не предпринимать никаких действий, а только выдать ответ ОК. Используется для тестирования. (No operation).
QUIT		Требует выдать ответ ОК и закрыть текущее соединение.
VERFY		Требует от приемника подтвердить, что ее аргумент является действительным именем пользователя.
SEND		Начинает почтовую транзакцию, доставляющую данные на один или несколько терминалов (а не в почтовый ящик).

SOML		Начинает транзакцию MAIL или SEND, доставляющую данные на один или несколько терминалов или в почтовые ящики.
SAML		Начинает транзакцию MAIL и SEND, доставляющие данные на один или несколько терминалов и в почтовые ящики.
EXPN		Команда SMTP-приемнику подтвердить, действительно ли аргумент является адресом почтовой рассылки, и если да, вернуть адрес получателя сообщения (expand).
HELP		Команда SMTP-приемнику вернуть сообщение-справку о его командах.

Таблица 4.2

Коды ответа протокола SMTP

Код	Описание
211	Ответ о состоянии системы или помощь.
214	Сообщение-подсказка (помощь).
220 <имя_домена>	Служба готова к работе.
221 <имя_домена>	Служба закрывает канал связи.
250	Запрошенное действие почтовой транзакции успешно завершилось.
251	Данный адресат не является местным; сообщение будет передано по маршруту <forward-path>.
354	Начинай передачу сообщения. Сообщение заканчивается комбинацией CRLF-точка-CRLF.
421 <имя_домена>	Служба недоступна; соединение закрывается.
450	Запрошенная команда почтовой транзакции не выполнена, так как почтовый ящик недоступен.
451	Запрошенная команда не выполнена; произошла локальная ошибка при обработке сообщения.
452	Запрошенная команда не выполнена; системе не хватило ресурсов.
500	Синтаксическая ошибка в тексте команды; команда не опознана.

501	Синтаксическая ошибка в аргументах или параметрах команды.
502	Данная команда не реализована.
503	Неверная последовательность команд.
504	У данной команды не может быть аргументов.
550	Запрошенная команда не выполнена, так как почтовый ящик недоступен.
551	Данный адресат не является местным; попробуйте передать сообщение по маршруту <forward-path>.
552	Запрошенная команда почтовой транзакции прервана; дисковое пространство, доступное системе, переполнилось.
553	Запрошенная команда не выполнена; указано недопустимое имя почтового ящика.
554	Транзакция не выполнена.

Синтаксис команд протокола SMTP (SP – пробел):

HELO <SP> <domain> <CRLF>
 MAIL <SP> FROM:<reverse-path> <CRLF>
 RCPT <SP> TO:<forward-path> <CRLF>
 DATA <CRLF>
 RSET <CRLF>
 SEND <SP> FROM:<reverse-path> <CRLF>
 SOML <SP> FROM:<reverse-path> <CRLF>
 SAML <SP> FROM:<reverse-path> <CRLF>
 VRFY <SP> <string> <CRLF>
 EXPN <SP> <string> <CRLF>
 HELP <SP> <string> <CRLF>
 NOOP <CRLF>
 QUIT <CRLF>

4.7. Контрольные вопросы

1. Назначение протокола SMTP.
2. Опишите модель работы протокола SMTP.
3. Каковы особенности кодировки почтовых сообщений?
4. Опишите типовую последовательность действий почтового клиента при отправке сообщения.
5. Какие операции должны выполняться во время почтовой транзакции? Какими командами она открывается и закрывается?
6. По каким принципам формируются команды и коды ответов протокола SMTP?

5. Взаимодействие прикладных программ с помощью протокола электронной почты POP3

Цель работы: изучение принципов организации взаимодействия прикладных программ с помощью протокола электронной почты POP3 и приобретение практических навыков создания клиентских почтовых приложений, использующих протокол POP3.

Необходимое оборудование: IBM PC-совместимый компьютер, подключенный к глобальной сети Internet, с установленным программным обеспечением Borland Delphi версии 5.0 и выше.

5.1. Модель протокола POP3, его назначение и стадии POP3-сессии

Post Office Protocol (POP) – протокол доставки почты пользователю из почтового ящика почтового сервера POP. Многие концепции, принципы и понятия протокола POP выглядят и функционируют подобно SMTP. Команды POP практически идентичны командам SMTP, отличаясь в некоторых деталях. Сервер POP находится между агентом пользователя и почтовыми ящиками.

В настоящее время существуют две версии протокола POP – POP2 и POP3, обладающие примерно одинаковыми возможностями, однако несовместимые друг с другом. У POP2 и POP3 разные номера портов протокола (109 и 110 соответственно). Протокол POP3 не является расширением или модификацией POP2 – это совершенно другой протокол. POP2 определен в документе RFC 937 (Post Office Protocol-Version 2, Butler, et al, 1985), а POP3 – в RFC 1225 (Post Office Protocol-Version 3, Rose, 1991). POP3 разработан с учетом специфики доставки почты на персональные компьютеры и имеет соответствующие операции для этого.

Конструкция протокола POP3 обеспечивает возможность пользователю обратиться к своему почтовому серверу и изъять накопившуюся для него почту. Пользователь может получить доступ к POP-серверу из любой точки доступа к Internet. При этом он должен запустить специальный почтовый агент, работающий по протоколу POP3, и настроить его для работы со своим почтовым сервером. Сообщения доставляются клиенту по протоколу POP, а посылаются при помощи SMTP. То есть на компьютере пользователя существуют два отдельных агента-интерфейса к почтовой системе – доставки (POP) и отправки (SMTP).

В протоколе POP3 оговорены три стадии процесса получения почты: авторизация, транзакция и обновление. После того как сервер и клиент POP3 установили соединение, начинается стадия авторизации. На стадии авторизации клиент идентифицирует себя для сервера. Если авторизация прошла успешно, сервер открывает почтовый ящик клиента и начинается стадия транзакции. В ней клиент либо запрашивает у сервера информацию (например, список почтовых сообщений), либо просит его совершить определенное действие (например, выдать почтовое сообщение). На стадии обновления сеанс связи заканчивается.

В этом режиме POP3 сервер освобождает все занятые ресурсы и завершает работу. После этого TCP-соединение закрывается. У POP3 сервера может быть INACTIVITY AUTOLOGOUT таймер. Этот таймер должен быть, по крайней мере, с интервалом 10 минут. Это значит, что если клиент и сервер не взаимодействуют друг с другом, сервер автоматически прерывает соединение и при этом не переходит в режим обновления.

5.2. Формат сообщений

Почтовое сообщение состоит из двух частей: заголовка и тела письма. Между ними расположена пустая строка. Каждое поле заголовка состоит в свою очередь из имени поля и значения, которые разделяются двоеточием. Полей заголовка может быть множество, но наиболее часто встречаются только некоторые из них. Поля *To:* и *From:* указывают на получателя и отправителя письма соответственно, поле *Subject:* используется для передачи темы письма. *Reply-To:* применяется для указания обратного адреса, а *CC:* – копии сообщения. *Received:* – такую строку добавляет каждый почтовый сервис, через который прошло письмо. По нему легко отследить путь письма. *Date:* означает дату создания письма. Используется специальный стандарт с указанием часового пояса отправителя относительно Гринвича.

Существует множество нестандартизованных полей заголовка. Они начинаются с символа X. Это могут быть такие поля, как *X-Mailer:* – почтовая программа отправителя; *X-MSMail-Priority:*, *X-Priority:* – приоритет (важность) письма и т.д. Разработчик также может создавать и использовать собственные поля, начинающиеся с X.

Дополнением к традиционной электронной почте является ее расширение MIME (Multipurpose Internet Mail Extensions). Оно не требует каких-либо переделок в почтовых серверах, но позволяет снять с электронной почты привычные ограничения и пре-

доставляет возможность пересылать любую информацию. Например, MIME позволяет создавать многосекционные письма. В каждой секции описывается тип информации, находящейся в ней (*Content-Type:*) и вид кодирования (*Content-Transfer-Encoding:*). Секции могут содержать как простой текст (*Content-Type: text/plain*) и текст в формате HTML (*Content-Type: text/html*), так и произвольные файлы в прикреплениях письма (*attach*).

Так как электронная почта изначально разрабатывалась как служба по пересылке обычного текста, это наложило ограничение на передаваемую информацию. Для передачи двоичной информации требуется специальное кодирование. Используется несколько методов кодирования: для текста с использованием только латинских букв, цифр и некоторых спецсимволов обычно используется *Content-Transfer-Encoding: 7bit*. Если сюда еще добавить русские буквы, то используется *Content-Transfer-Encoding: 8bit*. Для передачи HTML-кода применяют *Content-Transfer-Encoding: quoted-printable*. Чтобы отличить код от текста, перед ним ставится символ «=». Например, фраза «Привет!» будет выглядеть в закодированном виде как «=CF=F0=E8=E2=E5=F2!». Для двоичных файлов (архивы, изображения и т.д.) используют *Content-Transfer-Encoding: base64*, когда каждые 6 байт преобразуются в 8 печатных символов.

5.3. Процесс получения сообщений. Команды и ответы протокола POP3

Команды POP3 состоят из ключевых слов, за некоторыми следует один или более аргументов. Все команды заканчиваются парой CRLF. Ключевые слова и аргументы состоят из ASCII символов. Ключевое слово и аргументы разделены одиночным пробелом. Ключевое слово состоит от 3-х до 4-х символов, а аргумент может быть длиной до 40 символов.

Ответы в POP3 состоят из индикатора состояния и ключевого слова, за которым может следовать дополнительная информация. Ответ заканчивается парой CRLF. Существует только два индикатора состояния: «+OK» – положительный и «-ERR» – отрицательный. Синтаксис и описание команд и ответов протокола POP3 приведены в пункте 5.5 «Справочные данные»

Стадия авторизации. Как только будет установлено TCP соединение с POP3 сервером, он отправляет приглашение, заканчивающееся парой CRLF, например:

S: +OK POP3 server ready.

Теперь POP3 сессия находится в режиме авторизации. Клиент должен идентифицировать себя на сервере, используя команды USER и PASS. Сначала необходимо отправить команду USER, после которой в качестве аргумента следует имя пользователя. Если сервер отвечает положительно, то необходимо отправить команду PASS, аргументом которой является пароль. Если после отправки команды USER или PASS сервер отвечает негативно, то можно попытаться авторизоваться снова или выйти из сессии с помощью команды QUIT. После успешной авторизации сервер открывает и блокирует maildrop (почтовый ящик). В ответе на команду PASS сервер информирует, сколько сообщений находится в почтовом ящике, и передает их общий размер. Синтаксис, описание и возможные ответы на команды приведены в справочном разделе.

Стадия транзакции. После успешной идентификации пользователя на сервере POP3 сессия переходит в режим транзакции, где пользователь может передавать соответствующие команды (см. справочный раздел). После каждой из таких команд следует ответ сервера. В режиме транзакции доступны действия: получение статистической информации о корреспонденции, получение списка писем, чтение выбранного письма, чтение определенного количества строк текста выбранного письма, чтение идентификаторов всех или выбранного сообщения, пометка сообщения как удаленного, сброс отметок об удалении и др.

Стадия обновления. Когда клиент передает команду QUIT в режиме транзакции, то сессия переходит в режим обновления. В этом режиме сервер удаляет все сообщения, помеченные для удаления. После этого TCP-соединение закрывается с соответствующим ответом сервера.

5.4. Задание на лабораторную работу

1. Ознакомиться с теоретическим и справочным материалом.
2. Получить у преподавателя доменное имя или IP-адрес POP3-сервера, с которого возможно получение почтовых сообщений в учебных аудиториях, имя почтового ящика и пароль.
3. Разработать приложение почтового клиента, позволяющее выполнять получение текстовых сообщений по протоколу POP3 из почтового ящика.

Приложение должно поддерживать все приведенные в справочном разделе команды, кроме UIDL. Приложение использует для обмена с сервером транспортный протокол TCP/IP. Процесс обмена командами и ответами должен отображаться на экране.

Предполагается, что в заголовке и тексте сообщения не используются коды русских букв. То есть в приложении допускается отсутствие функций работы с другими методами кодирования, кроме Content-Transfer-Encoding: 7bit.

5.5. Справочные данные

Ниже приведено описание команд и ответов протокола POP3.

USER <SP> <name> <CRLF>

Возможные ответы:

+OK name is a valid mailbox;

-ERR never heard of mailbox name.

Описание. Первый шаг авторизации на сервере. Если ответом на эту команду является строка индикатора «+OK», клиент может отправлять команду PASS – ввод пароля – или QUIT – завершить сессию. Если ответом является строка «-ERR», клиент может либо повторить команду USER, либо закрыть сессию. Аргумент: name – строка, идентифицирующая почтовый ящик системы.

PASS <SP> <passw> <CRLF>

Возможные ответы:

+OK maildrop locked and ready;

-ERR invalid password;

-ERR unable to lock maildrop.

Описание. Второй шаг авторизации на сервере. Если ответом на эту команду является строка индикатора «+OK», авторизация прошла успешно и сессия переходит в состояние транзакции. Если ответом является строка «-ERR», то пользователь либо ввел неправильный пароль, либо неверно указал почтовый ящик. Аргумент: passw – строка-пароль.

QUIT <CRLF>

Возможные ответы:

+OK.

Описание. Завершение сессии. При появлении команды QUIT в состоянии транзакции сессия переходит в состояние обновления и осуществляется удаление помеченных сообщений.

STAT <CRLF>

Возможные ответы:

+OK <n> <s>.

Описание. Команда STAT используется для просмотра текущего состояния почтового ящика: n – количество сообщений, s – их общий объем. Сообщения, помеченные как удаленные, не учитываются.

LIST [*<SP>* *<mes>*] *<CRLF>*

Возможные ответы:

+OK scan listing follows;

<n1> *<s1>*;

<n2> *<s2>*... -ERR no such message.

Описание. Команда LIST может передаваться как с аргументом *mes* – номером сообщения, так и без аргумента. Если команда содержит аргумент и сообщение с указанным номером существует, ответом на нее будет «информационная строка», которая содержит номер сообщения *n1* и размер сообщения *s1* в байтах. Если аргумент не указан – ответом будет список информационных строк *n_i s_i* обо всех сообщениях в данном почтовом ящике. Сообщения, помеченные как удаленные, в этом списке не приводятся.

RETR *<SP>* *<mes>* *<CRLF>*

Возможные ответы:

+OK message follows;

<заголовок и тело сообщения>;

-ERR no such message.

Описание. Используется для передачи клиенту запрашиваемого сообщения. Аргумент команды: *mes* – номер сообщения.

DELE *<SP>* *<mes>* *<CRLF>*

Возможные ответы:

+OK message deleted;

-ERR no such message.

Описание. По команде DELE сообщение с указанным в аргументе *mes* номером помечается как удаленное. Перенумерация остальных сообщений не производится. Физическое удаление помеченных сообщений происходит только при переходе сессии в стадию обновления.

NOOP *<CRLF>*

Возможные ответы:

+OK.

Описание. Используется для проверки состояния сервера. Никаких действий не производится. Ответ всегда положительный.

RSET *<CRLF>*

Возможные ответы:

+OK.

Описание. По команде RSET со всех сообщений снимаются пометки об удалении.

TOP *<SP>* *<mes>* *<SP>* *<n>* *<CRLF>*

Возможные ответы:

+OK;

< заголовок сообщения + первые n строк из тела сообщения>;

-ERR no such message.

Описание. По команде TOP передается заголовок и n первых строк сообщения с номером mes. Если количество строк в сообщении меньше указанного в параметре n, пользователю передается все сообщение. Если n=0, то передается только заголовок.

UIDL [<SP> <mes>] <CRLF>

Возможные ответы:

+OK unique-id listing follows;

<№ сообщения1 + ID1, № сообщения2 + ID2, ...>;

-ERR no such message.

Описание. Если указан номер сообщения mes, то передается уникальный идентификатор для этого сообщения. Если аргумент не был передан, то передаются идентификаторы всех сообщений, кроме помеченных для удаления.

5.6. Контрольные вопросы

1. Охарактеризуйте модель протокола POP3.
2. Назовите и поясните назначение стадий POP3-сессии.
3. Опишите формат почтового сообщения.
4. Назовите особенности и методы кодирования информации в почтовых сообщениях.
5. По каким принципам формируются команды и коды ответов протокола POP3?
6. Какие операции выполняются на стадиях авторизации и обновления?
7. Какие операции выполняются на стадии транзакции?

6. Взаимодействие прикладных программ с помощью протокола передачи данных FTP

Цель работы: изучение принципов организации взаимодействия прикладных программ с помощью протокола передачи данных FTP и приобретение практических навыков создания клиентских приложений, использующих протокол FTP.

Необходимое оборудование: IBM PC-совместимый компьютер с установленным программным обеспечением: Borland Delphi версии 5.0 и выше, FTP- сервер Babyftp 1.0.

6.1. Назначение и модели работы протокола FTP

FTP (File Transfer Protocol – протокол передачи данных) – один из старейших протоколов в Internet и входит в его стандарты. Первые спецификации FTP относятся к 1971 году. С тех пор FTP

претерпел множество модификаций и значительно расширил свои возможности. FTP может использоваться как в программах пользователей, так и в виде специальной утилиты операционной системы.

FTP предназначен для решения задач разделения доступа к файлам на удаленных узлах, прямого или косвенного использования ресурсов удаленных компьютеров, обеспечения независимости клиента от файловых систем удаленных узлов, эффективной и надежной передачи данных.

Обмен данными в FTP происходит по TCP-каналу. Обмен построен на технологии «клиент–сервер». FTP не может использоваться для передачи конфиденциальных данных, поскольку не обеспечивает защиты передаваемой информации и передает между сервером и клиентом открытый текст. FTP-сервер может потребовать от FTP-клиента аутентификации. Однако пароль и идентификатор пользователя будут переданы от клиента на сервер открытым текстом.

Простейшая модель работы протокола FTP представлена на рис. 6.1. FTP-соединение инициируется интерпретатором протокола пользователя. Управление обменом осуществляется по каналу управления в стандарте протокола TELNET. Команды FTP генерируются интерпретатором протокола пользователя и передаются на сервер. Ответы сервера отправляются пользователю также по каналу управления

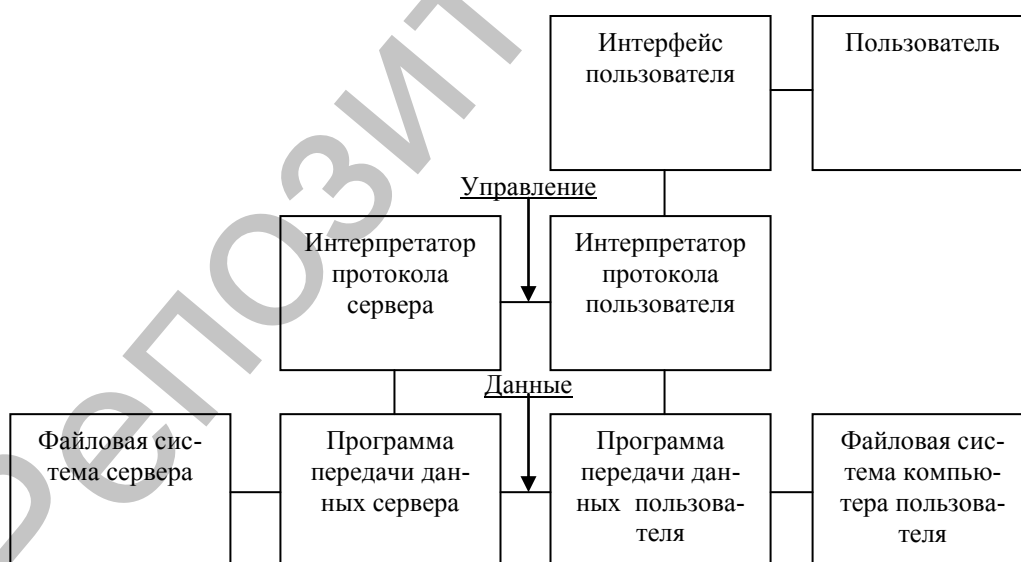


Рис. 6.1. Модель работы протокола FTP.

Команды FTP определяют параметры канала передачи данных и самого процесса передачи. Они также определяют и характер работы с удаленной и локальной файловыми системами.

Сессия управления инициализирует канал передачи данных. При организации канала передачи данных последователь-

ность действий другая, отличная от организации канала управления. В этом случае сервер инициирует обмен данными в соответствии с согласованными в сессии управления параметрами.

Канал данных устанавливается для того же узла, что и канал управления, через который ведется настройка канала данных. Канал данных может быть использован как для приема, так и для передачи данных.

Алгоритм работы протокола FTP состоит в следующем:

1. Сервер FTP использует в качестве управляющего соединения на TCP порт 21, который всегда находится в состоянии ожидания соединения со стороны пользователя FTP.

2. После того как устанавливается управляющее соединение модуля «Интерпретатор протокола пользователя» с модулем сервера – «Интерпретатор протокола сервера», пользователь (клиент) может отправлять на сервер команды. FTP-команды определяют параметры соединения передачи данных: роль участников соединения (активный или пассивный), порт соединения (как для модуля «Программа передачи данных пользователя», так и для модуля «Программа передачи данных сервера»), тип передачи, тип передаваемых данных, структуру данных и управляющие директивы, обозначающие действия, которые пользователь хочет совершить (например, сохранить, считать, добавить или удалить данные или файл и другие).

3. После того как согласованы все параметры канала передачи данных, один из участников соединения (пассивный модуль) становится в режим ожидания открытия соединения на определенный порт. После этого другая сторона (активный модуль) открывает соединение на указанный порт и начинается обмен данными.

4. После окончания передачи данных соединение между «Программой передачи данных сервера» и «Программой передачи данных пользователя» закрывается, но управляющее соединение «Интерпретатора протокола сервера» и «Интерпретатора протокола пользователя» остается открытым. Пользователь, не закрывая сессии FTP, может еще раз открыть канал передачи данных.

Возможна ситуация, когда данные должны передаваться на третий узел. В этом случае имеет место другая модель работы протокола FTP, при которой пользователь организует канал управления с двумя серверами и прямой канал данных между ними. Команды управления идут через пользователя, а данные – напрямую между серверами. Подробное рассмотрение этой модели работы протокола FTP выходит за рамки данной лабораторной работы.

6.2. Особенности управления процессом обмена данными

Основу передачи данных FTP составляет механизм установления соединения между соответствующими портами и выбора параметров передачи. Каждый участник FTP-соединения должен поддерживать порт передачи данных по умолчанию. По умолчанию «Программа передачи данных пользователя» использует тот же порт, что и для передачи команд, а «Программа передачи данных сервера» использует порт L-1, где L – управляющий порт. Однако для ускорения процессов обмена данными участники соединения часто используют другие порты передачи данных.

Передача данных может вестись в активном или пассивном режиме. Если доступ к ftp-серверу осуществляется через прокси-сервер, то возможна работа только в пассивном режиме.

Установление соединения передачи данных при активном режиме работы происходит следующим образом:

1. Клиент создает сокет на выбранном им порте P и активизирует его (переводит в режим ожидания соединения).

2. Клиент направляет серверу по управляющему соединению команду PORT (описание приведено ниже), в которой указывает свой IP-адрес и выбранный для передачи данных порт P.

3. В случае успешного получения и обработки команды PORT сервер по управляющему соединению отправляет клиенту положительный отклик. Затем сервер пытается соединиться с портом P клиента со своего локального порта данных A (может быть выбран сервером произвольно).

4. Клиент обнаруживает событие соединения на порт P и начинает процесс передачи данных с использованием управляющего соединения для отправки команд FTP-сервиса и соединения передачи данных для получения либо отправки данных.

Установление соединения передачи данных при пассивном режиме работы происходит следующим образом:

1. Клиент отправляет по управляющему соединению команду PASV, указывающую серверу о намерении клиента работать в пассивном режиме.

2. В случае успешного получения и обработки команды PASV сервер создает сокет на выбранном им порте P, активизирует его (переводит его в режим ожидания соединения) и отправляет клиенту по управляющему соединению отклик, в котором указываются параметры соединения (IP-адрес сервера и порт P).

3. Клиент, получив положительный отклик с параметрами соединения, создает на произвольном порте A сокет и активизирует его, то есть пытается соединиться с портом P сервера.

4. В случае успешного установления соединения клиент начинает процесс передачи данных с использованием управляющего соединения для отправки команд FTP-сервиса и соединения передачи данных для получения либо отправки данных.

Одновременно с передачей данных по установленному соединению в обоих режимах по каналу «Интерпретатор протокола сервера» – «Интерпретатор протокола пользователя» могут передаваться уведомления о получении данных. Протокол FTP требует, чтобы управляющее соединение было открыто, пока по каналу обмена данными идет передача.

Как правило, сервер FTP ответственен за открытие и закрытие канала передачи данных. Сервер FTP должен самостоятельно закрыть канал передачи данных в следующих случаях:

1. Сервер закончил передачу данных в формате, который требует закрытия соединения.
2. Сервер получил от пользователя команду «прервать соединение».
3. Пользователь изменил параметры порта передачи данных.
4. Было закрыто управляющее соединение.
5. Возникли ошибки, при которых невозможно возобновить передачу данных.

6.3. Команды и ответы протокола FTP

Все команды протокола FTP отправляются «Интерпретатором протокола пользователя» в текстовом виде – по одной команде в строке. Каждая строка команды – идентификатор и аргументы – заканчиваются символами CRLF. Имя команды отделяется от аргумента символом пробела.

Обработчик команд возвращает код обработки каждой команды, состоящий из трех цифр. Коды обработки составляют определенную иерархическую структуру и, как правило, определенная команда может вернуть только определенный набор кодов. За кодом обработки команды следует символ пробела и текст пояснения. Описание команд и основных кодов ответов приведено в пункте 6.5 «Справочные данные».

Команды протокола FTP, которыми обмениваются «Интерпретатор протокола сервера» и «Интерпретатор протокола пользователя», можно разделить на три группы.

1. *Команды управления доступом к системе* обеспечивают авторизацию пользователя в системе, выход из нее и настройку некоторых текущих параметров соединения.

2. Команды управления потоком данных устанавливают параметры передачи данных. Все параметры, описываемые этими командами, имеют значение по умолчанию, поэтому команды управления потоком используются только тогда, когда необходимо изменить значение параметров передачи, используемых по умолчанию. Команды управления потоком могут подаваться в любом порядке, но все они должны предшествовать командам FTP-сервиса.

3. Команды FTP-сервиса определяют действия, которые необходимо произвести с указанными файлами. Как правило, аргументом команд этой группы является имя файла.

6.4. Задание на лабораторную работу

1. Ознакомиться с теоретическим и справочным материалом.

2. Запустить приложение простейшего ftp-сервера «babyftp.exe». Сервер поддерживает работу только с анонимными пользователями (имя пользователя – «anonymous», пароль – любой адрес электронного почтового ящика). Параметры доступа настраиваются в окне «Settings».

3. Разработать приложение ftp-клиента, реализующие следующие функции:

- создание и удаление папок;
- перемещение по папкам;
- чтение списка содержимого папки;
- получение, отправка, удаление и переименование файлов.

Для получения списка содержимого папки необходимо использовать активный режим. Для получения и отправки файлов – пассивный. Приложение использует для обмена с сервером транспортный протокол TCP/IP. Процесс обмена командами и ответами должен отображаться на экране.

Примечание: для отправки файлов на сервер необходимо использовать метод `SendStreamThenDrop` класса `TCustomWinSocket`. Данный метод обеспечивает автоматическое закрытие соединения после завершения передачи файла. В этом случае для связывания метода с файлом необходимо создать файловый поток (класс `TFileStream`).

6.5. Справочные данные

Команды протокола FTP.

1. Команды управления доступом к системе:

USER. Как правило, эта команда открывает сессию FTP между клиентом и сервером. Аргументом команды является имя

(идентификатор) пользователя для работы с файловой системой. Эта команда может подаваться не только в начале, но и в середине сессии, если, например, пользователь желает изменить идентификатор, от имени которого будут проводиться действия. При этом все переменные, относящиеся к старому идентификатору, освобождаются. Если во время изменения идентификатора происходит обмен данными, обмен завершается со старым идентификатором пользователя.

PASS. Данная команда подается после ввода идентификатора пользователя и содержит в качестве аргумента пароль пользователя.

CWD. Команда обеспечивает возможность работы с различными каталогами удаленной файловой системы. Аргументом команды является строка, указывающая путь каталога удаленной файловой системы, в котором желает работать пользователь.

REIN. Команда реинициализации. Эта команда очищает все переменные текущего пользователя, сбрасывает параметры соединения. Если в момент подачи команды происходит передача данных, передача продолжается и завершается с прежними параметрами.

QUIT. Команда закрывает управляющий канал. Если в момент подачи команды происходит передача данных, канал закрывается после окончания передачи данных.

2. Команды управления потоком данных:

PORT. Команда указывает серверу адрес и порт, которые будут использоваться клиентом для прослушивания соединения. Синтаксис команды: «PORT<SP>h1,h2,h3,h4,p1,p2». Аргументами команды являются 32-битный IP адрес и 16-битный номер порта соединения. Эти значения разбиты на шесть 8-битных полей и представлены в десятичном виде, где hN – байты адреса (от старшего к младшему), а pN – байты порта (от старшего к младшему).

PASV. Эта команда отправляется серверу для указания, что он должен «слушать» соединение. Ответом на данную команду является строка, содержащая адрес и порт узла, находящегося в режиме ожидания соединения. Формат представления данных соответствует формату команды PORT.

Команды *TYPE*, *STRU*, *MODE* определяют, соответственно, тип передаваемых данных (ASCII, Image и другие), структуру или формат передачи данных (File, Record, Page), способ передачи (Stream, Block и другие). Использование этих команд очень важно при построении взаимодействия в гетерогенных средах и весьма отличающихся операционных и файловых систем взаимодействующих узлов.

3. Команды FTP-сервиса:

RETR. Эта команда указывает модулю «Программа передачи данных сервера» передать копию файла, заданного параметром этой команды, модулю передачи данных на другом конце соединения.

STOR. Команда указывает модулю «Программа передачи данных сервера» принять данные по каналу передачи данных и сохранить их как файл, имя которого задано параметром этой команды. Если такой файл уже существует, он будет перезаписан, если нет, будет создан новый.

Таблица 6.1

Основные коды ответов на команды протокола FTP

Код	Описание
110	Комментарий
125	Канал открыт, обмен данными начат
150	Статус файла правилен, подготавливается открытие канала
200	Команда корректна
220	Слишком много подключений к FTP-серверу (можете попробовать позднее). В некоторых версиях указывает на успешное завершение промежуточной процедуры
221	Успешное завершение по команде quit
225	Канал сформирован, но информационный обмен отсутствует
226	Закрытие канала, обмен завершен успешно
230	Пользователь идентифицирован, продолжайте
250	Запрос прошел успешно
331	Имя пользователя корректно, нужен пароль
332	Для входа в систему необходима аутентификация
421	Процедура невозможна, канал закрывается
425	Открытие информационного канала невозможно
426	Канал закрыт, обмен прерван
450	Запрошенная функция не реализована, файл недоступен, например, занят
451	Локальная ошибка, операция прервана
452	Ошибка при записи файла (недостаточно места)
500	Синтаксическая ошибка, команда не может быть интерпретирована (возможно, она слишком длинна)

501	Синтаксическая ошибка (неверный параметр или аргумент)
503	Неудачная последовательность команд
504	Команда не применима для такого параметра

RNFR и *RNTO*. Команды должны следовать одна за другой. Первая команда содержит в качестве аргумента старое имя файла, вторая – новое. Последовательное применение этих команд переименовывает файл.

ABOR. Команда предписывает серверу прервать выполнение предшествующей сервисной команды (например, передачу файла) и закрыть канал передачи данных.

DELE. Удаление указанного файла.

MKD. Создание указанного в аргументе каталога.

RMD. Удаление указанного в аргументе каталога.

LIST. Получение списка файлов в указанном каталоге. Передача списка осуществляется по соединению «Программа передачи данных сервера» – «Программа передачи данных клиента».

6.6. Контрольные вопросы

1. Сколько логических соединений необходимо для работы протокола FTP? Каким образом они используются?
2. Какие стадии можно выделить в FTP-сессии?
3. В чем заключается особенность работы протокола FTP в активном режиме?
4. В чем заключается особенность работы протокола FTP в пассивном режиме?
5. Какая сторона соединения отвечает за закрытие канала передачи данных? В каких случаях это происходит?
6. Опишите структуру команды и ответа протокола FTP.
7. На какие группы можно разбить команды протокола FTP? Каково назначение команд каждой из групп?

СПИСОК ИСПОЛЬЗУЕМОЙ ЛИТЕРАТУРЫ

1. Применение цифровой обработки сигналов / под ред. Э. Оппенгеймера; пер. с англ. – М.: Мир, 1980.
2. Шелухин О.И., Лукьянцев Н.Ф. Цифровая обработка и передача речи. – М.: «Радио и связь», 2000.
3. Синепол В.С., Цикин И.А. Системы компьютерной видеоконференцсвязи. – М.: ООО «Мобильные коммуникации», 1999.
4. Компьютерные сети: учебный курс / пер. с англ. – М.: Издательский отдел «Русская Редакция» ТОО «Channel Trading Ltd.», 1997.
5. Новиков Ю.В., Карпенко Д.Г. Аппаратура локальных сетей. – М.: Изд-во «Эком», 1998.
6. Гроднев И.И. Волоконно-оптические линии связи. – М.: Радио и связь, 1990.
7. Берлин Б.З. и др. Волоконно-оптические системы связи на ГТС. Справочник / Б.З. Берлин, А.С. Брискер, В.С. Иванов; под ред. А.С. Брискера и А.Н. Голубева. – М.: Радио и связь, 1994.
8. Уайндер С. Справочник по технологиям и средствам связи; пер. с англ. – М.: Мир, 2000.
9. Радиотехнические системы передачи информации: учеб. пособие для вузов / В.А. Борисов, В.В. Калмыков, Я.М. Ковальчук, И.А. Цикин и др.; под ред. В.В. Калмыкова. – М.: Радио и связь, 1990.
10. Левин Л.С., Плоткин М.А. Цифровые системы передачи информации. – М.: Радио и связь, 1982.
11. Андрианов В.И., Соколов А.В. / Средства мобильной связи. – СПб.: ВНУ-Санкт-Петербург, 1998.
12. В.А. Ключников. Современные информационные технологии. – Минск: ООО «БТН-информ», 2001.
13. А.И. Фролов. Сети ЭВМ и телекоммуникации. – Орел, 2006.
14. Баканов В.М. Сети ЭВМ и телекоммуникации: учеб.-метод. пособие. – М.: МГУПИ, 2008.

Репозиторий ВГУ