

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ
Государственное образовательное учреждение высшего профессионального образования
«ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Ю.В. Алхимов

СОВРЕМЕННЫЕ КОММУНИКАЦИОННЫЕ СИСТЕМЫ

Учебное пособие для магистров
по направлению 551500 (магистерская программа 551523 –
Приборы и методы контроля качества и диагностики)

Издательство
Томского политехнического университета
2008

УДК 681.31

А

Алхимов Ю.В.

А Современные коммуникационные системы: учебное пособие / Ю.В. Алхимов, В.К. Кулешов. – Томск: Изд. ТПУ, 2008. – 200 с.

Учебное пособие посвящено современным методам передачи информации и построения компьютерных сетей обмена данными.

Работа предназначена для студентов и аспирантов обучающихся по направлению 551500 – Приборостроение, а также может быть полезно студентам других специальностей.

УДК 681.31

Рекомендовано к печати Редакционно-Издательским
Советом Томского политехнического университета

© Томский политехнический университет, 2008

© Алхимов Ю.В., Кулешов В.К., 2008

© Оформление. Издательство Томского
политехнического университета, 2008

ВВЕДЕНИЕ

Развитие средств коммуникаций во все времена было одним из главных направлений развития науки и техники. Данное учебное пособие представляет собой обзор современных систем передачи информации и сетевых технологий и стандартов. В нем наряду с проводными технологиями построения коммуникационных сетей большое внимание уделено беспроводным технологиям, определяющим развитие средств коммуникации сегодня. Множество факторов, включая возрастающую конкуренцию и внедрение цифровых технологий, привели к беспрецедентному росту на рынке беспроводных коммуникаций.

Открытие и начало применения радио в конце XIX века было естественным следствием развития исследований по электричеству. Фактическое изобретение радиосвязи должно быть приписано Никола Тесла, который провел общественную демонстрацию явления в 1893 году.

7 мая 1895 года на заседании Русского физико-химического общества в Петербурге А.С. Попов продемонстрировал действие своего прибора, явившегося, по сути дела, первым в мире радиоприемником. День 7 мая стал днем рождения радио. Ныне он ежегодно отмечается в нашей стране. А. С. Попов продолжал настойчиво совершенствовать приемную аппаратуру. Он ставил своей непосредственной задачей построить прибор для передачи сигналов на большие расстояния. Вначале радиосвязь была установлена на расстоянии 250 м. Неустанно работая над своим изобретением, Попов вскоре добился дальности связи более 600 м. При участии А. С. Попова началось внедрение радиосвязи на флоте и в армии России. На маневрах Черноморского флота в 1899 году ученый установил радиосвязь на расстоянии свыше 20 км. В начале 1900 года радиосвязь была успешно использована во время спасательных работ в Финском заливе – благодаря сообщению, переданному по радио, ледокол «Ермак» снял со льдины рыбаков, которых шторм унес в море. В 1901 году дальность радиосвязи была уже 150 км. За границей усовершенствование подобных приборов проводилось фирмой, организованной итальянским инженером Гульельмо Маркони. В 1901 году он послал телеграфные сигналы через Атлантический океан на расстояние в 1800 миль.

Радио, начавшее свою практическую историю спасением людей, стало новым прогрессивным видом связи XX и XXI века. За прошлое столетие, усовершенствования в беспроводных технологиях привели к радио, телевидению, мобильному телефону и спутникам связи. Все типы информации можно теперь послать почти в каждый уголок мира. В настоящее

время, основное внимание ученых сосредоточено на спутниковой связи, беспроводной организации сетей и технологиях сотовой связи.

Спутники связи были запущены в 1960-х. Те первые спутники могли обеспечить только 240 голосовых каналов. Сегодня спутники несут приблизительно одну треть трафика голосовых сообщений и обеспечивают передачу всех телевизионных сигналов между странами. Современные спутники обычно вводят задержку распространения сигналов, которые они обрабатывают, в четверть секунды. Новые спутники находящиеся на более низких орбитах и с меньшей задержкой сигнала скоро будут развернуты, чтобы обеспечить службы передачи данных, типа доступа в Интернет.

Беспроводные технологии позволяют развивать глобальные и локальные сети без развития кабельных сетей. В настоящее время быстрыми темпами развивается стандарт IEEE 802.11 как стандарт для беспроводных локальных сетей. Для обеспечения беспроводной связи между устройствами широко используется стандарт Bluetooth.

Сотовый или мобильный телефон это современный эквивалент проводного телефона, обеспечивающий двухстороннюю коммуникацию. Радиотелефоны первого поколения использовали аналоговую технологию. Эти устройства были тяжелы, и стабильность связи была не всегда удовлетворительной, но они успешно демонстрировали преимущества мобильной связи перед проводной. Текущее поколение беспроводных устройств построено, используя цифровые технологии вместо аналоговых. Цифровые сети могут нести намного больше трафика и обеспечивать лучший прием и безопасность чем аналоговые сети. Кроме того, цифровые технологии сделали возможным введение дополнительных служб, типа идентификации вызывающего. Беспроводные устройства следующего поколения также будут цифровыми и будут иметь соединения с Интернет, используя новые частотные диапазоны и более высокие скорости передачи и обработки информации.

Влияние беспроводных радио технологий было и продолжит быть очень значительным. Очень немногие изобретения были в состоянии «сократить» мир в такой степени.

Стандарты, которые определяют, как устройства радиосвязи взаимодействуют между собой, быстро сближаются и скоро позволят создать глобальную беспроводную сеть, которая сможет предоставить большое разнообразие служб.

В настоящее время мы наблюдаем настоящую «сотовую революцию». В 1990 году число пользователей сотовой связи было приблизительно 11 миллионов. Сейчас это число более 1 миллиарда. Устройства следующего поколения, с доступом к Интернет, только усилят эту тен-

денцию. По некоторым оценкам число устройств подключенных к беспроводному Интернету скоро превысит число проводных Интернет пользователей.

Телефоны в настоящее время самый очевидный показатель успеха радио. С 1996 года число новых подписчиков мобильного телефона превысило число новых подписчиков на стационарные телефоны. Есть много причин, почему это случилось. Мобильные телефоны удобны, они двигаются с людьми. Кроме того, по природе они обеспечивают связь независимо от местоположения разговаривающего. Мобильный телефон общается с базовыми станциями, которые располагаются в стационарных местах, образуя структуру как бы похожую на пчелиные соты.

Технические новшества внесли свой вклад в успех мобильных телефонов. Телефонные трубки стали меньше и легче, срок службы аккумулятора увеличился, а цифровые технологии улучшили прием и позволили лучше использовать спектр радиосигнала. Как со многими типами цифрового оборудования, уменьшились затраты, связанные с мобильными телефонами.

Во многих географических областях мобильные телефоны единственный экономичный способ обеспечить телефонное обслуживание населения. Операторы сотовой связи могут установить базовые станции быстро и с меньшими затратами чем при прокладывании медного кабеля в земляных траншеях.

Мобильные телефоны только одно из проявлений сотовой революции. Появляется все больше и больше новых типов беспроводных устройств. Эти новые устройства имеют доступ к Интернет. Они включают органайзеры и телефоны, имеют доступ к сети, мгновенную передачу сообщений, электронную почту, и другие службы, доступные в Интернет. Беспроводные устройства в автомобилях позволяют пользователям загружать карты и указания как оптимально достичь точки назначения. Скоро беспроводные устройства смогут вызвать помощь, если произошел несчастный случай, или информировать пользователя о самой дешевой топливозаправочной станции в непосредственной близости от его местоположения. Будут доступны и другие удобства. Например, холодильники могут однажды быть в состоянии заказать, через Интернет продукты, которые закончились.

Если вначале беспроводные технологии были ориентированы в основном на голосовую связь, то теперь внимание переключилось на передачу данных. Ожидается, что беспроводные технологии передачи данных превратятся скоро в многомиллиардный рынок. Большая часть этого рынка – беспроводной Интернет. Беспроводные пользователи будут использовать Интернет по-другому, чем стационарные пользовате-

ли. Беспроводные устройства ограничили возможности устройств отображения информации (дисплеев) и ввода данных по сравнению с типичными стационарными устройствами типа персонального компьютера. Поэтому изменится сам стиль общения с беспроводными устройствами. Поскольку беспроводные устройства привязаны к конкретной базовой станции, информация может быть приспособлена к географическому местоположению пользователя. Информация будет в состоянии найти пользователей, вместо пользователей, ищущих информацию.

Сегодня нет единственной сотовой беспроводной сети. Устройства поддерживают один или два из несметного числа технологий и вообще работают только в пределах границ сети единственного оператора. Чтобы это преодолеть, нужно проделать много работы по разработке новых и сближению уже существующих стандартов сотовой связи

Международный Телекоммуникационный Союз (ITU) работает, чтобы развить семью стандартов для беспроводных устройств следующего поколения. Новые стандарты будут использовать более высокие частоты, чтобы увеличить пропускную способность сетей. Новые стандарты также помогут преодолеть несовместимость, которая существовала в сетях первого и второго поколений.

Основные беспроводные системы второго поколения – Глобальная Система для Мобильной Связи (Global System for Mobile Communications, GSM), Персональная Служба Связи (Personal Communications Service, PCS) IS-136 и IS-95. Стандарт IS-136 использует коллективный доступ с разделением времени использования (Time Division Multiple Access, TDMA), в то время как IS-95 использует доступ с разделением кода (Code Division Multiple Access, CDMA). GSM и IS-136 используют специализированные каналы для служб передачи данных на скорости в 9.6 кБит/с.

ITU развивает стандарты IMT-2000 (International Mobile Telecommunications-2000). Эта семья стандартов предназначена обеспечить единую глобальную беспроводную сеть. Стандарты развиваются вокруг полосы частот 2 ГГц. Новые стандарты и полоса частот обеспечат скорости передачи данных до 2 МБит/с.

В дополнение к стандартам описывающим, использование диапазонов частот, методов кодирования и передачи данных, также необходимо определить стандарты, описывающие, как мобильные устройства будут взаимодействовать с Интернет. Несколько международных комитетов и консорциумов промышленных фирм работают в этом направлении. Форум Беспроводного Прикладного Протокола (Wireless Application Protocol, WAP) развивает общий протокол, который позволяет устройствам с ограниченными дисплеем и возможностями ввода

данных взаимодействовать с Интернет. Комитет IETF (Internet Engineering Task Force) развивает мобильный стандарт IP, который приспособливает общепринятый в проводных сетях протокол IP к работе в мобильной окружающей среде.

Интернет все более и более использует мультимедиа. Графика, видео и аудио в большом количестве присутствуют на страницах Всемирной паутины. Деловая связь следует за той же тенденцией. Например, электронная почта часто включает объемные мультимедиа вложения. Чтобы полностью отвечать современным потребностям, беспроводные сети требуют тех же самых высоких скоростей передачи данных как и стационарные сети. Более высокие скорости передачи данных доступны в широкополосной беспроводной технологии.

Широкополосные беспроводные службы имеют те же самые преимущества беспроводного обслуживания: удобство и меньшая стоимость. Операторы могут развернуть обслуживание быстрее и без развертывания кабельной сети.

Существует много инициативных групп, развивающих широкополосные стандарты. Стандарты охватывают сегменты от беспроводных локальных сетей до маленьких беспроводных домашних сетей. Скорости передачи данных изменяются от 2 Мбит/с до более чем 100 Мбит/с. Многие из этих технологий уже доступны, и еще многие станут доступными в ближайшее время.

Беспроводные локальные сети (WLANs) обеспечивают сетевое обслуживание там, где это трудно или слишком дорого обеспечить с использованием проводных сетей. Основные стандарты WLAN: IEEE 802.11b и HiperLAN. Стандарт IEEE предусматривает скорость передачи данных в 11 Мбит/с. Европейский стандарт HiperLAN определяет максимальную скорость передачи 24 Мбит/с, и в будущем до 54 Мбит/с.

Потенциальная проблема с 802.11b это совместимость с Bluetooth. Bluetooth – спецификация беспроводной связи, которая определяет обмен данными между устройствами типа портативных компьютеров, карманных компьютеров и мобильных телефонов. Bluetooth и 802.11b используют одинаковую полосу частот. Технологии, если их применить одновременно в одном и том же приборе, мешают друг другу.

Инициативная группа HomeRF развивает стандарты, чтобы определить беспроводной обмен данными между устройствами типа домашних компьютеров и внешних устройств. HomeRF в настоящее время работает со скоростями передачи до 2 Мбит/с, но новый пересмотр стандарта повысит скорости до 10 Мбит/с.

Беспроводные технологии удобны и часто менее дороги, чем проводные, но эти технологии не идеальны. Есть ограничения, политиче-

ские и технические трудности, которые могут, в конечном счете, препятствовать беспроводным технологиям достигать их полного потенциала. Основных проблем две – несовместимые стандарты и ограничения самих устройств, использующих беспроводные технологии.

Уже упоминалось, что есть три стандарта для цифровой сотовой связи. Устройство, использующее IS-136, не будет работать на территории, на которой использована технология IS-95. Также упоминалась невозможность использовать Bluetooth и 802.11b в одном и том же устройстве. Это только два примера проблем, которые возникают, когда всеобъемлющего стандарта не существует. Нехватка такого стандарта сдерживает развитие беспроводных технологий до истинного идеала: доступ к данным в любое время и в любом месте.

Ограничения, присущие мобильным устройствам также ограничивают свободный обмен данными. Маленький жидкокристаллический монитор на мобильном телефоне не годится для того, чтобы отображать большое количество информации. Кроме того, мобильные беспроводные устройства не могут обратиться к огромному большинству сайтов WWW в Интернет. Браузеры в мобильных устройствах используют специальный язык WML вместо стандартного HTML.

Вероятно, ни одно беспроводное устройство не будет в состоянии обеспечить все потребности. Беспроводной потенциал может быть реализован, но не в одном единственном устройстве. Радио преуспеет, потому что это будет разнообразие устройств, которые могут реализовать все разнообразие потребностей.

ГЛАВА 1. ПОНЯТИЕ ИНФОРМАЦИИ

Предмет нашего курса – изучение средств передачи информации на расстояние и построение на их основе сетей передачи данных. Таким образом при изучении курса мы будем иметь дело с понятием информации.

1.1. Что такое информация

Любая наука начинается со строгих определений используемых ею понятий и терминов. Поэтому было бы вполне разумным начать изложение основ теории информации именно с ее точного определения. Определить какое-либо понятие – значит выразить его через другие понятия, уже определенные ранее. Сложность ситуации, однако, в том, что информация является одной из исходных категорий мироздания, и, следовательно, определение «информации вообще» невозможно свести к каким-то более простым, более «исходным» терминам. Что касается частных трактовок понятия «информация», то следует отметить значительное их расхождение в различных научных дисциплинах, в технике и на бытовом уровне. Такое положение не следует считать каким-то необычным – можно привести много аналогичных примеров, когда термин имеет и используется во множестве значений: движение, энергия, система, связь, язык и пр. Неоднозначность преодолевается тем, что в каждой «узкой» дисциплине дается свое определение термина. Его следует считать частным – и именно оно используется. Но это, безусловно, не дает основания переносить такое определение и применять его вне рамок данной дисциплины. Например, в теоретической механике «связь» определяется как некое внешнее воздействие, ограничивающее возможности перемещения (степени свободы) тела; нет смысла такую трактовку пытаться применять, скажем, в телеграфии или социальных науках.

Аналогична ситуация и с термином «информация». На бытовом уровне и во многих научных дисциплинах он ассоциируется с понятиями: сведения, знания, данные, известие, сообщение, управление и др. Общим во всех перечисленных примерах является то, что в них существенным и значимым для использования является содержательная сторона информации. С позиций «здравого смысла» это представляется вполне естественным. Однако оценка смысла и ценности одной и той же информации различными людьми, вообще говоря, будет различной; объективная количественная мера смысловой стороны информации отсутствует.

Отделив информацию от ее семантически-содержательной основы, мы получаем возможность построить определение информации и параллельно ввести ее объективную количественную меру. Открытие та-

кого способа определения информации является одной из главных заслуг теории информации.

Обсудим особенность, которой обладает любая информация, – это то, что информация – категория нематериальная. Следовательно, для существования и распространения в нашем материальном мире она должна быть обязательно связана с какой-либо материальной основой – без нее информация не может проявиться, передаваться и сохраняться, например, восприниматься и запоминаться нами.

Материальный объект или среду, которые служат для представления или передачи информации, будем называть ее **материальным носителем**.

Однако не с любым процессом можно связать информацию. В частности, стационарный процесс, т. е. процесс с неизменными в течение времени характеристиками, информацию не переносит. Примером может служить постоянный электрический ток, ровное горение лампы, или равномерный гул – они содержат лишь ту информацию, что процесс идет, т. е. что-то функционирует.

Изменение характеристики носителя, которое используется для представления информации, называется **сигналом**, а значение этой характеристики, отнесенное к некоторой шкале измерений, называется **параметром сигнала**.

В табл. 1.1 приведены примеры процессов, используемых для передачи информации, и связанных с ними сигналов.

Таблица 1.1

Информационные процессы и связанные с ними сигналы

Способ передачи	Процесс	Параметры сигнала
Звук	Звуковые волны	Высота и громкость звука
Радио, телевидение	Радиоволны	Частота, амплитуда или фаза радиоволны
Изображение	Световые волны	Частота и амплитуда световых волн
Телефон, компьютерная сеть	Электрический ток	Частота и амплитуда электрических колебаний в линии связи

Однако одиночный сигнал, как мы увидим в дальнейшем, не может содержать много информации. Поэтому для передачи информации используется ряд следующих друг за другом сигналов.

Последовательность сигналов называется **сообщением**.

Таким образом, от источника к приемнику информация передается в виде сообщений. Можно сказать, что сообщение выступает в качестве

материальной оболочки для представления информации при передаче. Следовательно, сообщение служит переносчиком информации, а информация является содержанием сообщения.

Обсудим следующее исходное понятие – информационный процесс. Вообще термин «процесс» применяется в тех случаях, когда некоторое качество, характеризующее систему или объект, меняется с течением времени в результате внешних воздействий или каких-то внутренних причин. Какие атрибуты могут изменяться с течением времени у нематериальной информации? Очевидно, только ее содержание и материальная оболочка, посредством которого информация представлена, т. е. сообщение. В связи с этим примем следующее определение:

Информационный процесс – это изменение с течением времени содержания информации или представляющего его сообщения.

Различных видов информационных процессов оказывается немного:

- порождение (создание) новой информации;
- преобразование информации (т. е. порождение новой информации в результате обработки имеющейся);
- уничтожение информации;
- передача информации (распространение в пространстве).

С передачей информации связана еще одна пара исходных сопряженных понятий – источник и приемник информации.

Источник информации – это субъект или объект, порождающий информацию и представляющий ее в виде сообщения.

Приемник информации – это субъект или объект, принимающий сообщение и способный правильно его интерпретировать.

В этих определениях сочетание «субъект или объект» означает, что источники и приемники информации могут быть одушевленными (человек, животные) или неодушевленными (технические устройства, природные явления). Для того чтобы объект (или субъект) считался источником информации, он должен не только ее породить, но и иметь возможность инициировать какой-то нестационарный процесс и связать информацию с его параметрами, т. е. создать сообщение. Например, если человек что-то придумал, но держит это в своем мозге, он не является источником информации; однако он им становится, как только свою идею изложит на бумаге (в виде текста, рисунка, схемы и пр.) или выскажет словами.

В определении приемника информации важным представляется то, что факт приема сообщения еще не означает получение информации; информация может считаться полученной только в том случае, если приемнику известно правило интерпретации сообщения. Другими словами, понятия «приемник сообщения» и «приемник информации» не

тождественны. Например, слыша речь на незнакомом языке, человек оказывается приемником сообщения, но не приемником информации.

1.2. Формы представления информации

В предыдущем параграфе было сказано, что передача информации производится с помощью сигналов, а самым сигналом является изменение некоторой характеристики носителя с течением времени. При этом в зависимости от особенностей изменения этой характеристики (т. е. параметра сигнала) с течением времени выделяют два типа сигналов: непрерывные и дискретные.

Сигнал называется **непрерывным** (или аналоговым), если его параметр может принимать любое значение в пределах некоторого интервала.

Если обозначить Z – значение параметра сигнала, а t – время, то зависимость $Z(t)$ будет непрерывной функцией (рис. 1.1, а).

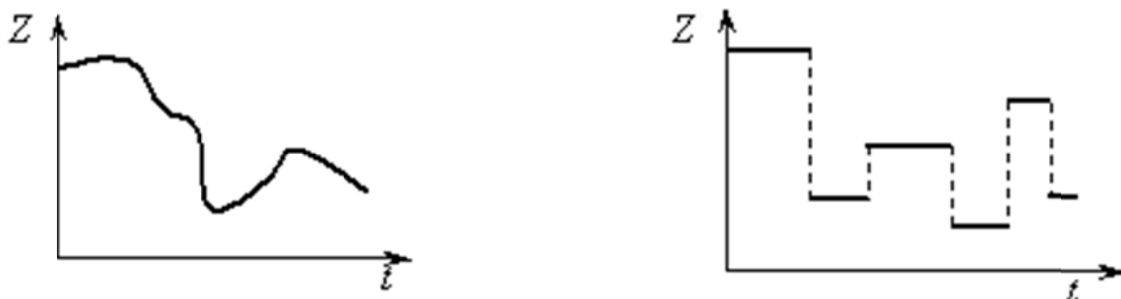


Рис. 1.1. Непрерывные и дискретные сигналы

Примерами непрерывных сигналов являются речь и музыка, изображение, показание термометра (параметр сигнала – высота столба спирта или ртути имеет непрерывный ряд значений) и пр.

Сигнал называется **дискретным**, если его параметр может принимать конечное число значений в пределах некоторого интервала.

Пример дискретных сигналов представлен на рис. 1.1, б. Как следует из определения, дискретные сигналы могут быть описаны дискретным и конечным множеством значений параметров $\{Z\}$. Примерами устройств, использующих дискретные сигналы, являются часы (электронные и механические), цифровые измерительные приборы, книги, табло и пр.

Поскольку последовательность сигналов есть сообщение, качество прерывности-непрерывности сигналов переносится и на сообщение – существуют понятия «непрерывное сообщение» и «дискретное сообщение». Очевидно, что дискретным будет считаться сообщение, построенное из дискретных сигналов.

Принципиальным и важнейшим различием непрерывных и дискретных сигналов является то, что дискретные сигналы можно обозна-

чить, т. е. приписать каждому из конечного числа возможных значений сигнала знак, который будет отличать данный сигнал от другого.

Знак – это элемент некоторого конечного множества отличных друг от друга сущностей.

Природа знака может любой – жест, рисунок, буква, сигнал светофора, определенный звук и т. д. Природа знака определяется носителем сообщения и формой представления информации в сообщении.

Вся совокупность знаков, используемых для представления дискретной информации, называется набором знаков. Таким образом, набор есть дискретное множество знаков.

Набор знаков, в котором установлен порядок их следования, называется **алфавитом**.

Следовательно, алфавит – это упорядоченная совокупность знаков. Порядок следования знаков в алфавите называется лексикографическим. Благодаря этому порядку между знаками устанавливаются отношения «больше–меньше»: для двух знаков ξ и ψ принимается, что $\xi < \psi$, если порядковый номер у ξ в алфавите меньше, чем у ψ .

Вернемся к обсуждению информационных процессов, связанных с преобразованием одних сигналов в другие. Ясно, что технически это осуществимо. Ранее сигналы и их последовательности – сообщения – были названы нами «материальными оболочками для информации», и, естественно, встает вопрос: что происходит с его содержимым, т. е. с информацией, при изменении «оболочки»? Попробуем найти ответ на него.

Поскольку имеются два типа сообщений, между ними, очевидно, возможны четыре варианта преобразований (рис. 1.2).

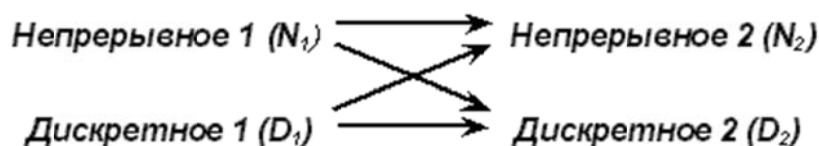


Рис. 1.2. Типы преобразований

Осуществимы и применяются на практике все четыре вида преобразований. Рассмотрим примеры устройств и ситуаций, связанных с такими преобразованиями, и одновременно попробуем проследить, что при этом происходит с информацией.

Примерами устройств, в которых осуществляется преобразование типа $N_1 \rightarrow N_2$, являются микрофон (звук преобразуется в электрические сигналы); магнитофон и видеоманитофон (чередование областей намагничивания ленты превращается в электрические сигналы, которые затем преобразуются в звук и изображение); телекамера (изображение и звук превращаются в

электрические сигналы); радио- и телевизионный приемник (радиоволны преобразуются в электрические сигналы, а затем в звук и изображение); аналоговая вычислительная машина (одни электрические сигналы преобразуются в другие). Особенностью данного варианта преобразования является то, что оно всегда сопровождается частичной потерей информации. Потери связаны с помехами (шумами), которые порождает само информационное техническое устройство и которые воздействуют извне. Эти помехи пришиваются к основному сигналу и искажают его. Поскольку параметр сигнала может иметь любые значения (из некоторого интервала), то невозможно отделить ситуации: был ли сигнал искажен или он изначально имел такую величину. В ряде устройств искажение происходит в силу особенностей преобразования в них сообщения, например в черно-белом телевидении теряется цвет изображения; телефон пропускает звук в более узком частотном интервале, чем интервал человеческого голоса; кино- и видеоизображение оказываются плоскими, они утратили объемность.

Теперь обсудим общий подход к преобразованию типа $N \rightarrow D$. С математической точки зрения перевод сигнала из аналоговой формы в дискретную означает замену описывающей его непрерывной функции времени $Z(t)$ на некотором отрезке $[t_1, t_2]$ конечным множеством (массивом) $\{Z_i, t_i\}$ ($i=0 \dots n$, где n – количество точек разбиения временного интервала). Подобное преобразование называется дискретизацией непрерывного сигнала и осуществляется посредством двух операций: развертки по времени и квантования по величине сигнала.

Развертка по времени состоит в том, что наблюдение за значением величины Z производится не непрерывно, а лишь в определенные моменты времени с интервалом Δt :

$$\Delta t = \frac{t_n - t_0}{n}.$$

Квантование по величине это отображение вещественных значений параметра сигнала в конечное множество чисел, кратных некоторой постоянной величине – шагу квантования (ΔZ).

Совместное выполнение обеих операций эквивалентно нанесению масштабной сетки на график $Z(t)$, как показано на рис. 1.3. Далее, в качестве пар значений $\{Z_i, t_i\}$ выбираются узлы сетки, расположенные наиболее близко к $Z(t_i)$. Полученное таким образом множество узлов называется дискретным представлением исходной непрерывной функции. Таким образом, любое сообщение, связанное с $Z(t)$, может быть преобразовано в дискретное, т. е. представлено посредством некоторого алфавита.

При такой замене довольно очевидно, что чем меньше n (больше Δt), тем меньше число узлов, но и точность замены $Z(t)$ значениями Z_i будет

меньшей. Другими словами, при дискретизации может происходить потеря части информации, связанной с особенностями функции $Z(t)$. На первый взгляд кажется, что увеличением количества точек n можно улучшить соответствие между получаемым массивом и исходной функцией, однако полностью избежать потерь информации все равно не удастся, поскольку n величина конечная. Ответом на эти сомнения служит так называемая теорема отсчетов, доказанная в 1933 г. В.А. Котельниковым (по этой причине ее иногда называют его именем), значение которой для решения проблем передачи информации было осознано лишь в 1948 г. после работ К. Шеннона. Теорема, которую мы примем без доказательства, но результаты будем в дальнейшем использовать, гласит:

Непрерывный сигнал можно полностью отобразить и точно воссоздать по последовательности измерений или отсчетов величины этого сигнала через одинаковые интервалы времени, меньшие или равные половине периода максимальной частоты, имеющейся в сигнале.

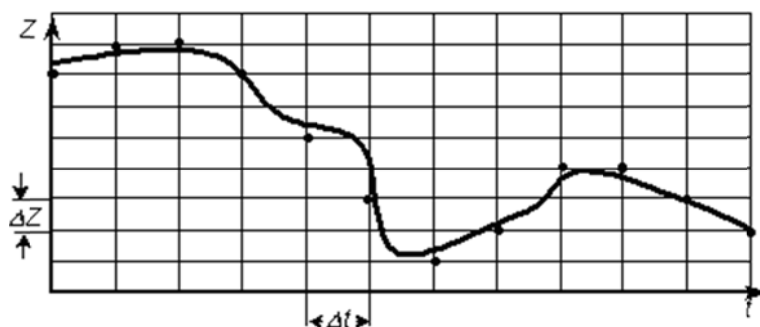


Рис. 1.3. Дискретизация аналогового сигнала за счет операций развертки по времени и квантования по величине

Комментарии к теореме:

1. Теорема касается только тех линий связи, в которых для передачи используются колебательные или волновые процессы. Это не должно восприниматься как заметное ограничение, поскольку действие большинства практических устройств связи основано именно на этих процессах.
2. Любое подобное устройство использует не весь спектр частот колебаний, а лишь какую-то его часть; например, в телефонных линиях используются колебания с частотами от 300 Гц до 3400 Гц. Согласно теореме отсчетов определяющим является значение верхней границы частоты. Обозначим его ν_m .

Смысл теоремы в том, что дискретизация не приведет к потере информации и по дискретным сигналам можно будет полностью восстановить исходный аналоговый сигнал, если развертка по времени выполнена в соответствии со следующим соотношением:

$$\Delta t \leq \frac{1}{2\nu_m}. \quad (1.1)$$

Например, для точной передачи речевого сигнала с частотой до $\nu_m = 4000$ Гц при дискретной записи должно производиться не менее 8000 отсчетов в секунду; в телевизионном сигнале $\nu_m = 4$ МГц, следовательно, для его точной передачи потребуется около 8000000 отсчетов в секунду.

Однако, помимо временной развертки, дискретизация имеет и другую составляющую – квантование. Какими соображениями определяется шаг квантования ΔZ ? Любой получатель сообщения (человек или устройство) всегда имеют конечную предельную точность распознавания величины сигнала. Например, человеческий глаз в состоянии различить около 16 миллионов цветовых оттенков; это означает, что при квантовании цвета нет смысла делать большее число градаций. При передаче речи достаточной оказывается гораздо меньшая точность – около 1%; следовательно, для амплитуды звуковых колебаний $\Delta Z \leq 0,01 \cdot Z_{\max}$, а алфавит для обозначения всех градаций громкости должен содержать 100 знаков. Мы приходим к заключению, что шаг квантования определяется чувствительностью приемного устройства.

К достоинствам дискретной формы следует отнести:

- высокую помехоустойчивость;
- простоту и, как следствие, надежность и относительную дешевизну устройств по обработке информации;
- точность обработки информации, которая определяется количеством обрабатываемых элементов и не зависит от точности их изготовления;
- универсальность устройств.

1.3. Измерение информации

Начнем с простой ситуации, когда опыт имеет n равновероятных исходов. Очевидно, что неопределенность каждого из них зависит от n , т. е. мера неопределенности является функцией числа исходов ($f(n)$)

Можно указать некоторые свойства этой функции:

- 1) $f(1) = 0$, поскольку при $n = 1$ исход опыта не является случайным и, следовательно, неопределенность отсутствует;
- 2) $f(n)$ возрастает с ростом n , поскольку чем больше число возможных исходов, тем более затруднительным становится предсказание результата опыта.

Для определения явного вида функции $f(n)$ рассмотрим два независимых опыта α и β с количествами равновероятных исходов, соответственно n_α и n_β . Пусть имеет место сложный опыт, который состоит в

одновременном выполнении опытов α и β ; число возможных его исходов равно $n_\alpha \cdot n_\beta$, причем все они равновероятны. Очевидно, что неопределенность исхода такого сложного опыта будет больше неопределенности опыта α , поскольку к ней добавляется неопределенность β . Мера неопределенности сложного опыта равна $f(n_\alpha \cdot n_\beta)$. С другой стороны, меры неопределенности отдельных опытов α и β составляют, соответственно, $f(n_\alpha)$ и $f(n_\beta)$. В первом случае (сложный опыт) проявляется общая (суммарная) неопределенность совместных событий, во втором – неопределенность каждого из событий в отдельности. Однако из независимости α и β следует, что в сложном опыте они никак не могут повлиять друг на друга и, в частности, α не может оказать воздействия на неопределенность β , и наоборот. Следовательно, мера суммарной неопределенности должна быть равна сумме мер неопределенности каждого из опытов, т. е. мера неопределенности аддитивна:

$$f(n_\alpha \cdot n_\beta) = f(n_\alpha) + f(n_\beta). \quad (1.2)$$

Теперь задумаемся о том, каким может быть явный вид функции $f(n)$, чтобы он удовлетворял свойствам (1) и (2) и соотношению (1.2)? Легко увидеть, что такому набору свойств удовлетворяет функция $\log(n)$, причем можно доказать, что она единственная из всех существующих классов функций. Таким образом, за меру неопределенности опыта с n равновероятными исходами можно принять число $\log(n)$.

Следует заметить, что выбор основания логарифма в данном случае значения не имеет, в силу известной формулы преобразования логарифма от одного основания к другому. Поскольку это так, мы имеем возможность выбрать удобное для нас (из каких-то дополнительных соображений) основание логарифма. Таким удобным основанием оказывается 2, поскольку в этом случае за единицу измерения принимается неопределенность, содержащаяся в опыте, имеющем лишь два равновероятных исхода, которые можно обозначить, например, ИСТИНА (True) и ЛОЖЬ (False) и использовать для анализа таких событий аппарат математической логики.

Единица измерения неопределенности при двух возможных равновероятных исходах опыта называется **бит**.

Таким образом, нами установлен явный вид функции, описывающей меру неопределенности опыта, имеющего n равновероятных исходов:

$$f(n) = \log_2 n \quad (1.3)$$

Эта величина получила название энтропия. В дальнейшем будем обозначать ее H .

Вновь рассмотрим опыт с n равновероятными исходами. Поскольку каждый исход случаен, он вносит свой вклад в неопределенность всего

опыта, но так как все n исходов равнозначны, разумно допустить, что и их неопределенности одинаковы. Из свойства аддитивности неопределенности, а также того, что общая неопределенность равна $\log_2 n$, следует, что неопределенность, вносимая одним исходом составляет

$$\frac{1}{n} \log_2 n = -\frac{1}{n} \log_2 \frac{1}{n} = -p \log_2 p,$$

где $p = \frac{1}{n}$ – вероятность любого из отдельных исходов.

Обобщая это выражение на ситуацию, когда опыт α имеет n неравновероятных исходов $A_1, A_2 \dots A_n$, получим:

$$H(\alpha) = -\sum_{i=1}^n p(A_i) \cdot \log_2 p(A_i). \quad (1.4)$$

Введенная таким образом величина, как уже было сказано, называется энтропией опыта α .

Энтропия является мерой неопределенности опыта, в котором проявляются случайные события, и равна средней неопределенности всех возможных его исходов.

1. Как следует из (1.4), $H = 0$ только в двух случаях:
 - а) какая-либо из $p(A_j) = 1$; однако, при этом следует, что все остальные $p(A_i) = 0$ ($i \neq j$), т. е. реализуется ситуация, когда один из исходов является достоверным (и общий итог опыта перестает быть случайным);
 - б) все $p(A_i) = 0$, т. е. никакие из рассматриваемых исходов опыта невозможны, поскольку нетрудно показать, что $\lim_{p \rightarrow 0} (p \cdot \log p) = 0$.

Во всех остальных случаях, очевидно, что $H > 0$.

2. Очевидным следствием (1.2) будет утверждение, что для двух независимых опытов α и β

$$H(\alpha \wedge \beta) = H(\alpha) + H(\beta). \quad (1.5)$$

3. Энтропия сложного опыта, состоящего из нескольких независимых, равна сумме энтропий отдельных опытов.

Пусть имеется два опыта с одинаковым числом исходов n , но в одном случае они равновероятны, а в другом – нет. Каково соотношение энтропий опытов? Примем без доказательства следующее утверждение:

$$-\sum_{i=1}^n p(A_i) \cdot \log_2 p(A_i) \leq \log_2 n. \quad (1.6)$$

При прочих равных условиях наибольшую энтропию имеет опыт с равновероятными исходами.

Другими словами, энтропия максимальна в опытах, где все исходы равновероятны. Здесь усматривается аналогия (имеющая глубинную первооснову) с понятием энтропии, используемой в физике.

Разность $H(\alpha)$ и $H(\beta)$, очевидно, показывает какие новые сведения мы получаем, произведя опыт. Эта величина называется информацией.

$$I(\alpha, \beta) = H(\beta) - H_\alpha(\beta). \quad (1.7)$$

Данное выражение открывает возможность численного измерения количества информации, поскольку оценивать энтропию мы уже умеем. Из него легко получить ряд следствий:

Следствие 1. Поскольку единицей измерения энтропии является бит, то в этих же единицах может быть измерено количество информации.

Следствие 2. Можно считать, что энтропия равна информации относительно опыта, которая содержится в нем самом.

Можно построить уточнение: энтропия опыта равна той информации, которую мы получаем в результате его осуществления.

Следствие 2 и представление энтропии в виде (1.4) позволяют записать:

$$I = -\sum_{i=1}^n p(A_i) \cdot \log_2 p(A_i), \quad (1.8)$$

т. е. информация опыта равна среднему значению количества информации, содержащейся в каком-либо одном его исходе.

Легко получить следствие формулы для случая, когда все n исходов равновероятны. В этом случае все $p(A_i) = \frac{1}{n}$ и, следовательно

$$I = \sum_{i=1}^n \frac{1}{n} \cdot \log_2 n = \log_2 n.$$

Эта формула была выведена в 1928 г. американским инженером Р. Хартли и носит его имя. Она связывает количество равновероятных состояний (n) и количество информации в сообщении (I). Ее смысл в том, что, если некоторое множество содержит n элементов и x принадлежит данному множеству, то для его выделения (однозначной идентификации) среди прочих требуется количество информации, равное $\log_2 n$.

Информация это содержание сообщения, понижающего неопределенность некоторого опыта с неоднозначным исходом; убыль связанной с ним энтропии является количественной мерой информации. В случае равновероятных исходов информация равна логарифму отношения числа возможных исходов до и после (получения сообщения).

ГЛАВА 2. СРЕДСТВА СВЯЗИ

2.1. Что такое средство связи

Ранее источник информации был определен как объект или субъект, порождающий информацию и представляющий ее в виде сообщения, т. е. последовательности сигналов. При этом человек в информационном взаимодействии с окружающей средой ограничен возможностями собственных органов чувств. Однако спектр процессов, на основе которых производится передача информации, может быть расширен за счет использования средств связи:

Средства связи это совокупность устройств, обеспечивающих преобразование первичного сообщения от источника информации в сигналы заданной физической природы, их передачу, прием и представление в форме удобной потребителю.

Средств связи существует множество: почта, телефон, радио, телевидение, компьютерные сети и пр. Однако при всем разнообразии их конкретной реализации можно выделить общие элементы, представленные на рис. 2.1.



Рис. 2.1. Общая схема передачи информации
(ИИ – источник информации, К – кодер, ПрдС – передатчик сообщения, ПрС – приемник сообщения, ДК – декодер, При – приемник информации)

Источник информации (ИИ) выдает ее в виде первичного сообщения, представленного последовательностью первичных сигналов. Для дальнейшей передачи эти сигналы преобразуются в сигналы такой физической природы, которые могут распространяться в заданном материальном носителе – формируется вторичное сообщение. Примерами преобразователей являются: мегафон или телефонный аппарат, преобразующие голосовые сигналы в электрические; радиопередатчик, преобразующие голосовые сигналы в радиоволны; телекамера, преобразующая изображение в последовательность электрических импульсов; модем, переводящий

высокочастотные компьютерные сигналы в аналоговые низкочастотные и обратно, и пр. В общем случае при преобразовании выходные сигналы не полностью воспроизводят все особенности первичного сообщения, а лишь его существенные стороны, т. е. при преобразовании часть информации теряется. Например, полоса пропускания частот при телефонной связи от 300 до 3400 Гц, в то время как частоты, воспринимаемые человечески ухом, лежат в интервале 16...20000 Гц (т. е. телефонные линии «обрезают» высокие частоты, что приводит к искажениям звука); в черно-белом телевидении при преобразовании теряется цвет изображения. Именно в связи с возможными потерями встает задача выработки таких способов представления и последующего преобразования первичного сообщения, которые обеспечивали бы возможно более полную сохранность исходной информации и, одновременно, согласование со скоростью передачи информации по данной линии связи.

При необходимости перед преобразованием или в процессе его может осуществляться кодирование первичного сообщения кодером (К). Кодирование (точнее, первичное кодирование) может осуществляться непосредственно источником информации, например, человеком при работе на передатчике с использованием азбуки Морзе. Возможно совмещение кодера с преобразователем, например, при работе человека за клавиатурой телеграфного аппарата или компьютера он вводит знаки естественного языка, а уже устройством они переводятся в коды, которые затем передаются.

Непосредственная передача осуществляется передатчиком вторичного сообщения (ПрдС). Он инициирует некоторый нестационарный процесс, обеспечивающий распространение сигналов в канале связи.

Канал связи это материальная среда, а также физический или иной процесс, посредством которого осуществляется передача сообщения, т. е. распространение сигналов в пространстве с течением времени.

Каналы связи в зависимости от характера сигналов, передаваемых по ним подразделяются на дискретные и аналоговые. Примером дискретного канала является компьютерная сеть; аналогового – телефонная линия.

В табл. 2.1 приведены примеры некоторых каналов связи.

Любой реальный канал связи подвержен внешним воздействиям, а также в нем могут происходить внутренние процессы, в результате которых искажаются передаваемые сигналы и, следовательно, связанная с ними информация. Такие воздействия называются шумами (помехами). Источники помех могут быть внешними, например, так называемые «наводки» от мощных потребителей электричества или атмосферных явлений, приводящие к появлению нарушений в радиосвязи; одновременное действие нескольких близко расположенных однотипных ис-

точников (одновременный разговор нескольких человек). К помехам могут приводить и внутренние особенности данного канала, например, физические неоднородности носителя; паразитные явления в шинах; процессы затухания сигнала в линии связи из-за большой удаленности.

Таблица 2.1

Примеры каналов связи

Канал связи	Среда	Носитель сообщения	Процесс, используемый для передачи сообщения
Почта, курьеры	Среда обитания человека	Бумага	Механическое перемещение носителя
Телефон, компьютерные сети	Проводник	Электрический ток	Перемещение электрических зарядов
Радио, телевидение	Электромагнитное поле	Электромагнитные волны	Распространение электромагнитных волн
Зрение	Электромагнитное поле	Световые волны	Распространение световых волн
Слух	Воздух	Звуковые волны	Распространение звуковых волн
Обоняние, вкус	Воздух, пища	Химические вещества	Химические реакции
Осязание	Поверхность кожи	Объект, воздействующий на органы осязания	Теплопередача, давление

Если уровень помех оказывается соизмерим с интенсивностью несущего сигнала, то передача информации по данному каналу оказывается вообще невозможной. Однако и при относительно низких уровнях шумов они могут вызывать искажения передаваемых сигналов и, следовательно, частичную потерю связанной с ними информации. Существуют и применяются методы защиты от помех, например, экранирование электрических линий связей; улучшение избирательности приемного устройства и т. д. Другим способом защиты от помех является использование специальных методов кодирования информации.

После прохождения вторичного сообщения по каналу связи оно попадает в приемное устройство (ПрмС), где одновременно преобразуется в форму, необходимую для дальнейшей интерпретации. Если перед передачей применялось кодирование, после приема вторичное сообщение направляется в декодер (ДК) и лишь затем – к получателю (потребителю) информации (ПрИ). При этом декодер может быть совмещен с преобразователем (например, телеграфный аппарат или компьютер) или

с приемником информации (радист, принимающий сигналы азбуки Морзе и интерпретирующий их).

Понятие линия связи охватывает все элементы представленной на рис. 2.1 схемы от источника до приемника информации, т. е. **ЛИНИЯ СВЯЗИ** – это совокупность средств связи и канала связи, посредством которых осуществляется передача информации от источника к приемнику.

Характеристиками любой линии связи являются скорость, с которой возможна передача сообщения в ней, а также степень искажения сообщения в процессе передачи. Из этих параметров вычленим те, что относятся непосредственно к каналу связи, т. е. характеризуют среду и процесс передачи. При этом мы затронем только вопросы передачи по дискретному каналу связи.

Дискретный канал – канал связи, используемый для передачи дискретных сообщений.

Упрощенная схема передачи информации по дискретному каналу связи представлена на рис. 2.2.



Рис. 2.2. Дискретный канал связи

Источник дискретных сообщений (ИДС) использует для представления информации первичный алфавит $\{A\}$. Первичный кодер (ПК) кодирует знаки первичного алфавита n элементарными сигналами с алфавитом $\{a\}$. Действие помех в процессе передачи может состоять в том, что алфавит принимаемых сигналов будет отличаться от алфавита входных сигналов как их числом так и характеристиками. Пусть это будет алфавит $\{b\}$, содержащий m элементарных сигналов. Несовпадение алфавитов сигналов приводит к тому, что на выходе канала появляются такие комбинации элементарных сигналов, которые не могут быть интерпретированы как коды знаков первичного алфавита. Другими словами, алфавит приемника вторичного сообщения (ПрмДС) $\{B\}$ может не совпасть с алфавитом $\{A\}$. Для простоты будем считать, что декодер вторичных сигналов совмещен с приемником.

Вводя количественные характеристики процесса передачи информации, постараемся выделить из них те, которые зависят только от свойств канала, и те, которые определяются особенностями источника дискретного сообщения.

Дискретный канал считается заданным, если известны:

- время передачи одного элементарного сигнала τ .
- исходный алфавит элементарных сигналов $\{a\}$, т. е. все его знаки a_i ($i = 1 \dots n$, где n – число знаков алфавита $\{a\}$);
- n значений вероятностей появления элементарных сигналов на входе $p(a_i)$; эти вероятности называются априорными (поскольку они определяются не свойствами канала, а источником сообщения, т. е. являются внешними по отношению к каналу и самому факту передачи сообщения);
- алфавит сигналов на выходе канала $\{b\}$, т. е. все знаки b_j ($j = 1 \dots m$, где m – число знаков алфавита $\{b\}$); в общем случае $n \neq m$;
- значения условных вероятностей $p_{a_i}(b_j)$, каждая из которых характеризует вероятность появления на выходе канала сигнала b_j при условии, что на вход был послан сигнал a_i ; поскольку эти вероятности определяются свойствами самого канала передачи, они называются апостериорными; очевидно, количество таких вероятностей равно $n \cdot m$:

$$\begin{aligned} & p_{a_1}(b_1), p_{a_1}(b_2), \dots, p_{a_1}(b_m) \\ & p_{a_2}(b_1), p_{a_2}(b_2), \dots, p_{a_2}(b_m) \\ & \dots \\ & p_{a_n}(b_1), p_{a_n}(b_2), \dots, p_{a_n}(b_m) \end{aligned}$$

Очевидно также, что для каждой строки выполняется условие нормировки:

$$\sum_{i=1}^m p_{a_i}(b_i) = 1 (i = 1 \dots n).$$

Как мы увидим в дальнейшем, все остальные характеристики дискретного канала могут быть определены через перечисленные параметры.

Дискретный канал называется **однородным**, если для любой пары i и j условная вероятность $p_{a_i}(b_j)$ с течением времени не изменяется (т. е. влияние помех все время одинаково).

Дискретный канал называется **каналом без памяти**, если $p(a_i)$ и $p_{a_i}(b_j)$ не зависят от места знака в первичном сообщении (т. е. отсутствуют корреляции знаков).

2.2. Параметры дискретных каналов связи

Будем считать, что для передачи используются колебательные или волновые процессы – с практической точки зрения такие каналы представляют наибольший интерес (в частности, к ним относятся компьютерные линии связи).

Введем ряд величин, характеризующих передачу информации по каналу.

2.2.1. Ширина полосы пропускания

Любой преобразователь, работа которого основана на использовании колебаний (электрических или механических) может формировать и пропускать сигналы из ограниченной области частот. Пример с телефонной связью приводился выше. То же следует отнести и к радио и телевизионной связи – весь частотный спектр разделен на диапазоны (ДВ, СВ, КВ1, КВ2, УКВ, ДМВ), в пределах которых каждая станция занимает свой под диапазон, чтобы не мешать вещанию других.

Интервал частот, используемый данным каналом связи для передачи сигналов, называется **шириной полосы пропускания**.

Для построения теории важна не сама ширина полосы пропускания, а максимальное значение частоты из данной полосы (ν_m), поскольку именно им определяется длительность элементарного импульса τ :

$$\tau_0 = \frac{1}{\nu_m}. \quad (2.1)$$

Другими словами, каждые τ секунд по каналу можно передавать импульс или паузу, связывая с их последовательностью определенные коды. Использовать сигналы большей длительности, чем τ , в принципе, возможно (например, 2τ) – это не приведет к потере информации, хотя снизит скорость ее передачи по каналу. Использование же сигналов более коротких, чем τ , может привести к информационным потерям, поскольку информационный параметр сигнала будет принимать какие-то промежуточные значения между заданными дискретными (например, 0 и 1), что затруднит их интерпретацию. Следовательно, по дискретному каналу за единицу времени можно передавать не более ν_m элементарных сигналов.

Если канал является аналоговым, то ν_m характеризует число полных колебаний параметра за единицу времени, с каждым из которых можно связать два элементарных сигнала; по этой причине связь τ и ν_m оказывается иной:

$$\tau = \frac{1}{2\nu_m}.$$

В дальнейшем, как уже указывалось, мы будем рассматривать лишь дискретный канал и, следовательно, использовать (2.1).

Возможны частные случаи, когда передача ведется на единственной частоте, создаваемой, например, тактовым генератором; тогда, очевидно, ν_m равна тактовой частоте.

Зная τ , можно найти количество элементарных сигналов, передаваемое по каналу за единицу времени:

$$L = \frac{1}{\tau}$$

(очевидно, если известна v_m , то $L = v_m$). Если код знака первичного алфавита состоит из k_i элементарных сигналов, время его передачи по каналу составит $t_i = k_i \cdot \tau$, а среднее время передачи кодовой комбинации одного знака первичного алфавита будет равно $t = K(A, a) \cdot \tau$.

2.2.2. Пропускная способность канала связи

С передачей одного элементарного сигнала связано некоторое количество информации I_s . Если общее число различных элементарных сигналов n , а вероятности их появления $p(a_i)$ ($i = 1 \dots n$), то согласно формуле Шеннона

$$I_s = - \sum_{i=1}^n p(a_i) \cdot \log_2 p(a_i)$$

Однако, как обсуждалось выше, оптимальным будет такой вариант кодирования, при котором появление всех элементарных сигналов (знаков вторичного алфавита) оказывается равновероятным. В таком случае

$$I_s = I_{smax} = \log_2 n$$

Это значение является предельным (наибольшим) для информационного содержания элементарного сигнала выбранного вторичного алфавита. Поскольку такое количество информации передается за время τ , можно ввести величину, характеризующую предельную интенсивность информационного потока через канал – пропускную способность канала C :

$$C = \frac{I_s^{max}}{\tau} = L \cdot I_s^{max}. \quad (2.2)$$

Данная величина является характеристикой канала связи, поскольку зависит только от его особенностей. Это выражение служит определением пропускной способности как идеального канала (без помех), так и реального канала с помехами – просто, как мы увидим далее, информационное содержание элементарного сигнала в реальном канале оказывается меньше $\log_2 n$.

Если I_{smax} выражено в битах, а τ – в секундах, то единицей измерения C будет бит/с. Раньше такая единица называлась бод, однако, название не прижилось, и по этой причине пропускная способность канала связи измеряется в бит/с. Производными единицами являются:

$$1 \text{ Кбит/с} = 10^3 \text{ бит/с}, 1 \text{ Мбит/с} = 10^6 \text{ бит/с}, 1 \text{ Гбит/с} = 10^9 \text{ бит/с}.$$

При отсутствии в канале связи помех $I_{smax} = \log_2 n$; тогда

$$C_0 = L \cdot \log_2 n = \frac{\log_2 n}{\tau} = v_m \cdot \log_2 n \quad (2.3)$$

– максимально возможное значение пропускной способности (это обстоятельство отражено индексом «0»); в реальном канале $I_{\text{smax}} \leq \log_2 n$ и, следовательно, $C \leq C_0$.

2.2.3. Скорость передачи информации

Если источник выдает L элементарных сигналов в единицу времени, а средняя длина кода одного знака составляет $K(A, a)$, то, очевидно, отношение $L/K(A, a)$ будет выражать число знаков первичного алфавита, выдаваемых источником за единицу времени. Если с каждым из них связано среднее количество информации $I(A)$, то можно найти общее количество информации, передаваемой источником за единицу времени. Эта величина называется скоростью передачи или энтропией источника (будем обозначать ее J).

$$J = \frac{L}{K(A, a)} \cdot I(A) = \frac{I(A)}{\tau \cdot K(A, a)}. \quad (2.4)$$

Энтропия источника, в отличие от пропускной способности, является характеристикой источника, а не канала связи.

Размерностью J , как и C , является бит/с. Каково соотношение этих характеристик? Рассмотрим канал без помех. Тогда выразив L из (2.3) и подставив в (2.4), получим:

$$J = \frac{I(A) \cdot C_0}{K(A, a) \cdot \log_2 n}.$$

Согласно первой теореме Шеннона при любом способе кодирования

$$K(A, a) \geq \frac{I(A)}{\log_2 n},$$

хотя может быть сколь угодно близкой к этому значению. Следовательно, всегда $J \leq C_0$, т. е. скорость передачи информации по каналу связи не может превысить его пропускной способности.

Как показано в теории Шеннона, данное утверждение справедливо как при отсутствии в канале помех (шумов) (идеальный канал связи), так и при их наличии (реальный канал связи).

Пример 2.1.

Первичный алфавит состоит из трех знаков с вероятностями $p_1 = 0,2$; $p_2 = 0,7$; $p_3 = 0,1$. Для передачи по каналу без помех используются равномерный двоичный код. Частота тактового генератора 500 Гц. Какова пропускная способность канала и скорость передачи?

Поскольку код двоичный, $n = 2$; из (2.3) $C_0 = 500$ бит/с. Число знаков первичного алфавита $N = 3$. Из

$$I(A) = I_1 = -0,2 \cdot \log_2 0,2 - 0,7 \cdot \log_2 0,7 - 0,1 \cdot \log_2 0,1 = 1,16 \text{ бит}$$

$K(A, 2) \geq \log_2 N = 2$. Следовательно, из (2.4) получаем:

$$J = \frac{I_1}{K(A, 2) \cdot \tau} = \frac{v \cdot I_1}{K(A, 2)} = \frac{500 \cdot 1,16}{2} = 290 \text{ бит/с.}$$

ГЛАВА 3. КАБЕЛЬНЫЕ ЛИНИИ СВЯЗИ

3.1. Типы кабельных линий связи

Средой передачи информации называются те линии связи (или каналы связи), по которым производится обмен информацией между компьютерами. В подавляющем большинстве компьютерных сетей (особенно локальных) используются проводные или кабельные каналы связи, хотя в настоящее время все большее распространение находят и беспроводные сети. Информация в локальных сетях чаще всего передается в последовательном коде, то есть бит за битом. Понятно, что такая передача медленнее и сложнее, чем при использовании параллельного кода. Однако надо учитывать то, что при более быстрой параллельной передаче увеличивается количество соединительных кабелей в число раз, равное количеству разрядов параллельного кода (например, в 8 раз при 8-разрядном коде). Это совсем не мелочь, как может показаться на первый взгляд. При значительных расстояниях между абонентами сети стоимость кабеля может быть вполне сравнима со стоимостью компьютеров и даже превосходить ее. К тому же проложить один кабель (реже два разнонаправленных) гораздо проще, чем 8, 16 или 32. Значительно дешевле обойдется также поиск повреждений и ремонт кабеля.

Передача на большие расстояния при любом типе кабеля требует сложной передающей и приемной аппаратуры: для этого надо формировать мощный сигнал на передающем конце и детектировать слабый сигнал на приемном конце. При последовательной передаче для этого требуется всего один передатчик и один приемник. При параллельной же передаче количество передатчиков и приемников возрастает пропорционально разрядности используемого параллельного кода. Поэтому даже при разработке сети незначительной длины (порядка десятка метров) выбирают последовательную передачу.

К тому же при параллельной передаче чрезвычайно важно, чтобы длины отдельных кабелей были точно равны друг другу, иначе в результате прохождения по кабелям разной длины между сигналами на приемном конце образуется временной сдвиг, который может привести к сбоям в работе или даже к полной неработоспособности сети. Например, при скорости передачи 100 Мбит/с и длительности бита 10 нс этот временной сдвиг не должен превышать 5...10 нс. Такую величину сдвига дает разница в длинах кабелей в 1...2 метра. При длине кабеля 1000 метров это составляет 0,1...0,2 %.

Правда, в некоторых высокоскоростных локальных сетях все-таки используют параллельную передачу по 2...4 кабелям, что позволяет при

заданной скорости передачи применять более дешевые кабели с меньшей полосой пропускания, но допустимая длина кабелей при этом не превышает сотни метров. Примером может служить сегмент 100BASE-T4 сети Fast Ethernet.

Промышленностью выпускается огромное количество типов кабелей. Все выпускаемые кабели можно разделить на три большие группы:

- кабели на основе витых пар проводов (twisted pair), которые делятся на экранированные (shielded twisted pair, STP) и неэкранированные (unshielded twisted pair, UTP);
- коаксиальные кабели (coaxial cable);
- оптоволоконные кабели (fiber optic).

Каждый тип кабеля имеет свои преимущества и недостатки, так что при выборе типа кабеля надо учитывать как особенности решаемой задачи, так и особенности конкретной сети, в том числе и используемую топологию. В настоящее время действует стандарт на кабели EIA/TIA 568 (Commercial Building Telecommunications Cabling Standard), принятый в 1995 году и заменивший все действовавшие ранее фирменные стандарты.

Витые пары проводов используются в самых дешевых и на сегодняшний день, пожалуй, самых популярных кабелях. Кабель на основе витых пар представляет собой несколько пар скрученных изолированных медных проводов в единой диэлектрической (пластиковой) оболочке. Он довольно гибкий и удобный для прокладки.

Обычно в кабель входит две витые пары (рис. 3.1) или четыре витые пары.

Неэкранированные витые пары характеризуются слабой защищенностью от внешних электромагнитных помех, а также слабой защищенностью от подслушивания с целью, например, промышленного шпионажа. Перехват передаваемой информации возможен как с помощью контактного метода (посредством двух иголок, воткнутых в кабель), так и с помощью бесконтактного метода, сводящегося к радиоперехвату излучаемых кабелем электромагнитных полей. Для устранения этих недостатков применяется экранирование.

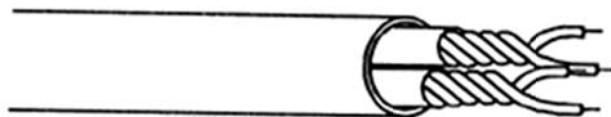


Рис. 3.1. Кабель с витыми парами

В случае экранированной витой пары STP каждая из витых пар помещается в металлическую оплетку-экран для уменьшения излучений

кабеля, защиты от внешних электромагнитных помех и снижения взаимного влияния пар проводов друг на друга (crosstalk – перекрестные наводки). Естественно, экранированная витая пара гораздо дороже, чем неэкранированная, а при ее использовании необходимо применять и специальные экранированные разъемы, поэтому встречается она значительно реже, чем неэкранированная витая пара.

Основные достоинства неэкранированных витых пар – простота монтажа разъемов на концах кабеля, а также простота ремонта любых повреждений по сравнению с другими типами кабеля. Все остальные характеристики у них хуже, чем у других кабелей. Например, при заданной скорости передачи затухание сигнала (уменьшение его уровня по мере прохождения по кабелю) у них больше, чем у коаксиальных кабелей. Если учесть еще низкую помехозащищенность, то становится понятным, почему линии связи на основе витых пар, как правило, довольно короткие (обычно в пределах 100 метров). В настоящее время витая пара используется для передачи информации на скоростях до 100 Мбит/с и ведутся работы по повышению скорости передачи до 1000 Мбит/с.

Коаксиальный кабель представляет собой электрический кабель, состоящий из центрального провода и металлической оплетки, разделенных между собой слоем диэлектрика (внутренней изоляции) и помещенных в общую внешнюю оболочку (рис. 3.2).



Рис. 3.2. Коаксиальный кабель

Коаксиальный кабель до недавнего времени был распространен наиболее широко, что связано с его высокой помехозащищенностью (благодаря металлической оплетке), а также более высокими, чем в случае витой пары, допустимыми скоростями передачи данных (до 500 Мбит/с) и большими допустимыми расстояниями передачи (до километра и выше). К нему труднее механически подключиться для несанкционированного прослушивания сети, он также дает заметно меньше электромагнитных излучений вовне. Однако монтаж и ремонт коаксиального кабеля существенно сложнее, чем витой пары, а стоимость его выше (он дороже примерно в 1,5–3 раза по сравнению с кабелем на основе

витых пар). Сложнее и установка разъемов на концах кабеля. Поэтому его сейчас применяют реже, чем витую пару.

Основное применение коаксиальный кабель находит в сетях с топологией типа «шина». При этом на концах кабеля обязательно должны устанавливаться терминаторы для предотвращения внутренних отражений сигнала, причем один (и только один!) из терминаторов должен быть заземлен. Без заземления металлическая оплетка не защищает сеть от внешних электромагнитных помех и не снижает излучение передаваемой по сети информации во внешнюю среду. Но при заземлении оплетки в двух или более точках из строя может выйти не только сетевое оборудование, но и компьютеры, подключенные к сети. Терминаторы должны быть обязательно согласованы с кабелем, то есть их сопротивление должно быть равно волновому сопротивлению кабеля. Например, если используется 50-омный кабель, для него подходят только 50-омные терминаторы.

Реже коаксиальные кабели применяются в сетях с топологией «звезда» и «пассивная звезда» (например, в сети Arcnet). В этом случае проблема согласования существенно упрощается, так как внешних терминаторов на свободных концах не требуется.

Волновое сопротивление кабеля указывается в сопроводительной документации. Чаще всего в локальных сетях применяются 50-омные (например, RG-58, RG-11) и 93-омные кабели (например, RG-62). 75-омные кабели, распространенные в телевизионной технике, в локальных сетях не используются. Вообще, марок коаксиального кабеля значительно меньше, чем кабелей на основе витых пар. Он не считается перспективным. Не случайно в сети Fast Ethernet не предусмотрено применение коаксиальных кабелей. Однако во многих случаях классическая шинная топология очень удобна – она не требует применения дополнительных устройств-концентраторов.

Существует два основных типа коаксиального кабеля:

- тонкий (thin) кабель, имеющий диаметр около 0,5 см, более гибкий;
- толстый (thick) кабель, имеющий диаметр около 1 см, значительно более жесткий. Он представляет собой классический вариант коаксиального кабеля, который уже почти полностью вытеснен более современным тонким кабелем.

Тонкий кабель используется для передачи на меньшие расстояния, чем толстый, так как в нем сигнал затухает сильнее. Зато с тонким кабелем гораздо удобнее работать: его можно оперативно проложить к каждому компьютеру, а толстый требует жесткой фиксации на стене помещения. Подключение к тонкому кабелю (с помощью разъемов BNC типа) проще и не требует дополнительного оборудования, а для под-

ключения к толстому кабелю надо использовать специальные довольно дорогие устройства, прокалывающие его оболочки и устанавливающие контакт как с центральной жилой, так и с экраном. Толстый кабель примерно вдвое дороже, чем тонкий. Поэтому тонкий кабель применяется гораздо чаще.

Типичные величины задержки распространения сигнала в коаксиальном кабеле составляют для тонкого кабеля около 5 нс/м, а для толстого – около 4,5 нс/м.

Существуют варианты коаксиального кабеля с двойным экраном (один экран расположен внутри другого и отделен от него дополнительным слоем изоляции). Такие кабели имеют лучшую помехозащищенность и защиту от прослушивания, но они немного дороже обычных.

В настоящее время считается, что коаксиальный кабель устарел, в большинстве случаев его вполне может заменить витая пара или оптоволоконный кабель. Новые стандарты на кабельные системы уже не включают его в перечень типов кабелей

Оптоволоконный (он же волоконно-оптический) кабель – это принципиально иной тип кабеля по сравнению с рассмотренными двумя типами электрического или медного кабеля. Информация по нему передается не электрическим сигналом, а световым. Главный его элемент – это прозрачное стекловолокно, по которому свет проходит на огромные расстояния (до десятков километров) с незначительным ослаблением.

Структура оптоволоконного кабеля очень проста и похожа на структуру коаксиального электрического кабеля (рис. 3.3), только вместо центрального медного провода здесь используется тонкое (диаметром порядка 1...10 мкм) стекловолокно, а вместо внутренней изоляции применяется стеклянная или пластиковая оболочка, не позволяющая свету выходить за пределы стекловолокна. В данном случае мы имеем дело с режимом полного внутреннего отражения света от границы двух веществ с разными коэффициентами преломления (у стеклянной оболочки коэффициент преломления значительно ниже, чем у центрального волокна). Металлическая оплетка кабеля обычно отсутствует, так как экранирование от внешних электромагнитных помех здесь не требуется, однако иногда ее все-таки применяют для механической защиты от окружающей среды (такой кабель иногда называют броневым, он может объединять под одной оболочкой несколько оптоволоконных кабелей).

Оптоволоконный кабель обладает исключительными характеристиками по помехозащищенности и секретности передаваемой информации. Никакие внешние электромагнитные помехи в принципе не способны исказить световой сигнал, а сам этот сигнал принципиально не порождает внешних электромагнитных излучений. Подключиться

к этому типу кабеля для несанкционированного прослушивания сети практически невозможно, так как это требует нарушения целостности кабеля. Теоретически возможная полоса пропускания такого кабеля достигает величины 10¹² Гц, что несравнимо выше, чем у любых электрических кабелей. Стоимость оптоволоконного кабеля постоянно снижается и сейчас примерно равна стоимости тонкого коаксиального кабеля. Однако в данном случае необходимо применение специальных оптических приемников и передатчиков, преобразующих световые сигналы в электрические и обратно, что порой существенно увеличивает стоимость сети в целом.

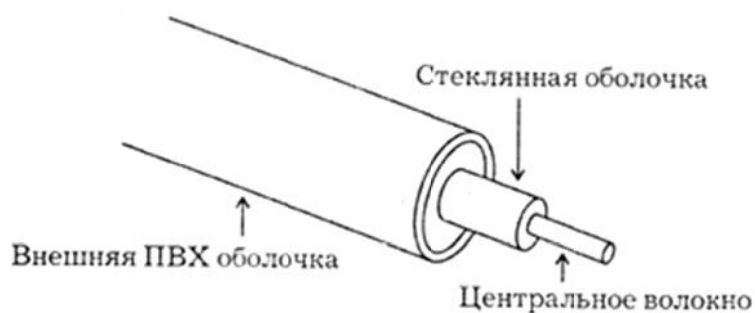


Рис. 3.3. Структура оптоволоконного кабеля

Типичная величина затухания сигнала в оптоволоконных кабелях на частотах, используемых в локальных сетях, составляет около 5 дБ/км, что примерно соответствует показателям электрических кабелей на низких частотах. Но в случае оптоволоконного кабеля при росте частоты передаваемого сигнала затухание увеличивается очень незначительно, и на больших частотах (особенно свыше 200 МГц) его преимущества перед электрическим кабелем неоспоримы, он просто не имеет конкурентов.

Однако оптоволоконный кабель имеет и некоторые недостатки.

Самый главный из них – высокая сложность монтажа (при установке разъемов необходима микронная точность, от точности скола стекловолокна и степени его полировки сильно зависит затухание в разьеме). Для установки разъемов применяют сварку или склеивание с помощью специального геля, имеющего такой же коэффициент преломления света, что и стекловолокно. В любом случае для этого нужна высокая квалификация персонала и специальные инструменты. Поэтому чаще всего оптоволоконный кабель продается в виде заранее нарезанных кусков разной длины, на обоих концах которых уже установлены разъемы нужного типа.

Хотя оптоволоконные кабели и допускают разветвление сигналов (для этого выпускаются специальные разветвители на 2...8 каналов), как правило, их используют для передачи данных только в одном направлении, между одним передатчиком и одним приемником. Ведь любое раз-

ветвление неизбежно сильно ослабляет световой сигнал, и если разветвлений будет много, то свет может просто не дойти до конца сети.

Оптоволоконный кабель менее прочен, чем электрический, и менее гибкий (типичная величина допустимого радиуса изгиба составляет около 10...20 см). Чувствителен он и к ионизирующим излучениям, из-за которых снижается прозрачность стекловолокна, то есть увеличивается затухание сигнала. Чувствителен он также к резким перепадам температуры, в результате которых стекловолокно может треснуть. В настоящее время выпускаются оптические кабели из радиационно стойкого стекла (стоят они, естественно, дороже).

Оптоволоконные кабели чувствительны также к механическим воздействиям (удары, ультразвук) – так называемый микрофонный эффект. Для его уменьшения используют мягкие звукопоглощающие оболочки.

Применяют оптоволоконный кабель только в сетях с топологией «звезда» и «кольцо». Никаких проблем согласования и заземления в данном случае не существует. Кабель обеспечивает идеальную гальваническую развязку компьютеров сети. В будущем этот тип кабеля, вероятно, вытеснит электрические кабели всех типов или, во всяком случае, сильно потеснит их. Запасы меди на планете истощаются, а сырья для производства стекла более чем достаточно.

3.2. Параметры кабельных линий связи

Кабельные линии связи относятся к электрическим цепям, в которых электромагнитное поле и потери распределены вдоль всей цепи. В результате напряжения и токи на различных участках даже неразветвленной цепи отличаются друг от друга, т. е. являются функциями двух независимых переменных: времени t и пространственной координаты x . Такие цепи называются цепями с распределенными параметрами. Смысл данного названия заключается в том, что у цепей данного класса каждый бесконечно малый элемент их длины характеризуется сопротивлением, индуктивностью, а между проводниками – соответственно емкостью и проводимостью (рис. 3.4).

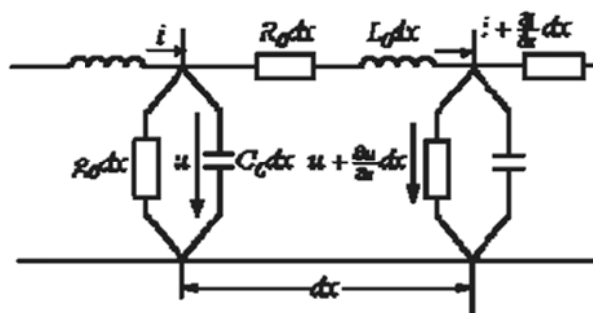


Рис. 3.4. Линия с распределенными параметрами

Для оценки, к какому типу отнести цепь: с сосредоточенными или распределенными параметрами – следует сравнить ее длину l с длиной электромагнитной волны $\lambda = v \cdot \tau$. Если $l > 0,05 \dots 0,1\lambda$, т. е. при $\tau = 0,02$ с и $v = 3 \cdot 10^5$ км/с $l > 300 \dots 600$ км. При частоте сигнала 100 МГц длина волны $\lambda = 3$ м, то есть уже при $l > 0,15 \dots 0,3$ м к линии следует подходить как к цепи с распределенными параметрами.

Для исследования процессов в цепи с распределенными параметрами (другое название длинная линия) введем дополнительное условие о равномерности распределения вдоль линии ее параметров: индуктивности, сопротивления, емкости и проводимости. Такую линию называют однородной.

Под первичными параметрами линии будем понимать сопротивление R_0 , индуктивность L_0 , проводимость g_0 и емкость C_0 , отнесенные к единице ее длины. Для получения уравнений однородной линии разобьем ее на отдельные участки бесконечно малой длины dx со структурой, показанной на рис. 3.4.

Пусть напряжение и ток в начале такого элементарного четырехполюсника равны u и i , а в конце соответственно

$$u + \frac{\partial u}{\partial x} dx \text{ и } i + \frac{\partial i}{\partial x} dx .$$

Разность напряжений в начале и конце участка определяется падением напряжения на резистивном и индуктивном элементах, а изменение тока на участке равно сумме токов утечки и смещения через проводимость и емкость. Таким образом, по законам Кирхгофа

$$\begin{aligned} -\frac{\partial u}{\partial x} dx &= R_0 dx \cdot i + L_0 dx \cdot \frac{\partial i}{\partial t} \\ -\frac{\partial i}{\partial x} dx &= g_0 dx \cdot u + C_0 dx \cdot \frac{\partial u}{\partial t} \end{aligned}$$

После сокращения на dx

$$\begin{aligned} -\frac{\partial u}{\partial x} &= R_0 \cdot i + L_0 \cdot \frac{\partial i}{\partial t} \\ -\frac{\partial i}{\partial x} &= g_0 \cdot u + C_0 \cdot \frac{\partial u}{\partial t} \end{aligned} \tag{3.1}$$

Теорию цепей с распределенными параметрами в установившихся режимах будем рассматривать для случая синусоидального тока. Далее будет показано, что любой сигнал может быть представлен как суперпозиция синусоидальных сигналов (ряд или интеграл Фурье). Тогда

полученные соотношения при $f=0$ можно распространить и на цепи постоянного тока, а воспользовавшись разложением в ряд Фурье – на линии периодического несинусоидального тока.

Вводя комплексные величины и заменяя $\frac{\partial}{\partial t}$ на $j\omega$ на основании (3.1) получаем

$$\begin{aligned} -\frac{d\dot{U}}{dx} &= (R_0 + j\omega L_0)\dot{I} = Z_0\dot{I}; \\ -\frac{d\dot{I}}{dx} &= (g_0 + j\omega C_0)\dot{U} = Y_0\dot{U}, \end{aligned} \quad (3.2)$$

где Z_0 и Y_0 – соответственно комплексные сопротивление и проводимость на единицу длины линии.

Продифференцировав первое уравнение в (3.2) по x и подставив выражение $d\dot{I}/dx$ из второго уравнения, получим

$$\frac{d^2\dot{U}}{dx^2} = Z_0 Y_0 \dot{U}.$$

Характеристическое уравнение

$$p^2 - Z_0 Y_0 = 0,$$

откуда $p_{1,2} = \sqrt{Z_0 Y_0} = \pm\gamma = \pm(\alpha + j\beta)$.

Таким образом,

$$\dot{U} = \dot{A}_1 e^{-\gamma x} + \dot{A}_2 e^{\gamma x} = \dot{A}_1 e^{-\alpha x} e^{-j\beta x} + \dot{A}_2 e^{\alpha x} e^{j\beta x}, \quad (3.3)$$

где $\gamma = \alpha + j\beta$ – постоянная распространения; α – коэффициент затухания; β – коэффициент фазы.

Для тока согласно первому уравнению в (3.2) можно записать

$$\dot{I} = -\frac{1}{Z_0} \frac{d\dot{U}}{dx} = \frac{\gamma}{Z_0} (A_1 e^{-\gamma x} - A_2 e^{\gamma x}) = \frac{1}{Z_C} (\dot{A}_1 e^{-\gamma x} - \dot{A}_2 e^{\gamma x}), \quad (3.4)$$

где $Z_C = \sqrt{Z_0/Y_0}$ – волновое сопротивление.

Волновое сопротивление Z_C и постоянную распространения γ называют вторичными параметрами линии, которые характеризуют ее свойства как устройства для передачи энергии или информации.

Определяя $\dot{A}_1 = A_1 e^{j\varphi_1}$ и $\dot{A}_2 = A_2 e^{j\varphi_2}$, на основании (3.3) запишем

$$u(x,t) = \sqrt{2} A_1 e^{-\alpha x} \sin(\omega t - \beta x + \varphi_1) + \sqrt{2} A_2 e^{\alpha x} \sin(\omega t + \beta x + \varphi_2). \quad (3.5)$$

Аналогичное уравнение согласно можно записать для тока.

Слагаемые в правой части соотношения (3.5) можно трактовать как бегущие волны: первая движется и затухает в направлении возрастания x , вторая – убывания. Действительно, в фиксированный момент времени каждое из слагаемых представляет собой затухающую (вследствие потерь энергии) гармоническую функцию координаты x , а в фиксированной точке – синусоидальную функцию времени.

Волну, движущую от начала линии в сторону возрастания x , называют прямой, а движущуюся от конца линии в направлении убывания x – обратной.

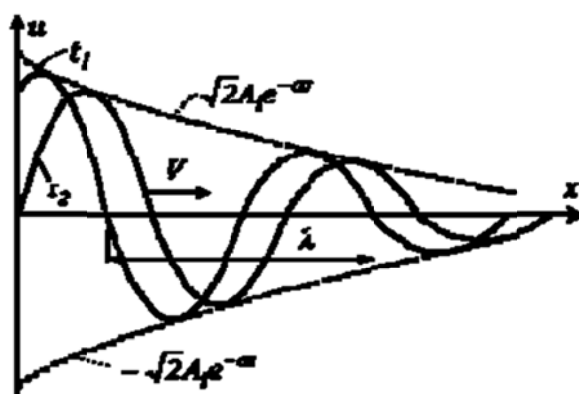


Рис. 3.5. Прямая волна напряжения

На рис. 3.5 представлена затухающая синусоида прямой волны для моментов времени t_1 и t_2 ($t_2 > t_1$). Перемещение волны характеризуется фазовой скоростью. Это скорость перемещения по линии неизменного фазового состояния, т. е. скорость, с которой нужно перемещаться вдоль линии, чтобы наблюдать одну и ту же фазу волны:

$$(\omega t + \beta x + \varphi_2) = const.$$

Продифференцировав это выражение по времени, получим

$$v = \frac{\omega}{\beta}. \quad (3.6)$$

Длиной волны λ называется расстояние между двумя ее ближайшими точками, различающимися по фазе на 2π рад. В соответствии с данным определением $\omega t - \beta(x + \lambda) + \varphi_1 = \omega t - \beta x - 2\pi + \varphi_1$, откуда

$$\lambda = \frac{2\pi}{\beta}$$

и с учетом (3.6)

$$\lambda = vt = \frac{v}{f}.$$

В соответствии с введенными понятиями прямой и обратной волн распределение напряжения вдоль линии в любой момент времени можно трактовать как результат наложения двух волн: прямой и обратной, перемещающихся вдоль линии с одинаковой фазовой скоростью, но в противоположных направлениях:

$$\dot{U} = \dot{U}_{np} + \dot{U}_{обр}.$$

где $\dot{U}_{np} = \dot{A}_1 e^{-\gamma x}$ и $\dot{U}_{обр} = \dot{A}_2 e^{\gamma x}$.

Представление напряжения в виде суммы прямой и обратной волн означает, что положительные направления напряжения для обеих волн выбраны одинаково: от верхнего провода к нижнему.

Аналогично для тока на основании (6) можно записать

$$\dot{I} = \dot{I}_{np} - \dot{I}_{обр},$$

где $\dot{I}_{np} = \dot{A}_1 e^{-\gamma x} / Z_C$ и $\dot{I}_{обр} = \dot{A}_2 e^{\gamma x} / Z_C$.

Положительные направления прямой и обратной волн тока различны: положительное направление прямой волны совпадает с положительным направлением тока (от начала к концу линии), а положительное направление обратной волны ему противоположно.

Для прямых и обратных волн напряжения и тока выполняется закон Ома

$$\dot{I}_{np} = \frac{\dot{U}_{np}}{Z_C} \text{ и } \dot{I}_{обр} = \frac{\dot{U}_{обр}}{Z_C}.$$

Рассмотрим теоретически важный случай бесконечно длинной однородной линии.

В случае бесконечно длинной линии в выражениях (3.3) и (3.4) для напряжения и тока слагаемые, содержащие $e^{\gamma x}$, должны отсутствовать, т. к. стремление $x \rightarrow \infty$ лишает эти составляющие физического смысла. Следовательно, в рассматриваемом случае $\dot{A}_2 = 0$. Таким образом, в решении уравнений линии бесконечной длины отсутствуют обратные волны тока и напряжения. В соответствии с вышесказанным

$$\begin{aligned} \dot{U} &= \dot{A}_1 e^{-\gamma x}, \\ \dot{I} &= \frac{\dot{A}_1 e^{-\gamma x}}{Z_C}. \end{aligned} \quad (3.7)$$

На основании соотношений (3.7) можно сделать важный вывод, что для бесконечно длинной линии в любой ее точке, в том числе и на вхо-

де, отношение комплексов напряжения и тока есть постоянная величина, равная волновому сопротивлению:

$$Z_c = \frac{\dot{U}}{\dot{I}} = Z_{ex}.$$

Таким образом, если такую линию мысленно рассечь в любом месте и вместо откинутой бесконечно длинной части подключить сопротивление, численно равное волновому сопротивлению, то режим работы оставшегося участка конечной длины не изменится. Отсюда можно сделать два вывода:

1. Уравнения бесконечно длинной линии распространяются на линию конечной длины, нагруженную на сопротивление, равное волновому сопротивлению. В этом случае также имеют место только прямые волны напряжения и тока.
2. У линии, нагруженной на волновое сопротивление, входное сопротивление также равно волновому сопротивлению.

Режим работы длинной линии, нагруженной на сопротивление, равное волновому, называется согласованным, а сама линия называется линией с согласованной нагрузкой.

Отметим, что данный режим практически важен для передачи информации, поскольку характеризуется отсутствием отраженных (обратных) волн, обуславливающих помехи.

Согласованная нагрузка полностью поглощает мощность волны, достигшей конца линии. Эта мощность называется натуральной. Поскольку в любом сечении согласованной линии сопротивление равно волновому сопротивлению, угол сдвига φ между напряжением и током неизменен. Таким образом, если мощность, получаемая линией от генератора, равна $P_1 = U_1 I_1 \cos \varphi$, то мощность в конце линии длиной l равна

$$P_2 = U_2 I_2 \cos \varphi = U_1 e^{-\alpha l} I_1 e^{-\alpha l} \cos \varphi = P_1 e^{-2\alpha l},$$

откуда КПД линии

$$\eta = \frac{P_2}{P_1} = e^{-2\alpha l}$$

и затухание

$$\alpha l = \frac{1}{2} \ln \frac{P_1}{P_2}.$$

Формула для затухания справедлива также и для волоконно-оптических линий связи.

ГЛАВА 4. ПЕРЕДАЧА ДАННЫХ ПО ВИТЫМ ПАРАМ

4.1. Способы передачи по витым парам

Согласно стандарту EIA/TIA 568, существуют пять категорий кабелей на основе неэкранированной витой пары (UTP):

- Кабель категории 1 – это обычный телефонный кабель (пары проводов не витые), по которому можно передавать только речь, но не данные. Данный тип кабеля имеет большой разброс параметров (волнового сопротивления, полосы пропускания, перекрестных наводок).
- Кабель категории 2 – это кабель из витых пар для передачи данных в полосе частот до 1 МГц. Кабель не тестируется на уровень перекрестных наводок. В настоящее время он используется очень редко. Стандарт EIA/TIA 568 не различает кабели категорий 1 и 2.
- Кабель категории 3 – это кабель для передачи данных в полосе частот до 16 МГц, состоящий из витых пар с девятью витками проводов на метр длины. Кабель тестируется на все параметры и имеет волновое сопротивление 100 Ом. Это самый простой тип кабелей, рекомендованный стандартом для локальных сетей. Сейчас он имеет наибольшее распространение.
- Кабель категории 4 – это кабель, передающий данные в полосе частот до 20 МГц. Используется редко, так как не слишком заметно отличается от категории 3. Стандартом рекомендуется вместо кабеля категории 3 переходить сразу на кабель категории 5. Кабель категории 4 тестируется на все параметры и имеет волновое сопротивление 100 Ом. Кабель был разработан для работы в сетях по стандарту IEEE 802.5.
- Кабель категории 5 – самый совершенный кабель в настоящее время, рассчитанный на передачу данных в полосе частот до 100 МГц. Состоит из витых пар, имеющих не менее 27 витков на метр длины (8 витков на фут). Кабель тестируется на все параметры и имеет волновое сопротивление 100 Ом. Рекомендуется применять его в современных высокоскоростных сетях типа Fast Ethernet и TPFDI. Кабель категории 5 примерно на 30...50 % дороже, чем кабель категории 3.
- Кабель категории 6 – перспективный тип кабеля для передачи данных в полосе частот до 200 МГц.
- Кабель категории 7 – перспективный тип кабеля для передачи данных в полосе частот до 600 МГц.

Согласно стандарту EIA/TIA 568, полное волновое сопротивление наиболее совершенных кабелей категорий 3, 4 и 5 должно составлять $100 \text{ Ом} \pm 15 \%$ в частотном диапазоне от частоты 1 МГц до максимальной частоты кабеля. Как видим, требования не очень жесткие: величина волнового сопротивления может находиться в диапазоне от 85 до 115 Ом. Здесь же отметим, что волновое сопротивление экранированной витой пары STP должно быть по стандарту равно $150 \text{ Ом} \pm 15 \%$. Для согласования импедансов кабеля и оборудования в случае их несовпадения применяют согласующие трансформаторы. Встречается также экранированная витая пара с волновым сопротивлением 100 Ом, но довольно редко.

Существует два способа передачи сигналов по витым парам: несбалансированная передача (несимметричные цепи) и балансная передача (симметричные цепи).

Кабели на основе медных витых пар, применяемые в структурированных кабельных системах (СКС), позволяют строить как симметричные, так и несимметричные цепи.

Все виды ЛВС используют балансную передачу сигналов по витым парам. Несимметричные цепи применяются для построения систем пожарных и охранных сигнализаций и для передачи постоянных питающих напряжений, например, от УАТС к телефонам на рабочих местах.

При несбалансированной передаче используется несимметричная цепь, то есть один из проводников заземляется с одной или с двух сторон. Сигналы передаются по остальным проводникам и изменяются относительно земли. По своей природе несимметричные цепи очень чувствительны к внешнему электромагнитному излучению (ЭМИ). На рис. 4.1 видно, что на входе приемника на сигнальном проводнике присутствует сумма напряжений сигнала U_C и наводок U_H от внешнего ЭМИ.

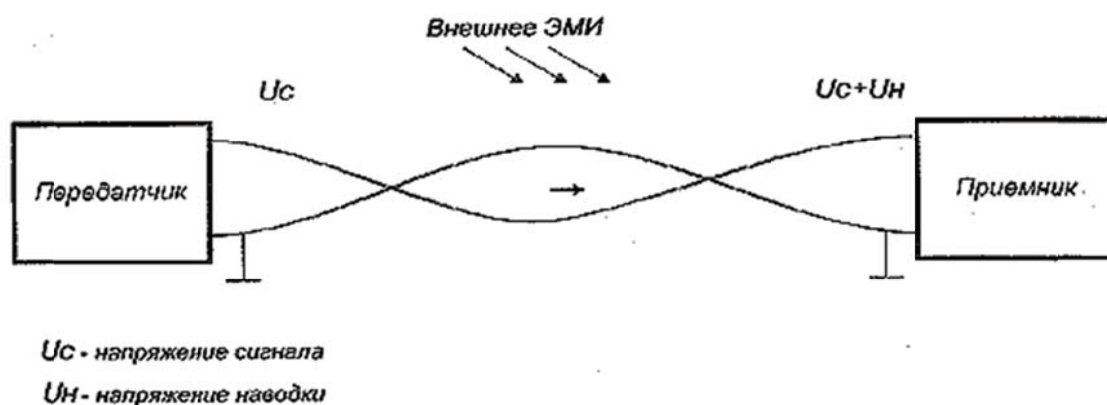


Рис. 4.1. Несимметричная цепь

Токи наводок на заземляющем проводнике стекают на землю, поэтому на нем U_H равно нулю. С другой стороны, сигнальный провод яв-

ляется источником излучения электромагнитной энергии во внешнее пространство. Это приводит к значительному затуханию сигнала в процессе его распространения.

Некоторое улучшение характеристик несимметричных цепей достигается в случае использования общего заземленного экрана, однако такое решение существенно повышает стоимость и трудоемкость монтажа кабельной системы.

Еще одним недостатком несимметричных цепей является отсутствие гальванической развязки передатчика и приемника. При неисправностях в системе заземления или в защитной изоляции сетевого оборудования высокое напряжение от систем электропитания может попасть как на заземляющий, так и на сигнальный провод, что часто сопровождается выходом из строя приемопередающих устройств на одном или обоих концах линии связи.

Достоинством несимметричных цепей является то, что для передачи N сигналов требуется только $N+1$ проводников (N сигнальных плюс один общий заземляющий).

Несимметричные цепи применяют для передачи низкочастотных сигналов на короткие расстояния. Широко известным ее примером является интерфейс RS-232.

Схема симметричной цепи, в которой используется балансный принцип передачи информации, изображена на рис. 4.2.

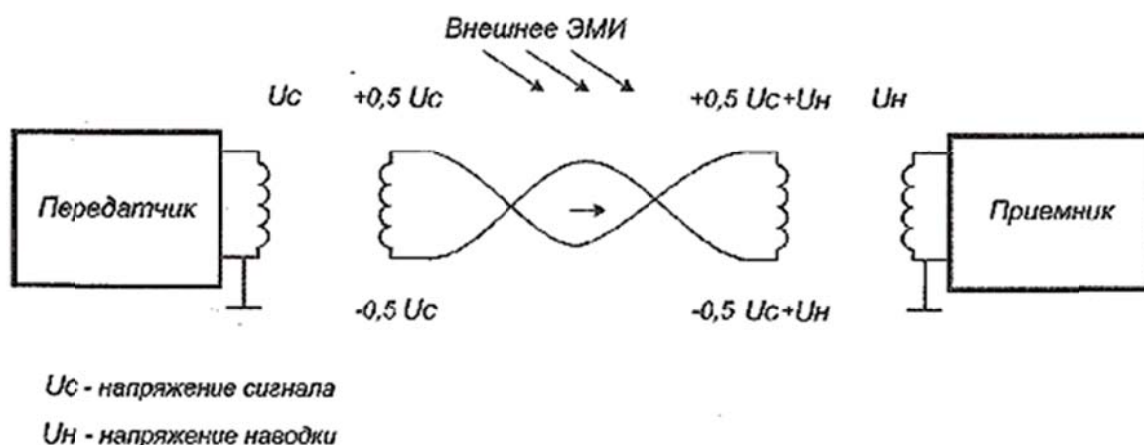


Рис. 4.2. Симметричная цепь

В симметричных цепях приемник и передатчик гальванически развязаны друг от друга согласующими трансформаторами. Во вторичные обмотки передается только разность потенциалов на первичной обмотке. Из рисунка выше видно, что токи наводки в полностью симметричной цепи приводят к противофазному изменению напряжения U_n

на первичной обмотке трансформатора приемника, так что результирующий мешающий сигнал не передается во вторичную обмотку.

Поэтому, в отличие от несимметричных, симметричные цепи значительно более устойчивы к внешним мешающим влияниям.

Соответственно ЭМИ проводников имеет равные по величине и противоположные по направлениям векторы электромагнитного поля ввиду противоположного направления токов в них. Поэтому в идеальной симметричной цепи суммарные векторы напряженности излучаемого электромагнитного поля равны нулю и излучаемое ЭМИ отсутствует.

Симметричные цепи позволяют передавать значительно более широкополосные сигналы, по сравнению с несимметричными, и поэтому широко используются для построения СКС.

Основными недостатками симметричных цепей с балансной передачей являются, во-первых, необходимость использования для приема и передачи N сигналов $2 \times N$ проводников (на каждый сигнал 2 провода) и, во-вторых, невозможность передачи постоянной составляющей сигнала.

4.2. Первичные электрические параметры витой пары

Электрические свойства витой пары, как и любой другой направляющей системы электромагнитных колебаний, полностью характеризуются ее первичными параметрами: сопротивлением R и индуктивностью L проводников, а также емкостью C и проводимостью g изоляции.

Эти параметры (R и g) обуславливают потери энергии: первый – тепловые потери в проводе и экране (при его наличии), второй – потери в изоляции.

Параметры L и C определяют реактивность витой пары как направляющей системы и, следовательно, ее частотные свойства.

Конкретные значения первичных параметров зависят от конструкции кабеля и, в частности, от геометрии отдельных его компонентов, их взаимного расположения, материала проводников, изоляции и внешних покровов и т. д.

4.2.1. Емкость

Конструктивно витая пара представляет собой два проводника, отделенных друг от друга слоем твердой изоляции и воздушным промежутком.

Такая структура может рассматриваться как конденсатор, где роль обкладок играют проводники, а функции диэлектрика выполняют расположенные между ними изоляционный материал и воздух, и обладает заметной емкостью, величина которой линейно возрастает по мере увеличения длины.

Электрическая емкость между проводниками витой пары ограничивает ширину полосы пропускания кабеля и приводит к искажениям высокочастотной части спектра передаваемого сигнала.

Емкость не зависит от частоты. Однако из-за особенностей методов, применяемых в процессе ее определения, при указании величины емкости часто указывается значение частоты сигнала, на которой проводятся измерения.

По стандарту EIA/TIA 568 для кабелей категории 3 на длине 100 м емкость не должна превышать 6,6 нФ, а для кабелей категорий 4 и 5 – 5,6 нФ. Некоторые фирмы выпускают по своим собственным техническим условиям кабели с существенно меньшей емкостью.

4.2.2. Активное сопротивление

Активное сопротивление зависит от материала провода, его длины и сечения, а также от температуры. Проводники витых пар, применяемых в СКС, изготавливаются из меди, обладающей низким удельным сопротивлением. Чем меньше сечение провода, чем больше его длина и чем выше температура, тем выше активное сопротивление и соответственно затухание витой пары.

Согласно требованиям стандарта EIA/TIA 568 при температуре 20°C сопротивление постоянному току любого проводника витой пары длиной 100 м не должно превышать 9,38 Ом.

С увеличением частоты сигнала активное сопротивление провода возрастает. Это вызвано тем, что, во-первых, в результате поверхностного эффекта происходит вытеснение тока к поверхности проводника и, во-вторых, ток протекает в основном по поверхности, обращенной ко второму проводнику (эффект близости). Оба эти эффекта приводят к уменьшению эффективного сечения проводника и, в конечном итоге, к увеличению сопротивления. Для минимизации вредного влияния этих эффектов проводники витых пар выполняются в виде монолитного провода, а не скрученными из нескольких тонких проводов.

Применение проводников из нескольких тонких проводов возможно только в соединительных шнурах, где требуется в первую очередь высокая гибкость и устойчивость к многократным изгибам, а повышенное затухание сказывается не столь сильно из-за небольшой общей длины.

4.2.3. Индуктивность

Витая пара состоит из двух изолированных проводников, каждый из которых при протекании через него тока накапливает энергию, то есть обладает свойством индуктивности. По мере увеличения частоты за счет поверхностного эффекта происходит уменьшение индуктивности.

4.2.4. Проводимость изоляции

Результирующая проводимость изоляции витой пары может быть записана в виде суммы двух составляющих: $g = g_0 + g_f$, где g_0 учитывает токи утечки, связанные с несовершенством диэлектрика, а g_f учитывает затраты энергии на диэлектрическую поляризацию.

Под поляризацией понимают переориентацию под действием электромагнитного поля связанных диполей, имеющих в диэлектрике. Переменное электромагнитное поле вызывает вибрацию диполей, которая приводит к повышению температуры диэлектрика. Нагрев диэлектрика, в свою очередь, облегчает вибрацию и повышает проводимость, что сопровождается ростом затухания сигнала. Особенно много диполей, образованных атомами хлора, содержится в поливинилхлориде, который является типовым изоляционным материалом для кабелей.

4.3. Вторичные параметры кабелей на основе витой пары

Вторичные параметры витой пары рассчитываются на основе первичных или, что значительно чаще, определяются экспериментально.

Вторичные параметры нормируются в технических условиях на витую пару и позволяют простыми средствами выполнить инженерный расчет линий связи, построенных на основе симметричного кабеля, и оценить их пригодность для передачи сигналов тех или иных приложений.

Волновое сопротивление, или импеданс, – это сопротивление, которое встречает электромагнитная волна при распространении вдоль любой однородной (то есть без отражений) направляющей системы, в том числе и витой пары. Оно свойственно данному типу кабеля и зависит только от его первичных параметров и частоты.

Волновое сопротивление связано с первичными параметрами следующим простым соотношением:

$$Z_c = \sqrt{\frac{(R + j\omega L)}{(g + j\omega C)}}$$

Волновое сопротивление численно равно входному сопротивлению линии бесконечной длины, которая имеет оконечную нагрузку, равную ее собственному волновому сопротивлению. Оно измеряется в Омах и определяет количественное соотношение между электрической и магнитной составляющей электромагнитной волны. В общем случае волновое сопротивление является комплексной величиной, его модуль падает по мере роста частоты и на высоких частотах стремится к фиксированному активному сопротивлению:

$$Z_{C\infty} = \lim_{\omega \rightarrow \infty} \sqrt{\frac{(R + j\omega L)}{(g + j\omega C)}} = \sqrt{L/C}.$$

Кабели на витых парах на звуковых частотах, то есть при передаче телефонных сигналов, имеют сопротивление около 600 Ом, по мере увеличения частоты оно быстро падает и на частотах свыше 1 МГц вплоть до верхней граничной частоты конкретного кабеля не должно отличаться от 100 Ом более чем на + 15 %.

При распространении по витой паре электромагнитный сигнал постепенно теряет свою энергию. Этот эффект называется ослаблением, или затуханием. Затухание принято оценивать в децибелах как разность между уровнями сигналов на выходе передатчика и входе приемника. Один децибел соответствует изменению мощности в 1,26 раза или напряжения в 1,12 раза.

Принято различать собственное и рабочее затухание кабеля.

Под собственным затуханием кабеля понимается затухание при работе в идеальных условиях. В обобщенном виде его величину теоретически можно определить как действительную часть коэффициента распространения γ , который связан с первичными параметрами следующим простым соотношением:

$$\gamma = \sqrt{(R + j\omega L)(g + j\omega C)}.$$

Экспериментально собственное затухание кабеля можно определить как разность уровней входного и выходного сигналов в том случае, если сопротивление источника сигнала и нагрузки равны между собой и волновому сопротивлению кабеля.

В процессе реальной эксплуатации это условие выполняется не во всех случаях, что обычно сопровождается увеличением затухания. Такое затухание называется рабочим.

Из изложенного следует важный практический вывод о том, что для минимизации рабочего затухания и его приближения к собственному сопротивление источника сигнала и нагрузка должны быть равны волновому сопротивлению, то есть, по терминологии электротехники, должна быть обеспечена согласованная нагрузка как источника сигнала, так и самого кабеля.

Из формулы выше следует, что затухание является частотозависимой величиной и, как все входящие в него параметры, зависит от длины кабеля. Результаты анализа формулы показывают, что затухание связано с длиной витой пары линейной зависимостью на всех частотах.

Для упрощения выполнения инженерных расчетов удобно пользоваться параметром коэффициента затухания или погонного затухания α ,

который численно равен затуханию кабеля фиксированной длины (применительно к кабелю типа витой пары это обычно 100 м).

Величины коэффициента затухания α , длины L и затухания A связаны между собой следующим простым соотношением:

$$A \text{ [дБ]} = \alpha \text{ [дБ/100 м]} \times L \text{ [м]}/100$$

Чем меньше величина затухания, тем более мощным оказывается сигнал на входе приемника и тем устойчивее при прочих равных условиях связь. Затухание вызывается активным сопротивлением и потерями в диэлектрической изоляции. Определенный вклад в затухание вносят также излучение электромагнитной энергии и отражения.

Любой проводник, по которому течет переменный ток, является источником излучения в окружающее пространство. Оно отбирает у сигнала энергию и ведет к возрастанию затухания сигнала. Это явление резко возрастает с увеличением частоты сигнала. При $\lambda < a$, где λ – длина волны электромагнитного сигнала, a – расстояние между проводниками, большая часть энергии идет на излучение в окружающее пространство и передача в неэкранированной направляющей системе становится невозможной. Для стандартной витой пары величина параметра a имеет значение порядка 2 мм, то есть критическая частота для нее будет равна 15 ГГц, что на два порядка ниже рабочих частот самых совершенных витых пар (150 МГц). С ростом частоты потери на электромагнитное излучение возрастают. Для минимизации потерь на излучение применяют балансную передачу и скрутку проводников в пары.

Как было отмечено выше, в идеальной симметричной цепи электромагнитное излучение отсутствует. На практике таких идеальных симметричных цепей не существует. Дело в том, что в такой цепи проводники должны бесконечно плотно прилегать друг к другу и в пределе быть стянутыми в бесконечно тонкую линию, суммарный протекающий ток через которую равен нулю. Проводники с меньшим диаметром и более тонкой изоляцией плотнее прилегают друг к другу. Однако чрезмерное уменьшение сечения проводника и уменьшение толщины изоляции ведет к повышению затухания за счет роста активного сопротивления и увеличения проводимости изолирующих покровов.

Из схемы на рис. 3.4 можно сделать вывод о том, что затухание с ростом частоты имеет тенденцию к росту. Это обусловлено как ростом сопротивления продольной ветви в основном за счет индуктивности L , так и падением сопротивления поперечной ветви, которое обусловлено главным образом наличием емкости C . По стандарту EIA/TIA 568 на длине 100 м и при температуре 20° С частотная характеристика $A(f)$

максимально допустимого затухания, начиная с 0,772 МГц, для кабелей категорий 3, 4 и 5 определяется согласно выражению

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

где A , дБ – максимальное допустимое затухание; f , МГц – частота сигнала; k_1, k_2, k_3 – константы, определяемые в зависимости от категории кабеля (табл. 4.1)

Таблица 4.1

Категория кабеля	K1	K2	K3
3	2,320	0,238	0,000
4	2,050	0,043	0,057
5	1,967	0,023	0,050

На рис. 4.5 показаны частотные зависимости предельно допустимых затуханий кабелей различных категорий, вычисленные по формуле.

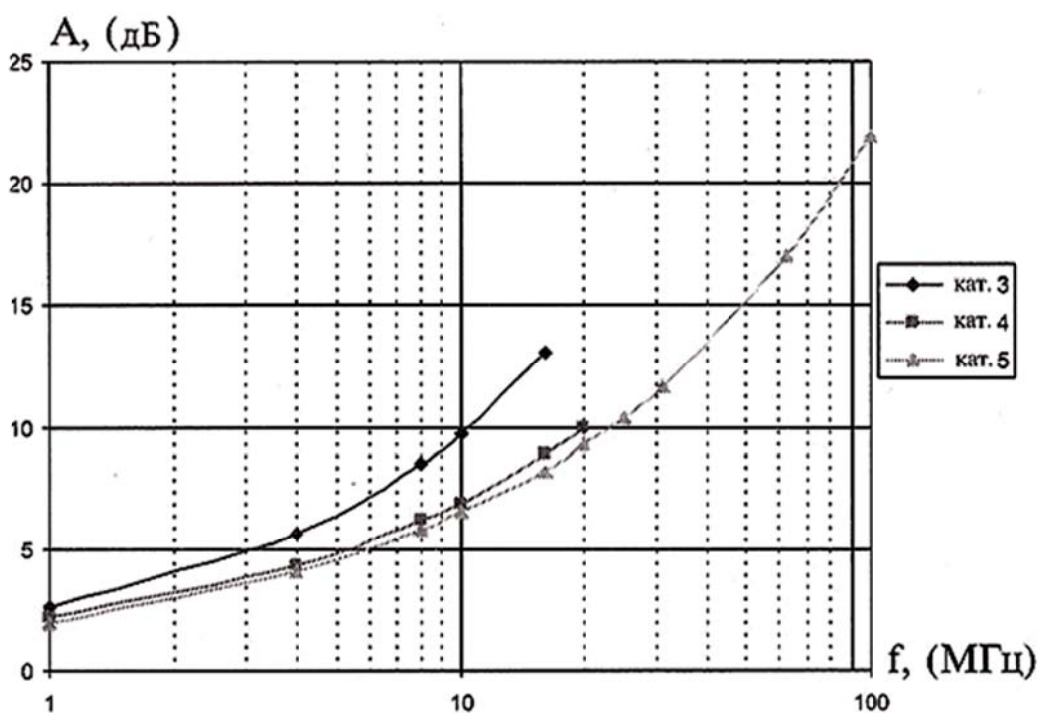


Рис. 4.5. Максимальное допустимое затухание кабелей категории 3, 4 и 5 на длине 100 м при $t=20\text{ }^{\circ}\text{C}$

Аппроксимация по формуле оказалась очень удачной и достаточно часто используется многими производителями кабельной продукции для описания характеристик их изделий. При этом принимаются свои значения коэффициентов k_1 – k_3 , а область действия распространяется на частоты до 400 и даже 550 МГц.

Кроме аналитического задания величины затухания стандарт EIA/TIA – 568 определяет этот параметр также в табличной форме с расширением нормируемых значений в область нижних частот. Это бывает полезным при выполнении инженерных расчетов трактов связи, предназначенных для поддержки работы некоторых приложений, а также позволяет сразу же получить необходимую информацию без выполнения вычислений.

При передаче сигнала часть его энергии вследствие неидеальности балансировки витой пары переходит в электромагнитное излучение, которое вызывает наведенные токи в соседних парах. Этот эффект называется переходными наводками. Наводки, накладываясь на полезные сигналы, передаваемые по соседним парам, могут приводить к ошибкам приема и в конечном итоге снижают качество связи.

Разность между уровнями передаваемого сигнала и создаваемой им помехи на соседней паре называется переходным затуханием. В зависимости от места и метода измерения этого параметра различают несколько видов переходного затухания. Рис. 4.6 показывает два типа наводок: на ближнем конце и на дальнем конце. На рисунке через I_1 обозначены токи наводок, создаваемые различными участками влияющей витой пары во влияемой паре.

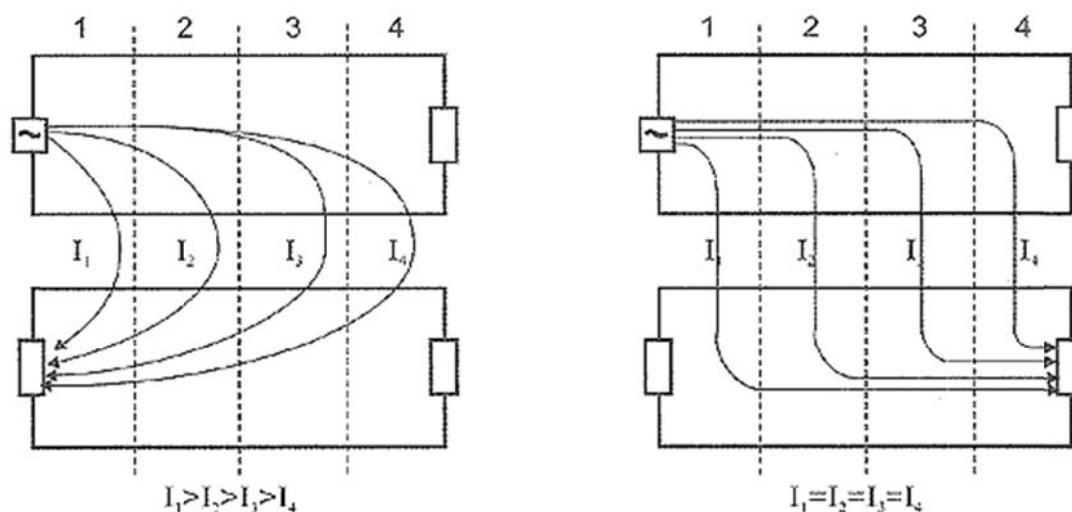


Рис. 4.6. Переходные наводки на ближнем (слева) и дальнем (справа) концах соседней пары

Если источник сигнала и точка измерения находятся на одном конце, то говорят о переходном затухании на ближнем конце, если на разных – то о переходном затухании на дальнем конце. В технике СКС первое из них традиционно имеет заимствованное из англоязычной технической литературы обозначение NEXT (Near End Crosstalk), а второе – FEXT (Far End Crosstalk).

Чем выше значение NEXT и FEXT, тем меньший уровень имеет наводка в соседних парах, и соответственно тем более качественным является кабель. С практической точки зрения представляет интерес частотная зависимость переходного затухания на ближнем и дальнем концах, а также зависимость этих параметров от длины линии. Влияющая пара и пара, подверженная влиянию, проложены параллельно под общей защитной оболочкой. За счет этого их проводники могут рассматриваться как обкладки конденсатора. Это означает, что с ростом частоты переходное затухание падает. Стандарт нормирует минимальные значения переходного затухания на ближнем конце при длине кабеля 100 м. Для определения минимально допустимого параметра NEXT на частотах, превышающих 0,772 МГц, используется следующее аппроксимирующее выражение:

$$\text{NEXT}(f) = \text{NEXT}(0,772) - 15 \lg(f/0,772),$$

где $\text{NEXT}(0,772)$ – минимально допустимое переходное затухание на ближнем конце на частоте 0,772 МГц, которое для кабелей категорий 3, 4 и 5 принимается равным 43, 58 и 64 дБ соответственно; f , МГц – частота сигнала.

Результаты расчетов по формуле выше приведены на рис. 4.7.

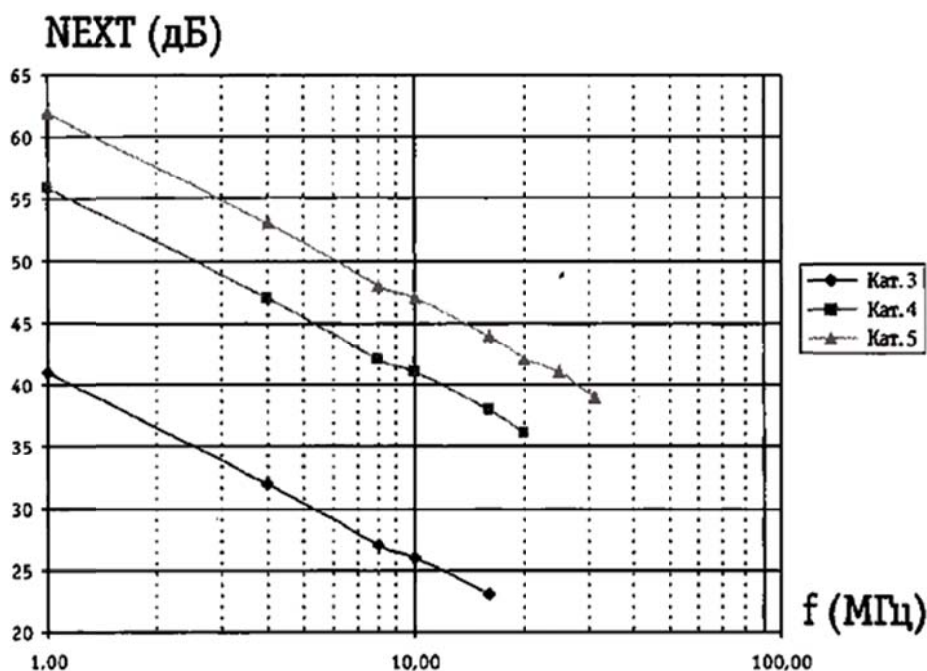


Рис. 4.7. Максимально допустимые значения NEXT для кабелей категории 3, 4 и 5 на длине 100 м

Дополнительно стандарт нормирует значения NEXT на частотах менее 0,772 МГц, что бывает необходимо для некоторых приложений. Нормируемые значения в этом случае представляются в табличной форме.

Суммирование отдельных составляющих одной частоты переходной помехи на ближнем конце происходит с различными фазами (по напряжению). Поэтому реальный график частотной зависимости величины NEXT имеет вид кривой с резкими перепадами величин переходного затухания на близких частотах. Стандарты нормируют только минимальную величину параметра NEXT, и кабель считается соответствующим требованиям стандарта, если во всем рабочем частотном диапазоне реальная величина NEXT не падает ниже определенного нормами значения.

Типовая зависимость переходного затухания на ближнем и дальнем концах от длины линии показана на рис. 4.8.

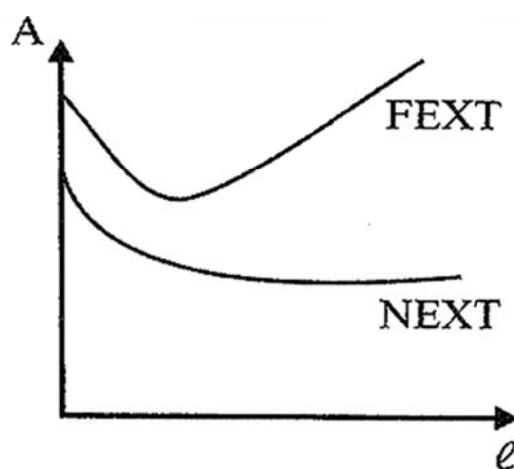


Рис. 4.8. Зависимость переходного затухания на дальнем и ближнем концах от длины линии

Переходное затухание на ближнем конце с увеличением длины линии сначала несколько уменьшается, а затем стабилизируется. Качественное объяснение этого эффекта состоит в том, что, начиная с определенной длины линии, токи помех с отдаленных участков приходят на ближний конец настолько ослабленными, что практически не увеличивают взаимного влияния между цепями, и величина NEXT остается постоянной. Отсюда следует, что значения NEXT для двух концов одной пары могут существенно различаться между собой, поэтому все стандарты требуют его измерения с обеих сторон.

График зависимости переходного затухания на дальнем конце от длины линии носит экстремальный характер. Вначале, пока длина линии мала, увеличение ее протяженности увеличивает мощность помехи. По мере увеличения длины начинает проявляться рост затухания помеховых составляющих, и FEXT монотонно возрастает.

Для улучшения параметра NEXT в симметричных кабелях применяют различный шаг скрутки витых пар. Кроме ослабления электромагнитной связи отдельных пар такое решение не позволяет им плотно

прилегать друг к другу по всей длине, что дополнительно увеличивает переходное затухание.

Известно, что сетевое оборудование различного назначения по-разному использует симметричный кабель как среду передачи. Поэтому в зависимости от приложения и метода использования кабеля нормирование величины переходных помех или, что эквивалентно, переходного затухания выполняется по-разному.

Наиболее популярными ЛВС в настоящее время являются сети Ethernet. При использовании полнодуплексного режима передатчик и приемник работают одновременно, и эта аппаратура использует для работы две витые пары одного кабеля. Этот случай в схематическом виде изображен на рис. 4.9.

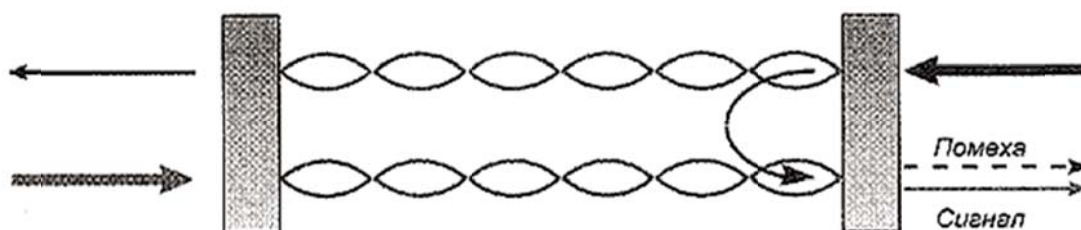


Рис. 4.9. Определение NEXT

При этом ослабленный после прохождения по витой паре информационный сигнал взаимодействует на входе приемника с мощной переходной помехой работающего на этом же конце передатчика. Поэтому достаточно нормировать следующий параметр:

$$NEXT = P_C - \max P_n,$$

где P_C уровень сигнала, P_n — уровень создаваемой им переходной помехи.

Величина $\max P_n$ берется в наихудшем случае, так как заранее неизвестно, какие две пары будут использоваться сетевым оборудованием для организации информационного обмена.

В последнее время при построении сетей обозначилась тенденция использования для передачи информации одновременно нескольких пар. С другой стороны, сигналы нескольких приложений все чаще передаются в одном многопарном кабеле. В данной ситуации нормирование только параметра NEXT оказывается недостаточным, так как на приемник одновременно действует несколько источников помех. Для учета этого обстоятельства используется более сложная расчетная модель, которая для 4-х парного кабеля имеет вид, изображенный на рис. 4.10 (все пары действуют на одну), и нормируется параметр так называемой суммарной мощности (power sum).

Из-за разного расстояния между парами, различного шага скрутки и т. д. разность между величинами NEXT и PS-NEXT оказывается равной не 4,8 дБ, а примерно 2 дБ.

Наконец, в новейших перспективных приложениях типа Gigabit Ethernet вход приемника и выход передатчика развязаны с помощью дифференциальной схемы. Это позволяет одновременно использовать одну витую пару для приема и передачи сигналов. В этой ситуации дополнительно к переходным помехам на ближнем конце необходимо учитывать также помехи на дальнем конце и соответственно нормировать величину переходного затухания на дальнем конце:

$$FEXT = P_C - \max P_{\text{п}}$$

где, $P_{\text{п}}$ – уровень переходной помехи на дальнем конце.

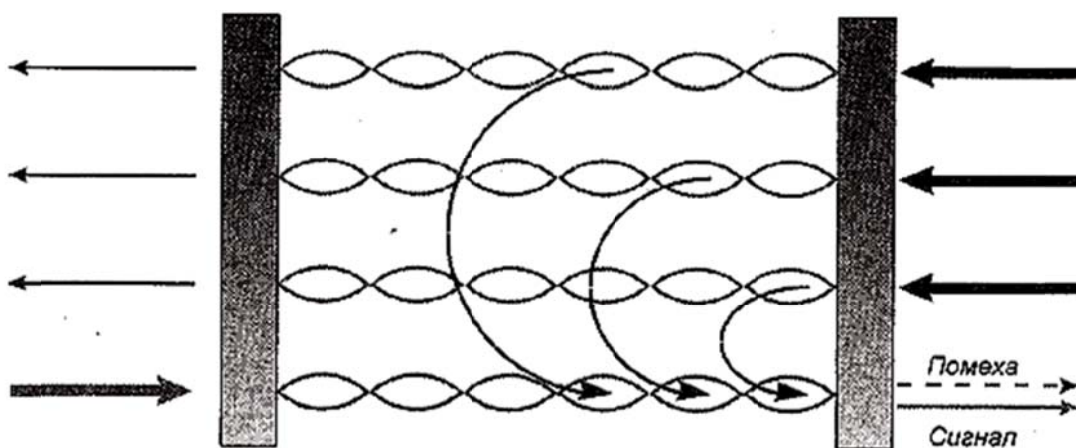


Рис. 4.10. Определение PS-NEXT

Аналогично переходной помехе на ближнем конце можно также ввести параметр PS-FEXT. Аналогично переходной помехе на ближнем конце может нормироваться и значение суммарной переходной помехи на дальнем конце. Переходная помеха на дальнем конце обычно оказывается меньшей по сравнению с переходной помехой на ближнем конце. Однако в отличие от помех на ближнем конце эти помеховые составляющие достаточно часто суммируются синфазно или с небольшой разностью фаз, что может дополнительно увеличить их мощность.

И, наконец, некоторые производители начинают нормировать так называемую глобальную переходную помеху GXT (global crosstalk), которая равна сумме наведенных переходных помех на обоих концах кабеля.

В настоящий момент официальными редакциями стандартов задаются только величины NEXT и PS-NEXT (последнее значение приводится для многопарных и комбинированных кабелей), нормирование величин FEXT и GXT производится ограниченным количеством фирм.

4.4. Защищенность

Для оценки качества передачи информации в технике проводной связи широко используется параметр защищенности от помех, или просто защищенности, который представляет собой разность между уровнями полезного сигнала и помехи в рассматриваемой точке.

Для расчетной модели уровень сигнала составляет $P_c = P_{\text{пер}} - A$, а уровень переходной помехи $P_{\text{пт}} = P_{\text{пер}} - \text{NEXT}$. Защищенность согласно определению будет равна:

$$\text{ACR} = \text{NEXT} - A,$$

то есть зависит только от величин затухания и переходного затухания.

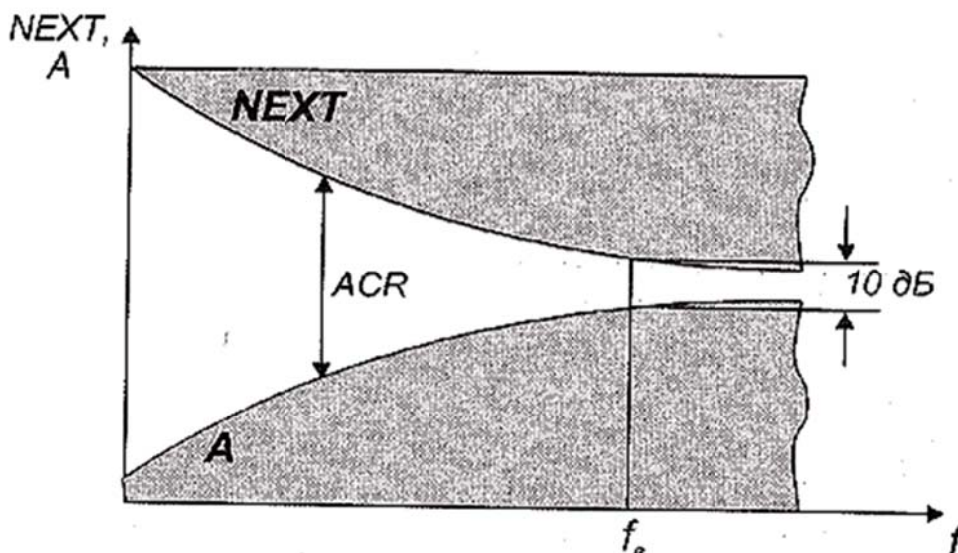


Рис. 4.11. Определение параметра защищенности

Параметр ACR определяет величину превышения помехи полезным сигналом и поэтому является интегральной характеристикой качества кабеля. Используемая для обозначения защищенности аббревиатура ACR означает Attenuation to Crosstalk Ratio. По мере увеличения величины ACR при прочих равных условиях начинает возрастать отношение сигнал/шум, и соответственно растет устойчивость связи. Из-за того что NEXT и A зависят от частоты, параметр ACR также является частотно-зависимым. Стандарт ISO/IEC 11801 регламентирует минимально допустимые значения ACR для кабелей категории 5 на частотах 20 МГц и выше. TIA/EIA 568 специально не оговаривает предельных значений ACR на разных частотах, однако они могут быть вычислены по формуле $\text{ACR} = \text{NEXT} - A$. Результаты этих расчетов для кабелей категорий 3, 4 и 5 на длине 100 м представлены на рис. 4.12.

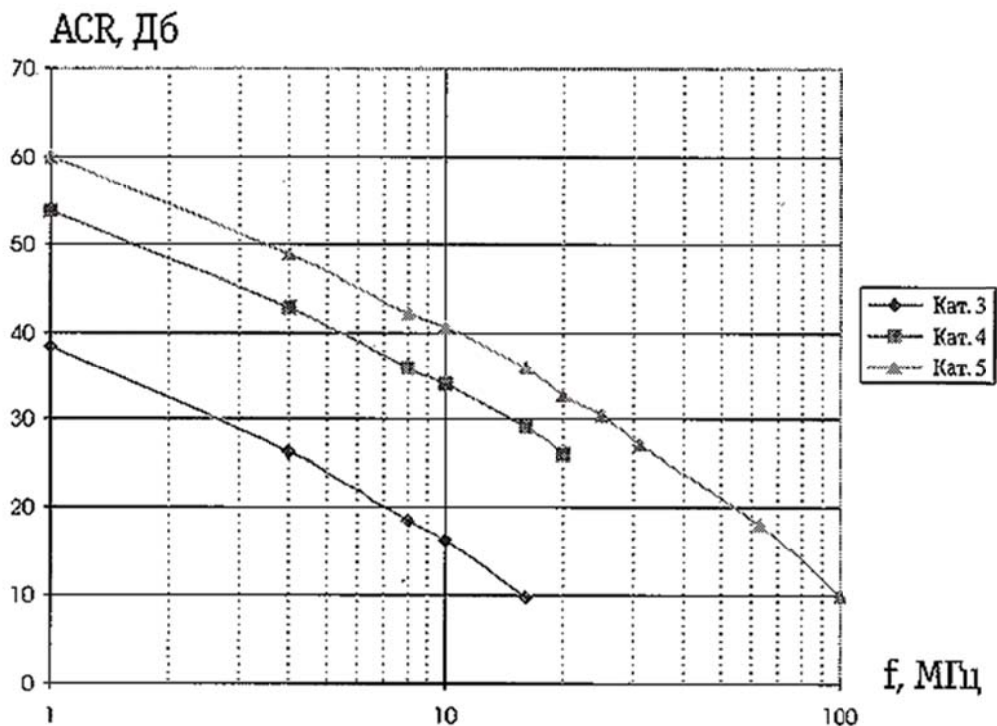


Рис. 4.12. Расчетные значения минимально допустимых параметров ACR по данным стандарта TIA/EIA 568 для кабелей категории 3, 4 и 5 на длине 100 м

Из этого рисунка видно, что, в худшем случае, сигнал на входе приемника должен превышать шумы наводок от соседней пары не менее чем на 10 дБ, что эквивалентно отношению сигнал/шум в 3,16 раз по напряжению или в 10 раз по мощности.

Введение параметра ACR позволяет конкретизировать понятие верхней граничной частоты кабеля. Считается, что кабели из витых пар с установленными на них оконечными разъемами обеспечивают устойчивую полнодуплексную работу любого приложения с такой верхней граничной частотой, на которой параметр ACR составляет 10 дБ. Это положение отдельно выделено на рис. 4.11.

Исключением из данного правила являются кабели категории 4, у которых на частоте 20 МГц $ACR = 26$ дБ. При этом верхнюю граничную частоту приложения не следует путать с максимальной частотой кабеля, на которой изготовитель сертифицирует его параметры, так как зачастую на ней значения ACR получаются отрицательными (особенно ярко это проявляется для неэкранированных конструкций с относительно невысоким NEXT). Необходимость сертификации параметров кабеля на этих частотах возникает для оценки возможности его использования для полудуплексной или однонаправленной (симплексной) передачи каких-либо сигналов, например телевизионных.

В случае высокочастотных приложений, которые в процессе работы используют для передачи информации все витые пары и одновременно в двух направлениях, нормирование только величины ACR оказывается недостаточным. Для расчета помеховой составляющей, создаваемой наводками на дальнем конце, используется аналогичная ACR величина ELFEXT=FEXT-A.

Применяемое для обозначения этого параметра сокращение ELFEXT означает Equal Level for Far End Crosstalk – эквивалентный уровень переходного затухания на дальнем конце.

4.5. Задержка распространения сигнала по витым парам

Параметр NVP (от англ. Nominal Velocity of Propagation) является мерой замедления скорости распространения электромагнитной волны вдоль витой пары, он численно равен отношению последней величины к скорости света в вакууме и выражается в виде десятичной дроби или в процентах. Достаточно редко для его обозначения применяется также сокращение VOP (velocity of propagation). Необходимость его учета связана с тем, что конечная скорость распространения приводит к появлению довольно значительной задержки прохождения сигналов, что может быть критичным для некоторых приложений, например для сетей Fast Ethernet.

Стандарты задают только самые общие требования к величине NVP и в настоящее время не определяют метод ее измерения.

Таблица 4.2

Величина NVP для витой пары категорий 3, 4, 5

Частота, МГц	NVP		
	Кат.3	Кат.4	Кат.5
1,00	0,4	0,6	0,65
10,00	0,6	0,6	0,65
100,00	-	-	0,65

Величина NVP витой пары зависит от диаметра проводников, расстояния между ними и от типа диэлектрика. Так, например, для проводников с полиэтиленовой изоляцией NVP составляет 0,65...0,7, с изоляцией из тефлона – 0,69...0,73, а применение поливинилхлоридных материалов уменьшает его значение до 0,60...0,64. Как и большинство первичных и вторичных параметров витой пары NVP имеет определенную зависимость от частоты. При типичном значении NVP в диапазоне 60...75 % для сегмента в 100 м время прохождения сигнала составляет 370...550 нс. Значение NVP меняется в зависимости от со-

стояния изоляции жил кабеля и температуры окружающей среды не более чем на 3 % от номинального значения.

В кабелях, содержащих несколько витых пар, каждая из них будет характеризоваться индивидуальным временем прохождения сигнала от передатчика к приемнику.

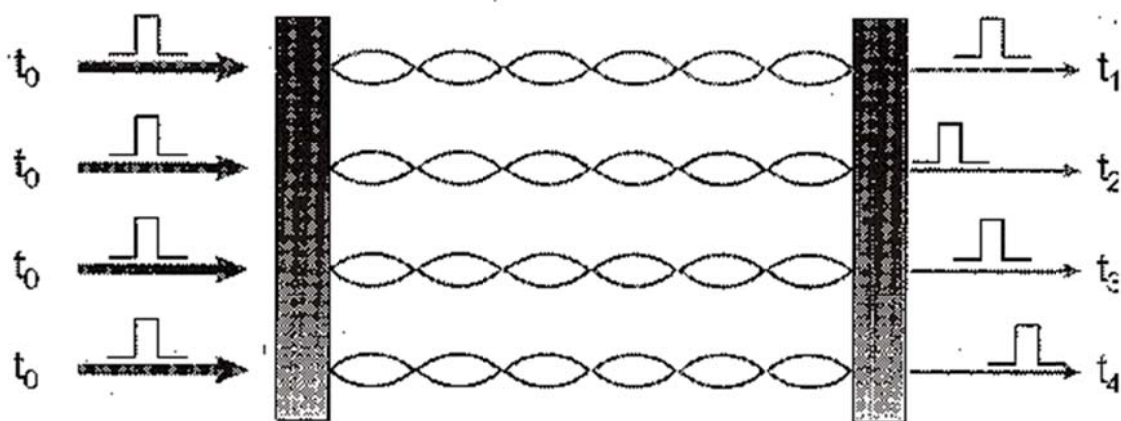


Рис. 4.13. К определению параметра skew

Разброс времени прохождения вызывается двумя основными причинами:

- разбросом скоростей распространения электромагнитной волны по витым парам;
- разностью электрических длин витых пар, обусловленной различным шагом скрутки.

Максимальная разность задержек прохождения между всеми парами обозначается параметром skew. Действующие редакции стандартов не предъявляют специальных требований к skew, поэтому этот параметр нормируют в технических условиях на свою продукцию только некоторые производители. Полезность нормирования значений skew объясняется, главным образом, следующими двумя причинами.

Во-первых, оборудование некоторых высокоскоростных локальных сетей использует для передачи сигналов все четыре пары одновременно. На передающей стороне сигнал распределяется по четырем парам, а на приемной опять собирается в исходный вид. Нормальное функционирование такой схемы передачи информации возможно только в том случае, если сигналы по всем четырем парам достигают приемника одновременно или по крайней мере не с очень большим разбросом. При нарушении этого условия возможен сбой синхронизации, и восстановление исходного сигнала на приемной стороне станет невозможным. Например, для сетей 100Base-T4 значение параметра skew не должно превышать 50 нс на максимальной длине сегмента.

Во-вторых, принцип действия портативных кабельных сканеров при измерении длины кабельных сегментов основан на измерении времени задержки между посылкой короткого импульса и приходом отраженного сигнала. При известной скорости распространения электромагнитных волн по кабелю можно вычислить его длину. Если скорости распространения по парам будут различны, то сканер будет выдавать разные значения длины кабеля по разным витым парам.

Если все пары изготовлены из одних и тех же материалов, по одной и той же технологии и в едином технологическом цикле, то разброс задержек распространения по ним сигналов будет весьма мал (как правило менее 10 нс). В последнее время некоторые производители стали выпускать кабели, в которых с целью повышения пожарной безопасности изоляция одной или двух пар изготовлена из тефлона, а остальных – из других изоляционных материалов. Хотя электрические характеристики таких кабелей соответствуют требованиям пятой категории, их skew может достигать 60 нс и более.

Влияние на величину skew различного шага скрутки пар кабеля менее значительно и, как правило, не превышает 10 нс.

Возможность изменения параметра NVP за счет изменения шага скрутки и выбора материала изоляции используется в некоторых конструкциях горизонтальных кабелей для минимизации величины skew. В этих изделиях изоляцией из «быстрого» тефлона покрываются проводники тех пар, которые дают наибольшее время задержки.

4.6. Структурные возвратные потери

Реальная линия всегда имеет более или менее сильные неоднородности, которые приводят к появлению отражений. Основными источниками неоднородностей являются производственные и эксплуатационные дефекты кабеля, разъемные соединители и оконечные нагрузки с сопротивлением, отличным от волнового. Электромагнитная волна, встречая такие неоднородности в процессе распространения по кабелю, частично отражается от них и возвращается к началу. При наличии нескольких неоднородностей волна претерпевает серию отражений. Результатом этого является возникновение в линии двух дополнительных паразитных потоков энергии: обратного, состоящего из суммы отраженных волн и попутного, возникающего в результате двойных отражений.

Обратный поток приводит к колебаниям входного сопротивления. Это затрудняет согласование с аппаратурой на концах линии, что сопровождается ростом рабочего затухания и приводит к искажениям передаваемого сигнала.

Интенсивность обратного отражения электромагнитных волн в местах неоднородности волнового сопротивления характеризуются параметром SRL (Structural Return Loss), который определяется как отношение мощности основного сигнала к мощности обратного потока энергии. Чем выше значение SRL, тем меньшую мощность имеют обратный и соответственно попутный потоки энергии, и тем более качественным является кабель.

Этот параметр из-за своей малости нормируется в логарифмических единицах и в стандартах указывается для длины кабеля 100 м. Например, для горизонтального кабеля категории 5 в наихудших условиях его значение до частот 20 МГц должно составлять не менее 23 дБ. В полосе частот 20...100 МГц минимально допустимая величина параметра SRL рассчитываются по формуле:

$$SRL = 23 - 10 \cdot \lg(f/20), \text{ где } f - \text{ частота в МГц.}$$

Норма на SRL должна выполняться для всех пар.

4.7. Шум от внешних источников

Вследствие неидеальности балансировки в витых парах могут возникать наводки от внешних источников электромагнитного излучения (ЭМИ). Имеется два основных вида внешних наводок:

- электромагнитная интерференция EMI (Electromagnetic Interference). Для нее характерны низкие частоты и высокие амплитуды. Эту наводку создают в основном электромоторы стартеры флуоресцентных ламп и силовые кабели;
- радиочастотная интерференция RFI (Radio Frequency Interference). Для нее характерны высокие частоты и низкие амплитуды. Основными источниками наводок этого вида являются сотовые радиотелефоны, радиовещание, телевидение и источники питания с высокочастотным преобразованием.

В целом витые пары весьма устойчивы к внешнему ЭМИ. Поэтому стандарты не предусматривают специальных требований к уровню шума, наведенного внешним ЭМИ.

ГЛАВА 5. ОПТОВОЛОКОННЫЕ ЛИНИИ СВЯЗИ

5.1. Типы оптических линий связи

Существуют два различных типа оптоволоконных кабелей:

- многомодовый, или мультимодовый, кабель, более дешевый, но менее качественный;
- одномодовый кабель, более дорогой, но имеющий лучшие характеристики.

Основные различия между этими типами связаны с разным режимом прохождения световых лучей в кабеле.

В одномодовом кабеле практически все лучи проходят один и тот же путь, в результате чего все они достигают приемника одновременно, и форма сигнала практически не искажается. Одномодовый кабель имеет диаметр центрального волокна около 1,3 мкм и передает свет только с такой же длиной волны (1,3 мкм). Дисперсия и потери сигнала при этом очень незначительны, что позволяет передавать сигналы на значительно большее расстояние, чем в случае применения многомодового кабеля. Для одномодового кабеля применяются лазерные приемопередатчики, использующие свет исключительно с требуемой длиной волны. Такие приемопередатчики пока еще сравнительно дороги и не слишком долговечны. Однако в перспективе одномодовый кабель должен стать основным благодаря своим прекрасным характеристикам.

В многомодовом кабеле траектории световых лучей имеют заметный разброс, в результате чего форма сигнала на приемном конце кабеля искажается. Центральное волокно имеет диаметр 62,5 мкм, а диаметр внешней оболочки – 125 мкм (это иногда обозначается как 62,5/125). Для передачи используется обычный (не лазерный) светодиод, что снижает стоимость и увеличивает срок службы приемопередатчиков по сравнению с одномодовым кабелем. Длина волны света в многомодовом кабеле равна 0,85 мкм. Допустимая длина кабеля достигает 2...5 км. В настоящее время многомодовый кабель – основной тип оптоволоконного кабеля, так как он дешевле и доступнее.

Задержка распространения сигнала в оптоволоконном кабеле не сильно отличается от задержки в электрических кабелях. Типичная величина задержки для наиболее распространенных кабелей составляет около 4...5 нс/м.

5.2. Передача сигналов по волоконным световодам

В волоконно-оптических кабельных системах сигналы передаются несущей оптического (обычно ближнего инфракрасного) диапазона волн по световодам из сверхчистого кварцевого стекла. Упрощенная структурная схема волоконно-оптической линии связи (ВОЛС) показана на рис. 5.1.

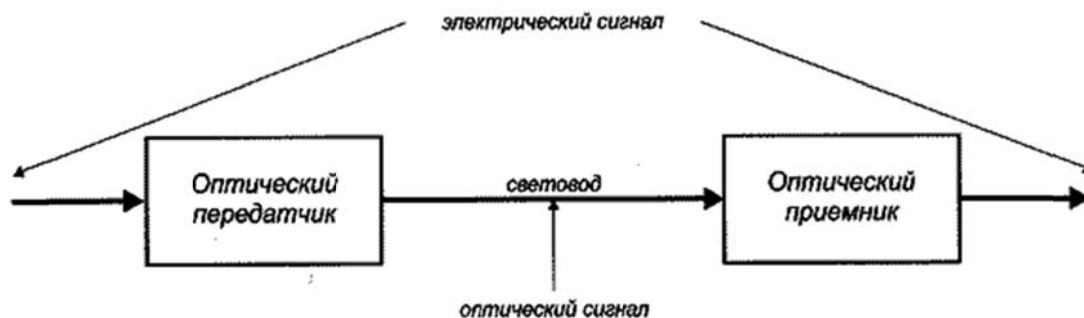


Рис. 5.1. Упрощенная структурная схема ВОЛС

Электрический сигнал поступает на вход оптического передатчика и модулирует интенсивность выходного сигнала излучателя. Оптический сигнал распространяется по волоконному световоду и поступает на вход оптического приемника, который осуществляет его демодуляцию и восстанавливает исходный электрический сигнал. Для обеспечения возможности нормальной эксплуатации оптический передатчик и приемник снабжаются розетками оптических разъемов. Схема управления и излучатель оптического передатчика, а также фотодиод и усилитель фототока со схемой согласования оптического приемника в подавляющем большинстве случаев интегрируются в одном корпусе.

Волоконный световод представляет собой тонкое стеклянное волокно цилиндрической формы, по которому происходит передача электромагнитного излучения оптического (обычно ближнего инфракрасного) диапазона длин волн, соответствующего частотам $10^{14} \div 10^{15}$ Гц.

Принцип действия волоконного световода основан на использовании известных процессов отражения и преломления оптической волны на границе раздела двух сред с различными оптическими свойствами. Оптические свойства материала зависят от показателя преломления n . Среда с большим показателем преломления называется оптически более плотной.

При падении луча на границу раздела двух сред в общем случае появляются преломленная и отраженная волны.

Согласно закону Снеллиуса угол падения $\varphi_{\text{п}}$ связан с углами отражения $\varphi_{\text{отр}}$ и преломления $\varphi_{\text{пр}}$ следующими соотношениями

$$\varphi_{\text{п}} = \varphi_{\text{отр}};$$

$$n_1 \cdot \sin \varphi_{\text{п}} = n_2 \cdot \sin \varphi_{\text{пр}}.$$

Если луч выходит из оптически более плотной среды в менее плотную ($n_1 > n_2$), то $\varphi_{\text{пр}} > \varphi_{\text{п}}$. По мере увеличения угла падения можно достичь такого состояния, когда преломленный луч начинает скользить по границе раздела двух сред без перехода в оптически более плотную среду. Угол падения, при котором наблюдается такой эффект, называется предельным углом полного внутреннего отражения $\varphi_{\text{пво}}$. Для всех углов падения, которые превышают предельный, будет иметь место только отражение, а преломленная волна будет отсутствовать. Это явление называется полным внутренним отражением, и оно положено в основу передачи оптического излучения по световоду.

Конструктивно волоконные световоды обычно имеют круглое поперечное сечение и образованы двумя основными элементами. В центре располагается сердцевина из оптически более плотного стекла, ее окружает оболочка из стекла с меньшей оптической плотностью. Диаметры сердцевины и оболочки принято измерять в микронах и указывать в технических характеристиках волокна в явном виде через знак косой черты следующим образом: «диаметр сердцевины»/ «диаметр оболочки», например, 62,5/125 обозначает световод с диаметром сердцевины в 62,5 мкм и с оболочкой диаметром 125 мкм.

На границе раздела сердцевины и оболочки происходит отражение оптических лучей, которые распространяются вдоль оси световода. Таким образом, сердцевина служит для передачи электромагнитной энергии, а оболочка предназначена для создания условий отражения на границе раздела двух сред и защиты от излучения энергии в окружающую среду.

Излучение внешнего источника, падающее на входной торец волоконного световода, возбуждает в нем несколько типов волн, которые называются модами. В свою очередь, моды делятся на направляемые, вытекающие и излучаемые.

К направляемым относятся такие моды, которые распространяются вдоль сердцевины волокна и обеспечивают передачу информации (лучи 1 и 2 на рис. 5.2).

Направляемые моды считаются основным типом электромагнитной волны и возбуждаются теми лучами, которые падают на торец волокна под углом, не превышающим предельный угол Θ_A , называемый апертурным углом. Основные типы современных световодов имеют апертурный угол в пределах от 11,5 до 17 градусов.

Лучи, которые падают на торец волокна под углом, превышающим Θ_A , достигают границы раздела сердцевина-оболочка и за счет прелом-

ления в оболочку теряют часть энергии, испытывая при этом большое затухание (луч 3 на рис. 5.2). Эти моды называются вытекающими.

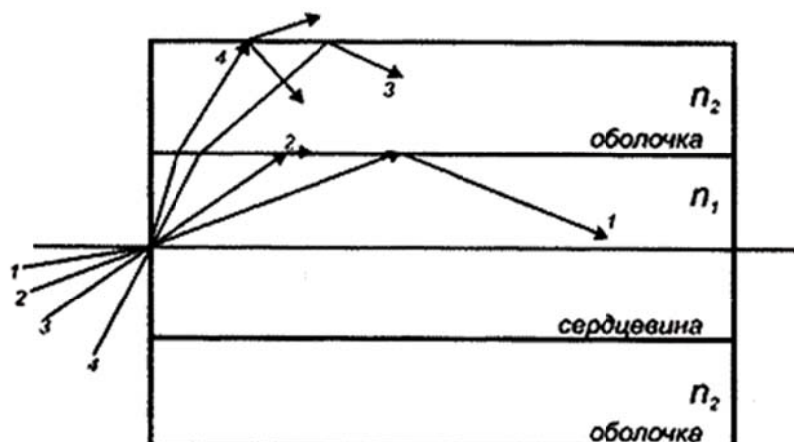


Рис. 5.2. Прохождение лучей в волоконном световоде

Наконец, при падении лучей под углами, существенно превышающими Θ_A , часть из них достигает внешней поверхности оболочки и излучается в окружающее пространство (луч 4 на рис. 5.2). Такие моды называются излучаемыми. Излучаемые моды возникают также в местах нерегулярностей световодов. Появление вытекающих и излучаемых мод приводит к росту потерь и искажению передаваемой информации.

5.3. Типы волоконных световодов

Диаметр оболочки наиболее распространенных световодов составляет 125 мкм. В области диаметров сердцевин наблюдается существенно большее разнообразие. В зависимости от диаметра сердцевины оптические волокна делятся на две группы: одномодовые и многомодовые, причем последние могут быть ступенчатыми и градиентными. В многомодовых световодах диаметр сердцевины выбирается много большим длины волны оптической несущей, и условия полного внутреннего отражения выполняются для нескольких типов волн (мод), количество которых в серийных волокнах обычно составляет от 1000 до 2000.

Показатель преломления оболочки, как правило, имеет постоянное значение, тогда как показатель преломления сердцевины может оставаться постоянным или же изменяться вдоль ее радиуса по определенному закону, который носит название профиля показателя преломления.

Простейшим типом волоконного световода является так называемый ступенчатый световод. В нем показатель преломления сердцевины остается постоянным вдоль ее радиуса. В ступенчатых многомодовых волокнах траектории лучей отдельных мод имеют вид зигзагообразных линий (см. рис. 5.3, а).

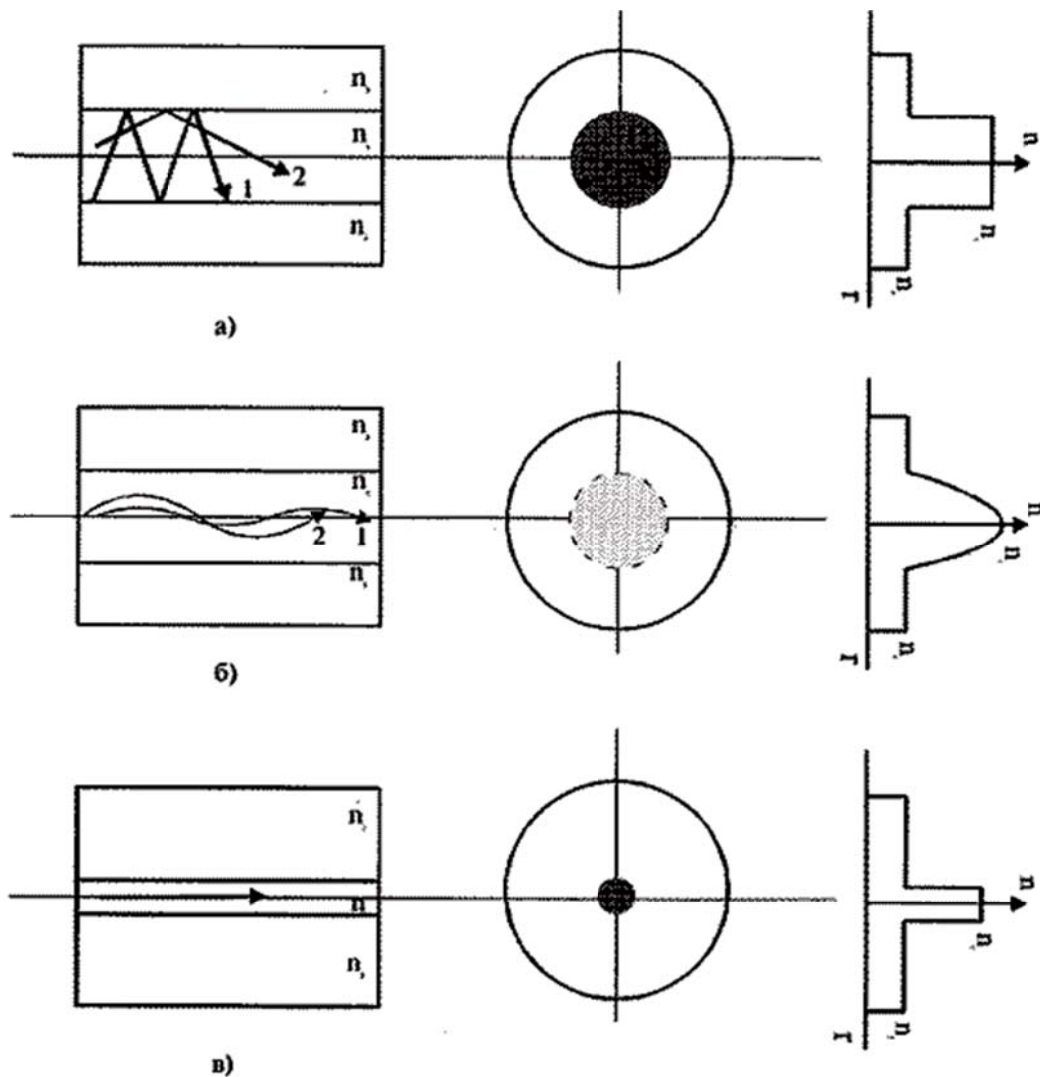


Рис. 5.3. Распространение световых лучей в световодах

От профиля показателя преломления в значительной степени зависят частотные свойства многомодовых световодов, поэтому на практике часто применяют отличные от ступенчатых профили. Так, например, в градиентном световоде показатель преломления за счет изменения количества легирующих добавок, главным образом германия, плавно снижается по мере удаления от оси по закону, близкому к квадратичной параболе (см. рис. 5.2, б). В таких волокнах траектории распространения большинства лучей представляют собой плавные кривые. В США для градиентных световодов наиболее популярны сердцевинны с диаметром 62,5 мкм, а в Европе и в России часто используются также волокна с диаметром сердцевинны 50 мкм.

В одномодовых световодах диаметр сердцевинны (7...10 мкм) соизмерим с длиной волны, и за счет этого в нем существует только одна направляемая мода (рис. 5.2, в).

5.4. Дисперсия электромагнитного излучения

Под дисперсией понимают увеличение длительности оптического импульса при прохождении им волоконного световода определенной длины, обычно 1 км.

Дисперсия определяет полосу пропускания световода и возникает из-за рассеяния во времени спектральных и модовых составляющих оптического сигнала. Наличие этого эффекта приводит к увеличению длительности оптического импульса в процессе его прохождения по оптическому кабелю, к появлению межсимвольной интерференции (наложению фронтов двух соседних импульсов друг на друга) и, в конечном счете, к увеличению вероятности ошибки принимаемого сигнала.

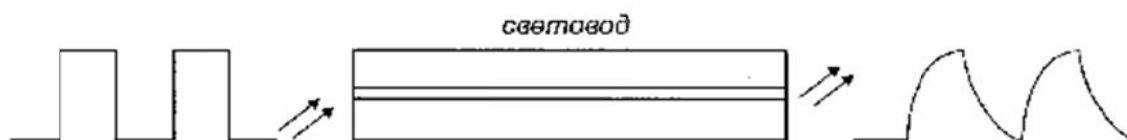


Рис. 5.4. Прохождение оптического импульса по световоду

Различают два вида дисперсии: межмодовую и хроматическую.

Хроматическая дисперсия обусловлена зависимостью условий распространения света по волокну от длины волны. Хроматическая дисперсия возникает из-за конечной спектральной ширины оптической несущей выходного сигнала передатчика и имеет две составляющие: материальную и волноводную. Материальная дисперсия обусловлена зависимостью показателя преломления сердцевины и оболочки от длины волны. Волноводная дисперсия определяется зависимостью условий распространения моды от длины ее волны. Степень влияния хроматической дисперсии на частотные характеристики канала зависит в первую очередь от спектральных свойств источника излучения. Для лазерных источников благодаря узкой полосе частот излучаемой несущей дисперсия характерна в меньшей степени. В некогерентных источниках (светодиодах) полоса несущей существенно шире, и хроматическая дисперсия проявляется значительно сильнее.

Межмодовая дисперсия возникает только в многомодовых световодах из-за наличия в них большого числа мод с различным временем распространения за счет разброса углов отражения и соответственно различной длины пути, который отдельные моды проходят в сердцевине волокна. Градиентный световод обладает свойством удержания мод вблизи центра сердцевины. Лучи, которые распространяются дальше от центра, проходят при этом больший путь, однако в материале с мень-

шей оптической плотностью, то есть с более высокой скоростью. Кроме того, по сравнению с волокном со ступенчатым профилем показателя преломления в нем существует меньшее количество мод. Совместное действие этих трех факторов приводит к тому, что градиентный световод имеет лучшие частотные свойства, и ширина его полосы пропускания более чем на порядок превышает аналогичный показатель ступенчатого волокна. Одновременно градиентные составляющие дисперсии световоды почти не отличаются от ступенчатых по таким эксплуатационным параметрам, как удельные потери, эффективность ввода излучения, сложность сращивания и т. д., и за счет этого практически полностью вытеснили последние из массового использования.

Результирующее значение расширения импульсов за счет межмодовой $\tau_{\text{мод}}$, материальной $\tau_{\text{мат}}$ и волноводной $\tau_{\text{вв}}$ дисперсии определяется по формуле

$$\tau = (\tau_{\text{мод}}^2 + (\tau_{\text{мат}} + \tau_{\text{вв}})^2)^{1/2}.$$

С учетом реального соотношения величин отдельных составляющих дисперсии для многомодовых волокон можно считать $\tau = \tau_{\text{мод}}$, а для одномодовых волокон $\tau = \tau_{\text{мат}} + \tau_{\text{вв}}$

Для одномодовых световодов параметр дисперсии учитывает ее зависимость от спектральных свойств источника излучения и поэтому имеет размерность пс/(нм·км). Современные волокна имеют величину дисперсии в пределах от 3 до 15...18 пс/(нм·км). Волноводная $\tau_{\text{вв}}$ и материальная $\tau_{\text{мат}}$ составляющие хроматической дисперсии в районе длин волн 1200...1600 нм имеют, как правило, противоположные знаки. Так как волноводная дисперсия зависит от профиля показателя преломления, то, варьируя этим параметром, для одномодовых световодов можно на заранее заданной длине волны или же в определенной спектральной полосе получить нулевую или близкую к нулевой дисперсию. Поэтому для одномодовых волокон вместо абсолютной величины дисперсии иногда указывают значение волны нулевой дисперсии и крутизну спектральной характеристики дисперсии в окрестностях этой длины волны (для серийных световодов эти параметры обычно составляют 1300...1310 нм и 3,5 пс/(нм·км) соответственно).

Частотные характеристики многомодовых волокон из-за преобладающего влияния межмодовой составляющей дисперсии мало зависят от спектральных свойств источника излучения. Поэтому их удобно оценивать эквивалентом дисперсии в частотной области, называемым коэффициентом широкополосности имеющим размерность МГц·км. В зависимости от длины волны типовые значения коэффициента широкополосности для современных световодов составляют 200...500 МГц·км.

5.5. Затухание сигналов в световодах

При расчете параметров линейно-кабельных сооружений оптической линии связи из-за экспоненциальной зависимости мощности сигнала от расстояния удобно пользоваться логарифмическими единицами, переход к которым позволяет свести основные расчеты мощности сигналов к операциям сложения и вычитания.

В качестве нулевого уровня в технике оптической связи принято значение 1 мВт, которое примерно соответствует максимальной мощности излучения типичного полупроводникового лазера и светодиода, а под уровнем мощности понимается величина

$$p = 10 \cdot \lg(P/1 \text{ мВт}) \text{ [дБм]},$$

где P – мощность оптического сигнала.

В табл. 5.1 дано соответствие уровней и мощностей оптических сигналов в типичных для волоконной оптики диапазонах их изменения.

Таблица 5.1

Уровень, дБм	Мощность	Уровень, дБм	Мощность
0	1 мВт	-30	1 мкВт
-10	100 мкВт	-33	500 нВт
-13	50 мкВт	-40	100 нВт
-20	10 мкВт	-50	10 нВт
-23	5 мкВт	-60	1 нВт

В процессе распространения по волокну оптический сигнал постепенно теряет свою энергию. Этот эффект называется затуханием. От величины затухания зависит максимальная дальность связи между двумя приемопередатчиками. В волоконно-оптической технике связи затухание принято измерять в децибелах.

Затухание в волоконных световодах обусловлено потерями на рассеяние и потерями на поглощение.

Потери на рассеяние обусловлены флуктуациями средней плотности и связанными с ними флуктуациями показателя преломления материала сердцевины. При попадании на нерегулярности оптические лучи меняют свое направление распространения и высвечиваются в окружающее пространство.

Потери на рэлеевское рассеяние (кривая 1 на рис. 5.5) и инфракрасное поглощение (кривая 2 на рис. 5.5) определяют нижний предел потерь волоконного световода.

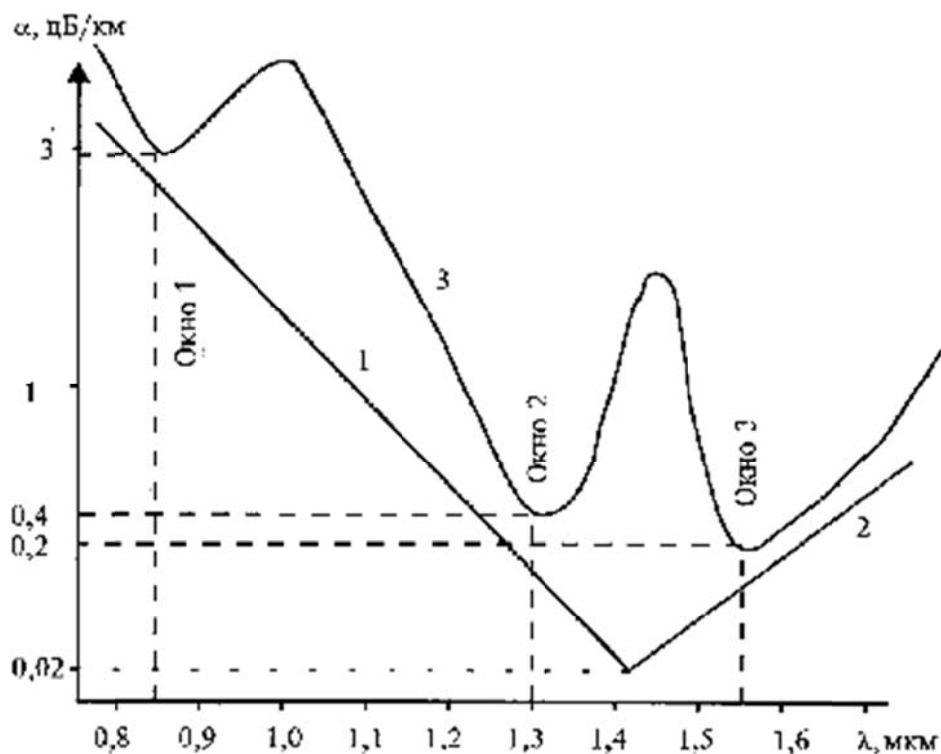


Рис. 5.5. Зависимость затухания кварцевого волоконного световода от длины волны

Как видно из рисунка, нижний теоретический предел потерь в кварцевом световоде находится в районе длины волны 1,4 мкм. Минимальная величина потерь никогда не достигается по двум основным причинам:

- в процессе изготовления оптического кабеля возникают так называемые кабельные потери, которые обусловлены деформациями волокна при наложении покрытий и защитных оболочек, а также скруткой элементов кабельного сердечника;
- материал сердцевины световода из-за конечной эффективности очистки содержит атомы и молекулы примесей, резонансное возбуждение которых приводит к резкому возрастанию потерь на определенных длинах волн. В диапазоне 0,8...1,7 мкм наибольший вклад в величину потерь вносят пары воды и металлы первой группы периодической системы элементов (никель, железо, хром и др.). Для получения приемлемого уровня потерь концентрация примесей не должна превышать 10^{-9} .

Таким образом, область эффективного использования кварцевых волоконных световодов ограничена диапазоном видимого и ближнего инфракрасного диапазонов (ИК) длин волн. Для среднего и дальнего ИК диапазонов необходимы новые материалы.

На рис. 5.5 схематично в виде кривой 3 показана спектральная зависимость коэффициента затухания реальных световодов с учетом фундаментальных и дополнительных потерь, вызываемых примесями. Из графика ясно следует, что работа по волоконно-оптическим кабелям эффективна не на всех длинах волн, а только в определенных участках спектра, где достигаются минимальные потери. Области минимальных потерь получили название окон прозрачности. Для кварцевых световодов практический интерес представляют три окна прозрачности. За их границы удобно принять значения, приведенные в стандарте ISO/IEC 11801 и перечисленные в табл. 5.2. Характеристики полупроводниковых излучателей и фотоприемников оптимизированы для работы в этих окнах.

Таблица 5.2

Характеристики окон прозрачности световодов

Окно прозрачности	Длина волны, мкм			Типовое затухание, дБ/км
	Минимальная	Центральная	Максимальная	
1	790	850	910	2...3
2 (MM)	1285	1300	1330	0,7...1,5
2(SM)	1288	1310	1339	0,4...1,0
3	1525	1550	1575	0,2...0,4

Из рис. 5.5 и табл. 5.2 видно, что переход из первого во второе окно прозрачности дает существенный выигрыш по величине затухания, тогда как работа в третьем окне большого выигрыша не приносит. С другой стороны, по мере увеличения рабочей длины волны начинает быстро расти стоимость активных оптоэлектронных компонентов. В линиях оптической связи локальных сетей, для обслуживания которых в основном используются волоконно-оптические тракты магистральных подсистем СКС, из-за сравнительно малой протяженности кабельных трасс стоимость оконечной аппаратуры является относительно большой величиной. Поэтому с учетом перечисленных выше обстоятельств в технике СКС в подавляющем большинстве случаев используют первое и второе окно прозрачности. Нормировка параметров одномодовых световодов, используемых при создании подсистем внешних магистралей, выполняется из соображений предпосылок применения в СКС одномодовых оптических кабелей, разработанных для городских и междугородных сетей связи. Линии дальней связи, стоимость которых определяется в первую очередь длиной участка регенерации, работают в основном во втором и третьем окне прозрачности, где кроме низкого затухания достигается также малая величина дисперсии.

5.6. Излучатели и фотоприемники

Излучатели предназначены для преобразования электрического входного сигнала в выходной оптический. Известны два основных класса излучателей, которые пригодны для использования в технике оптической связи: полупроводниковые светодиод и лазер.

Светодиоды изготавливаются в основном по арсенид-галлиевой технологии, генерируют некогерентное излучение и используются для работы по многомодовому оптическому кабелю на длинах волн 850 и 1300 нм. Типовая выходная мощность составляет порядка 1 мВт, однако из-за значительной ширины диаграммы направленности значение потерь ввода составляет, как правило, 10...17 дБ. Светодиод как излучатель волоконно-оптической линии связи эффективен при скоростях передачи до 100...155 Мбит/с.

Таблица 5.3

Световые излучатели

Диапазон	850 нм	1300 нм
Структура	GaAlAs	GaInAsP GaAlAsP
Мощность излучения, вводимая в световод 62,5/125 мкм, дБм	-13 + -18	-14 + -20
Ширина спектра излучения, нм	50	100
Время переключения, нс	4 + 6.5	0.6 + 3.0

Типовые параметры светодиодов, применяемых в активном волоконно-оптическом сетевом оборудовании, приведены в табл. 5.3.

Полупроводниковый лазер в отличие от светодиода генерирует близкое к монохроматическому излучение с шириной спектра 1...5 нм, работает на длинах волн 1300 и 1550 нм и используется для передачи по одномодовому волокну. Быстродействие современных лазеров позволяет модулировать их сигналами с частотами до 10 ГГц и выше. Типовая выходная мощность излучения составляет 5 мВт, потери на ввод обычно не превышают 2...5 дБ.

Диаграмма направленности полупроводникового лазера имеет меньшую ширину, что при прочих равных условиях позволяет ввести в волокно большую мощность. Одновременно полупроводниковый лазер имеет меньшую ширину спектральной характеристики, что имеет своим следствием снижение дисперсионных искажений передаваемого сигнала. При создании лазерных передатчиков приходится решать ряд сложных инженерных задач, что в конечном итоге увеличивает их стоимость примерно в несколько раз по сравнению со светодиодными.

Суммарная длина подсистем внешних и внутренних магистралей СКС при их построении на многомодовом кабеле должна составлять не более 2 км. Такая длина линии вполне обеспечивается светодиодными излучателями. Лазерные источники излучения используются в основном в сетевом оборудовании, подключенном к подсистеме внешних магистралей на одномодовом кабеле длиной 60 и более километров.

Фотоприемники обеспечивают обратное преобразование оптического излучения в электрический сигнал. Основными требованиями, предъявляемыми к фотоприемникам, являются высокая чувствительность на рабочей длине волны, низкий уровень собственных шумов, стабильность параметров, высокая надежность и низкое напряжение питания. Перечисленным требованиям в наиболее полной мере удовлетворяют лавинные и р-і-п фотодиоды, которые в основном и применяются в сетевом оборудовании оптической связи. Лавинные фотодиоды за счет эффекта внутреннего умножения имеют примерно на порядок более высокую чувствительность, однако требуют очень стабильного высоковольтного источника питания и имеют более высокую стоимость. Они находят применение в основном в линиях большой протяженности. В локальных системах связи, для обслуживания которых создаются СКС, высокая чувствительность является избыточной, и оптические приемники аппаратуры построены исключительно на р-і-п фотодиодах.

ГЛАВА 6. ОСНОВЫ БЕСПРОВОДНОЙ ПЕРЕДАЧИ

Эта глава описывает фундаментальные основы беспроводной передачи. Глава начинается с краткого обзора антенн и затем рассмотрены принципы распространения сигнала.

6.1. Антенны

Антенна может быть определена как электрический провод или система проводов, используемых или для того, чтобы излучить электромагнитную энергию в пространство или для того, чтобы собрать электромагнитную энергию из окружающего пространства. Для передачи сигнала электрическая энергия радиочастотного диапазона от передатчика преобразовывается в электромагнитную энергию антенной и излучается в окружающую среду (атмосфера, космос). Для приема сигнала электромагнитная энергия попадающая в антенну преобразуется в электрическую энергию, которая передается в приемник.

В двусторонней связи, одна и та же антенна часто используется и для передачи и для приема сигнала. Это возможно, потому что любая антенна передает энергию из окружающей среды на вход приемника с той же самой эффективностью, что передает энергию от выходов передатчика в пространство, если используется одна и та же частота сигнала. Характеристики антенны одинаковы, независимо от того посылает ли антенна или получает электромагнитную энергию.

6.1.1. Диаграмма направленности антенны

Антенна излучает мощность во всех направлениях, но обычно не делает это одинаково хорошо в каждом направлении. Общий способ характеризовать работу антенны – диаграмма направленности, которая является графическим представлением излучательных свойств антенны как функции пространственных координат. Самую простую диаграмму направленности имеет идеализированная антенна, известная как изотропная антенна. Изотропная антенна излучает мощность во всех направлениях одинаково. Фактически диаграмма направленности для изотропной антенны это сфера с антенной в центре. Однако, диаграммы направленности почти всегда изображаются как плоское поперечное сечение трехмерного образца. Образец для изотропной антенны показан на рис. 6.1, *a*. Расстояние от антенны до каждой точки на диаграмме направленности пропорционально мощности, излученной антенной в данном направлении. Рис. 6.1, *b* показывает диаграмму направленности другой идеализированной антенны. Это

направленная антенна, в которой излучение идет предпочтительно в одном направлении. Натуральная величина диаграммы направленности произвольна. Важным является только относительное расстояние от положения антенны в каждом направлении. Относительное расстояние определяет относительную мощность. Чтобы определить относительную мощность в данном направлении, нужно провести линию от положения антенны под соответствующим углом и найти точку пересечения этой линии с кривой диаграммы направленности. Рис. 6.1 показывает сравнение двух направлений передачи, А и В на двух диаграммах направленности. Изотропная антенна дает излучение равной мощности во всех направлениях и векторы А и В имеют равную длину. Для антенны Герца, вектор В длиннее вектора А, указывая, что в направлении В излучено больше мощности чем в направлении А, и относительные длины этих двух векторов пропорциональны мощности, излученной в этих двух направлениях.

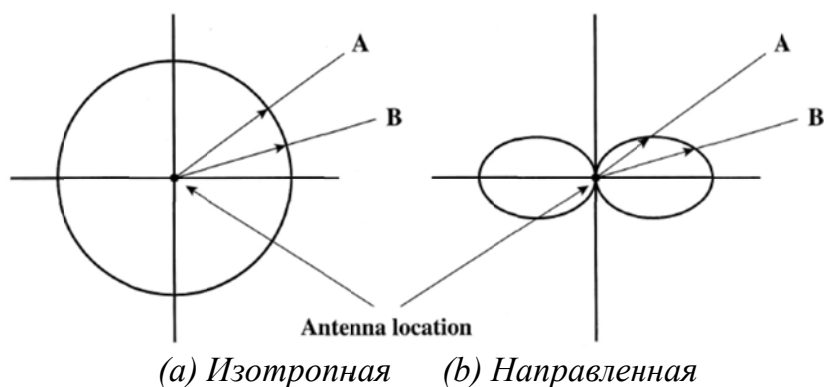


Рис. 6.1. Идеализированные лучевые диаграммы направленности

Лучевая диаграмма направленности дает удобное средство определения ширины луча антенны, которая является общей мерой направленности антенны. Ширина луча является углом, в пределах которого мощность, излученная антенной является по крайней мере половиной мощности, которая излучается в самом привилегированном направлении.

Когда антенна используется для приема, лучевая диаграмма становится лучевой диаграммой направленности в режиме приема. Самые удаленные точки диаграммы указывают лучшее направление для приема.

6.2. Типы антенн

6.2.1. Диполи

Две самых простых и самых широко используемых антенн – полуволновой диполь, или антенна Герца (рис. 6.2, а), и четвертьволновая вертикальная антенна (рис. 6.2, б). Полуволновой диполь состоит из двух коллинеарных проводов равной длины, отделенных небольшим

промежутком. Длина антенны равна половине длины волны сигнала, который может быть передан наиболее эффективно. Вертикальная четвертьволновая антенна это тип антенны, обычно используемый для автомобильных радиоприемников и портативных радиостанций.

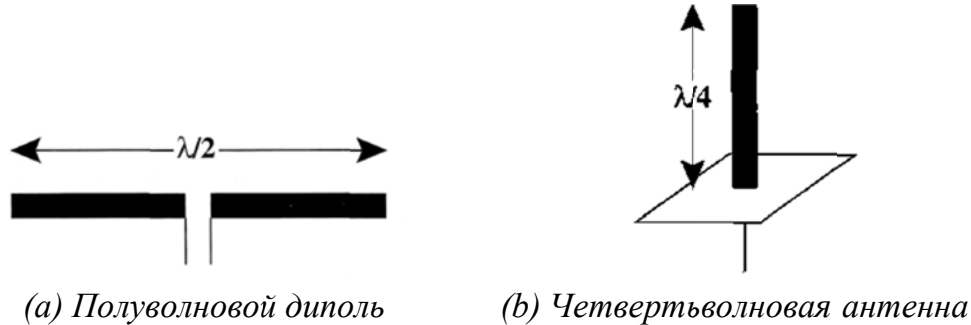
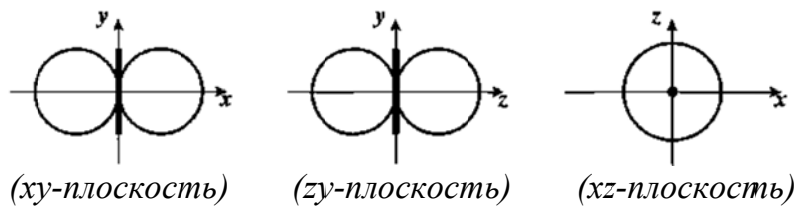
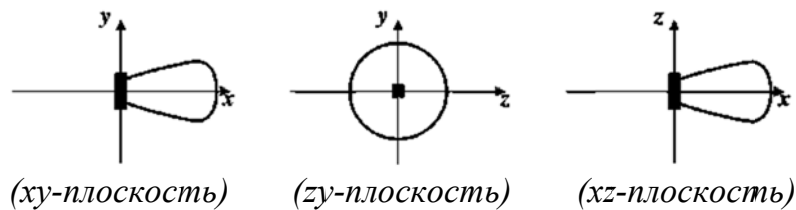


Рис. 6.2. Простые антенны



(a) Простой диполь



(b) Направленная антенна

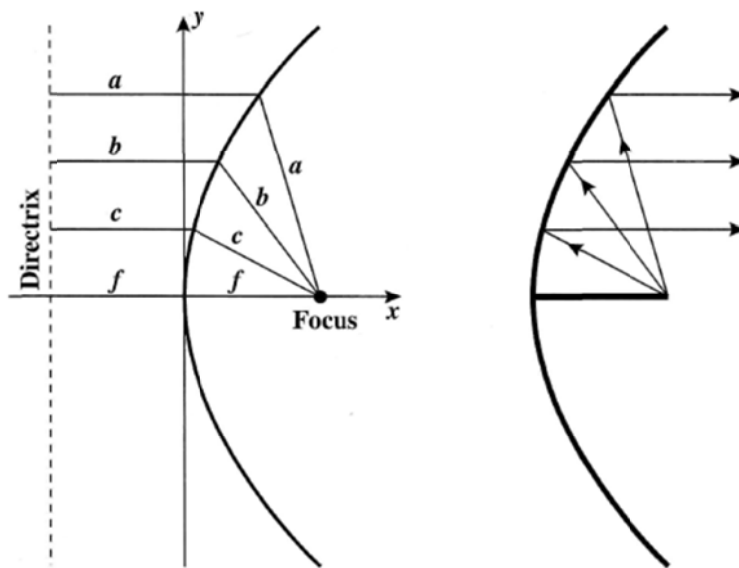
Рис. 6.3. Трехмерные диаграммы направленности

Полуволновой диполь имеет изотропную диаграмму направленности в одном измерении и диаграммы направленности в виде восьмерок в двух других измерениях (рис. 6.3, a). Для получения направленного излучения нужно использовать более сложные конфигурации антенны. Типичный направленный образец лучевой диаграммы направленности показан на рис. 6.3, b. В этом случае главное направление излучения антенны находится в направлении x .

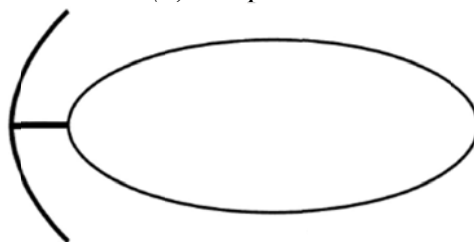
6.2.2. Параболическая антенна

Важный тип антенны это антенна с параболическим отражателем, которая используется в микроволновых и спутниковых применениях. Из курса школьной геометрии известно, что парабола – местоположение

всех точек, эквидистантных от неподвижной линии и неподвижной точки не на линии. Неподвижную точку называют фокусом, а неподвижную линию называют директрисой, или направляющей линией (рис. 6.4, *a*). Если парабола вращается вокруг своей оси, сгенерированную поверхность называют параболоидом. Поперечные сечения параболоида в плоскостях параллельных его оси формируют параболы. Поперечные сечения параболоида в плоскостях перпендикулярных его оси формируют окружности. Такие поверхности используются в автомобильных фарах, оптических и радио-телескопах, микроволновых антеннах из-за следующего свойства: Если источник электромагнитной энергии (или звук) помещен в центр параболоида, и если параболоид отражающая поверхность, то электромагнитная волна будет распространяться по линиям, параллельным оси параболоида. Рис. 6.4, *b* показывает этот эффект в поперечном сечении. Теоретически этот эффект создает пучок параллельных лучей без расхождения. Практически будет некоторое расхождение, потому что источник энергии не точечный. Обратное также верно. Если поступающие волны будут параллельны оси отражающей параболической антенны, то получающийся сигнал будет сконцентрирован в центре.



(*a*) Параболическая антенна (*b*) Отражательные свойства параболоида



(*c*) Диаграмма направленности параболической антенны

Рис. 6.4. Параболическая антенна

Рис. 6.4, с показывает типовую лучевую диаграмму направленности параболической антенны. Табл. 6.1 содержит данные о ширине излучаемого луча для антенн различных размеров на частоте 12 ГГц. Отметим что, чем больше диаметр антенны, тем более направленный луч у антенны.

Таблица 6.1

*Ширины луча антенны
для различных диаметров параболической антенны на частоте 12 ГГц*

Диаметр антенны (м)	Ширина луча (градусы)
0.5	3.5
0.75	2.33
1.0	1.75
1.5	1.166
2.0	0.875
2.5	0.7
5.0	0.35

6.2.3. Усиление антенны

Усиление антенны это мера ее направленности. Коэффициент усиления антенны определяется как выходная мощность, отдаваемая антенной в определенном направлении, по сравнению с мощностью, отдаваемой во всех направлениях идеальной всенаправленной антенной (изотропной антенной). Например, если антенна имеет усиление 3 дБ, то эта антенна лучше изотропной антенны в данном направлении на 3 дБ, или в 2 раза. Увеличение мощности, излучаемой в данном направлении, происходит за счет снижения мощности в других направлениях. Важно отметить, что усиление антенны определяет не общее повышение мощности излученного или принятого сигнала, а перераспределение мощности по разным направлениям.

Понятие, связанное с усилением антенны это понятие эффективной площади антенны. Эффективная площадь антенны связана с физическим размером антенны и с ее формой. Соотношение между коэффициентом усиления антенны и эффективной площадью определяется формулой

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

где G – коэффициент усиления антенны, A_e – эффективная площадь, f – частота, c – скорость света ($3 \cdot 10^8$ м\с), λ – длина волны.

Табл. 6.2 показывает коэффициент усиления антенны и эффективную площадь некоторых типичных форм антенн.

6.3. Распространение радиоволн

Сигнал, излученный от антенны идет по одному из трех направлений путей: наземная волна, пространственная волна или волна по линии соединяющей источник и приемник. Табл. 6.3 показывает, в каком частотном диапазоне преобладает каждый тип волн.

Таблица 6.2

Коэффициент усиления и эффективная площадь некоторых антенн

Тип антенны	Эффективная площадь, A_e (m^2)	Коэффициент усиления по мощности (относительно изотропной)
Изотропная	$\lambda^2/4\pi$	1
Бесконечно малый диполь или петля	$1.5 \lambda^2/4\pi$	1.5
Полуволновой диполь	$1.64 \lambda^2/4\pi$	1.64
Рупор, площадью A	$0.81 A$	$10 A/\lambda^2$
Параболическая, площадью A	$0.56 A$	$7 A/\lambda^2$
Турникет (два пересеченных, перпендикулярных диполя)	$1.15 \lambda^2/4\pi$	1.15

Распространение наземной волны (рис. 6.5, а) следует за контуром земли. Такая волна может распространяться на значительные расстояния, больше чем в пределах прямой видимости. Этот тип распространения радиоволн эффективен на частотах до 2 МГц. Несколько причин определяют тенденцию электромагнитной волны в этом диапазоне частот следовать за искривлением земли. Первая причина состоит в том, что электромагнитная волна наводит ток в поверхности земли. Действие этого тока состоит в том, что фронт импульса около земли замедляется, заставляя фронт импульса наклониться вниз и, следовательно, следовать за искривлением земли. Другая причина – дифракция, которая является явлением, имеющим отношение к поведению электромагнитных волн в присутствии препятствий.

Электромагнитные волны в этом частотном диапазоне рассеиваются атмосферой и не проникают через верхние слои атмосферы.

Самый известный пример связи посредством наземных волн это АМ радиовещание (амплитудно-модулированное вещание в длинноволновом диапазоне).

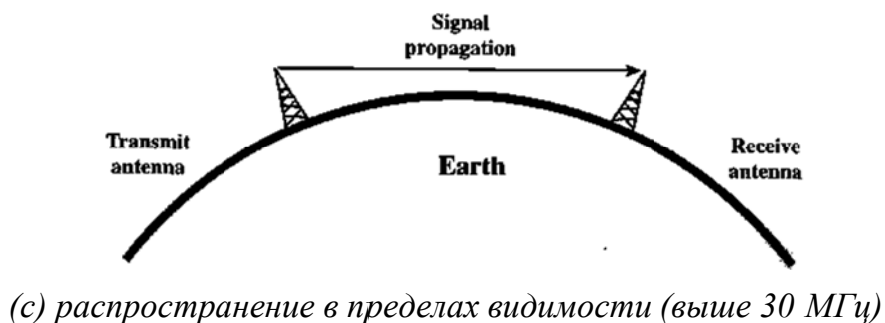


Рис. 6.5. Способы распространения радиоволн

Распространение пространственных волн используется для любительской радиосвязи, радио коротковолнового диапазона, и международного радиовещания. Распространение пространственных волн связано с отражением радиоволн от ионизованных верхних слоев атмосферы (ионосферы) и возврата этих волн к земле. Хотя кажется, что волна отражена от ионосферы, как будто ионосфера является твердой поверхностью отражения, фактически эффект вызван рефракцией. Рефракция описана далее.

Сигнал пространственной волны может испытывать множество переотражений между ионосферой и поверхностью земли (рис. 6.5, *b*). С этим способом распространения, сигнал может быть принят за тысячи километров от передатчика.

Таблица 6.3

Диапазоны частот

Диапазон	Частотный диапазон	Диапазон волн свободного пространства	Особенности распространения	Типичное использование
1	2	3	4	5
крайне низкие частоты	30...300 Гц	От 10 000 до 1 000 км	Наземные волны	Частоты линии питания; используемый некоторыми домашними системами управления.
звуковые частоты	300...3000 Гц	От 1 000 до 100 км	Наземные волны	телефонные аналоговые абонентские линии.
очень низкая частота	3...30 кГц	От 100 до 10 км	Наземные волны малое затухание; высокий уровень атмосферных помех	Навигация дальнего действия; подводная коммуникация
низкая частота	30...300 кГц	От 10 до 1 км	Наземные волны; немного менее надежный; поглощение в дневное время	Навигация дальнего действия; морские маяки связные радиостанции
средняя частота	300...3000 кГц	От 1 000 до 100 м.	Наземные волны и пространственные волны ночью; ослабление низко ночью, высоко днем; атмосферные помехи	Морское радио; пеленгация; вещание.
высокая частота	3...30 МГц	От 100 до 10 м.	пространственные волны изменяется со временем дня, сезона, и частоты.	Любительская связь; международное радиовещание, военная связь; дальняя коммуникация самолетов и кораблей

1	2	3	4	5
самая высокая частота	30...300 МГц	От 10 до 1 м.	В пределах прямой видимости; рассеивание из-за температурной инверсии; космический шум	Телевидение; FM радиопередача и двухстороннее радио
УВЧ диапазон (дециметровый диапазон)	300...3000 МГц	От 100 до 10 см	В пределах прямой видимости; космический шум	Телевидение УВЧ диапазона; мобильный телефон; радар; линии микроволновой связи; системы персональной подвижной связи
СВЧ (высокая частота высшего качества)	3...30 ГГц	От 10 до 1 см	В пределах прямой видимости; затухание из-за дождя выше 10 ГГц;	Спутниковая связь; радар; земные линии микроволновой связи; местная радиосвязь
Крайне высокие частоты (крайняя ВЧ)	30...300 ГГц	От 10 до 1 мм	В пределах прямой видимости; атмосферное ослабление из-за кислорода и водяного пара	Экспериментальный; местная радиосвязь
Инфракрасное излучение	300 ГГц к 400 THz	От 1 мм до 770 нанометров	В пределах прямой видимости	Инфракрасные ЛВС;
Видимый свет	400 THz к 900 THz	От 770 до 330 нанометров	В пределах прямой видимости	Оптическая связь

На частотах выше 30 МГц, ни наземная волна, ни пространственная волна не работают, и связь возможна только в пределах прямой видимости (рис. 6.5, с). Для спутниковой связи, сигнал свыше 30 МГц не отражается ионосферой, и поэтому сигнал может быть передан между наземной станцией и спутником на орбите, если он виден со станции (находится не за горизонтом). Для наземной связи передача и получение сигнала антенной должны быть в пределах эффективной линии видимо-

сти. Термин эффективный используется, потому что микроволны искривляются или преломляются атмосферой. Количество и направление изгиба зависит от условий, но вообще микроволны следуют за искривлением поверхности земли и поэтому распространяются дальше чем в пределах прямой видимости.

6.4. Рефракция

Кратко обсудим явление рефракции. Рефракция происходит, потому что скорость электромагнитной волны является функцией плотности среды, через которую она распространяется. В вакууме, электромагнитная волна распространяется со скоростью приблизительно $3 \cdot 10^8$ м/с. Эта постоянная, c , обычно называется скоростью света, но фактически относится к скорости света в вакууме. В воздухе, воде, стекле и других прозрачных или частично прозрачных средах, электромагнитные волны распространяются со скоростями меньше c .

Когда электромагнитная волна перемещается из среды одной плотности в среду другой плотности, ее скорость изменяется. Эффект вызывает изгиб направления волны на границе двух сред (рис. 6.6). Перемещаясь из менее плотной в более плотную среду, волна согнется к более плотной среде. Это явление легко наблюдать, частично погружая палку в воду. Результат очень напомнит иллюстрацию 5.6, с палкой, кажущейся более коротким и склонностью.

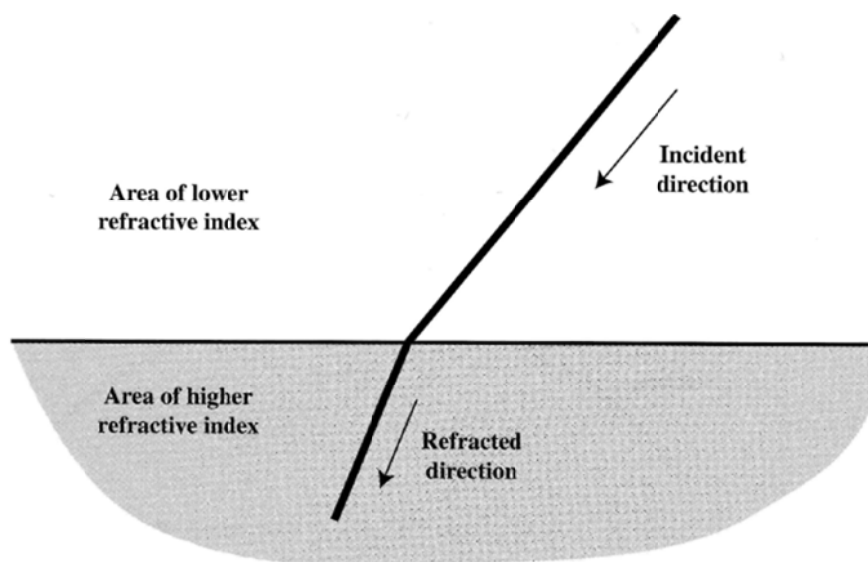
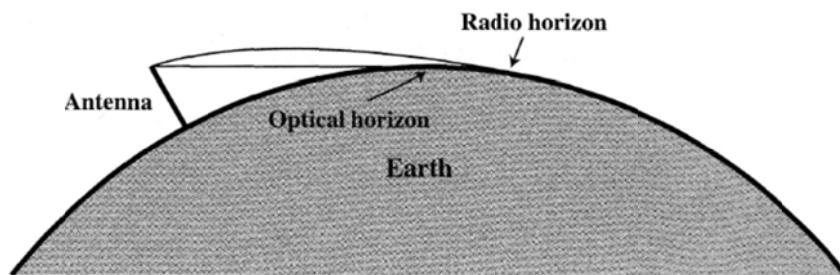


Рис. 6.6. Рефракция электромагнитной волны

Индекс рефракции одной среды относительно другого – синус угла падения, разделенного на синус угла рефракции. Индекс рефракции также равен отношению соответствующих скоростей в этих двух СМИ.

Абсолютный индекс рефракции среды вычислен по сравнению с тем из вакуума. Показатель преломления изменяется с длиной волны, так, чтобы преломляющие эффекты отличались для сигналов с различными длинами волны.



Оптическая иллюстрация 5.7 и Радиогоризонты

Хотя иллюстрация 5.6 показывает резкое, одноразовое изменение в руководстве, поскольку сигнал перемещается от одной среды до другого, непрерывный, постепенный изгиб сигнала произойдет, если это будет перемещаться через среду, в которой постепенно изменяется индекс рефракции. Под нормальными условиями распространения, показатель преломления атмосферы уменьшается с высотой так, чтобы радиоволны поехали более медленно около основания чем в более высоких высотах. Результат – небольшой изгиб радиоволн к земле.

Оптический и Линия радиосвязи Вида

Без прошедших препятствий, оптический луч обзора может быть выражен как

$$d = 3.57\sqrt{h}$$

где d – расстояние между антенной и горизонтом в километрах, и *концентратор* – высота антенны в метрах. Эффективное, или радио, луч обзора к горизонту выражен как (иллюстрация 5.7)

$$d = 3.57\sqrt{Kh}$$

где K – фактор регулирования, чтобы составлять рефракцию. Хорошее эмпирическое правило – $K = 4/3$. Таким образом, максимальное расстояние между двумя антеннами для распространения $3.57(\sqrt{Kh_1} + \sqrt{Kh_2})$ потери сигнала, где h_1 и h_2 – высоты этих двух антенн.

5.3. Передача прямой видимости

С любой системой коммуникаций, сигнал, который получен, будет отличаться от сигнала, который передан, из-за различных ухудшений передачи. Для аналоговых сигналов, эти ухудшения представляют различные случайные модификации, которые ухудшают качество сигнала.

Для цифровых данных, представлены ошибки в символе: двойной 1 преобразован в двойной 0, и наоборот. В этой секции мы исследуем различные ухудшения и комментарий к их эффекту на информационную пропускную способность линии связи. Наше беспокойство в этой книге – с передачей радио потери сигнала, и в этом контексте, самые существенные ухудшения следующие:

- Ослабление и искажение ослабления
- Освободите космическую потерю
- Шум
- Поглощение в атмосфере
- Многопутевой
- Рефракция

Ослабление

Сила сигнала уменьшается с расстоянием по любой среде передачи. Для ведомых СМИ, это сокращение силы, или ослабления, является вообще логарифмическим и таким образом типично выражается как постоянное число децибелов в расстояние единицы. Для неведомых СМИ, ослабление – более сложная функция расстояния и косметики атмосферы. Ослабление представляет три фактора для инженера передачи:

1. Полученный сигнал должен иметь достаточную силу так, чтобы электронная схемотехника в приемнике могла обнаружить и обработать сигнал.
2. Сигнал должен поддерживать уровень достаточно выше чем шум, который будет получен без ошибки.
3. Ослабление больше в верхних частотах, вызывая искажение.

С первыми и вторыми факторами имеют дело вниманием к мощности сигнала и использованию усилителей или ретрансляторов. Для связи пункта-к-пункту, мощность сигнала передатчика должна быть достаточно сильной, чтобы быть полученной понятно, но не настолько сильная, чтобы перегрузить схему передатчика или приемника, который вызвал бы искажение. Вне определенного расстояния, ослабление становится недопустимым образом большим, и ретрансляторы, или усилители используются, чтобы повысить сигнал равномерно. Эти проблемы более сложны, когда есть многократные приемники, где расстояние с передатчика на приемник является переменным.

Третий фактор известен как искажение ослабления. Поскольку ослабление изменяется как функция частоты, полученный сигнал искажен, уменьшая разборчивость. Определенно, частотные компоненты полученного сигнала имеют различные относительные силы чем ча-

стотные компоненты переданного сигнала. Чтобы преодолеть эту проблему, методы доступны для того, чтобы уравнивать ослабление поперек диапазона частот. Один подход состоит в том, чтобы использовать усилители, которые усиливают высокие частоты больше чем более низкие частоты.

Освободите Космическую Потерю

Для любого типа радиосвязи сигнал рассеивается с расстоянием. Поэтому, антенна с неподвижной областью получит меньше мощности сигнала дальше, это – от передающей антенны. Для спутниковой связи это – первичный способ пропадания сигнала. Даже если никакие другие источники ослабления или ухудшения не приняты, переданный сигнал уменьшает по расстоянию, потому что сигнал распространяется по большей и большей области. Эта форма ослабления известна как **свободная космическая потеря**, которая может быть специальной в терминах отношения *занятой* мощности излучения к мощности P_r , полученный антенной или, в децибелах, занимая 10 времен регистрация того отношения. Для идеальной изотропной антенны, свободная космическая потеря

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

где *занятая* = сигнализирует о мощности в передающей антенне P_r = сигнальная мощность в антенне получения X = длина волны поставщика услуг d = дальность распространения между антеннами c = скорость света ($3 \cdot 10^8$ м\с X), где d и X находятся в тех же самых единицах (например, метры). Это может быть переделано как

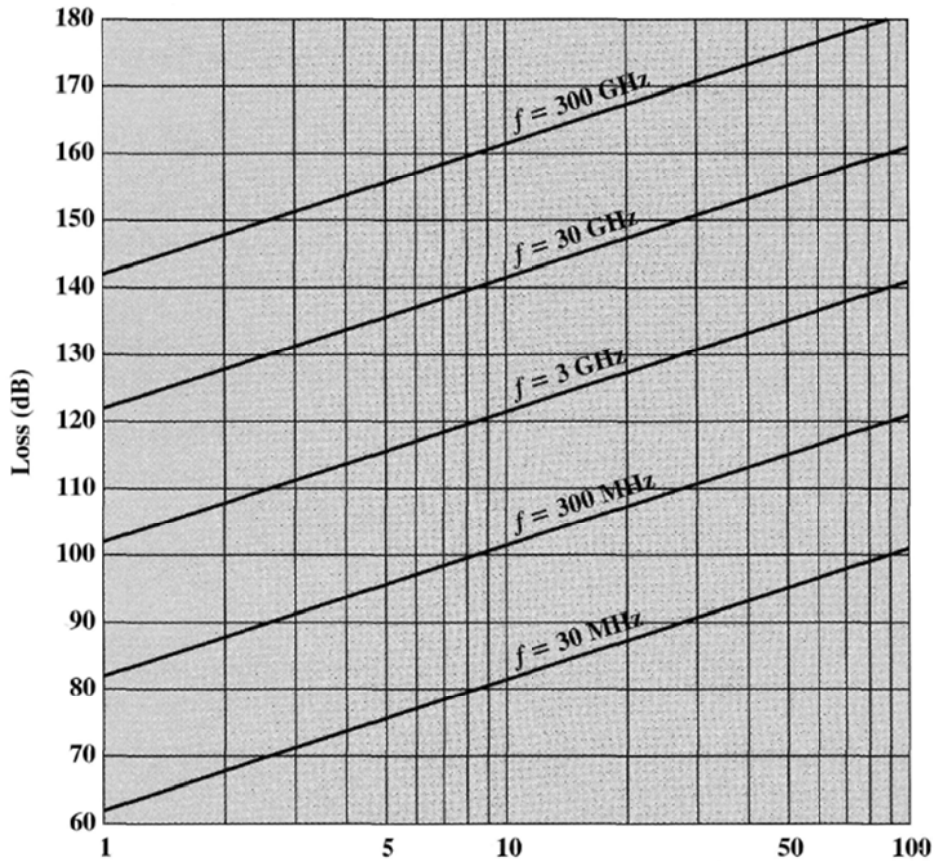
$$\begin{aligned} L_{dB} &= 10 \log \frac{P_i}{P_r} = 20 \log \left(\frac{4\pi d}{\lambda} \right) = -20 \log(\lambda) + 20 \log(d) + 21.98 \text{ dB} \\ &= 20 \log \left(\frac{4\pi f d}{c} \right) = 20 \log(f) + 20 \log(d) - 147.56 \text{ dB} \end{aligned} \quad (5.2)$$

Иллюстрация 5.8 иллюстрирует свободное космическое уравнение потери.

Для других антенн, мы должны принять во внимание выгоду антенны, которая выдает следующее свободное космическое уравнение потери:

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

где G_t = выгода передающей антенны; G_r = выгода антенны получения; A = эффективная область передающей антенны; *Площадь* = эффективная область антенны получения.



Расстояние (км) иллюстрация 5.8 Освобождает Космическую Потерю

Третья фракция получена из второй фракции, используя отношения между выгодой антенны и эффективной областью, определенной в Уравнении (5.1). Мы можем переделать это уравнение как

$$\begin{aligned}
 L_{dB} &= 20 \log (\lambda) + 20 \log (d) - 10 \log (A_r A_t) \\
 &= -20 \log (f) + 20 \log (d) - 10 \log (A_r A_t) + 169.54 \text{ dB}
 \end{aligned}
 \tag{5.3}$$

Таким образом, для тех же самых измерений антенны и разделения, чем дальше длина волны поставщика услуг (понижают несущую частоту/), тем выше – свободные космические потери на трассе. Интересно сравнить Уравнения (5.2) и (5.3). Уравнение (5.2) указывает, что как частотные увеличения, свободная космическая потеря также увеличивается, который предложил бы, что в верхних частотах, потери становятся более обременительными. Однако, Уравнение (5.3) показы, которым мы можем легко дать компенсацию за эту увеличенную потерю с прибылью антенны. Фактически, есть общий коэффициент усиления в верхних частотах, другие факторы, остающиеся постоянным. Уравнение (5.2) показы, что на неподвижном расстоянии увеличение частоты приводит к увеличенной потере, измеренной 20 регистрациями (f). Однако, если мы принимаем во внимание выгоду антенны, и устанавливаем площадь ан-

тенны, тогда изменение в потере измерено-20 регистрациями (/); то есть, есть фактически уменьшение в потере в верхних частотах.

Шум

Для любого случая передачи данных, полученный сигнал будет состоять из переданного сигнала, измененного различными искажениями, наложенными системой передачи, плюс дополнительные ложные сигналы, которые вставлены где-нибудь между передачей и приемом. Эти ложные сигналы упоминаются как шум. Шум – главный ограничивающий фактор в работе системы коммуникаций. Шум может быть разделен на четыре категории:

- Тепловые помехи
- Интермодуляционный шум
- Переходные помехи
- Импульсные помехи

Тепловые помехи происходят из-за теплового возмущения электронов. Это присутствует во всех электронных устройствах и средствах связи и – функция температуры. Тепловые помехи однородно распределены поперек частотного спектра и следовательно часто упоминаются как белый шум. Тепловые помехи не могут быть устранены и поэтому помещают верхнюю привязанную работу системы коммуникаций. Из-за слабости сигнала, полученного спутниковыми земными станциями, тепловые помехи особенно существенны для спутниковой связи.

Количество тепловых помех, которые будут найдены в ширине полосы 1 Гц в любом устройстве или проводе $N_0 = kT$ (W/Hz).

где N_0 = шум включают плотность в ваттах в 1 Гц ширины полосы k = Boltzmann's постоянный = 1.3803×10^{-23} J/K T = температура, в kelvins (абсолютная температура)

Шум, как предполагают, является независимым частоты. Таким образом тепловые помехи в подарке ватт в ширине полосы Герц B могут быть выражены как

$$N = kTB$$

или, в ваттах децибела,

$$\begin{aligned} N &= 10 \log k + 10 \log T + 10 \log B \\ &= -228.6 \text{ dBW} + 10 \log T + 10 \log B \end{aligned}$$

Когда сигналы в различных частотах разделяют ту же самую среду передачи, результатом может быть **интермодуляционный шум**. Интермодуляционный шум производит сигналы в частоте, которая является суммой или различием двух оригинальных частот или односторонних

магазинов тех частот. Например, смешивание сигналов в частотах f_1 и f_2 мог бы произвести энергию в частоте $f_x + f_2$. Этот полученный сигнал мог столкнуться с намеренным сигналом в частоте $f_1 + f_2$.

Интермодуляционный шум произведен, когда есть немного нелинейности в передатчике, приемнике, или прошедшей системе передачи. Обычно, эти компоненты ведут себя как линейные системы; то есть, продукция равна временам ввода постоянное. В нелинейной системе, продукция – более сложная функция входа. Такая нелинейность может быть вызвана составляющим сбоем, использование чрезмерной мощности сигнала, или только природа используемых усилителей. Именно под этими обстоятельствами сумма и сроки разностной частоты происходят.

Переходные помехи были испытаны любым, кто, используя телефон, был в состоянии услышать другую беседу; это – нежелательное сцепление между трактами передачи сигналов. Это может произойти электрическим сцеплением между соседними витыми парами или, редко, линии коаксиального кабеля, несущие многократные сигналы. Переходные помехи могут также произойти, когда ложные сигналы приняты микроволновыми антеннами; хотя очень направлено *antennas* используются, микроволновая энергия действительно распространялась в течение распространения. Как правило, переходные помехи имеют тот же самый порядок величины как, или меньше чем, тепловые помехи. Однако, в неимеющих лицензию диапазонах ИЗМА, переходные помехи часто доминируют.

Все типы шума, обсужденного пока имеют разумно предсказуемые и относительно постоянные величины. Таким образом возможно проектировать систему передачи, чтобы справиться с ними. **Импульсные помехи**, однако, непрерывны, состоя из нерегулярных импульсов или шумовых шипов короткой продолжительности и относительно высокой амплитуды. Это сгенерировано от разнообразия причин, включая внешние электромагнитные беспорядки, типа молнии, и ошибок и недостатков в системе коммуникаций.

Импульсные помехи – вообще только незначительное раздражение для аналоговых данных. Например, телефонная связь может быть развращена короткими щелчками и потрескиваниями без потери разборчивости. Однако, импульсные помехи – первичный источник ошибки в передаче цифровой информации. Например, острый шип энергии 0.01 с продолжительностей не разрушил бы никакой речевой информации, но смоеет приблизительно 560 битов данных, передаваемых в 56 кбит/с.

Выражение E_b/N_0

Глава 2 представляла отношение сигнал-шум (ОТНОШЕНИЕ СИГНАЛ – ШУМ). Есть параметр, связанный с ОТНОШЕНИЕМ СИГНАЛ – ШУМ, которое более удобно для того, чтобы определить скоро-

сти цифровых данных и коэффициенты ошибок, и это – стандартная качественная мера для работы системы цифровой связи. Параметр – отношение сигналной энергии в бит к шумовой плотности мощности в Герц, E_b/N_0 . Рассмотрите сигнал, цифровой или аналоговый, который содержит двоичные цифровые данные, переданные в определенной скорости передачи данных R . Напоминая то, что 1 ватт = 1 J/s, энергия в бит в сигнале дается $E_b = STb$, где S – сигналная мощность и Tb , – время, требуемое послать один бит. Скорость передачи данных R – только $R = 1/T$. Таким образом

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

или, в примечании децибела,

$$\begin{aligned} \left(\frac{E_b}{N_0}\right)_{\text{dB}} &= S_{\text{dBW}} - 10 \log R - 10 \log k - 10 \log T \\ &= S_{\text{dBW}} - 10 \log R + 228.6 \text{ dBW} - 10 \log T \end{aligned}$$

E_b/N_0 отношения важен, потому что коэффициент ошибок по элементам для цифровых данных – (уменьшающаяся) функция этого отношения. Учитывая ценность E_b/N_0 должен был достигнуть желательного коэффициента ошибок, параметры в предыдущей формуле могут быть отобраны. Отметьте, что как скорость передачи данных R увеличения, переданная сигналная мощность, относительно шума, должна увеличиться, чтобы поддержать необходимый E_b/N_0 .

Позвольте нам пробовать схватить этот результат интуитивно, рассматривая снова иллюстрацию 2.9. Сигнал здесь является цифровым, но рассуждение было бы тем же самым для аналогового сигнала. В нескольких экземплярах, шум достаточен, чтобы изменить ценность немного. Если бы скорость передачи данных была удвоена, то биты более сильно упаковались бы вместе, и тот же самый проход шума мог бы разрушить два бита. Таким образом, для постоянной сигналной и шумовой силы, увеличение скорости передачи данных увеличивает коэффициент ошибок.

Преимущество E_b/N_0 по ОТНОШЕНИЮ СИГНАЛ – ШУМ состоит в том, что последнее количество зависит от ширины полосы.

Мы можем связать E_b/N_0 с ОТНОШЕНИЕМ СИГНАЛ – ШУМ следующим образом. Мы имеем

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

Параметр N_0 – шумовая плотность мощности в ваттах/герц. Следовательно, шум в сигнале с шириной полосы *Бритиш Телеком* – $N = N_0 B$. Замена, мы имеем

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

Другая формулировка интереса имеет отношение со спектральной эффективностью E_b/N_0 . Отзыв, от Главы 2, результат Шаннона, что максимальная вместимость канала, в бит в секунду, повинует уравнению

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

где C – вместимость канала в бит в секунду, и B – ширина полосы канала в Герц. Это может быть переписано как

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

Используя Уравнение (5.4), и приравнивающий *Бритиш Телеком* C с B и R с C , мы имеем

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

Это – полезная формула, которая связывает достижимую спектральную *индикацию перегрузки* эффективности B с E_b/N_0 .

Поглощение в атмосфере

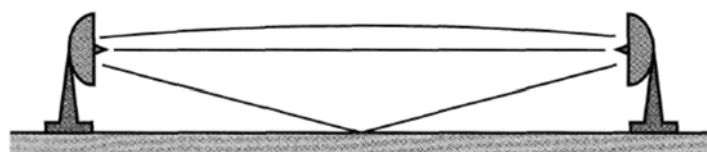
Дополнительная потеря между передачей и получением антенн – поглощение в атмосфере. Водный пар и кислород вносят вклад больше всего в ослабление. Максимальное затухание происходит около 22 ГГц из-за водного пара. В частотах ниже 15 ГГц, ослабление – меньше. Присутствие кислорода приводит к поглотительному пику около 60 ГГц, но вносит вклад меньше в частотах ниже 30 ГГц. Дождь и туман (приостановленные водные капельки) рассеивание причины радиоволн, которое приводит к ослаблению. Это может быть главной причиной пропадания сигнала. Таким образом, в областях существенного осадения, или длины пути должны быть сохранены коротким, или более низкие диапазоны частот должны использоваться.

Многопутевой

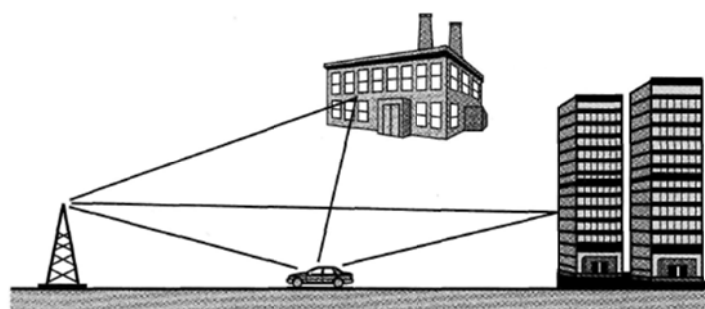
Для беспроводных средств обслуживания, где есть относительно свободный выбор того, где антенны должны быть расположены, они могут быть размещены так, чтобы, если есть количество, поблизости мешающее препятствия, есть прямая трасса в пределах прямой видимости с передатчика на приемник. Это вообще имеет место для многих спутниковых средств обслуживания и для микроволновой печи пункта-к-пункту. В других случаях, типа мобильной телефонной связи, в изобилии есть препятствия. Сигнал может быть отражен такими пре-

пятствиями так, чтобы многократные копии сигнала с изменением задержек могли быть получены. Фактически, в чрезвычайных случаях, может быть количество прямой сигнал. В зависимости от различий в длинах пути прямых и отраженных волн, составной сигнал может быть или больший или меньший чем прямой сигнал. Укреплением и отменой сигнала, следующего из сигнала после многократных трактов можно управлять для коммуникации между неподвижными, хорошо-расположенными антеннами, и между спутниками и установил станции основания. Одно исключение – то, когда тракт идет поперек воды, где ветер держит рефлексивную поверхность воды в движении. Для мобильной телефонной связи и коммуникации к антеннам, которые хорошо не расположены, многопутевые рассмотрения могут быть главными.

Иллюстрация 5.9 иллюстрирует в сроках генерала типы многопутевой интерференции, типичной в земной, установленной микроволновой печи и в мобильных коммуникациях. Для неподвижной микроволновой печи, в дополнение к прямому лучу обзора, сигнал может следовать за кривым трактом через атмосферу из-за рефракции, и сигнал может также размышлять от основания. Для мобильных коммуникаций, структуры и топографические особенности обеспечивают поверхности отражения.



(a) Микроволновая линия вида



(b) Подвижная радиостанция

Иллюстрация 5.9 Примеры Многопутевой Интерференции

Рефракция

Радиоволны преломляются (или согнуты), когда они размножаются через атмосферу. Рефракция вызвана изменениями в скорости сигнала с высотой или другими пространственными изменениями в атмосферных условиях. Обычно, скорость сигнала увеличивается с высотой, заставляя

радиоволны согнуться вниз. Однако, **при случае**, метеорологические условия могут привести к изменениям в скорости с высотой, которые отличаются значительно от типичных изменений. Это может привести к ситуации, в которой только часть фракции или количества волны прямой видимости достигает антенны получения.

1. Особенности систем подвижной радиосвязи

1.1. Доплеровское расширение спектра сигналов

Средства радиосвязи применяются в условиях движущегося со скоростью до 250...500 км/ч современного автомобильного и железнодорожного транспорта. При этом доплеровские смещения частоты взаимно перемещающихся приемников и передатчиков средств радиосвязи превышают нестабильность и точность настройки несущих частот радиосигналов. Кроме этого неравномерность скорости движения объектов приводит к непостоянству смещения доплеровского частот и, следовательно, к появлению шума из-за случайной частотной модуляции.

Доплеровское смещение частоты

$$f_d = v/\lambda = v/(c/f) = v \cdot f/c,$$

где: λ и f – длина волны и частота передатчика; v – скорость перемещения приемника относительно передатчика средств радиосвязи; $c = 3 \cdot 10^8$ м/с – скорость распространения света.

При частоте 2 ГГц и скорости 500 км/ч доплеровское смещение частоты составляет около 1000 Гц. В работе [11] показано, что ширина спектра шума случайной частотной модуляции равна приблизительно удвоенному доплеровскому рассеянию, т. е. 2000 Гц. Таким образом, спектр доплеровского шума попадает в полосу модулирующих звуковых частот и является помехой радиоканалов подвижной связи.

Доплеровские смещения частоты можно рассматривать, как эффект временной декорреляции сигналов. Время когерентности, в пределах которого коэффициент корреляции значений огибающей сигнала не менее 0,9, обратно пропорционально доплеровскому смещению частоты, т. е. определяется как

$$C_T = 1/f_d$$

и для приведенного примера равно 1 мс.

Применяют следующие способы борьбы с проявлениями в СПР доплеровского эффекта:

- синхронизацию несущих всех средств системной радиосвязи, которая в системах с TDMA является основным принципом обеспечения правильного функционирования систем;

- расширение полосы канала относительно полос узкополосных аналоговых систем, что лежит в основе функционирования систем с TDMA и CDMA.

Таким образом, в изученных нами современных системах СБД, МСБТ и СПД с TDMA и CDMA меры борьбы с проявлениями доплеровского эффекта за счет движения объектов с большими скоростями автоматически обеспечиваются. В сотовых и транкинговых системах, изучаемых в этой части курса, они должны обеспечиваться обязательно, так как эти системы предназначены для обслуживания подвижных абонентов.

1.2. Многолучевое распространение радиоволн

Используемые в сотовой связи с подвижными объектами (ПО) дециметровые радиоволны слабо огибают препятствия, распространяются в основном по прямой и испытывают многочисленные отражения от окружающих объектов и подстилающей поверхности (рис. 1.1). Следствиями такого многолучевого распространения является более быстрое, чем в свободном пространстве, убывание интенсивности принимаемого сигнала с расстоянием и интерференция радиоволн.

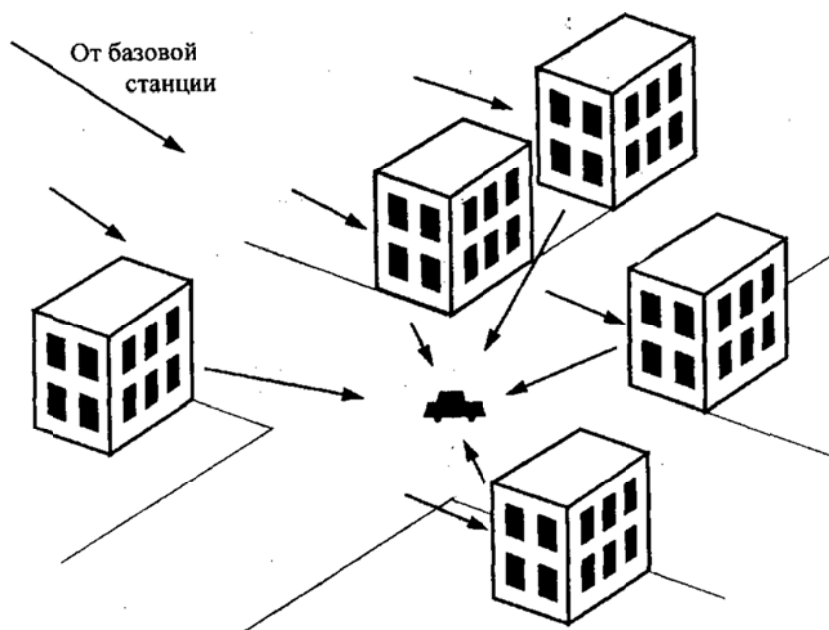


Рис. 1.1

При сложении нескольких сигналов, прошедших по разным путям и имеющих в точке приема в общем случае различные фазы, результирующий сигнал может быть как несколько выше среднего уровня, так и заметно ниже. Причем провалы, или замирания сигнала, образующиеся при взаимной компенсации сигналов вследствие неблагоприятного сочетания их фаз и амплитуд, могут быть достаточно глубокими.

При этом возникают также искажения результирующего сигнала типа межсимвольной интерференции в том случае, когда сигналы с соизмеримыми амплитудами настолько отличаются по разности хода, что символы одного сигнала «налезают» на соседние символы другого.

Колебания уровня (замирания) принимаемого сигнала практически всегда имеют две составляющие – быструю и медленную.

Быстрые замирания, являющиеся прямым следствием многолучевого распространения, описываются релеевским законом распределения, и потому их называют релеевскими замираниями и моделируют законом Релея-Райса [11]

$$W(r) = \left(\frac{r}{\sigma^2}\right) \cdot \exp\left(-\frac{r^2}{2\sigma^2}\right) \cdot \exp(-Q_0) \cdot I_0\left(\frac{r}{\sigma} \cdot \sqrt{2Q_0}\right), r \geq 0$$

где Q_0 – отношение мощности прямого (нефлуктуирующего) сигнала к средней мощности флюктуаций, σ^2 – средняя мощность или дисперсия флюктуаций сигнала, $I_0()$ – модифицированная функция первого рода нулевого порядка. Если прямой сигнал отсутствует, а принимаются только отраженные сигналы, то $Q_0 = 0$ и плотность распределения вероятностей описывается законом Релея [11]

$$W(r) = \left(\frac{r}{\sigma^2}\right) \cdot \exp\left(-\frac{r^2}{2\sigma^2}\right), r \geq 0$$

Диапазон изменений уровня сигнала при быстрых замираниях может достигать 40 дБ, из которых примерно 10 дБ – превышение над средним уровнем и 30 дБ – провалы ниже среднего уровня, причем более глубокие провалы встречаются реже, чем менее глубокие. При неподвижном абонентском аппарате интенсивность принимаемого сигнала, естественно, не меняется. При перемещении подвижной станции периодичность флюктуации в пространстве составляет около полуволны, т. е. порядка 10...15 см в линейной мере. Период флюктуации во времени зависит от скорости перемещения подвижной станции: например, при скорости 50 км/ч период флюктуации составляет около 10 мс, а при 100 км/ч – около 5 мс.

Медленные замирания обусловлены изменением условий затенения при перемещении подвижной станции и подчиняются логарифмически

$$W(r) = \frac{1}{\sqrt{2\pi} \cdot \sigma \cdot r} \cdot \exp\left[-\frac{(\ln r - a)^2}{2\sigma^2}\right], r > 0$$

нормальному закону распределения [11]

где r – амплитуда сигнала, a и σ^2 – параметры, характеризующие средний уровень амплитуды и глубину замираний, соответственно.

Интенсивность медленных флуктуации не превышает 5...10 дБ, а их периодичность соответствует перемещению подвижной станции на десятки метров. Фактически медленные замирания представляют собой изменение среднего уровня сигнала при перемещении подвижной станции, на которые накладываются быстрые замирания вследствие многолучевого распространения.

Основную неприятность при связи с подвижными объектами составляют быстрые замирания, поскольку они бывают достаточно глубокими, и при этом отношение сигнал/шум падает настолько сильно, что полезная информация может существенно искажаться шумами, вплоть до полной ее потери. Кроме этого методы борьбы с быстрыми замираниями радиоволн при связи с ПО эффективно устраняют и проявления медленных замираний.

1.3. Способы борьбы с быстрыми замираниями радиосигналов

Для борьбы с быстрыми замираниями радиоволн при связи с ПО используются следующие основные методы:

- разнесенный прием;
- скачки по частоте;
- эквалайзинг.

Метод разнесенного приема (английский термин diversity – разнесение) заключается в совместном использовании нескольких сигналов, различающихся (разнесенных) по какому-либо параметру. Причем разнесение должно выбираться таким образом, чтобы вероятность одновременных замираний всех используемых сигналов была много меньше, чем какого-либо одного из них. Иными словами, эффективность разнесенного приема тем выше, чем менее коррелированы замирания в составляющих сигналах.

Возможны пять вариантов разнесенного приема [11]:

- - с разнесением во времени (time diversity), при этом используются сигналы, сдвинутые во времени один относительно другого; этот метод сравнительно легко реализуем лишь в цифровой форме, и улучшение качества приема осуществляется в ущерб пропускной способности канала связи;
- - с разнесением по частоте (frequency diversity), при этом используются сигналы, передаваемые на нескольких частотах, поэтому «платой» является расширение используемой полосы частот;
- с разнесением по углу или по направлению (angle diversity или direction diversity), при этом прием производится на несколько ан-

тени с рассогласованными (не полностью перекрывающимися) диаграммами направленности, в этом случае сигналы с выходов разных антенн коррелированы тем слабее, чем меньше перекрытие диаграмм направленности, но при этом одновременно падает и эффективность приема (интенсивность принимаемого сигнала), по крайней мере для всех антенн, кроме одной;

- с разнесением по поляризации (polarization diversity), когда, например, две антенны принимают сигналы двух взаимно ортогональных поляризаций (практического значения этот вариант не имеет, поскольку в диапазоне СВЧ замирания на разных поляризациях сильно коррелированы);
- с разносом в пространстве (space diversity), т. е. с приемом сигналов на несколько пространственно разнесенных антенн, этот метод находит практическое применение и именно он обычно имеется в виду, когда говорят о разнесенном приеме.

Для метода пространственного разнесения очевидно, что выигрыш тем больше, чем больше число используемых антенн, однако при этом возрастает и сложность технического решения. Поэтому практическое применение находит простейшая система с двумя приемными антеннами базовых станций. В подвижных станциях разнесенный прием не применяют. С ростом расстояния между антеннами корреляция между флуктуациями уровня принимаемых ими сигналов падает, и чем больше разнос антенн, тем выше эффективность разнесенного приема. Но при этом возрастает и сложность технической реализации, поэтому практически разнос берется минимально возможным, при котором разнесенный прием уже достаточно эффективен. Реально с учетом, как аналитических оценок, так и эмпирических данных разнос обычно составляет около десятка длин волн, т. е. порядка нескольких метров.

Объединение сигналов с выходов двух антенн возможно:

- использованием одного более сильного из двух сигналов;
- додетекторное когерентное суммирование обоих сигналов;
- последетекторное суммирование сигналов с равными весами или со взвешиванием, обеспечивающим получение максимума отношения сигнал/шум.

Реализации метода скачков по частоте состоят в том, что несущая частота для каждого физического канала периодически изменяется. Поскольку релейские замирания являются частотно-селективными, то, если при работе на некоторой частоте имело место замирание, при изменении рабочей частоты на 0...300 кГц замирания с большой вероятностью не будет. Следовательно, при достаточно частых изменениях частоты существенно снижается вероятность длительных замираний и

групповых ошибок, а с одиночными ошибками можно успешно бороться при помощи помехоустойчивого канального кодирования.

Различают медленные и быстрые скачки по частоте. При медленных скачках период изменения частоты много больше длительности символа передаваемого сообщения, а при быстрых скачках – много меньше длительности символа.

Изменение частоты в пределах доступного диапазона может быть как регулярным (циклическим), так и нерегулярным (псевдослучайным). Режим работы со скачками по частоте не является обязательным и назначается по команде с центра коммутации радиосистемы.

В узкополосных TDMA-системах для компенсации межсимвольных искажений используется метод эквалайзинга. Термин *эквалайзинг* заимствован из английского языка (*equalizing* – буквально *выравнивание*) и имеет в данном случае смысл компенсации той разности хода между составляющими лучами при многолучевом распространении, которая приводит к межсимвольной интерференции. Эквалайзер по своей сути – это адаптивный фильтр, настраиваемый таким образом, чтобы сигнал на его выходе был в возможно большей степени очищен от межсимвольных искажений, содержащихся во входном сигнале.

Простейшая реализация эквалайзера (рис. 1.2) – трансверсальный фильтр. Покажем на простом примере, что такая схема может в некоторых ситуациях, существенно ослабить межсимвольные искажения. Предположим, что входной сигнал эквалайзера состоит из основного сигнала – некоторой последовательности однобитовых символов (единиц и нулей, первый график на рис. 1.3) и его копии, ослабленной в три раза и сдвинутой во времени на длительность t одного символа (второй график на рис. 1.3).

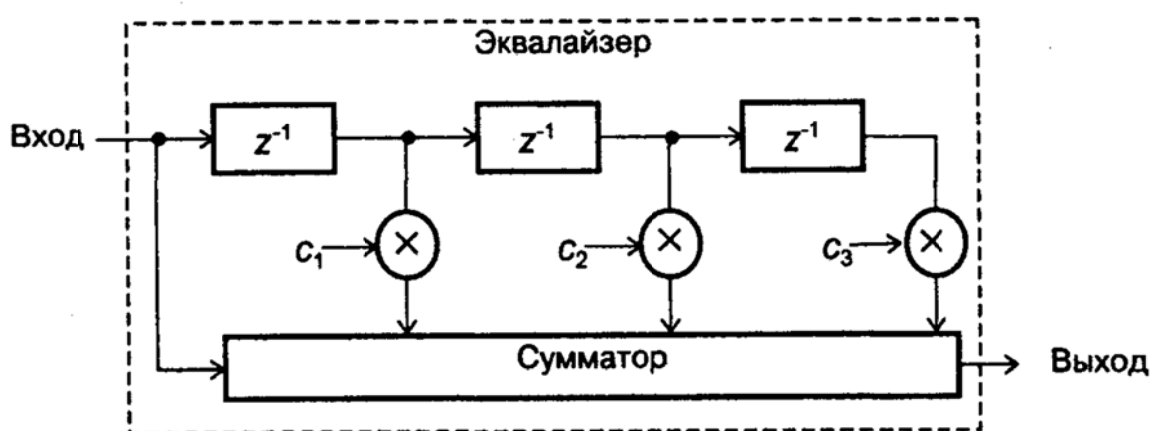


Рис. 1.2



Рис. 1.3

Если время дискретизации линии задержки фильтра равно τ , а значение коэффициента в первом отводе $C_1 = -1/3$, то при сложении входного сигнала и сигнала с первого отвода получим, что основной сигнал (первая составляющая входного сигнала) остается без изменений, а вторая составляющая входного сигнала компенсируется первой составляющей, задержанной на τ (сигнала с первого отвода линии задержки), вторая составляющая задержанного сигнала дает копию основного, но ослабленную уже в девять раз, задержанную на 2τ и с обратным знаком. Если во втором отводе линии задержки коэффициент $C_2 = 1/9$, то при сложении входного и двух задержанных сигналов получим неизменный основной сигнал и его копию, задержанную на 3τ и ослабленную в 27 раз. Таким образом, в рассматриваемом примере добавление каждого следующего элемента линии задержки с соответствующим значением коэффициента C , приводит к ослаблению искажающего сигнала втрое и к дополнительной задержке его во времени на τ .

В реальной жизни дело обстоит сложнее, чем в описанном примере, и число лучей может быть больше двух, и задержки едва ли будут кратны дискрету линии задержки, и амплитуды составляющих сигналов, так же как и их число и задержки, не будут заранее известны. Кроме того, при перемещении абонентского аппарата вся эта картина непрерывно изменяется. Поэтому настройка фильтра должна производиться адаптивно, в соответствии с конкретно складывающейся ситуацией с использованием обучающей последовательности C , передаваемой в каждом слоте.

Эквалайзер на основе трансверсального фильтра является линейным и не устраняет большие искажения сигналов. При этом используют более совершенные нелинейные эквалайзеры.

ГЛАВА 7. МОДЕЛЬ ВЗАИМОДЕЙСТВИЯ ОТКРЫТЫХ СИСТЕМ И ПРОБЛЕМЫ СТАНДАРТИЗАЦИИ СЕТЕЙ

Универсальный тезис о пользе стандартизации, справедливый для всех отраслей, в компьютерных сетях приобретает особое значение. Суть сети – это соединение разного оборудования, а значит, проблема совместимости является одной из наиболее острых. Без принятия всеми производителями общепринятых правил построения оборудования прогресс в деле «строительства» сетей был бы невозможен. Поэтому все развитие компьютерной отрасли в конечном счете отражено в стандартах – любая новая технология только тогда приобретает «законный» статус, когда ее содержание закрепляется в соответствующем стандарте.

В компьютерных сетях идеологической основой стандартизации является многоуровневый подход к разработке средств сетевого взаимодействия. Именно на основе этого подхода была разработана стандартная семиуровневая модель взаимодействия открытых систем, ставшая своего рода универсальным языком сетевых специалистов.

7.1. Многоуровневый подход. Протокол. Интерфейс. Стек протоколов

Организация взаимодействия между устройствами в сети является сложной задачей. Как известно, для решения сложных задач используется универсальный прием – декомпозиция, то есть разбиение одной сложной задачи на несколько более простых задач-модулей. Процедура декомпозиции включает в себя четкое определение функций каждого модуля, решающего отдельную задачу, и интерфейсов между ними. В результате достигается логическое упрощение задачи, а кроме того, появляется возможность модификации отдельных модулей без изменения остальной части системы.

При декомпозиции часто используют многоуровневый подход. Он заключается в следующем. Все множество модулей разбивают на уровни. Уровни образуют иерархию, то есть имеются вышележащие и нижележащие уровни. Множество модулей, составляющих каждый уровень, сформировано таким образом, что для выполнения своих задач они обращаются с запросами только к модулям непосредственно примыкающего нижележащего уровня. С другой стороны, результаты работы всех модулей, принадлежащих некоторому уровню, могут быть переданы

только модулям соседнего вышележащего уровня. Такая иерархическая декомпозиция задачи предполагает четкое определение функции каждого уровня и интерфейсов между уровнями. Интерфейс определяет набор функций, которые нижележащий уровень предоставляет вышележащему. В результате иерархической декомпозиции достигается относительная независимость уровней, а значит, и возможность их легкой замены.

Средства сетевого взаимодействия, конечно, тоже могут быть представлены в виде иерархически организованного множества модулей. При этом модули нижнего уровня могут, например, решать все вопросы, связанные с надежной передачей электрических сигналов между двумя соседними узлами. Модули более высокого уровня организуют транспортировку сообщений в пределах всей сети, пользуясь для этого средствами упомянутого нижележащего уровня. А на верхнем уровне работают модули, предоставляющие пользователям доступ к различным службам – файловой, печати и т. п. Конечно, это только один из множества возможных вариантов деления общей задачи организации сетевого взаимодействия на частные подзадачи.

Многоуровневое представление средств сетевого взаимодействия имеет свою специфику, связанную с тем, что в процессе обмена сообщениями участвуют две машины, то есть в данном случае необходимо организовать согласованную работу двух «иерархий». При передаче сообщений оба участника сетевого обмена должны принять множество соглашений. Например, они должны согласовать уровни и форму электрических сигналов, способ определения длины сообщений, договориться о методах контроля достоверности и т. п. Другими словами, соглашения должны быть приняты для всех уровней, начиная от самого низкого – уровня передачи битов – до самого высокого, реализующего сервис для пользователей сети.

На рис. 7.1 показана модель взаимодействия двух узлов. С каждой стороны средства взаимодействия представлены четырьмя уровнями. Процедура взаимодействия этих двух узлов может быть описана в виде набора правил взаимодействия каждой пары соответствующих уровней обеих участвующих сторон. нормализованные правила, определяющие последовательность и формат сообщений, которыми обмениваются сетевые компоненты, лежащие на одном уровне, но в разных узлах, называются *протоколом*.

Модули, реализующие протоколы соседних уровней и находящиеся в одном узле, также 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 со-

седнему уровню. В сущности, протокол и интерфейс выражают одно и то же понятие, но традиционно в сетях за ними закрепил разные области действия: протоколы определяют правила взаимодействия модулей одного уровня в разных узлах, а интерфейсы – модулей соседних уровней одном узле.

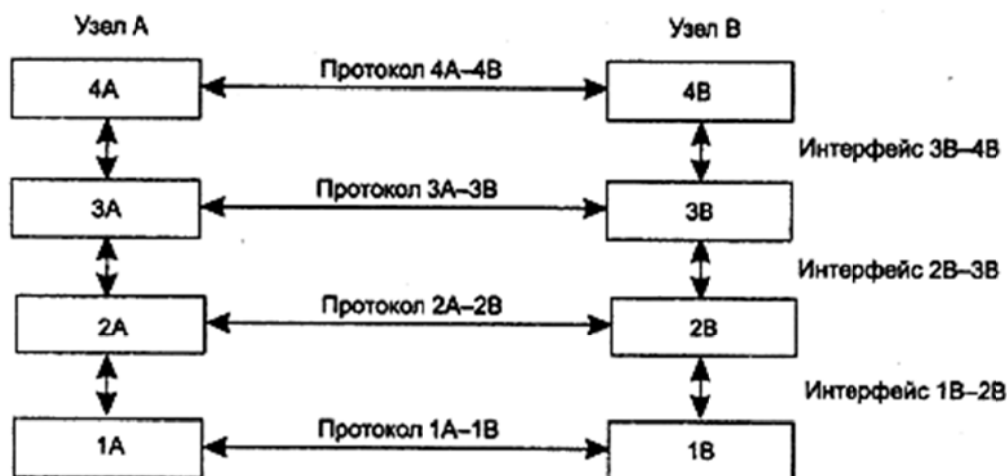


Рис. 7.1. Взаимодействие двух узлов

Средства каждого уровня должны обрабатывать, 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.

Понятно, что один и тот же алгоритм может быть запрограммирован с разной степенью эффективности. Точно так же и протокол может иметь несколько программных реализаций. Именно поэтому при сравнении протоколов следует учитывать не только логику их работы, но и качество программных решений. Более того, на эффективность взаимодействия устройств в сети влияет качество всей совокупности протоколов, составляющих стек, в частности, насколько рационально распределены функции между протоколами разных уровней и насколько хорошо определены интерфейсы между ними.

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 коммуника-

ционные устройства. В зависимости от типа устройства в нем должны быть встроенные средства, реализующие тот или иной набор протоколов.

7.2. Модель OSI

Из того, что протокол является соглашением, принятым двумя взаимодействующими объектами, в данном случае двумя работающими в сети компьютерами, совсем не следует, что он обязательно является стандартным. Но на практике при реализации сетей стремятся использовать стандартные протоколы. Это могут быть фирменные, национальные или международные стандарты.

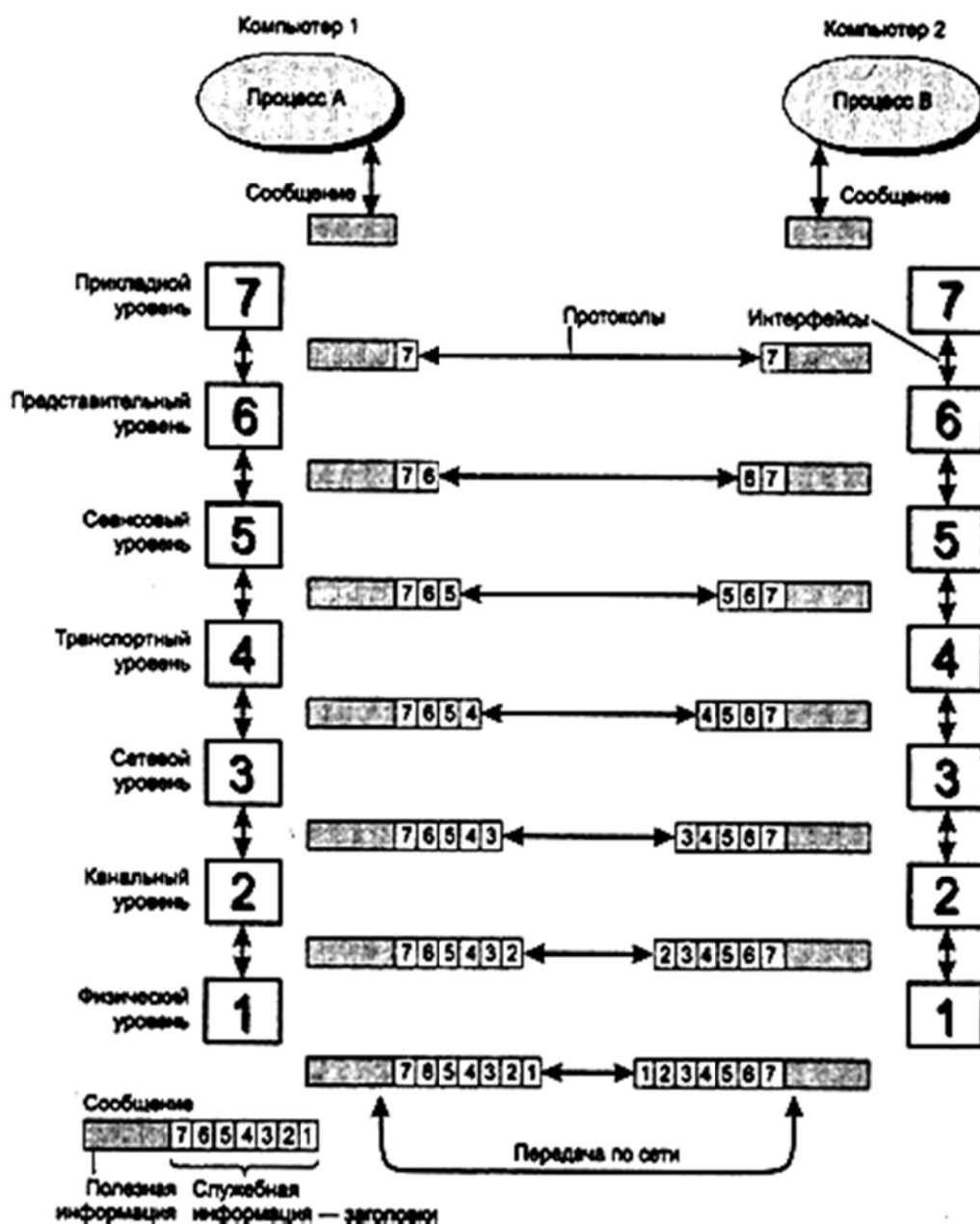


Рис. 3.2. Модель взаимодействия открытых систем ISO/OSI

В начале 80-х годов ряд международных организаций по стандартизации – ISO, ITU-T и некоторые другие – разработали модель, которая сыграла значительную роль в развитии сетей. Эта модель называется моделью взаимодействия открытых систем (Open System Interconnection, OSI) или моделью OSI. Модель OSI определяет различные уровни взаимодействия систем, дает им стандартные имена и указывает, какие функции должен выполнять каждый уровень. Модель OSI была разработана на основании большого опыта, полученного при создании компьютерных сетей, в основном глобальных, в 70-е годы. Полное описание этой модели занимает более 1000 страниц текста.

В модели OSI (рис. 7.2) средства взаимодействия делятся на семь уровней: прикладной, представительный, сеансовый, транспортный, сетевой, канальный и физический. Каждый уровень имеет дело с одним определенным аспектом взаимодействия сетевых устройств.

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.

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 службу; 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.

Итак, пусть приложение обращается с запросом к прикладному уровню, например к файловой службе. На основании этого запроса программное обеспечение прикладного уровня формирует сообщение стандартного формата. Обычное сообщение состоит из заголовка и поля данных. Заголовок содержит служебную информацию, которую необходимо передать через сеть прикладному уровню машины-адресата, чтобы сообщить ему, какую работу надо выполнить. В нашем случае заголовок, очевидно, должен содержать информацию о месте нахождения файла и о типе операции, которую необходимо над ним выполнить. Поле данных сообщения может быть пустым или содержать какие-либо данные, например те, которые необходимо записать в удаленный файл. Но для того чтобы доставить эту информацию по назначению, предстоит решить еще много задач, ответственное за которые несут нижележащие уровни.

После формирования сообщения прикладной уровень направляет его вниз по стеку представителю уровня. Протокол представительного уровня на основании информации, полученной из заголовка прикладного уровня, выполняет требуемые действия и добавляет к сообщению собственную служебную информацию – заголовок представительного уровня, в котором содержатся указания для протокола представительного уровня машины-адресата. Полученное в результате сообщение передается вниз сеансовому уровню, который в свою очередь добавляет свой заголовок, и т. д. (Некоторые реализации протоколов помещают служебную информацию не только в начале сообщения в виде заголовка, но и в конце, в виде так называемого «концевика».) Наконец, сообщение достигает нижнего, физического уровня, который собственно и передает его по линиям связи машине-адресату. К этому моменту сообщение «обрастает» заголовками всех уровней (рис. 7.3).

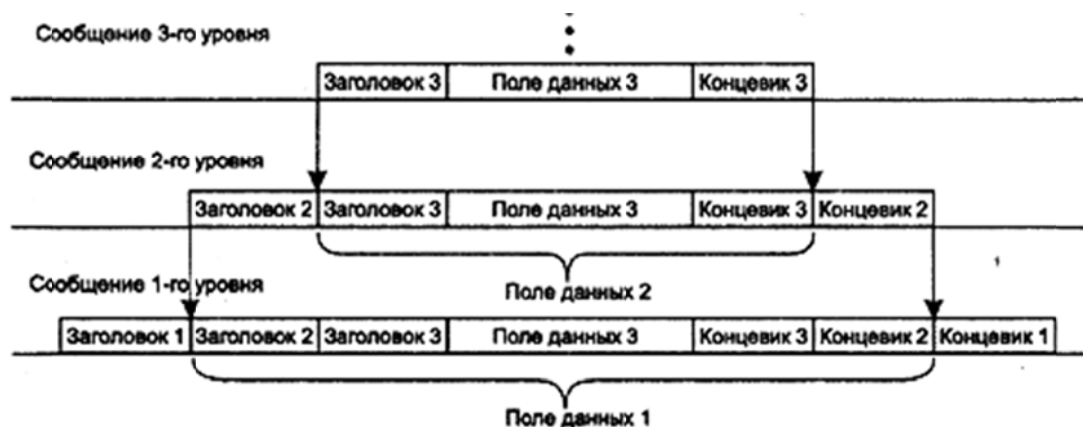


Рис. 7.3. Вложенность сообщений различных уровней

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.

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).

В модели OSI различаются два основных типа протоколов. В протоколах *с установлением соединения (connection-oriented)* перед обменом данными отправитель и получатель должны сначала установить соединение и, возможно, выбрать некоторые параметры протокола, которые они будут использовать при обмене данными. После завершения диалога они должны разорвать это соединение. Телефон – это пример взаимодействия, основанного на установлении соединения.

Вторая группа протоколов – *протоколы без предварительного установления соединения (connectionless)*. Такие протоколы называются также дейтаграммными протоколами. Отправитель просто передает сообщение, когда оно готово. Опускание письма в почтовый ящик – это пример связи без предварительного установления соединения. При взаимодействии компьютеров используются протоколы обоих типов.

7.3. Уровни модели OSI.

Физический уровень

Физический уровень (Physical layer) имеет дело с передачей битов по физическим каналам связи, таким, например, как коаксиальный кабель, витая пара, оптоволоконный кабель или цифровой территориальный канал. К этому уровню имеют отношение характеристики физических сред передачи данных, такие как полоса пропускания, помехозащищенность, волновое сопротивление и другие. На этом же уровне определяются характеристики электрических сигналов, передающих дискретную информацию, например, крутизна фронтов импульсов, уровни напряжения или тока передаваемого сигнала, тип кодирования, скорость передачи сигналов. Кроме этого, здесь стандартизируются типы разъемов и назначение каждого контакта.

Функции физического уровня реализуются во всех устройствах, подключенных к сети. Со стороны компьютера функции физического уровня выполняются сетевым адаптером или последовательным портом.

Примером протокола физического уровня может служить спецификация 10Base-T технологии Ethernet, которая определяет в качестве используемого кабеля неэкранированную витую пару категории 3 с волновым сопротивлением 100 Ом, разъем RJ-45, максимальную длину физического сегмента 100 метров, манчестерский код для представления данных в кабеле, а также некоторые другие характеристики среды и электрических сигналов.

Канальный уровень

На физическом уровне просто пересылаются биты. При этом не учитывается, что в некоторых сетях, в которых линии связи используются (разделяются) попеременно несколькими парами взаимодейству-

ющих компьютеров, физическая среда передачи может быть занята. Поэтому одной из задач канального уровня (Data Link layer) является проверка доступности среды передачи. Другой задачей канального Уровня является реализация механизмов обнаружения и коррекции ошибок. Для этого на канальном уровне биты группируются в наборы, называемые *кадрами (frames)*. Канальный уровень обеспечивает корректность передачи каждого кадра, помещая специальную последовательность бит в начало и конец каждого кадра, для его выделения, а также вычисляет контрольную сумму, обрабатывая все байты кадра определенным способом и добавляя контрольную сумму к кадру. Когда кадр приходит по сети, получатель снова вычисляет контрольную сумму полученных данных и сравнивает результат с контрольной суммой из кадра. Если они совпадают, кадр считается правильным и принимается. Если же контрольные суммы не совпадают, то фиксируется ошибка. Канальный уровень может не только обнаруживать ошибки, но и исправлять их за счет повторной передачи поврежденных кадров. Необходимо отметить, что функция исправления ошибок не является обязательной для канального уровня, поэтому в некоторых протоколах этого уровня она отсутствует, например, в Ethernet и frame relay.

В протоколах канального уровня, используемых в локальных сетях, заложена определенная структура связей между компьютерами и способы их адресации. Хотя канальный уровень и обеспечивает доставку кадра между любыми двумя узлами локальной сети, он это делает только в сети с совершенно определенной топологией связей, именно той топологией, для которой он был разработан. К таким типовым топологиям, поддерживаемым протоколами канального уровня локальных сетей, относятся общая шина, кольцо и звезда, а также структуры, полученные из них с помощью мостов и коммутаторов. Примерами протоколов канального уровня являются протоколы Ethernet, Token Ring, FDDI, 100VG-AnyLAN.

В локальных сетях протоколы канального уровня используются компьютерами, мостами, коммутаторами и маршрутизаторами. В компьютерах функции канального уровня реализуются совместными усилиями сетевых адаптеров и их драйверов.

В глобальных сетях, которые редко обладают регулярной топологией, канальный уровень часто обеспечивает обмен сообщениями только между двумя соседними компьютерами, соединенными индивидуальной линией связи. Примерами протоколов «точка-точка» (как часто называют такие протоколы) могут служить широко распространенные протоколы PPP и LAP-B. В таких случаях для доставки сообщений между конечными узлами через всю сеть используются средства сетевого уровня. Именно так организованы сети X.25. Иногда в глобальных

сетях функции канального уровня в чистом виде выделить трудно, так как в одном и том же протоколе они объединяются с функциями сетевого уровня. Примерами такого подхода могут служить протоколы технологий ATM и frame relay.

В целом канальный уровень представляет собой весьма мощный и законченный набор функций по пересылке сообщений между узлами сети. В некоторых случаях протоколы канального уровня оказываются самодостаточными транспортными средствами и могут допускать работу поверх них непосредственно протоколов прикладного уровня или приложений, без привлечения средств сетевого и транспортного уровней. Например, существует реализация протокола управления сетью SNMP непосредственно поверх Ethernet, хотя стандартно этот протокол работает поверх сетевого протокола IP и транспортного протокола UDP. Естественно, что применение такой реализации будет ограниченным – она не подходит для составных сетей разных технологий, например Ethernet и X.25, и даже для такой сети, в которой во всех сегментах применяется Ethernet, но между сегментами существуют петлевидные связи. А вот в двухсегментной сети Ethernet, объединенной мостом, реализация SNMP над канальным уровнем будет вполне работоспособна.

Тем не менее для обеспечения качественной транспортировки сообщений в сетях любых топологий и технологий функций канального уровня оказывается недостаточно, поэтому в модели OSI решение этой задачи возлагается на два следующих уровня – сетевой и транспортный.

Сетевой уровень

Сетевой уровень (Network layer) служит для образования единой транспортной системы, объединяющей несколько сетей, причем эти сети могут использовать совершенно различные принципы передачи сообщений между конечными узлами и обладать произвольной структурой связей. Функции сетевого уровня достаточно разнообразны. Начнем их рассмотрение на примере объединения локальных сетей.

Протоколы канального уровня локальных сетей обеспечивают доставку данных между любыми узлами только в сети с соответствующей типовой топологией, например топологией иерархической звезды. Это очень жесткое ограничение, которое не позволяет строить сети с развитой структурой, например, сети, объединяющие несколько сетей предприятия в единую сеть, или высоконадежные сети, в которых существуют избыточные связи между узлами. Можно было бы усложнять протоколы канального уровня для поддержания петлевидных избыточных связей, но принцип разделения обязанностей между уровнями приводит к другому решению. Чтобы с одной, стороны сохранить простоту

процедур передачи данных для типовых топологий, а с другой допустить использование произвольных топологий, вводится дополнительный сетевой уровень.

На сетевом уровне сам термин *сеть* наделяют специфическим значением. В данном случае под сетью понимается совокупность компьютеров, соединенных между собой в соответствии с одной из стандартных типовых топологий и использующих для передачи данных один из протоколов канального уровня, определенный для этой топологии.

Внутри сети доставка данных обеспечивается соответствующим канальным уровнем, а вот доставкой данных между сетями занимается сетевой уровень, который и поддерживает возможность правильного выбора маршрута передачи сообщения даже в том случае, когда структура связей между составляющими сетями имеет характер, отличный от принятого в протоколах канального уровня.

Сети соединяются между собой специальными устройствами, называемыми маршрутизаторами. *Маршрутизатор* – это устройство, которое собирает информацию о топологии межсетевых соединений и на ее основании пересылает пакеты сетевого уровня в сеть назначения. Чтобы передать сообщение от отправителя, находящегося в одной сети, получателю, находящемуся в другой сети, нужно совершить некоторое количество *транзитных передач между сетями*, или *хопов* (от *hop* – прыжок), каждый раз выбирая подходящий маршрут. Таким образом, маршрут представляет собой последовательность маршрутизаторов, через которые проходит пакет.

На рис. 7.4 показаны четыре сети, связанные тремя маршрутизаторами. Между узлами А и В данной сети пролегают два маршрута: первый через маршрутизаторы 1 и 3, а второй через маршрутизаторы 1, 2 и 3.

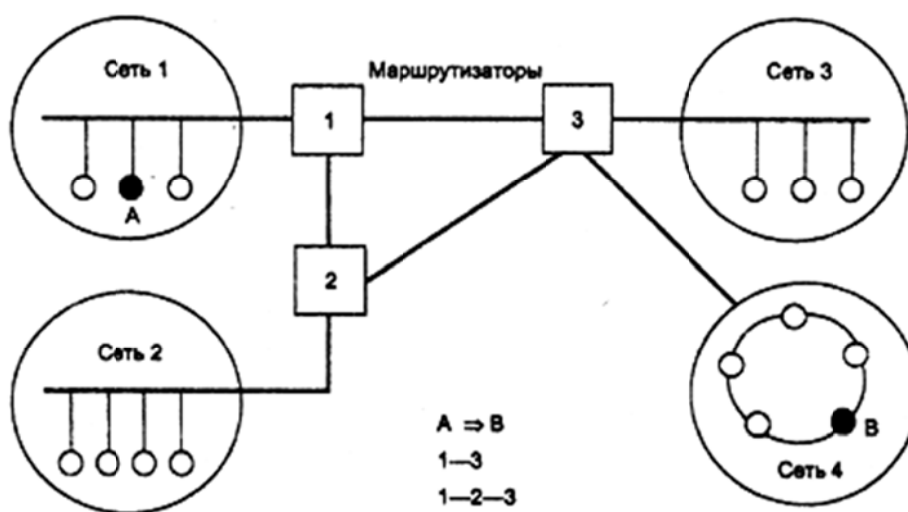


Рис. 7.4. Пример составной сети

Проблема выбора наилучшего пути называется *маршрутизацией*, и ее решение является одной из главных задач сетевого уровня. Эта проблема осложняется тем, что самый короткий путь не всегда самый лучший. Часто критерием при выборе маршрута является время передачи данных по этому маршруту; оно зависит от пропускной способности каналов связи и интенсивности трафика, которая может изменяться с течением времени. Некоторые алгоритмы маршрутизации пытаются приспособиться к изменению нагрузки, в то время как другие принимают решения на основе средних показателей за длительное время. Выбор маршрута может осуществляться и по другим критериям, например надежности передачи.

В общем случае функции сетевого уровня шире, чем функции передачи сообщений по связям с нестандартной структурой, которые мы сейчас рассмотрели на примере объединения нескольких локальных сетей. Сетевой уровень решает также задачи согласования разных технологий, упрощения адресации в крупных сетях и создания надежных и гибких барьеров на пути нежелательного трафика между сетями.

Сообщения сетевого уровня принято называть *пакетами (packets)*. При организации доставки пакетов на сетевом уровне используется понятие «номер сети». В этом случае адрес получателя состоит из старшей части – номера сети и младшей – номера узла в этой сети. Все узлы одной сети должны иметь одну и ту же старшую часть адреса, поэтому термину «сеть» на сетевом уровне можно дать и другое, более формальное определение: сеть – это совокупность узлов, сетевой адрес которых содержит один и тот же номер сети.

На сетевом уровне определяются два вида протоколов. Первый вид – *сетевые протоколы (routed protocols)* – реализуют продвижение пакетов через сеть. Именно эти протоколы обычно имеют в виду, когда говорят о протоколах сетевого уровня. Однако часто к сетевому уровню относят и другой вид протоколов, называемых протоколами обмена маршрутной информацией или просто *протоколами маршрутизации (routing protocols)*. С помощью этих протоколов маршрутизаторы собирают информацию о топологии межсетевых соединений. Протоколы сетевого уровня реализуются программными модулями операционной системы, а также программными и аппаратными средствами маршрутизаторов.

На сетевом уровне работают протоколы еще одного типа, которые отвечают за отображение адреса узла, используемого на сетевом уровне, в локальный адрес сети. Такие протоколы часто называют протоколами *разрешения адресов – Address Resolution Protocol, ARP*. Иногда их относят не к сетевому уровню, а к канальному, хотя тонкости классификации не изменяют их сути.

Примерами протоколов сетевого уровня являются протокол межсетевого взаимодействия IP стека TCP/IP и протокол межсетевого обмена пакетами IPX стека Novell.

Транспортный уровень

На пути от отправителя к получателю пакеты могут быть искажены или утеряны. Хотя некоторые приложения имеют собственные средства обработки ошибок, существуют и такие, которые предпочитают сразу иметь дело с надежным соединением. Транспортный уровень (Transport layer) обеспечивает приложениям или верхним уровням стека – прикладному и сеансовому – передачу данных с той степенью надежности, которая им требуется. Модель OSI определяет пять классов сервиса, предоставляемых транспортным уровнем. Эти виды сервиса отличаются качеством предоставляемых услуг: срочностью, возможностью восстановления прерванной связи, наличием средств мультиплексирования нескольких соединений между различными прикладными протоколами через общий транспортный протокол, а главное – способностью к обнаружению и исправлению ошибок передачи, таких как искажение, потеря и дублирование пакетов.

Выбор класса сервиса транспортного уровня определяется, с одной стороны, тем, в какой степени задача обеспечения надежности решается самими приложениями и протоколами более высоких, чем транспортный, уровней, а с другой стороны, этот выбор зависит от того, насколько надежной является система транспортировки данных в сети, обеспечиваемая уровнями, расположенными ниже транспортного – сетевым, канальным и физическим. Так, например, если качество каналов передачи связи очень высокое и вероятность возникновения ошибок, не обнаруженных протоколами более низких уровней, невелика, то разумно воспользоваться одним из облегченных сервисов транспортного уровня, не обремененных многочисленными проверками, квитированием и другими приемами повышения надежности. Если же транспортные средства нижних уровней изначально очень ненадежны, то целесообразно обратиться к наиболее развитому сервису транспортного уровня, который работает, используя максимум средств для обнаружения и устранения ошибок, – с помощью предварительного установления логического соединения, контроля доставки сообщений по контрольным суммам и циклической нумерации пакетов, установления тайм-аутов доставки и т. п.

Как правило, все протоколы, начиная с транспортного уровня и выше, реализуются программными средствами конечных узлов сети – компонентами их сетевых операционных систем. В качестве примера

транспортных протоколов можно привести протоколы TCP и UDP стека TCP/IP и протокол SPX стека Novell.

Протоколы нижних четырех уровней обобщенно называют сетевым транспортом или транспортной подсистемой, так как они полностью решают задачу транспортировки сообщений с заданным уровнем качества в составных сетях с произвольной топологией и различными технологиями. Остальные три верхних уровня решают задачи предоставления прикладных сервисов на основании имеющейся транспортной подсистемы.

Сеансовый уровень

Сеансовый уровень (Session layer) обеспечивает управление диалогом: фиксирует, какая из сторон является активной в настоящий момент, предоставляет средства синхронизации. Последние позволяют вставлять контрольные точки в длинные передачи, чтобы в случае отказа можно было вернуться назад к последней контрольной точке, а не начинать все с начала. На практике немногие приложения используют сеансовый уровень, и он редко реализуется в виде отдельных протоколов, хотя функции этого уровня часто объединяют с функциями прикладного уровня и реализуют в одном протоколе.

Представительный уровень

Представительный уровень (Presentation layer) имеет дело с формой представления передаваемой по сети информации, не меняя при этом ее содержания. За счет уровня представления информация, передаваемая прикладным уровнем одной системы, всегда понятна прикладному уровню другой системы. С помощью средств данного уровня протоколы прикладных уровней могут преодолеть синтаксические различия в представлении данных или же различия в кодах символов, например кодов ASCII и EBCDIC. На этом уровне может выполняться шифрование и дешифрование данных, благодаря которому секретность обмена данными обеспечивается сразу для всех прикладных служб. Примером такого протокола является протокол Secure Socket Layer (SSL), который обеспечивает секретный обмен сообщениями для протоколов прикладного уровня стека TCP/IP.

Прикладной уровень

Прикладной уровень (Application layer) – это в действительности просто набор разнообразных протоколов, с помощью которых пользователи сети получают доступ к разделяемым ресурсам, таким как файлы, принтеры или гипертекстовые Web-страницы, а также организуют свою совместную работу, например, с помощью протокола электронной по-

чты. Единица данных, которой оперирует прикладной уровень, обычно называется *сообщением (message)*.

Существует очень большое разнообразие служб прикладного уровня. Приведем в качестве примера хотя бы несколько наиболее распространенных реализаций файловых служб: NCP в операционной системе Novell NetWare, SMB в Microsoft Windows NT, NFS, FTP и TFTP, входящие в стек TCP/IP.

7.4. Сетезависимые и сетезависимые уровни

Функции всех уровней модели OSI могут быть отнесены к одной из двух групп: либо к функциям, зависящим от конкретной технической реализации сети, либо к функциям, ориентированным на работу с приложениями.

Три нижних уровня – физический, канальный и сетевой – являются сетезависимыми, то есть протоколы этих уровней тесно связаны с технической реализацией сети и используемым коммуникационным оборудованием. Например, переход на оборудование FDDI означает полную смену протоколов физического и канального уровней во всех узлах сети.

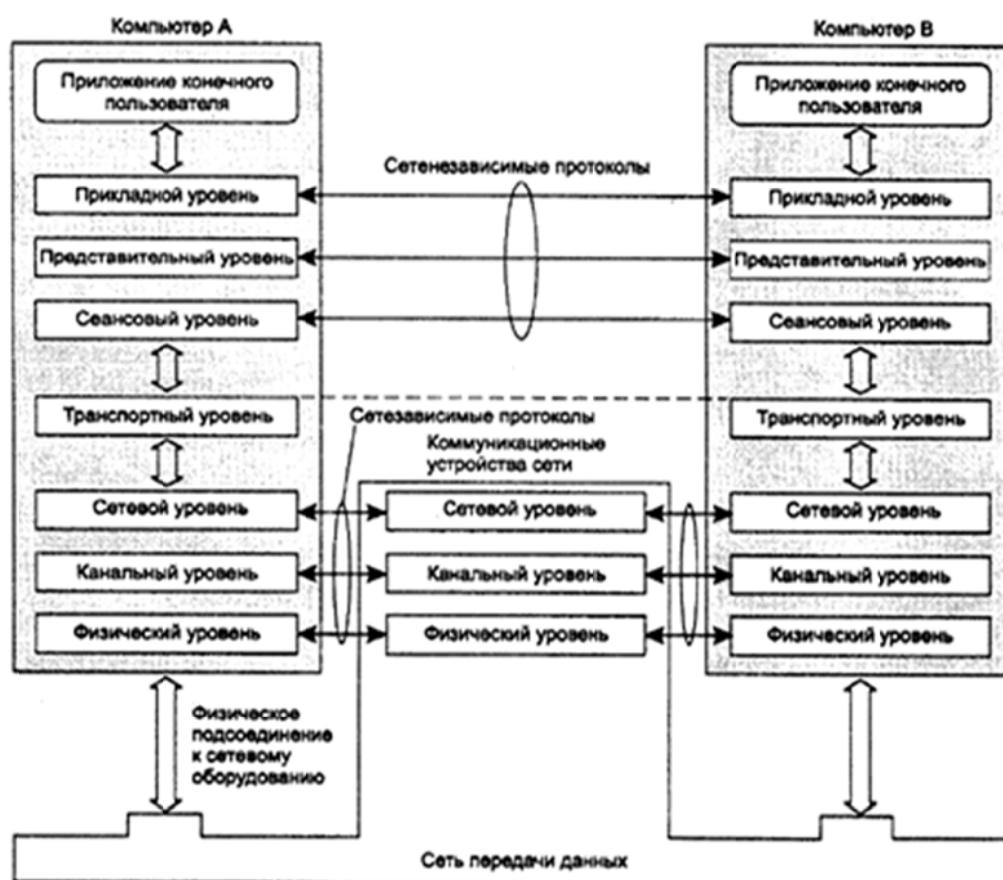


Рис. 7.5. Сетезависимые и сетезависимые уровни модели OSI

Три верхних уровня – прикладной, представительный и сеансовый – ориентированы на приложения и мало зависят от технических особенностей построения сети. На протоколы этих уровней не влияют какие бы то ни было изменения в топологии сети, замена оборудования или переход на другую сетевую технологию. Так, переход от Ethernet на высокоскоростную технологию 100VG-AnyLAN не потребует никаких изменений в программных средствах, реализующих функции прикладного, представительного и сеансового уровней.

Транспортный уровень является промежуточным, он скрывает все детали функционирования нижних уровней от верхних. Это позволяет разрабатывать приложения, не зависящие от технических средств непосредственной транспортировки сообщений.

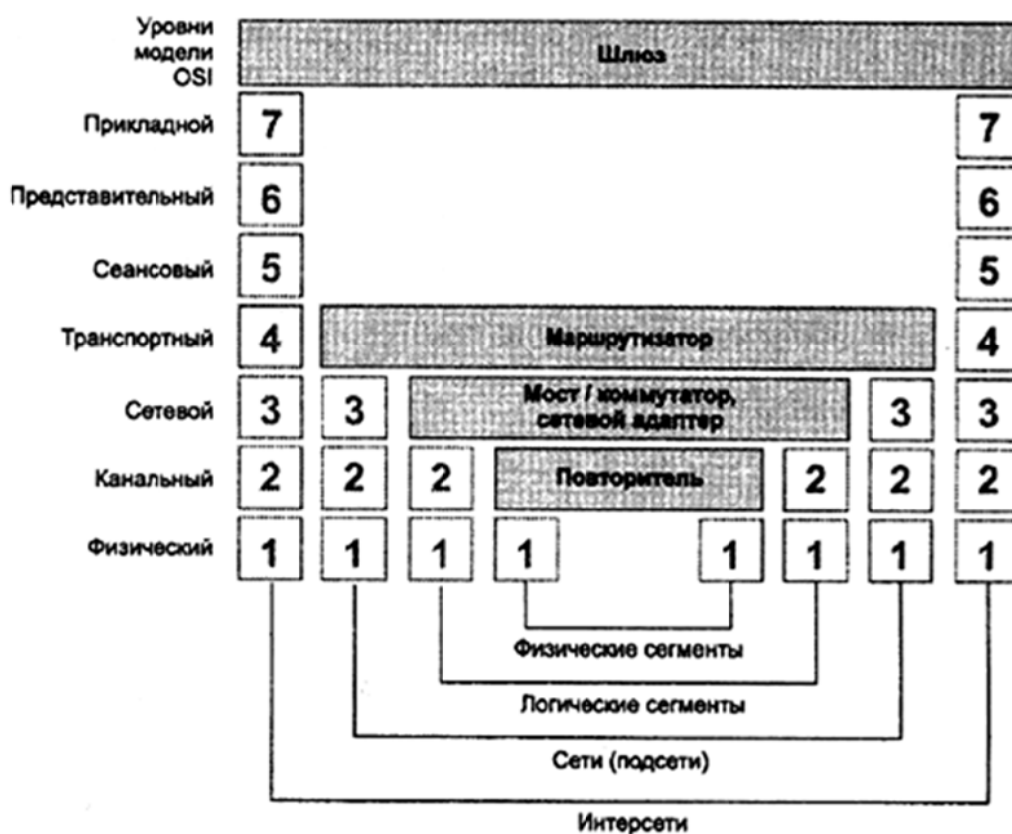


Рис. 7.6. Соответствие функций различных устройств сети уровням модели OSI

На рис. 7.5 показаны уровни модели OSI, на которых работают различные элементы сети. Компьютер с установленной на нем сетевой ОС взаимодействует с другим компьютером с помощью протоколов всех семи уровней. Это взаимодействие компьютеры осуществляют опосредованно через различные коммуникационные устройства: концентраторы, модемы, мосты, коммутаторы, маршрутизаторы, мультиплексоры. В зависимости от

типа коммуникационное устройство может работать либо только на физическом уровне (повторитель), либо на физическом и канальном (мост), либо на физическом, канальном и сетевом, иногда захватывая и транспортный уровень (маршрутизатор). На рис. 7.6 показано соответствие функций различных коммуникационных устройств уровням модели OSI.

Модель OSI представляет хотя и очень важную, но только одну из многих моделей коммуникаций. Эти модели и связанные с ними стеки протоколов могут отличаться количеством уровней, их функциями, форматами сообщений, службами, поддерживаемыми на верхних уровнях, и прочими параметрами.

7.5. 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. Эти протоколы часто не соответствуют рекомендуемому моделью OSI разбиению на уровни. В частности, функции сеансового и представительного уровня, как правило, объединены с прикладным уровнем. Такое несоответствие связано с тем, что модель OSI появилась как результат обобщения уже существующих и реально используемых стеков, а не наоборот.

Стек TCP/IP

Стек TCP/IP был разработан по инициативе Министерства обороны США более 20 лет назад 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. 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.

Сегодня стек TCP/IP представляет собой один из самых распространенных стеков транспортных протоколов вычислительных сетей. Действительно, только в сети Internet объединено около 10 миллионов компьютеров по всему миру, которые взаимодействуют друг с другом с помощью стека протоколов TCP/IP.

Стремительный рост популярности Internet привел и к изменениям в расстановке сил в мире коммуникационных протоколов – протоколы TCP/IP, на которых построен Internet, стали быстро теснить бесспорного лидера прошлых лет – стек IPX/SPX компании Novell. Сегодня в мире общее количество компьютеров, на которых установлен стек TCP/IP, сравнялось с общим количеством компьютеров, на которых работает стек IPX/SPX, и это говорит о резком переломе в отношении администраторов локальных сетей к протоколам, используемым на настольных компьютерах, так как именно они составляют подавляющее число мирового компьютерного парка и именно на них раньше почти везде работали протоколы компании Novell, необходимые для доступа к файловым серверам NetWare. Процесс становления стека TCP/IP в качестве стека номер один в любых типах сетей продолжается, и сейчас любая промышленная операционная система обязательно включает программную реализацию этого стека в своем комплекте поставки.

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 широко-вещательных 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 широко-вещательных 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.

Можно приводить и другие доводы за и против стека протоколов Internet, однако факт остается фактом – сегодня это самый популярный стек протоколов, широко используемый как в глобальных, так и локальных сетях.

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 широкоэмитательными 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 базовой системы ввода/вывода (BIOS) 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.

Модель OSI	IBM/Microsoft	TCP/IP	Novell	Стек OSI
Прикладной	SMB	Telnet, FTP, SNMP, SMTP, WWW	NCP, SAP	X.400 X.500 FTAM
Представительный				Представительный протокол OSI
Сеансовый	NetBIOS	TCP	SPX	Сеансовый протокол OSI
Транспортный				Транспортный протокол OSI
Сетевой	IP, RIP, OSPF		IPX, RIP, NLSP	ES-ES IS-IS
Канальный	802.3 (Ethernet), 802.5 (Token Ring), FDDI, Fast Ethernet, SLIP, 100VG-AnyLAN, X.25, ATM, LAP-B, LAP-D, PPP			
Физический	Коаксиал, экранированная и неэкранированная витая пара, оптоволокну, радиоволны			

Рис. 7.7. Соответствие популярных стеков протоколов модели ISO

На рис. 7.7 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. В большинстве случаев разработчики стеков отдавали предпочтение скорости работы сети в ущерб модульности – ни один стек, кроме стека OSI, не разбит на семь уровней. Чаще всего в стеке явно выделяются 3–4 уровня: уровень сетевых адаптеров, в котором реализуются протоколы физического и канального уровней, сетевой уровень, транспортный уровень и уровень служб, вбирающий в себя функции сеансового, представительного и прикладного уровней.

ГЛАВА 8. СТЕК ПРОТОКОЛОВ TCP/IP

8.1. История и перспективы стека TCP/IP

Transmission Control Protocol/Internet Protocol (TCP/IP) – это промышленный стандарт стека протоколов, разработанный для глобальных сетей.

Стандарты TCP/IP опубликованы в серии документов, названных Request for Comment (RFC). Документы RFC описывают внутреннюю работу сети Internet. Некоторые RFC описывают сетевые сервисы или протоколы и их реализацию, в то время как другие обобщают условия применения. Стандарты TCP/IP всегда публикуются в виде документов RFC, но не все RFC определяют стандарты.

Стек был разработан по инициативе Министерства обороны США (Department of Defence, DoD) более 20 лет назад для связи экспериментальной сети ARPAnet с другими сателлитными сетями как набор общих протоколов для разнородной вычислительной среды. Сеть ARPA поддерживала разработчиков и исследователей в военных областях. В сети ARPA связь между двумя компьютерами осуществлялась с использованием протокола Internet Protocol (IP), который и по сей день является одним из основных в стеке TCP/IP и фигурирует в названии стека.

Большой вклад в развитие стека TCP/IP внес университет Беркли, реализовав протоколы стека в своей версии ОС UNIX. Широкое распространение ОС UNIX привело и к широкому распространению протокола IP и других протоколов стека. На этом же стеке работает всемирная информационная сеть Internet, чье подразделение Internet Engineering Task Force (IETF) вносит основной вклад в совершенствование стандартов стека, публикуемых в форме спецификаций RFC.

Если в настоящее время стек TCP/IP распространен в основном в сетях с ОС UNIX, то реализация его в последних версиях сетевых операционных систем для персональных компьютеров (Windows NT 3.5, NetWare 4.1, Windows 95) является хорошей предпосылкой для быстрого роста числа установок стека TCP/IP.

Итак, лидирующая роль стека TCP/IP объясняется следующими его свойствами:

- Это наиболее заверченный стандартный и в то же время популярный стек сетевых протоколов, имеющий многолетнюю историю.
- Почти все большие сети передают основную часть своего трафика с помощью протокола TCP/IP.
- Это метод получения доступа к сети Internet.

- Этот стек служит основой для создания intranet-корпоративной сети, использующей транспортные услуги Internet и гипертекстовую технологию WWW, разработанную в Internet.
- Все современные операционные системы поддерживают стек TCP/IP.
- Это гибкая технология для соединения разнородных систем как на уровне транспортных подсистем, так и на уровне прикладных сервисов.
- Это устойчивая масштабируемая межплатформенная среда для приложений клиент-сервер.

8.2. Структура стека TCP/IP. Краткая характеристика протоколов

Так как стек TCP/IP был разработан до появления модели взаимодействия открытых систем ISO/OSI, то, хотя он также имеет многоуровневую структуру, соответствие уровней стека TCP/IP уровням модели OSI достаточно условно.

Структура протоколов TCP/IP приведена на рис. 8.1. Протоколы TCP/IP делятся на 4 уровня.

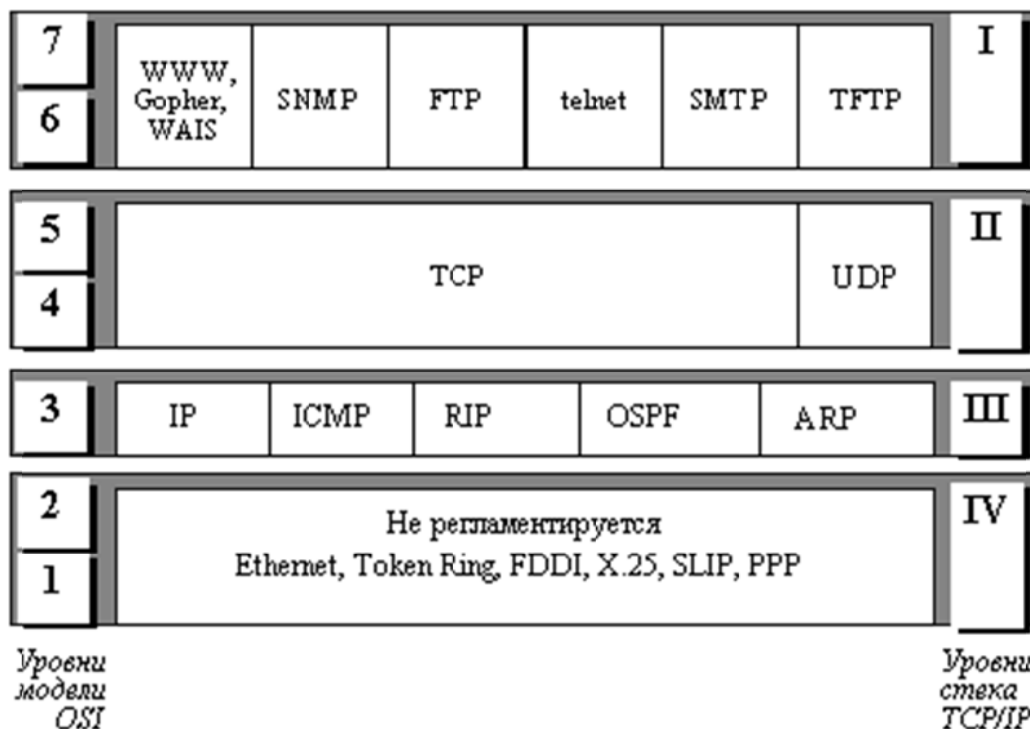


Рис. 8.1. Стек TCP/IP

Самый нижний (уровень IV) соответствует физическому и канальному уровням модели OSI. Этот уровень в протоколах TCP/IP не регламентируется, но поддерживает все популярные стандарты физического и канального уровня: для локальных сетей это Ethernet, Token Ring,

FDDI, Fast Ethernet, 100VG-AnyLAN, для глобальных сетей – протоколы соединений «точка-точка» SLIP и PPP, протоколы территориальных сетей с коммутацией пакетов X.25, frame relay. Разработана также специальная спецификация, определяющая использование технологии АТМ в качестве транспорта канального уровня. Обычно при появлении новой технологии локальных или глобальных сетей она быстро включается в стек TCP/IP за счет разработки соответствующего RFC, определяющего метод инкапсуляции пакетов IP в ее кадры.

Следующий уровень (*уровень III*) – это уровень межсетевого взаимодействия, который занимается передачей пакетов с использованием различных транспортных технологий локальных сетей, территориальных сетей, линий специальной связи и т. п.

В качестве основного протокола сетевого уровня (в терминах модели OSI) в стеке используется протокол **IP**, который изначально проектировался как протокол передачи пакетов в составных сетях, состоящих из большого количества локальных сетей, объединенных как локальными, так и глобальными связями. Поэтому протокол IP хорошо работает в сетях со сложной топологией, рационально используя наличие в них подсистем и экономно расходуя пропускную способность низкоскоростных линий связи. Протокол IP является дейтаграммным протоколом, то есть он не гарантирует доставку пакетов до узла назначения, но старается это сделать.

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

Следующий уровень (*уровень II*) называется основным. На этом уровне функционируют протокол управления передачей **TCP** (Transmission Control Protocol) и протокол дейтаграмм пользователя **UDP** (User Datagram Protocol). Протокол TCP обеспечивает надежную передачу сообщений между удаленными прикладными процессами за счет образования виртуальных соединений. Протокол UDP обеспечивает передачу прикладных пакетов дейтаграммным способом, как и IP, и

выполняет только функции связующего звена между сетевым протоколом и многочисленными прикладными процессами.

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

Протокол пересылки файлов **FTP** (File Transfer Protocol) реализует удаленный доступ к файлу. Для того, чтобы обеспечить надежную передачу, FTP использует в качестве транспорта протокол с установлением соединений – TCP. Кроме пересылки файлов протокол FTP предлагает и другие услуги. Так, пользователю предоставляется возможность интерактивной работы с удаленной машиной, например, он может распечатать содержимое ее каталогов. Наконец, FTP выполняет аутентификацию пользователей. Прежде, чем получить доступ к файлу, в соответствии с протоколом пользователи должны сообщить свое имя и пароль. Для доступа к публичным каталогам FTP-архивов Internet парольная аутентификация не требуется, и ее обходят за счет использования для такого доступа предопределенного имени пользователя Anonymous.

В стеке TCP/IP протокол FTP предлагает наиболее широкий набор услуг для работы с файлами, однако он является и самым сложным для программирования. Приложения, которым не требуются все возможности FTP, могут использовать другой, более экономичный протокол – простейший протокол пересылки файлов **TFTP** (Trivial File Transfer Protocol). Этот протокол реализует только передачу файлов, причем в качестве транспорта используется более простой, чем TCP, протокол без установления соединения – UDP.

Протокол **telnet** обеспечивает передачу потока байтов между процессами, а также между процессом и терминалом. Наиболее часто этот протокол используется для эмуляции терминала удаленного компьютера. При использовании сервиса telnet пользователь фактически управляет удаленным компьютером так же, как и локальный пользователь, поэтому такой вид доступа требует хорошей защиты. Поэтому серверы telnet всегда используют как минимум аутентификацию по паролю, а иногда и более мощные средства защиты, например, систему Kerberos.

Протокол **SNMP** (Simple Network Management Protocol) используется для организации сетевого управления. Изначально протокол SNMP

был разработан для удаленного контроля и управления маршрутизаторами Internet, которые традиционно часто называют также шлюзами. С ростом популярности протокол SNMP стали применять и для управления любым коммуникационным оборудованием – концентраторами, мостами, сетевыми адаптерами и т. д. и т. п. Проблема управления в протоколе SNMP разделяется на две задачи.

Первая задача связана с передачей информации. Протоколы передачи управляющей информации определяют процедуру взаимодействия SNMP-агента, работающего в управляемом оборудовании, и SNMP-монитора, работающего на компьютере администратора, который часто называют также консолью управления. Протоколы передачи определяют форматы сообщений, которыми обмениваются агенты и монитор.

Вторая задача связана с контролируруемыми переменными, характеризующими состояние управляемого устройства. Стандарты регламентируют, какие данные должны сохраняться и накапливаться в устройствах, имена этих данных и синтаксис этих имен. В стандарте SNMP определена спецификация информационной базы данных управления сетью. Эта спецификация, известная как база данных MIB (Management Information Base), определяет те элементы данных, которые управляемое устройство должно сохранять, и допустимые операции над ними.

8.3. Адресация в IP-сетях

Каждый компьютер в сети TCP/IP имеет адреса трех уровней:

- Локальный адрес узла, определяемый технологией, с помощью которой построена отдельная сеть, в которую входит данный узел. Для узлов, входящих в локальные сети – это MAC-адрес сетевого адаптера или порта маршрутизатора.
- IP-адрес, состоящий из 4 байт, например, 109.26.17.100. Этот адрес используется на сетевом уровне. Он назначается администратором во время конфигурирования компьютеров и маршрутизаторов. IP-адрес состоит из двух частей: номера сети и номера узла. Номер сети может быть выбран администратором произвольно, либо назначен по рекомендации специального подразделения Internet (Network Information Center, NIC), если сеть должна работать как составная часть Internet. Обычно провайдеры услуг Internet получают диапазоны адресов у подразделений NIC, а затем распределяют их между своими абонентами.

Номер узла в протоколе IP назначается независимо от локального адреса узла. Деление IP-адреса на поле номера сети и номера узла – гибкое, и граница между этими полями может устанавливаться весьма произвольно. Узел может входить в несколько IP-сетей. В этом случае

узел должен иметь несколько IP-адресов, по числу сетевых связей. Таким образом IP-адрес характеризует не отдельный компьютер или маршрутизатор, а одно сетевое соединение.

- Символьный идентификатор-имя, например, SERV1.IBM.COM. Этот адрес назначается администратором и состоит из нескольких частей, например, имени машины, имени организации, имени домена. Такой адрес, называемый также DNS-именем, используется на прикладном уровне, например, в протоколах FTP или telnet.

8.3.1. Три основных класса IP-адресов

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

128.10.2.30 – традиционная десятичная форма представления адреса,

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

На рис. 8.2 показана структура IP-адреса.

Класс А

0	N сети			N узла		
---	--------	--	--	--------	--	--

Класс В

1	0	N сети			N узла		
---	---	--------	--	--	--------	--	--

Класс С

1	1	0	N сети			N узла		
---	---	---	--------	--	--	--------	--	--

Класс D

1	1	1	0	адрес группы multicast			
---	---	---	---	------------------------	--	--	--

Класс E

1	1	1	1	0	зарезервирован		
---	---	---	---	---	----------------	--	--

Рис. 8.2. Структура IP-адреса

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

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

- Если первые два бита адреса равны 10, то сеть относится к классу В и является сетью средних размеров с числом узлов $2^8 \dots 2^{16}$. В сетях класса В под адрес сети и под адрес узла отводится по 16 битов, то есть по 2 байта.
- Если адрес начинается с последовательности 110, то это сеть класса С с числом узлов не больше 2^8 . Под адрес сети отводится 24 бита, а под адрес узла – 8 битов.
- Если адрес начинается с последовательности 1110, то он является адресом класса D и обозначает особый, групповой адрес – multicast. Если в пакете в качестве адреса назначения указан адрес класса D, то такой пакет должны получить все узлы, которым присвоен данный адрес.
- Если адрес начинается с последовательности 11110, то это адрес класса E, он зарезервирован для будущих применений.

В табл. 8.1 приведены диапазоны номеров сетей, соответствующих каждому классу сетей.

Таблица 8.1

Диапазоны номеров сетей

Класс	Наименьший адрес	Наибольший адрес
A	01.0.0	126.0.0.0
B	128.0.0.0	191.255.0.0
C	192.0.1.0.	223.255.255.0
D	224.0.0.0	239.255.255.255
E	240.0.0.0	247.255.255.255

**8.3.2. Соглашения о специальных адресах:
broadcast, multicast, loopback**

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

- если IP-адрес состоит только из двоичных нулей, то он обозначает адрес того узла, который сгенерировал этот пакет;
- если в поле номера сети стоят 0, то по умолчанию считается, что этот узел принадлежит той же самой сети, что и узел, который отправил пакет;
- если все двоичные разряды IP-адреса равны 1, то пакет с таким адресом назначения должен рассылаться всем узлам, находящимся в той же сети, что и источник этого пакета. Такая рассылка называется ограниченным широковещательным сообщением (limited broadcast);

- если в поле адреса назначения стоят сплошные 1, то пакет, имеющий такой адрес рассылается всем узлам сети с заданным номером. Такая рассылка называется широковещательным сообщением (broadcast);
- адрес 127.0.0.1 зарезервирован для организации обратной связи при тестировании работы программного обеспечения узла без реальной отправки пакета по сети. Этот адрес имеет название loopback.

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

В протоколе IP нет понятия широковещательности в том смысле, в котором оно используется в протоколах канального уровня локальных сетей, когда данные должны быть доставлены абсолютно всем узлам. Как ограниченный широковещательный IP-адрес, так и широковещательный IP-адрес имеют пределы распространения в интрасети – они ограничены либо сетью, к которой принадлежит узел – источник пакета, либо сетью, номер которой указан в адресе назначения. Поэтому деление сети с помощью маршрутизаторов на части локализует широковещательный шторм пределами одной из составляющих общую сеть частей просто потому, что нет способа адресовать пакет одновременно всем узлам всех сетей составной сети.

8.3.3. Отображение физических адресов на IP-адреса

В протоколе IP-адрес узла, то есть адрес компьютера или порта маршрутизатора, назначается произвольно администратором сети и прямо не связан с его локальным адресом, как это сделано, например, в протоколе IPX. Подход, используемый в IP, удобно использовать в крупных сетях и по причине его независимости от формата локального адреса, и по причине стабильности, так как в противном случае, при смене на компьютере сетевого адаптера это изменение должно бы было учитывать все адресаты всемирной сети Internet (в том случае, конечно, если сеть подключена к Internet'у).

Локальный адрес используется в протоколе IP только в пределах локальной сети при обмене данными между маршрутизатором и узлом этой сети. Маршрутизатор, получив пакет для узла одной из сетей, непосредственно подключенных к его портам, должен для передачи пакета сформировать

ровать кадр в соответствии с требованиями принятой в этой сети технологии и указать в нем локальный адрес узла, например его MAC-адрес. В пришедшем пакете этот адрес не указан, поэтому перед маршрутизатором встает задача поиска его по известному IP-адресу, который указан в пакете в качестве адреса назначения. С аналогичной задачей сталкивается и конечный узел, когда он хочет отправить пакет в удаленную сеть через маршрутизатор, подключенный к той же локальной сети, что и данный узел.

Для определения локального адреса по IP-адресу используется протокол разрешения адреса *Address Resolution Protocol, ARP*. Протокол ARP работает различным образом в зависимости от того, какой протокол канального уровня работает в данной сети – протокол локальной сети (Ethernet, Token Ring, FDDI) с возможностью широковещательного доступа одновременно ко всем узлам сети, или же протокол глобальной сети (X.25, frame relay), как правило не поддерживающий широковещательный доступ. Существует также протокол, решающий обратную задачу – нахождение IP-адреса по известному локальному адресу. Он называется реверсивный ARP – *RARP (Reverse Address Resolution Protocol)* и используется при старте бездисковых станций, не знающих в начальный момент своего IP-адреса, но знающих адрес своего сетевого адаптера.

В локальных сетях протокол ARP использует широковещательные кадры протокола канального уровня для поиска в сети узла с заданным IP-адресом.

Узел, которому нужно выполнить отображение IP-адреса на локальный адрес, формирует ARP запрос, вкладывает его в кадр протокола канального уровня, указывая в нем известный IP-адрес, и рассылает запрос широковещательно. Все узлы локальной сети получают ARP запрос и сравнивают указанный там IP-адрес с собственным. В случае их совпадения узел формирует ARP-ответ, в котором указывает свой IP-адрес и свой локальный адрес и отправляет его уже направленно, так как в ARP запросе отправитель указывает свой локальный адрес. ARP-запросы и ответы используют один и тот же формат пакета. Так как локальные адреса могут в различных типах сетей иметь различную длину, то формат пакета протокола ARP зависит от типа сети. На рис. 8.3 показан формат пакета протокола ARP для передачи по сети Ethernet.

В поле типа сети для сетей Ethernet указывается значение 1. Поле типа протокола позволяет использовать пакеты ARP не только для протокола IP, но и для других сетевых протоколов. Для IP значение этого поля равно 0800_{16} .

Длина локального адреса для протокола Ethernet равна 6 байтам, а длина IP-адреса – 4 байтам. В поле операции для ARP запросов указывается значение 1 для протокола ARP и 2 для протокола RARP.

Тип сети		Тип протокола
Длина локального адреса	Длина сетевого адреса	Операция
Локальный адрес отправителя (байты 0 – 3)		
Локальный адрес отправителя (байты 4 – 5)		IP-адрес отправителя (байты 0 – 1)
IP-адрес отправителя (байты 2 – 3)		Искомый локальный адрес (байты 0 – 1)
Искомый локальный адрес (байты 2 – 5)		
Искомый IP-адрес (байты 0 – 3)		

Рис. 8.3. Формат пакета протокола ARP

Узел, отправляющий ARP-запрос, заполняет в пакете все поля, кроме поля искомого локального адреса (для RARP-запроса не указывается искомым IP-адрес). Значение этого поля заполняется узлом, опознавшим свой IP-адрес.

В глобальных сетях администратору сети чаще всего приходится вручную формировать ARP-таблицы, в которых он задает, например, соответствие IP-адреса адресу узла сети X.25, который имеет смысл локального адреса. В последнее время наметилась тенденция автоматизации работы протокола ARP и в глобальных сетях. Для этой цели среди всех маршрутизаторов, подключенных к какой-либо глобальной сети, выделяется специальный маршрутизатор, который ведет ARP-таблицу для всех остальных узлов и маршрутизаторов этой сети. При таком централизованном подходе для всех узлов и маршрутизаторов вручную нужно задать только IP-адрес и локальный адрес выделенного маршрутизатора. Затем каждый узел и маршрутизатор регистрирует свои адреса в выделенном маршрутизаторе, а при необходимости установления соответствия между IP-адресом и локальным адресом узел обращается к выделенному маршрутизатору с запросом и автоматически получает ответ без участия администратора.

8.3.4. Отображение символьных адресов на IP-адреса: служба DNS

DNS (Domain Name System) – это распределенная база данных, поддерживающая иерархическую систему имен для идентификации узлов в сети Internet. Служба DNS предназначена для автоматического поиска IP-адреса по известному символьному имени узла. Спецификация DNS определяется стандартами RFC 1034 и 1035. DNS требует статической конфигурации своих таблиц, отображающих имена компьютеров в IP-адрес.

Протокол DNS является служебным протоколом прикладного уровня. Этот протокол несимметричен – в нем определены DNS-серверы и DNS-клиенты. DNS-серверы хранят часть распределенной базы данных о соответствии символьных имен и IP-адресов. Эта база данных распределена по административным доменам сети Internet. Клиенты сервера DNS знают IP-адрес сервера DNS своего административного домена и по протоколу IP передают запрос, в котором сообщают известное символьное имя и просят вернуть соответствующий ему IP-адрес.

Если данные о запрошенном соответствии хранятся в базе данного DNS-сервера, то он сразу посылает ответ клиенту, если же нет – то он посылает запрос DNS-серверу другого домена, который может сам обработать запрос, либо передать его другому DNS-серверу. Все DNS-серверы соединены иерархически, в соответствии с иерархией доменов сети Internet. Клиент опрашивает эти серверы имен, пока не найдет нужные отображения. Этот процесс ускоряется из-за того, что серверы имен постоянно кэшируют информацию, предоставляемую по запросам. Клиентские компьютеры могут использовать в своей работе IP-адреса нескольких DNS-серверов, для повышения надежности своей работы.

База данных DNS имеет структуру дерева, называемого доменным пространством имен, в котором каждый домен (узел дерева) имеет имя и может содержать поддомены. Имя домена идентифицирует его положение в этой базе данных по отношению к родительскому домену, причем точки в имени отделяют части, соответствующие узлам домена.

Корень базы данных DNS управляется центром Internet Network Information Center. Домены верхнего уровня назначаются для каждой страны, а также на организационной основе. Имена этих доменов должны следовать международному стандарту ISO 3166. Для обозначения стран используются трехбуквенные и двухбуквенные аббревиатуры, а для различных типов организаций используются следующие аббревиатуры:

- com – коммерческие организации (например, microsoft.com);
- edu – образовательные (например, mit.edu);
- gov – правительственные организации (например, nsf.gov);
- org – некоммерческие организации (например, fidonet.org);
- net – организации, поддерживающие сети (например, nsf.net).

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

включает имена всех доменов по направлению от хоста к корню. Пример полного DNS-имени:

fmpk.tpu.ru.

8.3.5. Автоматизация процесса назначения IP-адресов узлам сети – протокол DHCP

Как уже было сказано, IP-адреса могут назначаться администратором сети вручную. Это представляет для администратора утомительную процедуру. Ситуация усложняется еще тем, что многие пользователи не обладают достаточными знаниями для того, чтобы конфигурировать свои компьютеры для работы в интрасети и должны поэтому полагаться на администраторов.

Протокол *Dynamic Host Configuration Protocol (DHCP)* был разработан для того, чтобы освободить администратора от этих проблем. Основным назначением DHCP является динамическое назначение IP-адресов. Однако, кроме динамического, DHCP может поддерживать и более простые способы ручного и автоматического статического назначения адресов.

В ручной процедуре назначения адресов активное участие принимает администратор, который предоставляет DHCP-серверу информацию о соответствии IP-адресов физическим адресам или другим идентификаторам клиентов. Эти адреса сообщаются клиентам в ответ на их запросы к DHCP-серверу.

При автоматическом статическом способе DHCP-сервер присваивает IP-адрес (и, возможно, другие параметры конфигурации клиента) из пула наличных IP-адресов без вмешательства оператора. Границы пула назначаемых адресов задает администратор при конфигурировании DHCP-сервера. Между идентификатором клиента и его IP-адресом по-прежнему, как и при ручном назначении, существует постоянное соответствие. Оно устанавливается в момент первичного назначения сервером DHCP IP-адреса клиенту. При всех последующих запросах сервер возвращает тот же самый IP-адрес.

При динамическом распределении адресов DHCP-сервер выдает адрес клиенту на ограниченное время, что дает возможность впоследствии повторно использовать IP-адреса другими компьютерами. Динамическое распределение адресов позволяет строить IP-сеть, количество узлов в которой намного превышает количество имеющихся в распоряжении администратора IP-адресов.

DHCP обеспечивает надежный и простой способ конфигурации сети TCP/IP, гарантируя отсутствие конфликтов адресов за счет централизованного управления их распределением. Администратор управляет

процессом назначения адресов с помощью параметра «продолжительности аренды» (lease duration), которая определяет, как долго компьютер может использовать назначенный IP-адрес, перед тем как снова запросить его от сервера DHCP в аренду.

Примером работы протокола DHCP может служить ситуация, когда компьютер, являющийся клиентом DHCP, удаляется из подсети. При этом назначенный ему IP-адрес автоматически освобождается. Когда компьютер подключается к другой подсети, то ему автоматически назначается новый адрес. Ни пользователь, ни сетевой администратор не вмешиваются в этот процесс. Это свойство очень важно для мобильных пользователей.

Протокол DHCP использует модель клиент-сервер. Во время старта системы компьютер-клиент DHCP, находящийся в состоянии «инициализация», посылает сообщение discover (исследовать), которое широко-вещательно распространяется по локальной сети и передается всем DHCP-серверам частной интрасети. Каждый DHCP-сервер, получивший это сообщение, отвечает на него сообщением offer (предложение), которое содержит IP-адрес и конфигурационную информацию.

Компьютер-клиент DHCP переходит в состояние «выбор» и собирает конфигурационные предложения от DHCP-серверов. Затем он выбирает одно из этих предложений, переходит в состояние «запрос» и отправляет сообщение request (запрос) тому DHCP-серверу, чье предложение было выбрано.

Выбранный DHCP-сервер посылает сообщение DHCP-acknowledgment (подтверждение), содержащее тот же IP-адрес, который уже был послан ранее на стадии исследования, а также параметр аренды для этого адреса. Кроме того, DHCP-сервер посылает параметры сетевой конфигурации. После того, как клиент получит это подтверждение, он переходит в состояние «связь», находясь в котором он может принимать участие в работе сети TCP/IP. Компьютеры-клиенты, которые имеют локальные диски, сохраняют полученный адрес для использования при последующих стартах системы. При приближении момента истечения срока аренды адреса компьютер пытается обновить параметры аренды у DHCP-сервера, а если этот IP-адрес не может быть выделен снова, то ему возвращается другой IP-адрес.

В протоколе DHCP описывается несколько типов сообщений, которые используются для обнаружения и выбора DHCP-серверов, для запросов информации о конфигурации, для продления и досрочного прекращения лицензии на IP-адрес. Все эти операции направлены на то, чтобы освободить администратора сети от утомительных рутинных операций по конфигурированию сети.

Однако использование DHCP несет в себе и некоторые проблемы. Во-первых, это проблема согласования информационной адресной базы в службах DHCP и DNS. Как известно, DNS служит для преобразования символьных имен в IP-адреса. Если IP-адреса будут динамически изменяться сервером DHCP, то эти изменения необходимо также динамически вносить в базу данных сервера DNS. Хотя протокол динамического взаимодействия между службами DNS и DHCP уже реализован некоторыми фирмами (так называемая служба Dynamic DNS), стандарт на него пока не принят.

Во-вторых, нестабильность IP-адресов усложняет процесс управления сетью. Системы управления, основанные на протоколе SNMP, разработаны с расчетом на статичность IP-адресов. Аналогичные проблемы возникают и при конфигурировании фильтров маршрутизаторов, которые оперируют с IP-адресами.

Наконец, централизация процедуры назначения адресов снижает надежность системы: при отказе DHCP-сервера все его клиенты оказываются не в состоянии получить IP-адрес и другую информацию о конфигурации. Последствия такого отказа могут быть уменьшены путем использования в сети нескольких серверов DHCP, каждый из которых имеет свой пул IP-адресов.

8.4. Протокол межсетевого взаимодействия IP

Основу транспортных средств стека протоколов TCP/IP составляет протокол межсетевого взаимодействия – Internet Protocol (IP). К основным функциям протокола IP относятся:

- перенос между сетями различных типов адресной информации в унифицированной форме,
- сборка и разборка пакетов при передаче их между сетями с различным максимальным значением длины пакета.

Пакет IP состоит из заголовка и поля данных. Заголовок пакета имеет следующие поля:

- Поле Номер версии (VERS) указывает версию протокола IP. Сейчас повсеместно используется версия 4 и готовится переход на версию 6, называемую также IPng (IP next generation).
- Поле Длина заголовка (HLEN) пакета IP занимает 4 бита и указывает значение длины заголовка, измеренное в 32-битовых словах. Обычно заголовок имеет длину в 20 байт (пять 32-битовых слов), но при увеличении объема служебной информации эта длина может быть увеличена за счет использования дополнительных байт в поле Резерв (IP OPTIONS).
- Поле Тип сервиса (SERVICE TYPE) занимает 1 байт и задает приоритетность пакета и вид критерия выбора маршрута. Первые три

бита этого поля образуют подполе приоритета пакета (PRECEDENCE). Приоритет может иметь значения от 0 (нормальный пакет) до 7 (пакет управляющей информации). Маршрутизаторы и компьютеры могут принимать во внимание приоритет пакета и обрабатывать более важные пакеты в первую очередь. Поле Тип сервиса содержит также три бита, определяющие критерий выбора маршрута. Установленный бит D (delay) говорит о том, что маршрут должен выбираться для минимизации задержки доставки данного пакета, бит T – для максимизации пропускной способности, а бит R – для максимизации надежности доставки.

- Поле Общая длина (TOTAL LENGTH) занимает 2 байта и указывает общую длину пакета с учетом заголовка и поля данных.
- Поле Идентификатор пакета (IDENTIFICATION) занимает 2 байта и используется для распознавания пакетов, образовавшихся путем фрагментации исходного пакета. Все фрагменты должны иметь одинаковое значение этого поля.
- Поле Флаги (FLAGS) занимает 3 бита, оно указывает на возможность фрагментации пакета (установленный бит Do not Fragment – DF – запрещает маршрутизатору фрагментировать данный пакет), а также на то, является ли данный пакет промежуточным или последним фрагментом исходного пакета (установленный бит More Fragments – MF – говорит о том пакет переносит промежуточный фрагмент).
- Поле Смещение фрагмента (FRAGMENT OFFSET) занимает 13 бит, оно используется для указания в байтах смещения поля данных этого пакета от начала общего поля данных исходного пакета, подвергнутого фрагментации. Используется при сборке/разборке фрагментов пакетов при передачах их между сетями с различными величинами максимальной длины пакета.
- Поле Время жизни (TIME TO LIVE) занимает 1 байт и указывает предельный срок, в течение которого пакет может перемещаться по сети. Время жизни данного пакета измеряется в секундах и задается источником передачи средствами протокола IP. На шлюзах и в других узлах сети по истечении каждой секунды из текущего времени жизни вычитается единица; единица вычитается также при каждой транзитной передаче (даже если не прошла секунда). При истечении времени жизни пакет аннулируется.
- Идентификатор Протокола верхнего уровня (PROTOCOL) занимает 1 байт и указывает, какому протоколу верхнего уровня принадлежит пакет (например, это могут быть протоколы TCP, UDP или RIP).
- Контрольная сумма (HEADER CHECKSUM) занимает 2 байта, она рассчитывается по всему заголовку.

- Поля Адрес источника (SOURCE IP ADDRESS) и Адрес назначения (DESTINATION IP ADDRESS) имеют одинаковую длину – 32 бита, и одинаковую структуру.
- Поле Резерв (IP OPTIONS) является необязательным и используется обычно только при отладке сети. Это поле состоит из нескольких подполей, каждое из которых может быть одного из восьми predetermined типов. В этих подполях можно указывать точный маршрут прохождения маршрутизаторов, регистрировать проходимые пакетом маршрутизаторы, помещать данные системы безопасности, а также временные отметки. Так как число подполей может быть произвольным, то в конце поля Резерв должно быть добавлено несколько байт для выравнивания заголовка пакета по 32-битной границе.

Максимальная длина поля данных пакета ограничена разрядностью поля, определяющего эту величину, и составляет 65535 байтов, однако при передаче по сетям различного типа длина пакета выбирается с учетом максимальной длины пакета протокола нижнего уровня, несущего IP-пакеты. Если это кадры Ethernet, то выбираются пакеты с максимальной длиной в 1500 байтов, уместяющиеся в поле данных кадра Ethernet.

Протоколы транспортного уровня (протоколы TCP или UDP), пользующиеся сетевым уровнем для отправки пакетов, считают, что максимальный размер поля данных IP-пакета равен 65535, и поэтому могут передать ему сообщение такой длины для транспортировки через интернет. В функции уровня IP входит разбиение слишком длинного для конкретного типа составляющей сети сообщения на более короткие пакеты с созданием соответствующих служебных полей, нужных для последующей сборки фрагментов в исходное сообщение.

В большинстве типов локальных и глобальных сетей определяется такое понятие как максимальный размер поля данных кадра или пакета, в которые должен инкапсулировать свой пакет протокол IP. Эту величину обычно называют максимальной единицей транспортировки – *Maximum Transfer Unit, MTU*. Сети Ethernet имеют значение MTU, равное 1500 байт, сети FDDI – 4096 байт, а сети X.25 чаще всего работают с MTU в 128 байт.

Работа протокола IP по фрагментации пакетов в хостах и маршрутизаторах иллюстрируется рис. 8.4.

Пусть компьютер 1 связан с сетью, имеющей значение MTU в 4096 байтов, например, с сетью FDDI. При поступлении на IP-уровень компьютера 1 сообщения от транспортного уровня размером в 5600 байтов, протокол IP делит его на два IP-пакета, устанавливая в первом пакете признак фрагментации и присваивая пакету уникальный идентифика-

тор, например, 486. В первом пакете величина поля смещения равна 0, а во втором – 2800. Признак фрагментации во втором пакете равен нулю, что показывает, что это последний фрагмент пакета. Общая величина IP-пакета составляет 2800+20 (размер заголовка IP), то есть 2820 байтов, что уместается в поле данных кадра FDDI.

Далее компьютер 1 передает эти пакеты на канальный уровень К1, а затем и на физический уровень Ф1, который отправляет их маршрутизатору, связанному с данной сетью.

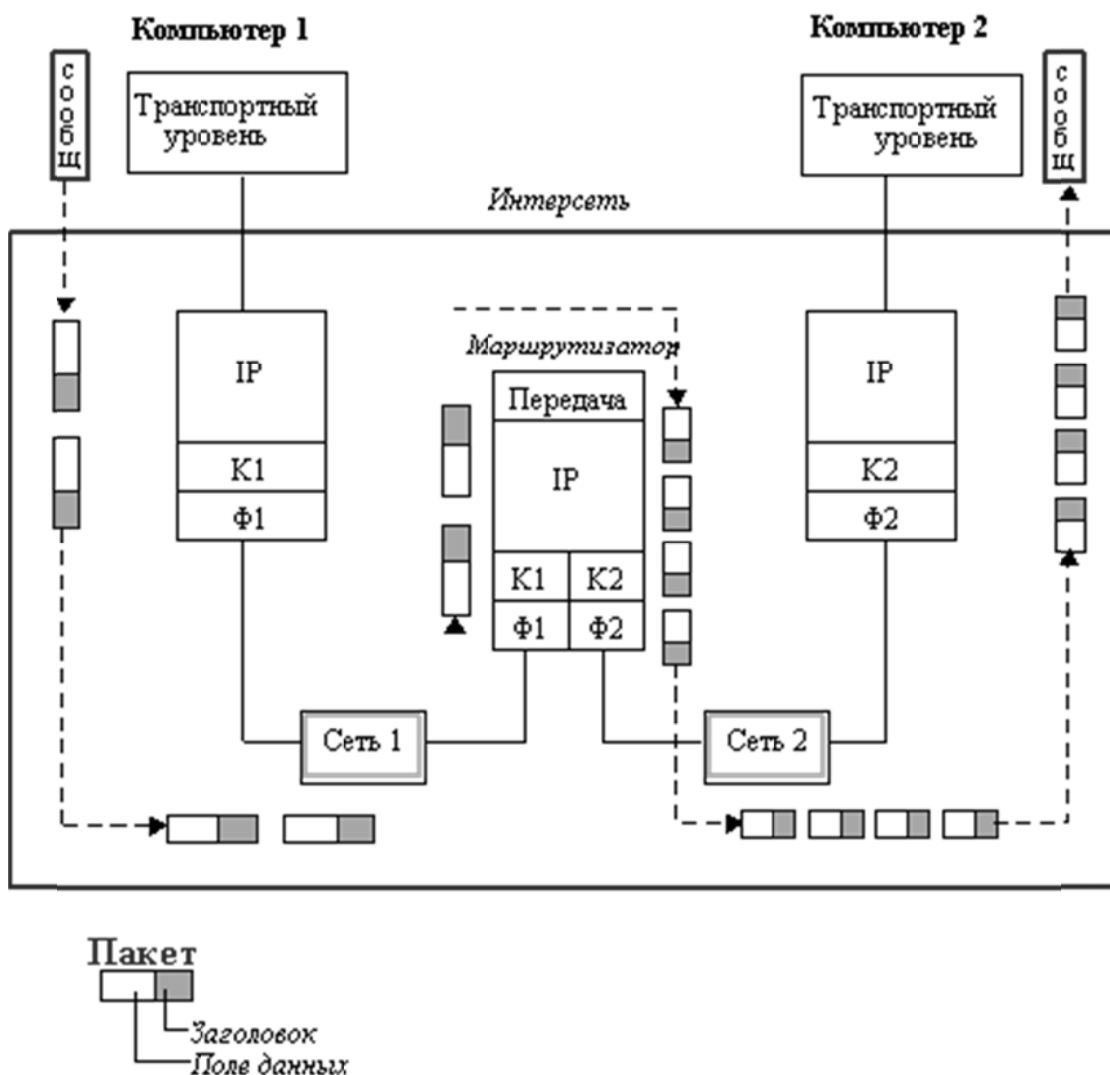


Рис. 8.4. Фрагментация IP-пакетов при передаче между сетями с разными максимальными размерами пакетов. К1 и Ф1 канальный и физический уровень сети 1, К2 и Ф2 канальный и физический уровень сети 2

Маршрутизатор видит по сетевому адресу, что прибывшие два пакета нужно передать в сеть 2, которая имеет меньшее значение MTU,

равное 1500. Вероятно, это сеть Ethernet. Маршрутизатор извлекает фрагмент транспортного сообщения из каждого пакета FDDI и делит его еще пополам, чтобы каждая часть уместилась в поле данных кадра Ethernet. Затем он формирует новые пакеты IP, каждый из которых имеет длину $1400 + 20 = 1420$ байтов, что меньше 1500 байтов, поэтому они нормально помещаются в поле данных кадров Ethernet.

В результате в компьютер 2 по сети Ethernet приходит четыре IP-пакета с общим идентификатором 486, что позволяет протоколу IP, работающему в компьютере 2, правильно собрать исходное сообщение. Если пакеты пришли не в том порядке, в котором были посланы, то смещение укажет правильный порядок их объединения.

Отметим, что IP-маршрутизаторы не собирают фрагменты пакетов в более крупные пакеты, даже если на пути встречается сеть, допускающая такое укрупнение. Это связано с тем, что отдельные фрагменты сообщения могут перемещаться по интернету по различным маршрутам, поэтому нет гарантии, что все фрагменты проходят через какой-либо промежуточный маршрутизатор на их пути.

При приходе первого фрагмента пакета узел назначения запускает таймер, который определяет максимально допустимое время ожидания прихода остальных фрагментов этого пакета. Если таймер истекает раньше прибытия последнего фрагмента, то все полученные к этому моменту фрагменты пакета отбрасываются, а в узел, пославший исходный пакет, направляется сообщение об ошибке с помощью протокола ICMP.

8.5. Маршрутизация с помощью IP-адресов

Рассмотрим теперь принципы, на основании которых в сетях IP происходит выбор маршрута передачи пакета между сетями.

Сначала необходимо обратить внимание на тот факт, что не только маршрутизаторы, но и конечные узлы – компьютеры – должны принимать участие в выборе маршрута. Пример, приведенный на рис. 8.5, демонстрирует эту необходимость. Здесь в локальной сети имеется несколько маршрутизаторов, и компьютер должен выбирать, какому из них следует отправить пакет.

Длина маршрута может существенно измениться в зависимости от того, какой маршрутизатор выберет компьютер для передачи своего пакета на сервер, расположенный, например, в Германии, если маршрутизатор 1 соединен выделенной линией с маршрутизатором в Копенгагене, а маршрутизатор 2 имеет спутниковый канал, соединяющий его с Токио.

В стеке TCP/IP маршрутизаторы и конечные узлы принимают решения о том, кому передавать пакет для его успешной доставки узлу

назначения, на основании так называемых таблиц маршрутизации (routing tables).

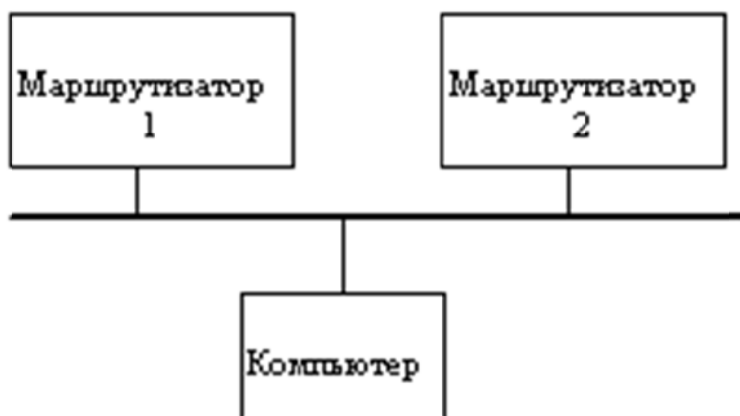


Рис. 8.5. Выбор маршрутизатора конечным узлом

Табл. 8.2 представляет собой типичный пример таблицы маршрутов, использующей IP-адреса сетей.

Таблица 8.2

Пример таблицы маршрутов

Адрес сети назначения	Адрес следующего маршрутизатора	Номер выходного порта Расстояние до сети назначения
56.0.0.0	198.21.17.7	1 20
56.0.0.0	213.34.12.4.	2 130
116.0.0.0	213.34.12.4	2 1450
129.13.0.0	198.21.17.6	1 50
198.21.17.0	–	2 0
213. 34.12.0	–	1 0
default	198.21.17.7	1 –

В табл. 8.2 в столбце «Адрес сети назначения» указываются адреса всех сетей, которым данный маршрутизатор может передавать пакеты. В стеке TCP/IP принят так называемый *одношаговый подход* к оптимизации маршрута продвижения пакета (next-hop routing) – каждый маршрутизатор и конечный узел принимает участие в выборе только одного шага передачи пакета. Поэтому в каждой строке таблицы маршрутизации указывается не весь маршрут в виде последовательности IP-адресов маршрутизаторов, через которые должен пройти пакет, а только один IP-адрес – адрес следующего маршрутизатора, которому нужно передать пакет. Вместе с пакетом следующему маршрутизатору передается ответственность за выбор следующего шага маршрутизации. Одноша-

говый подход к маршрутизации означает распределенное решение задачи выбора маршрута. Это снимает ограничение на максимальное количество транзитных маршрутизаторов на пути пакета.

(Альтернативой одношаговому подходу является указание в пакете всей последовательности маршрутизаторов, которые пакет должен пройти на своем пути. Такой подход называется маршрутизацией от источника – Source Routing. В этом случае выбор маршрута производится конечным узлом или первым маршрутизатором на пути пакета, а все остальные маршрутизаторы только отрабатывают выбранный маршрут, осуществляя коммутацию пакетов, то есть передачу их с одного порта на другой. Алгоритм Source Routing применяется в сетях IP только для отладки, когда маршрут задается в поле Резерв (IP OPTIONS) пакета.)

В случае, если в таблице маршрутов имеется более одной строки, соответствующей одному и тому же адресу сети назначения, то при принятии решения о передаче пакета используется та строка, в которой указано наименьшее значение в поле «Расстояние до сети назначения».

При этом под расстоянием понимается любая метрика, используемая в соответствии с заданным в сетевом пакете классом сервиса. Это может быть количество транзитных маршрутизаторов в данном маршруте (количество хопов от hop – прыжок), время прохождения пакета по линиям связи, надежность линий связи, или другая величина, отражающая качество данного маршрута по отношению к конкретному классу сервиса. Если маршрутизатор поддерживает несколько классов сервиса пакетов, то таблица маршрутов составляется и применяется отдельно для каждого вида сервиса (критерия выбора маршрута).

Для отправки пакета следующему маршрутизатору требуется знание его локального адреса, но в стеке TCP/IP в таблицах маршрутизации принято использование только IP-адресов для сохранения их универсального формата, не зависящего от типа сетей, входящих в интернет. Для нахождения локального адреса по известному IP-адресу необходимо воспользоваться протоколом ARP.

Конечный узел, как и маршрутизатор, имеет в своем распоряжении таблицу маршрутов унифицированного формата и на основании ее данных принимает решение, какому маршрутизатору нужно передавать пакет для сети N. Решение о том, что этот пакет нужно вообще маршрутизировать, компьютер принимает в том случае, когда он видит, что адрес сети назначения пакета отличается от адреса его собственной сети (каждому компьютеру при конфигурировании администратор присваивает его IP-адрес или несколько IP-адресов, если компьютер одновременно подключен к нескольким сетям). Когда компьютер выбрал следующий маршрутизатор, то он просматривают кэш-таблицу адресов

своего протокола ARP и, может быть, находит там соответствие IP-адреса следующего маршрутизатора его MAC-адресу. Если же нет, то по локальной сети передается широковещательный ARP-запрос и локальный адрес извлекается из ARP-ответа.

После этого компьютер формирует кадр протокола, используемого на выбранном порту, например, кадр Ethernet, в который помещает MAC-адрес маршрутизатора. Маршрутизатор принимает кадр Ethernet, извлекает из него пакет IP и просматривает свою таблицу маршрутизации для нахождения следующего маршрутизатора. При этом он выполняет те же действия, что и конечный узел.

Одношаговая маршрутизация обладает еще одним преимуществом – она позволяет сократить объем таблиц маршрутизации в конечных узлах и маршрутизаторах за счет использования в качестве номера сети назначения так называемого маршрута по умолчанию – *default*, который обычно занимает в таблице маршрутизации последнюю строку. Если в таблице маршрутизации есть такая запись, то все пакеты с номерами сетей, которые отсутствуют в таблице маршрутизации, передаются маршрутизатору, указанному в строке *default*. Поэтому маршрутизаторы часто хранят в своих таблицах ограниченную информацию о сетях интрасети, пересылая пакеты для остальных сетей в порт и маршрутизатор, используемые по умолчанию. Подразумевается, что маршрутизатор, используемый по умолчанию, передаст пакет на магистральную сеть, а маршрутизаторы, подключенные к магистрали, имеют полную информацию о составе интрасети.

Особенно часто приемом маршрутизации по умолчанию пользуются конечные узлы. Хотя они также в общем случае имеют в своем распоряжении таблицу маршрутизации, ее объем обычно незначителен, так как маршрутизация для компьютера – не основное занятие. Главная роль в маршрутизации пакетов в концепции протокола IP отводится, естественно, маршрутизаторам, которые должны обладать гораздо более полными таблицами маршрутизации, чем конечные узлы. Конечный узел часто вообще работает без таблицы маршрутизации, имея только сведения об IP-адресе маршрутизатора по умолчанию. При наличии одного маршрутизатора в локальной сети этот вариант – единственно возможный для всех конечных узлов. Но даже при наличии нескольких маршрутизаторов в локальной сети, когда проблема их выбора стоит перед конечным узлом, задание маршрута по умолчанию часто используется в компьютерах для сокращения объема их маршрутной таблицы.

Другим способом разгрузки компьютера от необходимости ведения больших таблиц маршрутизации является получение от маршрутизатора сведений о рациональном маршруте для какой-нибудь конкретной сети с помощью протокола ICMP.

Кроме маршрута default, в таблице маршрутизации могут встретиться два типа специальных записей – запись о специфичном для узла маршруте и запись об адресах сетей, непосредственно подключенных к портам маршрутизатора.

Специфичный для узла маршрут содержит вместо номера сети полный IP-адрес, то есть адрес, имеющий ненулевую информацию не только в поле номера сети, но и в поле номера узла. Предполагается, что для такого конечного узла маршрут должен выбираться не так, как для всех остальных узлов сети, к которой он относится. В случае, когда в таблице есть разные записи о продвижении пакетов для всей сети N и ее отдельного узла, имеющего адрес N,D, при поступлении пакета, адресованного узлу N,D, маршрутизатор отдаст предпочтение записи для N,D.

Записи в таблице маршрутизации, относящиеся к сетям, непосредственно подключенным к маршрутизатору, в поле «Расстояние до сети назначения» содержат нули.

Еще одним отличием работы маршрутизатора и конечного узла при выборе маршрута является способ построения таблицы маршрутизации. Если маршрутизаторы обычно автоматически создают таблицы маршрутизации, обмениваясь служебной информацией, то для конечных узлов таблицы маршрутизации создаются, как правило, вручную администраторами, и хранятся в виде постоянных файлов на дисках.

Существуют различные алгоритмы построения таблиц для одношаговой маршрутизации. Их можно разделить на три класса:

- алгоритмы фиксированной маршрутизации,
- алгоритмы простой маршрутизации,
- алгоритмы адаптивной маршрутизации.

Независимо от алгоритма, используемого для построения таблицы маршрутизации, результат их работы имеет единый формат. За счет этого в одной и той же сети различные узлы могут строить таблицы маршрутизации по своим алгоритмам, а затем обмениваться между собой недостающими данными, так как форматы этих таблиц фиксированы. Поэтому маршрутизатор, работающий по алгоритму адаптивной маршрутизации, может снабдить конечный узел, применяющий алгоритм фиксированной маршрутизации, сведениями о пути к сети, о которой конечный узел ничего не знает.

Фиксированная маршрутизация

Этот алгоритм применяется в сетях с простой топологией связей и основан на ручном составлении таблицы маршрутизации администратором сети. Алгоритм часто эффективно работает также для магистралей крупных сетей, так как сама магистраль может иметь простую

структуру с очевидными наилучшими путями следования пакетов в подсети, присоединенные к магистрали.

Различают одномаршрутные таблицы, в которых для каждого адресата задан один путь, и многомаршрутные таблицы, определяющие несколько альтернативных путей для каждого адресата. При использовании многомаршрутных таблиц должно быть задано правило выбора одного из них. Чаще всего один путь является основным, а остальные – резервными.

Простая маршрутизация

Алгоритмы простой маршрутизации подразделяются на три подкласса:

- Случайная маршрутизация – пакеты передаются в любом, случайном направлении, кроме исходного.
- Лавинная маршрутизация – пакеты передаются во всех направлениях, кроме исходного (применяется в мостах для пакетов с неизвестным адресом доставки).
- Маршрутизация по предыдущему опыту – таблицы маршрутов составляются на основании данных, содержащихся в проходящих через маршрутизатор пакетах. Именно так работают прозрачные мосты, собирая сведения об адресах узлов, входящих в сегменты сети. Такой способ маршрутизации обладает медленной адаптируемостью к изменениям топологии сети.

Адаптивная маршрутизация

Это основной вид алгоритмов маршрутизации, применяющихся маршрутизаторами в современных сетях со сложной топологией. Адаптивная маршрутизация основана на том, что маршрутизаторы периодически обмениваются специальной топологической информацией об имеющихся в интересах сетей, а также о связях между маршрутизаторами. Обычно учитывается не только топология связей, но и их пропускная способность и состояние.

Адаптивные протоколы позволяют всем маршрутизаторам собирать информацию о топологии связей в сети, оперативно отрабатывая все изменения конфигурации связей. Эти протоколы имеют распределенный характер, который выражается в том, что в сети отсутствуют какие-либо выделенные маршрутизаторы, которые бы собирали и обобщали топологическую информацию: эта работа распределена между всеми маршрутизаторами.

Рассмотрим на примере интереса, приведенной на рис. 8.6, каким образом происходит взаимодействие компьютеров через маршрутизаторы и доставка пакетов компьютеру назначения.

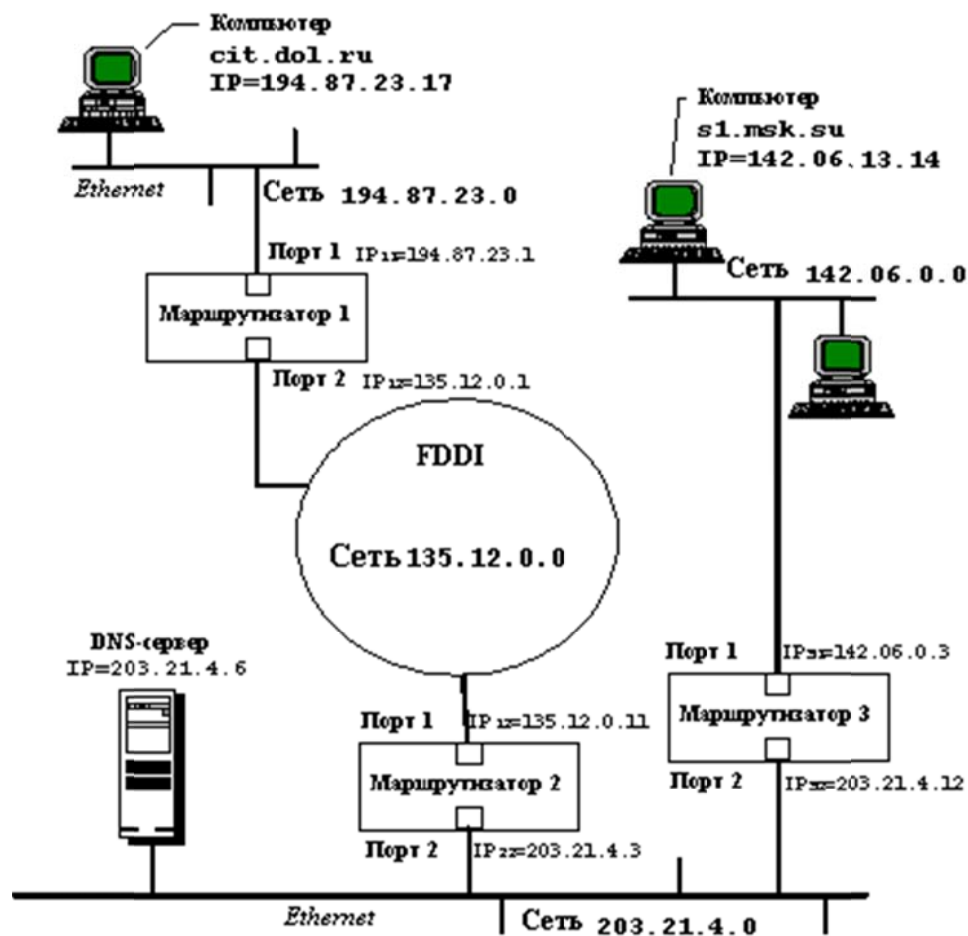


Рис. 8.6. Пример взаимодействия компьютеров через интернет

Пусть в приведенном примере пользователь компьютера cit.dol.ru, находящийся в сети Ethernet с IP-адресом 194.87.23.0 (адрес класса C), хочет взаимодействовать по протоколу FTP с компьютером s1.msk.su, принадлежащем сети Ethernet с IP-адресом 142.06.0.0 (адрес класса B). Компьютер cit.dol.ru имеет IP-адрес 194.87.23.1.17, а компьютер s1.msk.su – IP-адрес 142.06.13.14.

1. Пользователь компьютера cit.dol.ru знает символьное имя компьютера s1.msk.su, но не знает его IP-адреса, поэтому он набирает команду `> ftp s1.msk.su` для организации ftp-сеанса.

В компьютере cit.dol.ru должны быть заданы некоторые параметры для стека TCP/IP, чтобы он мог выполнить поставленную перед ним задачу.

В число этих параметров должны входить собственный IP-адрес, IP-адрес DNS-сервера и IP-адрес маршрутизатора по умолчанию. Так как к сети Ethernet, к которой относится компьютер cit.dol.ru, подключен только один маршрутизатор, то таблица маршрутизации конечным

узлам этой сети не нужна, достаточно знать IP-адрес маршрутизатора по умолчанию. В данном примере он равен 194.87.23.1.

Так как пользователь в команде `ftp` не задал IP-адрес узла, с которым он хочет взаимодействовать, то стек TCP/IP должен определить его самостоятельно. Он может сделать запрос к серверу DNS по имеющемуся у него IP-адресу, но обычно каждый компьютер сначала просматривает свою собственную таблицу соответствия символьных имен и IP-адресов. Такая таблица хранится чаще всего в виде текстового файла простой структуры – каждая его строка содержит запись об одном символьном имени и его IP-адресе. В ОС Unix такой файл традиционно носит имя `HOSTS`.

2. Будем считать, что компьютер `cit.dol.ru` имеет файл `HOSTS`, а в нем есть строка

```
142.06.13.14 s1.msk.su.
```

Поэтому разрешение имени выполняется локально, так что протокол IP может теперь формировать IP-пакеты с адресом назначения 142.06.13.14 для взаимодействия с компьютером `s1.msk.su`.

3. Протокол IP компьютера `cit.dol.ru` проверяет, нужно ли маршрутизировать пакеты для адреса 142.06.13.14. Так как адрес сети назначения равен 142.06.0.0, а адрес сети, к которой принадлежит компьютер, равен 194.87.23.0, то маршрутизация необходима.

4. Компьютер `cit.dol.ru` начинает формировать кадр Ethernet для отправки IP-пакета маршрутизатору по умолчанию с IP-адресом 194.87.23.1. Для этого ему нужен MAC-адрес порта маршрутизатора, подключенного к его сети. Этот адрес скорее всего уже находится в кэш-таблице протокола ARP компьютера, если он хотя бы раз за последнее включение обменивался данными с компьютерами других сетей. Пусть этот адрес в нашем примере был найден именно в кэш-памяти. Обозначим его MAC_{11} , в соответствии с номером маршрутизатора и его порта.

5. В результате компьютер `cit.dol.ru` отправляет по локальной сети кадр Ethernet, имеющий следующие поля:

DA (Ethernet)	...	DESTINATION IP
MAC_{11}		142.06.13.14	

6. Кадр принимается портом 1 маршрутизатора 1 в соответствии с протоколом Ethernet, так как MAC-узел этого порта распознает свой адрес MAC_{11} . Протокол Ethernet извлекает из этого кадра IP-пакет и передает его программному обеспечению маршрутизатора, реализующему протокол IP. Протокол IP извлекает из пакета адрес назначения и про-

сма­три­ва­ет за­пи­си сво­ей та­б­ли­цы мар­шру­ти­за­ции. Пу­сть мар­шру­ти­за­тор 1 име­ет в сво­ей та­б­ли­це мар­шру­ти­за­ции за­пись

142.06.0.0 135.12.0.11 2 1,

ко­то­рая го­во­рит о том, что па­ке­ты для се­ти 142.06. 0.0 нуж­но пе­ре­да­вать мар­шру­ти­за­то­ру 135.12.0.11, под­клю­чен­но­му к той же се­ти, что и порт 2 мар­шру­ти­за­то­ра 1.

7. Мар­шру­ти­за­тор 1 про­сма­три­ва­ет па­ра­мет­ры пор­та 2 и на­хо­дит, что он под­клю­чен к се­ти FDDI. Так как се­ть FDDI име­ет зна­че­ние мак­си­маль­но­го транс­пор­ти­ру­е­мо­го бло­ка MTU боль­ше, чем се­ть Ethernet, то фраг­мен­та­ция по­ля дан­ных IP-па­ке­та не тре­бу­ет­ся. По­это­му мар­шру­ти­за­тор 1 фор­ми­ру­ет кадр фор­ма­та FDDI, в ко­то­ром ука­зы­ва­ет MAC-а­дрес пор­та мар­шру­ти­за­то­ра 2, ко­то­рый он на­хо­дит в сво­ей кэш-та­б­ли­це про­то­ко­ла ARP:

DA (FDDI)	...	DESTINATION IP
MAC ₂₁		142.06.13.14	

8. Ана­ло­гич­но дей­ст­ву­ет мар­шру­ти­за­тор 2, фор­ми­ру­я кадр Ethernet для пе­ре­да­чи па­ке­та мар­шру­ти­за­то­ру 3 по се­ти Ethernet с IP-а­дресом 203.21.4.0:

DA (Ethernet)	...	DESTINATION IP
MAC ₃₂		142.06.13.14	

9. На­ко­нец, по­сле то­го, как па­ке­т по­сту­пил в мар­шру­ти­за­тор се­ти на­зна­че­ния – мар­шру­ти­за­тор 3, по­яв­ля­ет­ся воз­мож­ность пе­ре­да­чи это­го па­ке­та ком­пью­те­ру на­зна­че­ния. Мар­шру­ти­за­тор 3 ви­дит, что па­ке­т нуж­но пе­ре­дать в се­ть 142.06.0.0, ко­то­рая не­по­сред­ст­вен­но под­клю­че­на к его пер­во­му пор­ту. По­это­му он по­сы­ла­ет ARP-за­прос по се­ти Ethernet с IP-а­дресом ком­пью­те­ра s1.msk.su (счи­та­ем, что этой ин­фор­ма­ции в его кэш­е нет), по­лу­ча­ет от­вет, со­дер­жа­щий а­дрес MAC_{s1}, и фор­ми­ру­ет кадр Ethernet, до­ста­в­ля­ю­щий IP-па­ке­т по ло­каль­ной се­ти а­дреса­ту.

DA (Ethernet)	...	DESTINATION IP
MAC _{s1}		142.06.13.14	

8.6. Структуризация сетей IP с помощью масок

Час­то ад­ми­ни­ст­ра­то­ры се­тей ис­пы­ты­ва­ют не­удоб­ст­ва, из-за то­го, что ко­ли­че­ст­во цен­тра­ли­зо­ва­но вы­де­лен­ных им но­ме­ров се­тей не­до­статоч­но для то­го, что­бы струк­ту­ри­ро­вать се­ть над­ле­жа­щим об­ра­зом,

например, разместить все слабо взаимодействующие компьютеры по разным сетям.

В такой ситуации возможны два пути. Первый из них связан с получением от NIC дополнительных номеров сетей. Второй способ, употребляющийся более часто, связан с использованием так называемых *масок*, которые позволяют разделять одну сеть на несколько сетей.

Маска – это число, двоичная запись которого содержит единицы в тех разрядах, которые должны интерпретироваться как номер сети.

Например, для стандартных классов сетей маски имеют следующие значения:

- 255.0.0.0 – маска для сети класса А,
- 255.255.0.0 – маска для сети класса В,
- 255.255.255.0 – маска для сети класса С.

В масках, которые использует администратор для увеличения числа сетей, количество единиц в последовательности, определяющей границу номера сети, не обязательно должно быть кратным 8, чтобы повторять деление адреса на байты.

Пусть, например, маска имеет значение 255.255.192.0 (11111111 11111111 11000000 00000000). И пусть сеть имеет номер 129.44.0.0 (10000001 00101100 00000000 00000000), из которого видно, что она относится к классу В. После наложения маски на этот адрес число разрядов, интерпретируемых как номер сети, увеличилось с 16 до 18, то есть администратор получил возможность использовать вместо одного, централизованно заданного ему номера сети, четыре:

- 129.44.0.0 (10000001 00101100 00000000 00000000)
- 129.44.64.0 (10000001 00101100 01000000 00000000)
- 129.44.128.0 (10000001 00101100 10000000 00000000)
- 129.44.192.0 (10000001 00101100 11000000 00000000)

Например, IP-адрес 129.44.141.15 (10000001 00101100 10001101 00001111), который по стандартам IP задает номер сети 129.44.0.0 и номер узла 0.0.141.15, теперь, при использовании маски, будет интерпретироваться как пара:

- 129.44.128.0 – номер сети, 0.0. 13.15 – номер узла.

Таким образом, установив новое значение маски, можно заставить маршрутизатор по-другому интерпретировать IP-адрес. При этом два дополнительных последних бита номера сети часто интерпретируются как номера подсетей.

Еще один пример. Пусть некоторая сеть относится к классу В и имеет адрес 128.10.0.0 (рис. 8.7). Этот адрес используется маршрутизатором, соединяющим сеть с остальной частью интерсети. И пусть среди всех станций сети есть станции, слабо взаимодействующие между со-

бой. Их желательно было бы изолировать в разных сетях. Для этого сеть можно разделить на две сети, подключив их к соответствующим портам маршрутизатора, и задать для этих портов в качестве маски, например, число 255.255.255.0, то есть организовать внутри исходной сети с централизовано заданным номером две подсети класса С (можно было бы выбрать и другой размер для поля адреса подсети). Извне сеть по-прежнему будет выглядеть, как единая сеть класса В, а на местном уровне это будут две отдельные сети класса С. Приходящий общий трафик будет разделяться местным маршрутизатором между подсетями.

Необходимо заметить, что, если принимается решение об использовании механизма масок, то соответствующим образом должны быть сконфигурированы и маршрутизаторы, и компьютеры сети.

Протокол доставки пользовательских дейтаграмм UDP.

Задачей протокола транспортного уровня *UDP (User Datagram Protocol)* является передача данных между прикладными процессами без гарантий доставки, поэтому его пакеты могут быть потеряны, продублированы или прийти не в том порядке, в котором они были отправлены.

Зарезервированные и доступные порты UDP.

В то время, как задачей сетевого уровня является передача данных между произвольными узлами сети, задача транспортного уровня заключается в передаче данных между любыми *прикладными процессами*, выполняющимися на любых узлах сети. Действительно, после того, как пакет средствами протокола IP доставлен в компьютер-получатель, данные необходимо направить конкретному процессу-получателю. Каждый компьютер может выполнять несколько процессов, более того, прикладной процесс тоже может иметь несколько точек входа, выступающих в качестве адреса назначения для пакетов данных.

Пакеты, поступающие на транспортный уровень, организуются операционной системой в виде множества очередей к точкам входа различных прикладных процессов. В терминологии TCP/IP такие системные очереди называются *портами*. Таким образом, адресом назначения, который используется на транспортном уровне, является идентификатор (номер) порта прикладного сервиса. Номер порта, задаваемый транспортным уровнем, в совокупности с номером сети и номером компьютера, задаваемыми сетевым уровнем, однозначно определяют прикладной процесс в сети.

Назначение номеров портов прикладным процессам осуществляется либо централизованно, если эти процессы представляют собой популярные общедоступные сервисы, типа сервиса удаленного доступа к файлам TFTP (Trivial FTP) или сервиса удаленного управления telnet, либо локально для тех сервисов, которые еще не стали столь распро-

страненными, чтобы за ними закреплять стандартные (зарезервированные) номера.

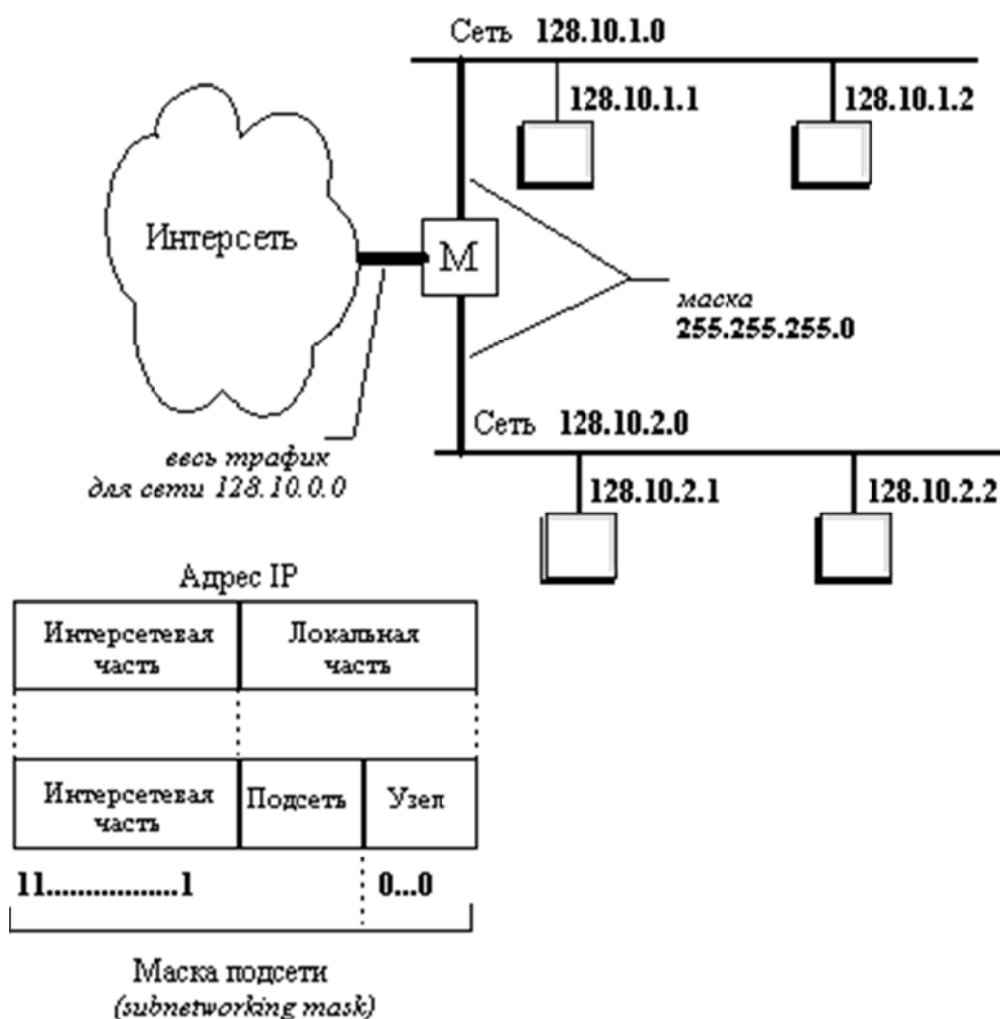


Рис. 8.7. Пример использования масок для структурирования сети

Централизованное присвоение сервисам номеров портов выполняется организацией *Internet Assigned Numbers Authority*. Эти номера затем закрепляются и публикуются в стандартах Internet. Например, упомятому выше сервису удаленного доступа к файлам TFTP присвоен стандартный номер порта 69.

Локальное присвоение номера порта заключается в том, что разработчик некоторого приложения просто связывает с ним любой доступный, произвольно выбранный числовой идентификатор, обращая внимание на то, чтобы он не входил в число зарезервированных номеров портов. В дальнейшем все удаленные запросы к данному приложению от других приложений должны адресоваться с указанием назначенного ему номера порта.

8.7. Сегменты ТСР

Единицей данных протокола ТСР является сегмент. Информация, поступающая к протоколу ТСР в рамках логического соединения от протоколов более высокого уровня, рассматривается протоколом ТСР как неструктурированный поток байт. Поступающие данные буферизуются средствами ТСР. Для передачи на сетевой уровень из буфера «вырезается» некоторая непрерывная часть данных, называемая сегментом.

В протоколе ТСР предусмотрен случай, когда приложение обращается с запросом о срочной передаче данных (бит PSH в запросе установлен в 1). В этом случае протокол ТСР, не ожидая заполнения буфера до уровня размера сегмента, немедленно передает указанные данные в сеть. О таких данных говорят, что они передаются вне потока – *out of band*.

Не все сегменты, посланные через соединение, будут одного и того же размера, однако оба участника соединения должны договориться о максимальном размере сегмента, который они будут использовать. Этот размер выбирается таким образом, чтобы при упаковке сегмента в IP-пакет он помещался туда целиком, то есть максимальный размер сегмента не должен превосходить максимального размера поля данных IP-пакета. В противном случае пришлось бы выполнять фрагментацию, то есть делить сегмент на несколько частей, для того, чтобы он влез в IP-пакет.

Аналогичные проблемы решаются и на сетевом уровне. Для того, чтобы избежать фрагментации, должен быть выбран соответствующий максимальный размер IP-пакета. Однако при этом должны быть приняты во внимание максимальные размеры поля данных кадров (MTU) всех протоколов канального уровня, используемых в сети. Максимальный размер сегмента не должен превышать минимальное значение на множестве всех MTU составной сети.

8.8. Порты и установление ТСР-соединений

В протоколе ТСР также, как и в UDP, для связи с прикладными процессами используются порты. Номера портам присваиваются аналогичным образом: имеются стандартные, зарезервированные номера (например, номер 21 закреплен за сервисом FTP, 23 – за telnet), а менее известные приложения пользуются произвольно выбранными локальными номерами.

Однако в протоколе ТСР порты используются несколько иным способом. Для организации надежной передачи данных предусматривается установление *логического соединения* между двумя прикладными процессами. В рамках соединения осуществляется обязательное подтверждение правильности приема для всех переданных сообщений, и при

необходимости выполняется повторная передача. Соединение в ТСР позволяет вести передачу данных одновременно в обе стороны, то есть полнодуплексную передачу.

Соединение в протоколе ТСР идентифицируется парой полных адресов обоим взаимодействующим процессам (оконечных точек). Адрес каждой из оконечных точек включает IP-адрес (номер сети и номер компьютера) и номер порта. Одна оконечная точка может участвовать в нескольких соединениях.

Установление соединения выполняется в следующей последовательности:

- При установлении соединения одна из сторон является инициатором. Она посылает запрос к протоколу ТСР на открытие порта для передачи (active open).
- После открытия порта протокол ТСР на стороне процесса-инициатора посылает запрос процессу, с которым требуется установить соединение.
- Протокол ТСР на приемной стороне открывает порт для приема данных (passive open) и возвращает квитанцию, подтверждающую прием запроса.
- Для того чтобы передача могла вестись в обе стороны, протокол на приемной стороне также открывает порт для передачи (active port) и также передает запрос к противоположной стороне.
- Сторона-инициатор открывает порт для приема и возвращает квитанцию. Соединение считается установленным. Далее происходит обмен данными в рамках данного соединения.

8.9. Концепция квитирования

В рамках соединения правильность передачи каждого сегмента должна подтверждаться квитанцией получателя. *Квитирование* – это один из традиционных методов обеспечения надежной связи. Идея квитирования состоит в следующем.

Для того, чтобы можно было организовать повторную передачу искаженных данных отправитель нумерует отправляемые единицы передаваемых данных (далее для простоты называемые кадрами). Для каждого кадра отправитель ожидает от приемника так называемую положительную квитанцию – служебное сообщение, извещающее о том, что исходный кадр был получен и данные в нем оказались корректными. Время этого ожидания ограничено – при отправке каждого кадра передатчик запускает таймер, и если по его истечению положительная квитанция не получена, то кадр считается утерянным. В некоторых протоколах приемник, в случае получения кадра с искаженными данными

должен отправить отрицательную квитанцию – явное указание того, что данный кадр нужно передать повторно.

Существуют два подхода к организации процесса обмена положительными и отрицательными квитанциями: с простоями и с организацией «окна».

Метод с простоями требует, чтобы источник, пославший кадр, ожидал получения квитанции (положительной или отрицательной) от приемника и только после этого посылал следующий кадр (или повторял искаженный). Из рис. 8.8 видно, что в этом случае производительность обмена данными существенно снижается – хотя передатчик и мог бы послать следующий кадр сразу же после отправки предыдущего, он обязан ждать прихода квитанции. Снижение производительности для этого метода коррекции особенно заметно на низкоскоростных каналах связи, то есть в территориальных сетях.

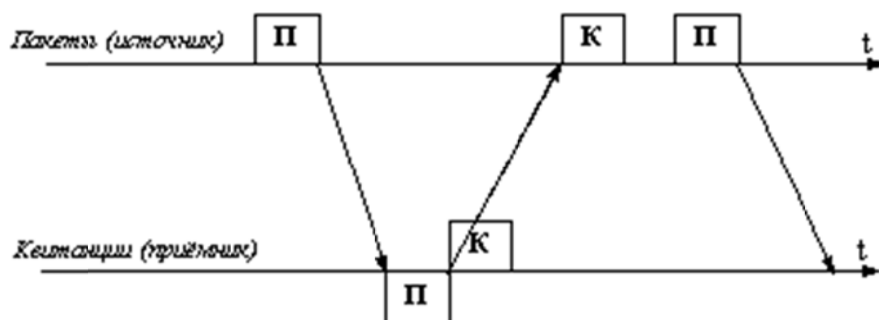


Рис. 8.8. Метод подтверждения корректности передачи кадров

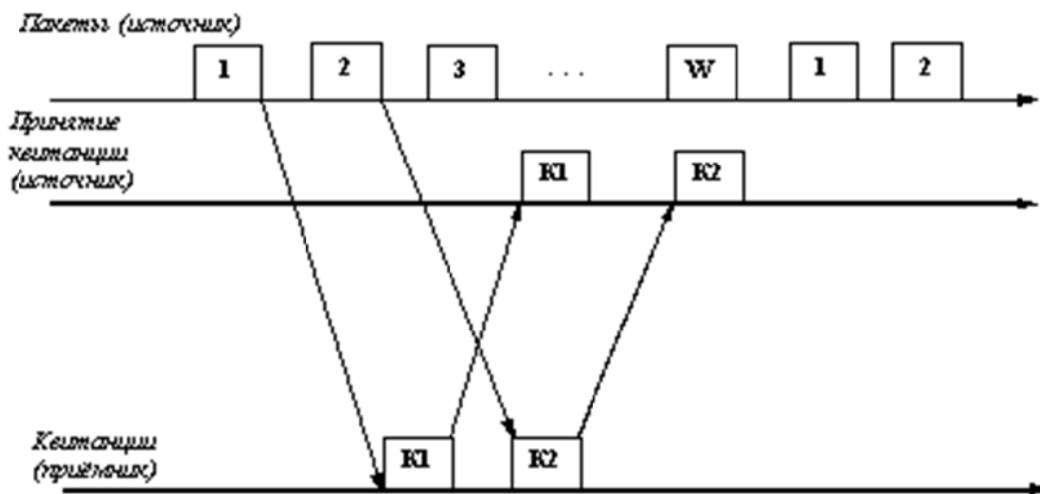


Рис. 8.9. Метод «окна» – непрерывная отправка пакетов

Во втором методе для повышения коэффициента использования линии источнику разрешается передать некоторое количество кадров в непрерывном режиме, то есть в максимально возможном для источника темпе, без получения на эти кадры ответных квитанций. Количество

кадров, которые разрешается передавать таким образом, называется размером окна. Рис. 8.9 иллюстрирует данный метод для размера окна в W кадров. Обычно кадры при обмене нумеруются циклически, от 1 до W . При отправке кадра с номером 1 источнику разрешается передать еще $W-1$ кадров до получения квитанции на кадр 1. Если же за это время квитанция на кадр 1 так и не пришла, то процесс передачи приостанавливается, и по истечению некоторого тайм-аута кадр 1 считается утерянным (или квитанция на него утеряна) и он передается снова.

Если же поток квитанций поступает более-менее регулярно, в пределах допуска в W кадров, то скорость обмена достигает максимально возможной величины для данного канала и принятого протокола.

Этот алгоритм называют алгоритмом скользящего окна. Действительно, при каждом получении квитанции окно перемещается (скользит), захватывая новые данные, которые разрешается передавать без подтверждения.

8.10. Формат сообщений TCP

Сообщения протокола TCP называются сегментами и состоят из заголовка и блока данных. Заголовок сегмента имеет следующие поля:

- Порт источника (SOURCE PORT) занимает 2 байта, идентифицирует процесс-отправитель;
- Порт назначения (DESTINATION PORT) занимает 2 байта, идентифицирует процесс-получатель;
- Последовательный номер (SEQUENCE NUMBER) занимает 4 байта, указывает номер байта, который определяет смещение сегмента относительно потока отправляемых данных;
- Подтвержденный номер (ACKNOWLEDGEMENT NUMBER) занимает 4 байта, содержит максимальный номер байта в полученном сегменте, увеличенный на единицу; именно это значение используется в качестве квитанции;
- Длина заголовка (HLEN) занимает 4 бита, указывает длину заголовка сегмента TCP, измеренную в 32-битовых словах. Длина заголовка не фиксирована и может изменяться в зависимости от значений, устанавливаемых в поле Опции;
- Резерв (RESERVED) занимает 6 битов, поле зарезервировано для последующего использования;
- Кодовые биты (CODE BITS) занимают 6 битов, содержат служебную информацию о типе данного сегмента, задаваемую установкой в единицу соответствующих бит этого поля:
- URG – срочное сообщение;
- ACK – квитанция на принятый сегмент;

- PSH – запрос на отправку сообщения без ожидания заполнения буфера;
- RST – запрос на восстановление соединения;
- SYN – сообщение используемое для синхронизации счетчиков переданных данных при установлении соединения;
- FIN – признак достижения передающей стороной последнего байта в потоке передаваемых данных.
- Окно (WINDOW) занимает 2 байта, содержит объявляемое значение размера окна в байтах;
- Контрольная сумма (CHECKSUM) занимает 2 байта, рассчитывается по сегменту;
- Указатель срочности (URGENT POINTER) занимает 2 байта, используется совместно с кодовым битом URG, указывает на конец данных, которые необходимо срочно принять, несмотря на переполнение буфера;
- Опции (OPTIONS) – это поле имеет переменную длину и может вообще отсутствовать, максимальная величина поля 3 байта; используется для решения вспомогательных задач, например, при выборе максимального размера сегмента;
- Заполнитель (PADDING) может иметь переменную длину, представляет собой фиктивное поле, используемое для доведения размера заголовка до целого числа 32-битовых слов.

8.11. Адресация в IPv6

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

- Unicast – индивидуальный адрес. Определяет отдельный узел – компьютер или порт маршрутизатора. Пакет должен быть доставлен узлу по кратчайшему маршруту.
- Cluster – адрес кластера. Обозначает группу узлов, которые имеют общий адресный префикс (например, присоединенных к одной физической сети). Пакет должен быть маршрутизирован группе узлов по кратчайшему пути, а затем доставлен только одному из членов группы (например, ближайшему узлу).
- Multicast – адрес набора узлов, возможно в различных физических сетях. Копии пакета должны быть доставлены каждому узлу набора, используя аппаратные возможности групповой или широковещательной доставки, если это возможно.

Как и в версии IPv4, адреса в версии IPv6 делятся на классы, в зависимости от значения нескольких старших бит адреса.

Большая часть классов зарезервирована для будущего применения. Наиболее интересным для практического использования является класс, предназначенный для провайдеров услуг Internet, названный *Provider-Assigned Unicast*.

Адрес этого класса имеет следующую структуру:

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

Каждому провайдеру услуг Internet назначается уникальный идентификатор, которым помечаются все поддерживаемые им сети. Далее провайдер назначает своим абонентам уникальные идентификаторы, и использует оба идентификатора при назначении блока адресов абонента. Абонент сам назначает уникальные идентификаторы своим подсетям и узлам этих сетей.

Абонент может использовать технику подсетей, применяемую в версии IPv4, для дальнейшего деления поля идентификатора подсети на более мелкие поля.

Описанная схема приближает схему адресации IPv6 к схемам, используемым в территориальных сетях, таких как телефонные сети или сети X.25. Иерархия адресных полей позволит магистральным маршрутизаторам работать только со старшими частями адреса, оставляя обработку менее значимых полей маршрутизаторам абонентов.

Под поле идентификатора узла требуется выделения не менее 6 байт, для того чтобы можно было использовать в IP-адресах MAC-адреса локальных сетей непосредственно.

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

ГЛАВА 9. ПРИНЦИП ПОСТРОЕНИЯ КОМПЬЮТЕРНЫХ СЕТЕЙ

9.1. Типы компьютерных сетей

Компьютерная сеть – это совокупность компьютеров и различных устройств, обеспечивающих информационный обмен между компьютерами в сети без использования каких-либо промежуточных носителей информации.

Все многообразие компьютерных сетей можно классифицировать по группе признаков:

- 1) территориальная распространенность;
- 2) ведомственная принадлежность;
- 3) скорость передачи информации;
- 4) тип среды передачи.

По территориальной распространенности сети могут быть локальными, глобальными, и региональными. Локальные – это сети, перекрывающие территорию не более 10 м², региональные – расположенные на территории города или области, глобальные на территории государства или группы государств, например, всемирная сеть Internet.

По принадлежности различают ведомственные и государственные сети. Ведомственные принадлежат одной организации и располагаются на ее территории. Государственные сети – сети, используемые в государственных структурах.

По скорости передачи информации компьютерные сети делятся на низко-, средне- и высокоскоростные.

По типу среды передачи разделяются на сети коаксиальные, на витой паре, оптоволоконные, с передачей информации по радиоканалам, в инфракрасном диапазоне.

Компьютеры могут соединяться кабелями, образуя различную топологию сети (звездная, шинная, кольцевая и др.).

Следует различать компьютерные сети и сети терминалов (терминальные сети). Компьютерные сети связывают компьютеры, каждый из которых может работать и автономно. Терминальные сети обычно связывают мощные компьютеры (майнфреймы), а в отдельных случаях и ПК с устройствами (терминалами), которые могут быть достаточно сложны, но вне сети их работа или невозможна, или вообще теряет смысл. Например, сеть банкоматов или касс по продаже авиабилетов. Строятся они на совершенно иных, чем компьютерные сети, принципах и даже на другой вычислительной технике.

В классификации сетей существует два основных термина: LAN и WAN.

LAN (Local Area Network) – локальные сети, имеющие замкнутую инфраструктуру до выхода на поставщиков услуг. Термин «LAN» может описывать и маленькую офисную сеть, и сеть уровня большого завода, занимающего несколько сотен гектаров. Зарубежные источники дают даже близкую оценку – около шести миль (10 км) в радиусе; использование высокоскоростных каналов.

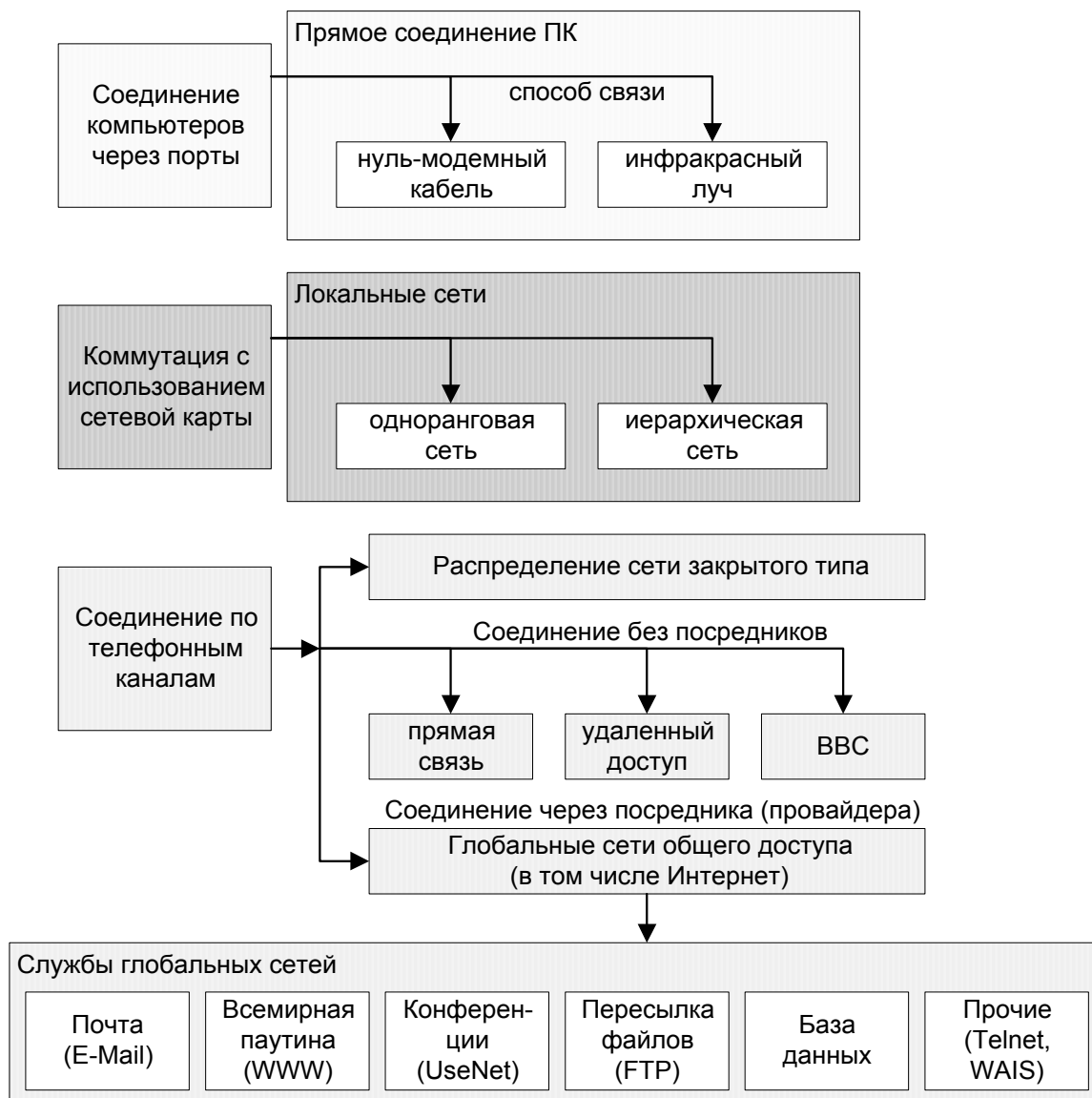


Рис. 9.1. Способы коммутации компьютеров и виды сетей

WAN (Wide Area Network) – глобальная сеть, покрывающая большие географические регионы, включающие в себя как локальные сети, так и прочие телекоммуникационные сети и устройства. Пример WAN –

сети с коммутацией пакетов (Frame Relay), через которую могут «разговаривать» между собой различные компьютерные сети.

Термин «корпоративная сеть» также используется в литературе для обозначения объединения нескольких сетей, каждая из которых может быть построена на различных технических, программных и информационных принципах.

Рассмотренные выше виды сетей являются сетями закрытого типа, доступ к ним разрешен только ограниченному кругу пользователей, для которых работа в такой сети непосредственно связана с их профессиональной деятельностью. Глобальные сети ориентированы на обслуживание любых пользователей.

На рис. 9.1 показаны способы коммутации компьютеров и виды сетей.

9.2. Локальные компьютерные сети (ЛКС)

9.2.1. Классификация ЛКС

Локальные вычислительные сети подразделяются на два кардинально различающихся класса: одноранговые (одноуровневые или Peer to Peer) сети и иерархические (многоуровневые).

Одноранговая сеть представляет собой сеть равноправных компьютеров, каждый из которых имеет уникальное имя (имя компьютера) и обычно пароль для входа в него во время загрузки ОС. Имя и пароль входа назначаются владельцем ПК средствами ОС. Одноранговые сети могут быть организованы с помощью таких операционных систем, как LANtastic, Windows'3.11, Novell NetWare Lite. Указанные программы работают как с DOS, так и с Windows. Одноранговые сети могут быть организованы также на базе всех современных 32-разрядных операционных систем – Windows'95 OSR2, Windows NT Workstation версии, OS/2) и некоторых других.

В иерархических локальных сетях имеется один или несколько специальных компьютеров – серверов, на которых хранится информация, совместно используемая различными пользователями.

Сервер в иерархических сетях – это постоянное хранилище разделяемых ресурсов. Сам сервер может быть клиентом только сервера более высокого уровня иерархии. Поэтому иерархические сети иногда называются сетями с выделенным сервером. Серверы обычно представляют собой высокопроизводительные компьютеры, возможно, с несколькими параллельно работающими процессорами, с винчестерами большой емкости, с высокоскоростной сетевой картой (100 Мбит/с и более). Компьютеры, с которых осуществляется доступ к информации на сервере, называются станциями или клиентами.

ЛКС классифицируются по назначению:

- Сети терминального обслуживания. В них включается ЭВМ и периферийное оборудование, используемое в монопольном режиме компьютером, к которому оно подключается, или быть общесетевым ресурсом.
- Сети, на базе которых построены системы управления производством и учрежденческой деятельности. Они объединяются группой стандартов MAP/TOP. В MAP описываются стандарты, используемые в промышленности. TOP описывают стандарты для сетей, применяемых в офисных сетях.
- Сети, которые объединяют системы автоматизации, проектирования. Рабочие станции таких сетей обычно базируются на достаточно мощных персональных ЭВМ, например фирмы Sun Microsystems.
- Сети, на базе которых построены распределенные вычислительные системы.

По классификационному признаку локальные компьютерные сети делятся на кольцевые, шинные, звездообразные, древовидные;

по признаку скорости – на низкоскоростные (до 10 Мбит/с), среднескоростные (до 100 Мбит/с), высокоскоростные (свыше 100 Мбит/с);

по типу метода доступа – на случайные, пропорциональные, гибридные;

по типу физической среды передачи – на витую пару, коаксиальный или оптоволоконный кабель, инфракрасный канал, радиоканал.

9.2.2. Структура ЛКС

Способ соединения компьютеров называется структурой или топологией сети. Сети Ethernet могут иметь топологию «шина» и «звезда». В первом случае все компьютеры подключены к одному общему кабелю (шине), во втором – имеется специальное центральное устройство (хаб), от которого идут «лучи» к каждому компьютеру, т. е. каждый компьютер подключен к своему кабелю.

Структура типа «шина», рис. 9.2, *а*, проще и экономичнее, так как для нее не требуется дополнительное устройство и расходуется меньше кабеля. Но она очень чувствительна к неисправностям кабельной системы. Если кабель поврежден хотя бы в одном месте, то возникают проблемы для всей сети. Место неисправности трудно обнаружить.

В этом смысле «звезда», рис. 9.2, *б*, более устойчива. Поврежденный кабель – проблема для одного конкретного компьютера, на работе сети в целом это не сказывается. Не требуется усилий по локализации неисправности.

В сети, имеющей структуру типа «кольцо», рис. 9.2, в, информация передается между станциями по кольцу с переприемом в каждом сетевом контроллере. Переприем производится через буферные накопители, выполненные на базе оперативных запоминающих устройств, поэтому при выходе их строя одного сетевого контроллера может нарушиться работа всего кольца.

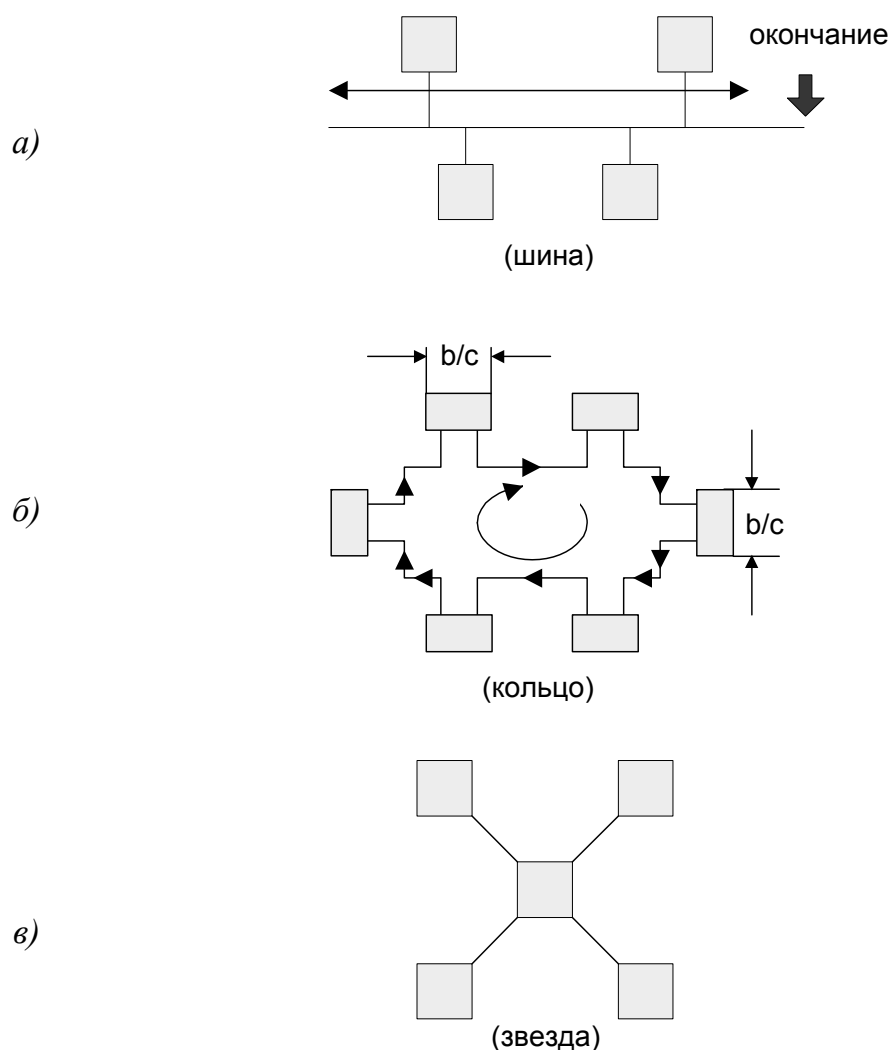


Рис. 9.2. Структура построения ЛКС: (а) шина, (б) кольцо, (в) звезда

Достоинство кольцевой структуры – простота реализации устройств, а недостаток – низкая надежность.

Все рассмотренные структуры – иерархические. Однако, благодаря использованию мостов, специальных устройств, объединяющих локальные сети с разной структурой, из вышеперечисленных типов структур могут быть построены сети со сложной иерархической структурой.

9.2.3. Физическая среда передачи в локальных сетях

Весьма важный момент – учет факторов, влияющих на выбор физической среды передачи (кабельной системы). Среди них можно перечислить следующие:

1. Требуемая пропускная способность, скорость передачи в сети.
2. Размер сети.
3. Требуемый набор служб (передача данных, речи, мультимедиа и т. д.), который необходимо организовать.
4. Требования к уровню шумов и помехозащищенности.
5. Общая стоимость проекта, включающая покупку оборудования, монтаж и последующую эксплуатацию.

Основная среда передачи данных ЛКС – неэкранированная витая пара, коаксиальный кабель, многомодовое оптоволокно. При примерно одинаковой стоимости одномодового и многомодового оптоволокна, оконечное оборудование для одномодового значительно дороже, хотя и обеспечивает большие расстояния. Поэтому в ЛКС используют, в основном, многомодовую оптику.

Основные технологии ЛКС: Ethernet, АТМ. Технологии FDDI (2 кольца), применявшаяся ранее для опорных сетей и имеющая хорошие характеристики по расстоянию, скорости и отказоустойчивости, сейчас мало используется, в основном, из-за высокой стоимости, как, впрочем, и кольцевая технология Token Ring, хотя обе они до сих пор поддерживаются на высоком уровне всеми ведущими вендорами, а в отдельных случаях (например, применение FDDI для опорной сети масштаба города, где необходима высокая отказоустойчивость и гарантированная доставка пакетов) использование этих технологий все еще может быть оправданным.

9.2.4. Типы ЛКС

Ethernet – изначально коллизийная технология, основанная на общей шине, к которой компьютеры подключаются и «борются» между собой за право передачи пакета. Основной протокол – CSMA/CD (множественный доступ с чувствительностью несущей и обнаружению коллизий). Дело в том, что если две станции одновременно начнут передачу, то возникает ситуация коллизии, и сеть некоторое время «ждет», пока «улягутся» переходные процессы и опять наступит «тишина». Существует еще один метод доступа – CSMA/CA (Collision Avoidance) – то же, но с исключением коллизий. Этот метод применяется в беспроводной технологии Radio Ethernet или Apple Local Talk – перед отправкой любого пакета в сети пробегает анонс о том, что сейчас будет происходить передача, и станции уже не пытаются ее инициировать.

Ethernet бывает полудуплексный (Half Duplex), по всем средам передачи: источник и приемник «говорит по очереди» (классическая коллизонная технология) и полнодуплексный (Full Duplex), когда две пары приемника и передатчика на устройствах говорят одновременно. Этот механизм работает только на витой паре (одна пара на передачу, одна пара на прием) и на оптоволокне (одна пара на передачу, одна пара на прием).

Ethernet различается по скоростям и методам кодирования для различной физической среды, а также по типу пакетов (Ethernet II, 802.3, RAW, 802.2 (LLC), SNAP).

Ethernet различается по скоростям: 10 Мбит/с, 100 Мбит/с, 1000 Мбит/с (Гигабит). Поскольку недавно ратифицирован стандарт Gigabit Ethernet для витой пары категории 5, можно сказать, что для любой сети Ethernet могут быть использованы витая пара, одномодовое (SMF) или многомодовое (MMF) оптоволокно. В зависимости от этого существуют различные спецификации:

- 10 Мбит/с Ethernet: 10BaseT, 10BaseFL, (10Base2 и 10Base5 существуют для коаксиального кабеля и уже не применяются);
- 100 Мбит/с Ethernet: 100BaseTX, 100BaseFX, 100BaseT4, 100BaseT2;
- Gigabit Ethernet: 1000BaseLX, 1000BaseSX (по оптике) и 1000BaseTX (для витой пары)

Существуют два варианта реализации Ethernet на коаксиальном кабеле, называемые «тонкий» и «толстый» Ethernet (Ethernet на тонком кабеле 0,2 дюйма и Ethernet на толстом кабеле 0,4 дюйма).

Тонкий Ethernet использует кабель типа RG-58A/V (диаметром 0,2 дюйма). Для маленькой сети используется кабель с сопротивлением 50 Ом. Коаксиальный кабель прокладывается от компьютера к компьютеру. У каждого компьютера оставляют небольшой запас кабеля на случай возможности его перемещения. Длина сегмента 185 м, количество компьютеров, подключенных к шине – до 30.

После присоединения всех отрезков кабеля с BNC-коннекторами (Bayonet-Neill-Concelnan) к T-коннекторам (название обусловлено формой разъема, похожей на букву «Т») получится единый кабельный сегмент. На его обоих концах устанавливаются терминаторы («заглушки»). Терминатор конструктивно представляет собой BNC-коннектор (он также надевается на T-коннектор) с впаянным сопротивлением. Значение этого сопротивления должно соответствовать значению волнового сопротивления кабеля, т. е. для Ethernet нужны терминаторы с сопротивлением 50 Ом.

Толстый Ethernet – сеть на толстом коаксиальном кабеле, имеющем диаметр 0,4 дюйма и волновое сопротивление 50 Ом. Максимальная длина кабельного сегмента – 500 м.

Прокладка самого кабеля почти одинакова для всех типов коаксиального кабеля.

Для подключения компьютера к толстому кабелю используется дополнительное устройство, называемое трансивером. Трансивер подсоединен непосредственно к сетевому кабелю. От него к компьютеру идет специальный трансиверный кабель, максимальная длина которого 50 м. На обоих его концах находятся 15-контактные DIX-разъемы (Digital, Intel и Xerox). С помощью одного разъема осуществляется подключение к трансиверу, с помощью другого – к сетевой плате компьютера.

Трансиверы освобождают от необходимости подводить кабель к каждому компьютеру. Расстояние от компьютера до сетевого кабеля определяется длиной трансиверного кабеля.

Создание сети при помощи трансивера очень удобно. Он может в любом месте в буквальном смысле «пропускать» кабель. Эта простая процедура занимает мало времени, а получаемое соединение оказывается очень надежным.

Кабель не режется на куски, его можно прокладывать, не заботясь о точном месторасположении компьютеров, а затем устанавливать трансиверы в нужных местах. Крепятся трансиверы, как правило, на стенах, что предусмотрено их конструкцией.

При необходимости охватить локальной сетью площадь большую, чем это позволяют рассматриваемые кабельные системы, применяется дополнительные устройства – репитеры (повторители). Репитер имеет 2-портовое исполнение, т. е. он может объединить 2 сегмента по 185 м. Сегмент подключается к репитеру через T-коннектор. К одному концу T-коннектора подключается сегмент, а на другом ставится терминатор.

В сети может быть не больше четырех репитеров. Это позволяет получить сеть максимальной протяженностью 925 м.

Существуют 4-портовые репитеры. К одному такому репитеру можно подключить сразу 4 сегмента.

Длина сегмента для Ethernet на толстом кабеле составляет 500 м, к одному сегменту можно подключить до 100 станций. При наличии трансиверных кабелей до 50 м длиной, толстый Ethernet может одним сегментом охватить значительно большую площадь, чем тонкий. Эти репитеры имеют DIX-разъемы и могут подключаться трансиверами, как к концу сегмента, так и в любом другом месте.

Очень удобны совмещенные репитеры, т. е. подходящие и для тонкого и для толстого кабеля. Каждый порт имеет пару разъемов: DIX и BNC, но они не могут быть задействованы одновременно. Если необходимо объединять сегменты на разном кабеле, то тонкий сегмент подключается к BNC-разъему одного порта репитера, а толстый – к DIX-разъему другого порта.

Репитеры очень полезны, но злоупотреблять ими не стоит, так как они приводят к замедлению работы в сети.

Ethernet на витой паре.

Витая пара – это два изолированных провода, скрученных между собой. Для Ethernet используется 8-жильный кабель, состоящий из четырех витых пар. Для защиты от воздействия окружающей среды кабель имеет внешнее изолирующее покрытие.

Основной узел на витой паре – hub (в переводе называется накопителем, концентратором или просто хаб). Каждый компьютер должен быть подключен к нему с помощью своего сегмента кабеля. Длина каждого сегмента не должна превышать 100 м. На концах кабельных сегментов устанавливаются разъемы RJ-45. Одним разъемом кабель подключается к хабу, другим – к сетевой плате. Разъемы RJ-45 очень компактны, имеют пластмассовый корпус и восемь миниатюрных площадок.

Хаб – центральное устройство в сети на витой паре, от него зависит ее работоспособность. Располагать его надо в легкодоступном месте, чтобы можно было легко подключать кабель и следить за индикацией портов.

Хабы выпускаются на разное количество портов – 8, 12, 16 или 24. Соответственно к нему можно подключить такое же количество компьютеров.

9.3. Технология Fast Ethernet IEEE 802.3U

Технология Fast Ethernet была стандартизирована комитетом IEEE 802.3. Новый стандарт получил название IEEE 802.3U. Скорость передачи информации 100 Мбит/с. Fast Ethernet организуется на витой паре или оптоволокне.

В сети Fast Ethernet организуются несколько доменов конфликтов, но с обязательным учетом класса повторителя, используемого в доменах.

Репитеры Fast Ethernet (IEEE 802.3U) бывают двух классов и различаются по задержке в мкс. Соответственно в сегменте (логическом) может быть до двух репитеров класса 2 и один репитер класса 1. Для Ethernet (IEEE 802.3) сеть подчиняется правилу 5-4-3-2-1.

Правило 5-4-3-2-1 гласит: между любыми двумя рабочими станциями не должно быть более 5 физических сегментов, 4 репитеров (концентраторов), 3 «населенных» физических сегментов, 2 «населенных» межрепитерных связей (IRL), и все это должно представлять собой один коллизийный домен (25,6 мкс).

Физически из концентратора «растет» много проводов, но логически это все один сегмент Ethernet и один коллизийный домен, в связи с ним любой сбой одной станции отражается на работе других. Поскольку все станции вынуждены «слушать» чужие пакеты, коллизия происходит в

пределах всего концентратора (на самом деле на другие порты посылается сигнал Jam, но это не меняет сути дела). Поэтому, хотя концентратор – это самое дешевое устройство и, кажется, что оно решает все проблемы заказчика, советуем постепенно отказаться от этой методики, особенно в условиях постоянного роста требований к ресурсам сетей, и переходить на коммутируемые сети. Сеть из 20 компьютеров, собранная на репитерах 100 Мбит/с, может работать медленнее, чем сеть из 20 компьютеров, включенных в коммутатор 10 Мбит/с. Если раньше считалось «нормальным» присутствие в сегменте до 30 компьютеров, то в нынешних сетях даже 3 рабочие станции могут загрузить весь сегмент.

В Fast Ethernet внутри одного домена конфликтов могут находиться не более двух повторителей класса II (рис. 9.3) или не более одного повторителя класса I (рис. 9.4)

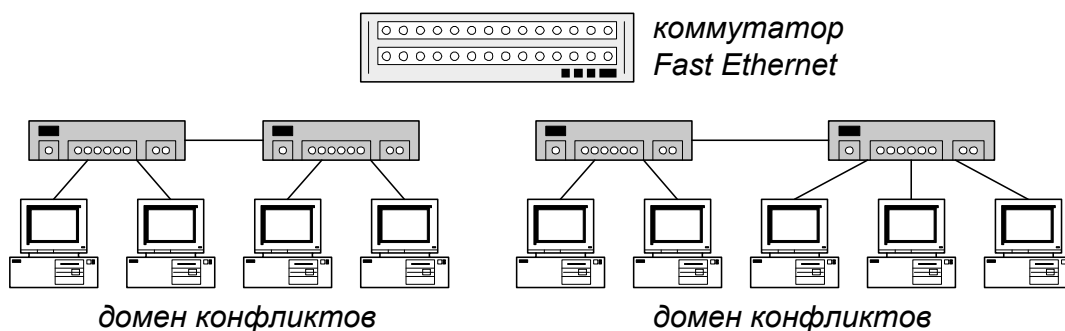


Рис. 9.3. Структура сети на повторителях класса 2 с использованием витой пары

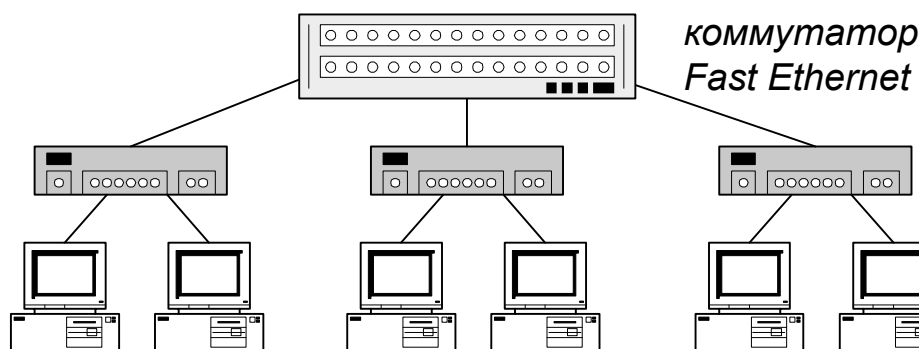


Рис. 9.4. Структура сети на повторителях класса 1 с использованием витой пары

Различные типы кабелей и устройств Fast Ethernet дают разную величину задержки RTD. Витая пара категории 5 – 1,11 бит-тайм на метр длины, оптоволоконный кабель 1 бит-тайм также на метр длины, сетевой адаптер – 50 бит-тайм, медиаконвертеры от 50 до 100, повторитель

класса I – 140, повторитель класса II – 92 бит-тайм. Задержку RTD между двумя сетевыми узлами рассчитать несложно, она равняется сумме соответствующих задержек их сетевых адаптеров и всех промежуточных сетевых компонентов (кабелей, повторителей).

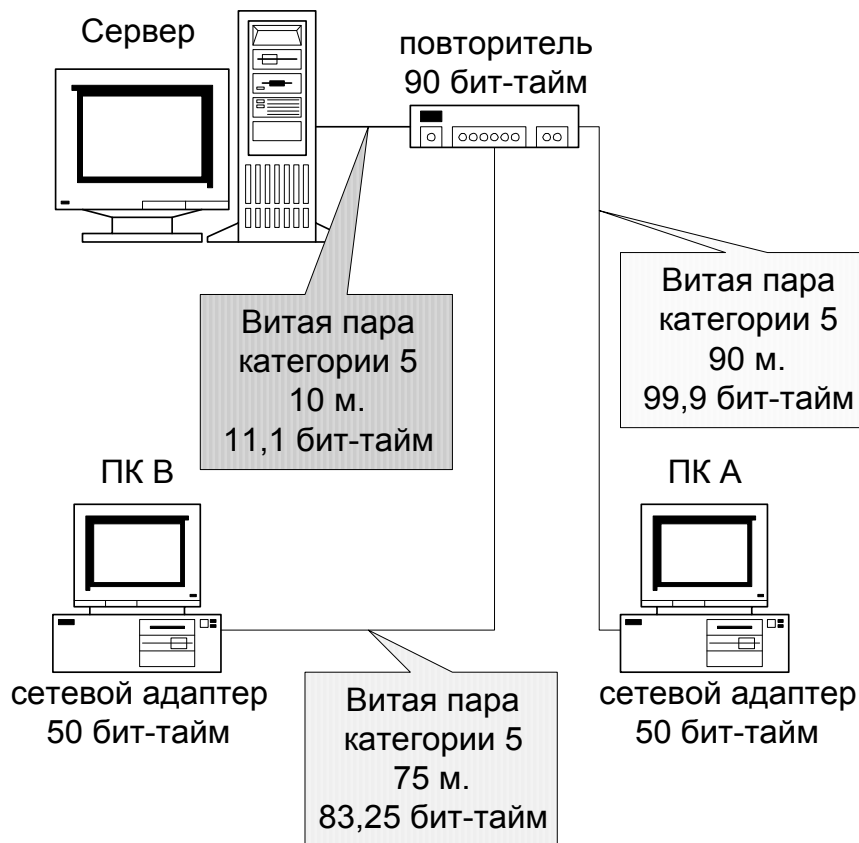


Рис. 9.5. Пример сети Fast Ethernet

В представленном на рис. 9.5 примере сети задержка сигнала на пути от ПК А до ПК В равна 373,15 бит-тайм. Разумеется, задержка RTD между любыми двумя узлами не должна превышать 512 бит-тайм. Отсюда вытекают ограничения на число повторителей (не более двух, класса II) и на физические размеры сетей Fast Ethernet. Максимальный размер сети на базе витой пары (спецификация 100Base-TX) с двумя повторителями класса II составляет 205 м, а два компьютера (устройства DTE) могут быть связаны между собой отрезком оптоволоконного кабеля длиной 412 м.

9.4. Технология Gigabit Ethernet

Следующий шаг в развитии технологии Ethernet – разработка проекта стандарта IEEE-802.32. Данный стандарт предусматривает скорость обмена информацией между станциями локальной сети 1 Гбит/с. Предполагая, что устройства Gigabit Ethernet будут объединять сегменты сетей с Fast Ethernet со скоростями 100 Мбит/с. Разрабатываются се-

тевые карты со скоростью 1 Гбит/с, а также серия сетевых устройств, таких как коммутаторы и маршрутизаторы.

В сети с Gigabit Ethernet будет использоваться управление трафиком, контроль перегрузок и обеспечение качества обслуживания (Quality Of Service – QOS). Стандарт Gigabit Ethernet – один из серьезных соперников развивающейся сегодня технологии ATM.

9.5. Технологии ATM

Сеть ATM имеет звездообразную топологию. Сеть ATM строится на основе одного или нескольких коммутаторов, являющихся неотъемлемой частью данной коммуникационной структуры.

Высокая скорость передачи и чрезвычайно низкая вероятность ошибок в волоконно-оптических системах выдвигают на первый план задачу создания высокопроизводительных систем коммутации на основе стандартов ATM.

Простейший пример такой сети – один коммутатор, обеспечивающий коммутацию пакетов, данных и несколько оконечных устройств.

ATM – это метод передачи информации между устройствами в сети маленькими пакетами фиксированной длины, названными ячейками (cells). Фиксация размеров ячейки имеет ряд существенных преимуществ по сравнению с пакетами переменной длины.

- Во-первых, ячейки фиксированной длины требуют минимальной обработки при операциях маршрутизации в коммутаторах. Это позволяет максимально упростить схемные решения коммутаторов при высоких скоростях коммутации.
- Во-вторых, все виды обработки ячеек по сравнению с обработкой пакетов переменной длины значительно проще, так как отпадает необходимость в вычислении длины ячейки.
- В-третьих, в случае применения пакетов переменной длины передача длинного пакета данных могла бы вызвать задержку выдачи в линию пакетов с речью или видео, что привело бы к их искажению.

Модель ATM имеет четырехуровневую структуру. Различают несколько уровней:

- пользовательский (User Layer) – включает уровни, начиная с сетевого и выше (IPX/SPX или TCP/IP);
- адаптации (ATM Adaptation Layer – AAL);
- ATM (ATM Layer);
- физический (Physical Layer).

Пользовательский уровень обеспечивает создание сообщения, которое должно быть передано в сеть ATM и соответствующим образом преобразовано.

Уровень адаптации (AAL) обеспечивает доступ пользовательских приложений к коммутирующим устройствам АТМ. Данный уровень формирует стандартные АТМ-ячейки и передает их на уровень АТМ для последующей обработки.

Физический уровень обеспечивает передачу ячеек через разнообразные коммутационные среды. Данный уровень состоит из двух подуровней – подуровня преобразования передачи, реализующего различные протоколы передачи по физическим линиям, и подуровня адаптации к среде передачи.

Оконечные устройства АТМ – сети, подключающиеся к коммутаторам через интерфейс, называемый UNI – интерфейс пользователя с сетью. UNI может быть интерфейсом между рабочей станцией, ПК, АТС, маршрутизатором, или каким угодно «черным ящиком» и АТМ-коммутатором.

ГЛАВА 10. ОБЩИЕ СВЕДЕНИЯ О СОТОВЫХ СИСТЕМАХ ПОДВИЖНОЙ РАДИОСВЯЗИ

10.1. Хронология развития сотовых систем

Начиная с 40-х годов ученые и инженеры разных стран пытались решить проблему ограниченности частотного ресурса, обусловленную ростом количества абонентов радиосистем. В середине 40-х годов исследовательский центр Bell Laboratories американской компании AT&T предложил идею разбиения всей обслуживаемой территории на небольшие участки, которые стали называться сотами (от англ. Cell – ячейка, клеточка, сота). Каждая сота должна была обслуживаться передатчиком с ограниченным радиусом действия и фиксированной частотой. Это позволило бы без всяких взаимных помех использовать ту же самую частоту повторно в другой ячейке (cote). Так появились сотовые системы подвижной радиосвязи (ССПР).

Поколения ССПР показаны на рис. 10.1.

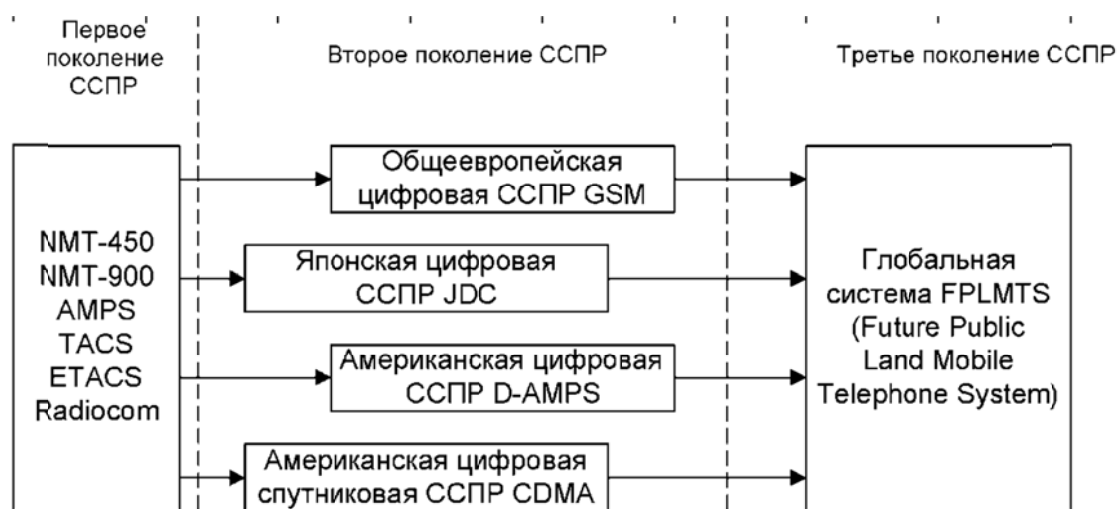


Рис. 10.1. Поколения сотовых систем

Еще в конце 70-х годов начались работы по созданию единого стандарта сотовой связи первого поколения (1G) для пяти североевропейских стран – Швеции, Финляндии, Исландии, Дании и Норвегии, который получил название NMT-450 (Nordic Mobile Telephone) и был предназначен для работы в диапазоне 450 МГц.

Впервые система сотовой связи стандарта NMT-450 вступила в эксплуатацию в Саудовской Аравии в 1981 г. На базе этого стандарта в 1985 г. был разработан стандарт NMT-900 диапазона 900 МГц, который

позволил расширить функциональные возможности ССПР и значительно увеличить их абонентскую емкость. С 1986 г. в скандинавских странах начали применять стандарт NMT-900.

В 1983 г. в США в районе Чикаго после ряда успешных полевых испытаний в исследовательском центре Bell Laboratories вступила в эксплуатацию радиосеть стандарта AMPS (Advanced Mobile Phone Service).

В 1985 г. в Великобритании был принят в качестве национального стандарт TACS (Total Access Communications System), разработанный на основе американского стандарта AMPS. В 1987 г. в связи с резким увеличением в Лондоне числа абонентов сотовой связи была расширена рабочая полоса частот. Новая версия этого стандарта сотовой связи получила название ETACS (Enhanced TACS).

Во Франции в 1985 г. был принят стандарт Radiocom-2000.

В 1992 г. в Санкт-Петербурге, а затем и в Москве появились российские проработки ССПР стандарта NMT-4501 (усовершенствованного стандарта NMT-450).

Все вышеперечисленные стандарты являются аналоговыми и относятся к первому поколению систем сотовой связи. В них используется аналоговый способ передачи информации с помощью частотной или фазовой модуляции. Этот способ имеет ряд существенных недостатков: возможность прослушивания разговоров другими абонентами, отсутствие эффективных методов борьбы с замираниями сигналов под влиянием окружающего ландшафта и зданий вследствие передвижения абонентов.

Использование новейших технологий и научных открытий в области связи и обработки сигналов позволило подойти к концу 80-х годов к новому этапу развития систем сотовой связи – созданию систем второго поколения (2G), основанных на цифровых методах обработки сигналов.

В целях разработки единого европейского стандарта цифровой сотовой связи для выделенного диапазона 900 МГц в 1982 г. Европейская конференция администраций почт и электросвязи (CEPT), объединяющая 26 стран, создала специальную группу Groupe Special Mobile (GSM) и дала название новому стандарту. Позднее в связи с широким распространением этого стандарта во всем мире GSM стали расшифровывать также, как Global System for Mobile Communications. Результатом работы этой группы стали опубликованные в 1990 г. требования к системе сотовой связи стандарта GSM. В 1992 г. в Германии вступила в эксплуатацию первая система сотовой связи этого стандарта.

США провозгласили свою концепцию D-AMPS ССПР поколения 2G. В отличие от Европы, в США не были выделены новые частотные диапазоны, поэтому система должна была работать в полосе частот, общей с аналоговой ССПР AMPS.

Одновременно американская компания Qualcomm начала активную разработку нового стандарта сотовой связи, основанного на технологии шумоподобных сигналов и кодовом разделении каналов CDMA. В 1993 г. в США после ряда успешных испытаний Промышленная ассоциация в области связи ТИА приняла стандарт CDMA-one как внутренний стандарт IS-95 цифровой сотовой связи. В 1995 г. в Гонконге была открыта первая радиосеть этого стандарта.

В Японии в 1991 г. Министерством почт и связи Японии был утвержден собственный стандарт сотовой связи JDC (Japanese Digital Cellular), близкий к американскому стандарту D-AMPS.

По сравнению с аналоговыми системами первого поколения цифровые ССПР второго поколения обеспечивают повышенное качество связи, а также взаимодействие с цифровыми коммутационными сетями и целый ряд дополнительных услуг. В них используются временное разделение каналов, шифрование сообщений и новый вид модуляции – GMSK (Gaussian Minimum Shift Keying).

Дальнейшее развитие сотовой подвижной связи (см. рис. 2.1) осуществляется в рамках создания проектов систем третьего поколения (3G).

В Европе концепция ССПР третьего поколения получила название UMTS (Universal Mobile Telecommunication System), в США – CDMA-2000. Они предусматривают объединение функциональных возможностей существующих цифровых систем связи ССПР, МСБД и СБАД в единую систему третьего поколения FPLMTS (Future Public Land Mobile Telephone System). Кроме этого, отличительной особенностью систем третьего поколения является межсетевой роуминг с спутниковыми системами радиосвязи «Иридиум», «Globalstar», «Inmarsat-P» и др., т. е. они фактически входят в состав этих систем радиосвязи и будут иметь архитектуру единой сети для связи абонентов, находящихся в различных условиях, включая движущийся транспорт, жилые помещения, офисы и т. д.

В настоящее время функционируют ССПР всех трех поколений.

10.2. Территориальная организация ССПР

Для ССПР, в которых зона обслуживания делится на ячейки (соты), характерно то, что в них обеспечивается сотовый принцип распределения частот по территории обслуживания или территориально-частотная организация ССПР.

Разделить обслуживаемую территорию на ячейки (соты) можно двумя способами, основанными на следующем:

- на использовании среднестатистических характеристик распространения радиоволн в регионе действия ССПР;

- на измерении или расчете энергетики распространения радиоволн для всех конкретных сот региона действия ССПР.

При реализации первого способа вся обслуживаемая территория разделяется на одинаковые по форме зоны, и с помощью закона статистической радиофизики определяются их допустимые размеры и расстояния до других зон, в пределах которых выполняются условия допустимого взаимного влияния. Для оптимального разделения территории на соты, т. е. без перекрытия или пропусков участков, могут быть использованы только три геометрические фигуры: треугольник, квадрат и шестиугольник. Наиболее подходящей фигурой является шестиугольник, так как при установке в его центре антенны с круговой диаграммой направленности будет обеспечен радиодоступ почти ко всем участкам соты.

При использовании рассмотренного способа интервал между зонами, в которых используются одинаковые рабочие каналы, обычно получается меньше требуемого для поддержания взаимных помех на допустимом уровне и количество сот и базовых станций увеличивается.

При втором способе разделения на зоны тщательно измеряют или рассчитывают зоны радиовидимости, в которых обеспечивается удовлетворительное обслуживание абонентов, определяют оптимальное место расположения базовой станции с учетом рельефа местности, для выравнивания зон рассматривают возможность использования направленных антенн, пассивных ретрансляторов и т. д. Все это позволяет использовать минимальное число сот и, следовательно, базовых станций. Второй способ более точный и экономически выгодный, но и более трудоемкий.

10.3. Частотная организация ССПР

Частотный план ССПР при трех наборах $F1$, $F2$, $F3$ частот базовых станций показан на рис. 10.2.

Группа сот с различными наборами частот называется кластером. На рис. 10.2, например, размерность кластера C равна трем. В существующих ССПР размерность кластера – от трех до пятнадцати.

Смежные базовые станции образуют кластерную группу станций. Если каждой базовой станции выделяется T каналов с шириной полосы каждого F , то общая ширина полосы, занимаемая кластерной группой станций и ССПР в целом, составит $Fc = F \cdot T \cdot C$.

Если $T = 1$, т. е. в соте находится один среднестатистический абонент, то $Fc = F_{min} = F \cdot C$. Поэтому величина C определяет минимально возможное число каналов в системе, и ее называют частотным параметром или коэффициентом повторения частот.

Уменьшение радиуса ячейки позволяет повысить эффективность использования выделенной полосы частот, увеличить абонентскую ем-

кость системы, уменьшить мощность передатчиков, чувствительность приемников базовых и подвижных станций, улучшить условия электромагнитной совместимости средств сотовой связи с другими радиоэлектронными средствами и системами и повысить безопасность работы абонентов с радиосредствами ССПР.

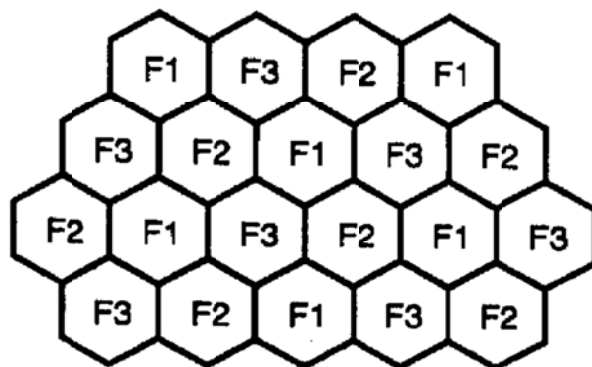


Рис. 2.2. Частотный план сотовой системы с тремя частотами

Радиус ячейки R следует уменьшать до оптимальной величины R_{min} , при которой в сайте находится один среднестатистический абонент. Поскольку наименьший кластер $C = 3$, то минимально возможная полоса ССПР $F_{cmin} = 3F$. Таким образом, полоса частот ССПР с сотнями сот и тысячами абонентов может теоретически иметь полосу частот всего в три раза превышающую полосу частот системы децентрализованной связи с тремя абонентами.

Однако в реальных случаях такой выигрыш не получается, так как из-за миграции абонентов количество каналов в соте выбирают по максимальному, не по среднестатистическому числу абонентов.

10.4. Способы борьбы с системными помехами ССПР

Системные помехи абонентской станции – это помехи от других абонентских и базовых станций рядом расположенных сот ССПР.

Первый способ борьбы с системными помехами заложен территориально-частотной организацией ССПР, поскольку базовые станции с одинаковыми наборами частот находятся не в соседних сотах, т. е. разнесены территориально.

Вторым способом снижения уровня системных помех может быть использование направленных секторных антенн с узкими диаграммами направленности. В секторе такой направленной антенны сигнал излучается преимущественно в одну сторону, а уровень излучения в противоположном направлении сокращается до минимума. Деление сот на секторы позволяет чаще применять частоты в сотах повторно.

Широко используемый способ повторного использования частот в организованных таким образом сотах основан на применении трехсекторных антенн для каждой базовой станции и трех соседних базовых станций с формированием ими девяти групп частот (рис. 10.3). В этом случае используются антенны с шириной диаграммы направленности 120° , размерность кластера $C = 3$, количество наборов частот $N = 3$, количество частот одной соты $T = N \cdot T_{сект} = 3T_{сект}$, и при минимальном количестве $T_{сект} = 1$ каналов набора ширина полосы ССПР $F_c = F_{cmin} = 3 \cdot 3F$, т. е. в три раза шире, чем при использовании антенн с круговой диаграммой направленности.

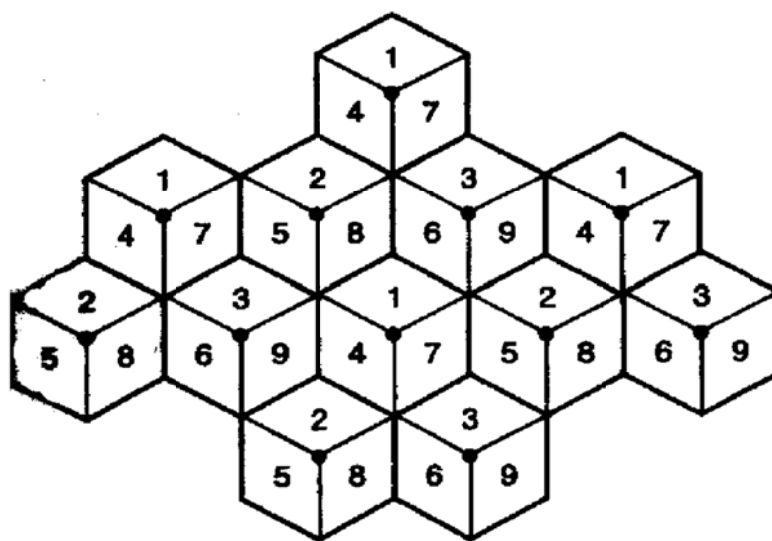


Рис. 10.3. Способ повторного использования частот

Высокую эффективность использования полосы частот и наибольшее число абонентов сети, работающих в этой полосе, обеспечивает способ повторения частот фирмы Motorola (рис. 10.4). В этом случае используются антенны с шириной диаграммы направленности 60° , размерность кластера $C = 4$, количество наборов частот $N = 6$, количество частот одного сайта $T = N \cdot T_{сект} = 6T_{сект}$, и при минимальном количестве $T_{сект} = 1$ каналов набора ширина полосы ССПР $F_c = F_{cmin} = 4 \cdot 6 \cdot F$, т. е. в восемь раз шире, чем при односекторных антеннах.

Оба примера (см. рис. 10.3, 10.4) использования многосекторных антенн показывают, что обеспечиваемое при этом снижение уровня системных помех достигается за счет расширения полосы рабочих частот ССПР.

Третий способ снижения системных помех направлен на исключение попадания в полосу абонентской станции комбинационных (интерференционных) колебаний третьего порядка частот $f_{комб} = 2f_i - f_j$, формируемых на нелинейностях входных каскадов абонентской станции

при приеме нескольких колебаний частот f_i, f_j от других абонентских или базовых станций ССПР.

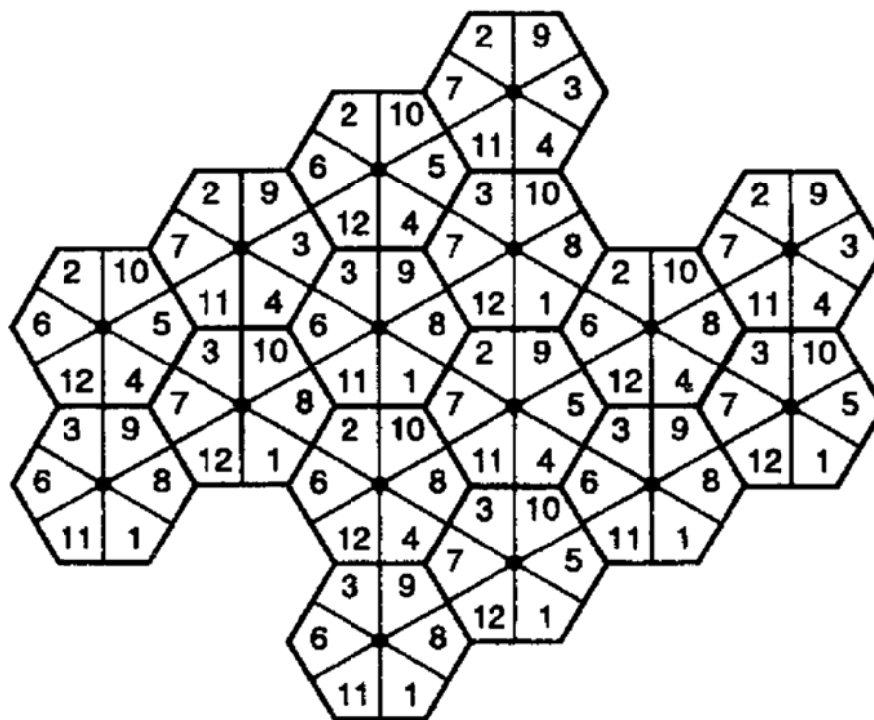


Рис. 10.4. Способ повторения частот

Полное число m радиоканалов, в которое входит группа из n каналов, комбинационные частоты любых пар которых не попадают на частоты этих n каналов, выражается

$$m = (n - 1) \cdot (n^2 - 2n + a)/4,$$

где $a = 4$ при четном n и $a = 3$ при нечетном n .

Так, например, при $n = 5$ рабочих каналах нужно иметь $m = 18$ частотных каналов и, следовательно, общую полосу частот $\beta = 3,6$ раза превышающую полосу рабочих частот. Чем меньше $\beta = m/n$ величина β , тем более эффективно использование частотного диапазона.

В ССПР основными процедурами являются эстафетная передача канала, роуминг, автоматическое установление входящего, исходящего вызовов и др. Несмотря на разнообразие стандартов сотовой связи, алгоритмы их функционирования, независимо от имеющихся особенностей, в основном сходны.

Процедура (рис. 10.5), называемая передачей управления вызовом или эстафетной передачей канала (в иностранной технической литературе – handover, или handoff) осуществляется, если по мере удаления абонента от базовой станции или в связи с ухудшением условий распро-

странения радиоволн уровень сигнала уменьшается, что ведет к ухудшению качества связи. Улучшение качества связи достигается путем автоматического переключения абонента на другой канал связи.

На рис. 10.5 MSC – центр коммутации и управления ССПР, BS – базовые станции, MS – абонентские станции. Для контроля качества связи базовая станция снабжена специальным приемником, периодически измеряющим уровень сигнала сотового телефона разговаривающего абонента и сравнивающим его с допустимым пределом. Если уровень сигнала меньше этого предела, то информация об этом автоматически передается в центр коммутации по служебному каналу связи. Центр коммутации выдает команду об измерении уровня сигнала сотового радиотелефона абонента на ближайшие к нему базовые станции. После получения информации от базовых станций об уровне этого сигнала центр коммутации переключает радиотелефон на ту из них, где уровень сигнала оказался наибольшим.

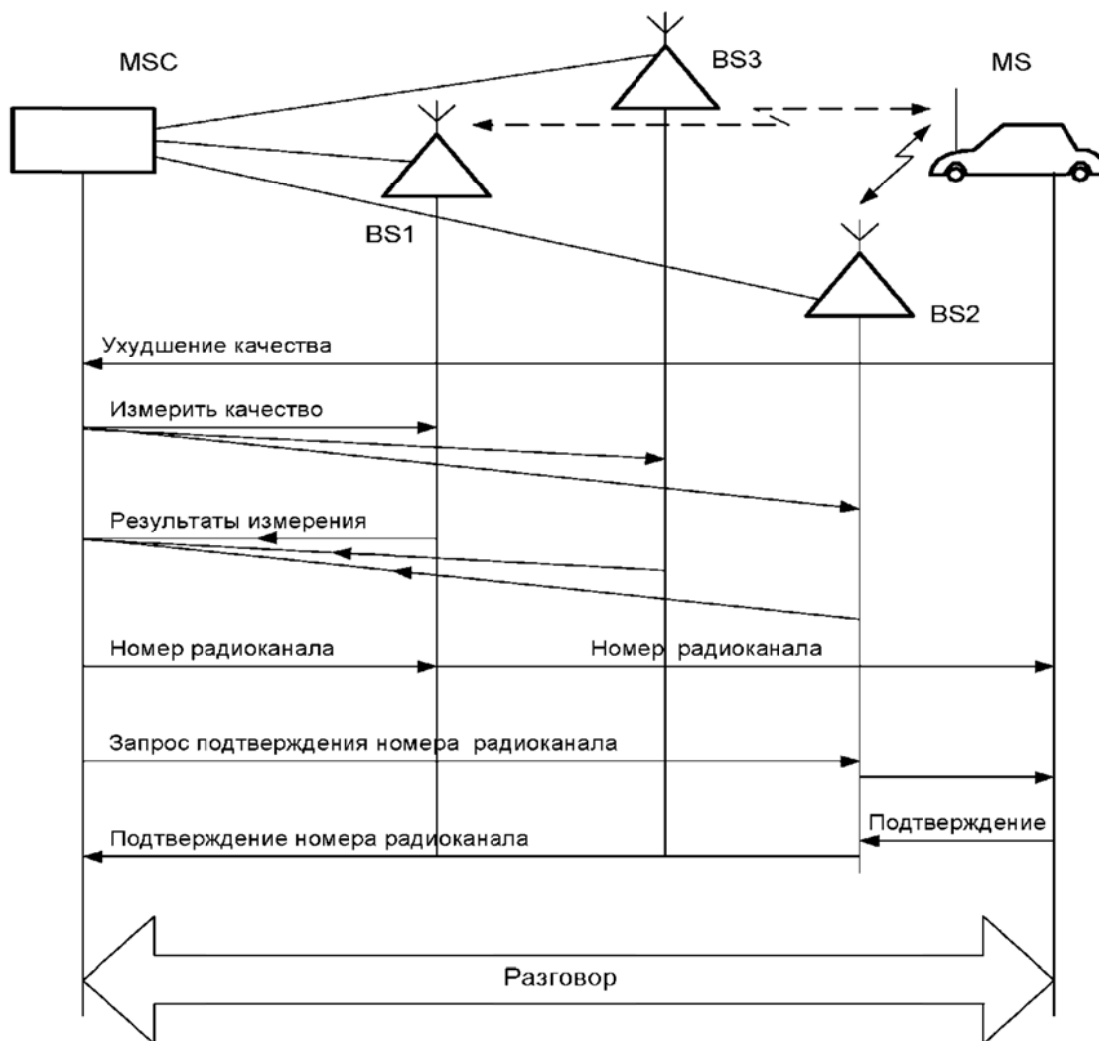


Рис. 10.5. Эстафетная передача канала

Иногда возникает ситуация, когда поток заявок на обслуживание, поступающий от абонентов сотовой сети, превышает количество каналов, имеющихся на всех близко расположенных базовых станциях. Это происходит тогда, когда все каналы станций заняты обслуживанием абонентов и нет ни одного свободного, и поступает очередная заявка на обслуживание от подвижного абонента. В этом случае как временная мера (до освобождения одного из каналов) используется принцип эстафетной передачи внутри соты или «жонглирование» абонентами. При этом происходит поочередное переключение каналов в пределах одной и той же базовой станции для обеспечения связью всех абонентов.

Одна из важных услуг сети сотовой связи – предоставление возможности использования одного и того же радиотелефона при поездке в другой город, область или даже страну, причем сотовая сеть позволяет не только самому абоненту звонить из другого города или страны, но и получать звонки от тех, кто не успел застать его дома. В сотовой радиосвязи такая возможность называется роуминг (от англ. roam – скитаться, блуждать).

Различают три вида роуминга:

- автоматический (именно с этой формой за рубежом обычно и связывают понятие роуминга), т. е. предоставление абоненту возможности выйти на связь в любое время в любом месте;
- полуавтоматический, когда абоненту для пользования данной услугой в каком-либо регионе необходимо предварительно поставить об этом в известность своего оператора;
- ручной, по сути, простой обмен одного радиотелефона на другой, подключенный к сотовой системе другого оператора.
- Для обеспечения автоматического и полуавтоматического роуминга необходимо выполнение трех условий:
- наличие в требуемых регионах сотовых систем стандарта, совместимого со стандартом компании, у которой был приобретен радиотелефон;
- наличие соответствующих организационных и экономических соглашений о роуминговом обслуживании абонентов для взаиморасчетов между операторами сетей;
- наличие каналов связи между системами, обеспечивающих передачу звуковой и другой информации для роуминговых абонентов.

При перемещении абонента в другую сеть ее центр коммутации запрашивает информацию в первоначальной сети и при наличии подтверждения полномочий абонента регистрирует его. Данные о местоположении абонента постоянно обновляются в центре коммутации первоначальной сети, и все поступающие туда вызовы автоматически переадресовываются в ту сеть, где в данный момент находится абонент.

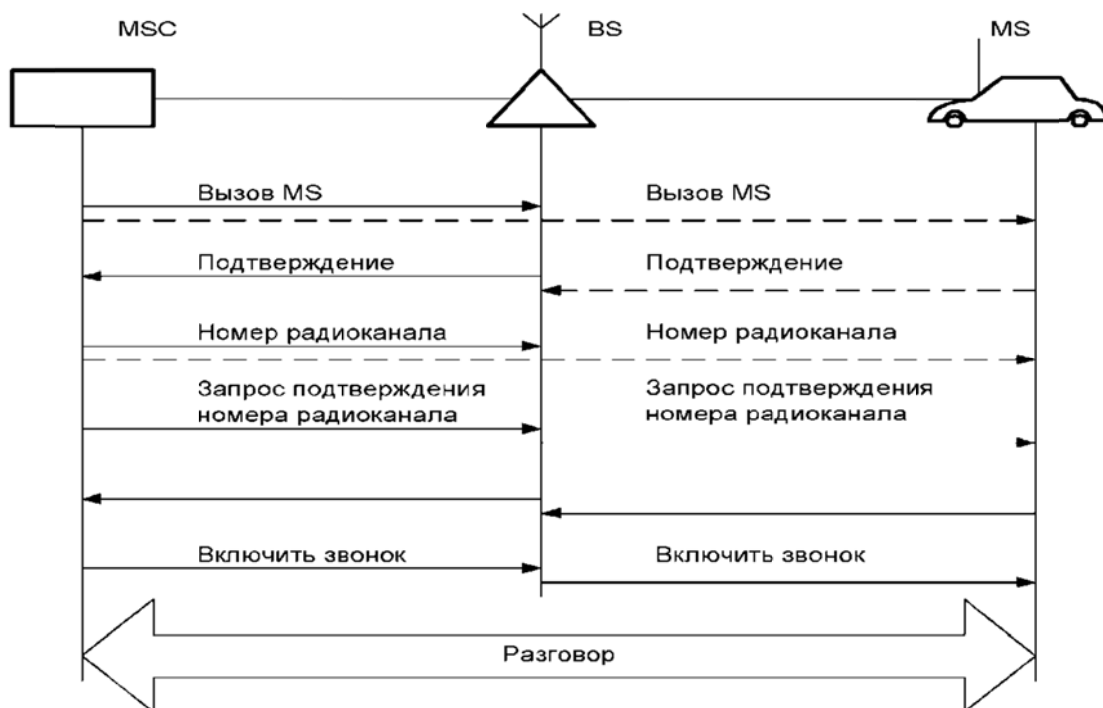


Рис. 10.6. Процедура установления входящего вызова

Процедура установления входящего вызова проиллюстрирована на рис. 10.6.

10.5. ССПР стандарта NMT

Структура ССПР NMT показана на рис. 10.7. Каждая из сот обслуживается многоканальным приемопередатчиком базовой станции. Она служит интерфейсом между сотовым телефоном и центром коммутации подвижной связи (MSC), где роль проводов обычной телефонной сети выполняют радиоканалы. Все базовые станции соединены с центром коммутации подвижной связи по выделенным проводным или радиорелейным каналам связи.

Центр коммутации подвижной связи – это автоматическая телефонная станция системы сотовой связи, обеспечивающая все функции управления сетью. Она осуществляет постоянное слежение за подвижными станциями, организует их эстафетную передачу, в процессе которой достигается непрерывность связи при перемещении подвижной станции из соты в соту, и переключение рабочих каналов в соте при появлении помех или неисправностей, а также производит соединение подвижного абонента с абонентом телефонной сети.

ССПР стандарта NMT – системы с централизованным управлением.

Число дуплексных каналов базовой станции кратно 8. Один из каналов является управляющим (control channel). По нему передаются

служебные цифровые сигналы. В некоторых ситуациях он может быть также *каналом вызова* (calling channel). На этом канале происходит непосредственное установление соединения при вызове подвижного абонента сети, а разговор начинается только после того, как будет найден свободный в данный момент канал и произойдет переключение на него. Когда на базовой станции все каналы связи заняты, канал вызова может также быть *трафиковым* каналом (traffic channel) и использоваться для ведения разговора. В служебных сигналах контрольного канала промаркированы:

- канал вызова каждой базовой станции;
- один или несколько свободных каналов;
- зоны обслуживания;
- страна, в которой находится подвижная станция;
- номер рабочего канала.

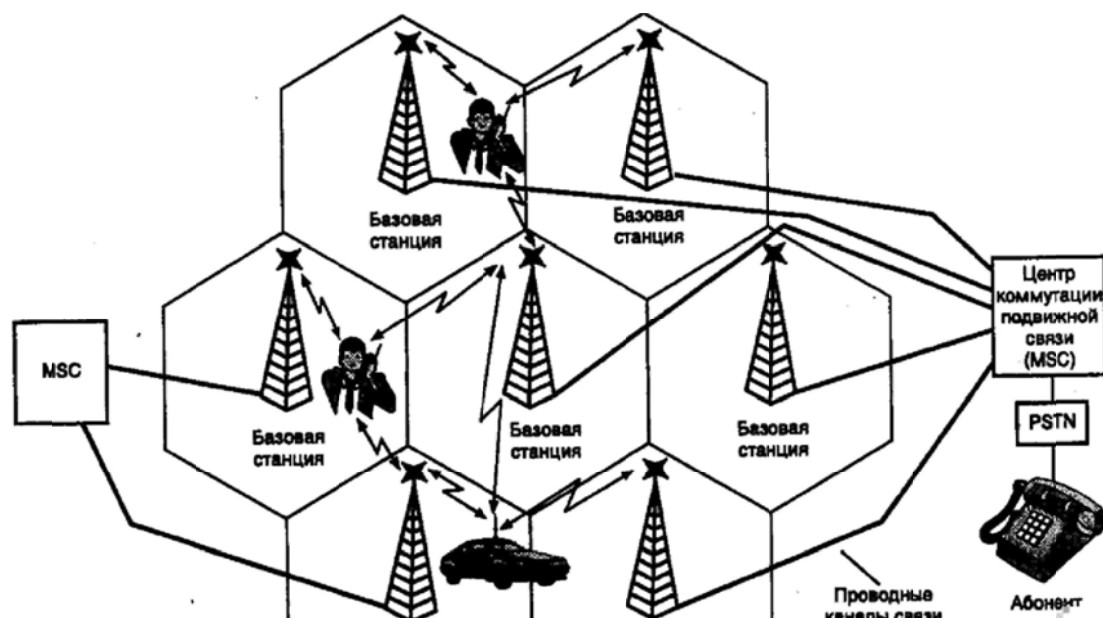


Рис. 10.7. Структура сети NMT

Все трафиковые сигналы аналоговые и используют фазовую модуляцию, а служебные сигналы являются цифровыми и передаются со скоростью 1200/1800 бит/с FSK модуляцией (Frequency Shift Keying). Принцип формирования FSK такой же, как при пакетной связи протокола AX25, но с частотами коммутации 1200 и 1800 Гц.

Принципы построения сотовых систем радиосвязи стандартов NMT-450 и NMT-900 практически совпадают. Оба стандарта сотовой связи базируются на спецификации стандарта NMT-450. Основные отличия более совершенного стандарта NMT-900 первоначально были связаны с введением в состав абонентского оборудования малогабарит-

ной носимой станции, совершенствованием управления и развитием услуг связи.

Кроме передачи речевых сообщений на местном, междугородном и международном уровнях ССПР стандартов NMT-450 и NMT-900 предоставляют абонентам широкий набор услуг, например:

- позволяют отправить телефаксы;
- обеспечивают доступ к различным базам данных при скорости передачи данных 4,8 кбит/с;
- дают возможность переадресовывать вызов на другой номер, ограничивают продолжительность разговоров, конференцсвязи трех абонентов, организовывать пользовательские группы с сокращенным набором номера;
- защищают доступ к сети с помощью системы SIS идентификации абонента (Subscriber Identification Security).

Так как общее число радиочастот, имеющихся в наличии в системах стандартов NMT-450 и NMT-900, ограничено, то для того, чтобы увеличить емкость системы связи предусматривается формирование малых зон связи («малые ячейки») в густонаселенных районах.

Выходная мощность передатчиков всех подвижных станций автоматически уменьшается по команде радиотелефонного коммутатора, когда станция входит в зону «малой ячейки». Эта же процедура используется для того, чтобы уменьшить помехи в случае, когда подвижные станции находятся близко от базовых станций с обычными зонами обслуживания.

В табл. 10.1 приведены основные характеристики ССПР стандартов NMT-450 и NMT-900.

Таблица 10.1

Характеристики стандартов NMT

Наименования параметров	NMT-450	NMT-900
Полоса частот:		
• для передачи подвижной станцией	453...457,5 МГц	890...915 МГц
• для приема подвижной станцией	463...467,5 МГц	935...960 МГц
Дуплексный разнос каналов приема и передачи	10 МГц	45 МГц
Частотный разнос каналов	25 (20) кГц	25 (12,5) кГц
Количество каналов	180 (225)	999 (1999)
Радиус соты	15...40 км	2...20 км
Мощность передатчика базовой станции	50 Вт	25 Вт
Мощность передатчика подвижной станции	15 Вт	6 Вт
	1,5 Вт	1 Вт
	0,15 Вт	0,1 Вт

10.6. ССПР стандарта AMPS

Стандарт AMPS во многом похож на стандарт NMT, но имеются и существенные отличия. Рассмотрим эти отличия.

В ССПР стандарта AMPS используется принцип разнесенного приема сигналов, поэтому базовые станции содержат по две антенны и двухканальные приемники, и осуществляется последетекторное сложение этих сигналов.

В ССПР стандарта AMPS используются прямой и обратный каналы управления, т. е. эти системы с комбинированным управлением, а ССПР стандарта NMT – системы с централизованным управлением.

Данные по прямому каналу управления в направлении от базовой станции к подвижной передаются непрерывным потоком, и при отсутствии информации для подвижной станции содержит только контрольный текст. В нерабочем состоянии приемное устройство подвижной станции сканирует каналы управления, выбирая канал с наиболее высоким уровнем сигнала.

В отличие от ССПР стандарта NMT в AMPS контроль достоверности принимаемых сообщений осуществляется и прямым и обратным каналом. При этом достоверность принимаемой информации служебных каналов увеличивается благодаря пяти повторам ее передачи.

В системах стандарта AMPS процедура эстафетной передачи канала подобна соответствующей процедуре стандарта NMT и отличается лишь тем, что центр коммутации идентифицирует *шесть* ближайших к абоненту базовых станций. Вся процедура эстафетной передачи занимает около 250 мс.

Как и в стандарте NMT трафиковые сигналы аналоговые и используют фазовую модуляцию, а служебные сигналы цифровые и передаются FSK модуляцией со скоростью 8 кбит/с.

Системы стандарта AMPS работают в диапазоне 824...849 МГц, 869...894 МГц с дуплексным разносом 45 МГц и имеют 666 дуплексных каналов при ширине полосы частот каждого канала 30 кГц. Узкополосная модификация аналогового стандарта NAMPS (Narrow Band AMPS) имеет полосу канала 10 кГц и соответственно в три раза большее количество каналов связи. Мощность передатчиков базовых станций составляет 45 Вт, автомобильных станций – 12 Вт, переносных аппаратов – 1 Вт.

Цифровые ССПР D-AMPS существуют двух модификаций: промежуточного стандарта IS-54, которые совмещают работу в аналоговом и цифровом режимах в том же диапазоне, что и AMPS, и стандарта IS-136, которые полностью цифровые и имеют дополнительный 1900 МГц диапазон. Кроме этого в ССПР D-AMPS расширен круг услуг и ассор-

тимент абонентских аппаратов, а также осуществляется управление базовыми станциями мощностью абонентских станций.

10.7. Цифровые ССПР стандарта GSM

10.7.1. Структура ССПР GSM

Структура ССПР стандарта GSM показана на рис. 10.8, на котором MSC (Mobile Switching Center) – центр коммутации подвижной связи; HLR – регистр положения; VLR – регистр перемещения; EIR (Equipment Identification Register) – регистр идентификации оборудования; AUC – центр аутентификации; BSS (Base Station System) – оборудование базовой станции; BSC (Base Station Controller) – контроллер базовой станции; BTS (Base Station Transmitter) – приемопередатчик базовой станции; SSS (Switch Station System) – приемопередатчик базовой станции; OMC (Operations and Maintenance Center) – центр управления и технического обслуживания; NMC (National Maintenance Center) – центр управления сетью; MS (Mobile Stations) – подвижные станции.

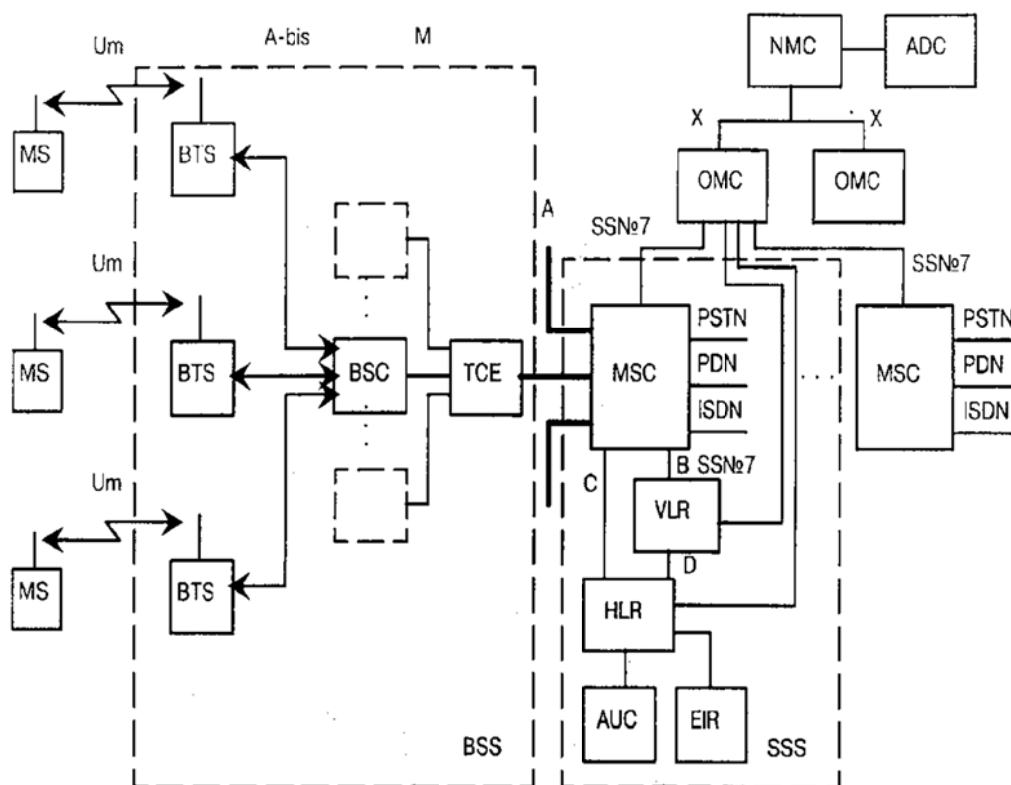


Рис. 10.7. Структура ССПР стандарта GSM

Несколько MS и одна BSS входят в одну соту, несколько сот и одна SSS входят в одну зону, несколько зон и один OMC входят в один регион, несколько регионов и один NMC входят во всю сеть/

Центр коммутации подвижной связи выполняет следующие задачи:

- обслуживает группу сот и обеспечивает все виды соединений, в которых нуждается в процессе работы подвижная станция;
- обеспечивает маршрутизацию вызовов и функции управления вызовами;
- осуществляет эстафетную передачу каналов и передачу каналов при появлении помех или неисправностях;
- формирует данные, необходимые для выписки счетов за предоставленные сетью услуги связи;
- составляет статистические данные, необходимые для контроля работы и оптимизации сети;
- поддерживает процедуры безопасности, применяемые для управления доступами к радиоканалам;
- регистрирует местоположения подвижных станций.

Центр коммутации для осуществления постоянного слежения за подвижными станциями использует регистры положения HLR и перемещения VLR.

Регистры *HLR* и *VLR* представляют собой справочную базу данных о постоянно прописанных в сети абонентах. В ней содержатся идентификационные номера и адреса, а также параметры подлинности абонентов, состав услуг связи, специальная информация о маршрутизации. В них ведется регистрация данных о роуминге абонента, включая данные о временном идентификационном номере подвижного абонента.

С помощью VLR достигается функционирование подвижной станции за пределами зоны, контролируемой HLR. VLR содержит такие же данные, как и HLR, но эти данные содержатся в VLR только до тех пор, пока абонент находится в зоне, контролируемой VLR.

В центре аутентификации *AUC* каждый подвижный абонент на время пользования системой связи получает стандартный модуль подлинности абонента (SIM), который содержит международный идентификационный номер (IMSI), свой индивидуальный ключ аутентификации и алгоритм аутентификации, с помощью которых и с учетом данных регистре EIR, проверяются полномочия абонента и осуществляется доступ последнего к сети связи.

База данных *EIR* кроме серийных номеров абонентских аппаратов содержит также следующие данные об абонентском оборудовании сети:

- белый список, содержащий номера IMEI, о которых есть сведения о том, что они закреплены за санкционированными подвижными станциями;
- черный список, содержащий номера IMEI подвижных станций, которые украдены или которым отказано в обслуживании по другой причине;

- серый список, содержащий номера IMEI подвижных станций, у которых существуют проблемы, выявленные по данным программного обеспечения, что не является основанием для внесения в «черный список».

Центр эксплуатации и технического обслуживания *ОМС* является центральным звеном сети GSM в регионе, который обеспечивает контроль и управление всеми компонентами региональной сети и контроль качества ее работы и включает в себя сбор статистических данных о нагрузке от компонентов сети GSM, изменения программного обеспечения и баз данных о конфигурации элементов сети, загрузка программного обеспечения в память других звеньев сети или *ОМС*.

ОМС соединяется с другими компонентами сети GSM по проводным каналам пакетной передачи протокола X25.

Центр управления сетью *НМС* позволяет обеспечивать рациональное иерархическое управление всей сетью GSM.

Оборудование *базовой станции BSS* состоит из контроллера базовой станции BSC, приемо-передающих устройств базовых станций BTS и транскодера TCE. Контроллер базовой станции может управлять несколькими приемо-передающими блоками. BSS обеспечивает следующие основные функции:

- управляет распределением радиоканалов;
- контролирует соединения;
- обеспечивает режим работы с прыгающей частотой;
- осуществляет модуляцию и демодуляцию сигналов;
- проводит кодирование и декодирование сообщений и речи;
- обеспечивает адаптацию скорости передачи речи и данных.

Транскодер *TCE* обеспечивает преобразование выходных сигналов канала передачи речи и данных MSC к виду, соответствующему радиointерфейсу GSM.

Стандарт GSM содержит интерфейсы для осуществления соединений с внешними сетями и между различным оборудованием сетей GSM.

Соединение с ГАТС, с сетью NMT-450, с общеевропейскими сетями GSM осуществляется MSC по линии связи 2 Мбит/с в соответствии с системой сигнализации SS№7.

Интерфейс между MSC и BSS (А-интерфейс) обеспечивает передачу сообщений для управления BSS, передачи вызова, управления передвижением. А-интерфейс объединяет каналы связи и линии сигнализации.

Интерфейс между MSC и HLR совмещен с VLR (В-интерфейс). Когда MSC необходимо определить местоположение подвижной станции, он обращается к VLR. Если подвижная станция инициирует процедуру местоопределения с MSC, он информирует свой VLR, который заносит

всю изменяющуюся информацию в свои регистры. Эта процедура происходит всегда, когда MS переходит из одной области местоположения в другую. Если абонент запрашивает специальные дополнительные услуги или изменяет некоторые свои данные, MSC также информирует VLR, который регистрирует изменения и при необходимости сообщает о них HLR.

Интерфейс между MSC и HLR (C-интерфейс) используется для обеспечения взаимодействия между MSC и HLR. MSC может послать сообщение HLR в конце сеанса связи для того, чтобы абонент мог оплатить разговор. Когда сеть фиксированной телефонной связи не способна исполнить процедуру установления вызова подвижного абонента, MSC может запросить HLR с целью определения местоположения абонента для того, чтобы послать вызов MS.

Интерфейс между HLR и VLR (D-интерфейс) используется для расширения обмена данными о положении подвижной станции, управления процессом связи. Основные услуги, предоставляемые подвижному абоненту, заключаются в возможности передавать или принимать сообщения независимо от местоположения. Для этого HLR должен пополнять свои данные. VLR сообщает HLR о положении MS, управляя и переприсваивая ей номера в процессе блуждания, посылает все необходимые данные для обеспечения обслуживания подвижной станции.

Интерфейс между MSC (E-интерфейс) обеспечивает взаимодействие между разными MSC при осуществлении процедуры эстафетной передачи абонента.

Интерфейс между BSC и BTS (A-bis интерфейс) служит для связи BSC с BTS. Передача осуществляется цифровыми потоками со скоростью 2,048 Мбит/с. Возможно использование физического интерфейса 64 кбит/с.

Интерфейс между BSC и OMC (O-интерфейс) предназначен для связи BSC с OMC и используется в сетях с пакетной коммутацией X25.

Внутренний BSC-интерфейс контроллера базовой станции обеспечивает связь между различным оборудованием BSC и оборудованием транскодирования (TCE), использует стандарт ИКМ-передачи 2,048 Мбит/с и позволяет организовать из четырех каналов со скоростью 16 кбит/с один канал со скоростью 64 кбит/с.

Интерфейс между MS и BTS (Um-радио интерфейс) определен Рекомендацией ETSI.

Совокупность интерфейсов стандарта GSM обеспечивает высокие параметры передачи сообщений, совместимость с существующими и перспективными информационными сетями, предоставляют абонентам широкий спектр услуг цифровой связи.

10.7.2. Службы сети стандарта GSM

ССПР стандарта GSM содержат два класса служб (рис. 10.8) сети PLMN (Public Land Mobile Network): основные службы и телеслужбы. На рис. 10.8 TE (Terminal Equipment) – терминальное оборудование, MT (Mobile Terminal) – подвижный терминал, IWF (Interworking Function) – функциональный межсетевой интерфейс.

Основные службы обеспечивают связь контроллеров MT между собой, в частности:

- асинхронную передачу данных в дуплексном режиме со скоростями 300, 600, 1200, 2400, 4800 и 9600 бит/с через линии ГАТС;
- синхронную передачу данных в дуплексном режиме со стандартными скоростями 1200...9600 бит/с через линии ГАТС, коммутируемые сети передачи данных общего пользования;
- асинхронную дуплексную передачу данных со стандартными скоростями 300...9600 бит/с через сети пакетной передачи данных общего пользования;
- синхронный дуплексный доступ к сети пакетной передачи данных со стандартными скоростями 2400...9600 бит/с.

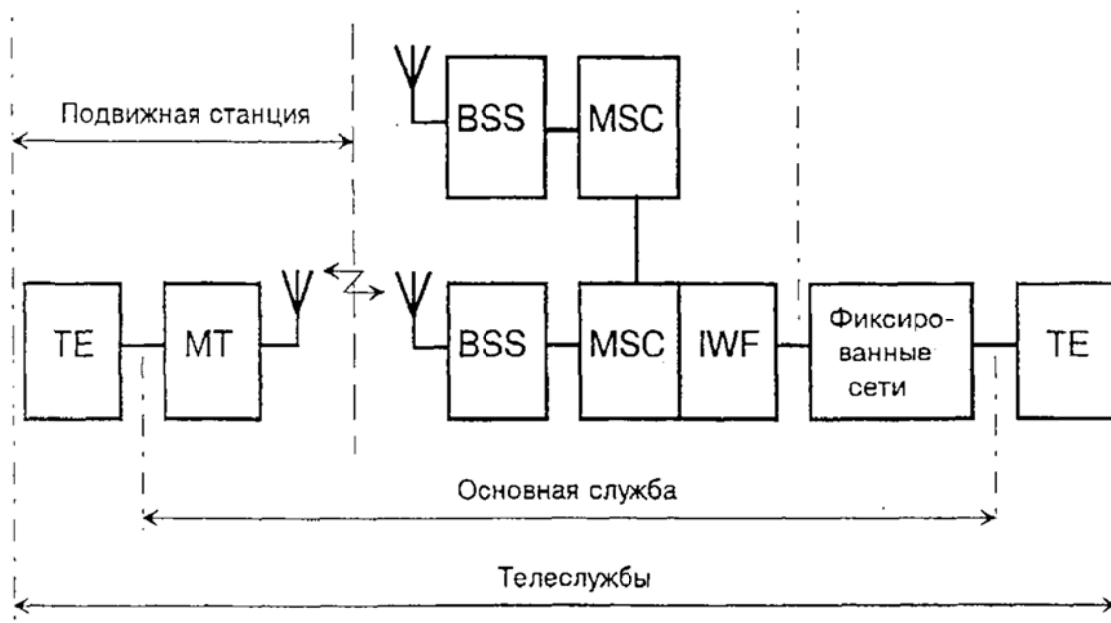


Рис. 10.8. Службы стандарта GSM

Телеслужбы обеспечивают связь между TE, управляемыми абонентами, а именно:

- спецсигнализацию (охрану квартир, сигналы бедствия и пр.);
- доступ к службам «Видеотекст» и «Телетекст»;

- доступ к оборудованию «Телефакс»;
- доступ к сети ИНТЕРНЕТ.
- особые услуги, например, оповещение о тарифных расходах, включение в закрытую группу пользователей, передача служебных буквенно-цифровых сообщений для отдельных групп пользователей.
- Поколение 2,5G стандарта GSM имеет следующие дополнительные услуги, позволяющие увеличить скорость передачи данных сетевых служб:
- HSCSD (High-Speed Circuit-Switched Data), при котором одному пользователю выделяется несколько физических каналов (временных окон TDMA-кадра), что увеличивает стандартную скорость 9600 бит/с в соответствующее число раз;
- GPRS (General Packet Radio Service) предусматривает организацию специального канала пакетной связи со скоростями до 115 кбит/с, но без помехоустойчивого кодирования;
- EDGE (Enhanced Data rates for Global Evolution) – введение наряду с гауссовской частотной манипуляцией с минимальным частотным сдвигом (GMSK) добавочного модуляционного формата восьмеричной ФМ, утраивающей скорость передачи при сохранении частотного ресурса.

Совместное применение этих мер позволяет получить скорость передачи данных до 384 кбит/с.

10.7.3. Организация логических каналов ССПР GSM

Различают физические и логические каналы связи. В стандарте GSM физические каналы реализованы в окнах TDMA. Сообщения и данные, представленные в цифровой форме, группируются и объединяются в логические каналы двух типов: каналы связи для передачи кодированной речи или данных TCH (Traffic Channel) и каналы управления для передачи сигналов управления и синхронизации CCH (Control Channel).

В стандарте GSM различают логические каналы связи двух основных видов:

- TCH/F (Full Rate Traffic Channel) – канал передачи сообщений с полной скоростью 22,8 кбит/с (другое обозначение V_m);
- TCH/H (Half Rate Traffic Channel) – канал передачи сообщений с половинной скоростью 11,4 кбит/с (другое обозначение L_m).

Один физический канал может представлять собой канал передачи сообщений с полной скоростью или два канала с половинной скоростью передачи. В первом случае канал связи занимает одно временное окно, во втором – два канала связи занимают то же самое временное окно, но с перемежением в соседних кадрах, т. е. каждый канал следует через кадр.

Для передачи кодированной речи и данных предназначены каналы связи следующих типов:

- TCH/FS (Full Rate Traffic Channel for Speech) – канал для передачи речи с полной скоростью;
- TCH/HS (Half Rate Traffic Channel for Speech) – канал для передачи речи с половинной скоростью;
- TCH/F 9,6 (Full Rate Traffic Channel for 9,6 kbit/s User Data) – канал передачи данных с полной скоростью 9,6 кбит/с;
- TCH/F 4,8 (Full Rate Traffic Channel for 4,8 kbit/s User Data) – канал передачи данных с полной скоростью 4,8 кбит/с;
- TCH/F 2,4 (Full Rate Traffic Channel for 2,4 kbit/s User Data) – канал передачи данных с полной скоростью 2,4 кбит/с;
- TCH/H 4,8 (Half Rate Traffic Channel for 9,6 kbit/s User Data) – канал передачи данных с половинной скоростью 4,8 кбит/с;
- CH/H 2,4 (Half Rate Traffic Channel for 9,6 kbit/s User Data) – канал передачи данных с половинной скоростью 2,4 кбит/с.

Различают четыре вида каналов управления:

- BCCH (Broadcast Control Channel) – каналы передачи сигналов управления;
- CCCH (Common Control Channel) – общие каналы управления;
- SDCCCH (Stand-alone Dedicated Control Channel) – индивидуальные каналы управления;
- ACCH (Associated Control Channel) – совмещенные каналы управления.
- Каналы передачи сигналов управления BCCH используются только в направлении с базовой станции на все подвижные станции. Они несут информацию, которая необходима подвижным станциям для работы в системе. Различают три вида каналов передачи сигналов управления BCCH:
- FCCH (Frequency Correction Channel) – канал подстройки частоты, который используется для синхронизации несущей в подвижной станции. По этому каналу передается немодулированная несущая;
- SCH (Synchronization Channel) – канал синхронизации, по которому передается информация на подвижную станцию о кадровой (временной) синхронизации;
- BCCH (Broadcast Control Channel) – канал управления передачей сигналов.

Используются три типа общих каналов управления CCCH:

- PCH (Paging Channel) – канал вызова, используется только в направлении от базовой станции к подвижной для ее вызова;
- RACH (Random Access Channel) – канал параллельного доступа, используется только в направлении от подвижной станции к базовой для запроса о назначении индивидуального канала управления;

- AGCH (Access Grant Channel) – канал разрешенного доступа, используется только для передачи с базовой станции на подвижную для выделения специального канала управления, обеспечивающего прямой доступ к каналу связи.

Выделенные индивидуальные каналы управления SDCCCH используются в двух направлениях для связи между базовой и подвижной станциями. Эти каналы предназначены для установки требуемого пользователем вида обслуживания. Различают два вида таких каналов:

- SDCCCH/4 (Stand-alone Dedicated Control Channel) – индивидуальный канал управления, состоит из четырех подканалов;
- SDCCCH/8 (Stand-alone Dedicated Control Channel) – индивидуальный канал управления, состоит из восьми подканалов.
- Различают два вида ACCH:
- FACCH (Fast Associated Control Channel) – быстрый совмещенный канал управления, служит для передачи команд при переходе подвижной станции из соты в соту, т. е. при «эстафетной передаче» подвижной станции;
- SACCH (Slow Associated Control Channel) – медленный совмещенный канал управления, по направлению «вниз» передает команды для установки выходного уровня мощности передатчика подвижной станции. По направлению «вверх» подвижная станция посылает данные, касающиеся уровня установленной выходной мощности базовой станции, а также измеренного приемником уровня радиосигнала и его качества.

10.7.4. Организация физических каналов ССПР GSM

Восемь физических каналов размещены в восьми временных окнах в пределах TDMA-кадра, при этом каждый физический канал использует одно и то же временное окно в каждом временном TDMA-кадре.

26 кадров длительностью 120 мс составляют мультикадр. Объединение каналов связи TCH с полной и половинной скоростями с медленным совмещенным каналом управления SACCH показано, например, на рис. 10.9 а, б, где на а – организация полноскоростного канала связи (TCH+SACCH); на б – организация полускоростного канала связи (TCH+SACCH); Т, t – данные или речь двух каналов связи TCH; А, а – данные двух каналов для передачи SACCH; «–» – пустой TDMA-кадр.

В полноскоростном канале связи в каждом 13-м TDMA-кадре мультикадра передается пакет информации канала SACCH; каждый 26-й TDMA-кадр мультикадра свободен. В полускоростном канале связи пакеты информации канала SACCH передается в каждом 13-м и 26-м TDMA-кадрах мультикадра.

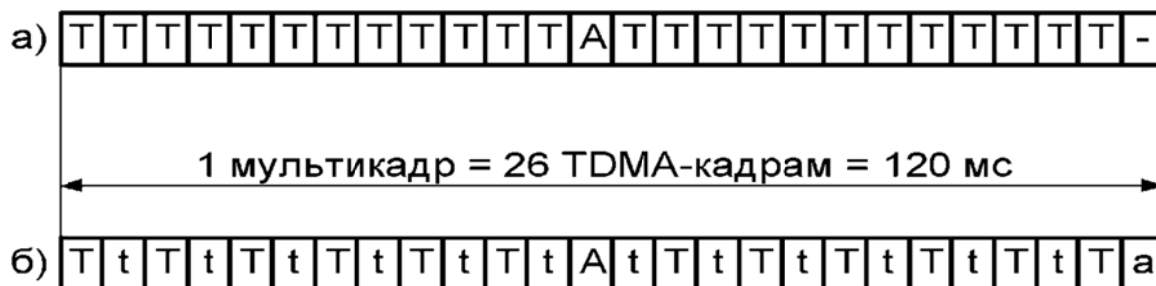


Рис. 10.9. Мультикадр

Для одного физического канала в каждом TDMA-кадре используется 114 бит. Так как в мультикадре для передачи канала связи TCH используется 24 TDMA-кадра из 26 и длительность мультикадра составляет 120 мс, общая скорость передачи информационных сообщений по TCH каналу составляет 22,8 кбит/с. Канал SACCH занимает в полноскоростном канале связи только один TDMA-кадр (114 бит) при скорости передачи по SACCH каналу 950 бит/с. Полная скорость передачи в объединенном TCH/SACCH канале с учетом пустого (свободного) 26-го TDMA-кадра составит $22,8 + 0,950 + 0,950 = 24,7$ кбит/с.

Как показано на рис. 10.9 за время 26-кадрового мультикадра (в одном физическом канале) может передаваться два полускоростных канала TCH, каждый по 12 TDMA-кадров (T и t). Пустой 26-й TDMA-кадр в полноскоростном канале TCH отводится для канала SACCH во втором полускоростном канале TCH. Для каждого полускоростного канала TCH скорость передачи составляет 11,4 кбит/с; полная скорость передачи в объединенном полускоростном канале TCH/SACCH остается прежней – 24,7 кбит/с.

Стандарт GSM разработан для создания ССПР в следующих полосах частот: 890...915 МГц – для передачи подвижными станциями (линия «вверх»), 935...960 МГц – для передачи базовыми станциями (линия «вниз»).

Каждая из полос, выделенных для сетей GSM, разделяется на частотные каналы. Разнос каналов составляет 200 кГц, что позволяет организовать в сетях GSM 124 частотных канала. Частоты, выделенные для передачи сообщений подвижной станцией на базовую и в обратном направлении, группируются парами, организуя дуплексный канал FDD с разносом 45 МГц. Эти пары частот сохраняются и при перескоках частоты.

В ССПР GSM используют абонентские станции 1 – 5-го классов с выходной мощностью 20, 8, 5, 2, 0.8 Вт, соответственно. При этом станции 1 – 3-го классов устанавливаются на транспортных средствах, а 4-го и 5-го представляют собой носимые модели.

Для передачи цифровых сообщений по физическим каналам используется гауссовская частотная манипуляция с минимальным частотным сдвигом (GMSK), которую отличают следующие свойства:

- постоянная по уровню огибающая, которая позволяет использовать эффективные передающие устройства с усилителями мощности в режиме класса С;
- компактный спектр на выходе усилителя мощности передающего устройства, обеспечивающий низкий уровень внеполосного излучения;
- хорошие характеристики помехоустойчивости канала связи.

10.7.5. Процедуры аутентификации и идентификации ССПР GSM

Аутентификация – процедура подтверждения подлинности (действительности, законности, наличия прав на пользование услугами сотовой связи) абонента системы подвижной связи. Необходимость введения этой процедуры вызвана неизбежным соблазном получения несанкционированного доступа к услугам сотовой связи, приводящим к многочисленным и разнообразным проявлениям особого рода мошенничества – фрода в сотовой связи. Слово аутентификация (английское authentication) происходит от греческого *authentikos* – подлинный, исходящий из первоисточника. В русском языке довольно часто используется родственный юридический термин – *аутентичные тексты*, например тексты договора на нескольких языках, имеющие равную силу.

Первоначально, в аналоговых системах сотовой связи первого поколения, процедура аутентификации имела простейший вид: подвижная станция передавала свой уникальный идентификатор (электронный серийный номер – Electronic Serial Number, ESN), и если таковой отыскивался среди зарегистрированных в домашнем регистре, то процедура аутентификации считалась успешно выполненной. Столь примитивная аутентификация оставляла большие возможности для фрода, поэтому со временем и в аналоговых системах, и тем более в системах сотовой связи второго поколения с использованием дополнительных возможностей цифровых методов передачи информации процедура аутентификации была значительно усовершенствована.

Идея процедуры аутентификации в цифровой системе сотовой связи заключается в шифровании некоторых паролей-идентификаторов с использованием квазислучайных чисел, периодически передаваемых на подвижную станцию с центра коммутации, и индивидуального для каждой подвижной станции алгоритма шифрования. Такое шифрование, с использованием одних и тех же исходных данных и алгоритмов, произ-

водится как на подвижной станции, так и в центре коммутации (или в центре аутентификации), и аутентификация считается закончившейся успешно, если оба результата совпадают.

В стандарте GSM процедура аутентификации связана с использованием модуля идентификации абонента (Subscriber Identity Module – SIM), называемого также SIM-картой (SIM-card) или смарт-картой (smart-card), о котором мы расскажем чуть подробнее, поскольку до сих пор такого повода нам не предоставлялось. Модуль SIM – это съемный модуль, напоминающий по внешнему виду пластиковую кредитную карточку и вставляемый в соответствующее гнездо абонентского аппарата. Модуль вручается абоненту одновременно с аппаратом и в принципе позволяет вести разговор с любого аппарата того же стандарта, в том числе с таксофонного. Модуль содержит персональный идентификационный номер абонента (Personal Identification Number – PIN), международный идентификатор абонента подвижной связи (International Mobile Subscriber Identity – IMSI), индивидуальный ключ аутентификации абонента K_i , индивидуальный алгоритм аутентификации абонента A3, алгоритм вычисления ключа шифрования A8. Для аутентификации используется зашифрованный отклик (signed response) S, являющийся результатом применения алгоритма A3 к ключу K_i и квазислучайному числу R, получаемому подвижной станцией от центра аутентификации через центр коммутации. Алгоритм A8 используется для вычисления ключа шифрования сообщений. Уникальный идентификатор IMSI для текущей работы заменяется временным идентификатором TMSI (Temporary Mobile Subscriber Identity – временный идентификатор абонента подвижной связи), присваиваемым аппарату при его первой регистрации в конкретном регионе, определяемом идентификатором LAI (Location Area Identity – идентификатор области местоположения), и сбрасываемым при выходе аппарата за пределы этого региона. Идентификатор PIN – код, известный только абоненту, который должен служить защитой от несанкционированного использования SIM-карты, например при ее утере. После трех неудачных попыток набора PIN-кода SIM-карта блокируется, и блокировка может быть снята либо набором дополнительного кода – персонального кода разблокировки (Personal unblocking key – PUK), либо по команде с центра коммутации.

Процедура аутентификации стандарта GSM схематически показана на рис. 10.10, где обозначены: R – случайное число; A3 – алгоритм аутентификации; A8 – алгоритм вычисления ключа шифрования; K_i – ключ аутентификации; K_s – ключ шифрования; S – зашифрованный отклик (Signed Response – SRES). Пунктиром отмечены элементы, не относящиеся непосредственно к процедуре аутентификации, но использу-

емые для вычисления ключа шифрования K_c . Вычисление производится каждый раз при проведении аутентификации.

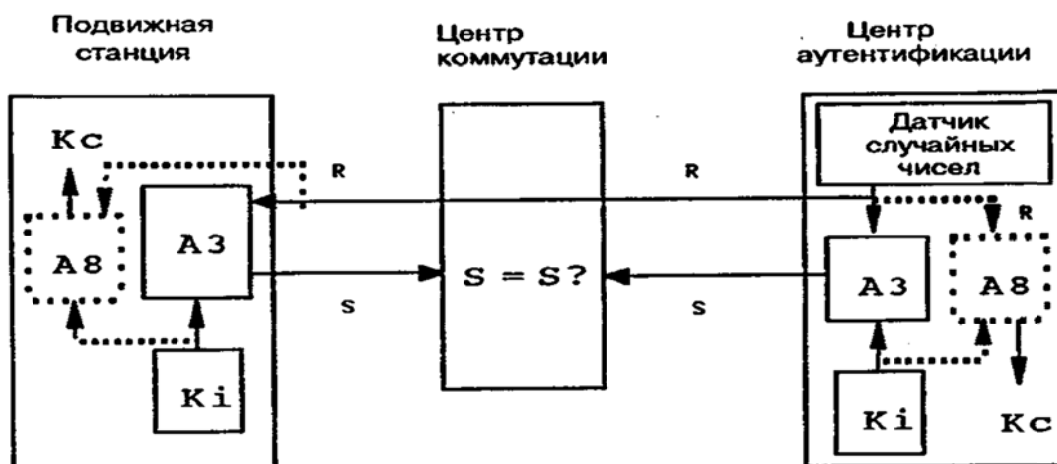


Рис. 10.10. Процедура аутентификации стандарта GSM

Идентификация – процедура отождествления подвижной станции (абонентского радиотелефонного аппарата), т. е. процедура установления принадлежности к одной из групп, обладающих определенными свойствами или признаками. Эта процедура используется для выявления утерянных, украденных или неисправных аппаратов. Слово идентификация (английское *identification*) происходит от средневекового латинского *identificare* – отождествлять.

Процедура идентификации заключается в сравнении идентификатора абонентского аппарата с номерами, содержащимися в соответствующих «черных списках» регистра аппаратуры, с целью изъятия из обращения украденных и технически неисправных аппаратов. Идентификатор аппарата делается таким, чтобы его изменение или подделка были трудными и экономически невыгодными.

10.8. Цифровые ССПР JDC

Структурная схема сети JDC или PDC показана на рис. 10.11.

В состав JDC входят три основных функциональных элемента: станция управления подвижной связью MCC (Mobile Communications Control Station), базовые и подвижные станции.

MCC является аналогом MSC в GSM и включает в себя подсистему управления связью подвижных абонентов и соединения с абонентами фиксированной телефонной сети (G-MCC), подсистему контроля перемещения абонентов (V-MCC), которая обеспечивает регистрацию местоположения абонентов и соединение вызовов, а также регистр поло-

жения (HLR), осуществляющий идентификацию подвижных абонентов и регистрацию зоны связи. Между MCC и основными ее элементами используется система сигнализации SS№7.

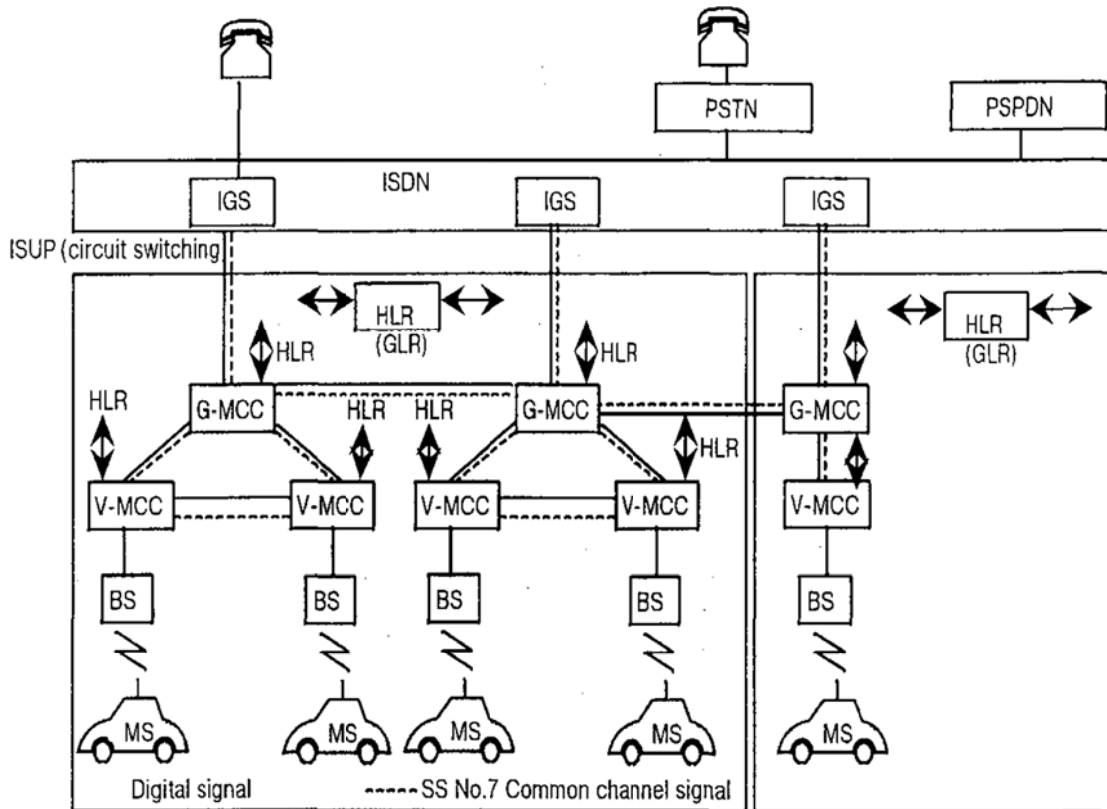


Рис. 10.10. Структурная схема сети JDC

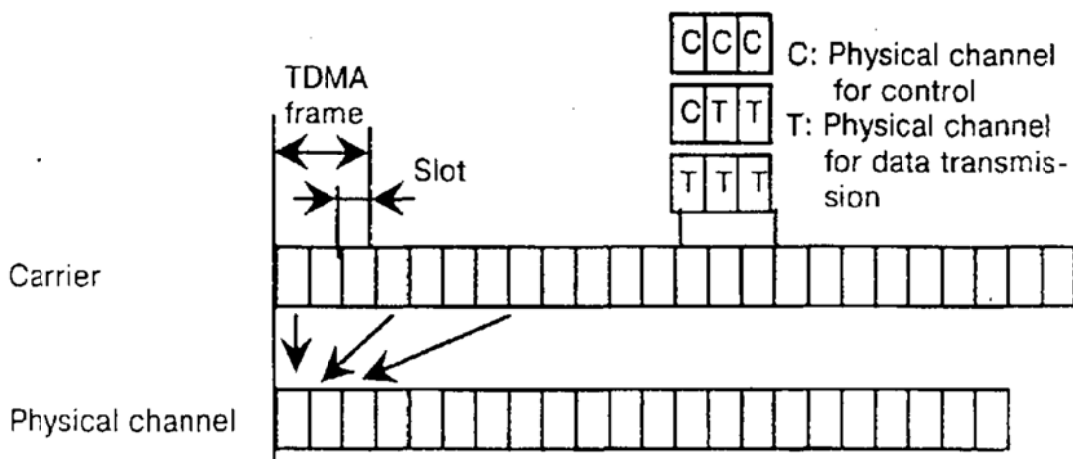


Рис. 10.11. Структура цифровых каналов связи JDC

Передача сообщений и управление связью осуществляются так же, как и в GSM, по каналам связи ТСН и управления ССН, которые формируются, передаются на радиоинтерфейс и передаются в физических

каналах. Полная структура цифровых каналов связи и управления в JDC показана на рис. 10.11. Назначение каналов аналогично стандарту GSM.

Физический канал в JDC формируется в одном из трех временных окон TDMA-кадра. Структура физического канала в JDC показана на рис. 10.12.

(1) Up ward

R	P	TCH (FACCH)	SW	CC	SF	SACCH (RCH)	TCH (FACCH)	G
4	2	112	20	8	1	15	112	6

(2) Down ward

R	P	TCH (FACCH)	SW	CC	SF	SACCH (RCH)	TCH (FACCH)
4	2	112	20	8	1	21	112

G	:Guard time	SACCH	:SACCH bits
R	:Ramp time	FACCH	:FACCH bits
P	:Preamble	RCH	:House keeping bits
SW	:Synchronization Word	SF	:Steal flag
CC	:Color Code		

Рис. 10.12. Структура физического канала в JDC

ССПР стандарта JDC обеспечивают:

- взаимодействие абонентов (терминалов) с фиксированными сетями связи (ГАТС, сетями пакетной связи и др.);
- взаимодействие подвижных абонентов (терминалов) с абонентами различных сотовых сетей;
- возможность шифрования передаваемых сообщений.

Основные технические показатели ССПР стандарта JDC схожи с показателями ССПР стандарта D-AMPS и отличаются дополнительным диапазоном 1400 МГц.

10.9. ССПР стандарта CDMA

ССПР CDMA-one стандарта IS-95 отличаются от одноименных СБАД только территориальной организацией с большим количеством базовых станций, с мелкими сотами и перекрытием больших регионов.

ССПР CDMA-1900 и CDMA-2000 являются очередным шагом эволюции систем стандарта CDMA. Их основные отличия от ССПР CDMA-one следующие:

- используется более высокочастотный диапазон 1850...1910 МГц и 1930...1990 МГц;
- для увеличения скорости передачи данных более 1 Мбит/с без потерь абонентской емкости в них расширен спектр сигналов от 1,25 МГц до 5 МГц двумя способами: традиционным DSSS с примене-

нием более высокочастотных ПСП и многочастотным, когда используются одновременно три полосы по 1,25 МГц на трех поднесущих;

- некогерентный вариант линии «вверх» от АС к БС заменен на когерентный, для чего организован обратный пилотный канал, позволяющий БС отслеживать частоту и фазу несущих каждой обслуживаемой АС;
- вместо бинарной ФМ в линии «вниз» от БС к АС для передачи данных используется КФМ, при этом длина посылки удваивается до 128, следовательно, объем ансамбля канализирующих ортогональных кодов Уолша и абонентская емкость становятся вдвое большими;
- в добавлении к действующему пилотному каналу линии «вниз» от БС к АС предусматриваются вспомогательные каналы управления диаграммой направленности антенны АС для осуществления пространственной селекции БС;
- помимо сверточных при высокоскоростной передаче применяют турбо-коды, повышающие помехоустойчивость радиолиний системы.

10.9.1. Функциональные возможности ССПР

Первоначально ССПР разрабатывались для обеспечения подвижных и стационарных абонентов речевой связью. При этом основные усилия разработчиков были направлены на:

- увеличение числа пользователей при ограниченном частотном диапазоне;
- расширение покрытия территории радиосвязью;
- повышение качества связи и скорости передачи данных;
- миниатюризацию и дизайн абонентских станций.

Однако со временем, когда основная часть вопросов была решена, радиоканалы ССПР стали использоваться для передачи неречевой информации и обеспечения дополнительных несвязных услуг. Ниже приведены некоторые примеры нестандартного использования ССПР.

Фотографии по телефону.

Две фирмы Siemens и Kodak внедрили новую услугу «фотокамера-телефон». Через специальную службу Photonet Online, созданную Kodak, владельцы мобильных телефонов получают доступ услуг мобильной фотографии. Пользователи смогут быстро пересылать свои фотоснимки друзьям, обмениваться фотографиями из интеллектуальных альбомов.

Штрих-коды по телефону

Сотрудники японской сотовой компании KDDI задумали оставить без работы театральные кассиров. Они решили доставлять билеты в ви-

де штрих-кода прямо на мобильные станции и обещают приступить к полномасштабной реализации такой услуги. Контроллер зрелищного учреждения сможет считать его специальным сканером, который в экстренном порядке принялась выпускать Hitachi.

Почерк по телефону

CDMA телефон SCH-i201 (Samsung) распознает английский текст, написанный от руки на дисплее аппарата. При желании можно нарисовать какую-нибудь небольшую схему (как пройти в библиотеку и т. д.), которую несложно отправить по протоколу SMS или электронной почтой в качестве прикрепленного файла. В трубке кроме традиционных реализована также функция голосового управления.

Сотовая автонавигация

Британский сотовый оператор Vodafone переключился на автомобили. Недавно эта компания изъявила желание оснащать новые форды системами связи и навигации. Они позволят водителям в оперативном режиме получать информацию о состоянии дорог, пробках, а в случае необходимости связываться со службами экстренной помощи.

Дистанционная медицина

Фирма Siemens выпустила сотовые телефонные трубки, у которых на задней стенке размещены четыре электрода. Приложив эти электроды к груди, и нажав кнопку, пользователь передает тщательно записанную кардиограмму в стационарный консультативный пункт. По этой же трубке он связывается с специалистами-кардиологами и получает медицинское заключение о состоянии своего здоровья.

Музыкально-навигационный телефон

Выпуск модели Videophone от Sanyo ознаменовал новый этап в развитии сотовых телефонов. В корпусе компактного мобильного телефона объединены видеокамера, MP3-плеер и GPS-приемник. В аппарате удачно сочетаются универсальность и качество.

Приведенные примеры применений ССПР показывают, что функциональные возможности таких «транспортных» средств для передачи информации без проводов далеко не полностью востребованы на данное время. В связи с тем, что аппаратура ССПР, как средств радиосвязи, имеет достаточно высокий уровень развития, основные перспективы совершенствования ССПР связаны именно с расширением сферы услуг и выполняемых функций этих радиосистем.

СПИСОК ЛИТЕРАТУРЫ

1. В.Г. Олифер, Н.А. Олифер Компьютерные сети. Принципы, технологии, протоколы: учебник для вузов. 2-е изд. – СПб.: Питер, 2004.
2. Джим Гейер Беспроводные сети. Первый шаг: пер. с англ. – М.: Издательский дом «Вильямс», 2005.
3. В.М. Вишневский, А.И. Ляхов, С.Л. Портной, И.В. Шахнович Широкополосные беспроводные сети передачи информации. – М.: Техносфера, 2005.
4. Шелухин О.И., Хизгилов В.А., Чивилев С.В. Системы радиодоступа / под ред. О.И. Шелухина. – М.: ГАСБУ, 1998.
5. Громаков Ю.А. Стандарты и системы подвижной радиосвязи. – М.: Эко-Трендз, 1997.
6. Карташевский В.Г., Семенов С.Н., Фирстова Т.В. Сети подвижной связи. – М.: Эко-Трендз, 2001.
7. Невдяев Л.М. Мобильная связь 3-го поколения / под ред. Ю.М. Горностаева. – М.: МЦНТИ, 2000.
8. Ратынский М.В. Основы сотовой связи / под ред. Д.В. Зимина. – 2-е изд. перераб. и доп. – М.: Радио и связь, 2000.
9. Системы мобильной связи: Учебное пособие для вузов / В.П. Ипатов, В.К. Орлов, И.М. Самойлов, В.Н. Смирнов; под ред. В.П. Ипатова. – М.: Горячая линия-Телеком, 2003.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	3
ГЛАВА 1. ПОНЯТИЕ ИНФОРМАЦИИ	9
1.1. Что такое информация	9
1.2. Формы представления информации	12
1.3. Измерение информации	16
ГЛАВА 2. СРЕДСТВА СВЯЗИ	20
2.1. Что такое средство связи	20
2.2. Параметры дискретных каналов связи	24
2.2.1. Ширина полосы пропускания	25
2.2.2. Пропускная способность канала связи	26
2.2.3. Скорость передачи информации	27
ГЛАВА 3. КАБЕЛЬНЫЕ ЛИНИИ СВЯЗИ	28
3.1. Типы кабельных линий связи	28
3.2. Параметры кабельных линий связи	34
ГЛАВА 4. ПЕРЕДАЧА ДАННЫХ ПО ВИТЫМ ПАРАМ	40
4.1. Способы передачи по витым парам	40
4.2. Первичные электрические параметры витой пары	43
4.2.1. Емкость	43
4.2.2. Активное сопротивление	44
4.2.3. Индуктивность	44
4.2.4. Проводимость изоляции	45
4.3. Вторичные параметры кабелей на основе витой пары	45
4.4. Защищенность	54
4.6. Структурные возвратные потери	58
4.7. Шум от внешних источников	59
ГЛАВА 5. ОПТОВОЛОКОННЫЕ ЛИНИИ СВЯЗИ	60
5.1. Типы оптических линий связи	60
5.2. Передача сигналов по волоконным световодам	61
5.3. Типы волоконных световодов	63
5.4. Дисперсия электромагнитного излучения	65
5.5. Затухание сигналов в световодах	67

5.6. Излучатели и фотоприемники	70
ГЛАВА 6. ОСНОВЫ БЕСПРОВОДНОЙ ПЕРЕДАЧИ	72
6.1. Антенны	72
6.1.1. Диаграмма направленности антенны	72
6.2. Типы антенн	73
6.2.1. Диполи	73
6.2.2. Параболическая антенна	74
6.2.3. Усиление антенны	76
6.3. Распространение радиоволн	77
6.4. Рефракция	81
ГЛАВА 7. МОДЕЛЬ ВЗАИМОДЕЙСТВИЯ ОТКРЫТЫХ СИСТЕМ И ПРОБЛЕМЫ СТАНДАРТИЗАЦИИ СЕТЕЙ	98
7.1. Многоуровневый подход. Протокол. Интерфейс. Стек протоколов	98
7.2. Модель OSI	101
7.3. Уровни модели OSI.	104
7.4. Сетезависимые и сетезависимые уровни	111
7.5. Стандартные стеки коммуникационных протоколов	113
ГЛАВА 8. СТЕК ПРОТОКОЛОВ TCP/IP	119
8.1. История и перспективы стека TCP/IP	119
8.2. Структура стека TCP/IP. Краткая характеристика протоколов	120
8.3. Адресация в IP-сетях	123
8.3.1. Три основных класса IP-адресов	124
8.3.2. Соглашения о специальных адресах: broadcast, multicast, loopback	125
8.3.3. Отображение физических адресов на IP-адреса	126
8.3.4. Отображение символьных адресов на IP-адреса: служба DNS	128
8.3.5. Автоматизация процесса назначения IP-адресов узлам сети – протокол DHCP	130
8.4. Протокол межсетевого взаимодействия IP	132
8.5. Маршрутизация с помощью IP-адресов	136
8.6. Структуризация сетей IP с помощью масок	144
8.7. Сегменты TCP	148

8.8.	Порты и установление TCP-соединений	148
8.9.	Концепция квитиования	149
8.10.	Формат сообщений TCP	151
8.11.	Адресация в IPv6	152
ГЛАВА 9. ПРИНЦИП ПОСТРОЕНИЯ КОМПЬЮТЕРНЫХ СЕТЕЙ		154
9.1.	Типы компьютерных сетей	154
9.2.	Локальные компьютерные сети (ЛКС)	156
9.2.1.	Классификация ЛКС	156
9.2.2.	Структура ЛКС	157
9.2.3.	Физическая среда передачи в локальных сетях	159
9.2.4.	Типы ЛКС	159
9.3.	Технология Fast Ethernet IEEE 802.3U	162
9.4.	Технология Gigabit Ethernet	164
9.5.	Технологии ATM	165
ГЛАВА 10. ОБЩИЕ СВЕДЕНИЯ О СОТОВЫХ СИСТЕМАХ ПОДВИЖНОЙ РАДИОСВЯЗИ		167
10.1.	Хронология развития сотовых систем	167
10.2.	Территориальная организация ССПР	169
10.3.	Частотная организация ССПР	170
10.4.	Способы борьбы с системными помехами ССПР	171
10.5.	ССПР стандарта NMT	176
10.6.	ССПР стандарта AMPS	179
10.7.	Цифровые ССПР стандарта GSM	180
10.7.1.	Структура ССПР GSM	180
10.7.2.	Службы сети стандарта GSM	184
10.7.3.	Организация логических каналов ССПР GSM	185
10.7.4.	Организация физических каналов ССПР GSM	187
10.7.5.	Процедуры аутентификации и идентификации ССПР GSM	189
10.8.	Цифровые ССПР JDC	191
10.9.	ССПР стандарта CDMA	193
10.9.1.	Функциональные возможности ССПР	194
СПИСОК ЛИТЕРАТУРЫ		196

Учебное издание

АЛХИМОВ Юрий Васильевич
КУЛЕШОВ В.К.

СОВРЕМЕННЫЕ КОММУНИКАЦИОННЫЕ СИСТЕМЫ

Учебное пособие

Издано в авторской редакции

Компьютерная верстка *К.С. Чечельницкая*
Дизайн обложки *О.Ю. Аршинова*

Подписано к печати 28.09.2011. Формат 60x84/16. Бумага «Снегурочка».


Печать XEROX. Усл. печ. л. 11,63. Уч.-изд. л. 10,52.

Заказ ____-11. Тираж 35 экз.



Национальный исследовательский Томский политехнический университет
Система менеджмента качества
Издательства Томского политехнического университета сертифицирована
NATIONAL QUALITY ASSURANCE по стандарту BS EN ISO 9001:2008



ИЗДАТЕЛЬСТВО  ТПУ. 634050, г. Томск, пр. Ленина, 30
Тел./факс: 8(3822)56-35-35, www.tpu.ru