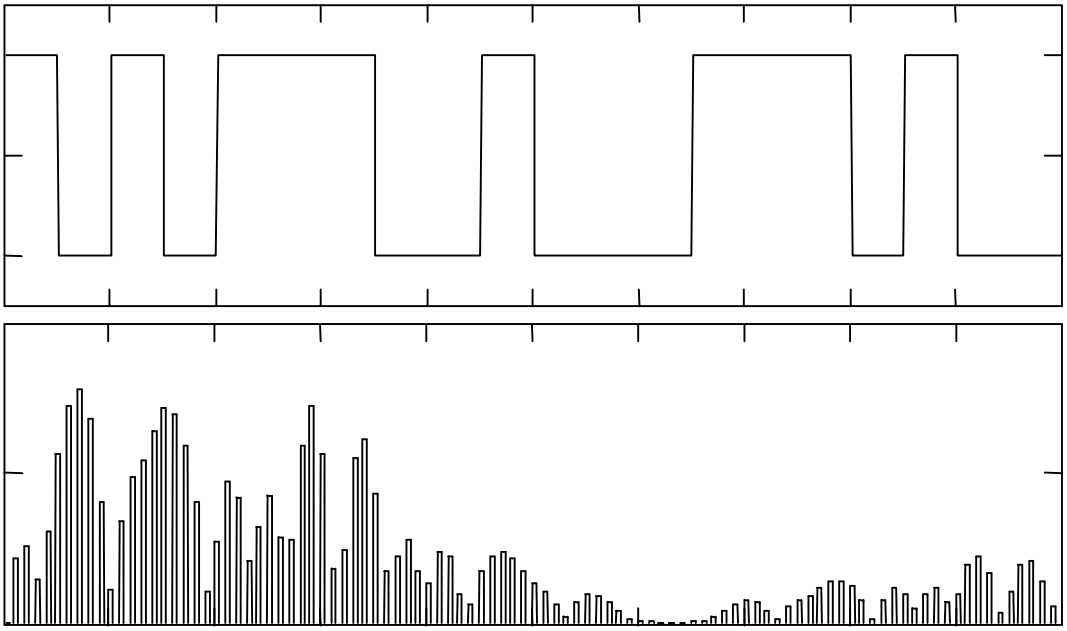
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ (ТУСУР)

С.Г. Михальченко, Е.Ю. Агеев

ЭКСПЛУАТАЦИЯ И РАЗВИТИЕ КОМПЬЮТЕРНЫХ СИСТЕМ И СЕТЕЙ

Раздел 1

Учебное пособие



ТОМСК — 2007

Федеральное агентство по образованию

ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ (ТУСУР)

Кафедра промышленной электроники

С.Г. Михальченко, Е.Ю. Агеев

ЭКСПЛУАТАЦИЯ И РАЗВИТИЕ КОМПЬЮТЕРНЫХ СИСТЕМ И СЕТЕЙ

Раздел 1

Учебное пособие

2007

Михальченко С.Г., Агеев Е.Ю.
Эксплуатация и развитие компьютерных систем и сетей: Учебное пособие. В 2-х разделах. — Томск: Томский государственный университет систем управления и радиоэлектроники, 2007. — Раздел 1. — 216 с.

Рассмотрены вопросы эксплуатации и развития компьютерных систем и сетей, вопросы доступа к среде передачи информации, способы коммутации и мультиплексирования, методы кодирования и адресации сетевых устройств. Изучаются вопросы установки, настройки и обслуживания аппаратного и программного обеспечения компьютерных информационных сетей.

Предназначено для студентов вузов, обучающихся по специальности «Промышленная электроника».

© Михальченко С.Г.,
Агеев Е.Ю., 2007
© ТУСУР, 2007

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ

Курс «Эксплуатация и развитие компьютерных систем и сетей» (ЭРКСС) совместно с курсами «Аппаратное и программное обеспечение ЭВМ» и «Операционные системы» составляет основу подготовки инженеров специальности «Промышленная электроника» в области компьютерных сетей и играет роль базы, без которой невозможна успешная деятельность инженера в области компьютерной техники и технологий.

Целью настоящего учебного пособия является помощь студентам при изучении курса ЭРКСС, практическое закрепление знаний по современным и классическим сетевым технологиям, способам построения компьютерных сетей и проведения необходимых расчетов при их проектировании.

Глава 1 настоящего пособия посвящена описанию основных принципов передачи данных. В ней рассматриваются следующие понятия: канал и линия связи, протоколы, пакеты, инкапсуляция сообщений и адресация компьютеров. Описывается эталонная модель взаимосвязи открытых сетей (OSI).

Вторая и третья главы посвящены изучению большинства применяемых в информационных сетях линий связи и характеристик канала передачи данных.

В главе 4 рассматриваются основные вопросы и способы модуляции и кодирования как физического, так канального и прикладного уровней компьютерных сетей.

Пятая глава настоящей работы описывает основные протоколы и стандарты локальных компьютерных сетей. Рассматриваются протоколы физического и канального уровня OSI.

В шестой главе детализировано рассматривается наиболее популярный в настоящее время стек сетевых протоколов TCP/IP.

Глава 7 посвящена изучению оборудования информационных сетей, описываются его основные функции, способы и цели применения. Рассматриваются способы настройки и конфигурирования.

Заключительная глава настоящей работы «Безопасность компьютерных сетей» описывает основные направления атак на компьютерные сети и способы противостояния этим атакам.

Таким образом, настоящее учебное пособие закрывает все основные темы, предложенные государственным стандартом РФ к изучению бакалаврами, студентами и магистрами специальности «промышленная электроника».

1 ПРИНЦИПЫ ПЕРЕДАЧИ ДАННЫХ

1.1 Телематика

Телематика — это научно-техническая дисциплина, изучающая методы и средства передачи информации на расстояния, существенно превышающие линейные размеры площади, занимаемой участниками связи. Название дисциплины произошло из частей слов *телекоммуникации* и *информатика*.

Передача данных по информационным сетям осуществляется аналогично связи программного продукта, расположенного на компьютере и некоторого *периферийного устройства* (ПУ). Программа, которой потребовалось выполнить обмен данными с ПУ, обращается к драйверу этого устройства, сообщая ему в качестве параметра адрес байта памяти, который нужно передать. Драйвер загружает значение этого байта в буфер контроллера ПУ, который начинает последовательно передавать биты в линию связи, представляя каждый бит соответствующим электрическим сигналом. Чтобы устройству управления ПУ стало понятно, что начинается передача байта, перед передачей первого бита информации контроллер ПУ формирует стартовый сигнал специфической формы, а после передачи последнего информационного бита — стоповый сигнал. Эти сигналы синхронизируют передачу байта.

Таким же образом, в самом простом случае, может быть реализовано взаимодействие компьютеров с помощью тех же самых средств, которые используются для взаимодействия компьютера с периферией, например, через последовательный интерфейс RS-232C. В этом случае происходит взаимодействие двух программ, работающих на каждом из компьютеров. Программа, работающая на одном компьютере, не может получить непосредственный доступ к ресурсам другого компьютера — его дискам, файлам, принтеру. Она может только *попросить* об этом программу, работающую на том компьютере, которому принадлежат эти ресурсы. Эти *просьбы* выражаются в виде сообщений, передаваемых по каналам связи между компьютерами. Сообщения могут содержать не только команды на выполнение некоторых действий, но и собственно информационные данные (например, содержимое некоторого файла).

Допустим, пользователю, работающему с текстовым редактором на персональном компьютере *A*, нужно прочитать часть некоторого файла, расположенного на диске персонального компьютера *B*. Предположим, что мы связали эти компьютеры по кабелю связи через COM-порты, которые реализуют интерфейс RS-232C (такое соединение часто называют нуль-модемным). В вычислительных сетях подобные функции передачи данных в линию связи выполняются, естественно, сетевыми адаптерами и их драйверами.

Схема передачи запросов для доступа к файлу на диске компьютера B приведена на рис. 1.1. Приложение *A* должно сформировать сообщение-запрос для приложения *B*. В запросе необходимо указать имя файла, тип операции (в данном случае — чтение), смещение и размер файловой области, содержащей нужные данные. Чтобы передать это сообщение компьютеру *B*, приложение *A* обращается к драйверу COM-порта, сообщая ему адрес в оперативной памяти, по которому драйвер находит сообщение и затем передает его байт за байтом приложению *B*. Приложение *B*, приняв запрос, выполняет его, то есть считывает требуемую область файла с диска с помощью средств локальной ОС в буферную область своей оперативной памяти, а далее с помощью драйвера COM-порта передает считанные данные по каналу связи в компьютер *A*, где они и попадают к приложению *A*.

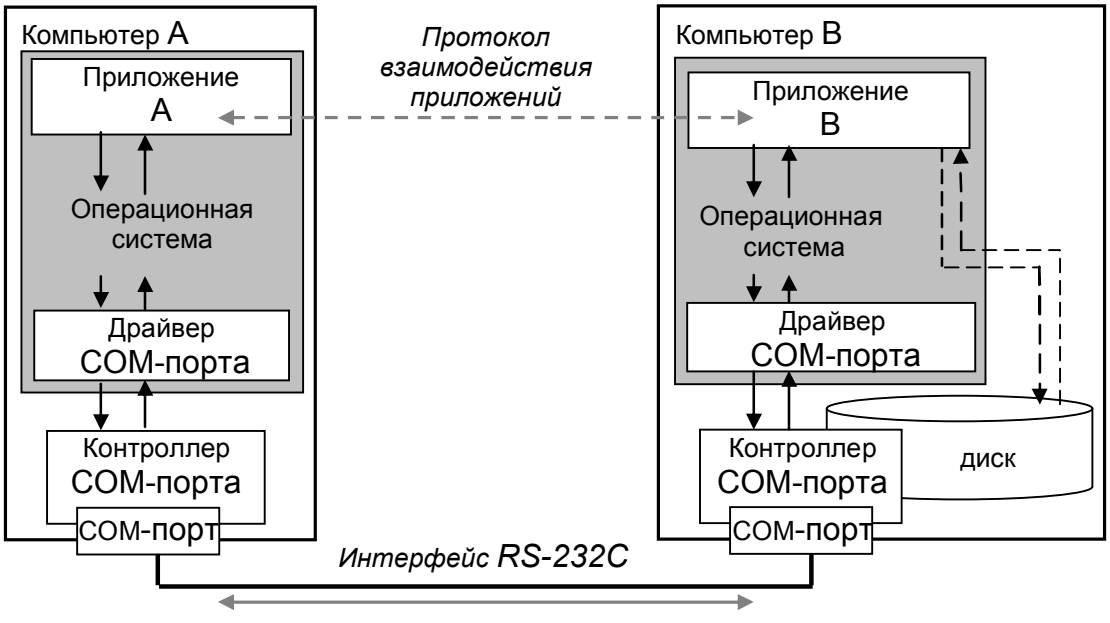


Рис. 1.1 — Взаимодействие двух приложений

Описанные функции приложения *A* могла бы выполнить сама программа текстового редактора, но включать эти функции в состав каждого приложения — текстовых редакторов, графических редакторов, систем управления базами данных и других приложений, которым нужен доступ к файлам, — не очень рационально (хотя существует большое количество программ, которые действительно самостоятельно решают все задачи по межмашинному обмену данными). Гораздо разумнее создать специальный программный модуль, который будет выполнять функции формирования сообщений-запросов и приема результатов для всех приложений компьютера. Такой служебный модуль называется ***клиентом***. На стороне же компьютера *B* должен работать другой модуль — ***сервер***, постоянно ожидающий прихода запросов на удаленный доступ к файлам, расположенным на диске этого компьютера. Сервер, приняв запрос из сети, обращается к локальному файлу и выполняет с ним заданные действия, возможно, с участием локальной ОС.

Программные клиент и сервер выполняют системные функции по обслуживанию запросов приложений компьютера *A* на удаленный доступ к файлам компьютера *B*. Чтобы приложения компьютера *B* могли пользоваться файлами компьютера *A*, описанную схему нужно симметрично дополнить клиентом для компьютера *B* и сервером для компьютера *A*.

Очень удобной и полезной функцией клиентской программы является способность отличить запрос к удаленному файлу от запроса к локальному файлу. Если клиентская программа умеет это делать, то приложения не должны заботиться о том, с каким файлом они работают (локальным или удаленным), клиентская программа сама распознает и перенаправляет (***redirect***) запрос к удаленной машине. Отсюда и название, часто используемое для клиентской части сетевой ОС, — *редиректор*.

С точки зрения аппаратного обеспечения, в локальных сетях задачи обмена данными возложены на сетевые адаптеры, в глобальных сетях — на аппаратуру передачи данных. Программные средства, реализующие передачу информации, включают классические элементы сетевой операционной системы: сервер, клиент и средства транспортировки сообщений по линии.

Программные и аппаратные средства, обеспечивающие связь (возможность передачи данных) между компьютерами изучает наука *телематика*. Современные телекоммуникационные технологии основаны на использовании информационных сетей.

1.2 Каналы и линии связи

Коммуникационная сеть — это система, состоящая из объектов, осуществляющих функции генерации, преобразования, хранения и потребления некоторого продукта (такие объекты называются пунктами или *узлами* сети), и линий передачи (связей, коммуникаций, соединений), осуществляющих передачу продукта между пунктами.

Отличительная особенность коммуникационных сетей — большие расстояния между узлами по сравнению с геометрическими размерами участков пространств, занимаемых ими. В качестве продукта могут фигурировать информация, энергия, вещество, и соответственно различают группы сетей информационных, энергетических, вещественных. В группах сетей возможно разделение на подгруппы. Так, среди вещественных сетей могут быть выделены сети транспортные, водопроводные, производственные и др. При функциональном проектировании сетей решаются задачи синтеза топологии, распределения продукта по узлам сети, а при конструкторском — выполняются размещение пунктов в пространстве и проведение (трассировка) соединений.

Информационная сеть — коммуникационная сеть, в которой продуктом генерирования, переработки, хранения и использования является информация.

Вычислительная сеть — информационная сеть, в состав которой входит вычислительное оборудование. Компонентами вычислительной сети могут быть ЭВМ и периферийные устройства, являющиеся источниками и приемниками данных, передаваемых по сети (см. рис. 1.2).

Аппаратура пользователя линии связи, вырабатывающая данные для передачи по линии связи и подключаемая к аппаратуре передачи данных называется ***оконечным оборудованием данных (ООД), Data Terminal Equipment*** (***DTE***). В качестве DTE могут выступать ЭВМ, промышленные контроллеры, принтеры и

другое вычислительное, измерительное и исполнительное оборудование автоматических и автоматизированных систем. Эту аппаратуру не включают в состав линии связи.

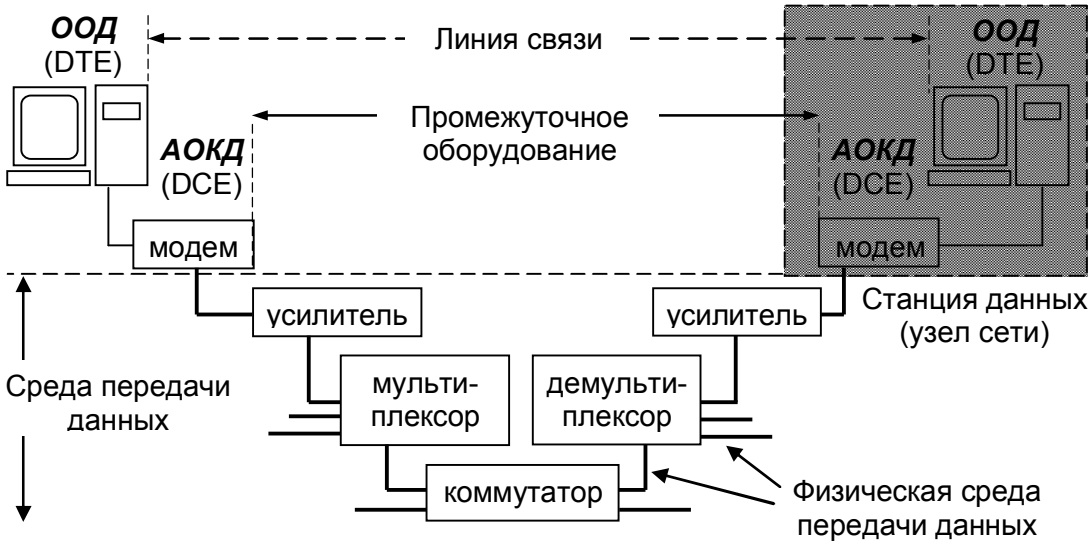


Рис. 1.2 — Среда передачи данных

Собственно пересылка данных происходит с помощью средств, объединяемых под названием ***среда передачи данных.***

Подготовка данных, передаваемых или получаемых DTE от среды передачи данных, осуществляется функциональным блоком, называемым ***аппаратурой окончания канала данных (АОКД) Data Circuit-Terminating Equipment*** (***DCE***). DCE может быть конструктивно отдельным или встроенным в DTE блоком. Примером DCE может служить модем.

Оборудование DTE и DCE вместе представляют *станцию данных*, которую часто называют ***узлом сети***. Под ***средой передачи данных*** понимается совокупность линий передачи данных и блоков взаимодействия (сетевого оборудования не входящего в станции данных) предназначенных для передачи данных между узлами сети. Среды передачи данных могут быть общего пользования или выделенными для конкретного пользователя.

Здесь понятие ***линии передачи данных (линии связи)*** — включает в себя средства, используемые в информационной сети для распространения данных в нужном направлении. Линия связи состоит из физической среды, по которой передаются электриче-

ские информационные сигналы, аппаратурой окончания канала данных (DCE) и промежуточного оборудования.

Промежуточное оборудование обычно используется на линиях связи большой протяженности. Промежуточная аппаратура решает две основные задачи — улучшение качества сигнала и создание постоянного составного канала связи между двумя абонентами сети.

В локальных сетях промежуточная аппаратура может совсем не использоваться, если протяженность физической среды (кабелей или радиоэфира) позволяет одному сетевому адаптеру принимать сигналы непосредственно от другого сетевого адаптера, без промежуточного усиления. В противном случае применяются устройства типа повторителей и концентраторов.

В глобальных сетях необходимо обеспечить качественную передачу сигналов на расстояния в сотни и тысячи километров. Поэтому без усилителей сигналов, установленных через определенные расстояния, построить территориальную линию связи невозможно. В глобальной сети необходима также и промежуточная аппаратура другого рода — мультиплексоры, демультиплексоры и коммутаторы. Эта аппаратура решает вторую указанную задачу, то есть создает между двумя абонентами сети составной канал из некоммутируемых отрезков физической среды — кабелей с усилителями.

Наличие промежуточной коммутационной аппаратуры избавляет создателей глобальной сети от необходимости прокладывать отдельную кабельную линию для каждой пары соединяемых узлов сети. Вместо этого между мультиплексорами и коммутаторами используется высокоскоростная физическая среда, например волоконно-оптический или коаксиальный кабель, по которому передаются одновременно данные от большого числа сравнительно низкоскоростных абонентских линий. Высокоскоростной канал так же называют ***уплотненным*** каналом.

Промежуточная аппаратура канала связи прозрачна для пользователя, он ее не замечает и не учитывает в своей работе. Для него важны только качество полученного канала, влияющее на скорость передачи дискретных данных. В действительности же промежуточная аппаратура образует сложную сеть, которую называют *первичной сетью,* так как сама по себе она никаких вы-

сокоуровневых служб (например, файловой или передачи голоса) не поддерживает, а только служит основой для построения компьютерных, телефонных или иных сетей (рис. 1.1).

Канал (***канал связи, channel***) — средства односторонней передачи данных. Примером канала может быть полоса частот, выделенная одному передатчику при радиосвязи. В некоторой линии можно образовать несколько каналов связи, по каждому из которых передается своя информация. При этом говорят, что *линия разделяется между несколькими каналами.*

1.3 Протоколы и пакеты

Для обеспечения совместимости аппаратно-программных средств взаимодействия узлов сети используется определенный *протокол* — совокупность соглашений относительно способа представления данных, обеспечивающего их передачу в нужных направлениях и правильную интерпретацию данных всеми участниками процесса информационного обмена.

Протокол — это набор семантических и синтаксических правил, определяющий поведение функциональных блоков сети при передаче данных.

Поскольку информационный обмен — процесс многофункциональный, то протоколы делятся на уровни. К каждому уровню относится группа родственных функций. Модули, реализующие протоколы соседних уровней и находящиеся в одном узле, взаимодействуют друг с другом в соответствии с четко определенными правилами и с помощью стандартизованных форматов сообщений. Эти правила принято называть *интерфейсом.*

Интерфейс определяет набор сервисов, предоставляемый данным уровнем соседнему уровню. В сущности, протокол и интерфейс выражают одно и то же понятие, но традиционно в сетях за ними закрепили разные области действия: протоколы определяют правила взаимодействия модулей одного уровня в разных узлах, а интерфейсы — модулей соседних уровней в одном узле. Средства каждого уровня должны отрабатывать, во-первых, свой собственный протокол, а во-вторых, интерфейсы с соседними уровнями.

Иерархически организованный набор протоколов, достаточный для организации взаимодействия узлов в сети, называется ***стеком коммуникационных протоколов***.

Коммуникационные протоколы могут быть реализованы как программно, так и аппаратно. Протоколы нижних уровней часто реализуются комбинацией программных и аппаратных средств, а протоколы верхних уровней — как правило, чисто программными средствами.

Протоколы реализуются не только компьютерами, но и другими сетевыми устройствами — концентраторами, мостами, коммутаторами, маршрутизаторами и т.д. Действительно, в общем случае связь компьютеров в сети осуществляется не напрямую, а через различные коммуникационные устройства. В зависимости от типа устройства в нем должны быть встроенные средства, реализующие тот или иной набор протоколов.

Компьютеры в сети обмениваются информацией друг с другом ***пакетами сообщений*** — адресованных информационных блоков. Пакеты составляют фундамент, на котором базируется работа ЛВС. Сетевой адаптер ЛВС осуществляет прием и передачу пакетов под управлением соответствующего программного обеспечения. Пакеты адресуются рабочим станциям, каждая из которых должна иметь уникальный адрес в ЛВС. Пакеты могут нести различную информацию в ЛВС:

- начало сеанса обмена данными;
- передача данных (возможно, записи из файла) другому ПК;
- подтверждение приема пакета данных;
- передача широковещательного сообщения всем адаптерам;
- конец сеанса обмена данными.

В различных системах компьютерных сетей пакеты определяются по-разному, но следующие элементы являются общими для всех:

- уникальный адрес отправителя;
- уникальный адрес получателя;
- признак, определяющий содержимое пакета;
- собственно данные или сообщение;
- контрольная сумма (CRC) для обнаружения ошибок при передаче.

В настоящее время стандартом в области идеологии построения протоколов является модель *OSI* (*Open System Interconnection*) — *ЭМВОС* (эталонная модель взаимосвязи открытых систем), принятая для описания общих принципов взаимодействия информационных систем. ЭМВОС признана всеми международными организациями как основа для стандартизации протоколов информационных сетей.

В ЭМВОС информационная сеть рассматривается как совокупность функций, которые делятся на уровни. Разделение на уровни позволяет вносить изменения в средства реализации одного уровня без перестройки средств других уровней, что значительно удешевляет и упрощает модернизацию средств по мере развития техники.

1.4 Эталонная модель взаимосвязи открытых сетей (OSI)

Как уже говорилось выше, организация ***ISO*** (***International Standards Organization*** — Международная организация по стандартизации) опубликовала модель архитектуры вычислительной сети, названной ***OSI*** (***Open System Interconnection*** — Связь открытых систем). Эталонная модель взаимодействия открытых систем OSI определяет общие принципы взаимодействия информационных систем и представляет собой основы стандартизации протоколов.

Модель OSI описывает только системные средства взаимодействия, реализуемые операционной системой, системными утилитами, системными аппаратными средствами. Модель не включает средства взаимодействия приложений конечных пользователей. Свои собственные протоколы взаимодействия приложения реализуют, обращаясь к системным средствам. Поэтому необходимо различать уровень взаимодействия приложений и прикладной уровень OSI.

Следует также иметь в виду, что приложение может взять на себя функции некоторых верхних уровней модели OSI. Например, некоторые СУБД имеют встроенные средства удаленного доступа к файлам. В этом случае приложение, выполняя доступ к удаленным ресурсам, не использует системную файловую служ-

бу; оно обходит верхние уровни модели OSI и обращается напрямую к системным средствам, ответственным за транспортировку сообщений по сети, которые располагаются на нижних уровнях модели OSI.

Модель OSI разделяет коммуникационные функции в ЛВС на семь уровней (см. рис. 1.3). Сообщение, предназначенное для передачи через компьютерную сеть, из прикладного уровня передается по интерфейсам уровней от одного до другого, проходит до физического уровня, пересылается на другую рабочую станцию, проходит от нижнего уровня в обратном порядке до достижения прикладной программы на другой рабочей станции через ее прикладной уровень. Поэтому интерфейсы между уровнями называются *точками доступа к службам* (***Service Access Points, SAP***). Уровни обмениваются между собой сообщениями, которые называются *протокольными блоками данных* (***Protocol Data Units, PDU***). Синтаксис протоколов определяет *формат*, а семантика — *значение* сообщений.

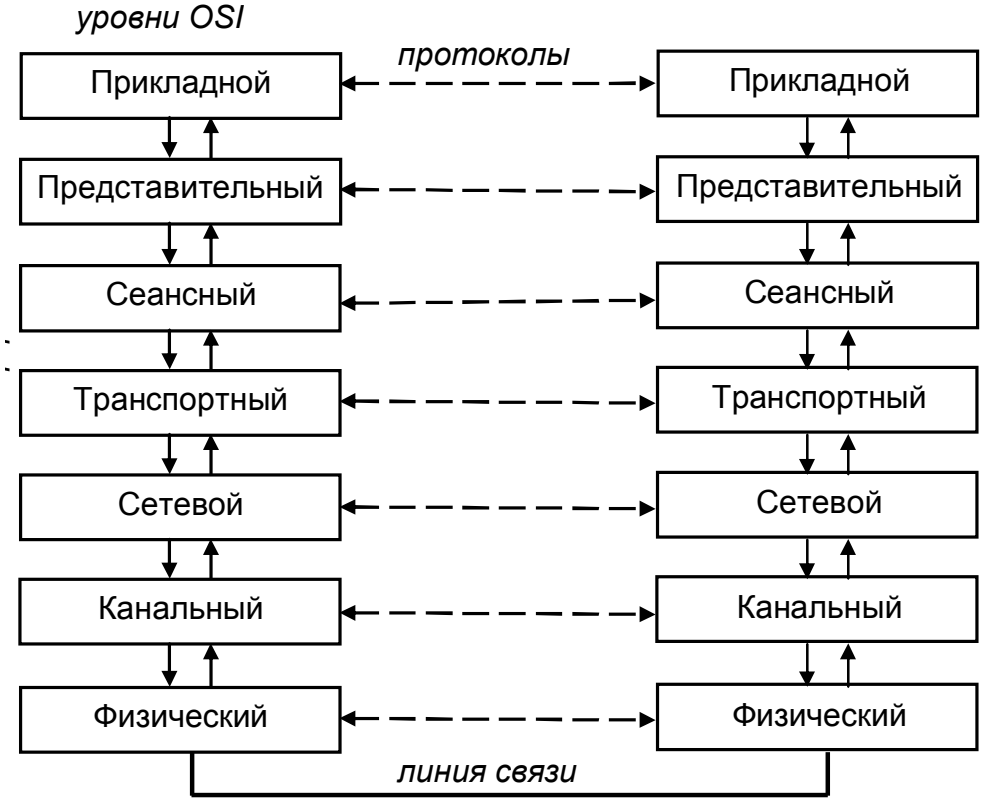


Рис. 1.3 — Эталонная модель взаимодействия открытых систем

Поскольку все нижестоящие уровни *прозрачны* для вышестоящих (включая и линию связи), то говорят, что приложения взаимодействуют по уровням. Понимать это нужно так, что при обращении ПО одного компьютера к другому, прикладной уровень как бы взаимодействует с прикладным уровнем, представительный — с представительным, физический — с физическим, и т.д.

Ниже дается описание и выполняемые уровнями функции:

Физический (Physical). Этот уровень модели OSI определяет физические, механические и электрические характеристики линий связи, составляющих ЛВС (кабелей, разъемов, оптоволоконных линий и т.п.). Этот уровень OSI обеспечивает средства для установления, поддержания и разъединения логических соединений между логическими объектами сети, реализует функции передачи битов данных через физические среды. Именно на физическом уровне осуществляется предоставление информации в виде электрических или оптических сигналов, преобразования формы сигналов, выбор параметров физических сред передачи данных. Можно считать этот уровень, отвечающим за аппаратное обеспечение. Хотя функции других уровней могут быть реализованы в соответствующих микросхемах, все же они относятся к программному обеспечению.

Физический уровень получает пакеты данных (*кадры*) от вышележащего канального уровня и преобразует их в оптические или электрические сигналы, соответствующие 0 и 1 бинарного потока. Эти сигналы посылаются через среду передачи на приемный узел. Механические и электрические/оптические свойства среды передачи определяются на физическом уровне и включают:

- тип кабелей и разъемов;
- разводку контактов в разъемах;
- схему кодирования сигналов для значений 0 и 1.

К числу наиболее распространенных спецификаций физического уровня относятся:

- EIA RS232C, CCITT V.24/ V.28 — COM-порт, механические, электрические характеристики несбалансированного последовательного интерфейса;
- EIA RS422/ 449, CCITT V.10 — механические, электрические и оптические характеристики сбалансированного последовательного интерфейса;

- IEEE 802.3 — Ethernet;
- IEEE 802.5 — Token Ring и др.

Канальный (Link). На этом уровне определяются правила использования физического уровня узлами сети. Электрическое представление данных в ЛВС (биты данных, методы кодирования данных и маркеры) распознаются на этом и только этом уровне. Здесь обнаруживаются и исправляются (путем требований повторной передачи данных) ошибки. Ввиду своей сложности канальный уровень подразделяется на два подуровня ***MAC*** (***Media Access Control*** — Контроль доступа к среде) и ***LLC*** (***Logical Link Control*** — Логический контроль связи).

Подуровень MAC связан с доступом к сети (передача маркера или обнаружение коллизий) и ее управлением. Подуровень LLC находится выше уровня MAC и связан с передачей и приемом пользовательских сообщений. Подуровень LLC оперирует функциями не связанными с особенностями передающей Среды. Канальный уровень предоставляет услуги по формированию ***кадра*** — пакета канального уровня — элементарного пакета среды передачи данных.

Наиболее часто используемые на 2 уровне протоколы:

- HDLC для последовательных соединений;
- IEEE 802.2 LLC (тип I и тип II) обеспечивают MAC для сред 802.x;
- Ethernet;
- Token Ring;
- FDDI;
- X.25;
- Frame relay.

Сетевой (Network). Этот уровень выполняет функции переключения и маршрутизации пакетов. Он отвечает за адресацию и доставку. На этом уровне происходит управление передачей пакетов через промежуточные узлы и сети, контроль нагрузки на сеть с целью снижения трафика. Маршрутизация (определение и реализация маршрутов, по которым передаются пакеты) сводится к определению логических каналов.

Логическим каналом называется виртуальное соединение двух или более объектов сетевого уровня, при котором возможен обмен данными между этими объектами. При этом не требуется

обязательного соответствия некоего физического соединения линии передачи данных между связываемыми пунктами. Логический канал абстрагирует сетевую связь от физической реализации соединения.

На сетевом уровне наиболее часто используются протоколы:

- IP — протокол сети Internet;
- IPX — протокол межсетевого обмена;
- X.25 (частично этот протокол реализован на уровне 2);
- CLNP — сетевой протокол без организации соединений.

Транспортный (Transport). Этот уровень предназначен для управления сквозными каналами в сети передачи данных: на транспортном уровне обеспечивается связь между оконечными пунктами (в отличие от сетевого уровня, на котором обеспечивается передача данных через промежуточные компоненты сети). К функциям транспортного уровня относятся мультиплексирование и демультиплексирование (сборка-разборка пакетов), обнаружение и устранение ошибок в передаче данных, реализация заказанного уровня услуг (например, заказанной скорости и надежности передачи).

На транспортном уровне пакеты обычно называют ***сегментами.*** Когда в процессе обработки находится более одного пакета, транспортный уровень контролирует очередность прохождения сегментов. Если приходит дубликат принятого ранее сегмента, то данный уровень опознает это и игнорирует этот сегмент.

Наиболее распространенные протоколы транспортного уровня:

- TCP — протокол управления передачей;
- NCP — Netware Core Protocol;
- SPX — упорядоченный обмен пакетами;
- TP4 — протокол передачи класса 4.

Сеансный (Session). Функции этого уровня состоят в координации связи между двумя прикладными программами, работающими на разных рабочих станциях. Это происходит в виде хорошо структурированного диалога. В число этих функций входит создание сеанса, управление передачей и приемом пакетов сообщений в течение сеанса и завершение сеанса. Этот уровень предназначен для организации и синхронизации диалога, ведущегося объектами (станциями) сети. На сеансном уровне опреде-

ляется тип связи (дуплекс или полудуплекс), начало и окончание заданий, последовательность и режим обмена запросами и ответами взаимодействующих партнеров.

Уровень представления (Presentation). Служит для преобразования данных из внутреннего формата компьютера в другой формат. Такая ситуация может возникнуть в ЛВС с неоднотипными ПК (IBM PC, Macintosh, DEC, Next, Burrough), которым необходимо обмениваться данными. Здесь реализуются функции представления данных (кодирование, форматирование, структурирование). Например, на этом уровне выделенные для передачи данные преобразуются из кода EBCDIC в ASCII и т.п.

Прикладной (Application). Этот уровень является пограничным между прикладной программой и процессами модели OSI. Включает средства управления прикладными процессами; эти процессы могут объединяться для выполнения поставленных заданий, обмениваться между собой данными. На прикладном уровне определяются и оформляются в блоки те данные, которые подлежат передаче по сети. Уровень включает, например, такие средства для взаимодействия прикладных программ, как прием и хранение пакетов в *почтовых ящиках* (*mail-box*).

К числу наиболее распространенных протоколов верхних уровней относятся:

- FTP — протокол переноса файлов;
- TFTP — упрощенный протокол переноса файлов;
- X.400 — электронная почта;
- Telnet;
- SMTP — простой протокол почтового обмена;
- CMIP — общий протокол управления информацией;
- SNMP — простой протокол управления сетью;
- NFS — сетевая файловая система;
- FTAM — метод доступа для переноса файлов.

Разделение на уровни позволяет вносить изменения в средства реализации одного уровня без перестройки средств других уровней, что значительно упрощает и удешевляет модернизацию средств по мере развития техники.

Один из факторов, который делает сетевую ОС каждого производителя *фирменной* (в отличие от открытой архитектуры) — это несовместимость с моделью OSI. В конкретных случаях мо-

жет возникнуть потребность в реализации лишь части названных функций, тогда соответственно в сети имеется лишь часть уровней. Так в простых (неразветвленных ЛВС) отпадает необходимость в средствах сетевого и транспортного уровней.

В табл. 1.1 показано соответствие некоторых, наиболее популярных протоколов уровням модели OSI. Часто это соответствие весьма условно, так как модель OSI — это только руководство к действию, причем достаточно общее, а конкретные протоколы разрабатывались для решения специфических задач, причем многие из них появились до разработки модели OSI.

Таблица 1.1 — **Эталонная модель OSI и некоторые популярные протоколы**

<table>
<tr><th>Модель OSI</th><th colspan="5">Протоколы</th></tr>
<tr><td>Прикладной</td><td rowspan="7">SNA</td><td rowspan="7">DECnet</td><td rowspan="3">NFS</td><td rowspan="4">NetWare</td><td rowspan="4">Windows</td></tr>
<tr><td>Представительский</td></tr>
<tr><td>Сеансовый</td></tr>
<tr><td>Транспортный</td><td>TCP/IP</td></tr>
<tr><td>Сетевой</td><td rowspan="2" colspan="3">Ethernet, Token Ring,
ARCNet, FDDI</td></tr>
<tr><td>Канальный</td></tr>
<tr><td>Физический</td></tr>
</table>

Большинство сетевых стандартов не соответствуют полностью модели OSI. Некоторые из них могут быть с ней соотнесены, другие — реализуют все уровни. Стандарты IBM SNA или DECnet поддерживаются в основном фирмами-разработчиками и формально не являются стандартами в силу своей закрытости, поэтому можно считать, что они соотносятся ко всем 7 уровням. Более открытые стандарты, такие как Unix, Nowell, Windows, представляют интерес для сторонних разработчиков, причем как сверху — на прикладном уровне, так и снизу — на сетевом.

1.5 Источники стандартов

Для правильного взаимодействия узлов различных вычислительных сетей их архитектура должна быть открытой. Этим целям служат унификация и стандартизация в области телекоммуникаций и вычислительных сетей. Работы по стандартизации

вычислительных сетей ведутся большим количеством организаций. В зависимости от статуса организаций различают следующие виды стандартов:

- ***стандарты отдельных фирм*** (например, стек протоколов DECnet фирмы Digital Equipment или графический интерфейс OPEN LOOK для Unix-систем фирмы Sun);
- ***стандарты специальных комитетов и объединений,*** создаваемых несколькими фирмами, например стандарты технологии ATM, разрабатываемые специально созданным объединением *ATM Forum*, насчитывающем около 100 коллективных участников, или стандарты союза *Fast Ethernet Alliance* по разработке стандартов 100 Мбит Ethernet;
- ***национальные стандарты,*** например, стандарт FDDI, представляющий один из многочисленных стандартов, разработанных Американским национальным институтом стандартов (ANSI), или стандарты безопасности для операционных систем, разработанные Национальным центром компьютерной безопасности (NCSC) Министерства обороны США;
- ***международные стандарты,*** например, модель и стек коммуникационных протоколов Международной организации по стандартам (ISO), многочисленные стандарты Международного союза электросвязи (ITU), в том числе стандарты на сети с коммутацией пакетов X.25, сети frame relay, ISDN, модемы и многие другие.

Некоторые стандарты, непрерывно развиваясь, могут переходить из одной категории в другую. В частности, фирменные стандарты на продукцию, получившую широкое распространение, обычно становятся международными стандартами де-факто, так как вынуждают производителей из разных стран следовать фирменным стандартам, чтобы обеспечить совместимость своих изделий с этими популярными продуктами. Например, из-за феноменального успеха персонального компьютера компании IBM фирменный стандарт на архитектуру IBM PC стал международным стандартом де-факто.

Более того, ввиду широкого распространения некоторые фирменные стандарты становятся основой для национальных и международных стандартов де-юре. Например, стандарт Ethernet, первоначально разработанный компаниями Digital Equipment,

Intel и Xerox, через некоторое время и в несколько измененном виде был принят как национальный стандарт IEEE 802.3, а затем организация ISO утвердила его в качестве международного стандарта ISO 8802.3.

Далее приводятся краткие сведения об организациях, наиболее активно и успешно занимающихся разработкой стандартов в области вычислительных сетей.

• ***Международная организация по стандартизации*** (***International Organization for Standardization, ISO,*** часто называемая также ***International Standards Organization***) представляет собой ассоциацию ведущих национальных организаций по стандартизации разных стран. Главным достижением ISO явилась модель взаимодействия открытых систем OSI, которая в настоящее время является концептуальной основой стандартизации в области вычислительных сетей. В соответствии с моделью OSI этой организацией был разработан стандартный стек коммуникационных протоколов OSI.

• ***Международный союз электросвязи*** (***International Telecommunications Union, ITU***) — организация, являющаяся в настоящее время специализированным органом Организации Объединенных Наций. Наиболее значительную роль в стандартизации вычислительных сетей играет постоянно действующий в рамках этой организации Международный консультативный комитет по телефонии и телеграфии (***МККТТ***) (***Consultative Committee on International Telegraphy and Telephony, CCITT***). В результате проведенной в 1993 году реорганизации ITU CCITT несколько изменил направление своей деятельности и сменил название — теперь он называется сектором телекоммуникационной стандартизации ITU (***ITU Telecommunication Standardization Sector, ITU-T***). Основу деятельности ITU-T составляет разработка международных стандартов в области телефонии, телематических служб (электронной почты, факсимильной связи, телетекста, телекса и т.д.), передачи данных, аудио- и видеосигналов. За годы своей деятельности ITU-T выпустил огромное число рекомендаций-стандартов.

• ***Институт инженеров по электротехнике и радиоэлектронике*** (***Institute of Electrical and Electronics Engineers, IEEE***) — национальная организация США, определяющая сете-

вые стандарты. В 1981 году рабочая группа 802 этого института сформулировала основные требования, которым должны удовлетворять локальные вычислительные сети. Группа 802 определила множество стандартов, из них самыми известными являются стандарты 802.1, 802.2, 802.3 и 802.5, которые описывают общие понятия, используемые в области локальных сетей, а также стандарты на два нижних уровня сетей Ethernet и Token Ring.

• ***Европейская ассоциация производителей компьютеров*** (***European Computer Manufacturers Association, ECMA***) — некоммерческая организация, активно сотрудничающая с ITU-T и ISO, занимается разработкой стандартов и технических обзоров, относящихся к компьютерной и коммуникационной технологиям. Известна своим стандартом ECMA-101, используемым при передаче отформатированного текста и графических изображений с сохранением оригинального формата.

• ***Ассоциация производителей компьютеров и оргтехники (Computer and Business Equipment Manufacturers Association, CBEMA)*** — организация американских фирм-производителей аппаратного обеспечения; аналогична европейской ассоциации ECMA; участвует в разработке стандартов на обработку информации и соответствующее оборудование.

• ***Ассоциация электронной промышленности*** (***Electronic Industries Association, EIA***) — промышленно-торговая группа производителей электронного и сетевого оборудования; является национальной коммерческой ассоциацией США; проявляет значительную активность в разработке стандартов для проводов, коннекторов и других сетевых компонентов. Ее наиболее известный стандарт — RS-232C.

• ***Министерство обороны США*** (***Department of Defense, DoD***) имеет многочисленные подразделения, занимающиеся созданием стандартов для компьютерных систем. Одной из самых известных разработок DoD является стек транспортных протоколов TCP/IP.

• ***Американский национальный институт стандартов*** (***American National Standards Institute, ANSI***) — эта организация представляет США в Международной организации по стандартизации ISO. Комитеты ANSI ведут работу по разработке стандартов в различных областях вычислительной техники. Так, комитет

ANSI X3T9.5 совместно с фирмой IBM занимается стандартизацией локальных сетей крупных ЭВМ (архитектура сетей SNA). Известный стандарт FDDI также является результатом деятельности этого комитета ANSI. В области микрокомпьютеров ANSI разрабатывает стандарты на языки программирования, интерфейс SCSI. Институт ANSI разработал рекомендации по переносимости для языков C, FORTRAN, COBOL.

• ***Internet***. Особую роль в выработке международных открытых стандартов играют стандарты Internet. Ввиду большой и постоянной растущей популярности Internet, эти стандарты становятся международными стандартами «де-факто», многие из которых затем приобретают статус официальных международных стандартов за счет их утверждения одной из вышеперечисленных организаций, в том числе ISO и ITU-T. Существует несколько организационных подразделений, отвечающих за развитие Internet и, в частности, за стандартизацию средств Internet.

Основным из них является ***Internet Society*** (***ISOC***) — профессиональное сообщество, которое занимается общими вопросами эволюции и роста Internet как глобальной коммуникационной инфраструктуры. Под управлением ISOC работает ***Internet Architecture Board*** (***IAB***) — организация, в ведении которой находится технический контроль и координация работ для Internet. IAB координирует направление исследований и новых разработок для стека TCP/IP и является конечной инстанцией при определении новых стандартов Internet. В IAB входят две основные группы: ***Internet Engineering Task Force*** (***IETF***) и ***Internet Research Task Force*** (***IRTF***). IETF — это инженерная группа, которая занимается решением ближайших технических проблем Internet. Именно IETF определяет спецификации, которые затем становятся стандартами Internet. В свою очередь, IRTF координирует долгосрочные исследовательские проекты по протоколам TCP/IP.

RFC

Все официальные стандарты сообщества internet публикуются в ***Request for Comment***, или в ***RFC***. В дополнение, существует множество RFC, которые не являются официальными стандартами, однако они публикуются с информационными целями.

Диапазон размеров RFC колеблется от 1 до почти 200 страниц. Каждый из них имеет собственный номер. Все RFC доступны бесплатно по электронной почте rfc-info@ISI.EDU или с использованием FTP через Internet.

Начиная изучать какую-либо проблему, наиболее полезным является просмотреть последние RFC. Для отслеживания новых публикаций существует индекс RFC, который содержит подробную информацию о том когда были заменены RFC на более новые и какая информация появилась во вновь вышедших RFC. Существует несколько важных RFC.

1. *Assigned Numbers RFC* содержит в себе все числа и константы, которые используются в протоколах Internet. В настоящее время самая последняя версия — RFC 1340. Также здесь описаны все заранее известные порты глобальной сети Internet.

2. Официальные стандарты протоколов Internet (*Internet Official Protocol Standards*), в настоящее время RFC 1600 [Postel 1994]. Этот RFC содержит информацию о состоянии стандартизации различных протоколов Internet. Каждый протокол имеет одно из следующих состояний стандартизации: стандарт, необязательный стандарт, рекомендованный стандарт, экспериментальный, информационный, или исторический. В дополнение, каждый протокол имеет уровень необходимости: необходим, рекомендуется, на выбор, с ограниченным использованием, или не рекомендуется.

3. Требования к хостам *Host Requirements RFC*. RFC 1122 описывает канальный уровень, сетевой уровень и транспортный уровень, RFC 1123 описывает прикладной уровень. Эти два RFC корректируют и интерпретируют различные важные RFC, вышедшие раньше, и часто являются исходной точкой при изучении деталей того или иного протокола.

4. Требования к маршрутизаторам *Router Requirements RFC*. Официальная версия — RFC 1009, однако новая версия близка к завершению. Это RFC напоминает RFC требования к компьютерам, однако содержит уникальные требования к маршрутизаторам.

1.6 Стандартные стеки коммуникационных протоколов

Важнейшим направлением стандартизации в области вычислительных сетей является стандартизация коммуникационных протоколов. В настоящее время в сетях используется большое количество стеков коммуникационных протоколов. Наиболее популярными являются стеки: TCP/IP, IPX/SPX, NetBIOS/SMB, DECnet, SNA и OSI. Все эти стеки, кроме SNA на нижних уровнях — физическом и канальном, — используют одни и те же хорошо стандартизованные протоколы Ethernet, Token Ring, FDDI и некоторые другие, которые позволяют использовать во всех сетях одну и ту же аппаратуру. Зато на верхних уровнях все стеки работают по своим собственным протоколам.

Стек OSI. Следует четко различать модель OSI и ***стек OSI***. В то время как модель OSI является концептуальной схемой взаимодействия открытых систем, стек OSI представляет собой набор вполне конкретных спецификаций протоколов. В отличие от других стеков протоколов стек OSI полностью соответствует модели OSI, он включает спецификации протоколов для всех семи уровней взаимодействия, определенных в этой модели. На нижних уровнях стек OSI поддерживает Ethernet, Token Ring, FDDI, протоколы глобальных сетей, X.25 и ISDN, — то есть использует разработанные вне стека протоколы нижних уровней, как и другие. Протоколы сетевого, транспортного и сеансового уровней стека OSI специфицированы и реализованы различными производителями, но распространены пока мало. Наиболее популярными протоколами стека OSI являются прикладные протоколы. К ним относятся: протокол *передачи файлов* ***FTAM***, протокол *эмуляции терминала* ***VTP***, протоколы *справочной службы* ***X.500***, *электронной почты* ***X.400*** и ряд других.

Протоколы стека OSI отличает большая сложность и неоднозначность спецификаций. Эти свойства явились результатом общей политики разработчиков стека, стремившихся учесть в своих протоколах все случаи жизни и все существующие и появляющиеся технологии. К этому нужно еще добавить и последствия большого количества политических компромиссов, неизбежных при принятии международных стандартов по такому злободневному вопросу, как построение открытых вычислительных

сетей. Из-за своей сложности протоколы OSI требуют больших затрат вычислительной мощности центрального процессора, что делает их наиболее подходящими для мощных машин, а не для сетей персональных компьютеров.

Одним из крупнейших производителей, поддерживающих OSI, является компания AT&T, ее сеть Stargroup полностью базируется на этом стеке.

Стек TCP/IP был разработан по инициативе Министерства обороны США для связи экспериментальной сети ARPAnet с другими сетями как набор общих протоколов для разнородной вычислительной среды. Большой вклад в развитие стека TCP/IP, который получил свое название по популярным протоколам IP и TCP, внес университет Беркли, реализовав протоколы стека в своей версии ОС UNIX. Популярность этой операционной системы привела к широкому распространению протоколов TCP, IP и других протоколов стека. Сегодня этот стек используется для связи компьютеров всемирной информационной сети Internet, а также в огромном числе корпоративных сетей.

Стек TCP/IP на нижнем уровне поддерживает все популярные стандарты физического и канального уровней. Для локальных сетей — это Ethernet, Token Ring, FDDI, для глобальных — протоколы работы на аналоговых коммутируемых и выделенных линиях SLIP, PPP, протоколы территориальных сетей X.25 и ISDN.

Основными протоколами стека, давшими ему название, являются протоколы ***IP*** и ***TCP***. Эти протоколы в терминологии модели OSI относятся к сетевому и транспортному уровням соответственно. Протокол IP обеспечивает продвижение пакета по составной сети, а TCP гарантирует надежность его доставки. За долгие годы использования в сетях различных стран и организаций стек TCP/IP вобрал в себя большое количество протоколов прикладного уровня. К ним относятся такие популярные протоколы, как протокол ***пересылки файлов FTP***, протокол ***эмуляции терминала TELNET***, ***почтовый протокол SMTP***, используемый в электронной почте сети Internet, ***гипертекстовые сервисы службы WWW*** и многие другие.

Хотя протоколы TCP/IP неразрывно связаны с Internet и каждый из многомиллионной армады компьютеров Internet работает на основе этого стека, существует большое количество локальных, корпоративных и территориальных сетей, непосредственно

не являющихся частями Internet, в которых также используют протоколы TCP/IP. Чтобы отличать их от Internet, эти сети называют сетями TCP/IP или просто IP-сетями.

Поскольку стек TCP/IP изначально создавался для глобальной сети Internet, он имеет много особенностей, дающих ему преимущество перед другими протоколами, когда речь заходит о построении сетей, включающих глобальные связи. В частности, очень полезным свойством, делающим возможным применение этого протокола в больших сетях, является его способность фрагментировать пакеты. Действительно, большая составная сеть часто состоит из сетей, построенных на совершенно разных принципах. В каждой из этих сетей может быть установлена собственная величина максимальной длины единицы передаваемых данных (кадра). В таком случае при переходе из одной сети, имеющей большую максимальную длину, в сеть с меньшей максимальной длиной может возникнуть необходимость деления передаваемого кадра на несколько частей. Протокол IP стека TCP/IP эффективно решает эту задачу. Другой особенностью технологии TCP/IP является гибкая система адресации, позволяющая более просто по сравнению с другими протоколами аналогичного назначения включать в интерсеть сети других технологий. Это свойство также способствует применению стека TCP/IP для построения больших гетерогенных сетей.

В стеке TCP/IP очень экономно используются возможности широковещательных рассылок. Это свойство совершенно необходимо при работе на медленных каналах связи, характерных для территориальных сетей.

Однако, как и всегда, за получаемые преимущества надо платить, и платой здесь оказываются высокие требования к ресурсам и сложность администрирования IP-сетей. Мощные функциональные возможности протоколов стека TCP/IP требуют для своей реализации высоких вычислительных затрат. Гибкая система адресации и отказ от широковещательных рассылок приводят к наличию в IP-сети различных централизованных служб типа ***DNS***, ***DHCP*** и т.п. Каждая из этих служб направлена на облегчение администрирования сети, в том числе и на облегчение конфигурирования оборудования, но в то же время сама требует пристального внимания со стороны администраторов.

Стек IPX/SPX. Этот стек является оригинальным стеком протоколов фирмы *Novell*, разработанным для сетевой операционной системы NetWare еще в начале 80-х годов. Протоколы сетевого и сеансового уровней ***Internetwork Packet Exchange*** (***IPX***) и ***Sequenced Packet Exchange*** (***SPX***), которые дали название стеку, являются прямой адаптацией протоколов ***XNS*** фирмы Xerox, распространенных в гораздо меньшей степени, чем стек IPX/SPX. Популярность стека IPX/SPX непосредственно связана с операционной системой ***Novell NetWare***, которая еще сохраняет мировое лидерство по числу установленных систем, хотя в последнее время ее популярность несколько снизилась и по темпам роста она отстает от Microsoft Windows NT.

Многие особенности стека IPX/SPX обусловлены ориентацией ранних версий ОС NetWare (до версии 4.0) на работу в локальных сетях небольших размеров, состоящих из персональных компьютеров со скромными ресурсами. Понятно, что для таких компьютеров компании Novell нужны были протоколы, на реализацию которых требовалось бы минимальное количество оперативной памяти (ограниченной в IBM-совместимых компьютерах под управлением MS-DOS объемом 640 Кбайт) и которые бы быстро работали на процессорах небольшой вычислительной мощности. В результате протоколы стека IPX/SPX до недавнего времени хорошо работали в локальных сетях и не очень — в больших корпоративных сетях, так как они слишком перегружали медленные глобальные связи широковещательными пакетами, которые интенсивно используются несколькими протоколами этого стека (например, для установления связи между клиентами и серверами). Это обстоятельство, а также тот факт, что стек IPX/SPX является собственностью фирмы Novell и на его реализацию нужно получать лицензию (то есть открытые спецификации не поддерживались), долгое время ограничивали распространенность его только сетями NetWare. Однако с момента выпуска версии NetWare 4.0 Novell внесла и продолжает вносить в свои протоколы серьезные изменения, направленные на их адаптацию для работы в корпоративных сетях. Сейчас стек IPX/SPX реализован не только в NetWare, но и в нескольких других популярных сетевых ОС, например SCO UNIX, Sun Solaris, Microsoft Windows NT.

Стек NetBIOS/SMB. Этот стек широко используется в продуктах компаний IBM и Microsoft. На физическом и канальном

уровнях этого стека используются все наиболее распространенные протоколы Ethernet, Token Ring, FDDI и другие. На верхних уровнях работают протоколы ***NetBEUI*** и ***SMB***.

Протокол NetBIOS (Network Basic Input/Output System) появился в 1984 году как сетевое расширение стандартных функций BIOS компьютеров IBM PC для сетевой программы PC Network фирмы IBM. В дальнейшем этот протокол был заменен так называемым протоколом расширенного пользовательского интерфейса ***NetBEUI — NetBIOS Extended User Interface***. Для обеспечения совместимости приложений в качестве интерфейса к протоколу NetBEUI был сохранен интерфейс NetBIOS. Протокол NetBEUI разрабатывался как эффективный протокол, потребляющий немного ресурсов и предназначенный для сетей, насчитывающих не более 200 рабочих станций. Этот протокол содержит много полезных сетевых функций, которые можно отнести к сетевому, транспортному и сеансовому уровням модели OSI, однако с его помощью невозможна маршрутизация пакетов. Это ограничивает применение протокола NetBEUI локальными сетями, не разделенными на подсети, и делает невозможным его использование в составных сетях. Некоторые ограничения NetBEUI снимаются реализацией этого протокола ***NBF*** (***NetBEUI Frame***), которая включена в операционную систему Microsoft Windows NT. Протокол ***SMB*** (***Server Message Block***) выполняет функции сеансового, представительного и прикладного уровней. На основе SMB реализуется файловая служба, а также службы печати и передачи сообщений между приложениями.

Стеки протоколов ***SNA*** фирмы IBM, ***DECnet*** корпорации Digital Equipment и ***AppleTalk/AFP*** фирмы Apple применяются в основном в операционных системах и сетевом оборудовании этих фирм.

1.7 Инкапсуляция сообщений и наслоение протоколов

В вычислительных сетях предусматривается подход, в котором протоколы как бы наслаиваются один на другой. При этом протокол низкого уровня, управляя адаптером, выполняет функции передачи сообщения и не подозревает о существовании файл-серверов и функций обслуживания или перенаправления

файлов. Протокол высокого уровня обслуживает файл-сервер и перенаправляет файлы, но ничего не знает о функционировании физического уровня. Функционируя совместно, эти протоколы собственно и образуют ЛВС. Принцип наслоения протоколов иллюстрируется на рис. 1.4.

Передача данных через разветвленные сети происходит при использовании ***инкапсуляции/деинкапсуляции*** порций данных. Так сообщение, пришедшее на транспортный уровень получает заголовки и передается на сетевой уровень. На сетевом уровне сегмент может быть разделен на части (пакеты сетевого уровня), если сеть не поддерживает передачу сегментов целиком. Пакет снабжается своим сетевым заголовком (т.е. производится инкапсуляция сегментов в пакеты). При передаче между узлами промежуточной ЛВС может потребоваться разделение пакетов на кадры (т.е. инкапсуляция пакетов в кадры). Наконец, сообщение достигает нижнего, физического уровня, который собственно и передает его по линиям связи машине-адресату. К этому моменту сообщение обрастаете заголовками всех уровней (рис. 1.4). На соответствующем уровне приемного узла сегменты деинкапсулируются и восстанавливается исходное сообщение.

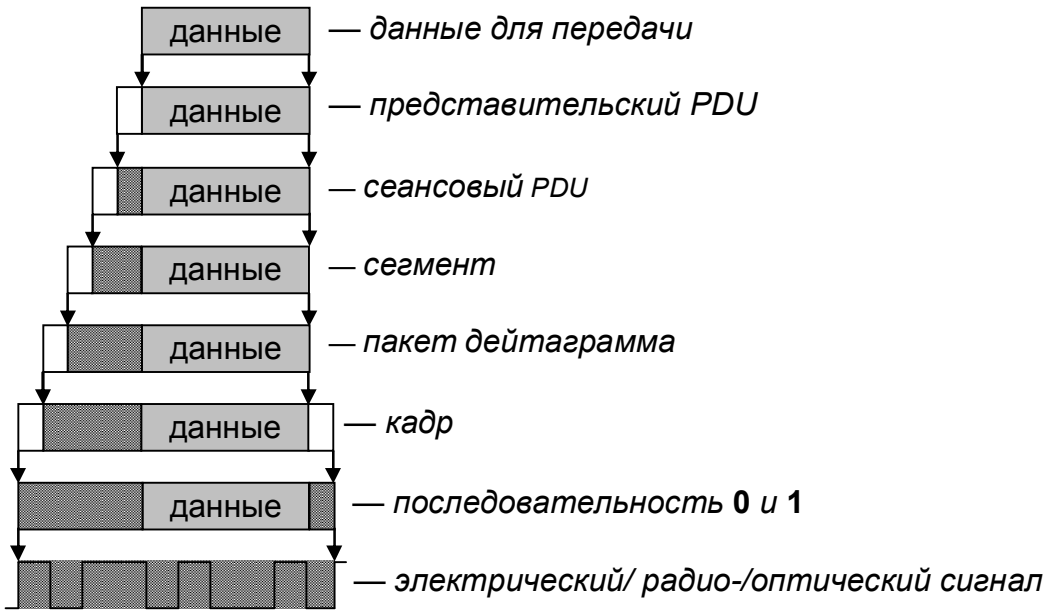


Рис. 1.4 — Несколько уровней пакетов сообщений

Когда сообщение по сети поступает на машину-адресат, оно принимается ее физическим уровнем и последовательно перемещается вверх с уровня на уровень. Каждый уровень анализирует и обрабатывает заголовок своего уровня, выполняя соответствующие данному уровню функции, а затем удаляет этот заголовок и передает сообщение вышележащему уровню.

Таким образом, пакеты более высокого уровня содержатся в пакетах более низкого уровня и на каждом уровне необходим специальный протокол. Наряду с термином ***сообщение*** (***message***) существуют и другие термины, применяемые сетевыми специалистами для обозначения единиц данных в процедурах обмена. В стандартах ISO для обозначения единиц данных, с которыми имеют дело протоколы разных уровней, используется общее название ***протокольный блок данных*** (***Protocol Data Unit, PDU***). Для обозначения блоков данных определенных уровней часто используются специальные названия: ***кадр*** (***frame***), ***пакет*** (***packet***), ***дейтаграмма*** (***datagram***), ***сегмент*** (***segment***).

1.8 Адресация компьютеров

Еще одной новой проблемой, которую нужно учитывать при объединении трех и более компьютеров, является проблема их адресации. К адресу узла сети и схеме его назначения можно предъявить несколько требований.

- Адрес должен уникально идентифицировать компьютер в сети любого масштаба.
- Схема назначения адресов должна сводить к минимуму ручной труд администратора и вероятность дублирования адресов.
- Адрес должен иметь иерархическую структуру, удобную для построения больших сетей. Эту проблему хорошо иллюстрируют международные почтовые адреса, которые позволяют почтовой службе, организующей доставку писем между странами, пользоваться только названием страны адресата и не учитывать название его города, а тем более улицы. В больших сетях, состоящих из многих тысяч узлов, отсутствие иерархии адреса может привести к большим издержкам — конечным узлам и коммуникационному оборудованию придется оперировать с таблицами адресов, состоящими из тысяч записей.

- Адрес должен быть удобен для пользователей сети, а это значит, что он должен иметь символьное представление например, *Servers* или *www.cisco.com.*
- Адрес должен иметь по возможности компактное представление, чтобы не перегружать память коммуникационной аппаратуры — сетевых адаптеров, маршрутизаторов и т.п.

Нетрудно заметить, что эти требования противоречивы — например, адрес, имеющий иерархическую структуру, скорее всего, будет менее компактным, чем неиерархический (такой адрес часто называют «плоским», то есть не имеющим структуры). Символьный же адрес, очевидно, потребует больше памяти, чем адрес-число.

Так как все перечисленные требования трудно совместить в рамках какой-либо одной схемы адресации, то на практике обычно используется сразу несколько схем, так что компьютер одновременно имеет несколько адресов-имен. Каждый адрес используется в той ситуации, когда соответствующий вид адресации наиболее удобен. А чтобы не возникало путаницы, и компьютер всегда однозначно определялся своим адресом, используются специальные вспомогательные протоколы, которые по адресу одного типа могут определить адреса других типов.

Наибольшее распространение получили три схемы адресации узлов сети.

- ***Аппаратные (hardware) адреса***. Эти адреса предназначены для сети небольшого или среднего размера, поэтому они не имеют иерархической структуры. Типичным представителем адреса такого типа является адрес сетевого адаптера локальной сети. Такой адрес обычно используется только аппаратурой, поэтому его стараются сделать по возможности компактным и записывают в виде двоичного или шестнадцатеричного значения. При задании аппаратных адресов обычно не требуется выполнение ручной работы, так как они либо встраиваются в аппаратуру компанией-изготовителем, либо генерируются автоматически при каждом новом запуске оборудования, причем уникальность адреса в пределах сети обеспечивает оборудование. Помимо отсутствия иерархии, использование аппаратных адресов связано еще с одним недостатком — при замене аппаратуры, например, сетевого адаптера, изменяется и адрес компьютера. Более того, при ус-

тановке нескольких сетевых адаптеров у компьютера появляется несколько адресов, что не очень удобно для пользователей сети.

- ***Символьные адреса или имена***. Эти адреса предназначены для запоминания людьми и поэтому обычно несут смысловую нагрузку. Символьные адреса легко использовать как в небольших, так и крупных сетях. Для работы в больших сетях символьное имя может иметь сложную иерархическую структуру, например *mvp@nw.ie.tusur.edu.ru.* Эта запись обозначает адрес электронной почты (абонент *mvp*) в сети *NetWare* (*nw*) одной из кафедр (*ie* — industrial electronics) Томского университета систем управления и электроники (*tusur*) и эта сеть относится к образовательной ветви (*edu* — education) Интернета России (*ru*).
- ***Числовые составные адреса.*** Символьные имена удобны для людей, но из-за переменного формата и потенциально большой длины их передача по сети не очень экономична. Поэтому, во многих случаях, для работы в больших сетях в качестве адресов используют числовые составные адреса фиксированного и компактного форматов. Типичным представителями адресов этого типа являются IP- и IPX-адреса. В них поддерживается двухуровневая иерархия, адрес делится на старшую часть — номер сети и младшую — номер узла. Такое деление позволяет передавать сообщения между сетями только на основании номера сети, а номер узла используется только после доставки сообщения в нужную сеть; точно так же, как название улицы используется почтальоном только после того, как письмо доставлено в нужный город. В последнее время, чтобы сделать маршрутизацию в крупных сетях более эффективной, предлагаются более сложные варианты числовой адресации, в соответствии с которыми адрес имеет три и более составляющих. Такой подход, в частности, реализован в новой версии протокола ***IPv6***, предназначенного для работы в сети Internet.

В современных сетях для адресации узлов применяются, как правило, одновременно все три приведенные выше схемы. Пользователи адресуют компьютеры символьными именами, которые автоматически заменяются в сообщениях, передаваемых по сети, на числовые номера. С помощью этих числовых номеров сообщения передаются из одной сети в другую, а после доставки сообщения в сеть назначения вместо числового номера используется аппаратный адрес компьютера. Сегодня такая схема характерна даже для не-

больших автономных сетей, где, казалось бы, она явно избыточна — это делается для того, чтобы при включении этой сети в большую сеть не нужно было менять состав операционной системы.

Проблема установления соответствия между адресами различных типов, которой занимается служба разрешения имен, может решаться как полностью централизованными, так и распределенными средствами. В случае централизованного подхода в сети выделяется один компьютер (***сервер имен***), в котором хранится таблица соответствия друг другу имен различных типов, например символьных имен и числовых номеров. Все остальные компьютеры обращаются к серверу имен, чтобы по символьному имени найти числовой номер компьютера, с которым необходимо обменяться данными.

При другом, распределенном подходе, каждый компьютер сам решает задачу установления соответствия между именами. Например, если пользователь указал для узла назначения числовой номер, то перед началом передачи данных, компьютер-отправитель посылает всем компьютерам сети сообщение (такое сообщение называется широковещательным) с просьбой опознать это числовое имя. Все компьютеры, получив это сообщение, сравнивают заданный номер со своим собственным. Тот компьютер, у которого обнаружилось совпадение, посылает ответ, содержащий его аппаратный адрес, после чего становится возможным отправка сообщений по локальной сети.

Распределенный подход хорош тем, что не предполагает выделения специального компьютера, который к тому же часто требует ручного задания таблицы соответствия имен. Недостатком распределенного подхода является необходимость широковещательных сообщений — такие сообщения перегружают сеть, так как они требуют обязательной обработки всеми узлами, а не только узлом назначения. Поэтому распределенный подход используется только в небольших локальных сетях. В крупных сетях распространение широковещательных сообщений по всем ее сегментам становится практически нереальным, поэтому для них характерен централизованный подход.

Наиболее известной службой централизованного разрешения имен является служба ***Domain Name System*** (***DNS***) сети Internet.

2 ЛИНИИ СВЯЗИ

2.1 Классификация сетей по типам линий связи

Классификация информационных сетей в зависимости от типа линий связи приведена на рис. 2.1.

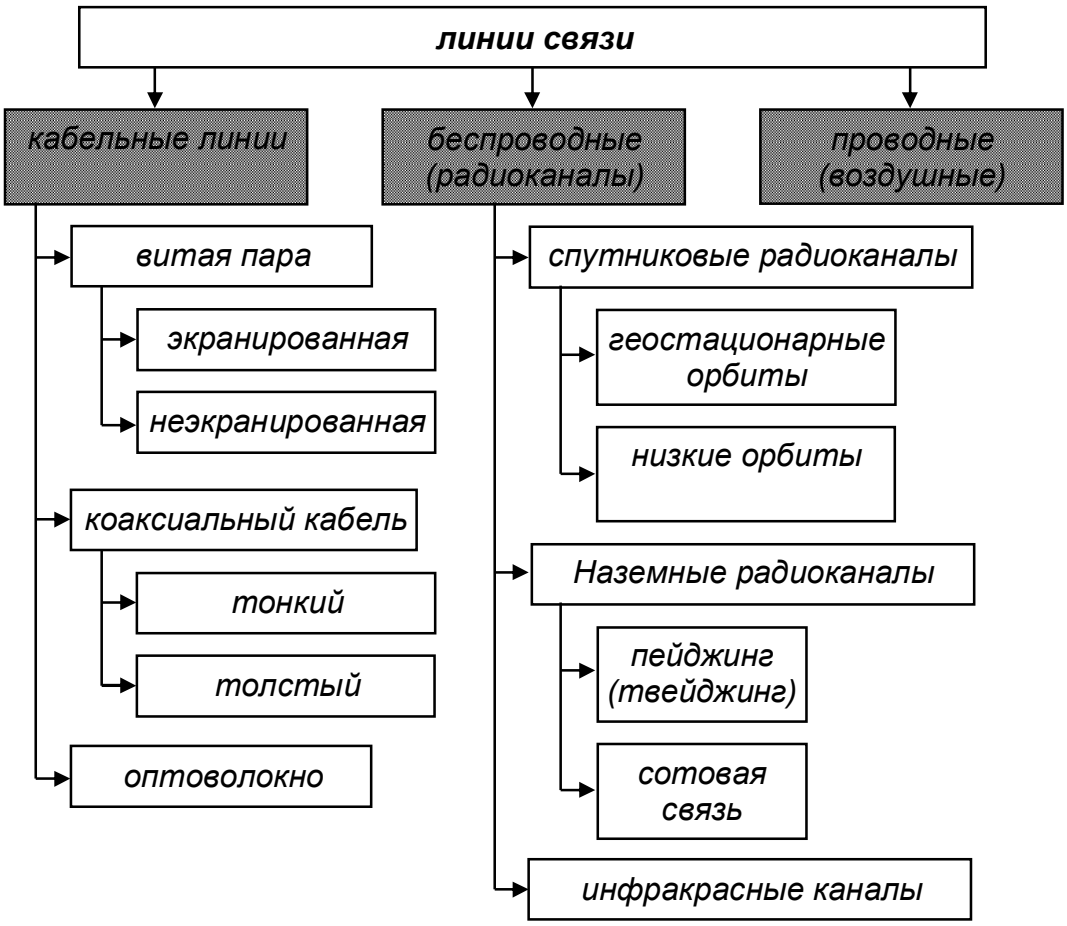


Рис. 2.1 — Типы линий связи

Физическая среда передачи данных (medium) может представлять собой кабель, то есть набор проводов, изоляционных и защитных оболочек и соединительных разъемов, а также земную атмосферу или космическое пространство, через которые распространяются электромагнитные волны. В зависимости от среды передачи данных линии связи разделяются на следующие (рис. 2.1):

- проводные (воздушные);
- кабельные (медные и волоконно-оптические);

- радиоканалы наземной и спутниковой связи.

Проводные (воздушные) линии связи представляют собой провода без изолирующих или экранирующих оплеток, проложенные между столбами и висящие в воздухе. По таким линиям связи традиционно передаются телефонные или телеграфные сигналы, но при отсутствии других возможностей эти линии используются и для передачи компьютерных данных. Скоростные качества и помехозащищенность этих линий очень плохие, и поэтому проводные линии связи быстро вытесняются кабельными системами.

Кабельные линии представляют собой достаточно сложную конструкцию. Кабель состоит из проводников, заключенных в несколько слоев изоляции: электрической, электромагнитной, механической, а также, возможно, климатической. Кроме того, кабель может быть оснащен разъемами, позволяющими быстро выполнять присоединение к нему различного оборудования. В компьютерных сетях применяются три основных типа кабеля: кабели на основе скрученных пар медных проводов, коаксиальные кабели с медной жилой, а также волоконно-оптические кабели.

Скрученная пара проводов называется ***витой парой (twisted pair)***. Витая пара существует в экранированном варианте ***(Shielded Twisted Pair, STP),*** когда пара медных проводов обертывается в изоляционный экран, и неэкранированном ***(Unshielded Twisted Pair, UTP),*** когда изоляционная обертка отсутствует. Скручивание проводов снижает влияние внешних помех на полезные сигналы, передаваемые по кабелю. Пару проводов часто используют как *сбалансированную линию*, в двух проводах которой передаются одни и те же уровни сигнала (по отношению к земле), но разной полярности. При приеме воспринимается разность сигналов, называемая *парафазным сигналом*, синфазные помехи при этом компенсируются. Помехоустойчивость существенно повышается, так как помехи в двух проводах обычно появляются в одной и той же фазе.

Коаксиальный кабель (coaxial) имеет несимметричную конструкцию и состоит из внутренней медной жилы и оплетки, отделенной от жилы слоем изоляции. Существует несколько ти-

пов коаксиального кабеля, отличающихся характеристиками и областями применения — для локальных сетей, для глобальных сетей, для кабельного телевидения и т.п. Наиболее популярными являются «толстый» диаметром 12,5 мм и «тонкий» диметром 6,25 мм. Толстый кабель имеет лучшую помехозащищенность и меньшее затухание, что позволяет использовать его на больших расстояниях, но он плохо гнется, что затрудняет его прокладку, и дороже тонкого.

Волоконно-оптический кабель (optical fiber) состоит из тонких (десятки микрон) волокон, по которым распространяются световые сигналы. Волоконно-оптические линии связи (ВОЛС) представляют собой кварцевый сердечник диаметром 5—60 мкм, покрытый отражающей оболочкой с внешним диаметром 125—200 мкм. Характеристики ВОЛС: работа на волнах 0,85—1,55 мкм, затухание 0,7 дБ/км, полоса частот до 2 ГГц. Предельные расстояния D для передачи данных по ВОЛС (без ретрансляции) зависят от длины волны излучения L: для L = 850 нм предельное расстояние D составит 5 км, а для L = 1300 нм D = 50 км, но аппаратурная реализация данного вида связи значительно дороже. Это наиболее качественный тип кабеля — он обеспечивает передачу данных с очень высокой скоростью (до 10 Гбит/с и выше) и, к тому же, лучше других типов передающей среды обеспечивает защиту данных от внешних помех.

Смешанные оптико-кабельные сети (чаще оптико-коаксиальные сети) (***hybrid fiber-coax, HFC***) — наиболее совершенные в смысле пропускной способности в настоящее время сети передачи данных. Оптоволоконный кабель прокладывается до группы строений, а далее идет разводка коаксиальным кабелем, спрягаются эти части при помощи оптического распределительного узла.

Беспроводные каналы связи. Передача информации по беспроводным каналам связи осуществляется на основе распространения электромагнитных колебаний. В табл. 2.1 приведены сведения о диапазонах частот электромагнитных колебаний, используемых в линиях связи.

Таблица 2.1

Диапазон	***Длина волны, м***	***Частота, ГГц***	***Применение***
дециметровый	1—0,1	0,3—3	ТВ, сотовые телефоны, спутниковая связь, радиоканал
сантиметровый	0,1—0,01	3—30	Радиорелейные линии, спутниковая связь, радиоканал
миллиметровый	0,01—0,001	30—300	радиоканал
инфракрасный	$0{,}001—7{,}5\cdot10^{-7}$	$300—4\cdot10^{5}$	ИК-каналы, ВОЛС
видимый свет	$(7{,}5—4)\cdot10^{-7}$	$(4—7{,}5)\cdot10^{5}$	ВОЛС

Радиоканалы наземной и спутниковой связи образуются с помощью передатчика и приемника радиоволн. Существует большое количество различных типов радиоканалов, отличающихся как используемым частотным диапазоном, так и дальностью канала. ***Диапазоны коротких, средних и длинных волн*** (***КВ, СВ и ДВ***), называемые также диапазонами амплитудной модуляции (***Amplitude Modulation, AM***) по типу используемого в них метода модуляции сигнала, обеспечивают дальнюю связь, но при невысокой скорости передачи данных. Более скоростными являются каналы, работающие на диапазонах ***ультракоротких волн*** (***УКВ***), для которых характерна частотная модуляция (***Frequency Modulation, FM***), а также диапазонах ***сверхвысоких частот*** (***СВЧ*** или ***microwaves***). В диапазоне СВЧ (свыше 4 ГГц) сигналы уже не отражаются ионосферой Земли, и для устойчивой связи, требуется наличие прямой видимости между передатчиком и приемником. Поэтому такие частоты используют либо спутниковые каналы, либо радиорелейные каналы, где это условие выполняется.

Проблема использования беспроводных каналов связи состоит в том, что чем выше рабочая частота, тем больше емкость (число каналов) системы связи, но тем меньше предельные расстояния, на которые возможна прямая передача между двумя

пунктами без ретрансляторов. Первая причина порождает тенденцию к освоению новых более высокочастотных диапазонов.

В территориальных сетях на региональном уровне часто используются ***радиорелейные линии связи*** (коммутация каналов, диапазон частот 15—23 ГГц, связь в пределах прямой видимости, что ограничивает дальность между соседними станциями до 50 км, антенны — последовательность ретрансляционных башен).

Радиоканалы входят необходимой составной частью в спутниковые и радиорелейные системы связи, применяемые в территориальных сетях, в сотовые системы мобильной связи, они используются в качестве альтернативы кабельным системам в локальных сетях и при объединении сетей отдельных ЛВС и предприятий в корпоративные сети. Радиосвязь используется в корпоративных сетях и ЛВС, если затруднена прокладка других каналов связи.

Радиоканал в локальных сетях реализует следующие функции:

- Выполняет роль моста между подсетями (двухточечное соединение с направленными антеннами, дальность в пределах прямой видимости, обычно — 15—20 км, с расположением антенн на крышах зданий). Мост имеет два адаптера: один — для формирования сигналов радиоканала, другой — для кабельной подсети.
- Радиоканал является общей средой передачи данных. В случае использования радиоканала в качестве общей Среды передачи данных в ЛВС сеть называют ***RadioEthernet*** (стандарт *IEEE 802/11*), она обычно используется внутри зданий. В состав аппаратуры входят приемопередатчики и антенны. Связь осуществляется на частотах от 1 до нескольких ГГц. Расстояние между узлами — несколько десятков метров
- Радиоканал служит соединением между центральными и терминальными узлами в сети с централизованным управлением. В варианте использования радиоканала для связи центрального и периферийного узлов центральный пункт имеет ненаправленную антенну, терминальные пункты при этом имеют направленные антенны. Дальность связи так же составляет десятки метров, а вне помещений — сотни метров.

Спутниковые каналы передачи данных. Спутники в системах связи могут находиться на геостационарных (около 36 тыс. км) или низких орбитах. При геостационарных орбитах заметны задержки при прохождении сигналов (туда и обратно около 500 мс). Возможно покрытие поверхности всего земного шара с помощью четырех спутников. В низкоорбитальных системах обслуживание конкретного пользователя происходит попеременно разными спутниками. Чем ниже орбита, тем меньше площадь покрытия — тем больше число наземных станций или спутников (обычно требуется около десятка спутников).

Основное оборудование, используемое в спутниковых системах связи, приведено на рис. 2.2.

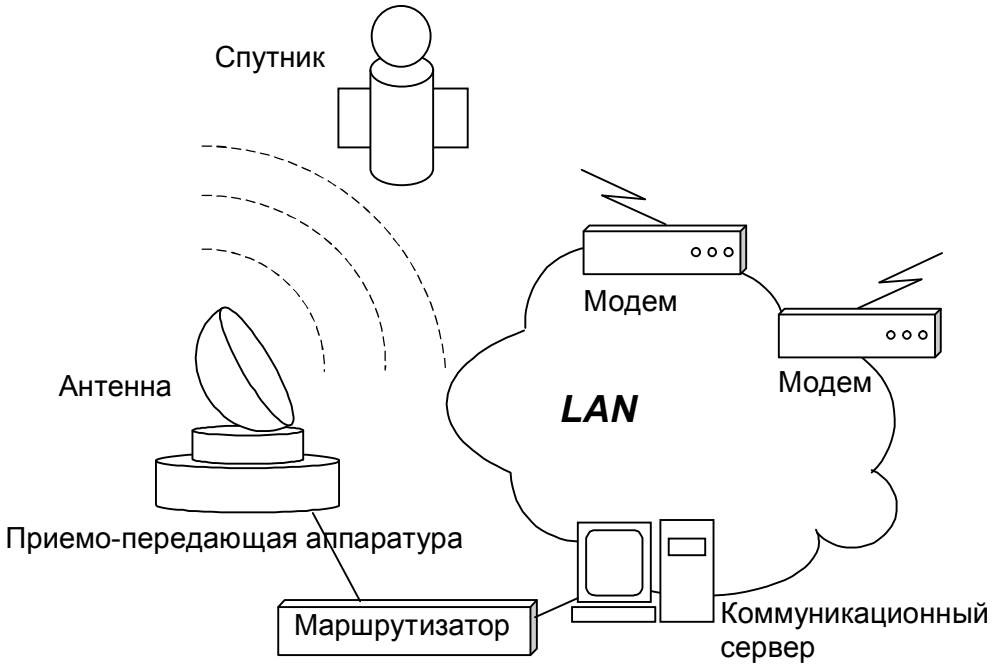


Рис. 2.2 — Схема спутниковой связи

Системы мобильной связи. Системы мобильной связи осуществляют передачу информации между пунктами, один из них или оба являются подвижными. Характерным признаком систем мобильной связи является применение радиоканала. К технологиям мобильной связи относятся также пейджинг, твейджинг, транкинг, для мобильной связи используются также и спутниковые каналы.

Пейджинг — система односторонней связи, при которой передаваемое сообщение поступает на пейджер (***pager***) пользователя, извещая его о необходимости предпринять или действие, или просто информируя его о тех или иных текущих событиях. Это наиболее дешевый вид мобильной связи.

Твейджинг — это двухсторонний пейджинг. В отличие от пейджинга возможно подтверждение получения сообщения и даже проведение некоторого подобия диалога.

Сотовые технологии обеспечивают телефонную связь между подвижными абонентами (ячейками). Связь осуществляется посредством базовых (стационарных) станций, выполняющих коммутирующие функции. Базовый коммутатор обслуживает некоторую зону.

Одной из наиболее широко распространенных технологий мобильной связи (в том числе и в России) является технология, соответствующая стандарту для цифровых сетей сотовой связи ***GSM*** (***Global System for Mobile Communications***), основанному на ***TDMA***.

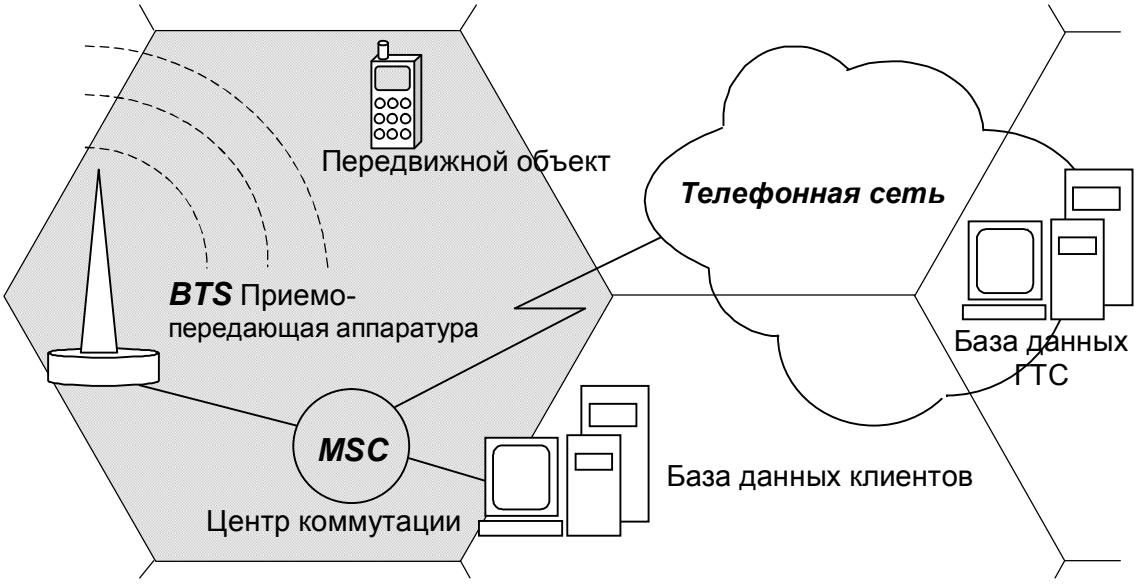


Рис. 2.3 — Схема сотовой телефонной связи

Архитектура GSM-системы аналогична архитектуре, приведенной на рис. 2.3. В каждой соте действует базовая станция ***BTS*** (***Base Transceiver Station***), обеспечивающая прием и передачу радиосигналов абонентам. BTS имеет частотный диапазон, отлич-

ный от диапазонов соседних сот. Мобильная ячейка прослушивает соседние BTS и сообщает контроллеру базовых станций ***BSC*** (***Base Station Controller***) о качестве приема. Контроллер BSC переключает мобильный объект на нужную BTS. Центр коммутации ***MSC*** (***Mobile services Switching Center***) осуществляет коммутацию и маршрутизацию, направляя вызовы нужному абоненту, в том числе и во внешние сети общего пользования. В базе данных хранятся сведения о местоположении пользователя, технических характеристиках мобильных станций, данные для идентификации пользователей.

Стандарт GSM может поддерживать интенсивный трафик (270 кбит/с), обеспечивает ***роуминг*** (т.е. автоматическое отслеживание перехода мобильного пользователя из одной соты в другую). Он так же допускает интеграцию речи и данных и связь с сетями общего пользования. Используются разновидности: сотовая связь ***GSM-900*** в частотном диапазоне 900 МГц (890—960 МГц), и микросотовая связь ***GSM-1800*** (DCS-1800) в диапазоне частот 1800 МГц (1710-1880 МГц). Название *микросотовая связь* обусловлено большим затуханием и, следовательно, меньшей площадью соты. Однако увеличение числа каналов выгодно при высокой плотности абонентов.

В перспективе предполагается использование широкополосного доступа *B-ISDN* на основе сотового стандарта ***UMTS*** (***Universal Mobile Telecommunication Systems***) с глобальным роумингом.

Мобильная связь для предприятий (т.е. ведомственная или профессиональная) может отличаться от сотовой связи индивидуальных пользователей. Такую ведомственную связь называют ***транкинговой*** (или ***транковой***). Для транкинговой связи характерны следующие особенности:

- связь внутри некоторой группы (бригады) и групповой вызов от центра ко всем членам группы;
- наличие приоритетности;
- скорость соединения должна быть выше, чем в обычных сотовых системах;
- возможность выхода в телефонную сеть общего пользования имеет меньшее значение, во многих случаях может вообще отсутствовать;

- преимущественная передача данных, в некоторых случаях голосовая связь не нужна; чаще используется полудуплексная передача.

В результате оперативность связи при уменьшении в цене.

2.2 Аналоговые и цифровые каналы

В зависимости от способа представления информации электрическими сигналами различают ***аналоговые*** и ***цифровые*** каналы передачи данных. Различие обеспечивается аппаратурой линии связи.

Типичным и наиболее распространенным типом аналоговых каналов являются телефонные каналы общего пользования (каналы тональной частоты), где полоса пропускания 0,3—3,4 кГц соответствует спектру человеческой речи.

Для передачи дискретной информации по каналам тональной частоты необходимы устройства преобразования сигналов, согласующие характеристики дискретных сигналов и аналоговых линий. Кроме того, в случае непосредственной передачи двоичных сигналов по телефонному каналу с полосой пропускания 0,3—3,4 кГц скорость передачи не превысит 3 кбит/с.

Действительно, пусть на передачу одного бита требуются два перепада напряжения, а длительность одного перепада $T_B = (3...4)/6.28 \cdot F_B$, где F_B — верхняя частота полосы пропускания. Тогда скорость передачи

$V < 1/(2 \cdot T_B)$.

Для передачи, например, цифрового сообщения по аналоговым линиям требуется изменение параметров сигнала в удобную форму для передачи сообщения, такое преобразование называется ***модуляцией,*** см. рис. 2.4.

Согласование параметров сигналов и среды при использовании аналоговых каналов осуществляется с помощью воплощения сигнала, выражающего передаваемое сообщение, в некотором процессе, называемом ***несущей*** и приспособленном к реализации в данной среде, в системах связи это электромагнитные колебания U некоторой несущей частоты:

$U = U_m \cdot \sin(\omega \cdot t + \varphi)$,

где U_m — амплитуда; ω — частота; φ — фаза колебаний несущей. Изменение параметров несущей по закону передаваемого сообщения называется ***модуляцией***. Если это изменение относится к амплитуде U_m, то модуляцию называют ***амплитудной (АМ),*** если к частоте ω — ***частотной (ЧМ)*** и если к фазе φ — ***фазовой (ФМ).*** При приеме сообщения предусматривается обратная процедура извлечения полезного сигнала из несущей, называемая демодуляцией. Модуляция и демодуляция выполняются в устройстве, называемом ***модемом***. Модем выполняет функции аппаратуры окончания канала данных.

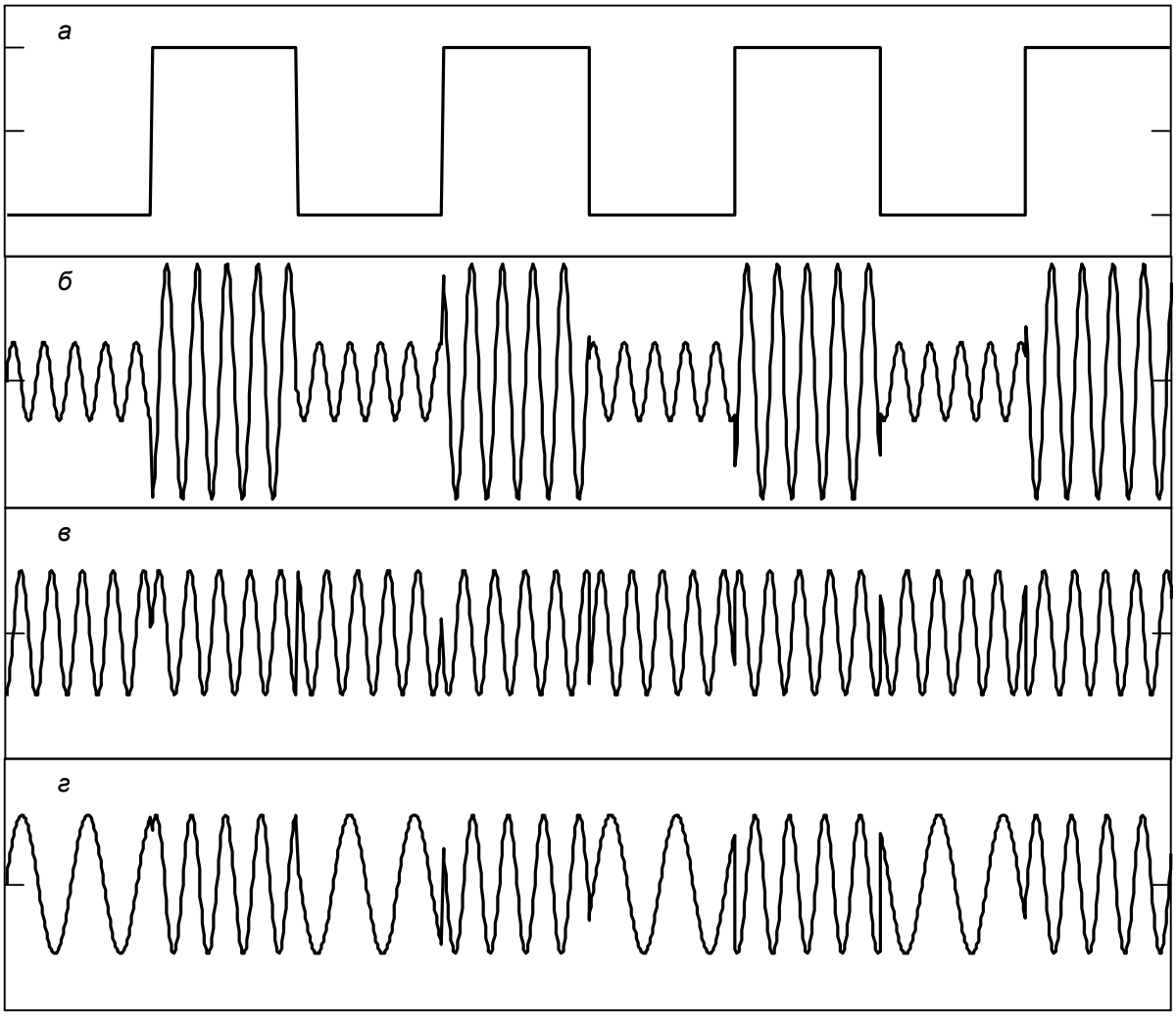


Рис. 2.4 — Виды аналоговой модуляции: цифрового сигнала (*а*): амплитудная (*б*), фазовая (*в*) и частотная (*г*)

Различают несколько технологий связи, основанных на ***цифровых каналах передачи данных***. Связь оконечного обору-

дования DTE с аппаратурой окончания канала данных DCE (например, компьютера с модемом или низкоскоростными периферийными устройствами) чаще всего осуществляется при помощи последовательных интерфейсов RS-232C, RS-422, RS-485 (их аналогами в системе стандартов ITU является V.24, V.11), а связь оконечного оборудования с цифровыми сетями передачи данных — при помощи интерфейсов X.21, X.35, G.703. Стандарты ITU серии V разрабатывались для передачи информации по телефонным линиям, а стандарты ITU серии X — для передачи данных.

В качестве *магистральных цифровых* каналов передачи данных в США и Японии применяют стандартную многоканальную систему ***T1*** (или ***DS-1****, Digital Signal-1*). Она включает 24 цифровых канала называемых ***DS-0.*** В каждом канале применена кодово-импульсная модуляция с частотой следования отсчетов 8 кГц и с квантованием сигналов по $2^8 = 256$ уровням, что обеспечивает скорость передачи 64 кбит/с на один канал или 1554 кбит/с на аппаратуру *T1*. Все 24 канала передают в *мультиплексор* (устройство формирования и передачи объединенного кадра) по байту, образуя 192-битный кадр с добавлением одного бита синхронизации. 24 кадра составляют суперкадр, в суперкадре имеются контрольный код и синхронизирующая комбинация. Сборку информации из нескольких линий и ее размещение в магистрали осуществляет мультиплексор. Канал *DS-0* (один слот) соответствует одному каналу, т.е. реализуется коммутация каналов. Некоторые мультиплексоры позволяют маршрутизировать потоки данных, направляя их в другие мультиплексоры, связанные с другими каналами *T1*, хотя собственно каналы *T1*, в общем случае, некоммутируемые. Применяются так же каналы ***T3*** (или ***DS-3***), состоящие из 28 каналов *T1* (45 Мбит/с).

В Европе более распространена аппаратура ***E1*** с 32 каналами по 64 кбит/с, т.е. с общей скоростью 2048 кбит/с. Применяются так же каналы ***E3*** (34 Мбит/с), состоящие из 16 каналов *E1*, но преимущественно в частных высокоскоростных сетях.

Аппаратура передачи дискретных компьютерных данных по аналоговым и цифровым линиям связи существенно отличается, так как в первом случае линия связи предназначена для передачи сигналов произвольной формы и не предъявляет никаких требований к способу представления единиц и нулей аппаратурой пе-

редачи данных, а во втором — все параметры передаваемых линией импульсов стандартизованы. Другими словами, на цифровых линиях связи протокол физического уровня определен, а на аналоговых линиях — нет.

2.3 Сети интегрального обслуживания и широкополосные каналы

Для передачи в единой среде голоса, звука, текста, изображения и данных применяются сети, называемые *сетями интегрального обслуживания*. *Несущей технологией*, в данном случае может выступать как цифровой, так и аналоговый сигнал. В многоканальной аппаратуре одна или несколько линий связи разделяется между сообщениями по частоте или времени. Для создания высокоскоростных каналов, которые мультиплексируют несколько низкоскоростных аналоговых абонентских каналов, при *аналоговом* подходе обычно используется техника *частотного мультиплексирования*.

В цифровых линиях связи передаваемые сигналы имеют конечное число состояний. Как правило, элементарный сигнал, то есть сигнал, передаваемый за один такт работы передающей аппаратуры, имеет 2 или 3 состояния, которые передаются в линиях связи импульсами прямоугольной формы. С помощью таких сигналов передаются как компьютерные данные, так и оцифрованные речь и изображение. В цифровых каналах связи используется промежуточная аппаратура, которая улучшает форму импульсов и обеспечивает их ресинхронизацию, то есть восстанавливает период их следования. Промежуточная аппаратура образования высокоскоростных *цифровых каналов* (мультиплексоры, демультиплексоры, коммутаторы) работает по принципу *временного мультиплексирования* каналов, когда каждому низкоскоростному каналу выделяется определенная доля времени (тайм-слот или квант) высокоскоростного канала.

Наиболее перспективными сетями интегрального обслуживания являются сети с цифровыми каналами передачи данных, например, сети ***ISDN (Integrated Services Digital Network).*** Сети ISDN могут быть как коммутируемыми, так и некоммутируемыми. Различают обычные ISDN со скоростями от 56 кбит/с до

1,54 Мбит/с и широкополосные ISDN (*Broadband ISDN* или иначе *B-ISDN*) со скоростями 155—2048 Мбит/с.

Функциональная схема одной из реализаций ISDN показана на рис. 2.5.

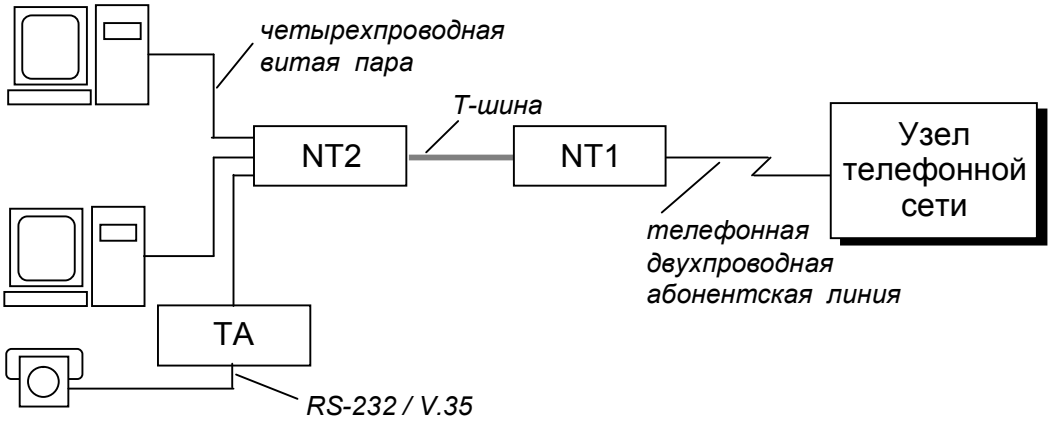


Рис. 2.5 — Функциональная схема ISDN

Применяют два варианта обычных сетей ISDN — базовый и специальный. В *базовом варианте ISDN* имеется два канала по 64 кбит/с (эти каналы называются *B-каналами*) и один служебный канал с 16 кбит/с (*D-канал*). В *специальном варианте ISDN* — 23 канала *B* и один или два служебных канала *D*. Каналы *B* используются как для передачи закодированной аналоговой информации (коммутация каналов), так и для передачи пакетов. Служебные каналы используются для сигнализации (передачи команд), например, вызова соединения и т.д. Приняты специальные сигнальные системы, устанавливающие перечень и форматы этих команд (основная такая система — *SS7 Signaling System — 7*).

Здесь оконечное оборудование соединяется с устройствами ISDN при помощи четырехпроводной витой пары, если же оконечное оборудование не имеет интерфейса ISDN, то оно подключается через адаптер TA. Устройство NT2 объединяет витые пары в одну Т-шину, которая имеет два провода от передатчика и два — к приемнику. Устройство NT1 реализует эхо-компенсацию и служит для интерфейса Т-шины с обычной телефонной двухпроводной абонентской линией. Очевидно, что для реализации технологий *T1*, *T3*, *E1*, *E3* и ISDN необходимо выбирать среду передачи данных с соответствующей полосой пропускания.

Цифровые абонентские линии xDSL. Для подключения клиентов к узлам магистральной сети с использованием на *последней миле* обычного телефонного кабеля наряду с каналами ISDN можно использовать цифровые абонентские линии xDSL. К их числу относятся ***HDSL*** (*High-bit-rate Digital Subcriber Loop*), ***SDSL*** (*Single Pair Symmetrical Digital Subcriber Loop*), ***ADSL*** (*Asymmetric Digital Subcriber Loop*).

Индивидуальный широкополосный доступ основывается на применении различных технологий семейства xDSL (табл. 2.2), развертываемых на базе существующей абонентской проводки городской телефонной сети. При этом абонентское оборудование представляет собой специальный DSL-модем, а оператор устанавливает у себя устройство концентрации трафика DSLAM.

Таблица 2.2

Технология	***Скорость передачи***	***Дальность***	***Топология/среда***
IDSL	128 Кбит/с	12 км	точка-точка, звезда / UTP Cat.3
HDSL	2 Мбит/с	6,5 км	точка-точка / UTP Cat.3
MSDSL	2 Мбит/с — 144 Кбит/с	6,5 км	точка-точка / UTP Cat.3
SDSL (G.shdsl)	2 Мбит/с — 144 Кбит/с	6 км	точка-точка, звезда / UTP Cat.3
ADSL	1 Мбит/с u/s 8 Мбит/с d/s	5,5 км	звезда / UTP Cat.3
VDSL	6,4 Мбит/с u/s 52 Мбит/с d/s	1,5 км	звезда / UTP Cat.3

Достоинства широкополосной технологии xDSL в высокой скорости доступа (т.к. полоса пропускания принадлежит пользователю целиком) и в использовании существующей инфраструктуры телефонной сети.

К недостаткам xDSL следует отнести высокую стоимость оборудования, требования провайдеру иметь специальное оборудование DSL (что для окупаемости заставляет иметь большое число пользователей и, следовательно, *высокую плотность* портов) и высокую чувствительность технологии к длине используе-

мой линии, т.к. пропускная способность обратно пропорциональна длине абонентской линии.

Указанные недостатки технологии xDSL относятся к индивидуальному типу пользовательского доступа, избежать их позволяет применение коллективного доступа, при котором оборудование xDSL размещается не на узле связи, а непосредственно в здании (помещении) группы абонентов. При этом внутри здания пользователей можно подключить и с помощью технологии отличной от xDSL, для чего выпускаются модульные коммутаторы, способные коммутировать наборные порты ADSL, VDSL, SHDSL, а также Ethernet 10/100 Base-TX и FX, HPNA 1.0 и 2.0.

Широкополосные линии беспроводного доступа функционируют описанным ниже способом. Оригинальный сигнал (аудио-, телевизионный или канал данных) поступает сначала на модулятор, затем на частотный конвертор, переносящий частоту исходного пакета в диапазон *несущего сигнала*. При этом каждый канал проецируется в свой частотный диапазон. Далее сигнал усиливается и через передающий модуль транслируется на всю область вещания базовой станции. Потом сигнал ретрансляторами усиливается и ретранслируется по всей области вещания. Наконец, ретранслированный сигнал поступает на приемник конечного пользователя и конвертируется в метровый или дециметровый диапазон для непосредственной подачи на телеприемник.

В компьютерных сетях сегодня применяются практически все описанные типы физических сред передачи данных, но наиболее перспективными являются волоконно-оптические. На них сегодня строятся как магистрали крупных территориальных сетей, так и высокоскоростные линии связи локальных сетей. Популярной средой является также витая пара, которая характеризуется отличным соотношением качества к стоимости, а также простотой монтажа. С помощью витой пары обычно подключают конечных абонентов сетей на расстояниях до 100 м от концентратора. Спутниковые каналы и радиосвязь используются чаще всего в тех случаях, когда кабельные связи применить нельзя — например, при прохождении канала через малонаселенную местность или же для связи с мобильным пользователем сети.

2.4 Классификация компьютерных сетей

Классификация компьютерных сетей по размерам

Вычислительные сети классифицируются по ряду признаков. В зависимости от расстояний между связываемыми узлами различают вычислительные сети:

- ***локальные*** (***ЛВС***) — охватывающие ограниченную территорию (обычно в пределах удаленности станций не более чем на несколько десятков или сотен метров друг от друга, реже на 1—2 км); локальные сети обозначают ***LAN*** (***Local Area Network***);
- ***корпоративные*** (***масштаба предприятия***) — совокупность связанных между собой ЛВС, охватывающих территорию, на которой размещено одно предприятие или учреждение в одном или нескольких близко расположенных зданиях;
- ***территориальные*** — охватывающие значительное географическое пространство; среди территориальных сетей можно выделить сети региональные и глобальные, имеющие соответственно региональные и мировые масштабы; региональные сети иногда называются сетями *MAN* (*Metropolitan Area Network*), а общее англоязычное название для территориальных сетей ***WAN*** (***Wide Area Network***);
- собо выделяют единственную в своем роде глобальную сеть ***Internet*** — это сеть сетей со своей технологией. В Internet существует понятие интрасетей (***Intranet***) — корпоративных сетей в рамках Internet.

Классификация компьютерных сетей по принадлежности

В зависимости от прав собственности на сети последние могут быть сетями общего пользования (***public***) или частными (***private***). Среди сетей общего пользования выделяют телефонные сети ***PSTN*** (***Public Switched Telephone Network***) и сети передачи данных ***PSDN*** (***Public Switched Data Network***).

Различают ***интегрированные сети***, ***неинтегрированные сети*** и ***подсети***. Интегрированная вычислительная сеть (***интерсеть***) представляет собой взаимосвязанную совокупность многих вычислительных сетей, которые в интерсети называются подсе-

тями. Интерсети необходимы так же для объединения технических средств автоматизированных систем проектирования в единые системы комплексной автоматизации — ***Computer Integrated Manufacturing*** (***CIM***). Обычно интерсети приспособлены для различных видов связи: телефонии, электронной почты, передачи видеоинформации, цифровых данных и т.п., и в этом случае они называются сетями интегрального обслуживания. Развитие интерсетей заключается в разработке средств сопряжения разнородных подсетей и стандартов для построения подсетей, изначально приспособленных к сопряжению. Подсети в интерсетях объединяются в соответствии с выбранной топологией с помощью блоков взаимодействия.

Классификация компьютерных сетей по топологии

В зависимости от *топологии* соединений узлов различают сети шинной (магистральной), кольцевой, звездной, иерархической и произвольной структуры (см. рис. 2.6).

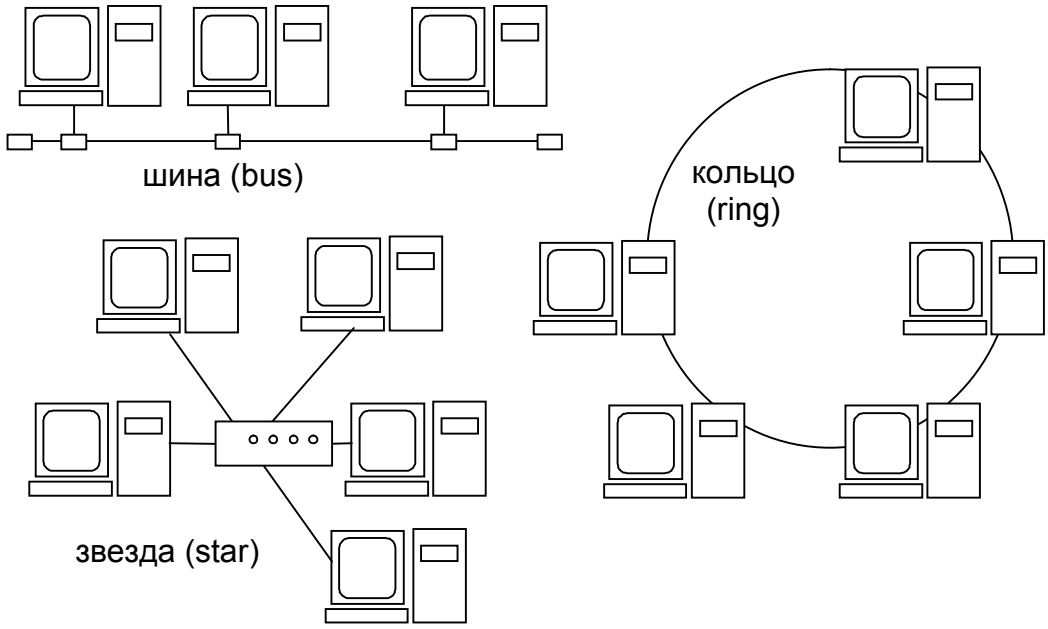


Рис. 2.6 — Основные топологические структуры сетей

- ***Шинная*** (***bus***) — локальная сеть, в которой связь между любыми двумя станциями устанавливается через один общий путь и данные, передаваемые любой станцией, одновременно

становятся доступными для всех других станций, подключенных к этой же среде передачи данных (последнее свойство называют широковещательностью).

- ***Кольцевая*** (***ring***) — узлы связаны кольцевой линией передачи данных (к каждому узлу подходят только две линии); данные проходя по кольцу, поочередно становятся доступными всем узлам сети.
- ***Звездная*** (***star***) — имеется центральный узел, от которого расходятся линии передачи данных к каждому из остальных узлов.

Комбинациями данных базовых топологических конструкций являются ***полносвязная*** топология (каждый компьютер соединен с каждым), ***иерархическая*** (древовидная) и ***смешанная*** топологии.

Классификация компьютерных сетей по способам управления

В зависимости от способа управления различают сети трех основных типов.

- ***Клиент/сервер*** — в которых выделяется один или несколько узлов (***серверов***), выполняющих управляющие или специальные обслуживающие функции, а остальные узлы (***клиенты***) являются терминальными, в них работают пользователи. Сети клиент/сервер различаются по характеру распределения функций между серверами, другими словами по типам серверов (например, файл-серверы, серверы баз данных и т.п.). При специализации серверов по определенным приложениям имеем сеть распределенных вычислений. Такие сети отличают также от централизованных сетей (сетей, построенных на мэйн-фреймах).
- ***Одноранговые*** — в них все узлы равноправны; поскольку в общем случае под клиентом понимается объект (устройство или программа), запрашивающий некоторые услуги, а под сервером — объект, предоставляющий эти услуги, то каждый узел в одноранговых может выполнять функции и клиента, и сервера.
- Наконец недавно появилась ***сетецентрическая концепция***, в соответствии с которой пользователь имеет лишь дешевое оборудование для обращения к удаленным компьютерам, а сеть обслуживает заказы на выполнение вычислений и получение ин-

формации, т.е. пользователю не нужно приобретать программное обеспечение для решения прикладных задач, ему нужно лишь платить за выполненные заказы. Подобные компьютеры называют *тонкими клиентами* или сетевыми компьютерами.

В зависимости от того, одинаковые или неодинаковые ЭВМ применяются в сети, различают сети однотипных ЭВМ или ***однородные***, и разнотипных ЭВМ — ***неоднородные*** (***гетерогенные***). В крупных автоматизированных системах, как правило, сети оказываются гетерогенными.

В зависимости от направления передачи линий различают ***симплексные*** (односторонняя передача), ***дуплексные*** (возможность передачи одновременной в обоих направлениях) и ***полудуплексные*** (возможность попеременной передачи в двух направлениях) линии связи. Различают так же одно- и многоканальные средства передачи данных в зависимости от числа каналов связи в аппаратуре передачи данных.

Установление соединения между отправителем и получателем с возможностью обмена сообщениями без заметных временных задержек характеризует режим работы ***on-line***, при существенных задержках с запоминанием информации в промежуточных узлах имеем режим ***off-line***.

2.5 Коммутация в информационных сетях

Существует два метода организации сквозной связи через сеть — *с соединением* и *без соединения*. В первом случае между передатчиком и приемником устанавливается определенный *маршрут* передачи информации. Во втором случае информация упаковывается в пакеты, которые могут прокладывать свой индивидуальный путь к приемнику через сеть (*дейтаграммы*).

Под ***коммутацией данных*** понимается их передача, при которой *физический канал* передачи данных может использоваться попеременно для обмена информацией между различными пунктами информационной сети в отличие от связи через некоммутируемые каналы, обычно закрепленные за определенными абонентами.

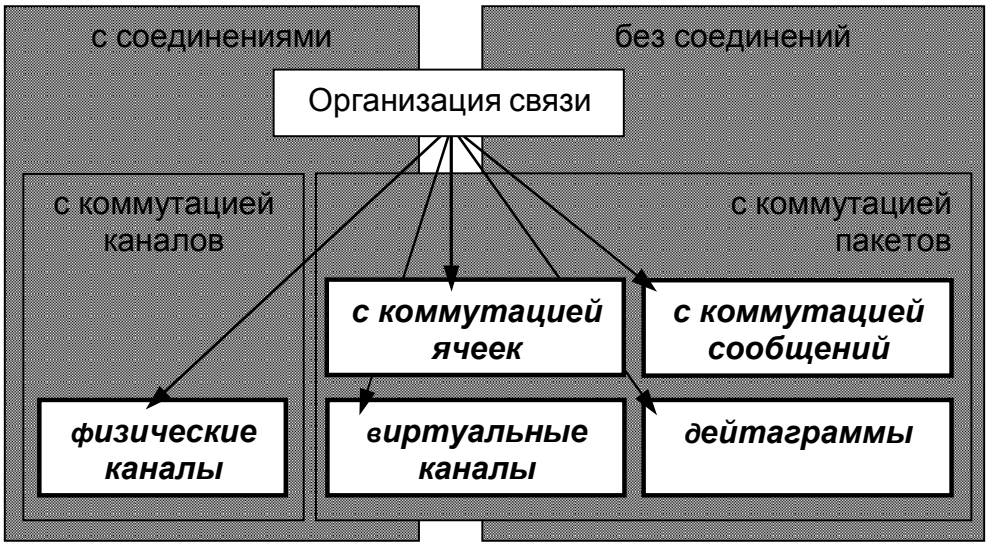


Рис. 2.7 — Классификация компьютерных сетей
по способам коммутации

Различают следующие способы коммутации данных:

- ***коммутация каналов*** — осуществляется соединение ООД двух или более станций данных и обеспечивается монопольное использование канала передачи данных до тех пор, пока соединение не будет разомкнуто (рис. 2.8). В этом случае требуется дополнительное время на установление и размыкание канала.

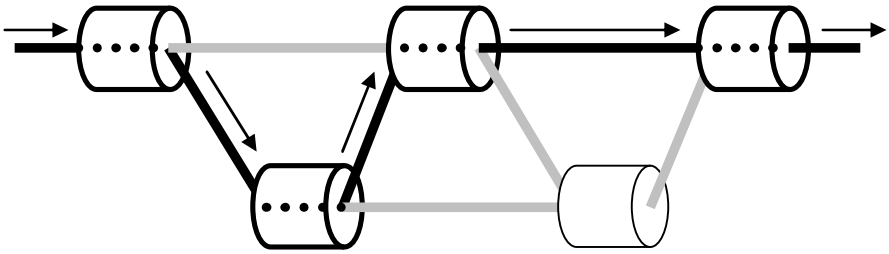


Рис. 2.8 — Коммутированный канал

Путь остается открытым в течение всего вызова, что чрезвычайно расточительно по отношению к пропускной способности канала. Но этот способ коммутации гарантирует определенную скорость передачи информации в течение всего сеанса не зависимо от общей нагрузки на сеть;

- ***коммутация пакетов*** — данные снабжаются адресной информацией и передаются по сети в отдельном блоке (*пакете*).

Пакет пересылается адресату через различные узлы сети. В этом случае канал используется только тогда, когда есть информация для передачи. Канал передачи данных занят только во время передачи пакета и по ее завершении освобождается для передачи других пакетов. Пакет передается через сеть поэтапно (без нарушения его целостности), сохраняясь в буферах промежуточного оборудования, пока соответствующее звено пути не освободится.

В зависимости от размера передаваемого в сеть пакета существует два типа пакетной коммутации — *коммутация сообщений* произвольной длины и *коммутация ячеек* фиксированного размера;

- ***коммутация сообщений*** — характеризуется тем, что создание физического канала между узлами необязательно и пересылка сообщений происходит без нарушения их целостности. Системы с коммутацией сообщений называют системами *хранения и отправки сообщений*. К недостаткам этого метода коммутации можно отнести большие задержки при передаче длинных сообщений и большой требуемый объем памяти промежуточных буферов;
- ***коммутация ячеек*** — все пакеты, называемые здесь ячейками, имеют фиксированную длину. Такой формат позволяет уменьшить объем работ, который сетевые узлы должны выполнять над пакетами. Сложность системы, таким образом, поддерживается на довольно низком уровне, а быстродействие — на высоком. Выбор оптимального размера ячеек — принципиальная для этого способа передачи и непростая задача: в слишком маленьких ячейках мал процент переносимой полезной информации (бо̀льшую ее часть занимает служебная информация), а слишком длинные ячейки увеличивают задержку и снижают скорость передачи и степень задействованности линии.

Асинхронный режим передачи ячеек (ATM) использует обычно ячейки с 48-байтовым информационным и 5-байтовым адресным полями. Для уменьшения размера адресной информации в ячейке обычно используются *виртуальные каналы*, в которых ячейки передаются по установленному маршруту в определенном порядке.

Методы коммутации пакетов разделяются на организацию связи с установлением соединений — *виртуальные каналы* и

коммутацию без установления соединений — *дейтаграммный* способ передачи;

- ***коммутация дейтаграмм*** — все пакеты рассматриваются и направляются через систему как отдельные объекты. Различные пакеты отправленные из источника *A* адресату *B* могут путешествовать по разным маршрутам и, возможно, прибудут к адресату не в исходной последовательности (рис. 2.9). Маршрут, по которому проходит пакет, будет в каждый конкретный момент времени зависеть от состояния сети. В заголовке пакета хранится его *порядковый номер* для восстановления их правильной последовательности у получателя. Заголовок дейтаграммы должен содержать адрес источника, адрес назначения и, возможно, информацию о выборе маршрута. Критерии, по которым выбирается маршрут в сети довольно разнообразны — это количество узлов на маршруте, протяженность и качество линий связи, их загруженность и даже стоимость аренды;

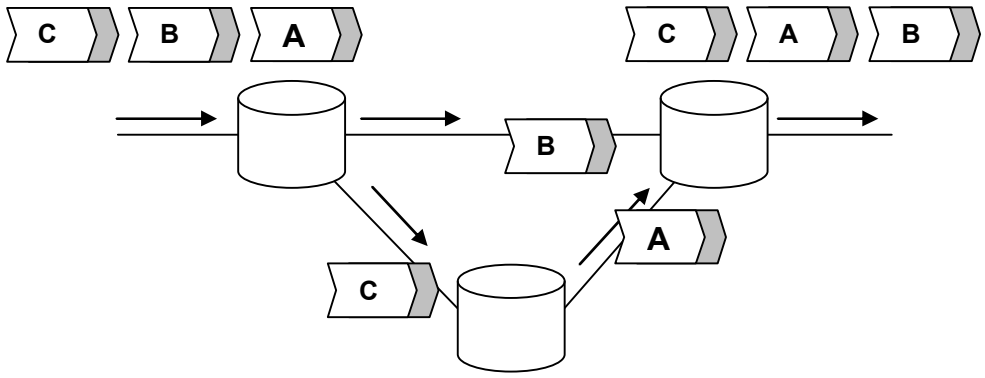


Рис. 2.9 — Коммутация дейтаграмм

- ***виртуальные каналы*** — немного напоминают физические каналы: первый пакет прокладывает путь между источником и пунктом назначения и по этому пути следуют все остальные пакеты сообщения. Пакеты прибывают в пункт назначения в определенном порядке (рис. 2.10). Такой способ передачи позволяет определить параметры информационного потока во всех промежуточных узлах, рассчитать требуемую нагрузку на сеть, а значит — эффективно управлять ею. Заголовок пакета виртуального канала проще, чем у дейтаграммы и, следовательно, занимает меньшую часть кадра.

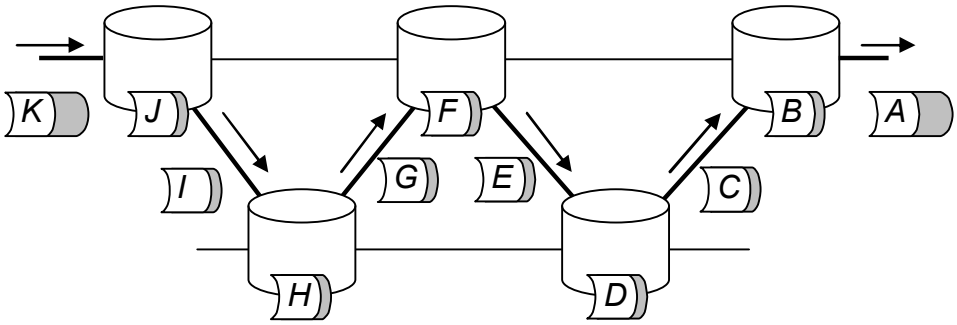


Рис. 2.10 — Виртуальный канал

Виртуальный канал — это логический двухточечный путь между двумя оконечными станциями. Выполняется обычно в три этапа: *создание канала (вызов)*, *передача данных* и *закрытие (отмена) канала*. Однако все эти преимущества умаляются тем, что коммуникационным службам легче пересекать сетевые границы (особенно границы между различными сетевыми технологиями), используя дейтаграммы, а не виртуальные каналы.

2.6 Мультиплексирование

Очевидно, что при передаче данных по сетям с различными коммутируемыми каналами возникают ситуации, когда требуется передача различных данных по одному каналу. Техника объединения нескольких коммутируемых каналов в один называется ***мультиплексированием*** (***multiplexing***). В настоящее время для мультиплексирования используются два основных способа:

- техника частотного мультиплексирования (***Frequency Division Multiplexing, FDM***), при которой в одном широкополосном канале располагается несколько узкополосных каналов, разнесенных по высокочастотным несущим. Для оптических кабелей эта же самая техника получила название разделения по длине волны (***Wave Division Multiplexing, WDM***);
- техника мультиплексирования с разделением времени (***Time Division Multiplexing, TDM***) ориентируется на дискретный (*цифровой*) характер передаваемых данных, при котором данные различных коммутируемых каналов объединяются *мультиплексором* и передаются в линию пакетами. На принимающем *де-*

мультиплексоре содержимое кадра распределяется по выходным каналам.

Мультиплексирование с разделением времени бывает двух основных типов — асинхронное ***Asynchronous*** *Time Division multiplexing (**ATD**)* и синхронное мультиплексирование с разделением времени ***Synchronous*** *Time Division multiplexing (**STD**).*

В системах синхронного временнòго мультиплексирования для передачи данных используются специальные кадры (*transmission frames*) разбитые на *интервалы* фиксированной длины. Эти интервалы разделяются между пользователями так, что позиция интервала в передающем кадре определяет пользователя, создавая тем самым *виртуальный канал*. Режим называется *синхронным*, потому что при передаче кадра определенный интервал прибывает в один и тот же момент времени от начала кадра (см. рис. 2.11).

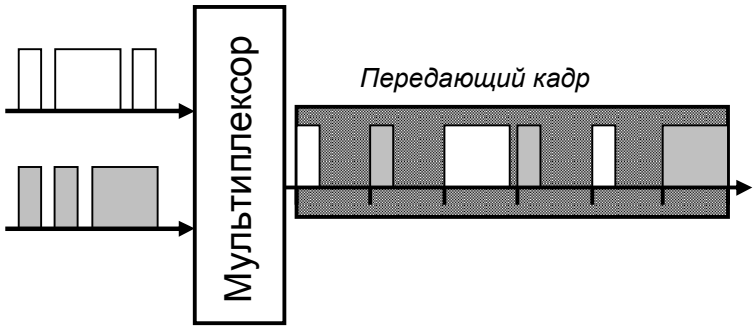


Рис. 2.11 — Синхронное мультиплексирование с разделением времени

Фиксированный временной интервал в составе нескольких кадров образует виртуальный канал, обслуживающий только одну передачу. Канал адресуется по номеру временного интервала. Все временные интервалы должны быть одинакового размера. Размеры интервала можно регулировать в зависимости от числа мультиплексированных в кадре каналов.

В асинхронной системе разделения времени нет распределения временных интервалов для каждого маршрута передачи. Вместо этого, пакеты пересылаются как можно быстрее по мере поступления (рис. 2.12). Пакеты, ожидающие передачи, просто хранятся в очереди. При возрастании нагрузки на систему все

большее число пакетов ожидает передачи в очереди, следовательно, пропускная способность системы ограничена, помимо скорости передачи, размерами буферов хранения передающих устройств. Эта система, очевидно, более выгодна при пульсирующем трафике.

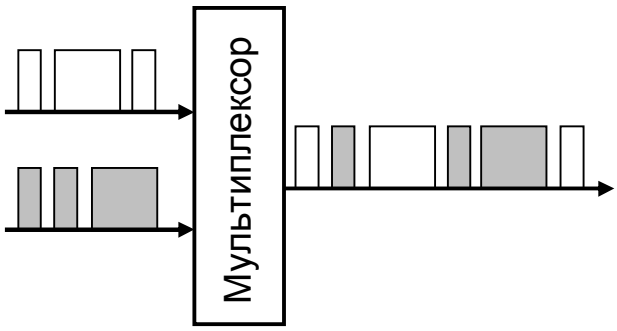


Рис. 2.12 — Асинхронное мультиплексирование с разделением времени

2.7 Качество работы сетевых служб

Цель коммуникационной системы состоит в переносе информации между двумя или несколькими конечными точками сети. Что касается приложений, программ и служб, работающих на прикладном уровне, то их задачей может быть передача речи, видеопотока, звука или файлов других типов, к сети же предъявляется требование просто передавать данные адресату в соответствии с некоторыми критериями качества. В стандарте ITU-T определено несколько критериев качества обслуживания (*Quality of service,* ***QoS***), которые разбиваются на две группы — *критерии связанные* и *несвязанные* с производительностью (эффективностью сетевых служб).

Параметры качества, не связанные с эффективностью служб

- ***Уровень обслуживания.*** Определяется степенью уверенности в том, что будут достигнуты уровни обслуживания, соответствующие *QoS*-критериям. Уровень обслуживания называется *детерминированным*, если гарантируется заданный QoS-уровень; *прогнозируемым*, если качество время от времени ухудшается из-

за статистического характера сетевых коммуникаций или *наилучшим из возможных*, если служба получает сетевые ресурсы в последнюю очередь и ничего не гарантирует в отношении уровня *QoS*.

- ***Стоимость.*** Это плата за использование сетевойслужбы согласованного качества.
- ***Приоритет.*** Определяет порядок работы служб — высокоприоритетные службы обслуживаются раньше низкоприоритетных.

Параметры качества, связанные с производительностью сети

- ***Задержка установки соединения (ISO), средняя задержка доступа или задержка установки вызова (ITU).*** Задержка между поступлением запроса на вызов соединения и подтверждением, что соединение было установлено.
- ***Вероятность отказа при установке соединения (ISO), вероятность блокировки (ITU).*** Вероятность того, что требуемое соединение не установлено (в пределах допустимой задержки установки).
- ***Пропускная способность*** (***throughput***) ***(ISO, ITU).*** Максимальное количество данных, которые могут быть успешно переданы за единицу времени по соединению в непрерывном режиме.
- ***Транзитная задержка (ISO), задержка кадра (ITU).*** Количество времени между отправлением *PDU*-блока и его получением.
- ***Вариации задержки (ISO), дрожание кадра (ITU).*** Расхождение между минимальной и максимальной кадровой (транзитной) задержкой.
- ***Скорость остаточных (не обнаруженных) ошибок (ISO), скорость кадровых ошибок (Frame Error Rate, FER).*** Вероятность того, что кадр или PDU-блок будет испорчен, потерян или дублирован в приемнике.
- ***Устойчивость (ISO), вероятность сброса соединения (ITU).*** Вероятность того, что соединение будет сброшено или восстановлено в пределах указанного интервала времени.

- ***Задержка в освобождении соединения (ISO, ITU).*** Задержка между выдачей конца запроса вызова и подтверждением освобождения соединения.

2.8 Беспроводные сети

Термин *беспроводные* применяется по отношению к телекоммуникационным сетям, в которых для передачи данных не используются кабельные линии. Средой передачи данных в беспроводных сетях служат электромагнитные волны, главным образом — радиоволны. Так как для перехвата радиосигналов потенциальному злоумышленнику достаточно настроить свой радиоприемник на соответствующую частоту, защита передаваемой информации в беспроводных сетях осуществляется с помощью различных методов шифрования передаваемого сигнала.

Беспроводные сети, также как сети на основе кабельных линий, можно разделить по размеру:

- на глобальные, Wireless WAN;
- муниципальные, Wireless MAN;
- локальные, Wireless LAN;
- персональные, Wireless PAN.

Глобальные беспроводные сети

Охватывают страны, континенты и весь земной шар. Основу глобальных беспроводных сетей составляют телефонные сети сотовой связи. В истории развития таких сетей выделяют несколько «поколений» (generation или просто G):

- Первое поколение, 1G — 80-е годы 20 века, первые сети сотовой связи, большое число отличающихся стандартов в разных странах: NMT, AMPS, TACS, C-450, Radiocom 2000, JTACS и т.д. Для передачи использовался аналоговый сигнал, защита от перехвата и прослушивания сообщений отсутствовала. Компьютерный трафик в таких сетях не передавался. В настоящее время сети первого поколения практически полностью вытеснены более современными типами мобильной связи;
- Второе поколение, 2G — 90-е годы 20 века, полностью цифровые сети сотовой связи. Наиболее распространенные стандарты: GSM, iDEN, D-AMPS. Используются две технологии мо-

дуляции TDMA — мультиплексирование с временным разделением каналов (используется в стандарте GSM) и CDMA — мультиплексирование с кодовым разделением каналов. Сообщения от одного абонента оказываются разделены на большое число «кусочков», собрать которые воедино может только специальная приемная станция, перехваченный сигнал потребует сложной расшифровки, поэтому передаваемый сигнал оказывается защищен от прослушивания, благодаря используемой при передаче модуляции. В сетях второго поколения появляются сервисы для передачи текстовых сообщений — SMS (Short Massage Service — служба коротких сообщений), компьютерного трафика — GPRS (General Packet Radio Services — пакетная радиосвязь общего назначения), мультимедийной информации (текста с аудио и видеофрагментами, картинками) — MMS, EMS. Скорость передачи данных по протоколу GPRS сравнима с проводным модемным соединением (Dial-Up) — до 115 Мбит/с.

- Третье поколение, 3G — развивается с началом 21 века (первая 3G мобильная сеть в Японии — 2001 г., в Европе — 2003 г.). Сети третьего поколения создаются в настоящее время. Главная задача, которую должны решить эти сети — обеспечение высокоскоростного доступа в Интернет и обмена видео в реальном времени — видеотелефония. Передача компьютерного трафика осуществляется с помощью более скоростных, по сравнению с GPRS, технологий: EDGE (Enhanced Data rates for Global Evolution — улучшенная для перехода к сетям 3G), скорость передачи данных до 384 кбит/с, и UMTS (Universal Mobile Telecommunications System — универсальная система мобильной связи), технология сетей 3G, скорость до 2 Мбит/с. В глобальных беспроводных сетях третьего поколения предполагается интеграция с локальными беспроводными сетями — WLAN.
- Четвертое поколение, 4G — будущее беспроводных глобальных сетей. В сетях четвертого поколения предполагается полная интеграция всех сервисов и типов сообщений как компьютерных сетей, так и сотовой связи, высокая степень защищенности передаваемого трафика на основе шифрования, передача видеосигналов высокой четкости, скорость передачи данных до 100 Мбит/с для мобильного абонента и до 1 Гбит/с для стационарной станции. Уже сейчас появляются технологии, претен-

дующие на роль переходных от сетей третьего поколения к четвертому, одна из них HSDPA (High-Speed Downlink Packet Access — высокоскоростная пакетная передача данных) — часть UMTS, появившаяся в ее последних реализациях. Технология HSDPA уже сейчас теоретически способна передавать данные со скоростью 14,4 Мбит/с, ведутся работы по дальнейшему увеличению скорости передачи. В качестве способа модуляции сигнала технология HSDPA использует модуляцию QPSK (Quadrature Phase Shift Keying — квадратурная с фазовым сдвигом) и 16QAM (Quadrature Amplitude Modulation — квадратурная амплитудная).

Муниципальные беспроводные сети

Беспроводные сети с радиусом действия порядка 10 км, WMAN. Введены стандартом IEEE 802.16 в 2002 г. Стандарт отводит для использования в таких сетях диапазон радиоволн от 10 до 66 ГГц, при этом передающая и приемная станции должны находиться в условиях прямой видимости. Оборудование, реализующее стандарт 2002 года, не производилось. Распространение получила технология широкополосного радиодоступа WiMAX, рис. 1, основанная на более поздних модификациях этого стандарта:

- IEEE 802.16d-2004, стандарт для неподвижных приемников и передатчиков, Fixed WiMAX;
- IEEE 802.16e-2005, стандарт для подвижных приемников и передатчиков, Mobile WiMAX.

Цифры 2004, 2005 — годы введения стандартов. Собственно говоря, IEEE 802.16e-2005 просто расширил существующий к этому времени стандарт на работу с мобильными устройствами. Это стало возможно благодаря тому, что вместо OFDM — модуляции (Orthogonal Frequency Division Multiplexing — мультиплексирование с ортогональным частотным разделением), использовавшейся в Fixed WiMAX, было предложено применить SOFDM (Scalable OFDM — масштабируемую OFDM). Кроме того, в новой версии стандарта вводилась поддержка нескольких приемных антенн — MIMO (Multiple-input multiple-output communications), тогда как в стандарте Fixed WiMAX сигналы от нескольких абонентов принимались на одну антенну. Одно из главных преимуществ технологии WiMAX — нет необходимости соблюдения ус-

ловия прямой видимости. Частотный диапазон сетей WiMAX — от 2 до 11 ГГц, большинство существующих устройств WiMAX работают в диапазоне 5—6 ГГц. Теоретически предельная скорость передачи данных до абонента — 70 Мбит/с. Реально получаемые результаты в настоящее время значительно скромнее, порядка 10 Мбит/с. Для защиты передаваемых данных могут использоваться различные методы шифрования, например, 3DES. В мире развернуто около сотни сетей WiMAX, однако высокая стоимость оборудования, слабая поддержка производителями и быстрый рост скорости передачи данных в глобальных беспроводных сетях делают перспективы такого вида сетей трудно предсказуемыми.

В России сети WiMAX стали появляться в 2006 г. Научно-производственная фирма «Микран» при ТУСУРе, первая, и пока единственная из российских производителей, выпустила оборудование WiMAX стандарта IEEE 802.16d-2004 — несколько десятков базовых и абонентских станций, под торговой маркой WiMIC-6000. Стоимость абонентской станции 40—50 тыс. руб., базовой станции — порядка 200 тыс. руб.

Локальные беспроводные сети

Основой для построения локальный беспроводных сетей (WLAN) стал стандарт IEEE 802.11, принятый в 1997 г. Этот стандарт предполагал работу беспроводных устройств как в инфракрасном диапазоне волн, так и в микроволновом частотном диапазоне — 2,4—2,485 ГГц при скорости передачи данных 1 Мбит/с и 2 Мбит/с. В качестве метода доступа был выбран метод множественного доступа с распознаванием коллизий CSMA/CS, применяющийся в сетях Ethernet. ИК-диапазон остался частью стандарта, но практического воплощения не получил. Устройства, работающие в радиодиапазоне были выпущены несколькими кампаниями. В 1999 г. стандарт расширяется введением двух более скоростных версий:

- IEEE 802.11a, фактически новый стандарт, использующий более высокочастотный диапазон — 5 ГГц, OFDM модуляцию и обеспечивающий скорость передачи до 54 Мбит/с;
- IEEE 802.11b, являющийся развитием стандарта IEEE 802.11. При работе в том же частотном диапазоне скорость пере-

дачи данных увеличивается до 11 Мбит/с за счет использования модуляции CCK (Complementary code keying). В оригинальной версии стандарта применялась технология DSSS (Direct-Sequencing Spread Spectrum — расширение спектра сигнала прямой последовательностью) с двоичной относительной фазовой манипуляцией DBPSK и квадратурной относительной фазовой модуляцией DQPSK.

Организация Wireless Ethernet Compatibility Alliance, занимающаяся вопросами сертификации оборудования, выпускаемого по стандарту IEEE 802.11, в том же 1999 г. по аналогии с термином Hi-Fi создает торговую марку ***Wi-Fi*** (иногда, расшифровываемую, как Wireless Fidelity — «Беспроводная надёжность»), изменяя собственное название на ***Wi-Fi Alliance.*** Оборудование стандарта IEEE 802.11b, благодаря незначительным отличиям от предыдущей версии стандарта, получило во всем мире широкое распространение. Первые устройства стандарта IEEE 802.11a появляются значительно позже, только в 2001 г, а в 2003 г. был принят еще один стандарт беспроводной связи IEEE 802.11g, который можно рассматривать, как дальнейшее развитие стандарта IEEE 802.11b. При работе в том же частотном диапазоне (2,4—2,485 ГГц), этот стандарт, благодаря применению OFDM модуляции, позволяет достигать скорости передачи данных 54 Мбит/с. Стандарт IEEE 802.11g поддерживает и работу устройств стандарта IEEE 802.11b, допуская создание «смешанных» беспроводных сетей. Совместимость обеспечивается автоматическим переходом сетевого адаптера стандарта IEEE 802.11g на работу по стандарту IEEE 802.11b при включении в сеть соответствующего устройства. При этом скорость работы сети ограничивается максимальными для стандарта IEEE 802.11b — 11 Мбит/с.

В 2004 г. началась работа над новым стандартом локальных беспроводных сетей с еще более высокой пропускной способностью — IEEE 802.11n. Утверждение этого стандарта ожидается в 2008—2009 гг., но уже сейчас некоторые производители начинают выпуск своих версий оборудования нового стандарта. Адаптеры этого стандарта будут работать в двух диапазонах: 2,4 и 5 ГГц. При использовании технологии MIMO — нескольких приемо-передающих антенн в одном сетевом адаптере, пространственной селекции сигналов и новой схемы кодирования, плани-

руется достигнуть максимальной пропускной способности 74 Мбит/с.

Необходимо отметить, что стандарты IEEE 802.11a и IEEE 802.11b не совместимы, т.к. работают в разных диапазонах волн. Однако многие производители предлагают устройства с поддержкой всех трех версий стандарта IEEE 802.11a, b и g. Такие устройства фактически совмещают в одном корпусе два сетевых адаптера, один — стандартов IEEE 802.11b/g, другой стандарта IEEE 802.11a. На основе таких универсальных, двухдиапазонных точек доступа возможно создание беспроводной сети, сетевые адаптеры в которой работают и по тому и по другому стандарту. Некоторые производители, кроме того, предлагают оборудование стандарта 802.11g+ (SuperG), указывая что сетевой адаптер может передавать данные на скорости в 100, 108 или даже 125 Мбит/с. Это расширение стандарта 802.11g, реализованное самими производителями соответствующего оборудования. Многие из них в том или ином виде реализовали такой расширенный режим работы сетевого адаптера, но, поскольку стандарта на такой режим еще не существует, нет никакой гарантии, что решения различных производителей смогут взаимодействовать друг с другом.

Сети Wi-Fi — название, объединяющее беспроводные сети, работающие по любому из типов стандарта IEEE 802.11, благодаря наличию встроенных адаптеров практически во всех выпускаемых мобильных устройствах: ноутбуках, карманных компьютерах — КПК и др., а также простоте развертывания такой сети, стали сегодня одним из символов современности. Они работают в аэропортах, торговых центрах, университетах, кафе, гостиницах, при проведении конференций, симпозиумов и т.п.

Стандартами предусмотрено два режима работы сетевого адаптера в сети Wi-Fi:

- Режим одноранговой сети, когда компьютеры соединяются между собой с помощью беспроводных сетевых адаптеров, рис. 2.13. Этот режим называют режимом независимых базовых зон обслуживания ***IBSS*** (***Independent Basic Service Set***) или ***ad hoc*** (это словосочетание имеет латинское происхождение и означает *специальный*). Такой режим — аналог прямого соединения *компьютер — компьютер* с помощью обычных сетевых адаптеров и кроссоверного кабеля.

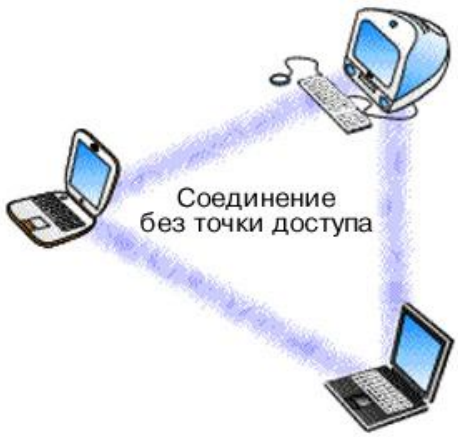


Рис. 2.13 — Беспроводное соединение
типа Ad-hoc

- Режим с выделенной точкой доступа (***Access Point, AP***) или режим инфраструктуры (***infrastructure***). В этом режиме в сети работает по крайней мере один беспроводной концентратор — точка доступа, обеспечивающая связь всех остальных беспроводных устройств между собой.

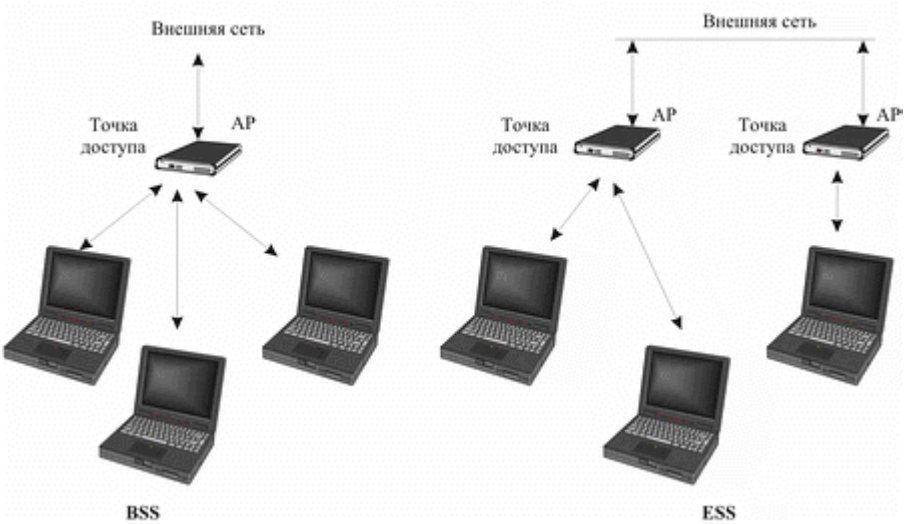


Рис. 2.14 — Беспроводная сеть с точкой доступа

Основной вид режима инфраструктуры — BSS (Basic Service Set) предусматривает работу через одну точку доступа. Расширенный вид — ESS (Extended Service Set) допускает работу в общей сети нескольких точек доступа, каждая из которых обеспечивает связь для своей группы клиентов (BSS), рис. 2.14. Меж-

ду собой точки доступа соединяются с помощью либо сегментов кабельной сети, либо радиомостов.

Диапазон действия Wi-Fi сети зависит от режима работы, поддерживаемого адаптером стандарта и условий распространения радиосигнала. Дело в том, что мощность излучения беспроводного адаптера или точки доступа, от которой, в первую очередь, зависит дальность связи, не может превышать 100 мВт. Обычно точка доступа имеет излучатель с мощностью, близкой к максимально допустимой, а мощность излучения сетевых адаптеров намного ниже. Ограничения на максимальную мощность связаны с тем, что электромагнитное излучение используемых длин волн опасно для человека, оно интенсивно поглощается биологическими тканями, вызывая их нагрев. Например, микроволновое излучение мощностью в несколько сотен ватт с близкой длиной волны (частота порядка 2 ГГц) используется в бытовых СВЧ-печах для приготовления пищи.

Микроволновые печи из-за близости длины волны к частотам сетевых адаптеров стандартов IEEE 802.11b/g могут быть источником интерференционных помех для развернутой Wi-Fi сети, значительно сужая радиус ее действия и уменьшая скорость передачи данных. Такого же рода помехи в диапазоне 2,4 ГГц создают некоторые сотовые телефоны и локальные беспроводные сети, развернутые по-соседству. Эффективным методом борьбы с помехами, создаваемыми соседней Wi-Fi сетью, является отстройка по частоте. Стандарт разрешает работу на 13 независимых частотных каналах, полоса частот каждого канала имеет ширину 5 МГц. Однако, для того чтобы две различных близко расположенных беспроводных сети не создавали помех друг для друга, они должны работать на частотах, различающихся не менее чем на 22 МГц. Это значит, что от соседней беспроводной сети необходимо отстроиться не менее чем на 5 каналов. Отсюда известное правило **1, 6, 11** — указывающее номера каналов, на которые должны быть настроены пересекающиеся беспроводные сети, рис. 2.15. Автоматически такая отстройка обычно не выполняется. Но программные утилиты, поставляемые вместе с беспроводным адаптером позволяют выполнить настройку вручную.

В режиме *ad hoc* расстояние между взаимодействующими станциями не превышает 10 метров в условиях прямой видимости, — в одной комнате. Наличие преград — стен, перекрытий и т.п, требует повышенной мощности сигнала и обязательного включения в беспроводную сеть точки доступа.

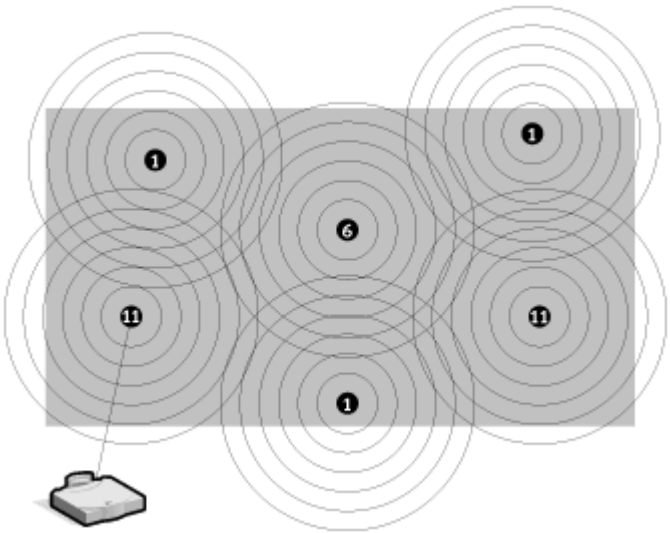


Рис. 2.15 — Правильная настройка номеров каналов пересекающихся беспроводных сетей.

В таблице 2.3 приведена примерная дальность действия беспроводной сети с одной точкой доступа для разных типов стандарта. Указана дальность для максимальной скорости передачи данных. По мере увеличения расстояния скорость передачи снижается до минимального значения 1 МГц, пока связь не будет потеряна окончательно.

Таблица 2.3

Стандарт	Максимальная дальность в помещении (метров)	Максимальная дальность на открытом пространстве (метров)
IEEE 802.11b	40	90
IEEE 802.11g	40	90
IEEE 802.11a	35	75

Использование направленных внешних антенн в условиях прямой видимости может многократно увеличить дальность связи (по некоторым данным до 8 км).

Безопасность в беспроводных сетях

Настройка беспроводного Wi-Fi адаптера для работы в сети требует задания не только тех же параметров, которые необходимы и при использовании обычного адаптера: IP-адреса и т.д., но и специфических параметров, характерных только для беспроводной сети. В первую очередь, это имя сети, — ***SSID***, ***Service Set IDentifier***, и, затем, ключи шифрования, т.к. весь передаваемый по радиоканалу объем данных предварительно шифруется. Задание ключей шифрования (а значит и включение самого режима шифрования) не обязательно, беспроводная сеть может быть построена как открытая система (*Open System*), однако, вероятность внедрения злоумышленника в такую систему сегодня столь велика, что использование шифрования настоятельно рекомендуется.

Для защиты беспроводных сетей первоначально был разработан и предложен стандарт ***WEP*** (***Wired Equivalency Privacy*** — эквивалент проводной безопасности). WEP-шифрование осуществляется 64 или 128-битовым ключом, часть которого: 40 или 104 бит, вводится пользователем, а дополнительные 24 бит представляют собой автоматически генерируемую случайную последовательность — вектор инициализации. Длина ключа в стандарте WEP оказалась слишком малой. Если злоумышленник может просто пассивно захватывать сетевые пакеты, ему достаточно получить порядка 4 млн. пакетов, чтобы определить ключ шифрования, это одна из самых простых атак. На интенсивно работающей сети захват 4 млн. пакетов занимает примерно один час. Прослушивать Wi-Fi сеть и захватывать чужие сетевые пакеты в принципе возможно, так как беспроводной адаптер имеет в своем составе приемо-передатчик соответствующего диапазона. Если заставить его работать не по стандарту, то он сможет захватывать все сигналы в наблюдаемом диапазоне частот.

Такой режим, его называют режимом мониторинга (*Monitoring*) обеспечивается соответствующей программой-драйвером. Он, естественно, не поддерживается производителями оборудования. Тем не менее, хакерами подобные драйверы создаются и они доступны в Интернете. В 2005 г. группа сотрудников ФБР с помощью *общедоступного программного обеспечения* продемонстрировала взлом беспроводной сети, защищенной WEP-шифрованием за 3 минуты.

Слабость WEP-шифрования стала очевидной вскоре после его появления, поэтому в 2003 г. Wi-Fi Alliance предложил его замену — шифрование ***WPA*** (***Wi-Fi Protected Access***). Этот алгоритм шифрования поддерживается всеми устройствами стандарта ***IEEE 802.11g*** (сетевые адаптеры стандарта IEEE 802.11b поддерживают только WEP-шифрование). WPA-шифрование может выполняться двумя способами:

- на основе стандарта IEEE 802.1x;
- на основе предварительно выданных ключей WPA-PSK.

Стандарт IEEE 802.1x определяет правила безопасной аутентификации в сетях любого типа (как проводных, так и беспроводных), в частности в нем описывается протокол аутентификации ***EAP*** (***Extensible Authentication Protocol***), использующий для аутентификации пользователей внешний сервер, так называемый ***RADIUS-***сервер (***Remote Authentication Dial-In User Service***). Этот вариант обычно используется в крупных беспроводных сетях, сетях предприятий. Шифрование

WPA-PSK (***Pre-Shared Key*** — заранее выданный ключ) основано на пароле, вводимом пользователем, и применяется для малых беспроводных сетей. Стойкость шифрования в этом случае определяется качеством пароля. И в том и в другом случае, после прохождения процедуры аутентификации ключи шифрования изменяются для каждого передаваемого пакета по сложному алгоритму. Протокол динамического изменения ключей шифрования носит название ***TKIP*** (***Temporal Key Integrity Protocol***).

WPA-PSK шифрование в случае применения простого пароля (123, qwerty, и т.п.) также может быть легко взломано в случае захвата злоумышленником сетевых пакетов, содержащих пароль, обмен которыми происходит на этапе подключения рабочей станции к точке доступа. Существуют даже программы, принудительно разрывающие соединение, для облегчения задачи перехвата пакетов аутентификации. Однако хороший пароль может служить достаточно надежной защитой беспроводного соединения.

Для дополнительной защиты беспроводного соединения можно использовать так называемую «фильтрацию по MAC-адресам», включив в конфигурации точки доступа ограничение на подключение к сети клиентов только с указанными физиче-

скими адресами сетевых адаптеров. Практически во всех современных точках доступа реализована такая возможность.

Методы защиты беспроводных соединений продолжают совершенствоваться. В 2004 г. был принят стандарт безопасности беспроводных сетей ***IEEE 802.11i,*** известный также как ***WPA2***, в нем на смену протоколу TKIP пришло более надежное шифрование ***AES*** (***Advanced Encryption Standard***).

Беспроводные персональные сети

Сети ***WPAN*** описываются стандартом ***IEEE 802.15.*** Они позволяют устанавливать беспроводные сетевые соединения с устройствами, используемые внутри *личного рабочего пространства* (***Personal Operating Space, POS***). Под личным рабочим пространством понимается пространство, окружающее пользователя, радиусом менее10 метров. В настоящий момент основной технологией сетей WPAN является Bluetooth, но в последние годы стали появляться альтернативные технологии. В будущем предполагается переход в персональных сетях на технологию сверширокополосного доступа ***UWB*** (***Ultra-Wide Band, Ultraband***), позволяющую значительно увеличить скорость обмена данными.

Bluetooth

Технология Bluetooth (стандарт ***IEEE 802.15.1***) разрабатывалась с целью устранения соединительных кабелей различных устройств, находящихся на небольших расстояниях друг от друга, например, компьютера и клавиатуры или микрофона и сотового телефона. Устройства Bluetooth работают в том же частотном диапазоне 2,4—2,485 ГГц, что и устройства стандарта 802.11b, поэтому они могут создавать взаимные помехи. И, как и WLAN, они также подвержены действию помех от микроволновых печей и сотовых телефонов. По уровню мощности устройства Bluetooth делятся на три класса:

- 1 класс, мощность до 100 мВт. Такие устройства обязательно имеют систему управления излучаемой мощностью для того, чтобы использовать только минимально необходимую в данный момент мощность. Дальность действия устройств первого класса — 100 метров. Они предназначены для использования в промышленности и мало распространены;

- 2 класс, максимальная мощность 2,5 мВт. Дальность действия — 10 метров;
- 3 класс, максимальная мощность 1 мВт. Дальность действия — 1 метр.

Весь частотный диапазон делится на 79 каналов по 1 МГц каждый. Технология Bluetooth использует модуляцию ***GFSK*** (***Gaussian Frequency Shift Keying***), при передаче каждого очередного пакета частота скачкообразно изменяется. Выбор канала передачи при этом осуществляется на основе псевдослучайной последовательности. Таким образом, если рядом работают несколько пар приёмник-передатчик, то они не мешают друг другу.

Этот алгоритм является также составной частью системы защиты конфиденциальности передаваемой информации: переход происходит по псевдослучайному алгоритму и определяется отдельно для каждого соединения. В стандарте Bluetooth предусмотрено шифрование передаваемых данных ключом длиной от 8 до 128 бит с возможностью выбора односторонней или двусторонней аутентификации, что позволяет устанавливать стойкость результирующего шифрования в соответствии с законодательством каждой страны (в некоторых странах запрещено использование сильной криптографии). В дополнение к шифрованию на уровне протокола может быть применено шифрование на уровне приложений — здесь уже применение сколь угодно стойких алгоритмов никто не ограничивает.

Каждое устройство в сети Bluetooth (ее называют *пикосеть* — ***piconetwork***) имеет уникальный 48-битовый адрес и может быть ведущим (***master***) или ведомым (***slave***). Инициатор соединения становится ведущим, ведомые отвечают ведущему. Ведущий может обмениваться информацией с семью активными ведомыми. До 255 ведомых могут в это время находится в неактивном состоянии (***parked***), но в любой момент времени ведущий может перевести любое из этих устройств в активное состояние. Стандарт определяет два типа соединений, которые поддерживают передачу речи и данных:

- Синхронная линия связи, ориентированная на соединение (***SCO, Synchronous Connection Oriented***). Обеспечивает создание постоянного канала «точка — точка» между ведущим и ведомым. Повторная передача пакетов в случае ошибок не выполняется;

- Асинхронная линия связи без соединения (***ACL, Asynchronous Connectionless Link***). Обычно используется для передачи данных, в случае возникновения ошибок они корректируются повторной отправкой пакетов.

В общем случае предполагается работа устройств в распределенной (***scatternet***) сети, такая сеть образуется, когда одно устройство является ведомым в нескольких пикосетях или ведущим в одной пикосети и ведомым в другой. Коммуникация между пикосетями осуществляется через устройство, одновременно входящее в несколько пикосетей. Распределенная сеть при асинхронных связях между пикосетями может объединять до 10 пикосетей. Сеть Bluetooth, состоящая только из одной пикосети считается специальной (***ad-hoc***).

Технология Bluetooth с момента появления первых устройств прошла несколько этапов развития:

- первая версия, получившая широкое распространение — ***Bluetooth 1.1***, появилась в 2000 г., максимальная скорость передачи в синхронном режиме 433 кбит/с, в асинхронном — 721 кбит/с;
- в разработанной в 2003 г. версии ***Bluetooth 1.2*** была добавлена технология адаптивной перестройки рабочей частоты (***AFH***, ***Adaptive Frequency-Hopping spread spectrum***), что улучшило помехоустойчивость. Появилась поддержка трехпроводного интерфейса UART;
- в версии ***Bluetooth 2.0***, выпущенной в 2004г., скорость передачи повысилась до 2,1 Мбит/с;
- в ***Bluetooth 2.1*** (2007 г.) осуществлена поддержка энергосберегающей технологии Sniff Subrating, которая позволяет увеличить продолжительность работы устройства от одного заряда аккумулятора как минимум в пять раз. Кроме того обновленная спецификация существенно упрощает и ускоряет установление связи между двумя устройствами, а также делает указанные соединения более защищенными, благодаря технологии Near Field Communication.

Ожидается, что версия ***Bluetooth 3.0***, которая должна прийти на смену Bluetooth 2.1, будет использовать технологию сверширокополосного доступа ***UWB*** (***Ultra Wideband***), что позволит достичь скорости передачи данных в 480 Мбит/с.

ZigBee

Технология радиодоступа ZigBee стала развиваться с 2004 г (первый стандарт ***IEEE 802.15.4***). Основной идеей было получить не такую быстродействующую, как Bluetooth, но более простую и дешевую технологию маломощной беспроводной связи с высокой степенью защиты передаваемых данных. Максимальная мощность передатчика ZigBee ограничена 1 мВт, используется тот же частотный диапазон 2,4 ГГц, скорость передачи данных не более 250 кбит/с. В настоящее время ZigBee Alliance поддерживают более 200 компаний, однако широкого распространения такие устройства не получили.

Wibree

В 2001 г. исследовательский центр компании Nokia начал работу по созданию сверхмаломощной (***ultra-low-power***) версии Bluetooth, дешевой, с очень низким потреблением энергии. В 2004 г. его проект был опубликован как «расширение Bluetooth» — *Bluetooth Low End Extension*. Впоследствии к работе над проектом подключились другие компании и в 2006 новая технология получила торговую марку ***Wibree***. Мощность передатчика Wibree не превышает 0,25 мВт, дальность связи — не более 10 м. Устройства Wibree будут выпускаться двух типов:

- Отдельное устройство — ***Stand-along***. Устройства такого типа будут использоваться для задач не требующих высоких скоростей обмена данными, но они должны работать длительно без замены питающей батареи, например, в датчиках сердечного ритма.
- Двухрежимное устройство — ***Dual-mode***. Обеспечит совместную работу устройств ***Stand-along Wibree*** и устройств ***Bluetooth***, такие устройства будут встраиваться в модули Bluetooth.

Размеры устройств Wibree по заявлениям разработчиков будут сравнимы с пуговицей. Первые устройства должны появиться в конце 2007 г.

UWB

Технология ***UWB*** основана на передаче множества закодированных импульсов негармонической формы очень малой мощ-

ности (0,05 мВт) и малой длительности в широком диапазоне частот (от 3,1 до 10,6 ГГц). Короткие сигналы UWB сравнительно устойчивы к многолучевому затуханию, возникающему при отражении волны от стен, потолка, зданий, транспортных средств. Высокоскоростные UWB-устройства хорошо подходят для работы с видеопотоками и приложениями, требующими быстрой пересылки данных. Стандарт UWB разрабатывается рядом компаний под руководством Intel в рамках спецификаций ***IEEE 802.15.4a.***

3 ХАРАКТЕРИСТИКИ КАНАЛА СВЯЗИ

3.1 Типы характеристик и способы их определения

К основным характеристикам канала передачи данных относятся:

- амплитудно-частотная характеристика;
- полоса пропускания;
- затухание;
- помехоустойчивость;
- перекрестные наводки на линии;
- пропускная способность;
- достоверность передачи данных;
- удельная стоимость.

В первую очередь разработчика вычислительной сети интересуют пропускная способность и достоверность передачи данных, поскольку эти характеристики прямо влияют на производительность и надежность создаваемой сети. Пропускная способность и достоверность — это характеристики, как линии связи, так и способа передачи данных. Поэтому если способ передачи (протокол) уже определен, то известны и эти характеристики.

Например, пропускная способность цифровой линии всегда известна, так как на ней определен *протокол физического уровня* (частота дискретизации и способ кодирования данных), который задает битовую скорость передачи данных — 64 Кбит/с, 2 Мбит/с и т.п.

Однако нельзя говорить о пропускной способности линии связи, до того как для нее определен протокол физического уровня. Именно в таких случаях, когда только предстоит определить, какой из множества существующих протоколов можно использовать на данной линии, очень важными являются остальные характеристики *среды передачи данных*, такие как полоса пропускания, перекрестные наводки, помехоустойчивость и другие характеристики.

Для определения характеристик среды часто используют анализ ее реакций на некоторые эталонные воздействия. Такой подход позволяет достаточно просто и однотипно определять характеристики линий связи любой природы, не прибегая к слож-

ным теоретическим исследованиям. Чаще всего в качестве эталонных сигналов для исследования реакций линий связи используются синусоидальные сигналы различных частот. Это связано с тем, что сигналы этого типа часто встречаются в технике и с их помощью можно представить любую функцию времени — как непрерывный процесс колебаний звука, так и прямоугольные импульсы, генерируемые компьютером.

3.2 Спектральный анализ сигналов на линиях связи

Из теории гармонического анализа известно, что любой периодический процесс можно представить в виде суммы синусоидальных колебаний различных частот и различных амплитуд. Набор всех составляющих синусоид (*гармоник*) называют *спектральным разложением* исходного сигнала (рис. 3.1). Непериодические сигналы можно представить в виде интеграла синусоидальных сигналов с непрерывным спектром частот.

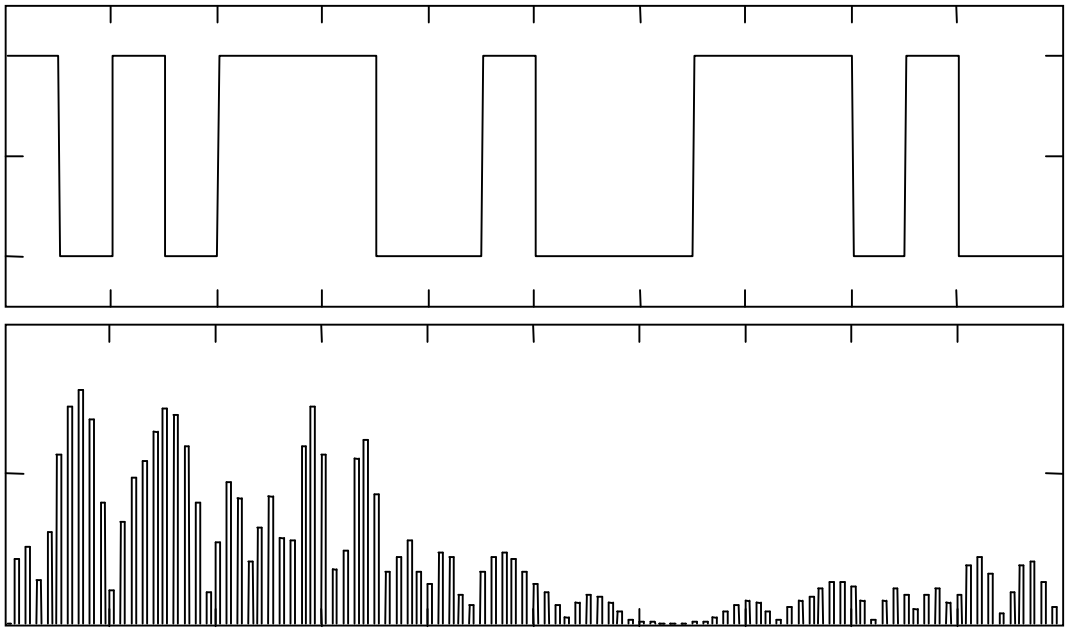


Рис. 3.1 — Цифровой сигнал и его спектр

Техника нахождения спектра любого исходного сигнала хорошо известна. Для некоторых сигналов, которые хорошо описываются аналитически (например, для последовательности прямо-

угольных импульсов одинаковой длительности и амплитуды), спектр легко вычисляется на основании формул Фурье.

Методика построения спектра базируется на разложении исходного сигнала в функциональный ряд, например в ряд Фурье. Коэффициенты ряда Фурье однозначно определяют гармонические составляющие (*гармоники*) исследуемого сигнала. Коэффициенты a_j, b_j ряда Фурье вычисляются по формулам

$$a_j = \frac{1}{T}\int_0^T F(x)\cdot\cos(j\cdot x)\,dx,\ b_j = \frac{1}{T}\int_0^T F(x)\cdot\sin(j\cdot x)\,dx. \qquad (3.1)$$

Так, например, спектр потенциального кода (рис. 3.1) построен как раз по этим формулам. Просуммировав ряд Фурье по всем гармоникам можно восстановить сигнал по его спектральному разложению.

$$S(x) = \frac{a_0}{2} + \sum_{j=1}^{n}\left(a_j\cdot\cos(j\cdot x) + b_j\cdot\sin(j\cdot x)\right). \qquad (3.2)$$

Для сигналов произвольной формы, встречающихся на практике, спектр можно найти с помощью специальных приборов — *спектральных анализаторов*, которые измеряют спектр реального сигнала и отображают амплитуды составляющих гармоник на экране или распечатывают их на принтере.

Искажение передающим каналом синусоиды какой-либо частоты приводит, в конечном счете, к искажению передаваемого сигнала любой формы, особенно если синусоиды различных частот искажаются неодинаково. Вследствие этого на приемном конце линии сигналы могут плохо распознаваться (рис. 3.2).

Линия связи искажает передаваемые сигналы из-за того, что ее физические параметры отличаются от идеальных. Так, например, медные провода всегда представляют собой некоторую распределенную по длине комбинацию активного сопротивления, емкостной и индуктивной нагрузки (рис. 3.3). В результате для синусоид различных частот линия будет обладать различным полным сопротивлением, а значит, и передаваться они будут поразному.

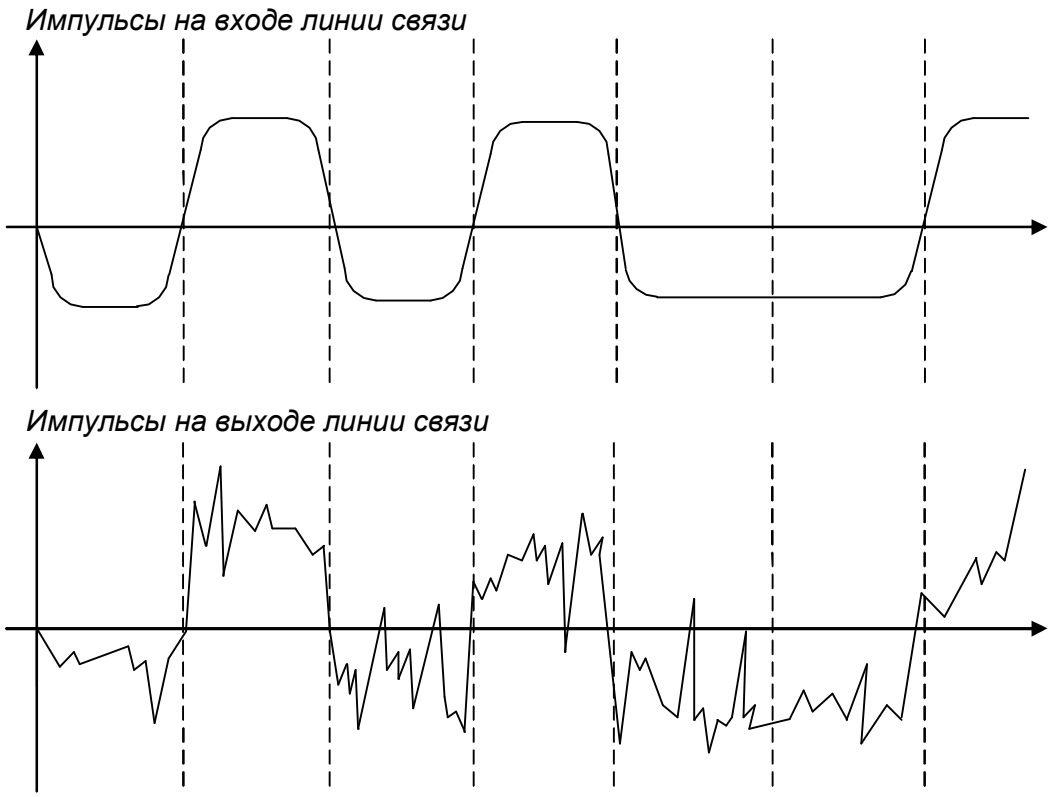


Рис. 3.2 — Искажение импульсов в линии связи

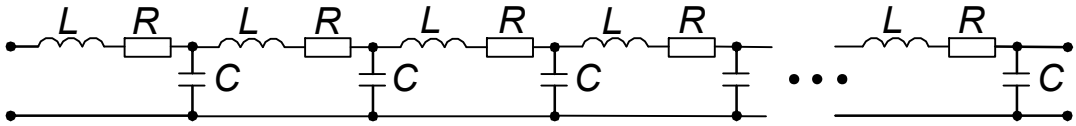


Рис. 3.3 — Представление линии как активно-емкостной нагрузки

Волоконно-оптический кабель также имеет отклонения, мешающие идеальному распространению света. Если линия связи включает промежуточную аппаратуру, то она также может вносить дополнительные искажения, так как невозможно создать устройства, которые бы одинаково хорошо передавали весь спектр синусоид, от нуля до бесконечности.

Кроме искажений сигналов, вносимых внутренними физическими параметрами линии связи, существуют и внешние помехи, которые вносят свой вклад в искажение формы сигналов на выходе линии. Эти помехи создают различные электрические двигатели, электронные устройства, атмосферные явления и т.д. Несмотря на защитные меры, предпринимаемые разработчиками

кабелей и усилительно — коммутирующей аппаратуры, полностью компенсировать влияние внешних помех не удается.

3.3 Амплитудно-частотная характеристика, полоса пропускания и затухание

Степень искажения синусоидальных сигналов линиями связи оценивается с помощью таких характеристик, как амплитудно-частотная характеристика, полоса пропускания и затухание на определенной частоте.

Амплитудно-частотная характеристика показывает, как затухает амплитуда синусоиды на выходе линии связи по сравнению с амплитудой на ее входе для всех возможных частот передаваемого сигнала (рис. 3.4). Вместо амплитуды в этой характеристике часто используют также такой параметр сигнала, как его мощность.

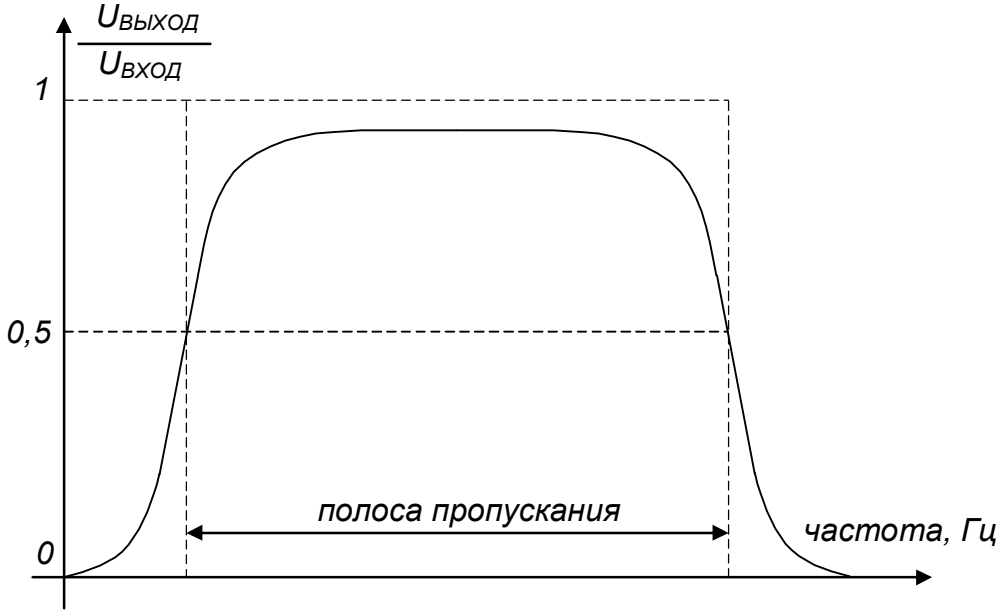


Рис. 3.4 — Амплитудно-частотная характеристика

Знание амплитудно-частотной характеристики реальной линии позволяет определить форму выходного сигнала практически для любого входного сигнала. Для этого необходимо найти спектр исследуемого входного сигнала, преобразовать амплитуду составляющих его гармоник в соответствии с амплитудно-

частотной характеристикой, а затем найти форму выходного сигнала, сложив преобразованные гармоники.

Несмотря на полноту информации, предоставляемой амплитудно-частотной характеристикой о линии связи, ее использование осложняется тем обстоятельством, что получить ее весьма трудно. Ведь для этого нужно провести тестирование линии эталонными синусоидами по всему диапазону частот от нуля до некоторого максимального значения, которое может встретиться во входных сигналах. Причем менять частоту входных синусоид нужно с небольшим шагом, а значит, количество экспериментов должно быть очень большим. Поэтому на практике вместо амплитудно-частотной характеристики применяются другие, упрощенные характеристики — полоса пропускания и затухание.

Полоса пропускания (bandwidth), *(полоса частот)* — это непрерывный диапазон частот, для которого отношение амплитуды выходного сигнала ко входному превышает некоторый заранее заданный предел, обычно *0.5*. То есть полоса пропускания определяет диапазон частот синусоидального сигнала, при которых этот сигнал передается по линии связи без значительных искажений. Знание полосы пропускания позволяет получить с некоторой степенью приближения тот же результат, что и знание амплитудно-частотной характеристики. Как мы увидим ниже, ширина полосы пропускания в наибольшей степени влияет на максимально возможную *скорость* передачи информации по линии связи.

Полоса пропускания связана со скоростью передачи информации. Различают ***бодовую*** (модуляционную) и ***информационную*** скорости. Бодовая скорость измеряется в бодах и определяется *числом изменений дискретного сигнала в единицу времени*, а информационная — *числом битов информации, переданных в единицу времени*. Именно бодовая скорость определяется полосой пропускания линии.

Если на бодовом интервале (между соседними изменениями сигнала) передается N бит, то число градаций несущей равно 2^N. Например, при числе градаций $16 = 2^4$ и скорости 1200 бод одному боду соответствует 4 бит/с и информационная скорость составит 4800 бит/с.

Затухание (attenuation) определяется как относительное уменьшение амплитуды или мощности сигнала при передаче по линии сигнала определенной частоты. Таким образом, затухание представляет собой одну точку из амплитудно-частотной характеристики линии. Часто при эксплуатации линии заранее известна основная частота передаваемого сигнала, то есть та частота, гармоника которой имеет наибольшую амплитуду и мощность. Поэтому достаточно знать затухание на этой частоте, чтобы приблизительно оценить искажения передаваемых по линии сигналов. Более точные оценки возможны при знании затухания на нескольких частотах, соответствующих нескольким основным гармоникам передаваемого сигнала.

Затухание *A* обычно измеряется в децибелах (*дБ*, *decibel — dB*) и вычисляется по следующей формуле:

$$A = 10 \cdot \log_{10} \frac{P_{ВЫХ}}{P_{ВХ}}, \qquad (3.3)$$

где $P_{ВЫХ}$ — мощность сигнала на выходе линии, $P_{ВХ}$ — мощность сигнала на входе линии. Так как мощность выходного сигнала кабеля без промежуточных усилителей всегда меньше, чем мощность входного сигнала, затухание кабеля всегда является отрицательной величиной.

Например, кабель на витой паре категории UTP5 характеризуется затуханием не ниже *–23.6 дБ* для частоты *100 МГц* при длине кабеля *100 м*. Частота *100 МГц* выбрана потому, что кабель этой категории предназначен для высокоскоростной передачи данных, сигналы которых имеют значимые гармоники с частотой примерно *100 МГц*. Кабель категории UTP3 предназначен для низкоскоростной передачи данных, поэтому для него определяется затухание на частоте *10 МГц* (не ниже *–11.5 дБ*). Часто оперируют с абсолютными значениями затухания, без указания знака.

Абсолютный ***уровень мощности*** (например, уровень мощности передатчика) также измеряется в децибелах. При этом в качестве базового значения мощности сигнала, относительно которого измеряется текущая мощность, принимается значение в 1 мВт. Таким образом, уровень мощности *p* вычисляется по следующей формуле:

$$p = 10 \cdot \log_{10} \frac{P}{1\ \text{мВт}}\ [\text{дБм}], \qquad (3.4)$$

где p — мощность сигнала в милливаттах, а *дБм* — это единица измерения уровня мощности (*децибел на* 1 *мВт*).

Таким образом, амплитудно-частотная характеристика, полоса пропускания и затухание являются универсальными характеристиками, и их знание позволяет сделать вывод о том, как через линию связи будут передаваться сигналы любой формы. Полоса пропускания зависит от типа линии связи и ее протяженности.

3.4 Пропускная способность

Пропускная способность (throughput) линии характеризует максимально возможную скорость передачи данных по линии связи. Пропускная способность канала связи зависит не только от его характеристик, таких как амплитудно-частотная характеристика, но и от спектра передаваемых сигналов. Если *значимые* гармоники сигнала (то есть те гармоники, амплитуды которых вносят основной вклад в результирующий сигнал) попадают в полосу пропускания линии, то такой сигнал будет хорошо передаваться данной линией связи и приемник сможет правильно распознать информацию, отправленную по линии передатчиком (рис. 3.5, *а*). Если же значимые гармоники выходят за границы полосы пропускания линии связи, то сигнал будет значительно искажаться, приемник будет ошибаться при распознавании информации, а значит, информация не сможет передаваться с заданной пропускной способностью (рис. 3.5, *б*).

Пропускная способность линий связи коммуникационного и сетевого оборудования измеряется в битах в секунду — [бит/с], а не в байтах в секунду, это связано с тем, что данные в линии передаются последовательно, то есть побитно, а не параллельно, байтами. Пропускная способность измеряется также в пакетах в секунду.

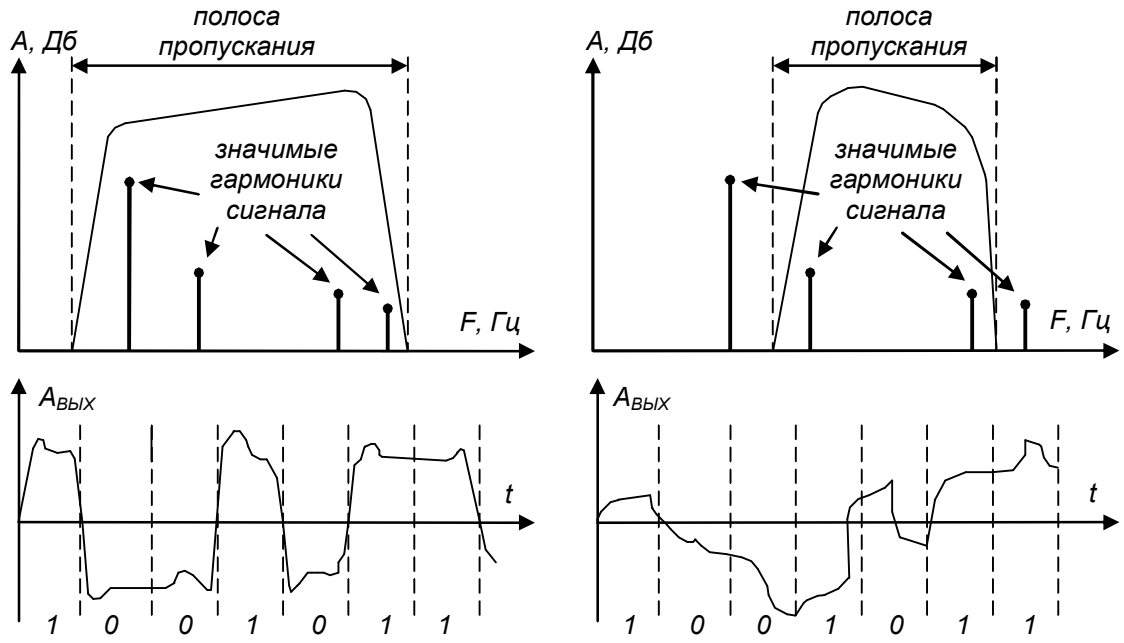


Рис. 3.5 — Полоса пропускания и спектр сигнала

Средняя пропускная способность вычисляется путем деления общего объема переданных данных на время их передачи, причем выбирается достаточно длительный промежуток времени — час, день или неделя. Средняя пропускная способность отдельного элемента или всей сети позволяет оценить работу сети на большом промежутке времени, в течение которого в силу закона больших чисел пики и спады интенсивности трафика компенсируют друг друга

Мгновенная пропускная способность отличается от средней тем, что для усреднения выбирается очень маленький промежуток времени — например, 10 мс или 1 с.

Максимальная пропускная способность — это наибольшая мгновенная пропускная способность, зафиксированная в течение периода наблюдения. Максимальная пропускная способность позволяет оценить возможности сети справляться с пиковыми нагрузками.

3.5 Физическое и логическое кодирование

Выбор способа представления дискретной информации в виде сигналов, подаваемых на линию связи, называется ***физическим*** или ***линейным кодированием***. От выбранного способа ко-

дирования зависит спектр сигналов и, соответственно, пропускная способность линии. Таким образом, для одного способа кодирования линия может обладать одной пропускной способностью, а для другого — другой. Например, витая пара категории UTP3 может передавать данные с пропускной способностью 10 Мбит/с при способе кодирования стандарта физического уровня 10Base-T и 33 Мбит/с при способе кодирования стандарта 100Base-T4. В примере, приведенном на рис. 3.6 а), принят следующий способ кодирования — *логическая единица* представлена на линии положительным потенциалом, а *логический ноль* — отрицательным (такой способ кодирования называется *потенциальным кодом*).

Если сигнал изменяется так, что можно различить только два его состояния, то любое его изменение будет соответствовать наименьшей единице информации — биту. Если же сигнал может иметь более двух различимых состояний, то любое его изменение будет нести несколько бит информации. Количество изменений информационного параметра несущего периодического сигнала в секунду измеряется в ***бодах*** (***baud***). Период времени между соседними изменениями информационного сигнала называется ***тактом*** работы передатчика.

Пропускная способность линии в битах в секунду в общем случае не совпадает с числом бод. Она может быть как выше, так и ниже числа бод, и это соотношение зависит от способа кодирования.

Если сигнал имеет более двух различимых состояний, то пропускная способность в битах в секунду будет выше, чем число бод. Например, если информационными параметрами *несущей* являются фаза и амплитуда синусоиды, причем различаются 4 состояния фазы в 0, 90, 180 и 270 градусов и два значения амплитуды сигнала, то информационный сигнал может иметь 8 различимых состояний, кодируемых тремя битами ($8 = 2^3$). В этом случае модем, работающий со скоростью 2400 бод (с тактовой частотой 2400 Гц) передает информацию со скоростью 7200 бит/с, так как при одном изменении сигнала передается 3 бита информации.

При использовании сигналов с двумя различимыми состояниями может наблюдаться обратная картина. Это часто происхо-

дит потому, что для надежного распознавания приемником пользовательской информации каждый бит в последовательности кодируется с помощью нескольких изменений информационного параметра несущего сигнала. Например, при кодировании единичного значения бита импульсом положительной полярности, а нулевого значения бита — импульсом отрицательной полярности физический сигнал дважды изменяет свое состояние при передаче каждого бита. При таком кодировании пропускная способность линии в два раза ниже, чем число бод, передаваемое по линии.

На пропускную способность линии оказывает влияние не только физическое, но и логическое кодирование. ***Логическое кодирование*** выполняется до физического кодирования и подразумевает замену бит исходной информации новой последовательностью бит, несущей ту же информацию, но обладающей, кроме этого, дополнительными свойствами, например возможностью для приемной стороны обнаруживать ошибки в принятых данных. Сопровождение каждого байта исходной информации одним битом четности — это пример примитивного, но очень часто применяемого способа логического кодирования при передаче данных с помощью модемов. Другим примером логического кодирования может служить шифрация данных, обеспечивающая их конфиденциальность при передаче через общественные каналы связи. При логическом кодировании чаще всего исходная последовательность бит заменяется более длинной последовательностью, поэтому пропускная способность канала по отношению к полезной информации при этом уменьшается.

3.6 Связь между пропускной способностью линии и полосой пропускания

Чем выше частота несущего периодического сигнала, тем больше информации в единицу времени передается по линии и тем выше пропускная способность линии при фиксированном способе физического кодирования. С другой стороны, с увеличением частоты периодического несущего сигнала увеличивается и ширина спектра этого сигнала. Здесь ***ширина спектра*** — это разность между максимальной и минимальной частотами набора значимых синусоид, кодирующих последовательность сигналов.

Линия передает этот спектр синусоид с теми искажениями, которые определяются ее полосой пропускания. Чем больше несоответствие между полосой пропускания линии и шириной спектра передаваемых информационных сигналов, тем больше сигналы искажаются и тем вероятнее ошибки в распознавании информации принимающей стороной, а значит, скорость передачи информации на самом деле оказывается меньше, чем можно было предположить.

Связь между полосой пропускания линии и ***максимально возможной пропускной способностью***, вне зависимости от принятого способа физического кодирования, установил *Клод Шеннон*, она выражается так:

$$C = F \cdot \log_2\left(1 + \frac{P_C}{P_{Ш}}\right), \qquad (3.5)$$

где C — максимальная пропускная способность линии в битах в секунду, F — ширина полосы пропускания линии в герцах, P_C — мощность сигнала, $P_{Ш}$ — мощность шума.

Из этого соотношения видно, что хотя теоретического предела пропускной способности линии с фиксированной полосой пропускания не существует, на практике такой предел имеется. Действительно, повысить пропускную способность линии можно за счет увеличения мощности передатчика или же уменьшения мощности шума (помех) на линии связи. Обе эти составляющие поддаются изменению с большим трудом. Повышение мощности передатчика ведет к значительному увеличению его габаритов и стоимости. Снижение уровня шума требует применения специальных кабелей с хорошими защитными экранами, что весьма дорого, а также снижения шума в передатчике и промежуточной аппаратуре, чего достичь весьма не просто. К тому же влияние мощностей полезного сигнала и шума на пропускную способность ограничено логарифмической зависимостью, которая растет далеко не так быстро, как прямо пропорциональная. Так, при достаточно типичном исходном отношении мощности сигнала к мощности шума в 100 раз повышение мощности передатчика в 2 раза даст только 15 % увеличения пропускной способности линии.

Близким по своей сути к формуле (3.5) Шеннона является следующее соотношение, полученное *Найквистом*, которое также определяет максимально возможную пропускную способность линии связи, но без учета шума на линии:

$$C = 2 \cdot F \cdot \log_2 M, \qquad (3.6)$$

где M — количество различимых состояний информационного параметра.

Если сигнал имеет 2 различимых состояния, то пропускная способность равна удвоенному значению ширины полосы пропускания линии связи (рис. 3.6, *а*). Если же передатчик использует более чем 2 устойчивых состояния сигнала для кодирования данных, то пропускная способность линии повышается, так как за один такт работы передатчик передает несколько бит исходных данных, например 2 бита при наличии четырех различимых состояний сигнала (рис. 3.6, *б*).

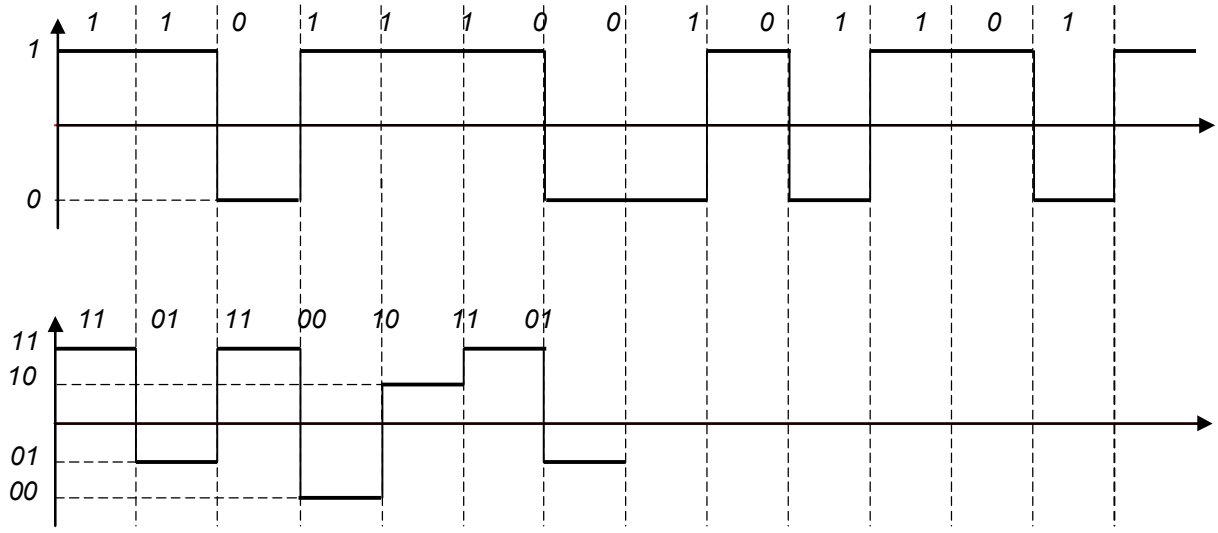


Рис. 3.6 — Повышение скорости передачи за счет дополнительных состояний

Хотя формула Найквиста явно не учитывает наличие шума, косвенно его влияние отражается в выборе количества состояний информационного сигнала, Для повышения пропускной способности канала хотелось бы увеличить это количество до значительных величин, но на практике мы не можем этого сделать из-за шума на линии. Например, для примера, приведенного на

рис. 3.6, можно увеличить пропускную способность линии еще в два раза, используя для кодирования данных не 4, а 16 уровней. Однако если амплитуда шума часто превышает разницу между соседними 16-ю уровнями, то приемник не сможет устойчиво распознавать передаваемые данные. Поэтому количество возможных состояний сигнала фактически ограничивается соотношением мощности сигнала и шума, а формула Найквиста определяет предельную скорость передачи данных в том случае, когда количество состояний уже выбрано с учетом возможностей устойчивого распознавания приемником.

Приведенные соотношения дают предельное значение пропускной способности линии, а степень приближения к этому пределу зависит от конкретных методов физического кодирования, рассматриваемых ниже.

3.7 Помехоустойчивость и достоверность

Помехоустойчивость линии определяет ее способность уменьшать уровень помех, создаваемых во внешней среде, на внутренних проводниках. Помехоустойчивость линии зависит от типа используемой физической среды, а также от экранирующих и подавляющих помехи средств самой линии. Наименее помехоустойчивыми являются радиолинии, хорошей устойчивостью обладают кабельные линии и отличной — волоконно-оптические линии, малочувствительные к внешнему электромагнитному излучению. Обычно для уменьшения помех, появляющихся из-за внешних электромагнитных полей, проводники экранируют и/или скручивают.

Перекрестные наводки на ближнем конце (***Near End Cross Talk — NEXT***) определяют помехоустойчивость кабеля к внутренним источникам помех, когда электромагнитное поле сигнала, передаваемого выходом передатчика по одной паре проводников, наводит на другую пару проводников сигнал помехи. Если ко второй паре будет подключен приемник, то он может принять наведенную внутреннюю помеху за полезный сигнал. Показатель *NEXT*, выраженный в децибелах, равен

$$NEXT = 10 \cdot \log_{10}\left(\frac{P_{ВЫХ}}{P_{НАВ}}\right), \qquad (3.7)$$

где $P_{ВЫХ}$ — мощность выходного сигнала, $P_{НАВ}$ — мощность наведенного сигнала.

Чем меньше значение *NEXT*, тем лучше кабель. Так, для витой пары категории UTP5 показатель *NEXT* должен быть меньше *–27 дБ* на частоте *100 МГц*.

Показатель *NEXT* обычно используется применительно к кабелю, состоящему из нескольких витых пар, так как в этом случае взаимные наводки одной пары на другую могут достигать значительных величин. Для одинарного коаксиального кабеля (то есть состоящего из одной экранированной жилы) этот показатель не имеет смысла, а для двойного коаксиального кабеля он также не применяется вследствие высокой степени защищенности каждой жилы. Оптические волокна также не создают сколько-нибудь заметных помех друг для друга.

В связи с тем, что в некоторых новых технологиях используется передача данных одновременно по нескольким витым парам, в последнее время стал применяться показатель ***PowerSUM***, являющийся модификацией показателя *NEXT*. Этот показатель отражает суммарную мощность перекрестных наводок от всех передающих пар в кабеле.

Достоверность передачи данных характеризует вероятность искажения для каждого передаваемого бита данных, этот показатель называют так же интенсивностью битовых ошибок (***Bit Error Rate***, ***BER***).

$$BER = \frac{N_{ИСК}}{N}, \qquad (3.8)$$

где N — количество переданной информации, а $N_{ИСК}$ — количество искаженных при передаче бит.

Величина BER для каналов связи без дополнительных средств защиты от ошибок (например, самокорректирующихся кодов или протоколов с повторной передачей искаженных кадров) составляет, как правило, 10^{-4}—10^{-6}, в оптоволоконных линиях связи — 10^{-9}. Значение достоверности передачи данных, например, в 10^{-4} говорит о том, что в среднем из 10 000 бит искажается значение одного бита.

Искажения бит происходят как из-за наличия помех на линии, так и по причине искажений формы сигнала ограниченной полосой пропускания линии. Поэтому для повышения достоверности передаваемых данных нужно повышать степень помехозащищенности линии, снижать уровень перекрестных наводок в кабеле, а также использовать более широкополосные линии связи.

3.8 Характеристики кабельных линий

Для построения линий связи в вычислительных сетях в настоящее время используются огромное количество различных видов кабелей. Тем не менее, в проекты стандартных вычислительных сетей закладываются на сегодня всего три вида кабелей:

- коаксиальный кабель (двух типов):
  - тонкий коаксиальный кабель (***thin coaxial cable***);
  - толстый коаксиальный кабель (***thick coaxial cable***);
- витая пара (двух основных типов):
  - неэкранированная витая пара (***unshielded twisted pair — UTP***);
  - экранированная витая пара (***shielded twisted pair — STP***);
- волоконно-оптический кабель (двух типов):
  - многомодовый кабель (***fiber optic cable multimode***);
  - одномодовый кабель (***fiber optic cable single mode***).

Наиболее популярным видом среды передачи данных на небольшие расстояния (до 100 м) становится неэкранированная витая пара, которая включена практически во все современные стандарты и технологии локальных сетей и обеспечивает пропускную способность до 100 Мбит/с (на кабелях категории *UTP5*).

При проектировании и монтаже вычислительных систем, как указывалось выше, в качестве стандартных систем передачи данных можно использовать довольно ограниченную номенклатуру кабелей: кабель с витыми парами (*UTP-кабель*) категорий 3, 4 или 5 с различными типами экранов или без них (STP — экранирование медной оплеткой, FTP — экранирование фольгой, SFTP — экранирование медной оплеткой и фольгой), тонкий коаксиальный кабель (RG-58) с разным исполнением центральной жилы (RG-58/U — сплошная медная жила, RG-58A/U — много-

жильный, RG-58C/U — специальное /военное/ исполнение кабеля RG-58A/U), толстый коаксиальный кабель и волоконно-оптический кабель. При этом каждый вид кабельной подсистемы накладывает те или иные ограничения на проект сети.

Таблица 3.1 — **Ограничения на максимальную длину сегмента**

100 м	у кабеля с витыми парами
185 м	у тонкого коаксиального кабеля
500 м	у толстого коаксиального кабеля
1000 м	у многомодового (mm) оптоволоконного кабеля
2000 м	у одномодового (sm) оптоволоконного кабеля (с применением специальных средств до 40—70—90 км)

Таблица 3.2 — **Максимальное количество узлов на сегменте**

2	у кабеля с витыми парами
30	у тонкого коаксиального кабеля
100	у толстого коаксиального кабеля
2	у оптоволоконного кабеля

Ниже более подробно рассматриваются различные типы кабелей, их технические характеристики и применение.

Характеристики коаксиального кабеля

Конструктивно коаксиальный кабель состоит из одножильного или многожильного проводника, окруженного диэлектрическим материалом, как правило, плотным или мягким пенополимером. Диэлектрик помещается в непрерывный алюминиевый экран, ламинированный полистером, а затем в луженую медную сетку. Вся конструкция помещается в оболочку из поливинилхлоридного или огнеупорного полимерного материала (рис. 3.7).

Рис. 3.7 — Коаксиальный кабель

Для коаксиального кабеля качество передачи сигнала определяется четырьмя электрическими параметрами, относящимися

к материалу диэлектрика и геометрическим размерам кабеля — импедансом, затуханием, емкостью и временной задержкой распространения сигнала или скоростью его распространения в передающей среде.

Импеданс (или характеристический импеданс) — сопротивление [0м] волновой передающей среды переменному электрическому току. Величина импеданса напрямую зависит от отношения размеров внутреннего и внешнего проводников и связана обратной зависимостью с диэлектрической постоянной кабеля. В отличие от сопротивления проводника импеданс не изменяется при изменении длины кабеля. Для того, чтобы система могла работать с максимальной эффективностью, номинальные импедансы передатчика, приемника и кабеля должны очень точно совпадать. Значения импеданса для кабелей определяют электрические требования к коммутационному оборудованию.

Затухание (*attenuation*) — потери или уменьшение уровня сигнала при прохождении его по передающей среде. Существует два типа потерь, определяющих величину затухания сигнала собственные потери в проводниках (центральном проводнике и экране) и диэлектрические потери. Оба типа потерь растут с увеличением частоты.

Емкость — отношение величины электрического заряда двух проводников к разнице потенциалов между ними или, говоря другими словами, — энергия, накапливаемая кабелем. Емкость измеряется в пФ на единицу длины. Как и импеданс, емкость коаксиального кабеля зависит от размеров внутреннего и внешнего проводников и диэлектрической константы диэлектрического материала. Емкость и импеданс обратно пропорциональны друг другу.

Время задержки распространения сигнала по длине кабеля прямо пропорционально квадратному корню диэлектрической константы. В вакууме электромагнитные волны распространяются со скоростью света. В кабеле волна распространяется несколько медленнее — со скоростью, обратно пропорциональной диэлектрической константе кабеля. Чем меньше диэлектрическая константа, тем ближе скорость распространения сигнала к скорости света. Более низким значениям диэлектрической константы соответствуют более высокие скорости передачи.

Фазовая задержка обусловлена тем, что более высокочастотные сигналы распространяются в передающей среде быстрее по сравнению с низкочастотными. В широкополосной сети информация обычно передается в виде цифрового кода, в котором низкочастотный тон определенной длительности представляет двоичную **1**, а высокочастотный тон представляет **0**. Вследствие того, что низкочастотные сигналы распространяются медленнее, они обладают тенденцией к отставанию от более быстрых высокочастотных сигналов и приходят к концу линии с фазовым сдвигом.

Импеданс и обратные потери

Импеданс (***impedance***) характеризует путь прохождения данных. Например, если сигнал передается с импедансом 50 Ом, то и структурированная проводка должна соответствовать импедансу 50 Ом. Любое отклонение от этой величины приведет к тому, что часть сигнала отразится назад к источнику данных. Отражение означает, что вместо того, чтобы продолжать свой путь дальше вперед, в действительности энергия отражается назад к передатчику; в конечном итоге это приводит к ослаблению распространяющегося в прямом направлении сигнала.

При распространении электрического сигнала по проводнику ток и напряжение находятся всегда в одном отношении друг к другу, это отношение называется импедансом и может быть выражено формулой:

$$Z_L = \sqrt{\frac{R' + j \cdot \omega \cdot L'}{G' + j \cdot \omega \cdot C'}}, \qquad (3.9)$$

в более упрощенной формуле, не учитывающей потерь в кабеле ($R' = 0,\ G' = 0$), волновое сопротивление может быть записано в виде:

$$Z_L = \sqrt{\frac{L'}{C'}}. \qquad (3.10)$$

Действительная величина Z_L остаётся независимой от частоты. Если частота становится очень высокой, то $\omega \cdot L >> R$ и $\omega \cdot C >> G$. Следовательно, импеданс является также постоянным для высоких частот.

Для величины импеданса справедливы следующие аппроксимации: — длинный тонкий проводник: L велико, C мало, Z_L велико — короткий толстый проводник: L мало, C велико, Z_L мало.

Изменение импеданса может быть вызвано множеством причин. Одна из них — несоблюдение технологии в процессе изготовления: любое отклонение от предусмотренного расстояния между проводниками или нарушение свойств изолирующего материала способно привести к изменению импеданса.

Другая распространенная причина — несоответствие компонентов. Например, несоответствие имеет место, когда шнур переключений с одним импедансом присоединяется к горизонтальной проводке с другим импедансом.

Такое несовпадение неизбежно вызовет отражение энергии в точке разрыва. Если импеданс обусловливает возможность несоответствия, то обратные потери характеризуют его последствия. Обратные потери (измеряемые в дБ) позволяют выяснить, какая доля сигнала теряется вследствие отражения.

Коэффициент отражения

Если участок кабеля имеет сопротивление Z_X, которое отличается от волнового сопротивления кабеля, например, на конце кабеля, часть энергии импульса будет отражаться на этом участке кабеля. Отношение между отражённой частью и частью импульса, распространяющегося за неоднородность, может быть описано посредством ***коэффициента отражения r*** (Z_L = волновое сопротивление кабеля, Z_X = сопротивление на источнике дефекта):

$$r = \frac{Z_X - Z_L}{Z_X - Z_L}. \qquad (3.11)$$

Обычно коэффициент отражения является комплексным параметром, т.е. отражение меняет не только амплитуду импульса, но и форму. Если коэффициент отражения чисто действительный, имеет место только изменение амплитуды импульса с сохранением его формы.

Относительно коэффициента отражения различают три простых специальных случая: ***согласование*** (если кабель нагружен на волновое сопротивление), ***разомкнутый конец*** и ***короткое***

замыкание. Можно легко рассчитать, какой вид отражения можно ожидать на определённом ответвлении.

Чтобы определить коэффициент отражения, необходимо знать сопротивление в точке x текущего положения импульса и сопротивление в точке $x + dx$. На однородном участке кабеля, сопротивление Z_L в точке x, равно сопротивлению Z_X в точке $x + dx$, тогда:

$$r = \frac{Z_X - Z_L}{Z_X - Z_L} = \frac{Z_L - Z_L}{Z_L - Z_L} = \frac{0}{2 \cdot Z_L} = 0. \qquad (3.12)$$

Как правило, нельзя считать участки кабеля, свободными от всех отражений. Производство, хранение, и прокладка кабеля всегда приводит к небольшим вариациям его характеристик. Эти вариации вызывают небольшие отражения.

- Если импульс не отражается от конца кабеля, значит, цепь нагружена согласованно. Если сопротивление нагрузки имеет величину, равную волновому сопротивлению, то импульс не отражается: $Z_X = Z_L$, следовательно, коэффициент отражения $r = 0$.

Хорошо согласованный с терминатором кабель полностью поглощает сигнал отражения, что служит гарантией правильности выбора терминатора (рис. 3.8).

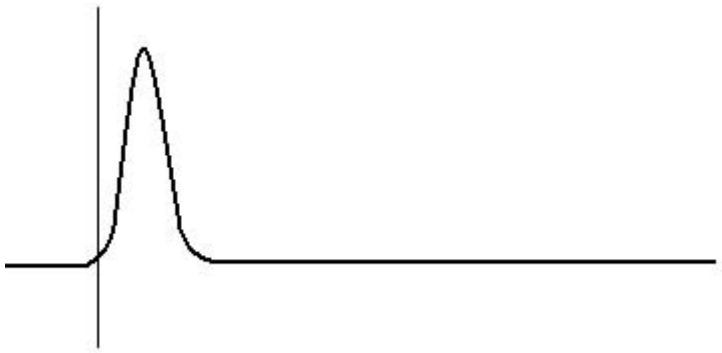


Рис. 3.8 — Рефлектограмма правильно согласованного кабеля

- Если же наоборот, кабель на конце разомкнут или оборван, импульс будет отражаться с полной амплитудой и без изменения формы. Отклик (эхо), вызываемое дефектом, имеет ту же полярность, что и передаваемый импульс.

$$r = \frac{Z_X - Z_L}{Z_X - Z_L} = \frac{\infty - Z_L}{\infty + Z_L} = \frac{\infty}{\infty} = 1. \qquad (3.13)$$

Если конец разомкнут, то импульс отражается, следовательно, Z_X бесконечно велико, а коэффициент отражения $r = 1$.

На рефлектограмме (рис. 3.9) представлен случай отражения сигнала от точки большого сопротивления (второй курсор), что соответствует обрыву кабеля. Состояние, описываемое рефлектограммой, получило название характерного обрыва (COMPLIT OPEN).

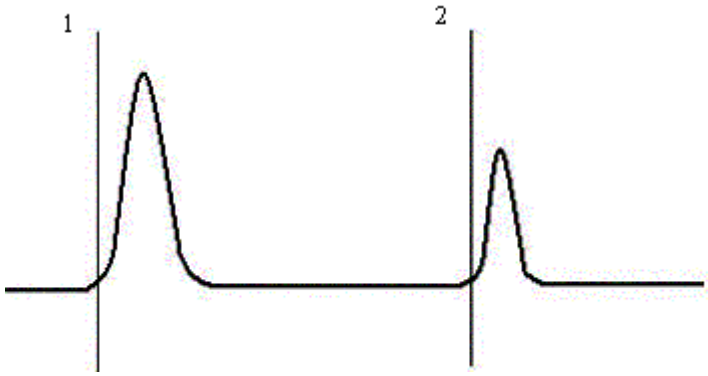


Рис. 3.9 — Рефлектограмма соответствующая обрыву кабеля

- Если кабель замкнут на конце накоротко, импульс будет отражаться с полной амплитудой и неизменённой формой, но с противоположной полярностью.

$$r = \frac{Z_X - Z_L}{Z_X - Z_L} = \frac{0 - Z_L}{0 + Z_L} = \frac{-Z_L}{Z_L} = -1. \qquad (3.14)$$

При коротком замыкании $Z_X = 0$, а $r = -1$.

Отражение со сменой полярности сигнала соответствует короткому замыканию в кабеле (рис. 3.10), малому сопротивлению неоднородности. Такое состояние получило название характерного короткого замыкания (DEAD SHORT).

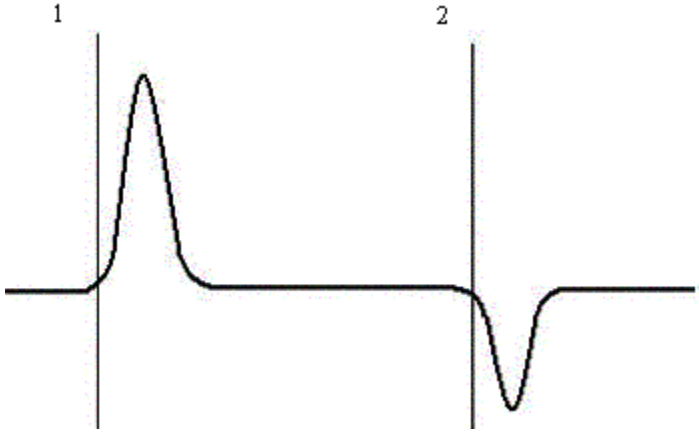


Рис. 3.10 — Рефлектограмма соответствующая короткому замыканию

Коннектор коаксиальной линии

Коаксиальная линия представляет собой среду передачи данных с топологией *общая шина*. Для подключения устройств к коаксиальному кабелю используются специальные Т-коннекторы (***connector***). Включение коннектора в линию должно быть, естественно хорошо сбалансировано по импедансу, чтобы не вызвать дополнительных отражений.

Если импульс распространяется в ответвление кабеля, результирующее сопротивление Z_X равно параллельному соединению входных сопротивлений Z_L двух кабелей.

$$Z_X = \frac{Z_L \cdot Z_L}{Z_L + Z_L} = \frac{Z_L^{\,2}}{2 \cdot Z_L} = \frac{1}{2} Z_L, \qquad (3.15)$$

таким образом, сопротивление, которое встречает импульс в месте разветвления, равно половине волнового сопротивления кабеля. Коэффициент отражения может быть рассчитан по формуле:

$$r = \frac{Z_X - Z_L}{Z_X - Z_L} = \frac{\frac{1}{2} Z_L - Z_L}{\frac{1}{2} Z_L + Z_L} = \frac{-\frac{1}{2} Z_L}{\frac{3}{2} Z_L} = -\frac{1}{3}. \qquad (3.16)$$

Идеальное ответвление не меняет формы импульса при его отражении. При этом отражается 33 % амплитуды импульса.

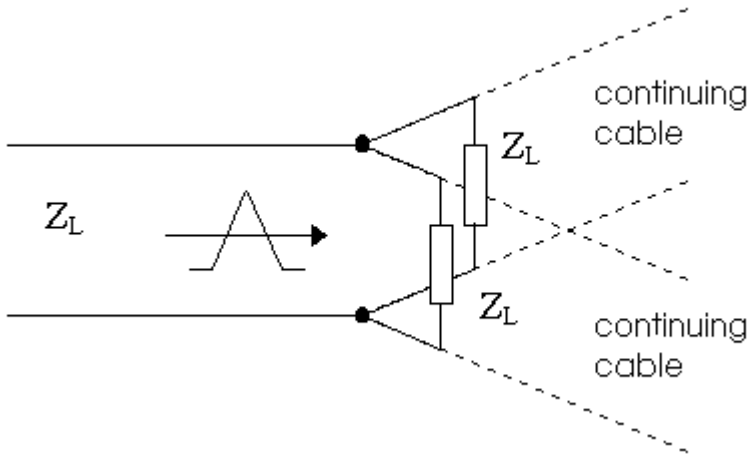


Рис 3.11 — Импульс, отражённый от ответвления

Слабое отрицательное отражение можно наблюдать при любом виде параллельного соединения. Если сопротивление соединяется параллельно с волновым сопротивлением кабеля, резуль-

тирующая величина будет всегда меньше волнового сопротивления, даже если параллельное сопротивление очень велико.

Открытый Т-коннектор внутри коаксиальной линии, обычно очень трудно идентифицировать. Если одно гнездо Т-коннектора открыто, то его можно рассматривать как параллельный открытый конец (см. рис. 3.13). Параллельный открытый конец не вызывает отражения.

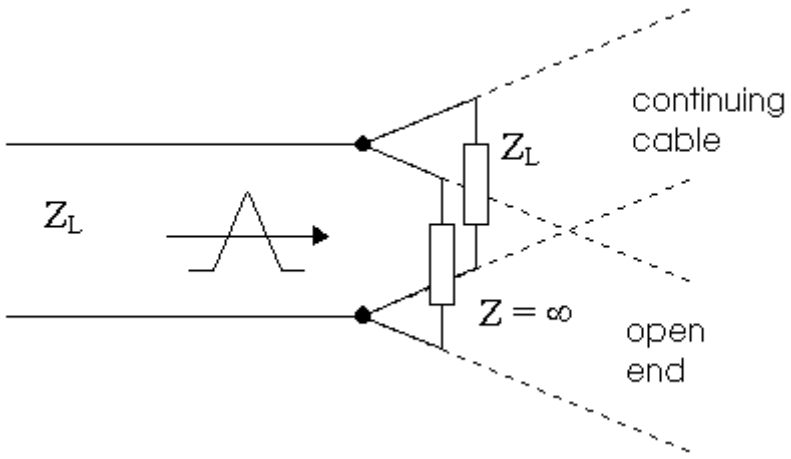


Рис 3.12 — Импульс, отражённый от открытого Т-коннектора

Величина сопротивления параллельного соединения в этом случае не влияет на результирующие сопротивление:

$$Z_X = \frac{\infty \cdot Z_L}{\infty + Z_L} \approx \frac{\infty \cdot Z_L}{\infty} = Z_L. \qquad (3.17)$$

Если импульс попадает на открытый Т-коннектор, то сопротивление, которое он встречает, близко к волновому сопротивлению кабеля. Идеальный Т-коннектор не отражает импульсов. Однако, как правило, на Т-коннекторе нарушается симметрия кабеля. Экран размыкается; и открытое гнездо действует подобно параллельному присоединению небольшой ёмкости.

Разъёмное соединение обычно может считаться действительным последовательным сопротивлением.

$$Z_X = \frac{Z_L \cdot Z_L}{Z_L + Z_L} \approx \frac{Z_L^{\,2}}{2 \cdot Z_L} = \frac{1}{2} Z_L. \qquad (3.18)$$

При известном сопротивлении контакта R коэффициент отражения может быть рассчитан по формуле:

$$Z_X = R + Z_L + R = Z_L + 2 \cdot R;$$

$$r = \frac{Z_X - Z_L}{Z_X + Z_L} \approx \frac{Z_L + 2 \cdot R - Z_L}{Z_L + 2 \cdot R + Z_L} = \frac{R}{Z_L + R}. \qquad (3.19)$$

Как правило, R очень мало по сравнению с Z_L. Т.о., соединитель вызывает очень небольшое (но всегда положительное) отражение.

Скорость сигнала и время распространения импульса

Импульс распространяется в кабеле со *скоростью сигнала* ***v***, которая является характеристикой кабеля. Эта скорость может быть примерно описана через относительную диэлектрическую проницаемость материала изоляции и вычислена по формуле:

$$\nu \approx \frac{1}{\sqrt{\varepsilon_L}} \cdot c, \qquad (3.20)$$

в этой формуле c обозначает скорость света в вакууме (c = *= 2,9979 м/мкс*).

В качестве величины диэлектрической проницаемости воздуха можно принять 1, для пластмасс значение диэлектрической проницаемости лежит в пределах от 2 до 4, а для воды — 80.

Обратная величина корня из диэлектрической проницаемости, которая входит в аппроксимационную формулу, может также использоваться для определения скорости сигнала. Эта величина описывает отношения скорости сигнала в кабеле к скорости света в вакууме. Это отношение обозначено как *коэффициент укорочения* ***g***. Для кабелей с пластмассовой изоляцией значение g лежит между 0.5 и 0.8.

Таблица 3.3 — **Типовые значения коэффициентов укорочения для некоторых типов кабеля**

Кабель РК-50-2-11	1.52
Кабель РК-100-7-1	1.20
Воздушная линия	1.00
Кабель П-270	3.00
Кабель П-274М	1.39
Кабель СБ. АБ	1.84

Полная длина пути зависит от затухания и дисперсии кабеля. Сигналы в начале линии могут быть зарегистрированы как

функция времени. Время t_X, необходимое импульсу для распространения от начала кабеля до неисправности и обратно, может быть измерено. Используя известную скорость сигнала (или значение ***g***) в кабеле, можно вычислить расстояние l_X между началом кабеля и неисправностью:

$$l_X = \frac{\nu}{2} \cdot t_X . \qquad (3.21)$$

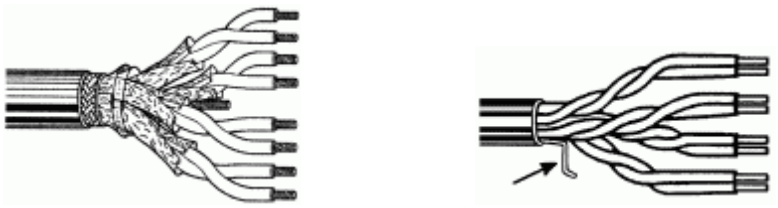
Из соображений практичности вместе с ***g*** часто используется величина **ν/2**, связанная с коэффициентом укорочения отношением:

$$\frac{\nu}{2} = \frac{c}{2 \cdot g} . \qquad (3.22)$$

Если эта величина известна, её можно непосредственно подставить в предшествующую формулу.

Характеристики витой пары

Наибольшей популярностью UTP пользуется в качестве горизонтальной проводки, а именно для подключения настольных систем к телекоммуникационным шкафам. Как следует из названия, UTP состоит из нескольких неэкранированных витых пар, окруженных общей оболочкой.



а *б*

Рис. 3.13 — Экранированная (*а*), неэкранированная (*б*) витая пара

Несмотря на наличие двух- и 25-парных кабелей, наибольшей популярностью пользуется четырехпарная проводка. Хотя в большинстве локально-сетевых сред, таких, как *10/100BaseTX*, используется только две из четырех пар, новые рассматриваемые протоколы, в частности *Gigabit Ethernet*, будут задействовать все четыре пары.

В таблице 3.4 рассмотрены основные стандарты передачи данных технологии типа Ethernet, типы используемых для передачи кабелей.

Таблица 3.4

	Среда передачи	***Макс. длина сегмента***	***Топология***	***Коннектор***
10 Base 2	50 Ом coaxial	185 м	bus	BNC
10 Base 5	50 Ом coaxial	500 м	bus	AUI
10 Base T	UTP 3 \4 \5 (две пары)	100 м	star	RJ-45
100 Base TX	UTP 5 (две пары)	100 м	star	RJ-45
100 Base FX	62.5 / 125 мкм MMF	400 м	star	MIC/ ST/ SC
1000 Base CX	STP	25 м	star	RJ-45
1000 Base T	UTP 5 (четыре пары)	100 м	star	RJ-45
1000 Base SX	62.5 / 50 мкм MMF	62.5—275 м 50—550 м	star	SC
1000 Base LX	62.5 / 50 мкм MMF 9 мкм SMF	62.5 мкм — 440 м 50 мкм — 550 м 9 мкм — 10 км	star	SC

На настоящее время наиболее широко распространена витая пара 5 категории, не смотря на то, что выпускаются UTP 6 и UTP 7. Когда витая пара категории 5 только появилась, лишь немногим системам был действительно необходим предоставляемый ею диапазон рабочих частот. Так, *Ethernet* на 10 Мбит/с и *Token Ring* на 4 Мбит/с разрабатывались в расчете на проводку UTP 3. Однако с появлением новых систем, таких, как *100 Base T* и *ATM* на 155 Мбит/с, потребность в UTP 5 стала очевидной. В последнее время уже новые протоколы, в частности *ATM* на

622 Мбит/с и *1000 Base T*, заставляют многих задуматься о достаточности UTP 5 для их реализации.

Сложные схемы кодирования. В целях оптимального распределения энергии по диапазону частот системами типа *100 Base T* используются многоуровневые схемы кодирования. Они имеют множество достоинств, в частности низкий уровень шумов. К сожалению, чем сложнее схема кодирования, тем чувствительнее система. Поэтому кабель не должен иметь разрывов импеданса, обладая при этом хорошей изоляцией.

Функционирование в полнодуплексном режиме. В системах наподобие *10 Base T* в каждый конкретный момент времени активна только одна пара. Новые системы могут работать в полнодуплексном режиме, т. е. сигналы передаются и принимаются одновременно. Это позволяет увеличить пропускную способность кабеля UTP фактически вдвое. Однако для этого кабель должен иметь стабильные характеристики импеданса с минимальным отражением и хорошую изоляцию от перекрестных помех между парами на ближнем/дальнем конце.

Использование нескольких пар. В обычных сетях активны только две из четырех пар. Между тем пропускную способность можно значительно увеличить за счет использования всех четырех пар кабеля UTP 5. Чтобы это стало возможным, кабель должен обеспечивать при прохождении сигнала как можно меньшие помехи между парами, когда активны все четыре пары. Это послужило толчком к сертификации кабелей UTP 5 на соответствие параметрам суммарного затухания.

Погонное затухание. Одной из наиболее серьезных проблем для любой кабельной инфраструктуры является затухание сигнала. К сожалению, при передаче информации от устройства к устройству качество сигнала ухудшается. Так, при прохождении расстояния в 100 м по кабелю UTP сигнал *100 Base T* обычно теряет значительную часть своей первоначальной мощности (см. рис. 3.14). Если эти потери окажутся чересчур велики, то принимающее устройство не сможет распознать передаваемые данные. Чтобы этого не случилось, комитеты по стандартизации налагают ограничения на допустимый размер потерь.

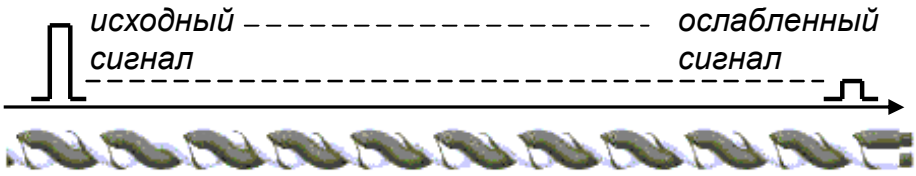


Рис. 3.14 — Затухание

Потери характеризуются термином *погонное затухание* или просто *затухание* (*attenuation*). В случае UTP затухание определяет величину потерь при прохождении сигнала по проводящей среде (3.23) и выражается в децибелах [дБ].

$$Att = -10 \cdot \lg\left(\frac{P_{ВЫХ}}{P_{ВХ}}\right) = -10 \cdot \lg\left(\frac{U_{ВЫХ}}{U_{ВХ}}\right). \qquad (3.23)$$

Использование децибел в качестве единицы измерения имеет свои преимущества. Например, нетрудно запомнить, что при затухании сигнала на 3 дБ он теряет 50 % своей мощности. В табл. 3.5 показано, как децибелы соотносятся с потерянной мощностью сигнала.

Таблица 3.5

Погонное затухание сигнала (дБ)	***Ослабление мощности сигнала (% потерь)***
3	50
6	75
10	90
15	97
20	99

Величина потерь зависит от конструкции кабеля, в том числе от размера проводника, состава, изоляции и/или материала оболочки, диапазона рабочих частот, скорости передачи и протяженности кабеля. Материал проводника (состав) также имеет большое значение. Например, медь имеет меньшие потери, чем сталь. Некоторые материалы, в частности серебро, имеют еще лучшие характеристики, нежели медь, однако многие из них слишком дороги для массового применения. Материал изоляции также может иметь влияние на затухание сигнала. В высококаче-

ственных кабелях UTP для изоляции проводника обычно используются материалы с низкими потерями, такие, как фторированный этиленпропилен или полиэтилен. Эти материалы обычно имеют меньшие потери, чем другие соединения, такие, как PVC. Материал оболочки также отражается на величине затухания. Именно поэтому многие производители отделяют оболочку от изолированных пар с помощью конструкции нежесткой трубы.

Разновидностью затухания является такая характеристика как ослабление отраженного сигнала (***RL — Return Lost***) так же измеряемая в децибелах [дБ] (3.24):

$$RL = -20 \cdot \lg\left(\frac{U_{OTP}}{U_{BX}}\right), \qquad (3.24)$$

здесь U_{BX} — амплитуда переданного, а U_{OTP} — амплитуда отраженного сигнала.

Кроме того, как известно, затухание в медной проводке UTP увеличивается с ростом частоты. Например, при 100 МГц затухание больше, чем при 1 МГц (при условии, что кабели имеют одинаковую длину). И, наконец, потеря сигнала зависит от протяженности кабеля. При прочих равных условиях — чем длиннее кабель, тем больше потери. По этой причине затухание выражается в децибелах на единицу длины.

По стандарту TIA\ EIA-568-A на длине 100 м и при температуре 20 °C частотная характеристика *A(f)* максимально допустимого затухания для кабелей категории UTP 3, UTP 4, UTP 5 определяется согласно следующему выражению:

$$A(f) = k_1 \cdot \sqrt{f} + k_2 \cdot \sqrt{f} + k_3 \cdot \sqrt{f}\,, \qquad (3.25)$$

где A, [дБ] — максимально допустимое затухание, f [МГц] — частота сигнала, k_1, k_2, k_3 — константы, определяемые в зависимости от категории кабеля.

Таблица 3.6

Категория кабеля	k_1	k_2	k_3
UTP 3	2,320	0,238	0,000
UTP 4	2,050	0,043	0,057
UTP 5	1,967	0,023	0,050

Характеристический импеданс соответствует входному импедансу однородной линии передачи бесконечной длины то есть линии передачи предельной длины, терминированной нагрузкой со значением ее собственного характеристического импеданса. В общем случае, характеристический импеданс — это комплексное число с резистивной и реактивной компонентами. Он является функцией частоты передаваемого сигнала и не зависит от длины линии. При очень высоких частотах характеристический импеданс асимптотически стремится к фиксированному резистивному сопротивлению. Типичное значение импеданса для витой пары — 100 м при частотах свыше 1 Гц.

Обратные потери (потери при отражении). Когда импеданс кабеля и нагрузки не совпадает, сигнал, распространяющийся по кабелю, частично будет отражаться в точке интерфейса кабель-нагрузка. Мощность отраженного сигнала носит название потерь при отражении или обратных потерь. Чем лучше совместимость импедансов, тем меньше отражаемая мощность и тем ниже обратные потери. UTP- кабель используется для соединения типа «точка-точка» и не поддерживает ответвлений (как коаксиальный кабель), в связи с чем, проблемы, связанные с отражением сигнала, частично снимаются.

Временная задержка распространения сигнала (*delay*)**.** Сигнал, распространяющийся от входной точки к выходной, приходит с временной задержкой, величина которой является отношением длины кабеля к скорости распространения сигнала *v* в передающей среде. В случае идеальной линии передачи, состоящей из двух проводников в вакууме, скорость распространения сигнала равна скорости распространения света в вакууме *c.* На практике скорость распространения сигнала в кабеле зависит от свойств диэлектрических материалов, окружающих проводники.

Значение задержки является частотно-зависимой величиной и, согласно проектам новых редакций стандартов на кабельные системы на основе витой пары не должно превышать

$$delay = 534 + \frac{36}{\sqrt{f}}, \text{ нс.} \qquad (3.26)$$

Перекос задержки. Другой привлекающий к себе значительное внимание параметр, характерный исключительно для ви-

тых пар — это *перекос задержки* (рис. 3.15). Перекос задержки характеризует синхронизацию путей передачи сигнала по разным парам в кабеле.

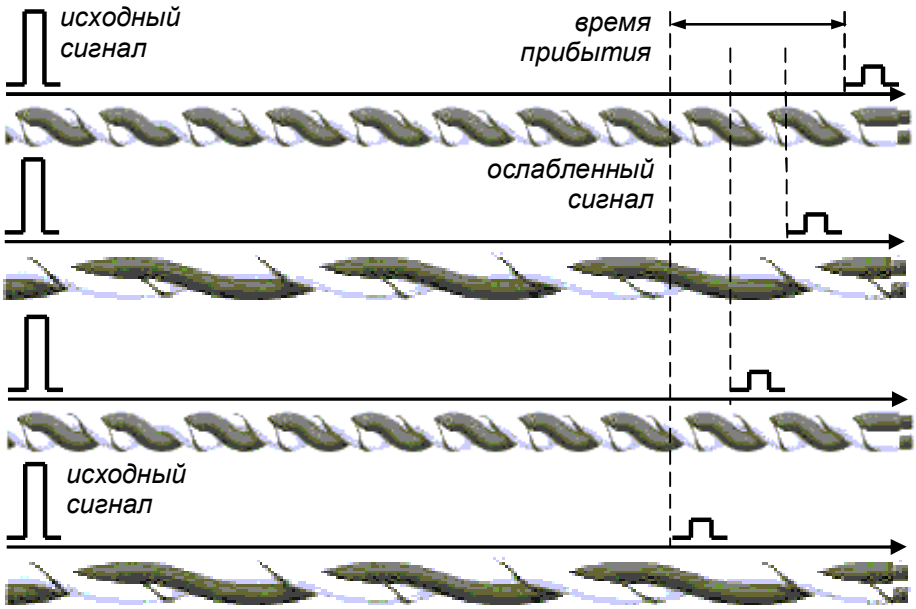


Рис. 3.15 — Перекос задержки

Когда все четыре пары активны, сигналы должны прибывать согласованно. Измеряемый в наносекундах [нс] перекос задержки характеризует разницу во времени поступления сигналов по разным парам кабеля. Если эта разница окажется чересчур велика, то принимающее устройство будет не в состоянии восстановить сигнал. В конечном итоге это приведет к ошибкам и потере данных.

***Перекрестные наводки* (*crosstalk*).** Витая пара называется активной, если по ней передается сигнал. Активная пара, естественно, создает электромагнитное поле. Это поле может оказывать влияние на другие находящиеся поблизости активные пары (см. рис. 3.16).

Один из наиболее сложных для понимания моментов в отношении перекрестных наводок связан с единицами измерения, а именно с децибелами. В случае погонного затухания чем больше величина в децибелах, тем выше потери сигнала. В случае перекрестной наводки все наоборот — чем больше величина в децибелах, тем меньше помехи (табл. 3.7).

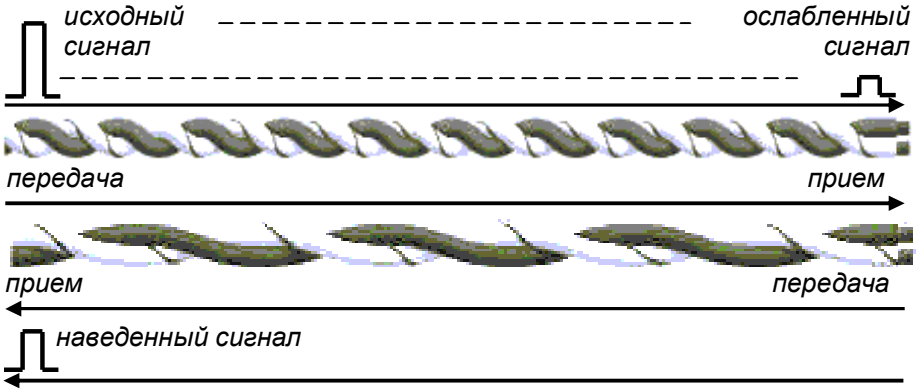


Рис. 3.16 — Перекрестные наводки

Таблица 3.7

Переходное затухание дБ	Напряжение на активной паре	Наведенное напряжение в соседней паре
3	1	0,7
6	1	0,5
10	1	0,3
20	1	0,1

Очевидно, появление шумов в соседних парах нежелательно. Как видно из диаграммы, чем больше величина переходного затухания в децибелах, тем меньше наведенное напряжение (т. е. шумы) в соседних парах. Погонное затухание характеризует потерю сигнала. Следовательно, чем больше величина в децибелах, тем выше потеря сигнала. Однако перекрестные наводки характеризуют потерю шума. В этом случае чем больше величина в децибелах, тем больше потери шума. И конечно, чем активнее затухает шум, тем лучше.

Виды перекрестных наводок

Перекрестные наводки на ближнем конце (***Near End Cross Talk — NEXT***). Такие системы, как *10BaseT Ethernet*, используют две пары для обмена данными: одну — для передачи, вторую — для приема (см. рис. 3.17) Сигнал имеет наибольшую мощность сразу же после момента передачи данных. И обратно, сигнал обладает наименьшей мощностью непосредственно перед моментом приема данных.

$$NEXT = -20 \cdot \lg\left(\frac{U_{ПОМ}}{U_{ВХ}}\right), \qquad (3.27)$$

где $U_{ВХ}$ — амплитуда переданного сигнала, а $U_{ПОМ}$ — амплитуда помехи на ближнем конце.

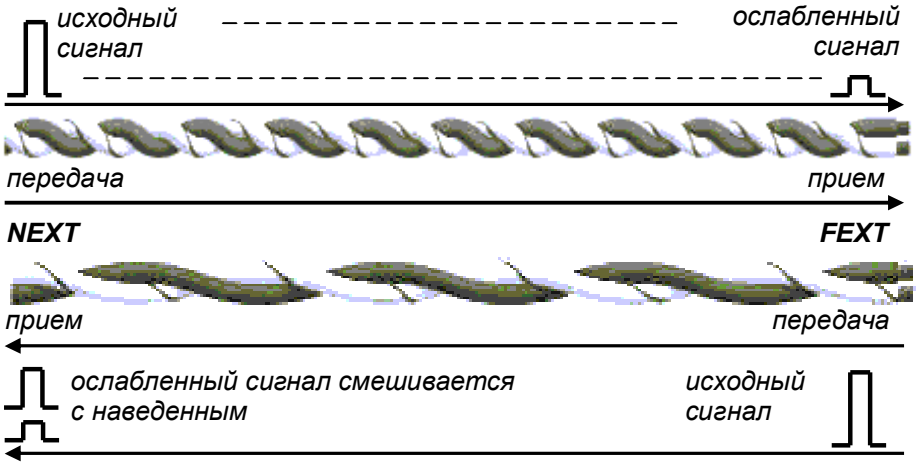


Рис. 3.17 — Перекрестные наводки на ближнем и дальнем конце

Наиболее часто термин *перекрестные наводки* используется вместе со словосочетанием на *ближнем конце*. Причина этого в том, что на ближнем конце, где сигнал имеет наибольшую мощность, он порождает мощное электромагнитное излучение (электромагнитные помехи). Рядом же с передатчиком по соседней паре идет ослабленный сигнал на приемник. Такая комбинация может иметь самые серьезные последствия для принимаемого сигнала, так как он оказывается под воздействием сильного соседнего поля. Это явление имеет место на ближнем конце, поэтому оно и выделяется.

Суммарные перекрестные наводки (***PSNEXT,*** *PowerSum NEXT*). Как отмечалось ранее, некоторые системы задействуют все четыре пары. При рассмотрении перекрестных наводок на ближнем конце мы исходили из того, что используются только две пары. Однако, если активны все четыре пары, как в стандарте на *Gigabit Ethernet*, они порождают значительные шумы. Для исследования паразитных электромагнитных влияний всех актив-

ных витых пар на выбранную используется такая характеристика, как суммарные перекрестные наводки.

Отношение сигнал\шум — отношение затухания сигнала к ослаблению перекрестной помехи ***ACR*** (*Attenuation-to-crosstalk ratio*). Положительное значение ACR означает, что сигнал преобладает над шумом, при отрицательном — что шум больше сигнала и прием сигнала, а тем более его распознавание, становится проблематичным. Для относительно надежной работы сети значение ACR должно быть не менее +2..+4 дБ.

Перекрестные наводки на дальнем конце (***Far End Cross Talk — FEXT***). Обычно данные передаются в одном направлении, а именно от передающего устройства к принимающему. Однако в некоторых системах данные передаются в двух направлениях. Такие системы называются полнодуплексными. В случае полнодуплексной передачи шумы возникают как на ближнем, так и на дальнем конце. Ввиду этого стандарт на допустимый уровень перекрестных наводок на дальнем конце введен во многие новые спецификации.

$$FEXT = -20 \cdot \lg\left(\frac{U_{ПОМ}}{U_{ВХ}}\right), \qquad (3.28)$$

где $U_{ВХ}$ — амплитуда переданного сигнала, а $U_{ПОМ}$ — амплитуда помехи на дальнем, в данном случае, конце.

Шум на дальнем конце измерить не так-то просто, потому что значительная доля шумов теряется или затухает по пути к тестовому устройству. Поэтому стандартной практикой является вычитание перекрестных наводок и учет только одних шумов. Величина «шумы минус затухание» получила название *приведенных перекрестных наводок на дальнем конце* или ***ELFEXT*** *(Equal Level FEXT)*.

$$ELFEXT = FEXT - Att\,. \qquad (3.29)$$

Сторонние перекрестные наводки. Этот термин используется для описания перекрестных помех между кабелями. Данный эффект наиболее заметен, когда активны несколько пар в кабеле. В этом случае излучаемая отдельным кабелем энергия может быть достаточно существенна.

Относительная скорость распространения сигнала (NVP). Параметр ***NVP*** (*Nominal Velocity of Propagation*) является

мерой замедления скорости распространения электромагнитной волны вдоль витой пары. Он численно равен отношению фактической скорости распространения к скорости света в вакууме и выражается в виде десятичной дроби или в процентах. Стандарты задают только самые общие требования к величине NVP (см. табл. 3.8), а их последние редакции не определяют метод измерения этого параметра.

Таблица 3.8

Частота, МГц	NVP		
	UTP 3	**UTP 4**	**UTP 5**
1	0,4	0,6	0,65
10	0,6	0,6	0,65
100	–	–	0,65

Эксплуатационные характеристики витой пары

Все кабели должны иметь витые пары проводов, применение кабелей с несвитыми попарно проводами не допускается. Это относится даже к коротким отрезкам плоского кабеля. При использовании экранированных кабелей на витой паре, сегменты последних рекомендуется заземлять на одном (и только на одном!) конце. На практике это удобнее производить на конце, подключенном к концентратору.

- минимальный радиус изгиба — 5 см;
- температура при работе и хранении:
  - –35...+60С — для кабеля в поливинилхлоридной оболочке;
  - –55...+200С — для кабеля в тефлоновой оболочке.
- температура при монтаже:
  - –20...+60С — для кабеля в поливинилхлоридной оболочке;
  - –35...+200С — для кабеля в тефлоновой оболочке.
- относительная влажность:
  - 0...+100% — для кабеля в поливинилхлоридной оболочке, допускается случайная конденсация;
  - не реагирует на влажность, конденсацию и водяные брызги — для кабеля в тефлоновой оболочке.

- возможность применения на открытом воздухе:
  – запрещено — для кабеля в поливинилхлоридной оболочке;
  – разрешено — для кабеля в тефлоновой оболочке.
- запрещено применение тонкого коаксиального кабеля для прокладки на открытом воздухе между двумя не связанными друг с другом зданиями (между зданиями, не имеющими общего контура заземления).

Типы разводки витой пары. Для подключения различных сетевых устройств посредством витой пары используется различные типы разводки кабеля. На рис. 3.18 приведены два основных типа правильной разводки витой пары прямая (*straight-through*) и перекрестная (*crossover*), называемая так же *нуль-хабовая.*

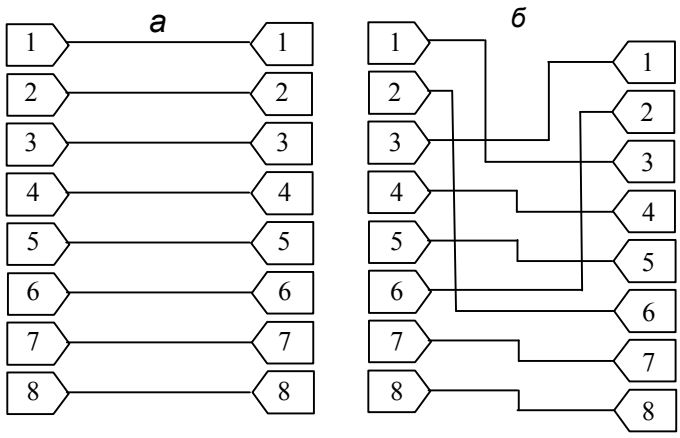


Рис. 3.18 — Витая пара straight-through (*а*)
и crossover (*б*) разводки

Различные типы разводки используются для организации среды передачи между различными сетевыми устройствами (см. табл. 3.9). На рис. 3.19 приведен пример типичной для локальной кабельной системы. Можно видеть, что различные фрагменты сети используют различные типы витой пары.

Таблица 3.9

***Straight-through* кабель**	***Crossover* кабель**
switch — router (коммутатор — маршрутизатор)	switch — switch (коммутатор — коммутатор)

Окончание табл. 3.9

***Straight-through* кабель**	***Crossover* кабель**
switch — PS or Server (коммутатор — ПК или сервер)	switch — hub (коммутатор — концентратор)
hub — PS or Server (концентратор — ПК или сервер)	hub — hub (концентратор — концентратор)
	router — router (маршрутизатор — маршрутизатор)
	PS — PS or Server (ПК — ПК или сервер)
	router — PC (маршрутизатор — ПК)

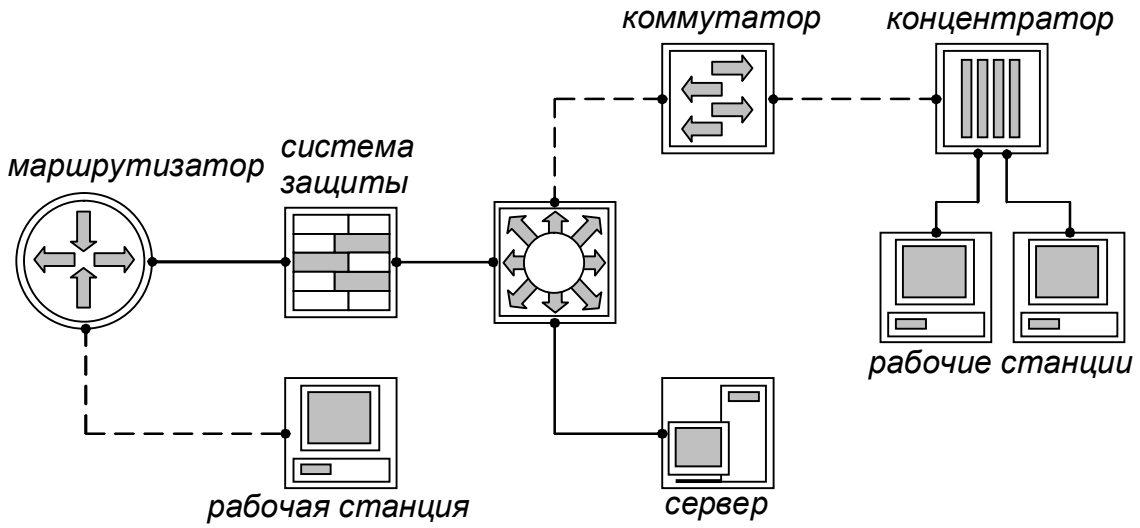


Рис. 3.19 — Использование straight-through (•———•) crossover (•– – –•) разводки витой пары

Структурированная кабельная система

Согласно зарубежным исследованиям (*LAN Technologies*), 70 % времени простоев обусловлено проблемами, возникшими вследствие низкого качества применяемых кабельных систем. Поэтому так важно правильно построить фундамент сети — кабельную систему. В последнее время в качестве такой надежной основы все чаще используется *структурированная кабельная система.*

Структурированная кабельная система (***Structured Cabling System, SCS***) — это набор коммутационных элементов (кабелей,

разъемов, коннекторов, кроссовых панелей и шкафов), а также методика их совместного использования, которая позволяет создавать регулярные, легко расширяемые структуры связей в вычислительных сетях.

При относительно высокой начальной стоимости, структурированные кабельные системы оправдывают капиталовложения за счет:

- длительного использования;
- допускают одновременное использование разных протоколов и сред передачи данных;
- модульности и возможности внесения изменений, а также наращивания мощности без влияния на существующие сети;
- позволяет обеспечить одновременный и быстрый доступ ко всем системам, проложенным в кабельных каналах;
- не зависят от поставщика сетевого оборудования;
- являясь единой сетью, позволяют создавать независимые участки сети;
- допускают использование ранее установленного оборудования;
- не зависят от изменений в информационных технологиях;
- обеспечивает зрительное восприятие разделения кабельных подсистем по функциональному признаку.

Структурированные кабельные системы — это реализация модульного представления о кабельных системах связи, рассматривающая последние в виде набора подсистем. Для того чтобы проектирование проистекало менее болезненно, и для того, чтобы в процессе эксплуатации ее было несложно модернизировать, расширить или даже перепрофилировать, кабельную подсистему желательно рассматривать в виде нескольких стандартизованных компонент — подсистем.

СКС выделяют пять таких подсистемы: горизонтальную подсистему; вертикальную подсистему; кампус (базовую подсистему — магистраль между зданиями); подсистему рабочей группы и административную подсистему.

Подсистема рабочей группы — это функционально-территориальная подсистема. Как правило, пользователь начинает думать о локальной вычислительной сети уже имея рабочие места, оснащенные компьютерами. Очень часто при этом некото-

рые компьютеры оказываются сопряженными или друг с другом, или с какими-то устройствами (обычно приборами, принтерами и модемами коллективного использования). То есть пользователь перед началом выполнения работ по проектированию ЛВС уже имеет кабельную подсистему той или иной степени сложности. Эту подсистему можно сохранить, если она в достаточной степени развита, или заменить на более приспособленную для решения задач данной рабочей группы.

Горизонтальная подсистема — это территориальная подсистема. Обычно основной объем работ по прокладкам кабеля приходится на нее. Подсистема рабочей группы и административная подсистема, как правило, являются ее составными частями. В зависимости от характеристик объекта, на котором она устанавливается (производственный цех, этаж административного здания, спортивный стадион, морской порт, выставочный павильон и т.п.), эту подсистему приходится проектировать на оптоволокне, защищенной или незащищенной витой паре, коаксиальном кабеле.

В чаще всего в горизонтальных подсистемах применяется оборудование, работающего со скоростью 100 Мбит/с. В тех же случаях, когда в ближайшей перспективе нет смысла в использовании сетевого оборудования с пропускной способностью выше 10 Мбит/с (оборудование 3-й категории), но есть перспектива развития сети, желательно сразу установить кабельную систему, способную работать со скоростью 100 Мбит/с (5-й категории). Это позволит во-первых, немного приподнять общую производительность сети благодаря уменьшению количества коллизий, а во-вторых, при дальнейшем развитии сети (переходе на оборудование 5-й категории) не придется производить никаких работ, связанных с заменой кабельного хозяйства.

Вертикальные подсистемы — территориальные подсистемы, служащие для подключения горизонтальных подсистем друг к другу. Обычно реализуются на базе коаксиального кабеля, защищенной витой пары или волоконно-оптического кабеля.

Административная подсистема. Эту кабельную подсистему, как правило, не выделяют в виде самостоятельной структуры. С одной стороны это правильно, но часто ее желательно обозначить перед заказчиком как отдельную структуру. Администт-

ративная подсистема кабельного монтажа — это функциональная подсистема. Ее назначение связывать подсистемы рабочих групп и горизонтальные подсистемы в единое целое. Она должна обеспечивать возможность установления резервных связей, подключение дополнительных рабочих мест и других подсистем. Нередко в рамках административной подсистемы требуется поддержка автономной системы энергоснабжения, голосовой и видеосвязи. Одно из основных требований к административной подсистеме — гибкость и возможность увеличения мощности.

Базовые подсистемы (*кампус*) служат для объединения вертикальных или административных подсистем друг с другом. В этом случае наиболее оправдано применение оптоволокна. В настоящее время на оптоволокне *Ethernet* работает со скоростями 10 Мбит/с и 100 Мбит/с, в ближайшем будущем ожидается появление оборудования со скоростью 660 Мбит/с (теоретическая пропускная способность оптических кабелей на сегодня оценивается цифрой 200 Гбит/с). Многие компании используют для организации базовых подсистем оборудование, поддерживающее *FDDI* стандарт — волоконный распределенный интерфейс данных, имеющий производительность 100 Мбит/с. В последнее время, с утверждением стандарта на *ATM*, в мире все шире начинает применяться этот тип оборудования.

На кабельных системах экономить неразумно. Лучше поставить на 2—3 компьютера меньше. Их всегда можно докупить чуть позже (и при этом дешевле), а кабельную подсистему придется менять или реконструировать и это будут выброшенные «на ветер» деньги.

3.9 Характеристики оптоволоконных каналов

Волоконно-оптические линии связи (***ВОЛС***) — это вид связи, при котором информация передается по оптическим диэлектрическим волноводам, известным под названием *оптическое волокно*. Оптическое волокно в настоящее время считается самой совершенной физической средой для передачи информации, а также самой перспективной средой для передачи больших потоков информации на значительные расстояния.

Физические особенности

• ***Широкополосность*** оптических сигналов, обусловленная чрезвычайно высокой частотой несущей ($F = 10^{14}$ Гц). Это означает, что по оптической линии связи можно передавать информацию со скоростью порядка 10^{12} бит/с (Терабит/с). Скорость передачи данных может быть увеличена за счет передачи информации сразу в двух направлениях, так как световые волны могут распространяться в одном волокне независимо друг от друга. Кроме того, в оптическом волокне могут распространяться световые сигналы двух разных поляризаций, что позволяет удвоить пропускную способность оптического канала связи. На сегодняшний день предел по плотности передаваемой информации по оптическому волокну не достигнут.

• Очень малое (по сравнению с другими средами) ***затухание*** светового сигнала в волокне. Лучшие образцы российского волокна имеют затухание 0.22 дБ/км на длине волны 1.55 мкм, что позволяет строить линии связи длиной до 100 км без регенерации сигналов. Для сравнения, лучшее волокно Sumitomo на длине волны 1.55 мкм имеет затухание 0.154 дБ/км. В оптических лабораториях США разрабатываются еще более «прозрачные», так называемые фторцирконатные волокна с теоретическим пределом порядка 0,02 дБ/км на длине волны 2.5 мкм. Лабораторные исследования показали, что на основе таких волокон могут быть созданы линии связи с регенерационными участками через 4600 км при скорости передачи порядка 1 Гбит/с.

Технические особенности

• Волокно изготовлено из кварца, основу которого составляет двуокись кремния, широко распространенного, а потому *недорогого* материала, в отличие от меди.

• Оптические волокна имеют диаметр около 100 мкм, то есть очень *компактны и легки*, что делает их перспективными для использования в авиации, приборостроении, в кабельной технике.

• Стеклянные волокна — не металл, при строительстве систем связи автоматически достигается *гальваническая развязка* сегментов. Применяя особо прочный пластик, на кабельных заводах изготавливают самонесущие подвесные кабели, не содержа-

щие металла и тем самым безопасные в электрическом отношении.

- Системы связи на основе оптических волокон *устойчивы к электромагнитным помехам*, а передаваемая по световодам информация защищена от несанкционированного доступа. Волоконно-оптические линии связи нельзя подслушать неразрушающим способом. Всякие воздействия на волокно могут быть зарегистрированы методом мониторинга (непрерывного контроля) целостности линии. Теоретически существуют способы обойти защиту путем мониторинга, но затраты на реализацию этих способов будут столь велики, что превзойдут стоимость перехваченной информации.
- Важное свойство оптического волокна — *долговечность*. Время жизни волокна, то есть сохранение им своих свойств в определенных пределах, превышает 25 лет, что позволяет проложить оптико-волоконный кабель один раз и, по мере необходимости, наращивать пропускную способность канала путем замены приемников и передатчиков на более быстродействующие.
- Существует способ скрытой передачи информации по оптическим линиям связи. При скрытой передаче сигнал от источника излучения модулируется не по амплитуде, как в обычных системах, а по фазе. Затем сигнал смешивается с самим собой, задержанным на некоторое время, большее, чем время когерентности источника излучения.

При таком способе передачи информация не может быть перехвачена амплитудным приемником излучения, так как он зарегистрирует лишь сигнал постоянной интенсивности.

Для обнаружения перехватываемого сигнала понадобится перестраиваемый интерферометр Майкельсона специальной конструкции. Причем, видимость интерференционной картины может быть ослаблена как 1:2N, где N — количество сигналов, одновременно передаваемых по оптической системе связи. Можно распределить передаваемую информацию по множеству сигналов или передавать несколько шумовых сигналов, ухудшая этим условия перехвата информации. Потребуется значительный отбор мощности из волокна, чтобы несанкционированно принять оптический сигнал, а это вмешательство легко зарегистрировать системами мониторинга.

Есть в волоконной технологии и свои недостатки:

- При создании линии связи требуются высоконадежные активные элементы, преобразующие электрические сигналы в свет и свет в электрические сигналы. Необходимы также оптические коннекторы (соединители) с малыми оптическими потерями и большим ресурсом на подключение-отключение. Точность изготовления таких элементов линии связи должна соответствовать длине волны излучения, то есть погрешности должны быть порядка доли микрона. Поэтому производство таких компонентов оптических линий связи очень дорогостоящее.
- Другой недостаток заключается в том, что для монтажа оптических волокон требуется прецизионное, а потому дорогое, технологическое оборудование. Как следствие, при аварии (обрыве) оптического кабеля затраты на восстановление выше, чем при работе с медными кабелями.

Преимущества от применения волоконно-оптических линий связи (ВОЛС) настолько значительны, что несмотря на перечисленные недостатки оптического волокна, эти линии связи все шире используются для передачи информации.

Оптическое волокно

Важнейший из компонентов *ВОЛС* — оптическое волокно. Для передачи сигналов применяются два вида волокна: ***одномодовое*** и ***многомодовое***. Свое название волокна получили от способа распространения излучения в них. Волокно состоит из сердцевины и оболочки с разными показателями преломления *n1* и *n2*.

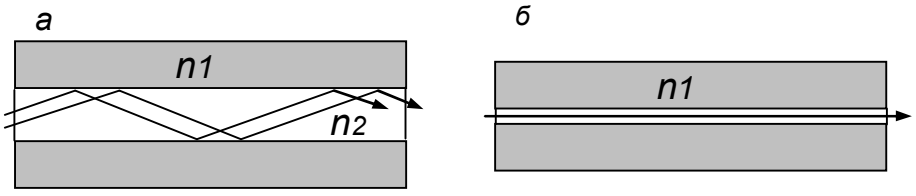


Рис. 3.20 — Многомодовое (*а*) и одномодовое (*б*) волокно

В одномодовом волокне диаметр световодной жилы порядка 8—10 мкм, то есть сравним с длиной световой волны. При такой

геометрии в волокне может распространяться только один луч (*одна мода*).

В многомодовом волокне размер световодной жилы порядка 50—60 мкм, что делает возможным распространение большого числа лучей (*много мод*).

Оба типа волокна характеризуются двумя важнейшими параметрами: затуханием и дисперсией.

Затухание обычно измеряется в дБ/км и определяется потерями на *поглощение* и на *рассеяние* излучения в оптическом волокне.

Потери на поглощение зависят от чистоты материала, потери на рассеяние зависят от неоднородностей показателя преломления материала.

Затухание зависит от длины волны излучения, вводимого в волокно. В настоящее время передачу сигналов по волокну осуществляют в трех диапазонах: 0.85 мкм, 1.3 мкм и 1.55 мкм, так как именно в этих диапазонах кварц имеет повышенную прозрачность.

Другой важнейший параметр оптического волокна — дисперсия. ***Дисперсия*** — это рассеяние во времени спектральных и модовых составляющих оптического сигнала. Существуют три типа дисперсии: *модовая, материальная* и *волноводная*. ***Модовая дисперсия*** присуща многомодовому волокну и обусловлена наличием большого числа мод, время распространения которых различно. ***Материальная дисперсия*** обусловлена зависимостью показателя преломления от длины волны. ***Волноводная дисперсия*** обусловлена процессами внутри моды и характеризуется зависимостью скорости распространения моды от длины волны.

Поскольку светодиод или лазер излучает некоторый спектр длин волн, дисперсия приводит к уширению импульсов при распространении по волокну и тем самым порождает искажение сигналов. При оценке пользуются термином ***полоса пропускания*** — это величина, обратная к величине уширения импульса при прохождении им по оптическому волокну расстояния в 1 км. Измеряется полоса пропускания в [МГц·км]. Из определения полосы пропускания видно, что дисперсия накладывает ограничение на дальность передачи и на верхнюю частоту передаваемых сигналов.

Если при распространении света по многомодовому волокну как правило преобладает модовая дисперсия, то одномодовому волокну присущи только два последних типа дисперсии. На длине волны 1.3 мкм материальная и волноводная дисперсии в одномодовом волокне компенсируют друг друга, что обеспечивает наивысшую пропускную способность.

Затухание и дисперсия у разных типов оптических волокон различны. Одномодовые волокна обладают лучшими характеристиками по затуханию и по полосе пропускания, так как в них распространяется только один луч. Однако, одномодовые источники излучения в несколько раз дороже многомодовых. В одномодовое волокно труднее ввести излучение из-за малых размеров световодной жилы, по этой же причине одномодовые волокна сложно сращивать с малыми потерями. Оконцевание одномодовых кабелей оптическими разъемами также обходится дороже.

Многомодовые волокна более удобны при монтаже, так как в них размер световодной жилы в несколько раз больше, чем в одномодовых волокнах. Многомодовый кабель проще оконцевать оптическими разъемами с малыми потерями (до 0.3 dB) в стыке. На многомодовое волокно расчитаны излучатели на длину волны 0.85 мкм — самые доступные и дешевые излучатели, выпускаемые в очень широком ассортименте. Но затухание на этой длине волны у многомодовых волокон находится в пределах 3—4 dB/км и не может быть существенно улучшено. Полоса пропускания у многомодовых волокон достигает 800 МГц·км, что приемлемо для локальных сетей связи, но не достаточно для магистральных линий.

Волоконно-оптический кабель

Вторым важнейшим компонентом, определяющим надежность и долговечность ВОЛС, является волоконно-оптический кабель (ВОК). На сегодня в мире несколько десятков фирм, производящих оптические кабели различного назначения. Наиболее известные из них: AT&T, General Cable Company (США); Siecor (ФРГ); BICC Cable (Великобритания); Les cables de Lion (Франция); Nokia (Финляндия); NTT, Sumitomo (Япония), Pirelli (Италия).

Определяющими параметрами при производстве ВОК являются условия эксплуатации и пропускная способность линии связи.

По условиям эксплуатации кабели подразделяют на:

- монтажные,
- станционные,
- зоновые,
- магистральные.

Первые два типа кабелей предназначены для прокладки внутри зданий и сооружений. Они компактны, легки и, как правило, имеют небольшую строительную длину.

Кабели последних двух типов предназначены для прокладки в колодцах кабельных коммуникаций, в грунте, на опорах вдоль ЛЭП, под водой. Эти кабели имеют защиту от внешних воздействий и строительную длину более двух километров.

Для обеспечения большой пропускной способности линии связи производятся ВОК, содержащие небольшое число (до 8) одномодовых волокон с малым затуханием, а кабели для распределительных сетей могут содержать до 144 волокон как одномодовых, так и многомодовых, в зависимости от расстояний между сегментами сети.

При изготовлении ВОК в основном используются два подхода:

- конструкции со свободным перемещением элементов,
- конструкции с жесткой связью между элементами.

По видам конструкций различают кабели повивной скрутки, пучковой скрутки, кабели с профильным сердечником, а также ленточные кабели. Существуют многочисленные комбинации конструкций ВОК, которые в сочетании большим ассортиментом применяемых материалов позволяют выбрать исполнение кабеля, наилучшим образом удовлетворяющее всем условиям проекта, в том числе — стоимостным.

Особый класс образуют кабели, встроенные в грозотрос.

Отдельно рассмотрим способы сращивания строительных длин кабелей.

Сращивание строительных длин оптических кабелей производится с использованием кабельных муфт специальной конструкции. Эти муфты имеют два или более кабельных ввода, при-

способления для крепления силовых элементов кабелей и одну или несколько сплайс-пластин. Сплайс-пластина — это конструкция для укладки и закрепления сращиваемых волокон разных кабелей.

Оптические соединители

После того, как оптический кабель проложен, необходимо соединить его с приемо-передающей аппаратурой. Сделать это можно с помощью оптических коннекторов (соединителей). В системах связи используются коннекторы многих видов. Сегодня мы рассмотрим лишь основные виды, получившие наибольшее распространение в мире.

Характеристики коннекторов представлены в табл. 3.10. Когда мы говорим, что данные виды коннекторов имеют наибольшее распространение, то это означает, что большинство приборов ВОЛС имеют розетки (адаптеры) под один из перечисленных видов коннекторов.

Таблица 3.10

Тип разъема	Фиксация	ЛВС	Телеком-муникации	Кабельное ТВ	Измерит. аппаратура
FC/PC	Резьба	+	+	+	
ST	BNC	+	+		
SMA	Резьба	+			+
SC	Push-Pull*	+	+	+	+
FDDI(MIC)	Push-Pull*	+			

*Фиксация «Push-Pull» обеспечивает подключение коннектора к розетке наиболее простым образом — на защелке. Защелка-фиксатор обеспечивает надежное соединение, при этом не нужно вращать накидную гайку. Важное преимущество разъемов с фиксацией Push-Pull — это высокая плотность монтажа оптических соединителей на распределительных и кроссовых панелях и удобство подключения.

Электронные компоненты систем оптической связи

Теперь коснемся проблемы передачи и приема оптических сигналов.

Первое поколение передатчиков сигналов по оптическому волокну было внедрено в 1975 году. Основу передатчика состав-

лял светоизлучающий диод, работающий на длине волны 0.85 мкм в многомодовом режиме.

В течение последующих трех лет появилось второе поколение — одномодовые передатчики, работающие на длине волны 1.3 мкм. В 1982 году родилось третье поколение передатчиков — диодные лазеры, работающие на длине волны 1.55 мкм.

По мере продолжения исследований появилось четвертое поколение оптических передатчиков, давшее начало когерентным системам связи — то есть системам, в которых информация передается модуляцией частоты или фазы излучения. Такие системы связи обеспечивают гораздо большую дальность распространения сигналов по оптическому волокну. Специалисты фирмы NTT построили безрегенераторную когерентную ВОЛС STM-16 на скорость передачи 2.48832 Гбит/с протяженностью в 300 км, а в лабораториях NTT в начале 1990 года ученые впервые создали систему связи с применением оптических усилителей на скорость 2.5 Гбит/с на расстояние 2223 км.

Появление оптических усилителей на основе световодов, легированных эрбием, способных усиливать проходящие по световоду сигналы на 30 dB, дало начало пятому поколению систем оптической связи. В настоящее время быстрыми темпами развиваются системы дальней оптической связи на расстояния в тысячи километров. Успешно эксплуатируются трансатлантические линии связи США-Европа TAT-8 и TAT-9, Тихоокеанская линия США-Гавайские острова-Япония TPC-3. Ведутся работы по завершению строительства глобального оптического кольца связи Япония-Сингапур-Индия-Саудовская Аравия-Египет-Италия.

В последние годы наряду с когерентными системами связи развивается альтернативное направление: *солитоновые системы связи*. ***Солитон*** — это световой импульс с необычными свойствами: он сохраняет свою форму и теоретически может распространяться по «идеальному» световоду бесконечно далеко. Солитоны являются идеальными световыми импульсами для связи. Длительность солитона составляет примерно 10 трилионных долей секунды (10 пс). Солитоновые системы, в которых отдельный бит информации кодируется наличием или отсутствием солитона, могут иметь пропускную способность не менее 5 Гбит/с на расстоянии 10 000 км.

Такую систему связи предполагается использовать на уже построенной трансатлантической линии ТАТ-8. Для этого придется поднять подводный ВОК, демонтировать все регенераторы и срастить все волокна напрямую. В результате на подводной магистрали не будет ни одного промежуточного регенератора.

Применение ВОЛС в вычислительных сетях

Наряду со строительством глобальных сетей связи оптическое волокно широко используется при создании локальных вычислительных сетей (ЛВС).

Сейчас наиболее популярны локальные и магистральные сети на базе Ethernet, Fast Ethernet, FDDI, ATM/SDH с применением оптических линий связи.

При установке протяженных сегментов сети не требуются повторители.

В оптических линиях связи очень низкий уровень шумов, что позволяет передавать информацию с коэффициентом ошибок не более 10**(–10).

Волоконно-оптические линии связи позволяют наращивать вычислительные возможности сети без замены кабельных коммуникаций. Для этого нужно просто установить более быстродействующие передатчики и приемники.

Кабель для связи сегментов сети стоит недорого, но работы по его прокладке могут составить самую крупную статью расходов по установке сети. Потребуется труд не только техников-кабельщиков, но и целой команды строителей (штукатуров, маляров, электриков), что обойдется недешево, если учесть возрастающую стоимость ручного труда.

В настоящее время оптическое волокно сложно использовать при строительстве общей шины, но его удобно использовать для связи «точка-точка», применяемой в топологии «звезда» и «кольцо».

Схема ВОЛС, применяемых, в частности, в ЛВС, устроена следующим образом:

Электрический сигнал идет от сетевого контроллера, устанавливаемого в рабочую станцию или сервер (например, сетевой контроллер Ethernet), затем поступает на электрический вход трансивера, который преобразует электрический сигнал в опти-

ческий. Оптический кабель (например, ОКГ-50-2) присоединяется к оптическим разъемам трансивера с помощью оптических соединителей (например, ST).

3.10 Характеристики беспроводных каналов

В качестве среды передачи данных в вычислительных сетях используются также электромагнитные волны различных частот — КВ, УКВ, СВЧ. Однако пока в локальных сетях радиосвязь используется только в тех случаях, когда оказывается невозможной прокладка кабеля, что объясняется прежде всего недостаточной надежностью сетевых технологий, построенных на использовании электромагнитного излучения. Для построения глобальных каналов этот вид среды передачи данных используется шире — на нем построены спутниковые каналы связи и наземные радиорелейные каналы, работающие в зонах прямой видимости в СВЧ-диапазонах.

Радиоканалы входят необходимой составной частью в спутниковые и радиорелейные системы связи, применяемые в территориальных сетях, в сотовые системы мобильной связи, они используются в качестве альтернативы кабельным системам в локальных сетях и при объединении сетей отдельных ЛВС и предприятий в корпоративные сети. Радиосвязь используется в корпоративных сетях и ЛВС, если затруднена прокладка других каналов связи.

Радиоканал в локальных сетях реализует следующие функции:

- Выполняет роль моста между подсетями (двухточечное соединение с направленными антеннами, дальность в пределах прямой видимости, обычно — 15—20 км, с расположением антенн на крышах зданий). Мост имеет два адаптера: один — для формирования сигналов радиоканала, другой — для кабельной подсети.
- Радиоканал является общей средой передачи данных. В случае использования радиоканала в качестве общей Среды передачи данных в ЛВС сеть называют *RadioEthernet* (стандарт *IEEE 802/11*), она обычно используется внутри зданий. В состав аппаратуры входят приемопередатчики и антенны. Связь осуще-

ствляется на частотах от 1 до нескольких ГГц. Расстояние между узлами — несколько десятков метров. В соответствии со стандартом IEEE 802/11 возможны два способа передачи двоичной информации в ЛВС, оба они имеют целью защитить информацию от нежелательного доступа. Первый способ имеет название *метода прямой последовательности* (*DSSS — Direct Sequence Spread Spectrum*). В нем вводится избыточность — каждый бит данных передается последовательностью из 11 элементов — *чипов*. Эта последовательность создается по алгоритму, известному участникам связи и дешифруется при приеме. Избыточность повышает помехоустойчивость, что позволяет снизить требования к мощности передатчика, а для сохранения высокой скорости — расширить полосу пропускания. Так в аппаратуре фирмы Aironet в диапазоне 2.4 ГГц имеется 4 канала шириной 22 МГц. Второй способ — *метод частотных скачков* (*FHSS — Frequency Hopping Spread Spectrum*). В этом методе полоса пропускания делится на 79 поддиапазонов. Передатчик периодически (с шагом 20—400 мс) переключается на новый поддиапазон, причем алгоритм изменения частот известен только участникам связи и может изменяться, что затрудняет несанкционированный доступ к данным.

- Радиоканал служит соединением между центральными и терминальными узлами в сети с централизованным управлением. В варианте использования радиоканала для связи центрального и периферийного узлов центральный пункт имеет ненаправленную антенну, терминальные пункты при этом имеют направленные антенны. Дальность связи так же составляет десятки метров, а вне помещений — сотни метров.

В оборудование беспроводных каналов передачи данных входят:

- *сетевые адаптеры и радиомодемы*, поставляемые вместе с комнатными антеннами и драйверами. Различаются способами обработки сигналов, характеризуются частотой передачи и пропускной способностью, дальностью связи. Сетевой адаптер вставляется в свободный разъем шины компьютера (например, ISA), а радиомодем подключается к цифровому оборудованию окончания через стандартный интерфейс (например, RS-232C, RS-449 или V.35);

- *радиомосты* используются для объединения между собой кабельных сегментов и отдельных локальных сетей в пределах прямой видимости и для организации магистральных каналов в опорных сетях, выполняют ретрансляцию и фильтрацию пакетов;
- *направленные и ненаправленные антенны*, антенные усилители и вспомогательное оборудование в виде кабелей, полосовых фильтров, грозозащитников и т.д.

Системы мобильной связи. Системы мобильной связи осуществляют передачу информации между пунктами, один из них или оба являются подвижными. Характерным признаком систем мобильной связи является применение радиоканала. К технологиям мобильной связи относятся также пейджинг, твейджинг, транкинг, для мобильной связи используются также и спутниковые каналы. *Сотовые технологии* обеспечивают телефонную связь между подвижными абонентами (ячейками). Связь осуществляется посредством стационарных станций, выполняющих коммутирующие функции. Базовый коммутатор обслуживает строго определенную зону. Доступ к радиоканалу в системах мобильной связи осуществляется одним из следующих способов.

- ***Случайный доступ*** (метод *АЛОХА*, назван так в связи с первым применением метода для связи между группой Гавайских островов). Применяется только при малых нагрузках. Его развитием стал метод *МДКН/ОК*, используемый в ЛВС и корпоративных сетях.
- ***Технология CDMA*** (*Code Division Multiple Access*) — выделение для каждого абонента своей кодовой комбинации, которой кодируются символы 1 и 0 передаваемых сообщений. Это широкополосная технология с возможностью одновременной передачи в отведенной полосе частот нескольких сообщений с различными кодами символов.
- ***Технология TDMA*** (*Time Division Multiple Access*) — временное мультиплексирование с выделением слота по требованию. Требования отсылаются в короткие интервалы времени (слоты запросов), при коллизиях запросы повторяются. Базовая станция выделяет свободные информационные слоты, сообщая их источнику и получателю.

3.11 Услуги спутниковой связи в России

Спутники в системах связи могут находиться на геостационарных (около 36 тыс. км) или низких орбитах. При геостационарных орбитах заметны задержки при прохождении сигналов (туда и обратно около 500 мс). Возможно покрытие поверхности всего земного шара с помощью четырех спутников. В низкоорбитальных системах обслуживание конкретного пользователя происходит попеременно разными спутниками. Чем ниже орбита, тем меньше площадь покрытия — тем больше число наземных станций или спутников (обычно требуется около десятка спутников).

Развитые корпоративные сети связи и передачи данных необходимы для успешной деятельности любого предприятия. Несмотря на то, что рынок услуг связи развивается быстро, в России до сих пор нет достаточного количества наземных сетей. Во многих регионах их строительство просто не оправдано экономически. Между тем использование спутниковых каналов и наземных станций позволяет быстро развернуть сети и сформировать сетевую инфраструктуру на большой территории. Наземные станции устанавливаются в любом удаленном от коммуникаций регионе, где спутниковая связь остается единственным доступным способом участия в современном бизнесе. Вместе с тем, применение спутниковой связи эффективно даже в регионах с развитыми наземными телекоммуникациями, так как затраты на эксплуатацию сети нередко значительно снижаются. Спутниковые решения для построения сетей на базе технологии VSAT обладают достаточно высокой гибкостью и могут быть легко адаптированы к требованиям бизнеса. В РФ услуги по строительству корпоративных широкополосных сетей с применением фиксированной спутниковой связи предлагают целый ряд компаний и организаций.

«Белком» (*http://www.belcom.ru*) предлагает решения по созданию корпоративных сетей для соединения представительств в СНГ с филиалами в любой точке мира для передачи данных, телефонии, Internet. С этой целью используются выделенные спутниковые каналы, каналы VSAT DAMA, комбинированные спутниковые и наземные каналы.

«ВИСАТ-ТЕЛ» (*http://www.vsat-tel.ru*) управляет проектами по созданию сетей спутниковой связи и предлагает решения для передачи данных, объединения локальных сетей, телефонной связи, видеоконференций, Internet с использованием VSAT на базе отечественного оборудования.

«Гейзер» (*http://www.geyser.ru*) формирует вторичные сети на основе сетей спутниковой связи VSAT. В числе услуг — поставка и монтаж оборудования, сопряжение с системами связи, адаптация сетей, создание корпоративных сетей передачи данных и телефонной связи, разработка интегрированных систем.

«ГИС-Инвест» (*http://www.gis.ru*) совместно с зарубежными партнерами специализируется на услугах доступа к Internet на базе спутниковых каналов и малых спутниковых терминалов.

«Зап-СибТрансТелеком» (ТТК) (*http://www.zsttk.ru*) занимается построением магистральной цифровой сети на базе оптических линий и системы спутниковой связи со станциями VSAT, предоставляет услуги связи операторам и корпоративным клиентам.

«Зонд-Холдинг» (*http://win.vsat.ru*) развивает сеть магистральных спутниковых каналов связи, предоставляет услуги строительства и эксплуатации сетей, аренду цифровых спутниковых каналов для передачи данных и телефонии с использованием спутников «Горизонт» и «Экспресс».

«Инжиниринг центр спутниковой связи» (*http://www.sbcec.ru*) владеет сетью KOBSTAR, охватывающей практически всю территорию РФ и СНГ, и обеспечивает поддержку телефонии, передачи данных, доступа в Internet через спутниковые и наземные каналы.

«Интерспутник» (*http://www.intersputnik.ru*) обеспечивает аренду спутниковой емкости для организации сетей связи на базе станций VSAT, сотрудничает с ФГУП «Космическая связь», «Информкосмос», «Локхид Мартин», KB Impuls Service (Германия), Gilat Satcom (Израиль). Совместно с партнерами компания предоставляет услуги по подключению корпоративных пользователей и провайдеров Internet к магистрали с использованием космического сегмента.

«Космическая связь» (*http://www.gpks.ru*) поддерживает каналы для организации через VSAT видеоконференций, обмена

данными между сетями, создания распределенных сетей IP и X.25, доступа к Internet и цифровым сетям общего пользования. Спутник покрывает практически всю территорию РФ, СНГ, Азии и Восточной Европы.

«Московский Телепорт» (MT) (*http://www.mteleport.ru*) выполняет заказы по строительству мультимедийных корпоративных сетей и сетей доступа к Internet. MT — партнер провайдера услуг спутниковой связи DeTeSat. Технологическая платформа SkyPerforer позволяет создавать сети до 10 узлов, а SkyWAN — до 250 узлов; IP Multicasting предоставляет высокоскоростной спутниковый канал. Спутники охватывают практически всю территорию РФ, СНГ и Европы.

«НэтЛайн» (*http://www.ntl.ru*) организует спутниковые и наземные цифровые каналы связи, резервные каналы и каналы «по требованию», корпоративные сети и системы связи.

«Ройлком» (*http://www.roilcom.ru*), оператор сетей спутниковой связи на территории РФ и США, предоставляет услуги связи, разрабатывает телекоммуникационные решения для территориально распределенных предприятий и корпоративных сетей связи.

«Санкт-Петербургский ТЕЛЕПОРТ» (*http://www.spb-teleport.ru*) совместно с компанией «Транстелеком» создает цифровую сеть связи, состоящую из системы фиксированной спутниковой связи «Транстелесат» на базе станций VSAT и оптических линий. Свободные ресурсы системы спутниковой связи доступны корпоративным потребителям в РФ и ближнем зарубежье.

«САТИС-ТЛ-94» (*http://www.satis-tl.com*) предоставляет все виды спутниковой связи: передачу данных, видеоконференции, подключение к Internet, телефонию, закрепленные каналы спутниковой связи в РФ и европейских странах, каналы по требованию (DAMA), SkyFrame, а также проектирует корпоративные сети передачи данных.

4 МОДУЛЯЦИЯ И КОДИРОВАНИЕ

4.1 Кодирование информации. Основные понятия

Информация — довольно широко исследуемое философское понятие. Истинность и значимость информации в нашем курсе не рассматривается. В контексте теории передачи данных информация будет трактоваться нами как некоторые ***сообщения***, составленные из *символов*. Объект, генерирующий информацию называется *источником* информации.

Количество информации передаваемой в сообщении определяется по формуле Хартли:

$$I(X) = -\log_2 P(X),$$

где *P(X)* — вероятность появления элемента *X* сообщения.

Из этой формулы следует, что единица измерения количества информации есть количество информации, содержащейся в одном бите двоичного кода при условии равной вероятности появления в нем 1 и 0. Один разряд десятичного кода содержит $I = -\log_2 0.1 = 3.32$ единицы информации, т.к. $P = 0{,}1$.

Энтропия источника информации есть мера неопределенности ожидаемой информации. Рассчитывается энтропия источника информации с независимыми сообщениями как среднее арифметическое количеств информации сообщений:

$$H(X) = -\sum_{k=1}^{N} P_k \cdot \log_2 P_k = \sum_{k=1}^{N} P_k \cdot I(X = a_k),$$

где P_k — вероятность появления k-го сообщения ($X = a_k$).

Пусть имеем два источника информации, один передает двоичный код с равновероятным появлением в нем 0 и 1: $P_0 = P_1 = 0.5$, другой имеет вероятность появления 1: $P_1 = 2^{-10}$, и вероятность 0: $P_0 = 1 - 2^{-10}$. Очевидно, что неопределенность в получении в очередном такте символа 1 или 0 от первого источника выше, чем от второго. Это подтверждается количественной оценкой энтропии: у первого источника $H_1 = 1$, у второго приблизительно $H_2 = -2^{-10} log_2 2^{-10}$, т.е. значительно меньше.

Итак информация — это уменьшение (снятие) неопределенности, энтропии. Теперь, если речь идет о последовательно получаемой информации (череде событий), то имеет смысл понятие

взаимная информация, т.е. информация о событии *X* изменится, если получено событие *Y*, связанное с событием *X*.

Информация о событии *X*, передаваемая событием *Y*, равна $I(X,Y) = H(X) - H(X|Y)$, где *H(X)* — энтропия события *X*, $H(X|Y)$ — энтропия события *X*, при условии, что *Y* — известно (рис. 4.1). Информация $I(X,Y)$ есть уменьшение первоначальной неопределенности (энтропии) *H(X)* когда происходит событие *Y*. Если события *X* и *Y* — независимы, то $H(X) = H(X|Y)$, а значит, $I(X,Y) = 0$, т.е. событие *Y* не дает никакой информации об *X*.

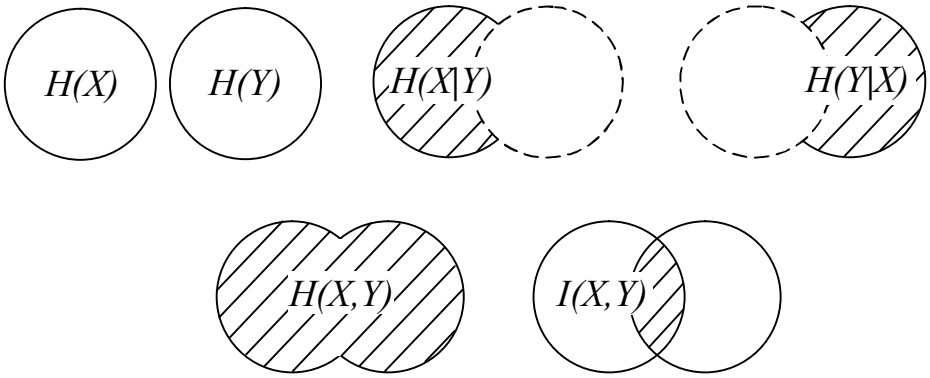


Рис. 4.1 — Отношение между энтропией
и информацией для двух событий

Источник информации — это производитель информационных символов (значений). Рассматриваются источники двух типов — *источник с памятью* и *источник без памяти*. ***Источник без памяти*** — это такой тип источника, при котором вероятность производимого символа не зависит от какого-либо из предыдущих символов. Если вероятность каждого символа не зависит от времени, то говорят, что источник *стационарный*. Для такого источника вероятность кажого символа S_i выражается следующим образом:

$$P(S_i = a_k) = P_k \ , \quad \forall i \, .$$

Энтропия такого источника определяется как

$$H(X) = \sum_{k=1}^{N} P_k \cdot \log_2 P_k \, ,$$

где $P_k > 0$. Набор допустимых символов $\{a_k\}$ называется алфавитом.

Источник с памятью — это обычно генератор не отдельных символов, а скорее группы символов, которые формируют различные сообщения. Например, каждая буква языка — это символ, но только некоторые группы символов образуют правильные (допустимые) комбинации — слова. Естественно, что каждый последующий символ в слове зависит от предыдущих, а последующее слово так же зависит от переданных ранее, таким образом, энтропию отдельного символа можно определять, угадывая вероятность следования определенного символа за предыдущим. Для больших участков текста правильное отгадывание последующего символа составляет немногим меньше 50 %, что определяет энтропию немного более, чем один бит на символ.

Кодирование. Реальные источники информации генерируют сообщения в неудобной для передачи форме (нет смысла обозначать каждое слово языка отдельным символом), практически удобнее отображать сообщения на управляемый набор кодовых символов (значений). Этот процесс называется ***кодированием***. Алфавит кодовых символов может совпадать с алфавитом источника информации, но не обязательно. В информационных системах это как правило *двоичный алфавит*.

Широко используются следующие двоичные коды: ***EBDCDIC*** (*Extended Binary Coded Decimal Interchange Code*), в котором символы кодируются восемью битами, (популярен благодаря его использованию в IBM, Российский эквивалент КОИ-8) и ***ASCII*** (*American Standards Committee for Information Interchange*) — семибитовый двоичный код (Российский эквивалент КОИ-7). Оба эти кода включают битовые комбинации для печатаемых символов и некоторых распространенных командных слов типа NUL, CR, ACK, NAK и др.

Для кодирования русского текста приходится вводить дополнительные битовые комбинации. Семибитовая кодировка здесь уже недостаточна. В восьмибитовой кодировке нужно под русские символы отводить двоичные комбинации, не занятые в общепринятом коде, чтобы сохранять неизменной кодировку латинской букв и других символов. Так возникли кодировка

КОИ-8, затем при появлении персональных компьютеров — альтернативная кодировка и при переходе к Windows — кодировка ***Windows-1251***. Множество используемых кодировок существенно усложняет проблему согласования почтовых программ в глобальных сетях.

Коды по отношению к *алфавиту кодирования* характеризуются такими величинами, как *минимальная разрядность кода*, *избыточностью* и *эффективностью*.

Минимальная разрядность кода определяет самый короткий возможный код, который полностью описывает источник. Сообщения такой длины не передают излишних (избыточных) символов, а потому, не генерируют никаких искажений и шума. Для двоичного кода длина должна быть больше или равна энтропии в битах.

Избыточность кода определяется как относительная разность между минимальной и средней длиной кодовых слов

$$R = \left(L_{\text{ср}} - L_{\min}\right) \Big/ L_{\text{ср}} .$$

Средняя длина вычисляется так:

$$L_{\text{ср}} = \sum_i P_i \cdot L_i ,$$

где L_i — длина i-того кодового слова, а P_i — вероятность этого слова.

Коэффициент избыточности сообщения определяется по формуле:

$$r = \left(I_{\max} - I\right) \Big/ I_{\max} ,$$

где I — количество информации в сообщении; $I_{\max}$ — максимально возможное количество информации в сообщении той же длины. Коэффициент избыточности сообщения характеризует количество неинформативных символов в сообщении. Например, последовательность *101010101010*... вообще не несет информации.

Пример избыточности дают сообщения на естественных языках. Так, у русского языка коэффициент избыточности r находится в пределах 0,3 … 0,5. Наличие избыточности позволяет ставить вопрос о сжатии информации без ее потери в передаваемых сообщениях.

Эффективность кода определяется как

$$\eta = \frac{L_{\min}}{L_{\text{ср}}}$$

и обычно выражается в процентах. Эффективность кода связана с избыточностью формулой $R = 1 - \eta$. Например, если код имеет эффективность 100%, то избыточность этого кода равна 0.

Для случая двоичного кода $L_{\min} = H$, так что величины избыточности R и эффективности η составляют:

$$R = \frac{L_{\text{ср}} - H}{L_{\text{ср}}}, \quad \eta = \frac{H}{L_{\text{ср}}}.$$

4.2 Кодирование информации. Классификация методов

С точки зрения теории передачи информации существует различные подходы к пониманию понятия кодирование (рис. 4.2). Вернее, различные процессы, производимые с информацией при ее передаче имеют одинаковое название «*кодирование*», но выполняют различные функции.

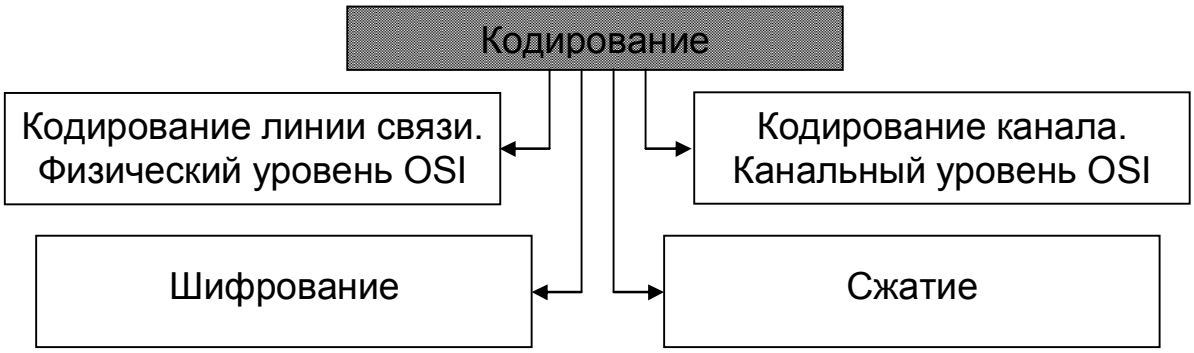


Рис. 4.2 — Кодирование информации

- ***Кодирование линии связи***. Состоит в преобразовании исходного сообщения к виду, удобному для транспортировки через среду передачи. Реализуется обычно на физическом уровне модели OSI для различных стеков сетевых протоколов, так как зависит от среды передачи.

Например, при передаче по телефонным воздушным линиям сообщение должно быть приведено к аналоговому сигналу с оп-

ределенными частотными и мощностными характеристиками. Передача по беспроводным сетям так же ведется в своем спектральном диапазоне, по оптоволоконным линиям передаются световые пучки и т.п. Для каждого способа передачи используется наиболее удобный способ предоставления информации, таким образом, одно и то же сообщение должно быть различным образом закодировано. К кодированию физического уровня относят такие коды как различные потенциальные и импульсные коды, модификации кодов RZ, NRZ, NRZI и т.п.

- ***Кодирование канала передачи***. Состоит в расширении кодового алфавита, позволяющем распознавать (и исправлять) ошибки после передачи через канал, котрый способен разрушать данные. В данном случае выбор способа кодирования не зависит от среды передачи, только лишь от степени искажения информации при передаче. Кодирование канального уровня OSI характеризуется *избыточностью* кода, т.е. число исходных комбинаций меньше числа закодированных. К кодам канального уровня можно отнести коды, обеспечивающие синхронизацию между передатчиком и приемником, распознающие и справляющие ошибки;
- ***Сжатие.*** Называемое так же ***кодированием источника*** состоит в максимально возможном сжатии информации до нескольких символов, для того, чтобы сэкономить ресурсы передачи или памяти. При помощи сжатия достигается максимальная *эффективность* кода: число исходных комбинаций как правило больше числа закодированных. Методы сжатия используются различными приложениями на разных уровнях OSI, при используются одинаковые алгоритмы.
- ***Шифрование***. Называется так же ***криптографическим кодированием***, состоит в таком преобразовании сообщения, чтобы другие пользователи не смогли понять его. Число исходных комбинаций в общем случае равно числу закодированных.

4.3 Кодирование на физическом уровне OSI

Любая информация передается по каналам связи в одной из двух форм — на основе синусоидального несущего сигнала или на основе последовательности прямоугольных импульсов. В со-

ответствии с распределением функций модели OSI, представление информации к виду удобному для передачи непосредственно по каналам связи (кодирование) должно производиться на канальном уровне. Но именно физический уровень обеспечивает передачу электрических сигналов, радиоволн или световых импульсов, а, следовательно, он определяет форму данных, передаваемых ему или принимаемых от него канальным уровнем.

При передаче в линию связи символы сообщения на физическом уровне представляются символами кода (например, двоичного) и реализуются сигналами в канале связи.

Кодирование в контексте данной темы нужно понимать как представление исходного сообщения последовательностью элементарных символов кода.

С точки зрения непрерывности или дискретности сигнала методы передачи данных физического уровня подразделяются на три основных группы (рис. 4.3): *аналоговая модуляция*, при которой цифровое или аналоговое сообщение кодируется непрерывным аналоговым сигналом, *дискретизация аналогового сигнала*, квантующая непрерывный сигнал несколькими дискретными значениями, и *цифровое кодирование*, переводящее один цифровой код в другой цифровой код, но более удобный для передачи.

При передаче дискретных данных по каналам связи применяются два основных типа физического представления информации: *аналоговая модуляция* (здесь кодирование осуществляется за счет изменения параметров аналогового сигнала — несущей) и *цифровое кодирование*. Помимо формы передаваемого сигнала, эти способы отличаются шириной спектра результирующего сигнала и сложностью аппаратуры, необходимой для их реализации.

При использовании прямоугольных импульсов спектр результирующего сигнала получается весьма широким. Это не удивительно, если вспомнить, что спектр идеального импульса имеет бесконечную ширину. Применение синусоиды приводит к спектру гораздо меньшей ширины при той же скорости передачи информации. Однако для реализации синусоидальной модуляции требуется более сложная и дорогая аппаратура, чем для реализации прямоугольных импульсов.

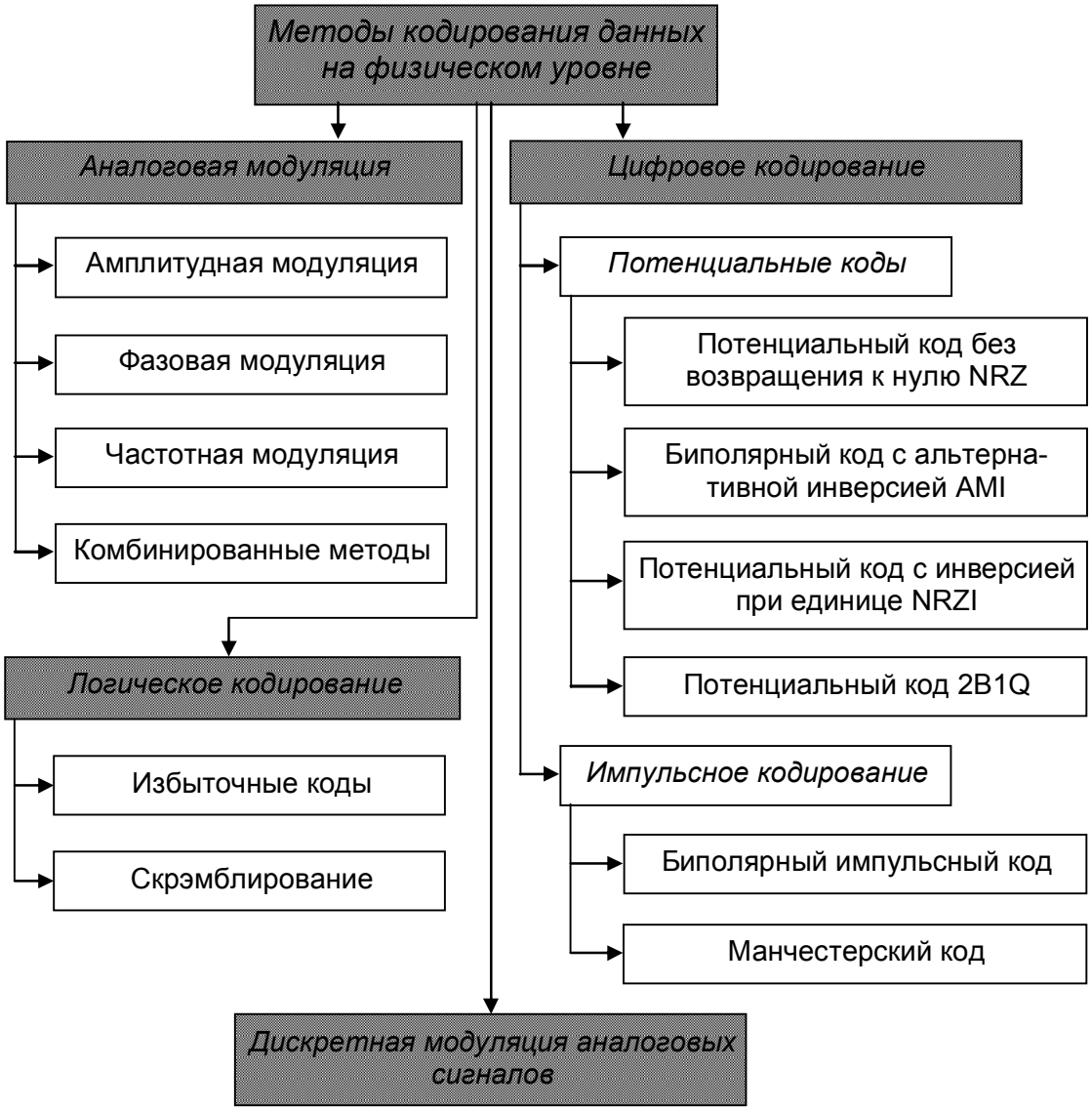


Рис. 4.3 — Методы передачи данных физического уровня

Аналоговая модуляция применяется для передачи дискретных данных по каналам с узкой полосой частот, типичным представителем которых является *канал тональной частоты,* предоставляемый в распоряжение пользователям общественных телефонных сетей. Этот канал передает частоты в диапазоне от 300 до 3400 Гц, таким образом, его полоса пропускания равна 3100 Гц. Хотя человеческий голос имеет гораздо более широкий спектр — примерно от 100 Гц до 10 кГц, — для приемлемого качества передачи речи диапазон в 3100 Гц является хорошим решением. Строгое ограничение полосы пропускания тонального канала связано с использованием аппаратуры уплотнения и коммутации каналов в телефонных сетях.

Дискретизация аналогового сигнала

Кодирование аналоговых сообщений. Чтобы хранить, передавать или обрабатывать информацию, находящуюся в аналоговой форме, ее необходимо специальным образом обработать — *оцифровать*. При этом аналоговый сигнал квантуется, т.е. производится оценка непрерывного сигнала ближайшими дискретными значениями (рис. 4.4).

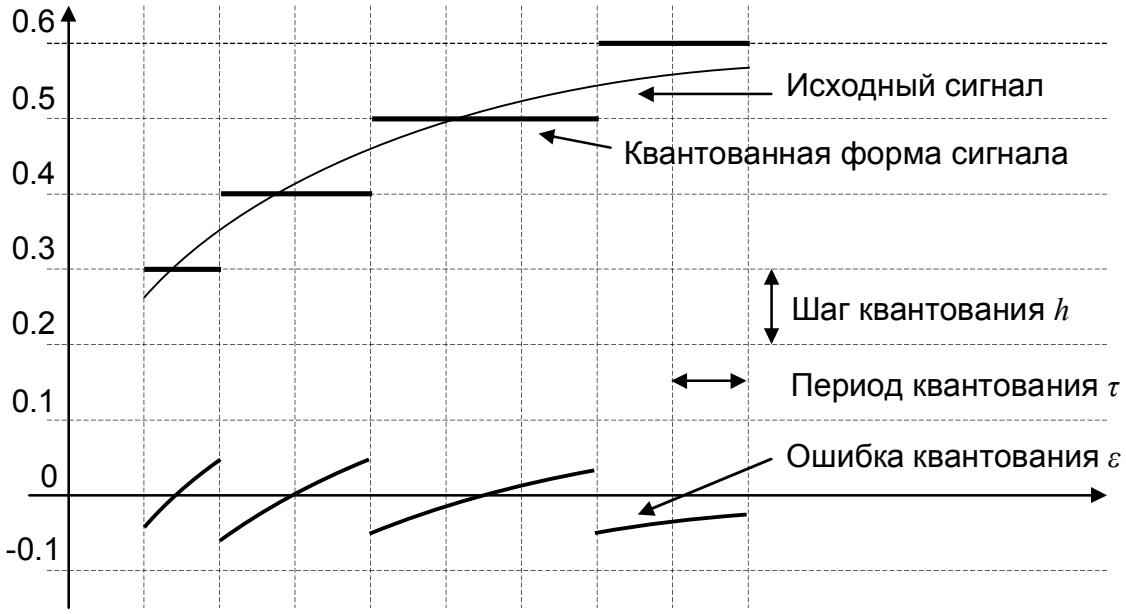


Рис. 4.4 — Дискретизация аналогового сигнала

Точность квантования. Разность между соседними квантованными значениями определяет *шаг квантования (h)*, а разницу между фактическим и квантованным значением называют *ошибкой квантования (*ε*).* Ошибка квантования ε по модулю не превосходит половину шага квантования *h*. С точки зрения передачи сигнала можно рассматривать ошибку квантования ε как шум при передаче квантованного сигнала.

Таким образом, показатель ***SNR*** (*Signal to Noise Ratio*) — *отношение сигнал-шум* можно вычислить как отношение мощности сигнала к мощности шума. Так как мощность пропорциональна квадрату амплитуды сигнала, то значение ошибки квантования можно вероятностно оценить, предположив, что величина ε в диапазоне *h* распределена равномерно:

$$\varepsilon_{ср} = \frac{1}{h}\int_{\frac{-h}{2}}^{\frac{h}{2}} \varepsilon^2 d\varepsilon = \frac{h^2}{12}.$$

Таким образом, мощность шума пропорциональна квадрату шага квантования *h*.

Очевидно, что величина шага квантования зависит от числа *q* разбиений исходной амплитуды сигнала на *q* дискретных значений (уровней квантования). Аналогично получаем, что ошибка квантования обратно пропорциональна квадрату *количества уровней квантования q*:

$$\varepsilon_{ср} = \frac{1}{2/q}\int_{\frac{-1}{q}}^{\frac{1}{q}} \varepsilon^2 d\varepsilon = \frac{1}{3\cdot q^2}.$$

Частота дискретизации. Очевидно, что адекватная дискретизация аналогового сигнала зависит от его частоты: чем выше частота исходного сигнала, тем с меньшими *периодами квантования* τ необходимо совершать замеры. Правило выбора предельного периода квантования τ при равномерной дискретизации с использованием модели сигнала с ограниченным спектром сформулировано академиком *В. А. Котельниковым*:

Любая непрерывная функция *F(t)*, спектр которой ограничен частотой $f_{\max}$ полностью определяется последовательностью своих дискретных значений F_k в моменты времени, отстоящие друг от друга на интервал τ:

$$\tau = \frac{1}{2\cdot f_{\max}} = \frac{\pi}{\omega_{\max}}.$$

Иными словами, теорема Котельникова определяет, что для адекватного представления непрерывной по времени функции *F(t)* при помощи дискретных отсчетов F_k частота дискретизации должна быть, по крайней мере, в два раза больше, чем самая высокочастотная гармоническая составляющая непрерывного сигнала $f_{\max}$.

Практически для передачи аналогового сигнала производится его *дискретизация* с частотой отсчетов *2F* и выполняется *им-*

пульсно-кодовая модуляция (***PCM — Pulse Code Modulation***) последовательности отсчетов. Импульсно-кодовая модуляция используется для передачи аналоговых сигналов по цифровым каналам связи. Этот вид модуляции сводится к измерению амплитуды аналогового сигнала в моменты времени, отстоящие друг от друга на τ, и к кодированию этих амплитуд цифровым кодом (величина τ определяется по теореме Котельникова).

В цифровых каналах ***ISDN*** за основу принята передача голоса с частотным диапазоном до 4 кГц, а кодирование производится восемью или семью битами. Отсюда, частота отсчетов (передачи байтов) равна 8 кГц, т.е. биты передаются с частотой 64 кГц (или 56 кГц при семибитовой кодировке).

При преобразовании амплитуды A аналогового сигнала в цифровой код $\{a_i\}$, $i = 1...n$ желательно учитывать нелинейность амплитудных характеристик приборов и иметь зависимость $\{a_i(A)\}$, $i = 1...n$, монотонно убывающей с ростом амплитуды.

Разновидностями импульсно-кодовой модуляции являются дельта-модуляция (***ДМ***), дифференциальная ДМ (***ДДМ***) и адаптивная ДМ (***АДДМ***). При использовании ДМ передаются разности амплитуд A_1 и A_2 соседних отсчетов, A_1 — амплитуда на входе модулятора, A_2 — амплитуда отсчета, которая соответствует переданному сигналу в предыдущем временном такте. Для представления разности передается только 1 бит (знак разности) поэтому нужна достаточно высокая частота отсчетов, чтобы не было «запаздывания» изменений передаваемого сигнала по сравнению с реальными изменениями. В ***ДДМ***, кроме того, знак разности $A_1 – A_2$ передается не постоянно, а только в момент пересечения величиной A_1 одного из уровней квантования. В ***АДДМ*** шаги отсчетов адаптируются к динамике изменения величины сигнала.

Цифровое кодирование

Как правило, среда не может передавать или принимать символы в их естественной форме. Требуется такое преобразование этой формы — кодирование, чтобы она наилучшим образом соответствовала используемой физической среде передачи. Такое преобразование называется *кодированием линии связи*, *цифровое кодирование* или *линейные коды*.

Символы в сообщениях могут относиться к алфавиту $\{a_i\}$, $i = 1...n$. Ставится задача кодирования любой последовательности символов алфавита $\{a_i\}$ элементами кода $\{b_j\}$, $j = 1...k$, $k \le n$, т. к. число k элементов кода ограничено сверху энергетическими соображениями. Так, если отношение сигнал/помеха для надежного различения уровня сигнала должно быть не менее заданного S, то наименьшая амплитуда для представления одного из k символов должна быть $S \cdot A$, где A — амплитуда помехи, а наибольшая амплитуда соответственно $S \cdot A \cdot k$. Поскольку мощность P передатчика пропорциональна квадрату амплитуды сигнала, то P должна превышать величину, пропорциональную $(SAk)^2$. В связи с этим распространено двоичное кодирование с $k = 2$. При двоичном кодировании сообщений с n типами букв, каждая из n букв кодируется определенной комбинацией *1* и *0* (например, код ASCII).

К цифровому кодированию физического уровня OSI предъявляются следующие требования:

- *требования по синхронизации* — код должен иметь такую форму, чтобы не возникало проблем с его синхронизацией. При необходимости код должен быть самосинхронизирующимся;
- *требования по наличию постоянной составляющей в спектре сигнала* — для передачи в спаренной среде переменного тока в сигнале должна присутствовать постоянная составляющая;
- *требования по энергетическому спектру* — спектр сигнала по мощности должен быть как можно меньше для увеличения КПД передачи;
- *требования по вероятности ошибок* — код должен обеспечивать, по возможности, наименьшую вероятность ошибок;
- *требования по прозрачности* — код должен работать при любой форме передаваемых символов;
- *требования по сложности* — кодирование и декодирование должно быть достаточно простым, чтобы уменьшить стоимость оборудования;
- *требования по уникальности* — процесс декодирования должен однозначно восстанавливать исходные данные.

На настоящее время разработано довольно много способов кодирования линии связи, но используются в реальном коммуникационном оборудовании лишь немногие — их мы и рассмотрим

далее. Существует несколько критериев классификации линейных кодов.

- По форме представления двоичных **0** и **1** линейные коды подразделяются на *коды с уровнями*, когда информация передается уровнями напряжения или тока сигнала, и *коды с переходами*, при которых информация передается *изменениями* уровня напряжения или тока (рис. 4.5).

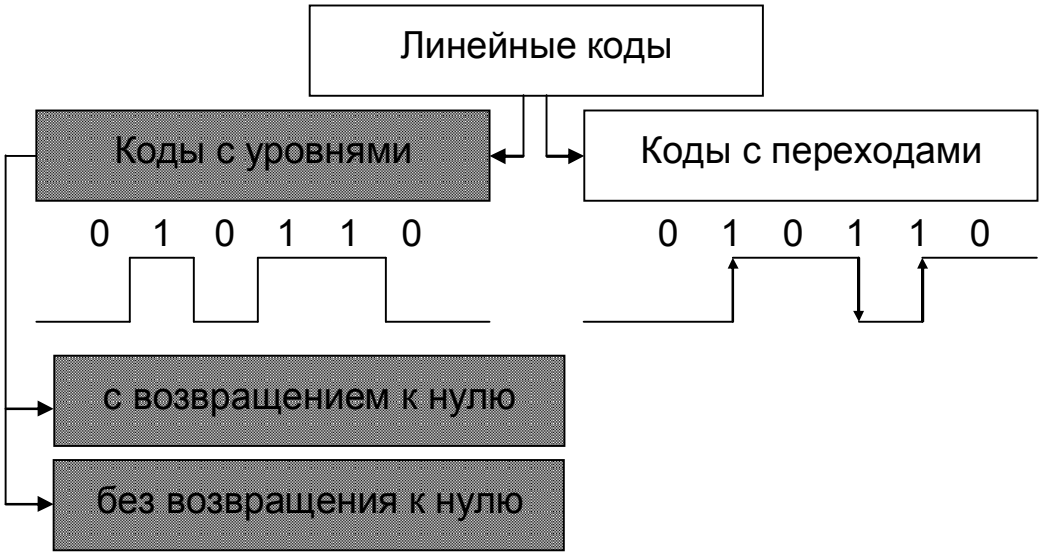


Рис. 4.5 — Линейные коды с уровнями и с переходами

- По наличию нулевого уровня линейные коды с уровнями подразделяются на коды *с возвращением к нулю* (*Return to Zero,* ***RZ***) — уровень возвращается к нулю в концекаждогокодового символа; и коды *без возвращения к нулю* (*Non Return to Zero,* ***NRZ***) — уровень импульсов в коде символов поддерживается на постоянном уровне (рис. 4.6).

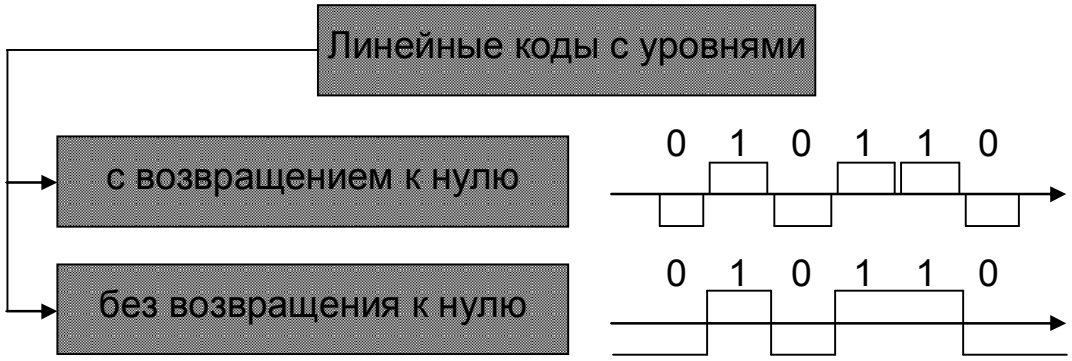


Рис. 4.6 — Линейные коды с уровнями с возвращением к нулю и без возвращения к нулю

• По полярности генерируемого на выходе сигнала линейные коды подразделяются на униполярные (имеющие только отрицательные или положительные значения) и биполярные (принимающие как положительные, так и отрицательные значения) см. рис. 4.7.

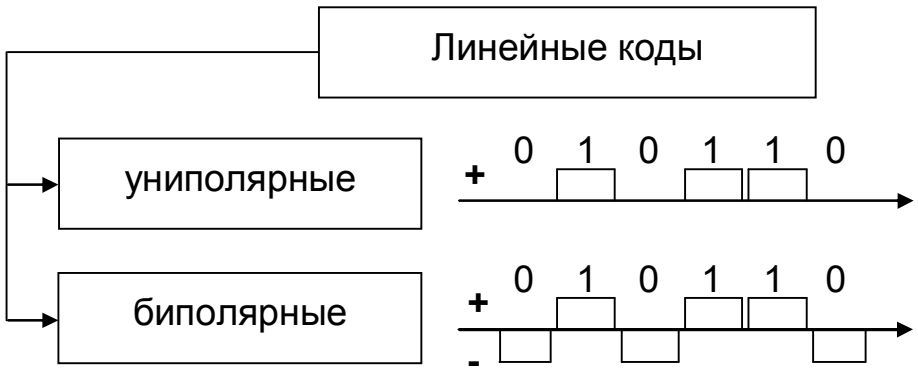


Рис. 4.7 — Униполярные и биполярные линейные коды

Потенциальный код с нулем. Самый очевидный способ кодирования состоит в задании бит двоичного сигнала последовательностью потенциалов (рис. 4.8, *а*): высокий уровень — для кодирования логической единицы и низкий — для нуля. Такой метод называется потенциальным кодом. Он так же используется при передаче данных между контроллерами компьютера.

Потенциальный код без возвращения к нулю. Другой метод потенциального кодирования (рис. 4.8, *б*), называется потенциальным кодированием без возвращения к нулю (***Non Return to Zero, NRZ***) из-за того, что при передаче последовательности единиц сигнал не возвращается к нулю.

Код NRZ довольно прост в реализации и имеет достаточно низкую частоту основной гармоники f_0, которая равна $N/2$ Гц, у других методов кодирования, например манчестерского, основная гармоника более высокочастотная.

Метод NRZ из-за двух потенциалов разной полярности обладает хорошей распознаваемостью ошибок, но не обладает свойством самосинхронизации, т.к. при передаче длинной последовательности единиц или нулей сигнал на линии не изменяется.

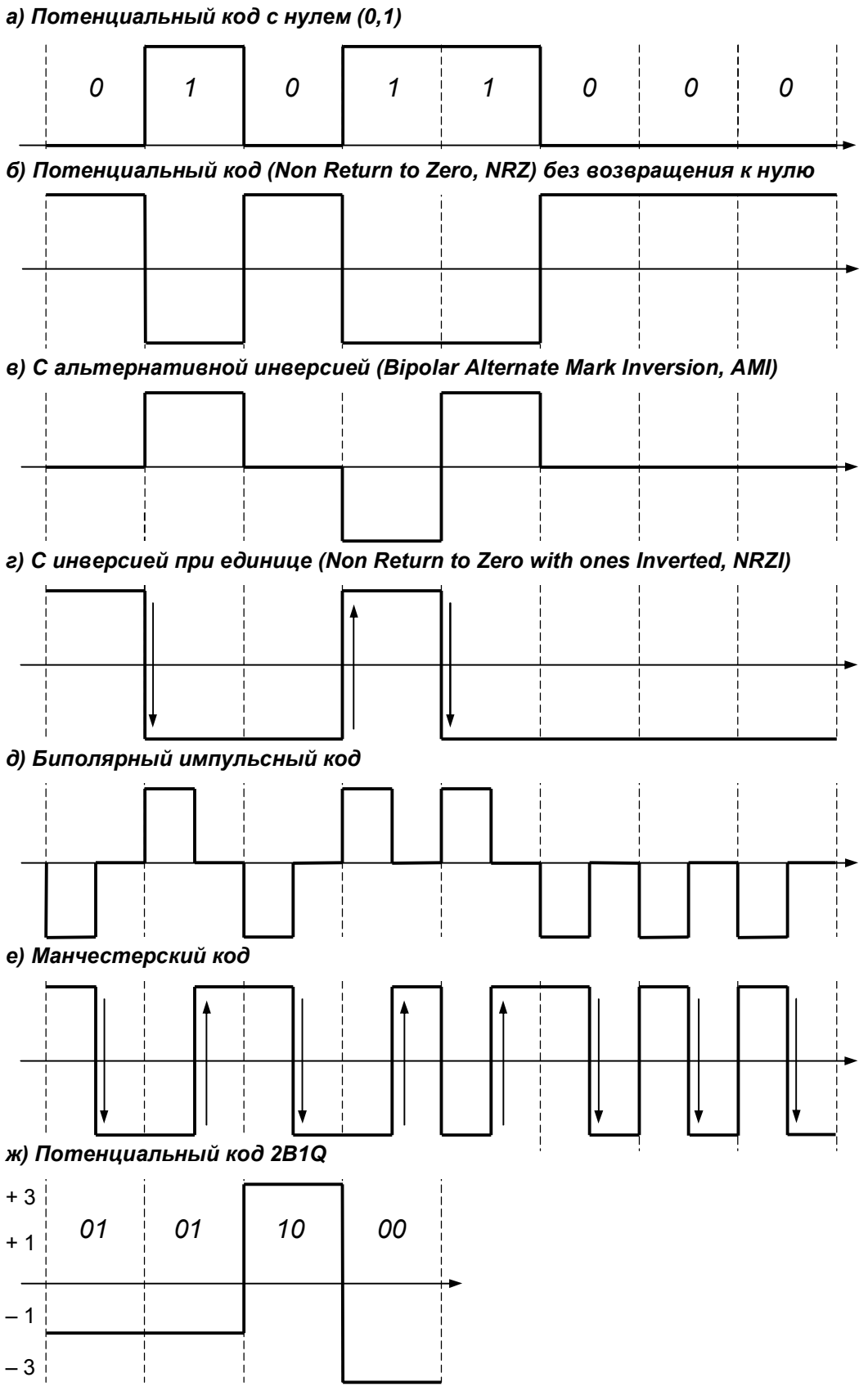


Рис. 4.8 — Методы кодирования физического уровня

Из-за этого приемник лишен возможности определять по входному сигналу моменты времени, когда нужно в очередной раз считывать данные. Даже при наличии высокоточного тактового генератора приемник может рассинхронизоваться с передатчиком, так как частоты двух генераторов никогда не бывают абсолютно одинаковыми. Поэтому при высоких скоростях обмена данными и длинных последовательностях единиц или нулей некоторое рассогласование тактовых частот может привести к ошибке в целый такт и, следовательно, некорректному считыванию бита. Другим серьезным недостатком метода NRZ является наличие низкочастотной составляющей спектра, которая приближается к нулю при передаче длинных последовательностей единиц или нулей. И поэтому линии связи, не обеспечивающие прямого гальванического соединения между приемником и источником, не используют этот вид кодирования. Из-за этого в чистом виде код NRZ в сетях не применяется, но используются его различные модификации, в которых устраняются как плохая самосинхронизация кода, так и постоянная составляющая.

Модификацией метода NRZ является метод ***биполярного кодирования с альтернативной инверсией*** (***Bipolar Alternate Mark Inversion, AMI***). В этом методе (рис. 4.8, *в*) используются три уровня потенциала — отрицательный, нулевой и положительный. Логический нуль кодируется нулевым потенциалом, а логическая единица — положительным или отрицательным уровнем, при этом потенциал каждой последующей единицы противоположен потенциалу предыдущей.

Код AMI ликвидирует проблемы постоянной составляющей и отсутствия самосинхронизации при передаче длинных последовательностей единиц, но не нулей. При передаче единиц сигнал на линии — это последовательность разнополярных импульсов (с тем же спектром, что и у кода NRZ, то есть без постоянной составляющей и с основной гармоникой $N/2$ Гц, где N — битовая скорость передачи данных). В случае длинной последовательности нулей сигнал вырождается в постоянный потенциал нулевой амплитуды.

В среднем, код AMI имеет более узкий спектр сигнала, чем NRZ, а значит, и более высокую пропускную способность: так, при передаче чередующихся единиц и нулей основная гармоника

f_0 имеет частоту $N/4$ Гц. Кроме того, код AMI предоставляет некоторые возможности по распознаванию ошибочных сигналов. Так, нарушение строгого чередования полярности сигналов говорит о ложном импульсе или исчезновении с линии корректного импульса. Сигнал с некорректной полярностью называется запрещенным сигналом (*signal violation*).

Так как в коде AMI используются не два, а три уровня сигнала на линии, то требуется увеличение мощности передатчика примерно на 3 дБ для обеспечения той же достоверности приема бит на линии.

Потенциальный код с инверсией при единице (***Non Return to Zero with ones Inverted, NRZI***) похож на AMI, но имеет только два уровня сигнала. При передаче нуля он передает тот же потенциал, который был установлен в предыдущем такте, то есть не меняет его, а при передаче единицы потенциал инвертируется на противоположный (рис. 4.8, *г*). Этот код удобен в тех случаях, когда использование третьего уровня сигнала весьма нежелательно, например, в оптических кабелях.

Улучшение потенциальных кодов. Для улучшения потенциальных кодов, подобных AMI и NRZI, используются два метода. Первый метод основан на ***добавлении в исходный код избыточных бит***, содержащих логические единицы. В этом случае длинные последовательности нулей прерываются, и код становится самосинхронизирующимся для любых передаваемых данных. Исчезает также постоянная составляющая, а значит, еще более сужается спектр сигнала. Но этот метод снижает полезную пропускную способность линии, так как избыточные единицы пользовательской информации не несут. Алгоритм добавления единиц известен, так что удаление их при приеме происходит в обратном порядке.

Другой метод основан на предварительном «перемешивании» исходной информации таким образом, чтобы вероятность появления единиц и нулей на линии становилась близкой. Устройства, или блоки, выполняющие такую операцию, называются ***скремблерами*** (*scramble* — свалка, беспорядочная сборка). При скремблировании используется известный алгоритм, поэтому приемник, получив двоичные данные, передает их на *дескремблер*, который восстанавливает исходную последовательность бит.

Избыточные биты при этом по линии не передаются. Оба метода относятся к логическому, а не физическому кодированию, так как они не определяют форму сигналов на линии.

Импульсные коды. Кроме потенциальных кодов в сетях используются импульсные коды, в которых данные представлены полным импульсом или же его частью — фронтом. Примером такого подхода является ***биполярный импульсный код***, в котором единица представлена импульсом одной полярности, а ноль — другой (рис. 4.8, *д*), а каждый импульс длится половину такта. У этого кода отличные самосинхронизирующими свойства, но при передаче длинной последовательности единиц или нулей может присутствовать постоянная составляющая. Кроме того, спектр у него шире, чем у потенциальных кодов. Так, при передаче всех нулей или единиц частота основной гармоники кода будет равна *N* Гц, что в два раза выше основной гармоники кода NRZ и в четыре раза выше основной гармоники кода AMI при передаче чередующихся единиц и нулей. Из-за слишком широкого спектра биполярный импульсный код используется редко.

В локальных сетях до недавнего времени самым распространенным методом кодирования был ***манчестерский код*** (рис. 4.8, *е*), применяемый в технологиях *Ethernet* и *Token Ring*.

В манчестерском коде для кодирования единиц и нулей используется изменение потенциала, то есть *фронт* импульса. При манчестерском кодировании каждый такт делится на две части. Информация кодируется перепадами потенциала, происходящими в середине каждого такта. Единица кодируется перепадом от низкого уровня сигнала к высокому, а ноль — обратным перепадом. В начале каждого такта может происходить служебный перепад сигнала, если нужно представить несколько единиц или нулей подряд. Так как сигнал изменяется по крайней мере один раз за такт передачи одного бита данных, то манчестерский код обладает хорошими самосинхронизирующими свойствами. Полоса пропускания манчестерского кода уже, чем у биполярного импульсного. У него также нет постоянной составляющей, а основная гармоника в худшем случае (при передаче последовательности единиц или нулей) имеет частоту *N* Гц, а в лучшем (при передаче чередующихся единиц и нулей) она равна *N/2* Гц, как и у кодов AMI или NRZ. Ширина полосы пропускания манчестер-

ского кода в полтора раза уже, чем у биполярного импульсного кода, а основная гармоника колеблется вблизи значения *3N/4* Гц. Манчестерский код имеет еще одно преимущество перед биполярным импульсным кодом — в нем используются два уровня сигнала, тогда как в импульсном коде — три.

Потенциальный код 2B1Q. На рис. 4.1, *ж* показан потенциальный код с четырьмя уровнями сигнала для кодирования данных. Это код ***2B1Q***, название которого отражает его суть — каждые два бита (*2B*) передаются за один такт сигналом, имеющим четыре состояния (*1Q*). Паре бит 00 соответствует потенциал –2.5 В, паре бит 01 соответствует потенциал –0,833 В, паре 11 — потенциал +0,833 В, а паре 10 — потенциал +2,5 В. Этот способ кодирования не защищает от длинных последовательностей одинаковых пар бит, так как при этом сигнал превращается в постоянную составляющую. При случайном чередовании бит спектр сигнала в два раза уже, чем у кода NRZ, так как при той же битовой скорости длительность такта увеличивается в два раза. Таким образом, с помощью кода 2B1Q можно по одной и той же линии передавать данные в два раза быстрее, чем с помощью кода AMI или NRZI. Однако для его реализации мощность передатчика должна быть выше, чтобы четыре уровня четко различались приемником на фоне помех.

Спектр сигнала

Методика построения спектра базируется на разложении исходного сигнала в функциональный ряд, например в ряд Фурье. Коэффициенты ряда Фурье однозначно определяют гармонические составляющие исследуемого сигнала. Коэффициенты a_j, b_j ряда Фурье вычисляются по формулам

$$a_j = \frac{1}{T}\int_0^T F(x)\cdot\cos(j\cdot x)\,dx,\quad b_j = \frac{1}{T}\int_0^T F(x)\cdot\sin(j\cdot x)\,dx$$

Так, например, на рис. 4.9 приведен потенциальный код и его спектральное разложение.

Если дискретные данные передаются с битовой скоростью *N* бит/с, то спектр состоит из постоянной составляющей нулевой частоты и бесконечного ряда гармоник с частотами *f, 3·f, 5·f, 7·f,...,* где *f = N/2*. Амплитуды этих гармоник убывают достаточно

медленно — с коэффициентами *1/3, 1/5, 1/7,...* от амплитуды гармоники *f*. В результате спектр потенциального кода требует для качественной передачи широкую полосу пропускания.

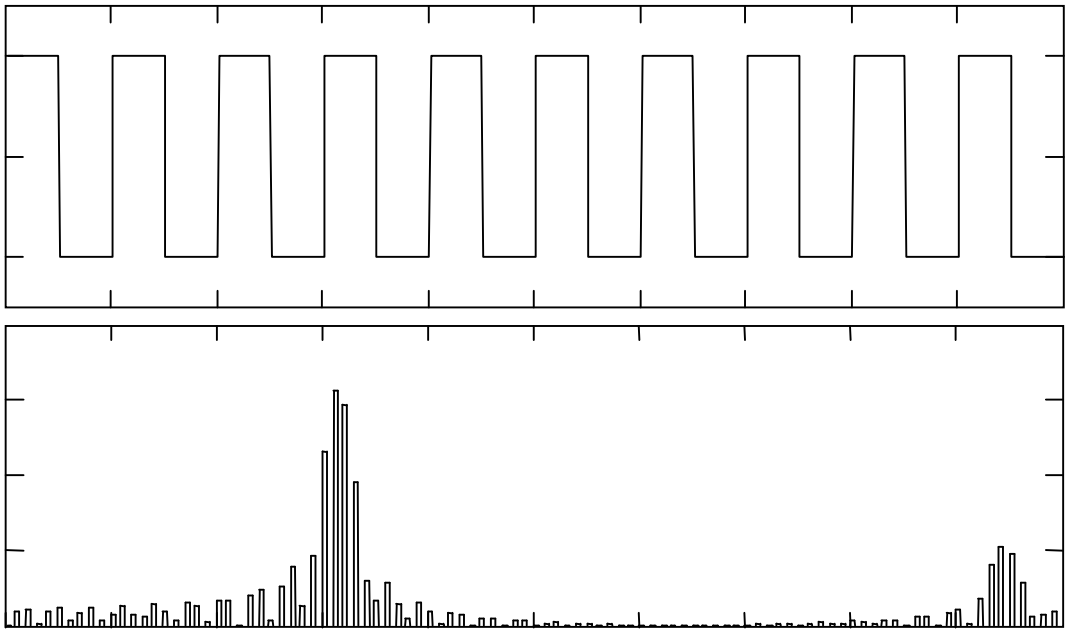


Рис. 4.9 — Равномерно чередующиеся логические нули и единицы

Кроме того, нужно понимать, что реально спектр сигнала постоянно меняется в зависимости от того, какие данные передаются по линии связи. Например, передача длинной последовательности нулей или единиц сдвигает спектр в сторону низких частот (см. рис. 4.10), а, в крайнем случае, — когда передаваемые данные состоят только из единиц (или только из нулей), спектр состоит из гармоники нулевой частоты. При передаче чередующихся единиц и нулей постоянная составляющая отсутствует. Поэтому спектр результирующего сигнала потенциального кода при передаче произвольных данных «плавает» в некоторой полосе частот от величины, близкой к 0 Гц, до примерно $7 \cdot f$ (гармониками с частотами выше $7 \cdot f$ можно пренебречь из-за их малого вклада в результирующий сигнал).

Для канала тональной частоты (300—3400 Гц) верхняя граница при потенциальном кодировании достигается для скорости передачи данных в 971 бит/с, а нижняя неприемлема для любых скоростей, так как полоса пропускания канала начинается с

300 Гц. Поэтому потенциальные коды на каналах тональной частоты никогда не используются.

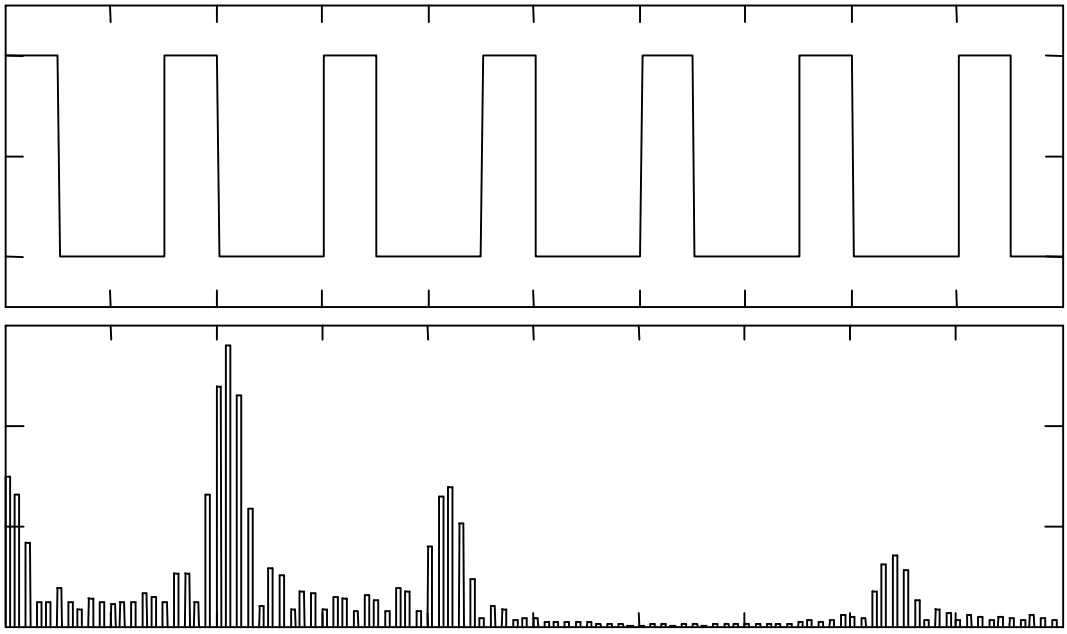


Рис. 4.10 — Дисбаланс между количеством логических нулей и единиц

В реальных системах передачи данных потенциальные коды, критичные к передаче последовательностей одинаковых символов — логических нулей или единиц (например, потенциальный, AMI, NRZI, 2B1Q) не используются. Вместо них применяются комбинированные методы, улучшающие спектральные характеристики потенциальных кодов при помощи добавления избыточных бит или скремблирования.

Поскольку негладкие функции (к которым относятся цифровые последовательности) теоретически имеет спектр бесконечной ширины, а для идеальной передачи цифрового сигнала необходимо учитывать все гармоники, то при реальной передаче данных часть информации о форме сигнала теряется. Для приемлемой распознаваемости сигнала достаточно корректно передать только значимые гармоники, например, для потенциального кода это частоты из диапазона от 0 до $7 \cdot f$ Гц, где $f = N/2$, а N [бит/с] — битовая скорость передачи данных.

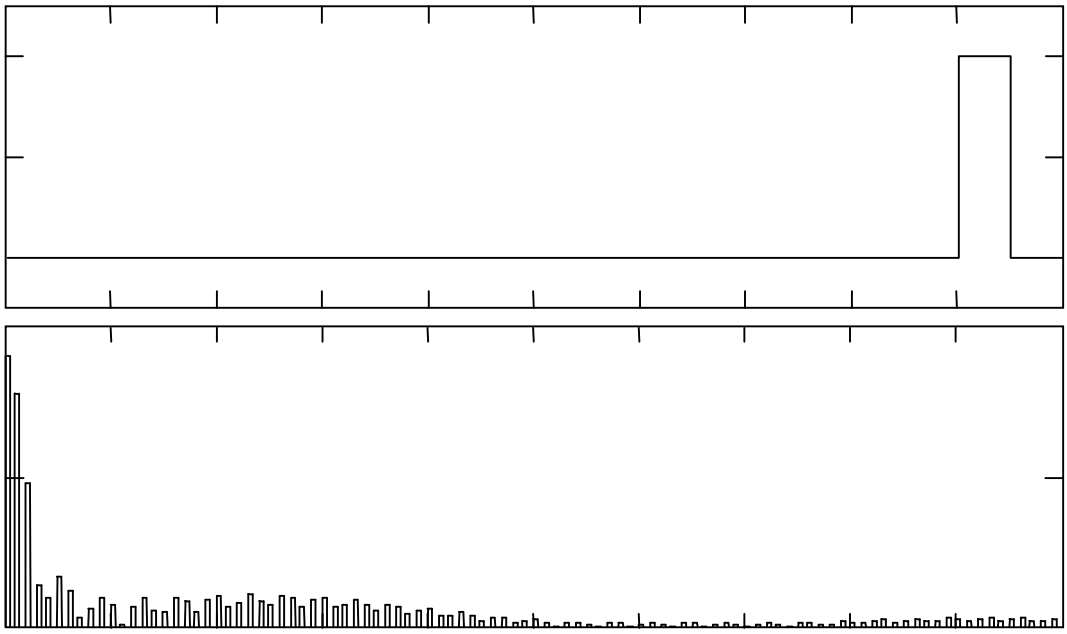


Рис. 4.11 — Дисбаланс между количеством логических нулей и единиц

Рассмотрим, для примера, потенциальный код, последовательность бит и спектральное разложение которого приведены на рис. 4.12.

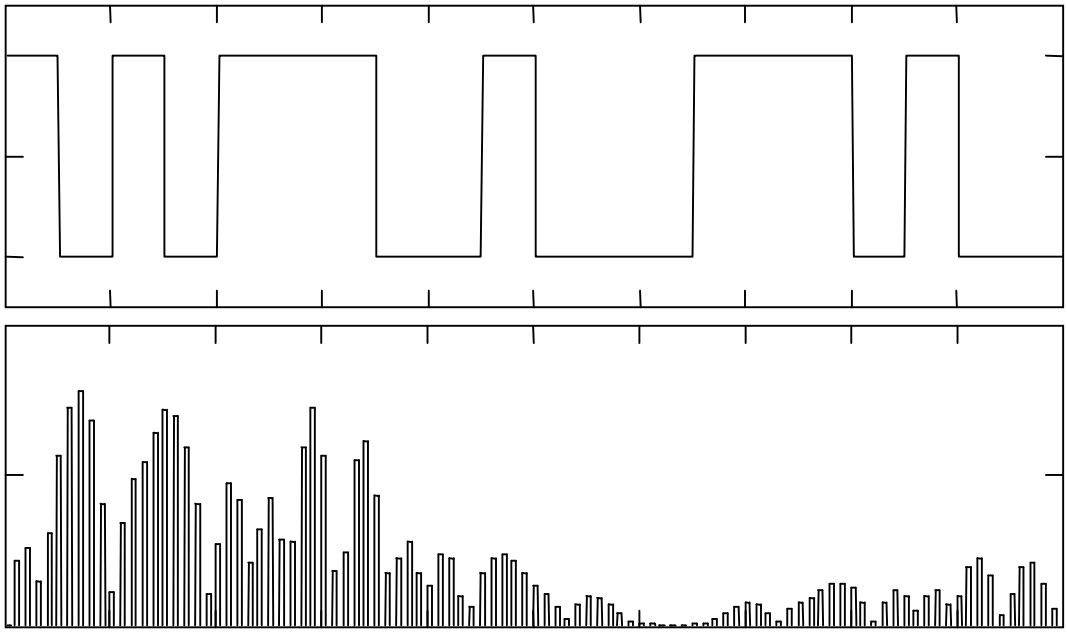


Рис. 4.12 — Исходный сигнал и его спектр

Восстановим теперь сигнал из набора гармоник — коэффициентов ряда Фурье. Просуммировав ряд Фурье по формуле

$$S(x) = \frac{a_0}{2} + \sum_{j=1}^{n} \left(a_j \cdot \cos(j \cdot x) + b_j \cdot \sin(j \cdot x) \right),$$

приблизительно получим результирующий сигнал, поступающий на приемник (рис. 4.13).

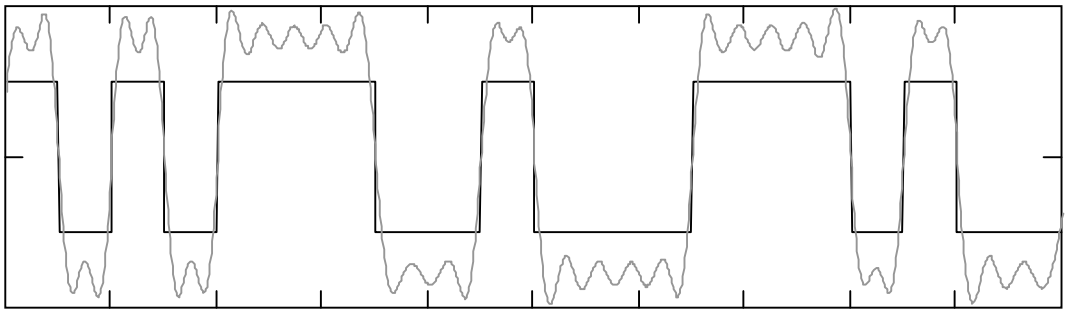


Рис. 4.13 — Исходный и результирующий сигнал

Если же какие-то из значимых гармоник сигнала не попадают в полосу пропускания, то эти гармоники будут искажены (а то и отброшены) линией связи. Таким образом, сигнал будет восстановлен не полностью.

При построении результирующего сигнала, приведенного на рис. 4.14, использовалась только та часть спектра, которая выделена серым цветом.

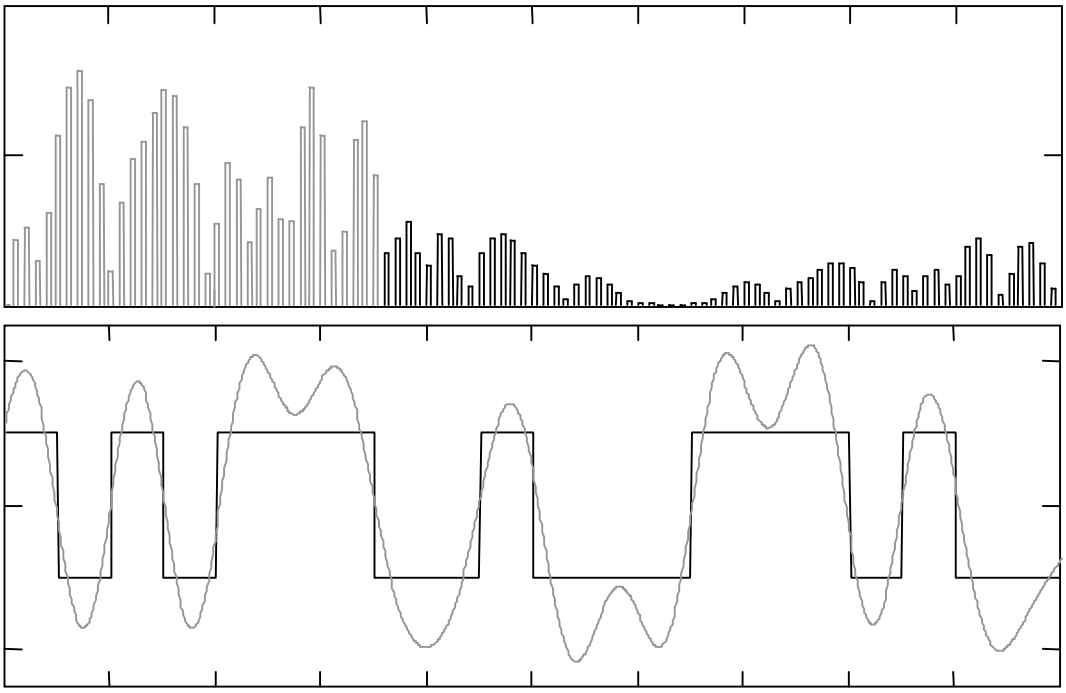


Рис. 4.14 — Исходный и результирующий сигнал
с отброшенными незначащими гармониками

Очевидно, что отброшенные гармоники (на рис. 4.14 выделены черным) не оказывают большого влияния, так как восстановленный сигнал может быть распознан с вероятностью близкой к единице.

Если же полоса пропускания канала передачи данных сдвинута в сторону, например, так, как показано на рис. 4.15 (серый цвет), то восстановленный сигнал, очевидно, корректно распознать не удастся.

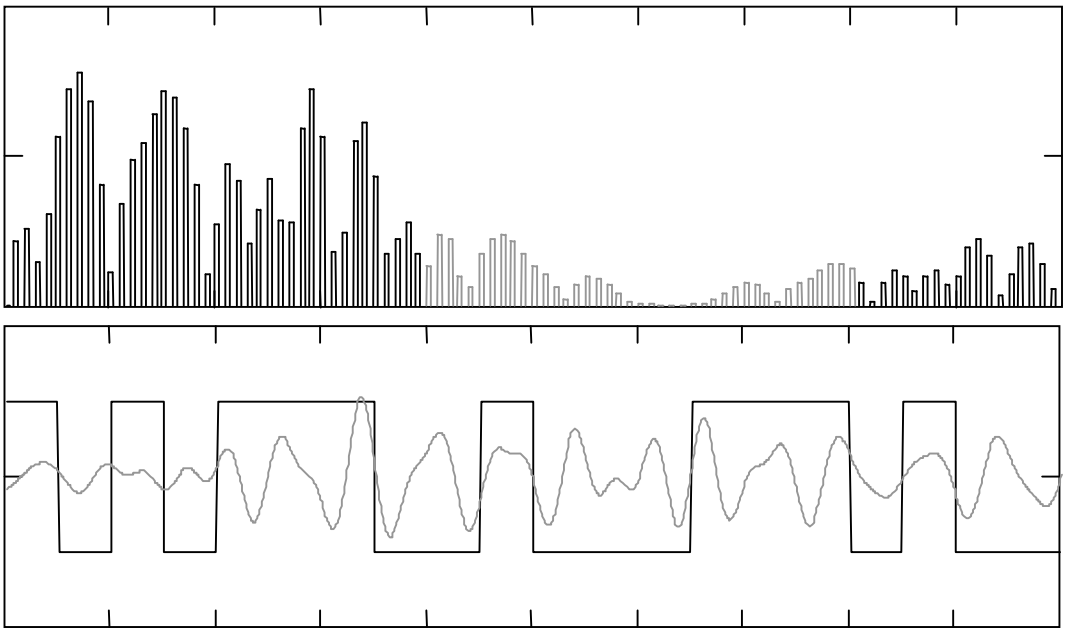


Рис. 4.15 — Исходный и результирующий сигнал
с узкой полосой пропускания канала

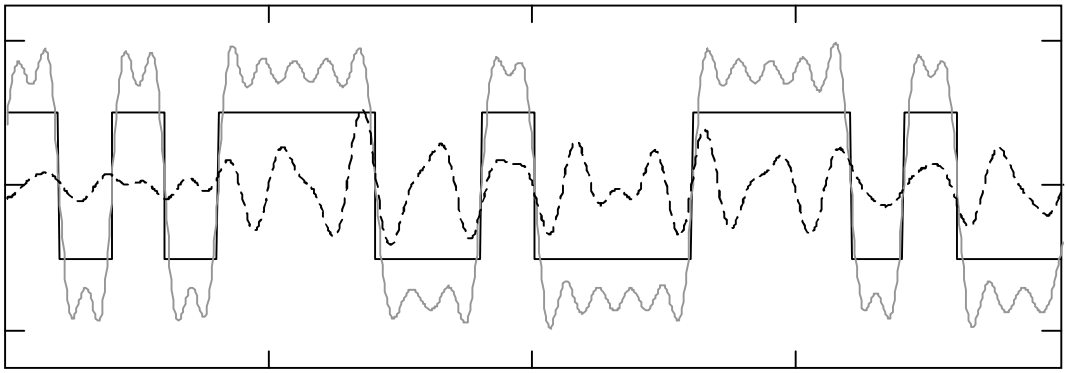


Рис. 4.16 — Сравнение результирующих сигналов
с различными полосами пропускания линии связи

Аналоговая модуляция сигнала

Аналоговая модуляция является таким способом физического кодирования, при котором информация кодируется изменением амплитуды, частоты или фазы синусоидального сигнала несущей частоты. Основные способы аналоговой модуляции показаны на рис. 4.19—4.21. На диаграмме (рис. 4.17*)* показана последовательность бит исходной информации, представленная потенциалами высокого уровня для логической единицы и потенциалом нулевого уровня для логического нуля. Такой способ кодирования называется *потенциальным кодом*, который часто используется при передаче данных между блоками компьютера.

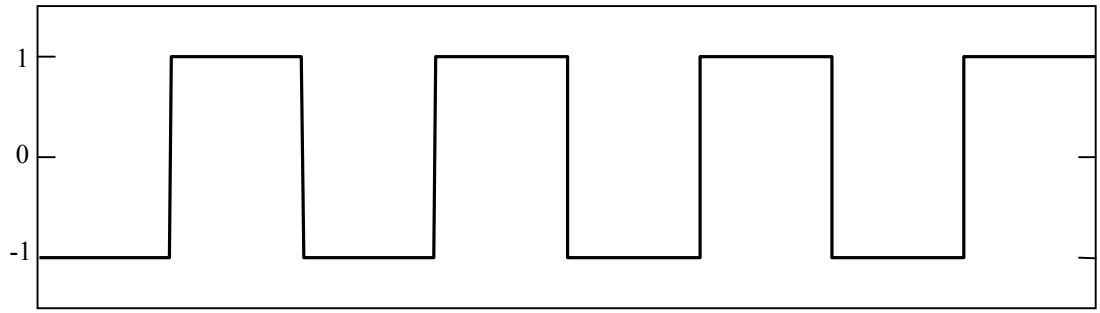


Рис. 4.17 — Потенциальный код

Потенциальный код обладает хорошей распознаваемостью, но, как видно и рис. 4.18, у него слишком широкий и медленно убывающий спектр.

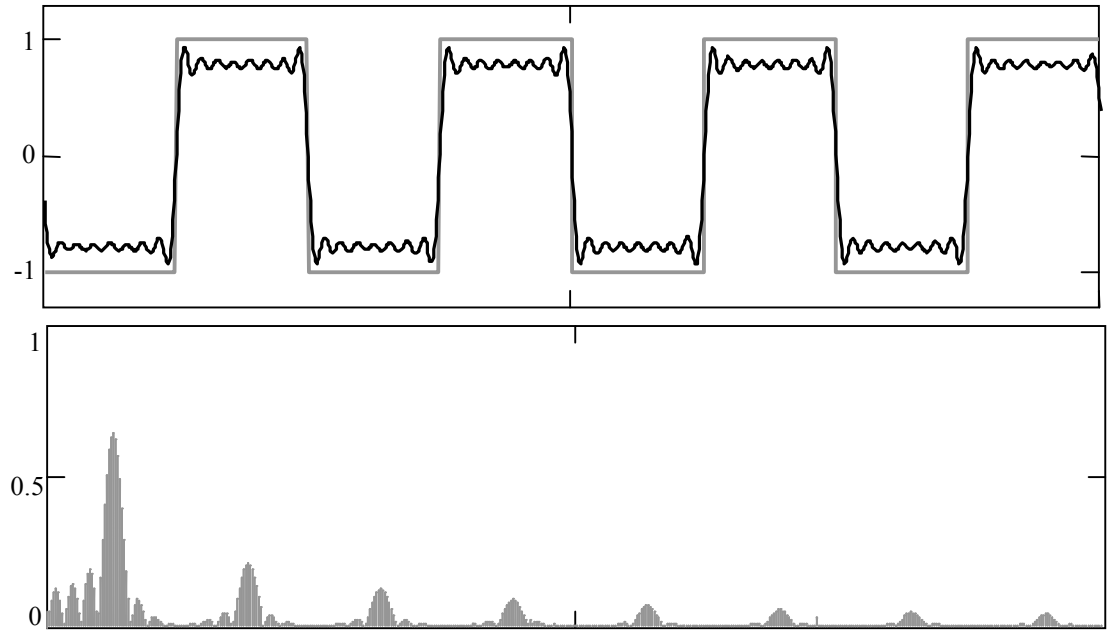


Рис. 4.18 — Спектр исходного потенциального кода

Для передачи дискретных данных по каналам с узкой полосой частот, типичным представителем которых является *канал тональной частоты* (общественные телефонные сети) применяется аналоговая модуляция. При амплитудной модуляции исходный цифровой сигнал кодируется непрерывным аналоговым гармоническим сигналом с фиксированной несущей гармоникой. На этот несущий сигнал накладывается информационная составляющая, которая и кодирует исходное информационное сообщение. Так как канал тональной частоты передает частоты в диапазоне от 300 до 3400 Гц (его полоса пропускания равна 3100 Гц), то и частоты информационной и несущей гармоник должны, разумеется, попадать в этот диапазон. Устройство, которое выполняет функции модуляции несущей синусоиды на передающей стороне и демодуляции на приемной стороне, носит название *модем* (модулятор-демодулятор).

Амплитудная модуляция. При амплитудной модуляции на выход модулятора поступает сигнал $V(t) = V_m \cdot \sin(\omega_V \cdot t - \varphi_V)$ и несущая $U(t) = U_m \cdot \sin(\omega_U \cdot t - \varphi_U)$. На выходе нелинейного элемента модулируются колебания:

$$U_{AM}(t) = U_m \cdot (1 + m \cdot \sin(\omega_V \cdot t - \varphi_V)) \cdot \sin(\omega_U \cdot t - \varphi_U), \quad (4.7)$$

где $m = V_m / U_m$ — коэффициент модуляции. На выходе модулятора в спектре сигнала присутствует несущая частота ω_U и две боковые частоты $\omega_U - \omega_V$ и $\omega_U + \omega_V$. Если сигнал занимает некоторую полосу частот, как это показано на рис. 4.18, то в спектре модулированного колебания появится две боковые полосы справа и слева от несущей (рис. 4.19).

При амплитудной модуляции во избежание искажений, называемых качанием фронта, необходимо выполнение условия $\omega_U >> \omega_V$. Соблюдение этого условия при стандартной (для среднескоростной аппаратуры передачи данных) несущей частоте 1700 Гц не может обеспечить информационные скорости выше 300 бит/с. Поэтому в модемах применяют дополнительное преобразование частоты: сначала производят модуляцию несущей, имеющей повышенную частоту, например $\omega_{НД}$= 10 кГц, затем с помощью фильтра выделяют спектр модулированного сигнала и с помощью преобразователя частоты переносят модулирующие

колебания на промежуточную частоту, например ω_U= 1700 Гц. Тогда при боковых полосах до 1400 Гц спектр сигнала согласуется с полосой пропускания телефонных линий.

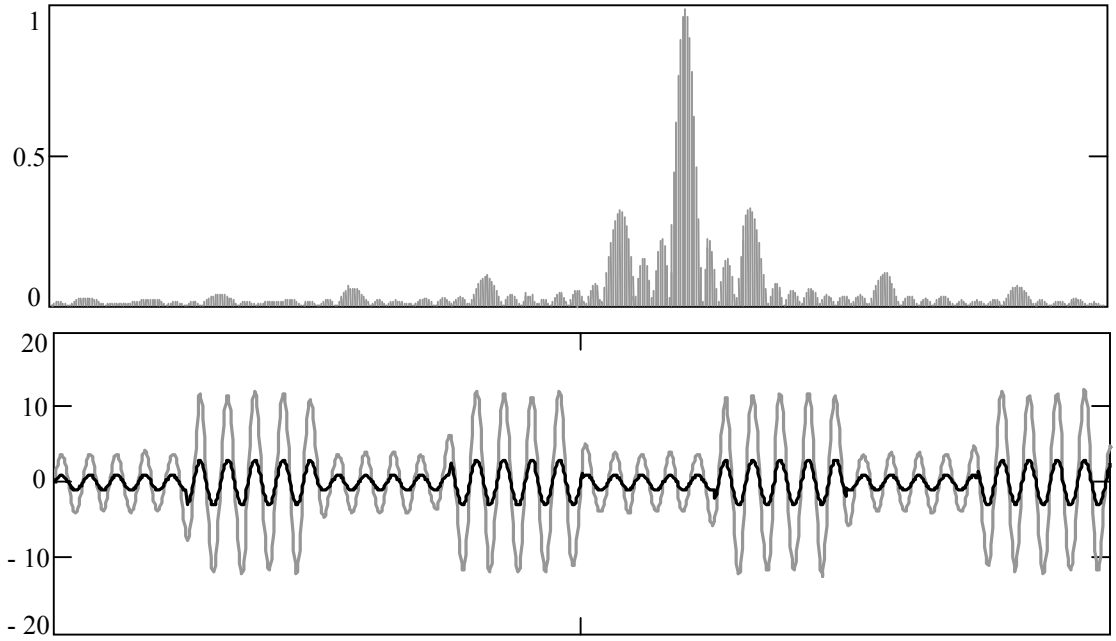


Рис. 4.19 — Спектр и восстановленный сигнал амплитудной модуляции

Однако достигаемые при этом скорости передачи данных остаются невысокими. Скорости передачи повышаются с помощью кавдратурно-амплитудной или фазовой модуляции за счет того, что вместо двоичных модулирующих сигналов используются дискретные сигналы с большим числом возможных значений.

В сравнительно простых модемах применяют ***частотную модуляцию*** (***FSK — Frequency Shift Keying***) (рис. 4.20) со скоростями передачи до 1200 бит/с. Так, если необходима дуплексная связь по двухпроводной линии, то возможно представление логических нуля и единицы в вызывном модеме частотами 980 Гц и 1180 Гц соответственно. При этом скорость передачи составляет 300 бод.

Спектр сигнала, модулирующего потенциальный код (рис. 4.17) частотным способом аналоговой модуляции и сама форма сигнала при частотной модуляции изображены на рис. 4.20.

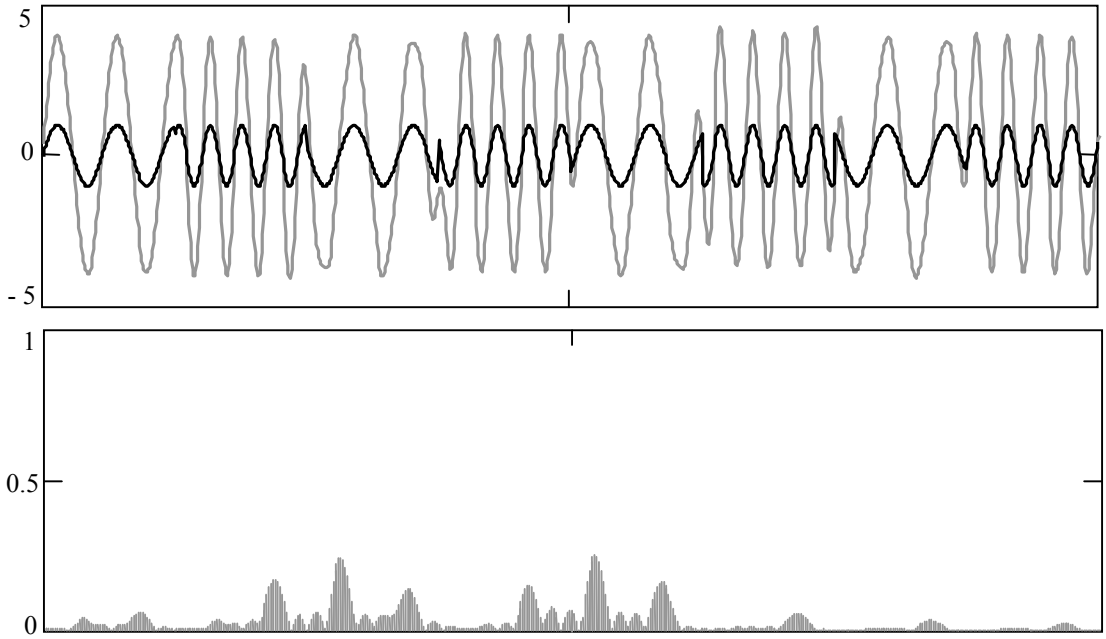


Рис. 4.20 — Частотная модуляция

Обычно для передачи сигнала об ошибке от приемника к передатчику нужен канал обратной связи. При этом требования к скорости передачи данных по обратному каналу могут быть невысокими. Тогда в полосе частот телефонного канала образуют обратный канал с частотной модуляцией, по которому со скоростью 75 бит/с передают единицу частотой 390 Гц и ноль — частотой 450 Гц.

Фазовая модуляция (***PSK — Phase Shift Keying***) двух уровней сигнала (нуля и единицы) осуществляется переключением между двумя несущими, сдвинутыми на полпериода друг относительно друга (рис. 4.21). Другой вариант фазовой модуляции — изменение фазы на $\pi/2$ в каждом такте при передаче нуля и на $3\pi/4$, если передается единица.

Исходный потенциальный код (рис. 4.17) после преобразования фазовой модуляции имеет спектр, приведенный на рис. 4.21.

Фазовая модуляция особенно часто применяется в комбинированных методах в сочетании с другими видами кодирования.

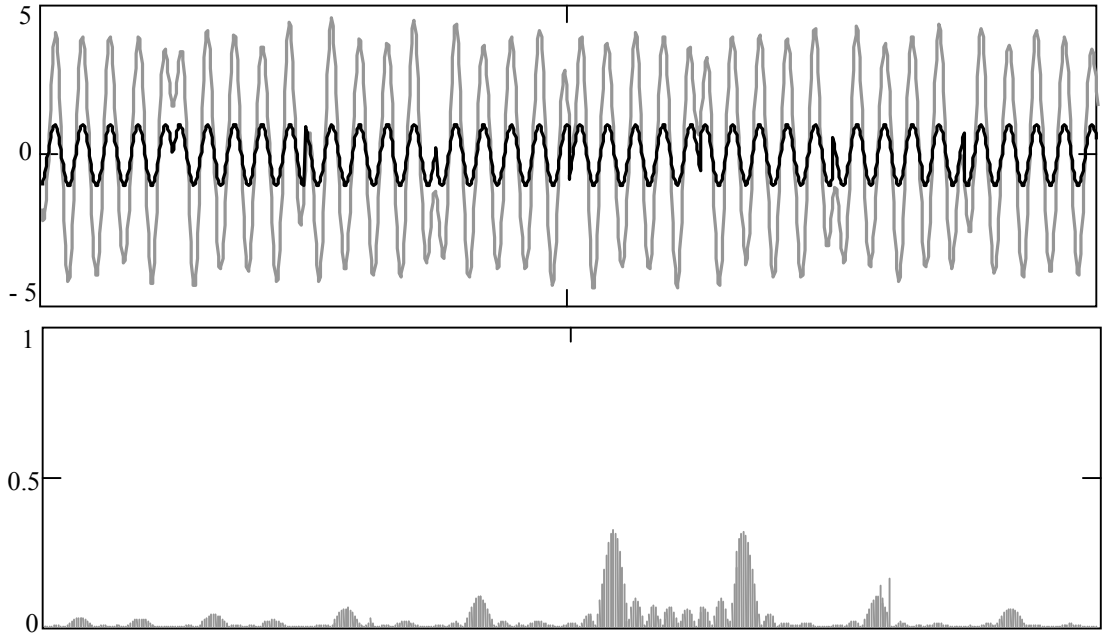


Рис. 4.21 — Фазовая модуляция

Спектр модулированного сигнала. Спектр результирующего модулированного сигнала зависит от типа модуляции и скорости модуляции, то есть желаемой скорости передачи бит исходной информации.

Рассмотрим неблагополучный, с точки зрения спектра, сигнал, изображенный на рис. 4.22 — при потенциальном кодировании; ввиду сравнительно большой последовательности логических единиц спектр этого сигнала смещается к нулю (рис. 4.23).

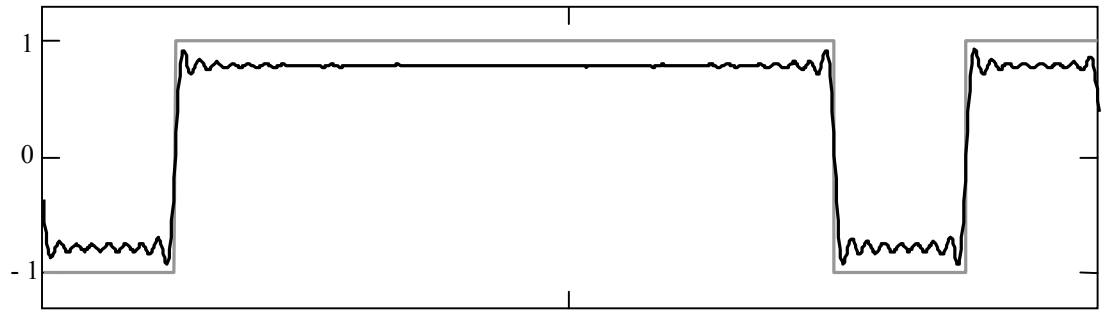


Рис. 4.22 — Сигнал со смещенным к нулю спектром

Из-за несогласованности спектра передаваемого сигнала и АЧХ аппаратуры передачи данных, приведенной на рис. 4.23,

сигнал не будет корректно распознан, ввиду отсутствия в спектре полученного сигнала некоторых значимых гармоник.

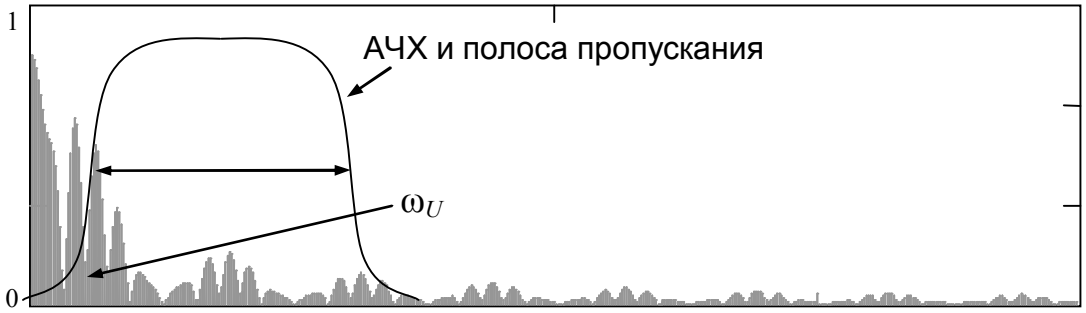


Рис. 4.23 — Спектр исходного сигнала не согласуется с полосой пропускания.

Применение аналоговой модуляции позволяет получить спектр сигнала сдвинутый вправо (рис. 4.24).

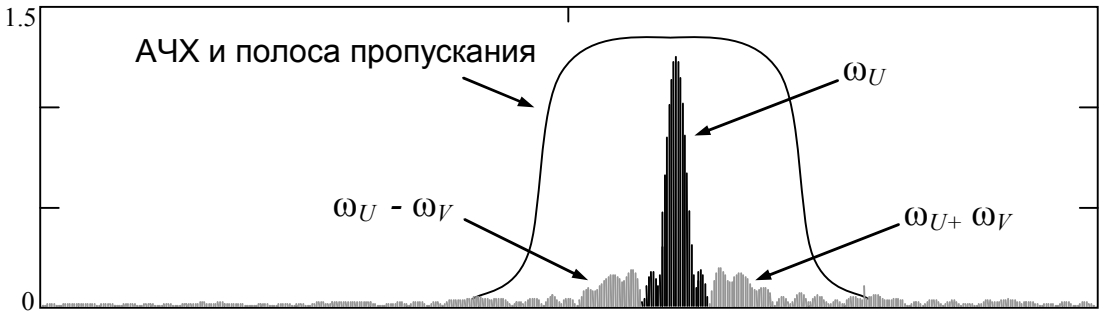


Рис. 4.24 — Спектр потенциального кода исходного сигнала при амплитудной модуляции

При амплитудной модуляции (рис. 4.24) спектр состоит из синусоиды несущей частоты ω_U и двух боковых гармоник: $\omega_U + \omega_V$ и $\omega_U - \omega_V$, где ω_V — частота изменения информационного параметра синусоиды, которая совпадает со скоростью передачи данных при использовании двух уровней амплитуды.

Частота ω_U определяет пропускную способность линии при данном способе кодирования. При небольшой частоте модуляции ширина спектра сигнала будет также небольшой (равной $2\omega_U$), поэтому сигналы не будут искажаться линией, если ее полоса пропускания будет больше или равна $2\omega_U$.

Для канала тональной частоты такой способ модуляции приемлем при скорости передачи данных не больше 3100/2 = = 1550 бит/с. Если же для представления данных используются 4 уровня амплитуды, то пропускная способность канала повышается до 3100 бит/с.

При фазовой и частотной модуляции спектр сигнала получается более сложным, чем при амплитудной модуляции, так как боковых гармоник здесь образуется более двух, но они также симметрично расположены относительно основной несущей частоты, а их амплитуды быстро убывают (рис. 4.20 и рис. 4.21). Поэтому эти виды модуляции также хорошо подходят для передачи данных по каналу тональной частоты.

Комбинированные аналоговые методы. Для повышения скорости передачи данных используют комбинированные методы модуляции, сочетающие те или иные методы аналоговой модуляции с методами цифрового кодирования. Наиболее распространенными являются методы ***квадратурной амплитудной модуляции*** (***Quadrature Amplitude Modulation, QAM***). Квадратурно-амплитудная модуляция, которую так же называют ***квадратурно-импульсной*** модуляцией) основанана передаче одним элементом модулированного сигнала *n* бит информации, где *n*= 4..8 (т.е. используется 16—256 дискретных значений амплитуды). Однако для надежного различения этих значений амплитуды требуется малый уровень помех (отношение сигнал/помеха не мене 12 дБ при *n*= 4).

При меньших отношениях сигнал/помеха лучше применять фазовую модуляцию с четырьмя или восемью дискретными значениями фазы для предоставления соответственно 2 или 3 бит информации. Тогда при скорости модуляции в 1200 бод (т.е. 1200 элементов аналогового сигнала в секунду, где элемент — часть сигнала между возможными сменами фаз) и четырехфазной модуляции скорость передачи данных равна 2400 бит/с. Используются также скорости передачи 4800 бит/с (при скорости модуляции 1600 бод и восьмифазной модуляции), 9600 бит/с и более при комбинации фазовой и амплитудной модуляций.

Однако из возможных комбинаций сигнала используются далеко не все. Например, в кодах *Треллиса* допустимы всего 6, 7 или 8 комбинаций из 16 для представления исходных данных, а

остальные комбинации являются запрещенными. Такая избыточность кодирования требуется для распознавания модемом ошибочных сигналов, являющихся следствием искажений из-за помех, которые на телефонных каналах, особенно коммутируемых, весьма значительны по амплитуде и продолжительны по времени.

Модемная связь. Физический уровень

Модуляция и демодуляция цифрового сигнала в аналоговый сигнала выполняется в устройстве, называемом ***модемом***. Модем выполняет функции аппаратуры окончания канала данных (DCE). В качестве оконечного оборудования обычно выступает компьютер, в котором имеется приемопередатчик — микросхема ***UART*** (***Universal Asynchronous Reciver/Transmitter).*** Приемопередатчик подключается к модему через один из последовательных портов компьютера и последовательный интерфейс RS-232C (скорость 9,6 Кбит/с на расстоянии до 15 м). Более высокая скорость (до 1000 Кбит/с на расстоянии до 100 м) обеспечивается интерфейсом RS-422, в котором используется две витые пары проводов с сопротивлениями на концах, образующие сбалансированную линию.

Для организации дуплексной (двунаправленной) модемной связи используются следующие способы:

- ***четырехпроводная линия связи*** — одна пара проводов для прямой, а другая — для обратной передачи;
- ***частотное разделение*** — прямая и обратная передачи ведутся на разныхчастотах, т.е. полоса для каждого направления сужается более чем вдвое по сравнению с полосой симплексной связи;
- ***эхо-компенсация*** — при установлении соединения с помощью посылки зондирующего сигнала определяются параметры (запаздывание и мощность) эха — отраженного собственного сигнала; в дальнейшем из принимаемого сигнала вычитается эхо собственного сигнала (см. рис. 4.25).

Протоколы модемной связи физического уровня определяют в телекоммуникационных технологиях способ модуляции, направленность передачи (дуплекс, полудуплекс, симплекс), ориентированность на выделенный или коммутируемый канал. Возможно отражение в протоколах и некоторых других специальных

характеристик передачи, например способа исправления ошибок и/или сжатия информации.

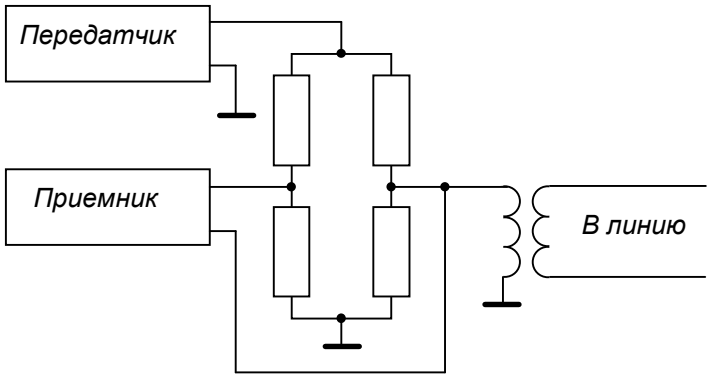


Рис. 4.25 — Эхо-компенсация

V.21. Протокол V.21 используется в простых модемах на 300 бит/с, в нем применена частотная модуляция с передачей по двухпроводной линии. Используется четыре частоты (980 и 1650 Гц — для представления **1** и 1180 и 1850 Гц для представления **0**) в прямом и обратном направлении передачи.

V.22. Протокол V.22 характеризуется скоростью 1200 бит/с, использует частотное разделение каналов (для дуплекса) и двукратная фазовая модуляция, т. е. фазовая модуляция с четырьмя значениями фазы. Несущие частоты 1200 и 2400 Гц.

V.27. Протокол V.27 использует трехкратную фазовую модуляцию (с восемью значениями фазы), достигается скорость 4800 бит/с по дуплексным выделенным каналам.

V.29. В протоколе V.29 скорость составляет 9600 бит/с, используется четырехпроводный выделенный канал для дуплекса (полудуплекс при коммутируемом двухпроводном канале). 16-позиционная квадратурно-импульсная модуляция. Эхо-компенсация.

V.32. В модемах, использующих спецификацию V.32, достигается скорость 9600 бит/с за счет фазовой модуляции и отфильтровывания эха собственного передатчика от принимаемых сигналов. Специальный контроллер автоматически снижает скорость передачи при наличии шумов в линии. Используется помехоустойчивое кодирование. На выделенных или коммутируемых

линиях применена 16-позиционная квадратурно-импульсная модуляция.

V.34. Высокоскоростные модемы строятся в соответствии с протоколом V.34, где скорости составляют от 2,4 до 28,8 Кбит/с с шагом 2,4 Кбит/с. Протокол предусматривает адаптацию передачи под конкретную обстановку, изменяя несущую в пределах 1600..2000 Гц, а так же автоматическое предварительное согласование способов модуляции в вызывающем и вызывном модемах. Дуплекс, 256-позиционная квадратурно-импульсная модуляция. Эхо-компенсация.

На выделенных телефонных линиях с интенсивным трафиком часто применяются четырехпроводные линии для дуплексной и двухпроводные для полудуплексной связи (протоколы V.23, V.26, V.27, V.29).

4.4 Цифровое кодирование (канальный уровень)

К задачам методов цифрового кодирования относят задачи выбора такого способа кодирования, который одновременно достигал бы нескольких целей:

- при одной и той же битовой скорости имел бы наименьшую ширину спектра результирующего сигнала;
- обеспечивал синхронизацию между передатчиком и приемником;
- обладал способностью распознавать ошибки;
- обладал низкой стоимостью реализации.

Более узкий спектр сигналов позволяет на одной и той же линии (с одной и той же полосой пропускания) добиваться более высокой скорости передачи данных. Кроме того, часто к спектру сигнала предъявляется требование отсутствия постоянной составляющей, то есть наличия постоянного тока между передатчиком и приемником. В частности, применение различных трансформаторных схем гальванической развязки препятствует прохождению постоянного тока.

Реальный канал способен ухудшать качество передаваемого сообщения, добавляя к нему ошибки. Для предотвращения этого приемнику необходимо обнаружить ошибку и, при возможности, исправить ее. Если же исправить обнаруженную ошибку прием-

ник не в состоянии, то запросить переданное сообщение заново. Распознавание и коррекцию искаженных данных сложно осуществить средствами физического уровня, поэтому чаще всего эту работу берут на себя протоколы, лежащие выше: канальный, сетевой, транспортный или прикладной. С другой стороны, распознавание ошибок на физическом уровне экономит время, так как приемник может не ждать полного помещения дефективного кадра в буфер, а отбраковывать его сразу при распознавании ошибочных бит внутри кадра. Тем не менее, задача обнаружения и исправления ошибок чаще всего относится к методам кодирования канального уровня.

Синхронизация передатчика и приемника нужна для того, чтобы приемник точно знал, в какой момент времени необходимо считывать новую информацию с линии связи. Эта проблема в сетях решается сложнее, чем при обмене данными между близко расположенными устройствами, например между блоками внутри компьютера или же между компьютером и принтером.

Асинхронное и синхронное кодирование

При посылке сообщения в линию связи передаваемые данные представляются электрическими сигналами, с точки зрения канального уровня, не имеет значения, какими именно. На физическом уровне сигнал кодируется потенциальными или импульсными кодами, что так же не принципиально для алгоритмов канального уровня. А вот правильно распознать полученные символы и правильно закодировать сообщение — это задача канального уровня.

Для правильного распознавания позиций символов в передаваемом сообщении получатель должен знать границы передаваемых элементов сообщения. На небольших расстояниях хорошо работает схема, основанная на отдельной тактирующей линии связи (рис. 4.26), так что информация снимается с линии данных только в момент прихода тактового импульса.

В сетях использование этой схемы вызывает трудности из-за неоднородности характеристик проводников в кабелях. На больших расстояниях неравномерность скорости распространения сигнала может привести к тому, что *i*-тый тактовый импульс придет настолько позже или раньше соответствующего *i*-того сигна-

ла данных, что бит данных будет пропущен или считан повторно. Другой причиной, по которой в сетях отказываются от использования тактирующих импульсов, является экономия проводников в дорогостоящих кабелях.

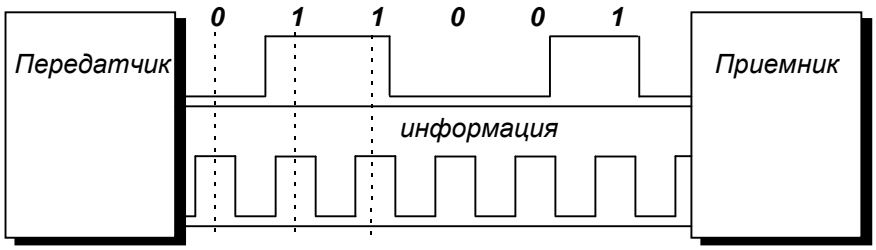


Рис. 4.26 — Синхронизация приемника и передатчика

Поэтому в сетях применяются так называемые ***самосинхронизирующиеся коды***, сигналы которых несут для передатчика указания о том, в какой момент времени нужно осуществлять распознавание очередного бита (или нескольких бит, если код ориентирован более чем на два состояния сигнала). Любой значительный перепад сигнала — ***фронт*** — может служить хорошим указанием для синхронизации приемника с передатчиком.

При использовании синусоид в качестве несущего сигнала результирующий код обладает свойством самосинхронизации, так как изменение амплитуды (фазы, частоты) несущего сигнала дает возможность приемнику определить момент появления входного кода.

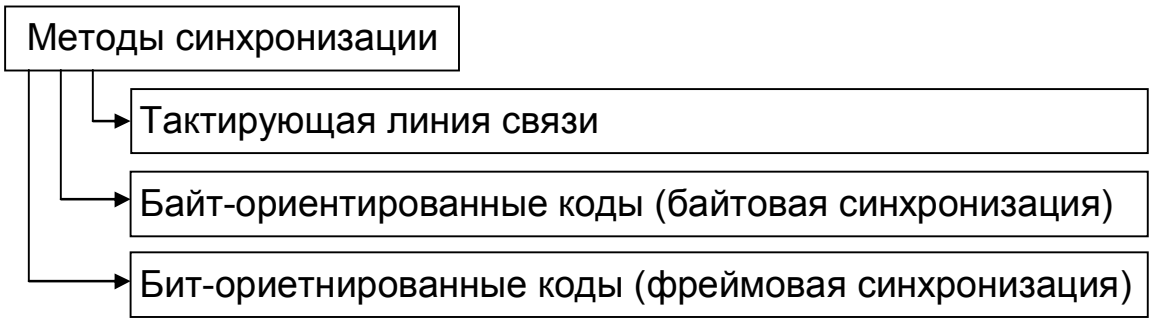


Рис. 4.27 — Методы синхронизации. Канальный уровень

В асинхронном режиме применяют коды, в которых явно выделены границы каждого символа (байта) специальными стар-

товыми и стоповыми символами. Подобные побайтно выделенные коды называют ***байт-ориентированными***, а способ передачи — *байтовой синхронизацией*. Однако это увеличивает количество битов, не относящихся собственно к сообщению.

В синхронном режиме синхронизм поддерживается во время передачи всего информационного блока без обрамления каждого байта. Такие коды называют ***бит-ориентированными***. Для входа в синхронизм нужно обозначить границы лишь всего передаваемого блока информации с помощью специальных начальной и конечной комбинации байтов (обычно это двухбайтовые комбинации). В этом случае синхронизация называется *блочной или фреймовой*. Для определения границ текстового блока (текст состоит только из печатаемых символов) можно использовать символы, отличающиеся от печатаемых. Для обрамления двоичных блоков применяют специальный символ (обозначим его DLE), который благодаря *стаффингу* становится уникальным. Уникальность заключается в том, что если DLE внутри блока, то сразу за ним вставляется еще один DLE. Приемник будет игнорировать каждый второй идущий подряд символ DLE. Если же DLE встречается без дублирования, то это граница блока.

Манчестерский код

Для кодирования информации наибольшее распространение получили *самосинхронизирующиеся коды*, так как при этом отпадает необходимость иметь дополнительную линию для передачи синхросигналов между узлами сети. В ЛВС чаще других применяют ***манчестерский код*** (рис. 4.28). Самосинхронизация обеспечивается благодаря формированию синхроимпульсов из перепадов, имеющихся в каждом такте манчестерского кода. В зависимости от переходов на каждой из двух синхросерий **A** и **B** манчестерский код способен передавать четыре символа: два из них **0** и **1** предназначены для двоичного кодирования сообщения, а другие два — **j** и **k** — для обеспечения синхронизации и обрамления блоков передаваемых данных.

Представленная на рис. 4.28 разновидность манчестерского кода используется при байт-ориентированном кодировании (при этом каждый байт, состоящий из символов **1** и **0**, обрамляется

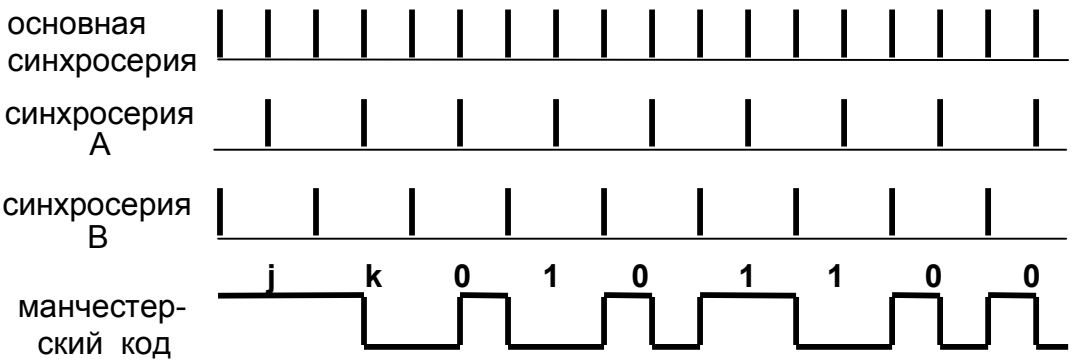


Рис. 4.28 — Манчестерское кодирование

символами **j** и **k**). В этом случае станция, получившая полномочия, начинает передавать серию сигналов **jkjkjkjk...** для того, чтобы станция-получатель могла войти в синхронизм с передающей станцией. После нескольких пар **jk** начинают передаваться байты самого сообщения. Различение четырех возможных значений сигнала выполняется в соответствии с таблицей:

	j	**k**	**1**	**0**
Синхросерия **A**	-	+	-	+
Синхросерия **B**	-	-	+	+

В случае бит-ориентированного манчестерского кода после входа в синхронизм не нужно обрамлять байты символами **j** и **k**, т.е. используется двузначное кодирование. Часто используется так же манчестерский код, в котором **1** представляется положительным, а **0** — отрицательным перепадом.

Протоколы канального уровня модемной связи

Протокол V.42 относится к стандартам, устанавливающим способы защиты от ошибок, а V.42bis, кроме того, регламентирует способы сжатия данных. Наряду с протоколами семейства V.42 для коррекции ошибок применяют протоколы **MNP** (***Microcom Network Protocol***).

Протокол V.42 является вариантом протокола **HDLC** (***High-level Data Link Control***) — бит-ориентированный базовый протокол. Установление соединения (вход в протокол) происходит в асинхронном байт-ориентированном режиме. Запрос на соединение осуществляется посылкой двухбайтовых сигналов *ODP*. Для

соединения необходимо согласте приемниеа в виде посылки ответа *ADP*. После чего образуется соединение и осуществляется переход в синхронный бит-ориентированный режим. В начале соединения передаются управляющие, а затем информационные кадры.

Протоколы канального уровня для модемной связи. Центральное место среди канальных протоколов телекоммуникаций занимают протоколы передачи файлов по телефонным каналам.

Функции канальных протоколов: управление потоком данных, координация работы передатчика с приемником. Различают протоколы по способам обнаружения и исправления ошибок, по реакции на возникновение ошибок (стартостопные и конвейерные), по способам защиты от несанкционированного доступа.

Стартостопный протокол характеризуется тем, что прежде чем посылать новый кадр информации, передатчик ждет подтверждения о правильном получении приемником предыдущего кадра, в ***конвейерных протоколах*** такое подтверждение может быть получено после передачи нескольких кадров. В последнем случае меньше задержка на ожидание подтверждений (квитанций), но больше затраты на повторную пересылку в случае ошибок.

Защита от несанкционированного доступа может реализовываться как аппаратно в модеме, так и в связной (коммутационной) программе.

В протоколах канального уровня различают командный режим и режим обмена данными. Некоторые действия, вызываемые командами ***командного режима***:

- имитация снятия трубки и ответ на вызов;
- имитация снятия трубки и набора номера (после того, как связь установится, модем переходит в режим обмена данными);
- переход из дуплексного режима в полудуплексный;
- отключение внутреннего динамика модема, и т.д.

Стандартом «де-факто» стал набор команд, реализуемых фирмой Hayes в своих модемах, это так называемые АТ-команды или Hayes-команды.

Основой для многих протоколов модемной связи стал протокол ***XModem***. Операции, выполняемые в режиме обмена данными приведены на рис. 4.29.

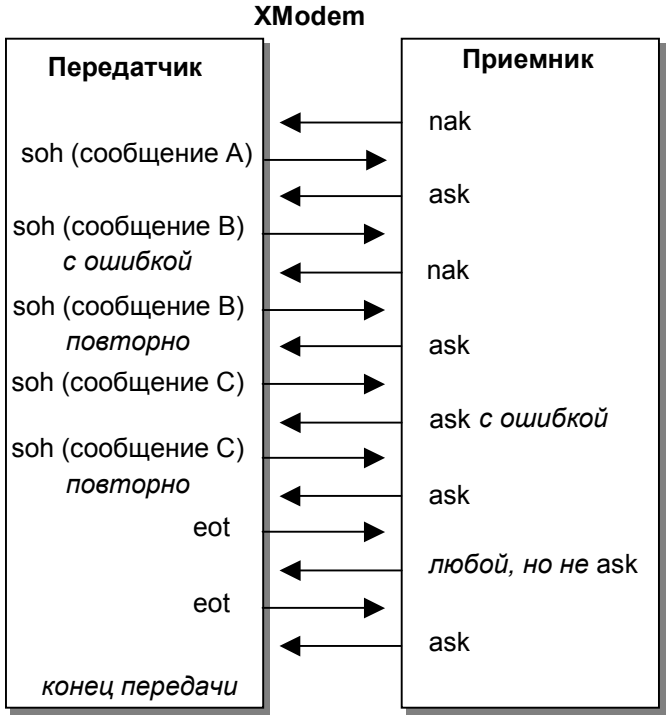


Рис. 4.29 — Процесс передачи сообщений
по протоколу XModem

В базовом варианте протокола XModem используется стартстопное управление, размер одного блока сообщения (пакета) равен 128 байт и 1 байт отводится под контрольную сумму. В варианте ***XModemCRC*** реализован более жесткий контроль ошибок за счет использования циклического кода с 16-разрядной проверяющей комбинацией. В варианте ***XModem1K*** дополнительно введено увеличение блока до 1024 байт при малой частоте ошибок.

В варианте ***YModem*** по сравнению с XModemCRC разрешена групповая передача файлов. В наиболее распространенном протоколе ***ZModem*** используется ***конвейерное управление*** (иначе называемое ***оконным управлением***) длина пакета автоматически меняется от 64 до 1024 байт в зависимости от качества канала. Если на приемном конце ZModem не поддерживается, то автоматически протокол переходит в YModem. Прерванная передача продолжается с места прерывания.

4.5 Способы контроля правильности передачи данных

При (канальном) кодировании происходит процесс преобразования элементов сообщения в соответствующие им числа (***кодовые символы***). Каждому элементу сообщения присваивается определенная совокупность кодовых символов, которая называется *кодовой комбинацией*. Совокупность кодовых комбинаций, обозначающих дискретные сообщения, образует *код*. Правило кодирования может быть выражено *кодовой таблицей*, в которой приводятся алфавит кодируемых сообщений и соответствующие им кодовые комбинации. Множество возможных кодовых символов называется *кодовым алфавитом*, а их количество ***m* — *основанием кода*** ($m = 2$ для двоичного кодирования). Кодирование ***N* *элементов сообщения*** сводятся к правилам записи N различных чисел в m-ичной системе счисления. Число разрядов ***n***, образующих кодовую комбинацию, называется *значностью кода*, или ***длиной кодовой комбинации***.

Коды, у которых все комбинации имеют одинаковую длину, называют ***равномерными***. Для равномерного кода число возможных комбинаций равно $\boldsymbol{m^n}$. Примером такого кода является пятизначный *код Бодо*, содержащий пять двоичных элементов ($m = 2$, $n = 5$). Число возможных кодовых комбинаций равно $2^5 = 32$, что достаточно для кодирования всех букв алфавита. Применение равномерных кодов не требует передачи разделительных символов между кодовыми комбинациями.

Неравномерные коды характерны тем, что у них кодовые комбинации отличаются друг от друга не только взаимным расположением символов, но и их количеством. Это приводит к тому, что различные комбинации имеют различную длительность. Типичным примером неравномерных кодов является *азбука Морзе*, в котором символы **0** и **1** используются только в двух сочетаниях — как одиночные (**1** и **0**) или как тройные (**111** и **000**). Сигнал, соответствующий одной единице, называется точкой, трем единицам — тире. Символ **0** используется как знак, отделяющий точку от тире, точку от точки и тире от тире. Совокупность **000** используется как разделительный знак между кодовыми комбинациями.

Управление правильностью (помехозащищенностью) передачи информации выполняется с помощью помехоустойчивого кодирования. Различают коды, обнаруживающие ошибки, и корректирующие коды, которые дополнительно к обнаружению еще и исправляют ошибки. Помехозащищенность достигается с помощью введения избыточности. Устранение ошибок с помощью ***корректирующих кодов*** (такое управление называют ***Forward Error Control***) реализуют в симплексных каналах связи. В дуплексных каналах достаточно применения кодов**,** ***обнаруживающих*** ошибки (***Feedback or Backward Error Control***), так как сигнализация об ошибке вызывает повторную передачу от источника. Это основные методы, используемые в информационных сетях.

Коды, у которых все возможные кодовые комбинации используются для передачи информации, называются *простыми*, или *кодами без избыточности*. В простых равномерных кодах превращение одного символа комбинации в другой, например **1** в **0** или **0** в **1**, приводит к появлению новой комбинации, т.е. к ошибке передачи.

Корректирующие коды

Корректирующие коды строятся так, что для передачи сообщения используются не все кодовые комбинации m^n, а лишь некоторая часть их (так называемые *разрешенные кодовые комбинации*). Тем самым создается возможность обнаружения и исправления ошибки при неправильном воспроизведении некоторого числа символов. Корректирующие свойства кодов достигаются избыточностью кодовых комбинаций.

Декодирование состоит в восстановлении сообщения по принимаемым кодовым символам. Устройства, осуществляющие кодирование и декодирование, называют соответственно *кодером* и *декодером*. Как правило, кодер и декодер выполняются физически в одном устройстве, называемым *кодеком*.

Расстоянием Хэмминга или ***кодовым расстоянием d(A, B)*** между двумя кодовыми последовательностями $A = \{a_1, a_2,..., a_n\}$ и $B = \{b_1, b_2,..., b_n\}$ будем называть число позиций, в которых символы этих последовательностей не совпадают. Кодовое расстояние вычисляется по формуле:

$$d(A,B) = A \otimes B = \sum_{i=1}^{n} (a_i \otimes b_i).$$

Пример: A = { **1**0**1**110 }, B = { **0**0**0**110 }, тогда $d(A, B) = 2$.

Говорят, что в канале произошла ***ошибка кратности q***, если в кодовой комбинации q символов приняты ошибочно. Легко видеть, что кратность ошибки есть не что иное, как расстояние Хэмминга между переданной и принятой кодовыми комбинациями, или, иначе, вес вектора ошибки. Рассматривая все разрешенные кодовые комбинации и определяя кодовые расстояния между каждой парой, можно найти наименьшее из них $d = min\ d(A; B)$, где минимум берется по всем парам разрешенных комбинаций. Это *минимальное кодовое расстояние* является важным параметром кода. Очевидно, что для простого кода $d = 1$.

Обнаруживающая способность кода характеризуется следующей теоремой. Если код имеет $d > 1$ и используется декодирование по методу обнаружения ошибок, то все ошибки кратностью $q < d$ обнаруживаются. Что же касается ошибок кратностью $q \geq d$, то одни из них обнаруживаются, а другие нет.

Исправляющая способность кода при этом правиле декодирования определяется следующей теоремой. Если код имеет $d > 2$ и используется декодирование с исправлением ошибок по наименьшему расстоянию, то все ошибки кратностью $q < d/2$ исправляются. Что же касается ошибок большей кратности, то одни из них исправляются, а другие нет.

Задача кодирования состоит в выборе кода, обладающего максимально достижимым d. Впрочем, такая формулировка задачи неполна. Увеличивая длину кода n и сохраняя число кодовых комбинаций M, можно получить сколь угодно большое значение d. Но такое «решение» задачи не представляет интереса, так как с увеличением n уменьшается возможная скорость передачи информации от источника. Если длина кода n задана, то можно получить любое значение d, не превышающее n, уменьшая число комбинаций M.

Поэтому задачу поиска наилучшего кода (в смысле максимального d) следует формулировать так: при заданных M и n найти код длины n, содержащий M комбинаций и имеющий наибольшее возможное d. В общем виде эта задача в теории кодиро-

вания не решена, хотя для многих значений n и M ее решения получены.

На первый взгляд помехоустойчивое кодирование реализуется весьма просто. В память кодирующего устройства (кодера) записываются разрешенные кодовые комбинации выбранного кода и правило, по которому с каждым из M сообщений источника сопоставляется одна из таких комбинаций. Данное правило известно и декодеру. Получив от источника определенное сообщение, кодер отыскивает соответствующую ему комбинацию и посылает ее в канал. В свою очередь, декодер, приняв комбинацию, искаженную помехами, сравнивает ее со всеми M комбинациями списка и отыскивает ту из них, которая ближе остальных к принятой.

Однако даже при умеренных значениях n такой способ не является удовлетворительным. Покажем это на примере. Пусть выбрана длина кодовой комбинации $n = 100$, а *скорость кода* примем равной 0.5 (число информационных и проверочных символов равно). Тогда число разрешенных комбинаций кода будет $2^{50} \approx 10^{15}$. Соответственно, размер таблицы будет $100 \times 10^{15} = 10^{17}$ бит, что составит примерно 10^{16} байт = 10000 Тбайт.

Таким образом, применение достаточно эффективных (а значит, и достаточно длинных) кодов при табличном методе кодирования и декодирования технически невозможно в настоящее время. Поэтому основное направление теории помехоустойчивого кодирования заключается в поисках таких классов кодов, для которых кодирование и декодирование осуществляются не перебором таблицы, а с помощью некоторых регулярных правил, определенных алгебраической структурой кодовых комбинаций.

Линейные коды

Линейными называются такие двоичные коды, в которых множество всех разрешенных блоков является линейным пространством относительно операции поразрядного сложения по модулю 2.

Если записать k линейно-независимых блоков в виде k строк, то получится матрица размером $n \times k$, которую называют *производящей* или ***порождающей матрицей*** кода G. Множество линейных комбинаций образует линейное пространство, содер-

жащее 2^k блоков, т.е. линейный код, содержащий 2^k блоков длиной n, обозначают (n, k). При заданных n и k существует много различных (n, k)-кодов с различными кодовыми расстояниями d, определяемых различными порождающими матрицами. Все они имеют *избыточность* $\varepsilon_k = 1 - k/n$ или *относительную скорость* $R_k = k/n$.

Чаще всего применяют ***систематические*** линейные коды, которые строят следующим образом. Сначала строится простой код длиной k, т.е. множество всех k-последовательностей двоичных символов, называемых *информационными*. Затем к каждой из этих последовательностей приписывается $r = n - k$ *проверочных* символов, которые получаются в результате некоторых линейных операций над информационными символами.

Простейший систематический код $(n, n - 1)$ строится добавлением к комбинации из $n - 1$ информационных символов одного проверочного, равного сумме всех информационных символов по модулю 2. Такой код $(n, n - 1)$ имеет $d = 2$ и позволяет обнаружить одиночные ошибки и называется *кодом с одной проверкой на четность*.

Преимуществом линейных, в частности систематических, кодов является то, что в кодере и декодере не нужно хранить большие таблицы всех кодовых комбинаций, а при декодировании не нужно производить большое количество сравнений.

Однако, для получения связи высокой надежности следует применять коды достаточно большой длины. Применение систематического кода в общем случае, хотя и позволяет упростить декодирование по сравнению с табличным способом, все же при значениях n порядка нескольких десятков не решает задачу практической реализации.

Совершенные и квазисовершенные коды

Совершенными (плотно упакованными) называют коды, в которых выполняются соотношения:

$$\sum_{j=0}^{m} C_n^j \cdot (b-1)^j = b^r - 1,$$

где m — максимальная кратность исправляемых ошибок, b — основание кода а r — число проверочных символов. Они отличают-

ся тем, что позволяют исправлять все ошибки кратностью m или меньше и ни одной ошибки кратности больше m.

Число известных совершенных кодов ограничено кодами Хэмминга значности $n = \frac{b^r - 1}{b - 1}$ и бинарным циклическим кодом Голея.

Квазисовершенными принято называть коды, исправляющие все ошибки кратности m и $A \le C_n^j \cdot (b-1)^j$ ошибок кратности $m+1$ при условии, что

$$\sum_{j=0}^{m} C_n^j \cdot (b-1)^j + A = b^r + 1.$$

Класс квазисовершенных кодов значительно шире, чем класс плотно упакованных кодов. Совершенные и квазисовершенные коды обеспечивают максимум вероятности правильного приема комбинации при равновероятных ошибках в канале связи.

Обнаруживающие коды

Итак, кодовым расстоянием $d(A, B)$ между двумя кодовыми последовательностями называют называть число позиций, в которых символы этих последовательностей не совпадают. Ошибкой кратности q называется ситуация, когда в кодовой комбинации q символов приняты ошибочно. Кратность ошибки, очевидно, есть кодовое расстояние между переданной и принятой кодовыми комбинациями. Рассматривая все разрешенные кодовые комбинации и определяя кодовые расстояния между каждой парой, можно найти наименьшее из них $d = min\ d(A; B)$, где минимум берется по всем парам разрешенных комбинаций.

Для всех корректирующих кодов минимальное кодовое расстояние $d > 1$. Для обнаружения ошибки кратности t требуется, чтобы $d = t + 1$, а для ее исправления — $d = 2t + 1$. С увеличением значения d растет корректирующая способность кода, которая количественно может быть выражена как вероятность обнаружения или исправления ошибок различных типов. В зависимости от назначения и возможностей помехозащищенных кодов различают коды *самокорректирующиеся* (позволяющие автоматически

исправлять ошибки) и *самоконтролирующиеся* (позволяющие автоматически обнаруживать наиболее вероятные ошибки).

Рассмотрим основные коды, обнаруживающие ошибки. Простейшим является систематический код ($n, n - 1$), получаемый добавлением к комбинации информационных символов одного проверочного, равного сумме всех информационных символов по модулю 2. Такой код позволяет обнаружить одиночные ошибки и называется *кодом с одной проверкой на четность*. Этот метод, разумеется, не является успешным для обнаружения ошибок кратности более 2.

Основным среди обнаруживающих методов являются, конечно, циклические коды.

Циклические коды

Поиск способов кодирования, при которых сложность декодера растет не экспоненциально, а лишь как некоторая степень n привел к появлению наиболее оптимальных — линейных кодов. В классе линейных систематических двоичных кодов это — *циклические коды*. Циклические коды просты в реализации и при невысокой избыточности обладают хорошими свойствами обнаружения ошибок. Циклические коды получили очень широкое распространение как в технике связи, так и в компьютерных средствах хранения информации. В зарубежных источниках циклические коды обычно называют *избыточной циклической проверкой* (*CRC*, *Cyclic Redundancy Check*).

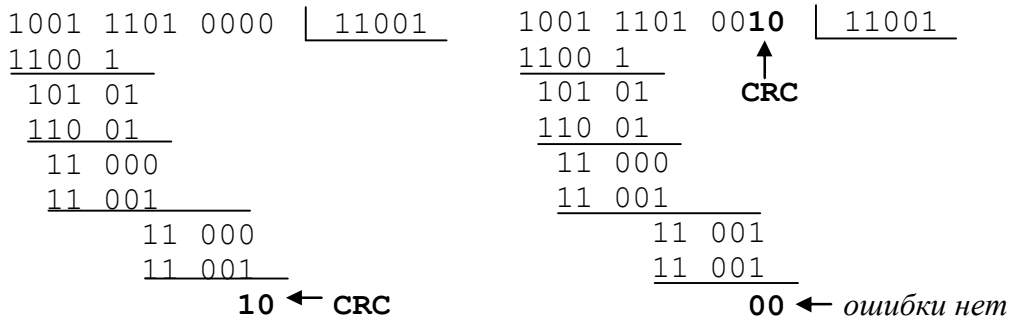


Рис. 4.30 — Циклический код. Пример обнаружения ошибки

Метод ***обнаруживающего циклического кодирования*** заключается в умножении исходного кода на порождающий многочлен $g(x)$, а декодирование — в делении на $g(x)$. Если остаток от

деления не равен нулю, то произошла ошибка. Сигнал об ошибке поступает на передатчик, что вызывает повторную передачу.

Методика построения циклических кодов позволяет разрабатывать и корректирующие коды, основанные на том же принципе.

Идея построения ***корректирующего циклического кода*** сводится к тому, что полином $s(x)$, представляющий информационную часть кодовой комбинации, нужно преобразовать в полином $p(x)$ степени не более $n - 1$, который без остатка делится на порождающий полином $g(x)$ (неприводимый многочлен) степени $n - k$.

Обнаружение ошибок при циклическом кодировании сводится к делению принятой кодовой комбинации на тот же образующий полином, который использовался для кодирования. Если ошибок в принятой комбинации нет (или они такие, что передаваемую комбинацию превращают в другую разрешенную), то деление на образующий полином производится без остатка. Наличие остатка свидетельствует о присутствии ошибок.

При использовании в циклических кодах декодирования с исправлением ошибок остаток от деления играет роль *синдрома*. Нулевой синдром указывает на то, что принятая комбинация является разрешенной. Всякому ненулевому синдрому соответствует определенная конфигурация ошибок, которая и исправляется.

Однако обычно в системах связи исправление ошибок при использовании циклических кодов не производится, а при обнаружении ошибок выдается запрос на повтор испорченной ошибками комбинации. Такие системы называются *системами с обратной связью* и будут рассмотрены ниже.

Линейный код $S(x)$ называют циклическим, если для любого кодового слова $[x_0\, x_1 \ldots\, x_{n-1}\, x_n]$ циклическая перестановка символов $[x_n\, x_0\, x_1 \ldots\, x_{n-1}]$ также является кодовым словом.

Перейдем от векторного описания кодов к *полиномиальному*. Последовательности символов основного алфавита $\{x_i\}$, составляющие сообщения и кодовые слова мы будем интерпретировать как коэффициенты полиномов. Например, считая, что коэффициенты записаны в порядке возрастания степени, сообщение [**1010**] можно интерпретировать в виде многочлена $1 + x^2$.

Многочлены $f(x)$ и $p(x)$ называются *сравнимыми* по модулю $p(x)$, если $f(x)$ делится на $p(x)$ нацело. Поэтому введена операция *mod* — остаток от деления. $(1 + x^2 + x^5)\ mod\ (1 + x^2) = x$. Заметим, что если степень $f(x)$ меньше степени $p(x)$, то результатом $f(x)\ mod\ p(x)$ будет просто $f(x)$

Замечательным свойством полиномиального представления кодов является возможность осуществить циклический сдвиг на одну позицию вправо простым умножением многочлена $p(x)$ степени $n - 1$ на единицу (многочлен x) и взятия остатка от деления на многочлен $x^n + 1$.

Пример 1

Осуществить единичный правый циклический сдвиг кодового слова [**1011**], используя полиномиальную интерпретацию. Вектору [**1011**] соответствует полином $1 + x^2 + x^3$. Умножим его на x, что дает $x + x^3 + x^4$. Найдем остаток от деления на $x^n + 1$, т.е. $(x + x^3 + x^4)\ mod\ (x^n + 1) = \quad (x + x^3 + x^4)\ mod\ (x^4 + 1) = \quad 1 + x + x^3$. Многочлен $1 + x + x^3$ соответствует вектору [**1101**], который получается из [**1011**] правым циклическим сдвигом на одну позицию.

Порождающий многочлен. Процедура циклического кодирования сводится к умножению многочлена-сообщения на подходящий многочлен, называемый *порождающим многочленом* данного кода.

Многочлен $g(x)$ называется порождающим многочленом линейного циклического кода длины n если $g(x)$ делит $1 + x^n$.

Таким образом, для получения порождающего многочлена $g(x)$ нам необходимо разложить полином $x^n + 1$ на множители и выделить многочлен такой степени, которая соответствует длине кодового слова.

Длина кодового слова n и длина k сообщения $g(x)$ связаны соотношением,

$$k = n - deg(g(x)),$$

где $deg(g(x))$ обозначает степень многочлена $g(x)$.

Пример 2

Найти порождающий многочлен (ПМ) линейного циклического кода длины $n = 15$, который осуществляет кодирование сообщений длины $k = 7$. Затем кодировать сообщение [**0110110**].

Для нахождения ПМ для кода длины 15 при длине сообщения 7, то нужно найти делитель полинома $x^{15} + 1$ степени $(n - k) = 15 - 7 = 8$.

Многочлен $x^{15} + 1$ разлагается на множители следующим образом:

$$x^{15} + 1 = (1 + x)(1 + x + x^2)(1 + x + x^2 + x^3 + x^4)(1 + x + x^4)\times$$
$$\times(1 + x^3 + x^4),$$

поэтому в качестве ПМ можно взять

$$g(x) = (1 + x + x^2 + x^3 + x^4)(1 + x + x^4), \text{ т.е.}$$
$$g(x) = 1 + x^4 + x^6 + x^7 + x^8.$$

Найдя ПМ, мы можем кодировать сообщение [**0110110**]. В полиномиальной интерпретации ему соответствует многочлен $p(x) = x + x^2 + x^4 + x^5$. Умножим $p(x)$ на порождающий многочлен g(x):

$$(x + x^2 + x^4 + x^5)(1 + x^4 + x^6 + x^7 + x^8) =$$
$$= x + x^2 + x^4 + x^6 + x^7 + x^8 + x^9 + x^{13}.$$

Итак, сообщение [**0110110**] кодировано в слово [**01101011110010**].

Алгоритм декодирования

Процедура декодирования многочленf $r(x)$ — полученного сообщения циклического кода приведена ниже (рис. 4.31).

Алгоритм декодирования

1. Построить *синдромный* многочлен $s(x) = r(x)\ mod\ g(x)$.

2. Для каждого $i \geq 0$, вычислять $s_i(x) = x^i\ s(x)\ mod\ g(x)$ до тех пор, пока не будет найден, полином $s_j(x)$ такой, что для него $wt(s_j) \leq t$, где t — максимальное число ошибок, исправляемых кодом, а $wt(s_j)$ — вес полинома $s(x)$.

3. Если $wt(s_i(x)) \leq t$, тогда построим полином ошибки $e(x) = x^{n-i}(s_i, 0)$, или иначе: $e(\underline{x}) = x^{n-j}\ s_j(x)\ mod\ (1 + x^n)$ и декодируем искомый полином $p(x) = r(x) - e(x)$. Конец.

4. Иначе, положим $i = i + 1$. Если $i = n$ то ошибка не исправляется. Конец.

5. Если $deg(s_{i-1}(x)) < n - k - 1$, тогда задать $s_i(x) = x\, s_{i-1}(x)$; иначе $s_i(x) = x\, s_{i-1}(x) - g(x)$. И вернуться к шагу (3).

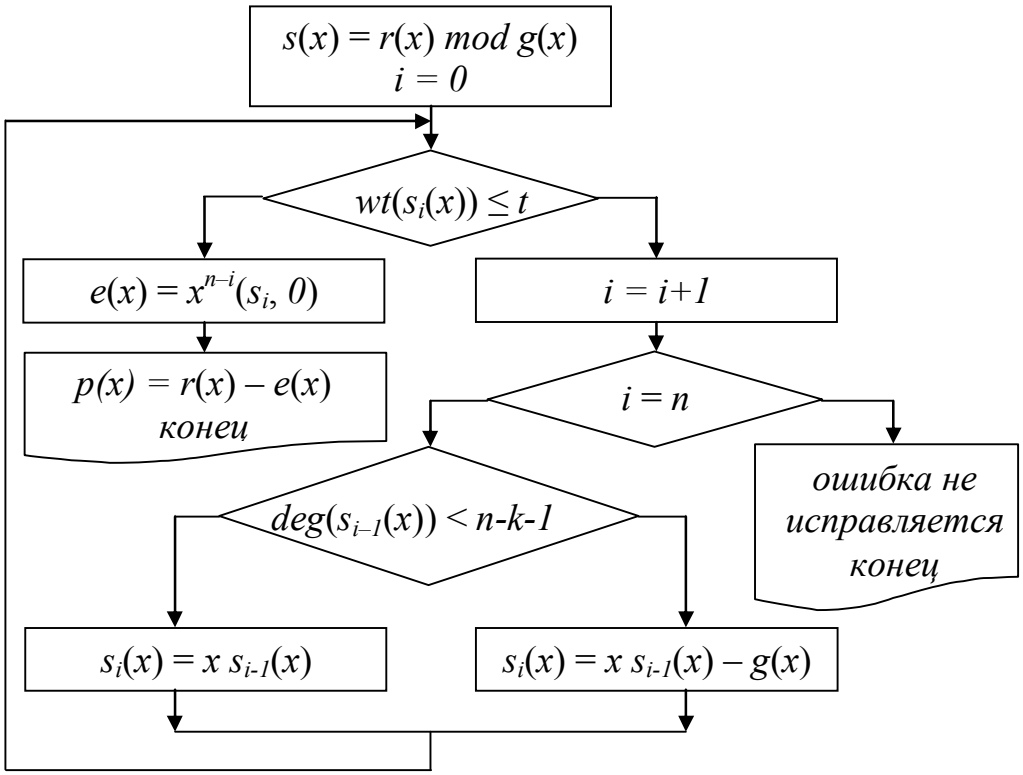


Рис. 4.31 — Процедура декодирования циклического кода

Пример 2 (продолжение)

Декодировать сообщение [**011010111010010**], которое было отправлено получено в первой части примера. Соответствующий вектору многочлен

$$r(x) = x + x^2 + x^4 + x^6 + x^7 + x^8 + x^{10} + x^{13}.$$

Найдем синдромный многочлен $s(x)$, (у нас

$$g(x) = 1 + x^4 + x^6 + x^7 + x^8).$$

$$(x + x^2 + x^4 + x^6 + x^7 + x^8 + x^{10} + x^{13}) \bmod (1 + x^4 + x^6 + x^7 + x^8) =$$
$$= x^2 + x^6 + x^8.$$

Для правильно принятого кодового слова синдром, очевидно, равен 0. В данном случае это не так — посланное сообщение было искажено помехой. В соответствии с описанной процедурой декодирования будем вычислять $s_i(x) = x^i\,(x^2 + x^6 + x^8)\ mod\ g(x)$

для последовательных возрастающих значений i пока не найдем многочлен степени меньшей или равной двум (число ошибок $t = 2$).

$s_1 = x\ s(x)\ mod\ g(x) = (x^3 + x^7 + x^9)\ mod\ g(x) = x^3 + x^4 + x^5 + x^6 + x^7$;

$s_2 = x^2\ s(x)\ mod\ g(x) = (x^4 + x^8 + x^{10})\ mod\ g(x) = 1 + x + x^2 + x^5$;

$s_3 = x^3\ s(x)\ mod\ g(x) = (x^5 + x^9 + x^{11})\ mod\ g(x) = x + x^2 + x^3 + x^6$;

$s_4 = x^4\ s(x)\ mod\ g(x) = (x^6 + x^{10} + x^{12})\ mod\ g(x) = x^2 + x^3 + x^4 + x^7$;

$s_5 = x^5\ s(x)\ mod\ g(x) = (x^7 + x^{11} + x^{13})\ mod\ g(x) =$

$= 1 + x^3 + x^5 + x^6 + x^7$;

$s_6 = x^6\ s(x)\ mod\ g(x) = (x^8 + x^{12} + x^{14})\ mod\ g(x) = x + 1.$

Многочлен $s6 = x + 1$ имеет вес 2, поэтому для нахождения многочлена ошибок вычисляем $e(x) = x^{15\ -\ 6}\ s_6(x)\ mod\ (1 + x^n) = x^9\ (x + 1)\ mod\ (1 + x^n) = x^{10} + x^9$, т.е. ошибка была в девятом и десятом битах.

Итак, если отправленное кодовое слово имеет не более (t) двух ошибок, то оно было таким:

$$p(x) = r(x) - e(x) = (x + x^2 + x^4 + x^6 + x^7 + x^8 + x^{10} + x^{13}) -$$
$$- (x^{10} + x^9) = x + x^2 + x^4 + x^6 + x^7 + x^8 + x^9 + x^{13}.$$

Этот многочлен соответствует вектору [**01101011110010**]. Чтобы восстановить исходное сообщение нам надо разделить кодовое слово $p(x)$ на ПМ $g(x)$ и получить $x + x^2 + x^4 + x^5$, значит, сообщение было [**0110110**].

Коды Хемминга

Наиболее известные из корректирующих кодов — коды Хемминга. В них исправляются ошибки кратности $r = [d - 1]/2$ и обнаруживаются ошибки кратности $d - 1$.

Коды Хемминга, применительно к двоичной системе счисления, предназначены либо для исправления одиночных ошибок (при $d = 3$), либо для исправления одиночных и обнаружения без исправления двойных ошибок ($d = 4$).

Исправление ошибок возможно благодаря избыточности кода Хемминга — к m информационным битам добавлено k контрольных, благодаря которым возможно определить и/или исправить ошибки. Количество необходимых контрольных бит можно определить из следующего неравенства: $m + k + 1 \le 2^k$. Обычно для характеристики кода Хэмминга используют пару (n, m), где n — длина передаваемого блока данных с контрольными битами, а

m — чистая длина данных. Например, (11, 7) означает, что передаваемая длина данных — 7 бит, количество контрольных бит равно 4, что составляет общую длину блока 11 бит. В отличие от других методов коррекции ошибки, где контрольные биты дописываются в конец или начало блока данных (либо вообще в другом пакете данных), биты кода Хэмминга записываются вместе с данными в строго определённых позициях.

В таком коде *n*-значное число имеет *m* информационных и *k* контрольных разрядов. Каждый из контрольных битов является знаком четности для определенной группы информационных знаков слова. При декодировании производится *k* групповых проверок на четность. В результате каждой проверки в соответствующий разряд регистра ошибки записывается **0**, если проверка была успешной, или **1**, если была обнаружена нечетность.

Группы для проверки образуются таким образом, что в регистре ошибки после окончания проверки получается *k*-разрядное двоичное число, показывающее номер позиции ошибочного двоичного разряда. Изменение этого разряда — исправление ошибки.

Позиция *i*-го контрольного знака имеет номер 2^{i-1}. При этом каждый контрольный знак входит лишь в одну группу проверки на четность.

Рассмотрим код Хемминга, предназначенный для исправления одиночных ошибок, т.е. код с минимальным кодовым расстоянием $d = 3$. Ошибка возможна в одной из *n* позиций. Следовательно, число контрольных знаков, а значит, и число разрядов регистра ошибок должно удовлетворять условию:

$$k \geq \log_2 (n+1),$$

тогда под информационные знаки остается *m* разрядов:

$$m \leq n - \log_2 (n+1).$$

Пример 1

Рассмотри механизм работы кода Хэмминга на примере передачи 7-битового кода {**1110011**}. Для контроля целостности блока данных такой длины, нам необходимо 4 бита кода Хэмминга, которые записываются в позициях 1, 2, 4 и 8:

Позиция бита	11	10	9	8	7	6	5	4	3	2	1
Значение бита	1	1	1	**x**	0	0	1	**x**	1	**x**	**x**

Контрольная сумма формируется путем выполнения операции «*исключающее ИЛИ*» над кодами позиций ненулевых битов. В данном случае это 11, 10, 9, 5 и 3:

11	1011
10	1010
09	1001
05	0101
03	0011
сумма	1110

Полученная контрольная сумма записывается в соответствующие разряды блока данных — младший бит в младший разряд:

Позиция бита	11	10	9	8	7	6	5	4	3	2	1
Значение бита	1	1	1	**1**	0	0	1	**1**	1	**1**	**0**

Код Хэмминга сформирован: {**11110011110**}

Теперь рассмотрим два случая ошибки:

1) ошибка в бите 7 — бит **0** заменён на бит **1**, таким образом, принят следующий код: {1111**1**011110}. Просуммировав номера позиций ненулевых битов получим сумму:

11	1011
10	1010
09	1001
08	1000
07	0111
05	0101
04	0100
03	0011
02	0010
сумма	0111

2) Ошибка в бите 5 — бит **1** заменён на бит **0**, принят следующий код: {111100**0**1110}. Просуммируем коды позиций с ненулевыми битами:

11	1011
10	1010
09	1001
08	1000
04	0100
03	0011
02	0010
сумма	0101

Найдена контрольная сумма в кодовых последовательностях, содержащих ошибку не равна 0. В обоих случаях контрольная сумма равна позиции бита, переданного с ошибкой. Теперь для исправления ошибки достаточно инвертировать бит, номер которого указан в контрольной сумме.

Прочие классы кодов

Наряду с циклическими кодами на практике используются другие типы кодов, обладающие различными свойствами. Подробное рассмотрение классов кодов выходит за рамки настоящего курса, поэтому приведем только их краткую характеристику.

Среди циклических кодов особое значение имеет класс кодов, предложенных Боузом и Рой—Чоудхури и независимо от них Хоквингемом. Коды Боуза—Чоудхури—Хоквингема (обозначаемые сокращением ***БЧХ***) отличаются сравнительно просто реализуемой процедурой декодирования.

Относительно простой является процедура *мажоритарного декодирования*, применимая для некоторого класса двоичных линейных, в том числе циклических кодов. Основана она на том, что в этих кодах каждый информационный символ можно несколькими способами выразить через другие символы кодовой комбинации.

Мощные коды (т.е. коды с длинными блоками и большим кодовым расстоянием d) при сравнительно простой процедуре декодирования можно строить, объединяя несколько коротких

кодов. Так строится, например, *итеративный код* из двух линейных систематических кодов (n_1, k_1) и (n_2, k_2). Минимальное кодовое расстояние для двумерного итеративного кода $d = d_1 d_2$, где d_1 и d_2 — соответственно минимальные кодовые расстояния для кодов 1-й и 2-й ступеней.

На итеративный код похож *каскадный код*, но между ними имеется существенное различие. Первая ступень кодирования в каскадном коде является линейным систематическим двоичным кодом (*внутренний код*), каждая комбинация которого рассматривается как один символ недвоичного кода второй ступени (*внешнего*). При приеме сначала декодируются (с обнаружением или исправлением ошибок) все строки (блоки) внутреннего кода, а затем декодируется блок внешнего m-ичного кода, причем исправляются ошибки и стирания, оставшиеся после декодирования внутреннего кода. В качестве внешнего кода используют обычно m-ичный *код Рида—Соломона*, который является подклассом кодов *БЧХ* и обеспечивает наибольшее возможное d при заданных n_2 и k_2, если $n_2 < m$. Каскадные коды во многих случаях наиболее перспективны среди известных блочных помехоустойчивых кодов.

Метод перемежения. Для каналов с группированием ошибок часто применяют *метод перемежения символов,* или *декорреляции ошибок*. Он заключается в том, что символы, входящие в одну кодовую комбинацию, передаются не непосредственно друг за другом, а перемежаются символами других кодовых комбинаций. Если интервал между символами, входящими в одну комбинацию, сделать больше максимально возможной длины группы ошибок, то в пределах комбинации группирования ошибок не будет. Группа ошибок распределится в виде одиночных ошибок на группу комбинаций. Одиночные ошибки будут легко обнаружены (и исправлены) декодером.

Системы с обратной связью. Нередко встречаются случаи, когда информация может передаваться не только от одного абонента к другому, но и в обратном направлении. В таких условиях появляется возможность использовать обратный поток информации для существенного повышения верности сообщений, переданных в прямом направлении. При этом не исключено, что по обоим каналам (прямому и обратному) в основном непосредственно передаются сообщения в двух направлениях (*дуплексная*

связь) и только часть пропускной способности каждого из каналов используют для передачи дополнительных данных, предназначенных для повышения верности.

Возможны различные способы использования *системы с обратной связью* в дискретном канале. Обычно их подразделяют на два типа: системы с информационной обратной связью и системы с управляющей обратной связью.

Системами с информационной обратной связью (***ИОС***) называются такие, в которых с приемного устройства на передающее поступает информация о том, в каком виде принято сообщение. На основании этой информации передающее устройство может вносить те или иные изменения в процесс передачи сообщения:

- повторить ошибочно принятые отрезки сообщения;
- изменить применяемый код (передав предварительно соответствующий условный сигнал и убедившись в том, что он принят);
- прекратить передачу при плохом состоянии канала до его улучшения.

В *системах с управляющей обратной связью* (***УОС***) приемное устройство на основании анализа принятого сигнала само принимает решение о необходимости повторения, изменения способа передачи, временного перерыва связи и передает об этом указание передающему устройству. Возможны и смешанные методы использования обратной связи, когда в некоторых случаях решение принимается на приемном устройстве, а в других случаях на передающем устройстве на основании полученной по обратному каналу информации.

Наиболее распространены системы с ***УОС*** при использовании одновременно с обнаружением ошибок. Такие системы часто называют системами с *автоматическим запросом ошибок* (***ARQ, Automatic Repeat reQuest***).

4.6 Алгоритмы сжатия

Наличие в сообщениях избыточности позволяет ставить вопрос о сжатии данных, т.е. о передаче того же количества информации с помощью последовательностей символов меньшей дли-

ны. Для этого используются специальные алгоритмы сжатия (*компрессии*), уменьшающие избыточность.

Основными техническими характеристиками процессов сжатия и результатов их работы являются:

- ***степень сжатия*** (*compress rating*) или отношение (*ratio*) объемов исходного и результирующего потоков;
- ***скорость сжатия*** — время, затрачиваемое на сжатие некоторого объема информации входного потока, до получения из него эквивалентного выходного потока;
- ***качество сжатия*** — величина, показывающая, на сколько сильно упакован выходной поток.

Степень сжатия оценивают ***коэффициентом сжатия***

$$K = \frac{n}{q},$$

где n — число минимально необходимых символов для передачи сообщения (практически это число символов на выходе эталонного алгоритма сжатия); q — число символов в сообщении, сжатом данным алгоритмом. Так, при двоичном кодировании n равно энтропии источника информации. Часто степень сжатия оценивают отношением длин кодов на входе и выходе алгоритма сжатия.

Все способы сжатия можно разделить на две категории: *обратимое* (*сжатие без потерь*) и *необратимое* (*с потерями*) сжатие (рис. 4.32).

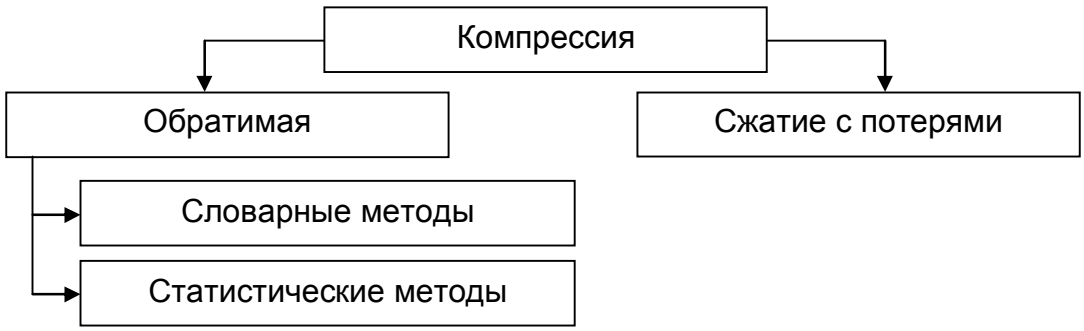


Рис. 4.32 — Методы сжатия

Под ***необратимым*** сжатием подразумевают такое преобразование входного потока данных, при котором происходит потеря малосущественной информации. В зависимости от формата

сжимаемых данных (изображение, звук или видеоряд), в выходном потоке невосстановимо отбрасывается часть информации. Обычно это невоспринимаемые человеческим глазом оттенки изображения на рисунках, искусственное сужение звукового канала до речевого диапазона в звукозаписи, ухудшение разрешения видео за счет движения и т.п. Кроме степени или величины сжатия, в таких алгоритмах возникает понятие качества, т.к. исходное изображение в процессе сжатия изменяется, то под качеством можно понимать степень соответствия исходного и результирующего изображения. Качество оценивается субъективно, исходя из формата информации, хотя имеются и соответствующие интеллектуальные алгоритмы и программы.

Обратимое сжатие всегда приводит к снижению объема выходного потока информации без изменения его информативности, т.е. — без потери информации. Более того, из выходного потока, при помощи восстанавливающего или декомпрессирующего алгоритма, можно получить входной, а процесс восстановления называется декомпрессией или распаковкой и только после процесса распаковки данные пригодны для обработки в соответствии с их внутренним форматом.

Методы обратимой компрессии делятся на *статистические* и *словарные*. ***Словарные*** методы заключаются в том, чтобы в случае встречи подстроки, которая уже была найдена раньше, кодировать ссылку, которая занимает меньше места, чем сама подстрока. Классическим словарным методом является метод Лемпела—Зива (LZ). Все используемые на сегодняшний день словарные методы являются лишь модификациями LZ.

Статистическое кодирование заключается в том, чтобы кодировать каждый символ, но использовать коды переменной длины. Примером таких методов служит метод Хаффмана. Обычно словарные и статистические методы комбинируются, поскольку у каждого свои преимущества.

Сжатие данных осуществляется либо на прикладном уровне с помощью программ сжатия, таких, как ARJ, ZIP, RAR, либо с помощью устройств защиты от ошибок (УЗО) непосредственно в составе аппаратуры ОКД — окончания канала данных, например, модемов по протоколам типа V.42bis.

Пропускная способность каналов связи более дорогостоящий ресурс, чем дисковое пространство, по этой причине сжатие данных до или во время их передачи еще более актуально. Здесь целью сжатия информации является экономия пропускной способности и в конечном итоге ее увеличение. Все известные алгоритмы сжатия сводятся к шифрованию входной информации, а принимающая сторона выполняет дешифровку принятых данных.

Методы обратимой компрессии

Очевидный способ сжатия числовой информации, представленной в коде ASCII, заключается в использовании сокращенного кода с четырьмя битами на символ вместо восьми, так как передается набор, включающий только 10 цифр, символы *точка*, *запятая* и *пробел*.

Алгоритмы RLE (***Run Length Encoding***). Суть методов данного типа состоит в замене цепочек или серий повторяющихся байтов или их последовательностей на один кодирующий байт и счетчик числа их повторений.

Пример:

`44 44 44 11 11 11 11 11 01 33 FF 22 22`

— исходная последовательность,

`03 44 04 11 00 03 01 33 FF 02 22`

— сжатая последовательность.

Первый байт указывает сколько раз нужно повторить символ (следующий байт). Если первый байт равен 00, то затем идет счетчик, показывающий сколько за ним следует неповторяющихся данных.

Данные методы, как правило, достаточно эффективны для сжатия растровых графических изображений (BMP, PCX, TIF, GIF), т.к. последние содержат достаточно много длинных серий повторяющихся последовательностей байтов. Недостатком метода RLE является довольно низкая степень сжатия.

Метод Шеннона—Фано

Данный метод так же довольно прост. Берутся исходные сообщения S_i и их вероятности появления P_i. Этот список делится на две группы с примерно равной интегральной вероятностью.

Каждому сообщению из группы **1** присваивается **0** в качестве первой цифры кода. Сообщениям из второй группы ставятся в соответствие коды, начинающиеся с **1**. Каждая из этих групп делится на две аналогичным образом и добавляется еще одна цифра кода. Процесс продолжается до тех пор, пока не будут получены группы, содержащие лишь одно сообщение. Каждому сообщению в результате будет присвоен код *S* с длиной — *lg(P(S))*. Это справедливо, если возможно деление на подгруппы с совершенно равной суммарной вероятностью. Если же это невозможно, некоторые коды будут иметь длину — *lg(P(S))+1*. Алгоритм Шеннона—Фано не гарантирует оптимального кодирования.

Арифметическое кодирование

Арифметическое кодирование является методом, позволяющим упаковывать символы входного алфавита без потерь при условии, что известно распределение частот этих символов и является наиболее оптимальным, т.к. достигается теоретическая граница степени сжатия.

Предполагаемая требуемая последовательность символов, при сжатии методом арифметического кодирования, рассматривается как некоторая двоичная дробь из интервала [0, 1). Результат сжатия представляется как последовательность двоичных цифр из записи этой дроби.

Идея метода состоит в следующем: исходный текст рассматривается как запись этой дроби, где каждый входной символ является *цифрой* с весом, пропорциональным вероятности его появления. Этим объясняется интервал, соответствующий минимальной и максимальной вероятностям появления символа в потоке.

Каждому слову во входном алфавите соответствует некоторый подинтервал из интервала [0, 1) а пустому слову соответствует весь интервал [0, 1). После получения каждого следующего символа интервал уменьшается с выбором той его части, которая соответствует новому символу. Кодом цепочки является интервал, выделенный после обработки всех ее символов, точнее, двоичная запись любой точки из этого интервала, а длина полученного интервала пропорциональна вероятности появления кодируемой цепочки.

Пример

Пусть алфавит состоит из двух символов: **1** и **0** с вероятностями соответственно 0,75 и 0,25. Применим данный алгоритм для кодирования цепочки **1 1 0 1 0**. Разобьем наш интервал вероятностей [0, 1) на части, длина которых пропорциональна вероятностям символов. В нашем случае это [0; 0,75) для символа **1** и [0,75; 1) для **0**.

Так как первый символ **1**, то оставляем для дальнейшего анализа левый интервал [0; 0,75), а правый отбрасываем. Разбиваем теперь вновь полученный интервал на части, пропорциональные вероятностям символов: [0; 0.5625) и [0.5625; 0.75), второй символ так же **1**, так что отбрасываем снова правый, а оставляем левый интервал — [0; 0.5625). Для третьего символа, равного **0**, отбрасываем левый интервал [0; 0.421875), а оставляем правый — [0.421875; 0.5625). Повторяем эту операцию для всех символов кодируемого сообщения (см. табл. 4.1).

Таблица 4.1

Шаг	Символ	Интервал	Полученный код
0		[0; 1)	
1	**1**	[0; 0.75)	**1**
2	**1**	[0; 0.5625)	**1 1**
3	**0**	[0.421875; 0.5625)	**1 1 0**
4	**1**	[0.421875; 0.52734375)	**1 1 0 1**
5	**0**	[0.5009765625; 0.52734375)	**1 1 0 1 0**

Границы k-того интервала $[a_k, b_k)$ позволяют рассчитать длину l_k k-того интервала $l_k = b_k - a_k$. Затем, в зависимости от кодируемого символа, вычисляются границы следующего интервала: при кодировании символа **1** — $[a_{k+1} = a_k;\ b_{k+1} = a_k + l_k \cdot p_{k+1})$, а при кодировании символа **0** — $[a_{k+1} = b_k - l_k \cdot p_{k+1};\ b_{k+1} = b_k)$, здесь p_{k+1} — вероятность k+1-го символа см. рис. 4.33.

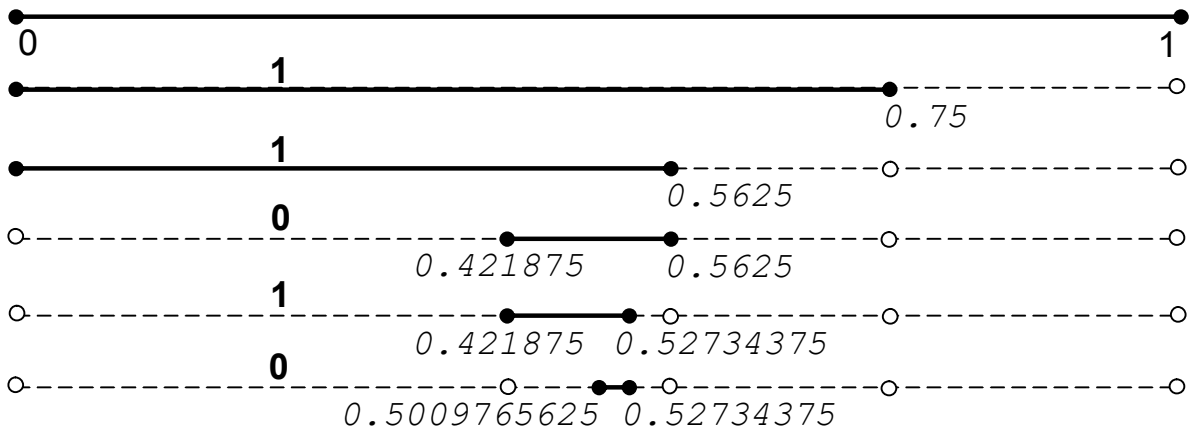


Рис. 4.33 — Арифметическое кодирование

В качестве кода можно взять любое число из интервала, полученного на последнем шаге, например, **`0.51`**.

Алгоритм декодирования работает аналогично кодирующему: получив на входе 0.51, производится разбиение на интервалы соответствующих вероятностей и выбирается тот из них, в который попадает код. Продолжая этот процесс, мы однозначно декодируем все символы сообщения. Для того чтобы декодирующий алгоритм мог определить конец цепочки, мы можем либо передавать ее длину отдельно, либо добавить к алфавиту дополнительный уникальный символ — *«конец цепочки»*.

Количество битов, необходимое для записи этого числа, примерно равно минус логарифму ширины интервала. Ширина интервала равна произведению вероятностей символов, т.е. вероятности p всего сообщения. Таким образом, длина кода равна — $\log(p)$, т.е. теоретическому пределу. На практике работают с переменными ограниченной длины, и точность вычислений будет ограничена, а значит, сжатие все же немного хуже.

Метод Хаффмана (***Huffman code***) или минимально-избыточный префиксный код (*minimum-redundancy prefix code*) относится к статистическим методам кодирования.

Обычно для хранения данных и передачи сообщений используются коды фиксированной длины, например, код ASCII. Множество символов представляется некоторым количеством кодовых слов равной длины, которая для кода ASCII равна 8 битам. При этом для всех сообщений с одинаковым количеством

символов требуется одинаковое количество битов при хранении и одинаковая ширина полосы пропускания при передаче.

Метод Хаффмана основан на кодировании более короткими кодовыми словами часто встречающиеся символы, а символы встречающиеся редко — более длинными. Подбирая кодовые последовательности таким образом, можно получить код с длиной, очень близкой к его энтропии (то есть информационной насыщенности). Кодовые слова при этом должны быть выбраны так, чтобы никакое из них не было префиксом другого кодового слова. Благодаря этому условию гарантируется возможность однозначного декодирования определенного закодированного текста.

Например, коде ASCII сообщение S = {DDABDECBCDBE} кодируется 96 битами следующим образом:

```
01000100-01000100-01000001-01000010-01000100-
01000101-01000011-01000010-01000011-01000100-
01000010-01000101.
```

В то время как при использовании кода со следующим представлением символов: E — 00, D — 10, B — 11, C — 010, A — 011, то же самое сообщение можно закодировать, используя только 27 бит:

```
10-10-011-11-10-00-010-11-010-10-11-00.
```

В своей статье, опубликованной в 1952 г., *Дэвид Хаффман* описал алгоритм поиска множества кодов, которые минимизируют ожидаемую длину сообщений при условии, что известны вероятности появления каждого символа. В этом методе символам, имеющим меньшую вероятность появления, ставятся в соответствие более длинные кодовые слова.

Пусть $A = \{a_1, a_2, ..., a_n\}$ — алфавит из n различных символов, $P = \{p_1, p_2, ..., p_n\}$ — соответствующий ему набор положительных весов (вероятностей).

Тогда набор бинарных кодов $C = \{c_1, c_2, ..., c_n\}$, такой что:

(1) c_i не является префиксом для c_j, при $i \neq j$;

(2) $\sum_{i=1}^{n} p_i \cdot |c_i|$ — минимальна

называется минимально-избыточным префиксным кодом (*minimum-redundancy prefix code*) или иначе кодом Хаффмана.

Для создания кода Хаффмана должны быть известны, или рассчитаны, вероятности $P = \{p_1, p_2, ..., p_n\}$ появления символов исходного алфавита $A = \{a_1, a_2, ..., a_n\}$ в кодируемом сообщении *S*. В этом случае либо пользуются стандартизованными таблицами вероятностей, составленными для каждого языка, либо используют в качестве вероятностей частоту появления символа в исходном сообщении.

Рассмотрим алгоритм Хаффмана на примере кодирования уже встречавшегося нам сообщения *S* = {DDABDECBCDBE}.

Составим список всех символов, встречающихся в исходном тексте, и определим количество появлений каждого символа в нем: символ A встречается 1 раз, B — 3, C — 2, D — 4 и E — 2 раза.

Алгоритм Хаффмана состоит из двух этапов — построения дерева вероятностей и построения кодов символов. На первом этапе составляется двоичный граф — дерево вероятностей, узлы которого строятся по простому правилу, а листьями являются символы, нуждающиеся в кодировании. На втором этапе ребра графа нумеруются символами **1** и **0**, а затем выписываются коды вершин графа.

- Составим список кодируемых символов (при этом будем рассматривать каждый символ как одноэлементное бинарное дерево, вес которого равен весу символа) как на рис. 4.34.

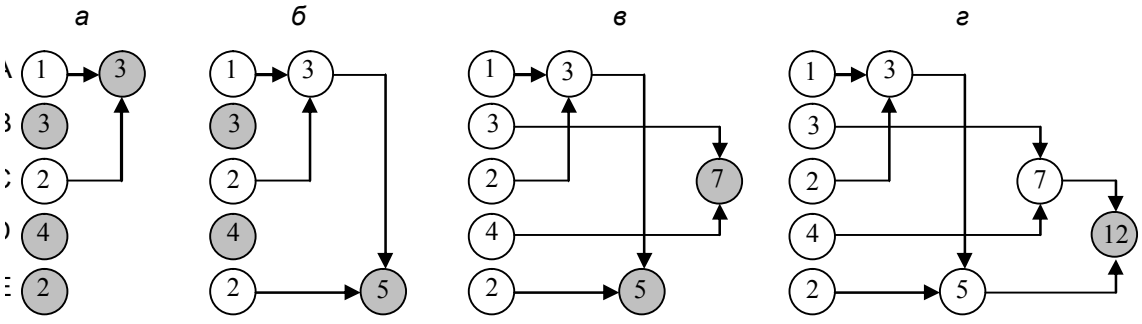


Рис. 4.34 — Этапы построения дерева Хаффмана

- Из списка выберем два узла с наименьшим весом (на рис. 4.34(*а*) это листья **A** и **C** с вероятностями 1 и 2 соответственно).

- Сформируем новый узел и присоединим к нему, в качестве дочерних, два узла выбранных из списка. При этом вес сформированного узла положим равным сумме весов дочерних узлов (в нашем случае — 3).
- Теперь два использованных узла объединены в дерево с общим весом — 3. Отныне не будем принимать во внимание при поиске наименьших вероятностей два объединенных узла, но будем рассматривать новое дерево как полноценную структуру с частотой появления, равной сумме частот появления двух соединившихся вершин (на рис. 4.34 вершины выпавшие из дальнейшего анализа — белого цвета, а рассматриваемые вершины — серого).
- Будем повторять эти действия до тех пор, пока не объединим все деревья в одно. Если в списке рассматриваемых (серых) вершин больше одной, то продолжаем — выбираем две вершины с наименьшей вероятностью (теперь это вершины с весом 3 и 2, см. рис. 4.34(*б*)) и объединяем в дерево с весом 5, затем (см. рис. 4.34(*в*)) вершины с вероятностями 3 и 4 объединяем в дерево с вершиной 7 и т.д.
- Будем повторять операцию объединения вершин до тех пор, пока не придем к одному дереву. Для проверки: очевидно, что в последней вершине будет записана длина кодируемого сообщения (рис. 4.34(*г*)).
- На этом построение дерева вероятностей закончено и веса его вершин нам больше не нужны. Занумеруем теперь ребра полученного графа, начиная с вершины и присваивая **1** верхнему (правому) ребру и **0** — нижнему (левому), как на рис. 4.35(*а*).

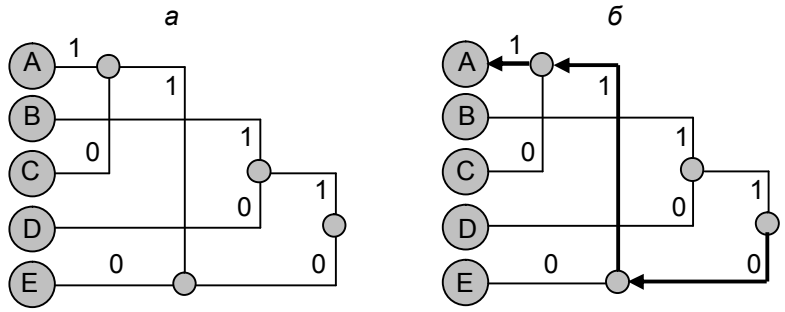


Рис. 4.35 — Составление кода Хаффмана

- Листовые узлы дерева Хаффмана соответствуют символам кодируемого алфавита (рис. 4.35). Глубина листовых узлов равна длине кода соответствующих символов.
- Код Хаффмана для каждого символа — это путь от корня дерева к соответствующему листовому узлу (на рис. 4.35 (*б*) стрелками обозначен путь для символа A). Его можно представить в виде битовой строки, в которой **0** соответствует выбору левого поддерева, а **1** — правого.

Выпишем коды $C = \{c_1, c_2, ..., c_n\}$ для всех символов в нашем примере:

E — **00,** D — **10,** B — **11,** C — **010,** A — **011.**

- Теперь у нас есть все необходимое для того чтобы закодировать сообщение *S*. Достаточно просто заменить каждый символ соответствующим ему кодом, получим строку

S′ = { 10100111110000101101010101100 }.

- Теперь декодируем «принятое» сообщение *S′*. Начиная с корня дерева, будем двигаться вниз, выбирая левое поддерево, если очередной символ в битовом потоке равен **0**, и правое — если **1**. Дойдя до листового узла, мы декодируем соответствующий ему символ. Следуя этому алгоритму, мы в точности получим исходное сообщение *S*.

```
10-10011111000010110101011100
10-10-011111000010110101011100
10-10-011-111000010110101011100
10-10-011-11-10000101101010101100
10-10-011-11-10-0001011010101100
10-10-011-11-10-00-01011010101100
10-10-011-11-10-00-010-11010101100
10-10-011-11-10-00-010-11-010101100
10-10-011-11-10-00-010-11-010-101100
10-10-011-11-10-00-010-11-010-10-1100
10-10-011-11-10-00-010-11-010-10-11-00
```

S = {DDABDECBCDBE}.

В теории кодирования информации доказывается, что код Хаффмана является префиксным, то есть код никакого символа не является началом кода какого-либо другого символа. Проверьте это на нашем примере. Из этого следует, что код Хаффмана

однозначно восстановим получателем, даже если не сообщается длина кода каждого переданного символа. Получателю пересылают только дерево Хаффмана в компактном виде, а затем входная последовательность кодов символов декодируется им самостоятельно без какой-либо дополнительной информации.

Оценим теперь ***степень сжатия***. В исходном сообщении S было 12 символов, на каждый из которых отводилось бы по $\lceil \log_2 |A| \rceil = \lceil \log_2 5 \rceil = 3$ бита минимального равномерного кода (или по 8 бит ASCII), таким образом, размер S равнялся бы 12×3 = 36 бит (или 96 бит ASCII). Размер закодированного сообщения S' можно получить воспользовавшись формулой $\sum_{i=1}^{n} p_i \cdot |c_i|$ из определения, или непосредственно, подсчитав количество бит в S'. И в том и другом случае мы получим 27 бит.

Передача кодового дерева. Для того чтобы закодированное сообщение удалось декодировать, декодеру необходимо иметь такое же кодовое дерево (в той или иной форме), какое использовалось при кодировании. Поэтому вместе с закодированными данными мы вынуждены сохранять соответствующее кодовое дерево. Ясно, что чем компактнее оно будет, тем лучше. Решить эту задачу можно несколькими способами. Самое очевидное решение — сохранить дерево в явном виде (т.е. как упорядоченное множество узлов и указателей того или иного вида). Это, пожалуй, самый расточительный и неэффективный способ. На практике он не используется.

Можно сохранить список частот символов (т. е. частотный словарь). С его помощью декодер без труда сможет реконструировать кодовое дерево. Хотя этот способ и менее расточителен, чем предыдущий, он не является наилучшим.

Наконец, можно использовать одно из свойств канонических кодов. Как уже было отмечено ранее, канонические коды вполне определяются своими длинами. Другими словами, все что необходимо декодеру — это список длин кодов символов. Учитывая, что в среднем длину одного кода для n-символьного алфавита можно закодировать $[\log_2(\log_2 n)]$ битами, получим очень эффективный алгоритм.

Предположим, что размер алфавита $n = 256$, и мы сжимаем обыкновенный текстовый файл. Скорее всего, мы не встретим все n символов нашего алфавита в таком файле. Положим тогда длину кода отсутствующих символов равной нулю. В этом случае сохраняемый список длин кодов будет содержать достаточно большое число нулей (длин кодов отсутствующих символов) сгруппированных вместе. Каждую такую группу можно сжать при помощи, например, *RLE*.

Алгоритмы, родственные методу Хаффмана

Обобщением этого способа является алгоритм, основанный на *словаре сжатия данных*. В нем происходит выделение и запоминание в словаре повторяющихся цепочек символов, которые кодируются цепочками меньшей длины.

Интересен *алгоритм «стопка книг»*, в котором код символа равен его порядковому номеру в списке. Появление символа в кодируемом потоке вызывает его перемещение в начало списка. Очевидно, что часто встречающиеся символы будут тяготеть к малым номерам, а они кодируются более короткими цепочками **1** и **0**.

Алгоритм Лемпеля—Зива (Lempel—Ziv)

Классический алгоритм Лемпеля—Зива — LZ77, названный так по году своего опубликования, предельно прост. Он формулируется следующим образом: «если в прошедшем ранее выходном потоке уже встречалась подобная последовательность байт, причем запись о ее длине и смещении от текущей позиции короче чем сама эта последовательность, то в выходной файл записывается ссылка (смещение, длина), а не сама последовательность».

Пример: фраза "КОЛОКОЛ_ОКОЛО_КОЛОКОЛЬНИ" закодируется как "КОЛО(-4,3)_(-5,4)О_(-14,7)ЬНИ".

Рассмотрим, еще, последовательность «AAAAAAA». С помощью алгоритма RLE она будет закодирована как «(A,7)», в то же время ее можно достаточно хорошо сжать и с помощью алгоритма LZ77: «A(–1,6)». Действительно, степень сжатия именно такой последовательности им хуже (примерно на 30—40 %), но сам по себе алгоритм LZ77 более универсален, и может намного лучше обрабатывать последовательности вообще несжимаемые методом RLE.

Данный алгоритм является несимметричным по времени, поскольку требует полного перебора буфера при поиске одинаковых подстрок. В результате нам сложно задать большой буфер из-за резкого возрастания времени компрессии. Однако потенциально построение алгоритма, в котором на (длину) и на (смещение) будет выделено по 2 байта, даст нам возможность сжимать все повторяющиеся подстроки размером до 32 Кбайт в буфере размером 64 Кбайта. Однако, минимальная подстрока, для которой нам выгодно проводить сжатие, должна состоять минимум из 5 байт, что и определяет малую ценность данного алгоритма. К достоинствам LZ можно отнести чрезвычайную простоту алгоритма декомпрессии.

Алгоритм Лемпеля—Зива—Велча (Lempel—Ziv—Welch, LZW)

Данный алгоритм отличают высокая скорость работы как при упаковке, так и при распаковке, достаточно скромные требования к памяти и простая аппаратная реализация. Недостаток — низкая степень сжатия по сравнению со схемой двухступенчатого LZ-кодирования, рассматриваемой ниже.

Предположим, что у нас имеется массив, хранящий строки текста и содержащий порядка от 2-х до 8-ми тысяч пронумерованных ячеек. Запишем в первые 256 ячеек строки, состоящие из одного символа, код которого равен номеру гнезда.

Алгоритм просматривает входной поток, разбивая его на подстроки и добавляя новые гнезда в конец словаря. Прочитаем несколько символов в строку S и найдем в словаре строку T_N — самый длинный префикс строки S.

Пусть он найден в гнезде с номером N. Выведем число N в выходной поток, переместим указатель входного потока на $length(T_N)$ символов вперед и добавим в словарь новое гнездо, содержащее строку T_N+C, где C — очередной символ на входе (сразу после T_N). Алгоритм преобразует поток символов на входе в поток индексов ячеек словаря на выходе.

При практической реализации этого алгоритма следует учесть, что любое гнездо словаря, кроме самых первых, содержащих односимвольные цепочки, хранит копию некоторого другого гнезда, к которой в конец приписан один символ. Вследствие

этого можно обойтись простой списочной структурой с одной связью.

Пример: входная строка ABCABCABCABCABCABC
кодируется как 1 2 3 4 6 5 7 7 7
со словарем A_1, B_2, C_3, AB_4, BC_5, CA_6, ABC_7

Двухступенчатый Алгоритм Лемпеля—Зива

Гораздо большей степени сжатия можно добиться при выделении из входного потока повторяющихся цепочек — блоков, и кодирования ссылок на эти цепочки с построением хеш-таблиц от первого до *n*-го уровня.

Метод, о котором и пойдет речь, принадлежит Лемпелю и Зиву и обычно называется *LZ-compression*.

Суть его состоит в следующем: упаковщик постоянно хранит некоторое количество последних обработанных символов в буфере. По мере обработки входного потока, вновь поступившие символы попадают в конец буфера (называемого также скользящим словарем — sliding dictionary), сдвигая предшествующие символы и вытесняя самые старые.

Затем, после построения хеш-таблиц, алгоритм выделяет (путем поиска в словаре) самую длинную начальную подстроку входного потока, совпадающую с одной из подстрок в словаре, и выдает на выход пару (length, distance). Здесь *length* — длина найденной в словаре подстроки, а *distance* — расстояние от нее до входной подстроки. В случае если такая подстрока не найдена, в выходной поток просто копируется очередной символ входного потока.

В первоначальной версии алгоритма предлагалось использовать простейший поиск по всему словарю. Однако, в дальнейшем, было предложено использовать двоичное дерево и хеширование для быстрого поиска в словаре, что позволило на порядок поднять скорость работы алгоритма.

Таким образом, двухступенчатый алгоритм Лемпеля—Зива преобразует один поток исходных символов в два параллельных потока длин и индексов в таблице (length + distance). Очевидно, что эти потоки являются потоками символов с двумя новыми алфавитами, и к ним можно применить один из упоминавшихся выше методов (RLE, кодирование Хаффмена или арифметиче-

ское кодирование). При реализации этого метода необходимо добиться согласованного вывода обоих потоков в один файл. Эта проблема обычно решается путем поочередной записи кодов символов из обоих потоков.

Пример: `abcabcabcabcabc`
создание хэш-таблицы: `1 a 1 b 1 c 3 3 6 3 9 3 12 3`
исключение большой группы повторяющихся последовательностей:
`1 a 1 b 1 c 12 3`
после чего RLE, метод Хаффмана, арифметическое кодирование и т.п.

Применение двухступенчатого алгоритма в компьютерных сетях

Еще раз, суть метода в замене потока символов кодами, записанными в памяти в виде словаря (таблица перекодировки). Соотношение между символами и кодами меняется вместе с изменением данных. Таблицы кодирования периодически меняются, что делает метод более гибким. Размер небольших словарей лежит в пределах 2—32 Кбайт, но более высоких коэффициентов сжатия можно достичь при очень больших словарях до 400 Кбайт. К полученным потокам данных применяют RLE, метод Хаффмана, арифметическое кодирование и т.п.

Реализация алгоритма возможна в двух режимах: *непрерывном* и *пакетном*. Первый использует для создания и поддержки словаря непрерывный поток символов. При этом возможен многопротокольный режим (TCP/IP, DECnet). Словари сжатия и декомпрессии должны изменяться синхронно, а канал должен быть достаточно надежен (X.25 или PPP), что гарантирует отсутствие искажения словаря при повреждении или потере пакета. При искажении одного из словарей оба ликвидируются и должны быть созданы вновь.

Пакетный режим сжатия также использует поток символов для создания и поддержания словаря, но поток здесь ограничен одним пакетом и по этой причине синхронизация словарей ограничена границами кадра. Для пакетного режима достаточно иметь словарь объемом, порядка 4 Кбайт. Непрерывный режим обеспечивает лучшие коэффициенты сжатия, но задержка полу-

чения информации (сумма времен сжатия и декомпрессии) при этом больше, чем в пакетном режиме.

При передаче пакетов иногда применяется сжатие заголовков, например, *алгоритм Ван Якобсона* (***RFC-1144***). Этот алгоритм используется при скоростях передачи менее 64 Кбит/с. При этом достижимо повышение пропускной способности на 50 % для скорости передачи 4800 бит/с. Сжатие заголовков зависит от типа протокола. При передаче больших пакетов на сверх высоких скоростях по региональным сетям используются специальные канальные алгоритмы, независящие от рабочих протоколов. Канальные методы сжатия информации не могут использоваться для сетей, базирующихся на пакетной технологии, SMDS (*Switched Multi-megabit Data Service),* ATM, X.25 и Frame Relay. Канальные методы сжатия дают хорошие результаты при соединении по схеме точка-точка, а при использовании маршрутизаторов возникают проблемы — ведь нужно выполнять процедуры сжатия/декомпрессии в каждом маршрутизаторе, что заметно увеличивает суммарное время доставки информации. Возникает и проблема совместимости маршрутизаторов, которая может быть устранена процедурой идентификации при у становлении виртуального канала.

Иногда для сжатия информации используют аппаратные средства. Такие устройства должны располагаться как со стороны передатчика, так и со стороны приемника. Как правило, они дают хорошие коэффициенты сжатия и приемлемые задержки, но они применимы лишь при соединениях точка-точка. Такие устройства могут быть внешними или встроенными, появились и специальные интегральные схемы, решающие задачи сжатия/декомпрессии. На практике задача может решаться как аппаратно, так и программно, возможны и комбинированные решения.

Если при работе с пакетами заголовки оставлять неизменными, а сжимать только информационные поля, ограничение на использование стандартных маршрутизаторов может быть снято. Пакеты будут доставляться конечному адресату, и только там будет выполняться процедура декомпрессии. Такая схема сжатия данных приемлема для сетей X.25, SMDS, Frame Relay и ATM.

Необратимая компрессия

Одна из серьезных проблем архивации с потерями заключается в том, что до сих пор не найден адекватный критерий оценки потерь качества.

Классический критерий — *среднеквадратичное отклонение* значений (*root mean square — RMS*)

$$d(x,y)=\sqrt{\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\left(x_{ij}-y_{ij}\right)^2}{n^2}}$$

не дает адекватной оценки: по нему даже при незаметном глазу понижении яркости всего на 5 % показатель ухудшения будет значительным. В то же время изображения со снегом резким изменением цвета отдельных точек, слабыми полосами или муаром будут признаны почти не изменившимися. Свои неприятные стороны есть и у других критериев.

Рассмотрим, например, *максимальное отклонение*:

$$\delta(x,y)=\max\left|x_{ij}-y_{ij}\right|.$$

Эта мера крайне чувствительна к резкому изменению отдельных пикселов: во всем изображении может существенно измениться только значение одного пиксела (что практически незаметно для глаза), однако согласно этой мере изображение будет сильно испорчено.

Мера, используемая на практике в настоящее время, называется мерой отношения сигнала к шуму (*peak-to-peak signal-to-noise ratio — PSNR*).

$$\eta(x,y)=10\cdot\log_{10}\frac{255^2\cdot n^2}{\sum_{i=1}^{n}\sum_{j=1}^{n}\left(x_{ij}-y_{ij}\right)^2}.$$

Данная мера, по сути, аналогична среднеквадратичному отклонению, однако пользоваться ей несколько удобнее за счет логарифмического масштаба шкалы. Ей присущи те же недостатки, что и среднеквадратичному отклонению.

Методы JPEG (*Joint Photographic Expert Group*), основанные на потере малосущественной информации (не различимые для глаза оттенки кодируются одинаково, коды укорачиваются).

В этих методах передаваемая последовательность пикселей делится на блоки 8×8, в каждом блоке производится преобразование Фурье, устраняющее высокие частоты. При разложении матрицы такой области в двойной ряд Фурье по косинусам значимыми оказываются только первые коэффициенты. Изображение восстанавливается по коэффициентам оставшихся частот.

Кроме того, благодаря несовершенству человеческого зрения, можно аппроксимировать коэффициенты более грубо без заметной потери качества изображения. Для этого используется квантование коэффициентов (quantization). В самом простом случае это арифметический побитовый сдвиг вправо. При этом преобразовании теряется часть информации, но могут достигаться большие коэффициенты сжатия.

Существенными положительными сторонами алгоритма является:

- возможность управлять степенью сжатия;
- коэффициент компрессии от 2 до 200 (задается пользователем);
- выходное цветное изображение может иметь 24 бита на пиксель.

К недостаткам можно отнести:

- При повышении степени сжатия изображение распадается на отдельные квадраты (8x8). Это связано с тем, что происходят большие потери в низких частотах при квантовании, и восстановить исходные данные становится невозможно.
- Проявляется эффект Гиббса — ореолы по границам резких переходов цветов.

Фрактальная архивация основана на том, что мы представляем изображение в более компактной форме с помощью коэффициентов системы итерируемых функций (*Iterated Function System — IFS*).

Строго говоря, IFS представляет собой набор трехмерных аффинных преобразований, в нашем случае переводящих одно изображение в другое. Преобразованию подвергаются точки в трехмерном пространстве (x_координата, y_координата, яркость).

Фактически, фрактальная компрессия — это поиск самоподобных областей — *фракталов* в изображении и определение для них параметров аффинных преобразований поворота, сдвига, от-

ражения и масштабирования. Алгоритм в явном виде ищет самоподобные области в нашем изображении. Поиск производится до тех пор, пока не будет найден *аттрактор* фрактала — неподвижная точка аффинных преобразований.

У фрактального метода очень высокая степень сжатия. Каждое преобразование кодируется буквально считанными байтами, в то время как изображение, построенное с их помощью, может занимать и несколько мегабайт.

Алгоритм имеет существенный недостаток — ему требуется очень (!) много времени. Даже в худшем случае (без применения сопутствующих оптимизирующих алгоритмов), потребуется перебор и сравнение всех возможных фрагментов изображения разного размера. Даже для небольших изображений при учете дискретности мы получим астрономическое число перебираемых вариантов.

Подавляющее большинство исследований в области фрактальной компрессии сейчас направлены на уменьшение времени архивации, необходимого для получения качественного изображения.

Сжатие видеоинформации

Видеоинформация представляется в виде видеороликов, т.е. наборов последовательно выводимых друг за другом взаимосвязанных изображений — кадров. Если скорость появления видеокадров превышает частоту слияния мельканий (порядка 25 кадров в секунду) то у пользователя создается впечатление непрерывного движения объектов (*full-motion video* — полнокадровое видео). Этот принцип был реализован в кино и в настоящее время остается основным при оцифровке видеоизображения.

Объем одной секунды видеоролика с частотой 30 кадров в секунду при разрешении 640x480 точек, представленных 8-разрядным кодом (256 цветов), составляет 9 Мбайт. При использовании 24-разрядной цветовой палитры (16 млн. цветов) и разрешения 1280x1024 эта цифра увеличится до 114 Мбайт.

Поэтому до сего времени ведутся интенсивные работы по разработки видеоформатов, хорошо сжимающих водеоизображения и позволяющих воспроизводить видеоинформацию в реальном времени без снижения качества изображений.

Методы, алгоритмы и устройства сжатия видеоданных объединяются под общим названием — *codec* (***CO****mpressor-****DEC****ompressor*). Задача видеокодека заключается в максимально возможном сжатии видеоизображения с возможностью его последующего восстановления (декомпрессии) с высокой скоростью и минимальными искажениями информации. Как правило, методы сжатия видео информации основаны на поиске лишней, избыточной информации и удаления ее с целью уменьшения объема. При этом могут использоваться различные алгоритмы сжатия. Некоторые основаны на внутрикадровом сжатии, т.е. сжимается информация по каждому отдельному кадру, другие базируются на межкадровом сжатии, при котором фиксируется динамика изменения информации по кадрам. В этом случае последующие кадры формируются на основе информации об изменении предыдущего кадра.

Видеоинформация формата QuickTime хранится в файлах с расширением *.mov. В формате QuickTime кроме видеоинформации может храниться аудиоинформация звукового сопровождения видеоданных. При частоте дискретизации 22,05 Кгц, разрядности 8 бит, в режиме «моно» одна секунда аудиоинформации занимает примерно 20—30 Кбайт. Одна секунда видеоизображения с таким же звуком занимает 150—200 Кбайт (236x168 — 320x240, 15 кадров в секунду).

В системах Windows распространен видеостандарт AVI (***A****udio* ***V****ideo* ***I****nterleaved*). Файлы этого стандарта имеют расширение *.avi. В AVI-файле применяется межкадровое сжатие. Оно содержит один ключевой кадр, относительно которого формируются остальные кадры видеоизображения. Утверждается, что AVI-файлы могут воспроизводиться с частотой 24 кадра в секунду. Но в большинстве случаев они записываются с частотой 15—18, а иногда и 10 кадров в секунду, чтобы уменьшить занимаемый объем. AVI-файл длительностью 1 секунда занимает от нескольких десятков до нескольких сотен килобайт (обычно 50—300 Кбайт).

В 1992 г. группа экспертов по движущимся изображениям (***M****oving* ***P****ictures* ***E****xperts* ***G****roup*) разработала новый стандарт видеокомпрессии — MPEG. Международная организация стандартизации (ISO) приняла его как стандарт компрессии MPEG-1 (ISO

11172). Для передачи телепрограмм по каналам связи используется формат MPEG-1. Он обладает разрешением 352x288 точек для стандарта PAL; 352x240 точек для стандарта NTSC и кино. Частота кадров: 25 (PAL), 29,97 (NTSC), 23,976 (кино). Скорость передачи данных: 384 Кбит/с — 5 Мбит/с.

Концепция сжатия видео в ***MPEG*** очень проста — определить, какая именно информация в потоке повторяется хотя бы в течение какого-то отрезка времени и принять меры к избежанию дублирования этой информации. Наиболее ценное достоинство ***MPEG*** кодирования, особенно удобное для передачи по различным сетям — возможность гибкой настройки качества изображения в зависимости от пропускной способности сети. Это и сделало ***MPEG-2*** фактическим стандартом для приема/передачи цифрового телевидения по различным сетям.

К сожалению, не существует возможности однозначно оценить качество кодирования некими приборами и измерениями. Единственный критерий здесь — человек и как он воспримет сжатую информацию. Поэтому правила сжатия видеоданных при ***MPEG*** кодировании вырабатывались на основе модели восприятия человеком видеоизображений (***HVS — Human Visual Sense***). Избыточность изображения согласно ***HVS*** определяется по трем основным критериям:

- Невидимые человеческим глазом детали изображения — места гашения по вертикали и горизонтали. Удаление этой информации вообще никак не сказывается на изображении.
- Статистическая избыточность. Подразделяется на пространственную и временную. Под пространственной избыточностью понимаются участки изображения, на которых смежные пиксели практически одинаковы. Под временной — не изменяемые во времени фрагменты изображения.
- Избыточность по цвету и яркости — рассчитывается исходя из ограниченной чувствительности человека к небольшим изменениям цветов и яркости деталей изображения.

Как реализуется формат MPEG. Для удобства кодирования видеоданных весь видеопоток разбивается на группы изображений ***GOP*** (***Group of Pictures***). Такая группа строится, как показано на рис. 4.36

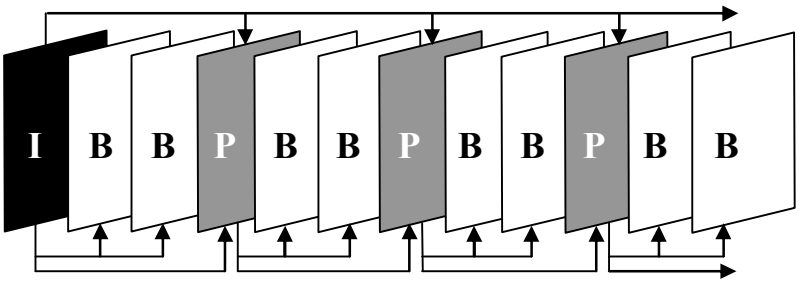


Рис. 4.36 — Кадры формата MPEG

В группе изображений используются следующие кадры:

- ***I*** — *Intra-кадры*, которые обычно называются опорными и содержат всю информацию об изображении. ***MPEG*** последовательности без этих кадров быть не может в принципе. При компрессии ***I-кадров*** происходит удаление только пространственной избыточности. Именно с этого кадра начинается декодирование изображения в последовательности;
- ***P*** — *Predictive кадры*. «Предсказанные» кадры, при формировании которых используется метод предсказания изображения на следующем кадре с учетом компенсации движения от последнего ***I*** или ***P-кадра*** перед формируемым. ***P-кадр*** также служит для дальнейшего предсказания изображения. ***P-кадр*** создается с помощью межкадровой компрессии, уменьшающей как пространственную, так и временную избыточность. Изображение ***P-кадра*** вычитается из следующего изображения и эта разница кодируется и вместе с вектором движения добавляется к сжатым данным;
- ***B*** — *Bi-directional*, «двунаправленные» кадры. Они названы так потому, что хранят наиболее существенную информацию с окружающих их ***I*** и ***P-кадров***. В кадры имеют наивысшую степень компрессии, но требуют предыдущего и последующего изображения для компенсации движения объектов на изображении.

Такую структуру ***MPEG*** потока обычно описывают в виде дроби M/N, для которой M сообщает общее число кадров в ***GOP***, а N — каким по счету будет очередной ***P-кадр*** после предыдущего. Таким образом, ***GOP***-последовательность, изображенная на рис. 4.36 выше, может быть записана как 12/3. Собственно поток данных ***MPEG*** состоит из 6 иерархических уровней:

- ***Блок*** — данные по яркости и цветности для блоков 8x8 пикселей изображения. Блоки анализируются по значениям Y (яркость), CB и CR (цветоразностные сигналы).
- ***Макроблок*** — как следует из названия, состоит из 4 простых блоков в окне 16x16 пикселей соответственно. В формате 4:2:0 макроблок содержит 4 блока яркостных данных Y и по одному CB и CR.
- ***Слой*** — содержит несколько смежных макроблоков.
- ***Кадр*** — состоит из группы слоев, содержащих изображение, которое, в свою очередь, может быть как I, так P или B.
- ***Группа изображений*** (она же ***GOP***) — содержит последовательность кадров. Может включать до 15 кадров и должна обязательно начинаться с I кадра.
- ***Видеопоследовательность*** — должна содержать минимум одну GOP, а также заголовок в начале последовательности и код конца последовательности.

Уровни и профили MPEG. Под профилем ***MPEG*** понимается подмножество структуры битового потока сжатого видеоизображения. В пределах такого подмножества возможен широкий разброс параметров потока и, соответственно, кодеров и декодеров для них.

Под уровнем понимается ряд ограничений, применяемых к параметрам ***MPEG***-потока, например, разрешение выходного изображения, частота кадров и т.п. Таблица 4.2 иллюстрирует максимальные значения ограничений, накладываемых на уровни и профили ***MPEG***.

Определяющую роль играют всего несколько наиболее важных параметров настройки MPEG-кодера (компрессора):

- ***Bit Rate*** (*Скорость потока*) — измеряется в мегабитах (обратите внимание, в мегабитах, а не в мегабайтах) в секунду. Чем выше ***Bit Rate***, тем выше качество изображения, но тем больше места занимает созданный кодером ***MPEG*** файл.
- ***Constant Bit Rate*** (*Постоянная скорость потока*) — параметр кодирования, указывающий на то, что скорость потока не должна зависеть от кодируемого изображения и быть постоянной величиной. Установка постоянной скорости позволяет точно определить размер итогового файла фильма, но не оказывает влияния на качество изображения только в случае одной и той же ди-

намики фильма на всем его протяжении. Например, в течение фильма непрерывно показываются автомобильные гонки или медленно и плавно картины в музее.

Таблица 4.2

Профиль/ уровень	***Основной I, P, B 4:2:0 (Main profile)***	***I,P,B 4:2:0 (SNR Scalable Profile)***	***Высокий I,P,B 4:2:0 или 4:2:2 (High Profile)***
Самый высокий (High Level)	1920x1152 80 Mbit/s	_	1920x1152 100 Mbit/s
Высокий (High 1440 Level)	1440x1152 60 Mbit/s	1440x1152 60 Mbit/s	1440x1152 80 Mbit/s
Основной (Main Level)	720x576 15 Mbit/s	720x576 15 Mbit/s	720x608 20 Mbit/s
Низкий (Low Level)	352x288 4 Mbit/s	352x288 4 Mbit/s	_

- ***Variable Bit Rate*** (Переменная скорость потока) — параметр кодирования, указывающий на то, что скорость потока должна зависеть от динамики изображения — расти на динамичных сценах и уменьшаться на статичных. ***Variable Bit Rate*** применяется сейчас наиболее широко, так как позволяет в подавляющем большинстве случаев добиться лучшего качества видео по сравнению с ***Constant Bit Rate*** при том же размере файла результата.
- ***Average Bit Rate*** (Средняя скорость потока) — параметр, совпадающий по значению с ***Constant Bit Rate*** при постоянной скорости потока и оговаривающий среднюю скорость потока с ***Variable Bit Rate***.

Методы разностного кодирования

К методам сжатия относят также ***методы разностного кодирования***, поскольку разности амплитуд отсчетов представляются меньшим числом разрядов, чем сами амплитуды. Разностное кодирование реализовано в методах дельта-модуляции и ее разновидностях.

*Предсказывающие (**предиктивные**) методы* основаны на экстраполяции значений амплитуд отсчетов, и если выполнено условие

$$\left|A_R - A_P\right| > d\,,$$

то отсчет должен быть передан, иначе он является избыточным; здесь A_R и A_P — амплитуды реального и предсказанного отсчетов, d — допуск (допустимая погрешность представления амплитуд). Иллюстрация предсказывающего метода с линейной экстраполяцией представлена рис. 4.37.

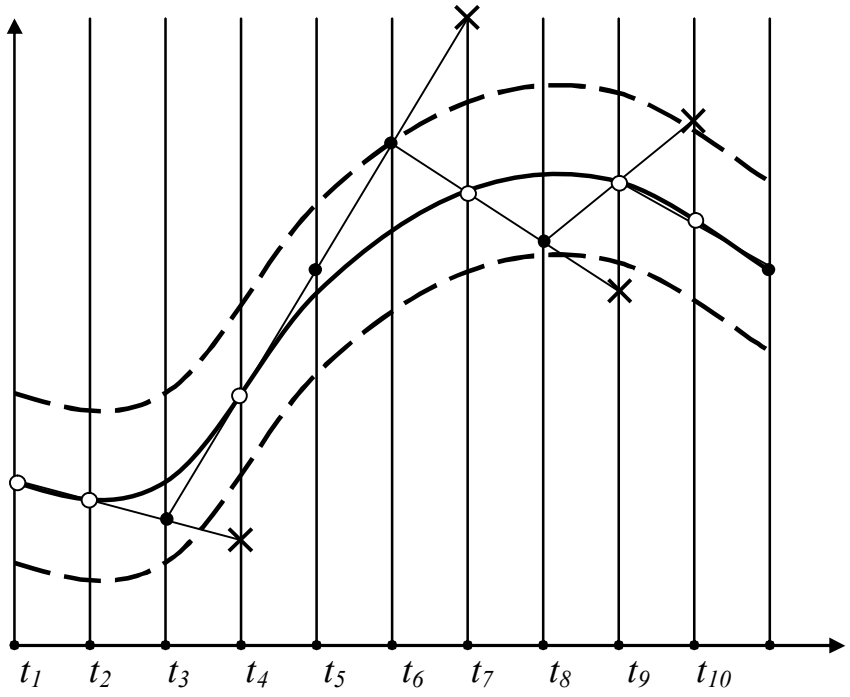


Рис. 4.37 — Предиктивное кодирование

Здесь точками показаны предсказываемые значения сигнала. Если точка выходит за пределы *коридора* (допуска *d*), показанного пунктирными линиями, то происходит передача отсчета. На рисунке передаваемые отсчеты отмечены светлыми кружками в моменты времени t_1, t_2, t_4, t_7, t_9, t_{10}. Если передачи отсчета нет, то на приемном конце рассчитывается экстраполированное значение (обозначенное на рисунке темным кружком).