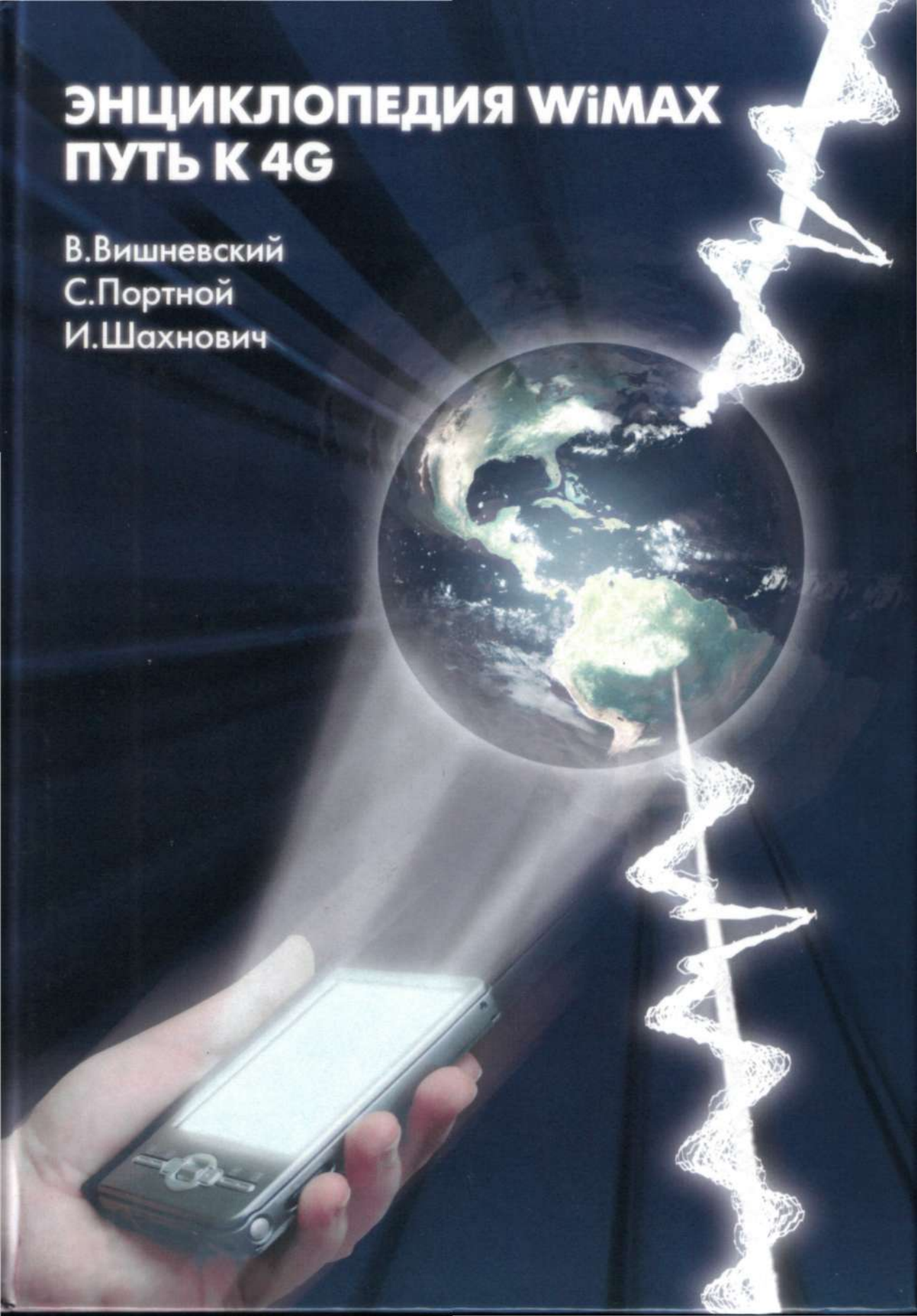
ЭНЦИКЛОПЕДИЯ WiMAX
ПУТЬ К 4G

В.Вишневский
С.Портной
И.Шахнович



МИР СВЯЗИ

В. Вишневский
С. Портной
И. Шахнович

Энциклопедия WiMAX
Путь к 4G

ТЕХНОСФЕРА
Москва
2009

Вишневский В. М., Портной С. Л., Шахнович И. В.
Энциклопедия WiMAX
Путь к 4G
Москва:
Техносфера, 2009. — 472 с. ISBN 978-5-94836-223-6

Книга написана известными специалистами в области беспроводных технологий. Издается при содействии и под эгидой WiMAX Forum.

В монографии описаны принципы построения, логическая и физическая структура беспроводных сетей передачи данных городского/регионального масштаба. Рассказано о беспроводных сетях IEEE 802.11, включая mesh-сети. Описана архитектура и принципы организации WiMAX-сетей (впервые в отечественной литературе). Детально изложена технология радиодоступа IEEE 802.16, включая мобильные сети (IEEE 802.16e). Описаны сотовые сети стандартов 3G и LTE (также впервые в отечественной литературе), а также технологии широкополосного цифрового теле- и радиовещания (DVB и DAB). Изложены принципы технологии MIMO. Приведены примеры реализации региональных WiMAX-сетей. Изложены теоретические основы передачи информации (теоремы Шеннона, Котельникова, Найквиста), методы кодирования и модуляции в беспроводных сетях.

В целом монография представляет собой уникальное справочное пособие по основным на сегодня технологиям широкополосного беспроводного доступа, охватывающего вопросы от архитектуры сетей до аппаратной реализации устройств и принципов сертификации оборудования. Сочетание как минимально необходимых теоретических основ беспроводных телекоммуникаций, так и описания конкретных стандартов, схемотехнических принципов построения поддерживающих их устройств и примеров реализации конкретных беспроводных сетей делает книгу полезной широкому кругу читателей, прежде всего — специалистам, занимающимся вопросами построения широкополосных беспроводных сетей, разработчикам телекоммуникационного оборудования, руководителям IT-отделов и аналогичных служб.

© 2009, Вишневский В. М., Портной С. Л., Шахнович И. В.
© 2009, ЗАО «РИЦ «Техносфера», оригинал-макет, оформление.

ISBN 978-5-94836-223-6

Содержание

Обращение к читателям

В настоящее время научно-техническое сообщество активно обсуждает концептуальные вопросы сетей связи следующих поколений. Поэтому как нельзя более своевременной стала книга известных российских специалистов В.М. Вишневского, С.Л. Портного, И.В. Шахновича «Энциклопедия WiMAX. Путь к 4G». Поставив перед собой непростую задачу — рассмотреть эволюцию беспроводных сетей передачи информации, авторы сумели отразить в монографии не только практически все технические аспекты этой проблемы (технологии, кодирование, системы модуляции, стандартизацию, безопасность и т. д.), но и вопросы частотного обеспечения, нормативно-правового регулирования и другие. В результате получилось уникальное инженерно-техническое издание, в котором сочетается строгий научный подход и простота изложения, что дало авторам полное право назвать ее «энциклопедией». И хотя статьи здесь расположены не в алфавитном порядке (как принято в энциклопедических изданиях), обширность материала и полнота сведений ставят эту книгу в один ряд с лучшими монографиями по телекоммуникациям. При этом авторам удалось избежать и другой крайности, которая подстерегала их при анализе многотомных спецификаций подкомитетов IEEE и других организаций — избыточности описания, ненужной детализации, мелких частных подробностей. Сегодня важно представить широкому кругу специалистов не столько подробную информацию о новых телекоммуникационных технологиях и оборудовании (ее, при желании, всегда можно отыскать в Интернете), сколько показать основные тенденции развития телекоммуникационных систем.

В сетях связи последующих поколений широкополосный доступ займет исключительно важное место, представляя собой одну из фундаментальных составляющих концепции NGN. К нему приковано особое внимание операторов связи и разработчиков нового оборудования. Деятельность по предоставлению услуг широкополосного доступа с большой долей вероятности станет новой крупной нишей телекоммуникационного рынка и уже одно это вызывает повышенный интерес к ним как со стороны различного рода инвесторов, так и государственных деятелей и политиков, регуляторов рынка. При рассмотрении проблем широкополосного доступа активно дискутируется вопрос о роли проводных, беспроводных и космических средств связи. Несмотря на различные взгляды на соотношение этих средств, эксперты едины во мнении, что широкополосные беспроводные сети на базе технологий сотовой связи третьего поколения, а также технологий Wi-Fi и WiMAX обладают сегодня исключительными преимуществами по оперативности развертывания, охвату территории,

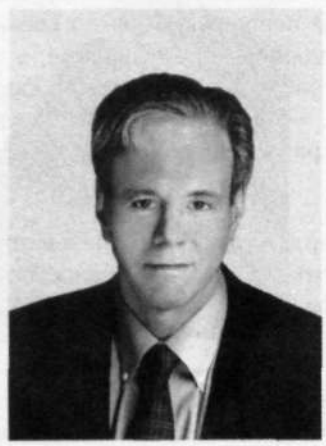
мобильности, представляя во многих случаях не только наиболее эффективное, но иногда и единственно возможное экономически оправданное решение. Это особо важно для Российской Федерации, с нашей необъятной территорией и значительным разбросом населения, особенно в сельской местности.

Завершается монография взглядом на путь к четвертому поколению сетей связи широкополосного беспроводного доступа. Здесь, конечно, не все так однозначно и просто, как представляется авторам. Научные споры по выбору технологий будущих сетей ведутся непрерывно, и доводы, приведенные в монографии, безусловно интересны всем, кому предстоит принять непростые решения.

Хочу выразить слова благодарности создателям представляемой монографии, выпуском которой они вносят ценный вклад в развитие телекоммуникаций России, повышение общего инженерного уровня персонала отрасли связи, разработчиков и производителей телекоммуникационного оборудования, а также в подготовку нового поколения специалистов отрасли. Появление в России книг, подобных «Энциклопедии WiMAX. Путь к 4G», — еще один, пусть маленький, но очень нужный шаг в продвижении нашей страны к современному информационному обществу.

Наум Семенович Мардер, доктор технических наук, заместитель Министра связи и массовых коммуникаций Российской Федерации

Address to the readers by the WiMAX Forum president Ron Resnick



Dear readers! Let me congratulate you on the new book «WiMAX encyclopedia. Way to 4G» appearance. Why is it so important nowadays?

The world population is 6.75 billion people, 4.05 of which use various types of telephone communication cellular networks, 3.5 billion use the Internet and only half of a billion use a broadband wireless access to diverse networks. Modern technologies and progress in general causes the rapid growth of broadband wireless access. However, the growth is restricted by various factors including necessity of large financial expenditures on the global networks infrastructure. It took mankind more than a century to develop a global telephone network; this very network was used for the Internet access at the first stage of its development. Howev-

Обращение к читателям Президента WiMAX-форума Рона Резника

Уважаемые читатели! Позвольте поздравить вас с появлением книги «Энциклопедия WiMAX. Путь к 4G». Почему она столь важна именно сегодня?

Население планеты Земля насчитывает 6,75 миллиарда человек, 4,05 из которых пользуются различными видами телефонной сотовой связи, 3,5 миллиарда пользуются Интернетом и только полмиллиарда обладают широкополосным мультимедийным доступом в различные сети. Современные технологии и прогресс человечества в целом диктуют стремительный рост широкополосного доступа. Однако этот рост сдерживают различные факторы, в том числе — необходимость огромных финансовых вложений в инфраструктуру всемирных сетей. Человечество затратило больше 100 лет для развития инфраструктуры всемирной телефонной сети общего пользования, именно эта сеть использовалась для доступа в Интернет на первом этапе его развития. Однако очень скоро стало ясно, что для широкополосного доступа нужны новые сети. Чтобы не строить их еще 100 лет, естественно использовать развитые технологии подвижной радиосвязи.

Таким образом, широкополосные беспроводные сети передачи информации становятся одним из основных направлений развития телекоммуникационной индустрии. А для стран, в которых большая территория сочетается с невысокой плотностью населения, беспроводные сети имеют особое значение,

er, it rapidly became apparent that novel networks are absolutely necessary for the broadband access. Not to spend another hundred of years on global network development it is natural to use the mobile radio communication technology.

Thus broadband wireless communication networks became one of the main directions of telecommunication industry development. For the countries with vast territories and low density of population, wireless networks have special significance since they make it possible to develop a large scale telecommunication infrastructure effectively and efficiently.

Even if WiMAX is not the only broadband wireless communication technology, it took in all the best from the borderline technologies like 3G, IEEE 802.11, DVB, DAB etc. That is why from our point of view it is WiMAX that is the best solution for broadband wireless internet access for people in both developing and developed countries. WiMAX technology enables both mobile and fixed line providers, including absolutely new ones starting from scratch, to progress rapidly. More than 475 WiMAX networks in 140 countries are developed already. There are numerous WiMAX providers in Russia and CIS. Nearly half of a billion people live in the regions where WiMAX providers operate or WiMAX licenses have effect. A year later the amount of people will reach the number of 800 million. International nonprofit association WiMAX Forum includes more than 400 companies — hardware components developers, OEM companies, all kinds of telecoms operators and WiMAX ecosystem.

That is why the book in your hands is so highly important and timely. WiMAX is a bunch of thriving technologies connected to almost all of the telecommunication areas. Most of the technologies from indispensable basics of communication theory to description of specific tech-

так как позволяют экономично и оперативно создавать телекоммуникационную инфраструктуру на обширных территориях.

Технологии WiMAX, хотя и не являются единственными технологиями широкополосной беспроводной связи, впитали в себя все лучшее, что есть в пограничных технологиях 3G, IEEE 802.11, DVB, DAB и др. И поэтому, на наш взгляд, именно они и являются наилучшим решением для обеспечения жителей Земли мобильным широкополосным доступом, как в развивающихся, так и в развитых странах. Технологии WiMAX позволяют быстро прогрессировать как мобильным, так и фиксированным операторам, в том числе — новым операторам, начинающим свою деятельность с чистого листа. Сегодня уже развернуто 475 WiMAX сетей в 140 странах. Много операторов в России и в других странах СНГ. Около полумиллиарда человек проживают в зоне действия операторов WiMAX или на территории, где действуют лицензии WiMAX. А через год таких людей будет 800 миллионов. Международная некоммерческая ассоциация WiMAX-форум объединяет более 400 компаний — разработчиков элементной базы, оборудования, операторов связи всех видов, а также «экосистему» WiMAX.

Вот почему книга, которую вы держите в руках, чрезвычайно важна и своевременна. Ведь WiMAX — это пакет бурно развивающихся технологий, связанных практически со всеми сферами телекоммуникаций. Практически все они полностью отражены в данной монографии — от необходимых основ теории передачи информации до описания конкретных технологий, аппаратуры и архитектуры сетей. И в этом отношении работа вполне по праву называется «Энциклопедия WiMAX».

nologies, equipment and network architecture are fully represented in this monograph. That is why this book is rightly called «WiMAX encyclopedia». I believe it is vital that authors didn't confine themselves to the only technology. There are descriptions of related and rival standards in the book as well, which make it possible for a thoughtful reader to arrive at a conclusion himself about the future of the global mobile telecommunications.

All stated above makes this book a unique and actual publication which will surely make a contribution to the development of the 4G networks. This book is equally useful for everyone who deals with development, using or adjusting of the broadband wireless systems, but first of all for those who deals with WiMAX. Go for it!

WiMAX Forum President
Ron Resnick

Мне представляется важным, что авторы не ограничились описанием только одной технологии. В книге рассказано о смежных и конкурирующих стандартах, что позволяет вдумчивому читателю прийти к самостоятельным выводам о перспективах развития глобальных мобильных телекоммуникаций.

Вот почему данная книга — это уникальная и своевременная работа, которая, безусловно, внесет свою лепту в дело развития сетей четвертого поколения. Она окажется одинаково полезной всем, кто так или иначе связан с разработкой, эксплуатацией и регулированием систем широкополосной беспроводной связи, но в первую очередь, конечно же, тем, кто связан с технологиями WiMAX. Дерзайте!

Президент WiMAX-форума
Рон Резник

Введение

Системы беспроводной передачи информации существуют столько же, сколько и сама человеческая цивилизация. Гонцы, стрелы, сигнальные костры, телеграф, искровые передатчики, спутниковые системы связи — все это звенья одной цепи. Изменялись технологии, но суть сетей передачи оставалась неизменной — организовать взаимодействие нескольких различных элементов так, чтобы информация без проводов в заданное время поступала из одной точки в другую. Однако, несмотря на почтенный возраст, беспроводные технологии в последние 15–20 лет развиваются чрезвычайно интенсивно, став одним из основных направлений развития телекоммуникационной индустрии.

Разделение на проводные и беспроводные технологии передачи информации в современном понимании началось в конце XIX века. К этому моменту уже окончательно оформились две ветви единого телекоммуникационного древа — передача голоса (телефония) и данных (телеграф). Однако проводная связь в ту эпоху оказалась проще, надежнее, защищеннее. Начался век проводных телекоммуникаций. Тысячи километров кабелей опутали землю, как паутина. Человечество потребляло все больше информации и все больше увязало в путах медной проволоки и кварцевого оптического волокна.

К концу XX века в технологии связи возникла новая волна — цифровая обработка. Вскоре практически любую информацию перед трансляцией, будь то речь или телевизионная картинка, стали преобразовывать в поток нулей и единиц. Настала эпоха цифровой связи. Благодаря цифровой обработке все теснее переплелись развивавшиеся параллельно технологии телефонии и передачи данных, чтобы с появлением пакетных сетей слиться практически воедино. Появился даже термин «мультимедиа», означающий объединение самых различных информационных технологий (голос, аудио/видео, данные) в единой технологической среде обработки и передачи. Взрывоподобное развитие Интернета лишь подтвердило тот факт, что цифровые сети для современной цивилизации стали столь же необходимы, как автострады, трубопроводы и линии электропередачи.

Локальные и региональные сети проникли во все сферы человеческой деятельности, включая экономику, науку, культуру, образование, промышленность и т. д. Технологию Ethernet (10 Мбит/с) сменили Fast Ethernet/Gigabit Ethernet/10Gigabit Ethernet (100/1000/10000 Мбит/с), в глобальных сетях свершился переход от неторопливой, но сверхнадежной технологии X.25 к Frame Relay, применению стека протоколов TCP/IP, к технологиям ATM и GigaEthernet. Без них невозможны столь привычные сегодня электронная почта, факсимильная и телефонная связь, доступ к удаленным базам данных в реальном масштабе времени, службы новостей, дистанционное обучение, телемедицина, телеконференции, телебиржи, телемагазины и т. д. Исчезни сегодня сети связи — и воцарит хаос. А ведь проводные линии связи так просто разрушить.

Наконец, в конце XX — начале XXI веков человечество начало вырываться из плена проводов. Уровень развития микроэлектроники позволил выпускать массовые дешевые средства беспроводной связи. Бум сотовой связи, сравнимый разве что с ростом производства персональных компьютеров, не замедляется вот уже четверть века. Мобильных телефонов во всем мире уже в несколько раз больше, чем обычных проводных телефонных аппаратов — номинально в абонентах

сотовых сетей числится 60% населения Земли. Фантастическими темпами развиваются технологии беспроводных локальных сетей, их догоняют персональные беспроводные сети и сети регионального масштаба. Видимо, все возвращается на круги своя — человечество тысячелетиями жило без проводов, хочет без них жить и впредь.

Но мало этого. Развиваются и персональные устройства обработки информации. Сегодня ноутбук, нетбук, наладонный компьютер — привычный аксессуар молодежи. А таким устройствам как пища нужна мультимедийная информация — всегда и везде. Поэтому сегодня мы переживаем период, когда на гребне третьей технологической волны зарождаются сети связи будущего — четвертого (4G) — поколения. Этот термин многие трактуют весьма вольготно. Но мы полагаем, что под 4G следует понимать вполне определенную группу технологий и стандартов, обслуживающих информационные потребности современных мобильных устройств обработки и отображения мультимедийной информации.

В последние десятилетия заметна миграция телекоммуникационных технологий в двух основных направлениях:

- от речевых услуг конечному пользователю к передаче скоростных потоков данных, которая в свою очередь уже делится на целый комплекс различных сервисов, включающих и речь, и данные, и видео;
- от неподвижных пользователей к кочующим (номадическим) и мобильным, что может обеспечить только беспроводная связь.

В монографии рассматривается область, которая является естественным пересечением двух указанных векторов развития телекоммуникационных технологий. Эту область сегодня называют сетями беспроводной связи 4G. Именно современным технологиям широкополосного беспроводного доступа и их переходу к 4G и посвящена данная монография. В качестве основной темы мы рассматриваем технологию WiMAX (Worldwide Interoperability for Microwave Access), но достаточно подробно останавливаемся на конкурирующих и смежных технологиях, таких как IEEE 802.11 (Wi-Fi) и стандарты 3GPP / 3GPP2, включая LTE.

Отметим, что современные телекоммуникационные технологии базируются на совокупности научных, технических и технологических достижений во многих областях, от микроэлектроники и схемотехники до теории связи, вычислительной техники и современных методов организации производства. Теория Максвелла оставалась мало кому понятной абстракцией до ее подтверждения опытами Герца. Кодовое разделение каналов и связь посредством шумоподобных сигналов не вышли бы из стен лабораторий и сложнейших военных систем, если бы не массовое появление дешевых процессоров цифровой обработки сигналов. Интернет остался бы ARPAnet'ом, если бы не лавиноподобное распространение персональных компьютеров и модемов. Сотовые телефоны и пейджеры, Wi-Fi-адаптеры и цифровое телевидение никогда не увидели бы свет без интеграции успехов в самых разных областях — технических, законодательных, организационных, научных и т. д.

Монография состоит из 11 глав, которые разбиты на четыре виртуальные части. Первая часть (главы 1–3) содержит необходимое описание основ теории передачи информации. В главе 1 мы предлагаем краткий исторический экскурс в историю беспроводной связи, а также вводим основные термины и понятия.

В главах 2 и 3 описаны методы кодирования информации, особенно корректирующего, а также методы модуляции и синтеза сигнально-кодовых конструкций. При рассмотрении методов корректирующего кодирования особое внимание уделяется новейшим и наиболее эффективным схемам, нашедшим свое применение практически во всех рассмотренных в монографии стандартах. К ним относятся три группы кодовых схем — каскадные коды Форни с внутренними сверточными кодами, декодируемыми по алгоритму Витерби, и внешними кодами Рида–Соломона; блоковые и сверточные турбокоды и низкоплотностные коды Галлагера (коды с малой плотностью проверок на четность). Сначала большое внимание уделяется синтезу сигнально-кодовых конструкций для гауссовского канала без памяти, и лишь затем — для каналов с переменными параметрами. Рассмотренные конструкции являются прообразом ортогонального частотного мультиплексирования (OFDM) и ортогонального частотного множественного доступа (OFDMA) — основы всех новейших стандартов беспроводной связи.

Вторая часть монографии (главы 4–6) содержит детальное описание той конкурентной среды, в которой развиваются технологии WiMAX. В главе 4 приводится описание стандартов цифрового видео- и радиовещания (DVB, DAB), причем впервые в отечественной литературе описан новейший стандарт второго поколения DVB-T2. В главе 5 подробно описаны сети стандарта IEEE 802.11, включая mesh-сети, а в главе 6 дается подробная справка о развитии стандартов сотовой связи, от первых аналоговых через GSM к стандартам третьего поколения 3G. Опять же, впервые в отечественной литературе приводится описание стандарта сотовой связи LTE.

Третья часть (главы 7–8) состоит из подробного описания самой группы стандартов IEEE 802.16 и соответствующих им технологий WiMAX. Впервые в отечественной литературе в одном месте рассмотрены новейшие стандарты IEEE 802.16 и соответствующие им спецификации архитектуры сети, вырабатываемые WiMAX-форумом, т. е. именно та полная группа объектов, которую и следует называть технологиями WiMAX.

Четвертая часть включает главы 9–11 и содержит технико-организационные основы технологий WiMAX, описание реализованных проектов и оборудования, а также сравнение с другими новейшими технологиями мобильного широкополосного доступа и выводы.

В целом, не являясь руководством разработчика и не заменяя описание стандартов, книга знакомит с основными понятиями и принципами современных беспроводных технологий радиосвязи и может стать основой для последующего глубокого изучения данного предмета. Особо отметим, что отдельные главы можно читать независимо друг от друга, поэтому монографию правомерно рассматривать как справочник по современным беспроводным технологиям передачи информации.

Авторы признательны Владиславу Михайловичу Тамаркину и Андрею Владимировичу Шурдаку за предоставленную информацию по безопасности сетей ШПД, а также Андрею Евгеньевичу Иванову и Светлане Николаевне Куприяхиной за помощь в работе.

Книга предназначена для широкого круга читателей — руководителей IT-отделов и подразделений, разработчиков аппаратуры, технических специалистов, связанных с телекоммуникациями, а также для студентов соответствующих специальностей и всех, кто интересуется современными технологиями связи.

ГЛАВА 1

БЕСПРОВОДНЫЕ СЕТИ ПЕРЕДАЧИ ИНФОРМАЦИИ. ИСТОРИЯ И ОСНОВНЫЕ ПОНЯТИЯ

Бурное развитие беспроводных сетей передачи информации в России и во всем мире, о котором многие говорят как о беспроводной революции в области передачи информации [1–3], связано с такими их достоинствами, как:

- гибкость архитектуры, т. е. возможности динамического изменения топологии сети при подключении, передвижении и отключении мобильных пользователей без значительных потерь времени;
- высокая скорость передачи информации (1–1000 Мбит/с и выше);
- быстрота проектирования и развертывания;
- высокая степень защиты от несанкционированного доступа;
- отказ от дорогостоящей и не всегда возможной прокладки или аренды оптоволоконного или медного кабеля.

Но прежде чем стать столь привычными, технологии беспроводной связи прошли более чем 150-летний путь развития. Чтобы более полно проникнутся величием того, что вскоре назовут связью 4G, кратко рассмотрим основные вехи строительства этого грандиозного здания, имя которому — глобальные беспроводные телекоммуникации.

1.1. Исторический очерк развития сетевых технологий

Беспроводные сети передачи информации, как следует из их названия, базируются на совокупности двух групп технологий — беспроводной передачи информации и сетевого взаимодействия. Исторически эти технологии зародились еще в позапрошлом веке. Родоначальником всех электронных сетей (систем) передачи данных, видимо, следует считать американского художника Самуэля Финли Бриза Морзе. В 1837 г. он разработал свою систему электросвязи по металлическому проводу и дал ей название «телеграф». Годом позже он дополнил ее знаменитой азбукой Морзе, т. е. механизмом кодирования источника, обязательным элементом всех современных сетей. 24 мая 1844 г. между Балтимором и Вашингтоном состоялся первый публичный сеанс телеграфной связи. Уже через 14 лет был проложен первый трансатлантический кабель, правда, просуществовал он лишь 26 дней.

В 1874 г. французский инженер Жан Морис Эмиль Бодо (Baudot) изобрел телеграфный мультиплексор, позволявший по одному проводу передавать до шести

телеграфных каналов. Значимость этого изобретения и авторитет Бодо были столь высоки, что, когда в 1877 г. другой французский инженер, Томас Муррэй, разработал первый в истории символьный телеграфный код с фиксированным размером символа (5 бит на символ), он назвал его кодом Бодо. Известный также под названием телексный код, он с незначительными изменениями применяется и сегодня (наиболее распространенная версия — стандартизированный Международным консультативным комитетом по телефонии и телеграфии (CCITT) Международный алфавит № 2). В честь Бодо названа и единица измерения скорости передачи телекоммуникационных символов (бод).

Следующий шаг сделали изобретатели телефона — профессор физиологии органов речи Бостонского университета Александр Грэйхем Белл при участии Томаса Ватсона (1875 г., приоритет от 14 февраля 1876 г.) и независимо от них — Элайша Грей в Чикаго. Последнему также принадлежит немалая роль в развитии сетевых технологий. Именно он в 1888 г. запатентовал Telautograph — первое устройство передачи факсимильных сообщений. Но это были лишь предпосылки сетей, а именно способы формирования канала связи и работы в нем. Сеть — это совокупность многих каналов, которыми необходимо управлять (коммутировать). В первых сетях, начиная с 1880 г., этим занимались телефонистки (вернее, телефонисты) методом установки штекеров в коммутационном поле.

С 1889 г. начался новый этап в развитии сетевых технологий — владелец бюро похоронных услуг из Канзас-Сити Элмон Браун Строуджер разработал систему автоматической коммутации каналов. Именно ему принадлежит приоритет в создании шагового искателя и декадно-шаговых АТС. Предание гласит, что Строуджер столкнулся с промышленной диверсией — жена его конкурента по цеху в Канзас-Сити работала телефонисткой и все звонки гробовщику направляла своему мужу. Видимо, это был один из первых в мире случаев электронного шпионажа. Он так возмутил Строуджера, что заставил изыскать способ избавиться от телефонисток на станции. Изобретение Строуджера оказалось столь удачным, что в 1891 г. он основал компанию Strowger Automatic Exchange (с 1901 г. — Automatic Electric, сегодня — отделение компании General Telephone and Electronics, GTE). Первая АТС этой компании емкостью 99 номеров была запущена в коммерческую эксплуатацию в 1892 г. (Ла-Порт, шт. Индиана). Примечательно, что на первых телефонных аппаратах для работы с АТС номер набирался посредством кнопок. В 1897 г. компания Строуджера представила прототип первого аппарата с дисковым номеронабирателем.

В 1885 г. произошло еще одно ключевое для сетевых технологий событие. Первые АТС обеспечивали одновременное соединение всех возможных пар абонентов. Очевидно, что при росте номерной емкости коммутационные матрицы становились невероятно дорогими и сложными. Впервые возникла проблема доступа к ограниченному коммутационному ресурсу. Ее разрешил российский инженер М. Ф. Фрейденберг, показавший, что для 10 тыс. абонентов достаточно обеспечить возможность одновременного соединения любых 500 пар. Отметим, что результат Фрейденберга справедлив и сегодня, для современных АТС: на 10 тыс. номеров допустимая вероятность предоставления соединения составляет 0,125. В 1895 г. М. Ф. Фрейденберг совместно с другим русским инженером С. М. Бердичевским-Апостоловым разработал и запатентовал в Великобритании АТС с предыскателем, выбиравшим свободный комплект линейных искателей при снятии абонентом трубки. Предыскатель и его принцип свободного поис-

ка стал основой для проектирования всех будущих АТС. Примерно с 1910 г. (к окончанию срока действия патента Строуджера) началось массовое внедрение электромеханических АТС. Работу, начатую М. Ф. Фрейденбергом, до логического завершения довел датский математик А. К. Эрланг, опубликовавший в 1909 г. ставшую классической работу «Теория вероятностей и телефонные переговоры» («The Theory of Probabilities and Telephone Conversations»), в которой предложил формулы для вычисления числа абонентов АТС, желающих одновременно вести разговоры.

Работы А. К. Эрланга положили начало нового научного направления — теории очередей (теории массового обслуживания), широко используемой первоначально для расчетов в телефонии, а затем при проектировании сетей передачи информации. Значительный вклад в развитие теории очередей внес выдающийся российский математик Александр Яковлевич Хинчин (Математическая теория стационарной очереди: Математический сборник, 1932, т. 39, № 4. О формулах Эрланга в теории массового обслуживания. Теория вероятностей и ее применения, 1962, т. 7, вып. 3.), выполнивший ряд оригинальных исследований для Московской телефонной сети.

В 1909 г. генерал-майор корпуса связи США доктор философии Джордж Оуэн Скваер изобрел способ посылки по телефонной линии нескольких радиограмм одновременно — родился метод частотного разделения каналов.

В 1928 г. американский физик-электрик и изобретатель Гарри Найквист в статье «Некоторые вопросы теории телеграфной передачи» («Certain Topics in Telegraph Transmission Theory») изложил принципы преобразования аналоговых сигналов в цифровые и сформулировал знаменитую теорему Найквиста. В СССР ее называли теоремой Котельникова, хотя Владимир Александрович опубликовал аналогичные результаты через пять лет после Найквиста. Но история все нивелирует — основополагающая теорема Клода Элвуда Шеннона о пропускной способности канала (1948) была сформулирована Котельниковым в его докторской диссертации годом раньше, в 1947 г. Однако у нас ее называют теоремой Шеннона.

В 1938 г. американец А. Х. Риверс патентует метод преобразования сигнала из аналоговой формы в цифровую для коммутации и передачи, названный импульсно-кодовой модуляцией (ИКМ). Этот метод впервые был практически реализован учеными из Bell Laboratories Клодом Шенноном, Джоном Р. Пирсом и Бернардом М. Оливером в быстродействующей цифровой передающей системе, позволившей транслировать несколько телефонных разговоров по одному каналу с высоким качеством, — появилась система с временным разделением (уплотнением) каналов.

Начиная с 1950-х годов сетевые и беспроводные технологии начали сближаться настолько тесно, что зачастую грань между ними провести уже трудно.

Беспроводные технологии также зарождались в XIX веке. Идея носилась в воздухе, вплотную к ней подошли такие ученые, как Г. Герц, О. Лодж, Э. Бранли. В 1892 г. английский ученый Вильям Крукс теоретически показал возможность и описал принципы радиосвязи. В 1893 г. сербский ученый Никола Тесла в США продемонстрировал передачу сигналов на расстояние. Тогда это событие не вызвало должного резонанса, возможно, потому, что Н. Тесла, работы которого существенно опережали время, интересовался беспроводной передачей на расстояние не информации, а энергии.

С 1878 г. над проблемой беспроводной связи работал преподаватель минных классов в Кронштадте Александр Степанович Попов. В 1884 г. он изобрел первую приемную антенну, создал прибор для регистрации грозовых разрядов на основе когерера — стеклянной трубки, заполненной металлическими опилками. Под воздействием электромагнитного поля проводимость этой трубки резко возрастала. 7 мая 1895 г. на заседании физического отделения Российского физико-химического общества состоялся его исторический доклад «Об отношении металлических порошков к электрическим колебаниям». Тогда А. С. Попов продемонстрировал свой прибор для регистрации грозовых разрядов («грозоотметчик») и высказал мысль о возможности его применения для беспроводной связи. Первая публичная демонстрация прототипа всех грядущих беспроводных систем состоялась 24 марта 1896 г. на заседании того же физико-химического общества. А. С. Попов передал на расстояние 250 м, возможно, первую в мире радиограмму, состоявшую из двух слов «Генрих Герц».

С 1894 г. успешно экспериментировал с физическими приборами для генерации и регистрации электромагнитных колебаний и двадцатилетний итальянский юноша Гульельмо Маркони, будущий нобелевский лауреат. В 1895 г. он установил связь на расстоянии порядка двух миль, а уже в 1896 г. запатентовал свое изобретение (в 1943 г. его патенты были аннулированы в пользу Н. Тесла [4]), в 1901-м установил радиосвязь через Атлантику.

В 1906 г. Ли де Форест создал первую электронную лампу (триод) — появилась возможность строить электронные усилители сигналов. С тех пор беспроводная связь развивалась — и продолжает по сей день — семимильными шагами, главным образом благодаря достижениям электроники. Отметим лишь основные вехи.

С 1920-х годов началось коммерческое радиовещание (посредством амплитудной модуляции). В 1933 г. Эдвин Ховард Армстронг изобрел частотную модуляцию (ЧМ), с 1936 г. началось коммерческое ЧМ-радиовещание. В 1946 г. компании AT&T и Bell System приступили к эксплуатации системы подвижной телефонной связи (MTS) для абонентов с автомобильными радиотелефонами (20 Вт). Для полудуплексной связи использовалось шесть каналов шириной по 60 кГц на частоте 150 МГц, однако из-за межканальной интерференции число каналов вскоре сократили до трех. Система позволяла соединяться с городской телефонной сетью.

12 августа 1960 г. был выведен на орбиту высотой 1500 км первый спутник связи — американский космический аппарат (КА) «Эхо-1» (Echo-1). Это был надувной шар с металлизированной оболочкой диаметром 30 м, выполнявший функции пассивного ретранслятора. Через два года, 10 июля и 13 декабря 1962 г., в США на низкие орбиты были запущены соответственно КА Telstar I и Relay-1 — первые спутники с активными ретрансляторами. Мощность их передатчиков не превышала 2 Вт. 19 августа 1964 г. впервые спутник связи был выведен на геостационарную орбиту. Это был также американский Syncom-3 (первые две попытки вывода в 1963 г. были неудачными). На следующий день был создан международный консорциум спутниковой связи Intelsat (International Telecommunications Satellite Organization), который стал крупнейшей международной организацией в области спутниковой связи. Сегодня ее услугами пользуются более чем в 200 странах, причем в начале 2001 г. 2/3 всего международного трафика передавалось через спутники Intelsat. 23 апреля

1965 г. был выведен на орбиту и начал успешно работать первый отечественный спутник связи «Молния-1» (также с третьей попытки). Мир вступил в эру спутниковой связи.

В истории сетевых технологий очередной этап начался в 1960-е годы и связан с массовым появлением компьютеров. Возникла потребность в передаче большого объема данных, зародилось понятие локальной вычислительной сети (ЛВС). Был разработан механизм коммутации сообщений (пакетов). В 1960-е годы над построением сети с коммутацией пакетов работали (параллельно, практически ничего не зная друг о друге) специалисты в трех организациях: в Массачусетском технологическом институте (MIT), корпорации RAND (тогда по сути центр стратегических исследований ВВС США, создавалась как подразделение компании Douglas Aircraft, с 1948 г. — независимая компания) и Национальной британской физической лаборатории (NPL). Пионерской работой в этой области явилась диссертация Леонарда Клейнрока на соискание степени доктора философии в MIT «Информационный поток в больших коммуникационных сетях» («Information Flow in Large Communication Nets», 1961). В 1964 г. была опубликована работа сотрудника корпорации RAND Пола Барана «О распределенных коммуникациях» («On Distributed Communications»). В ней были сформулированы принципы избыточной коммуникативности и показаны различные модели формирования коммуникационной системы, способной успешно функционировать при наличии значительных повреждений. В 1965 г. Лоуренс Робертс из MIT совместно с Томасом Меррилом связал компьютер TX-2 в Массачусетсе с ЭВМ Q-32 в Калифорнии по низкоскоростной коммутируемой телефонной линии. Так была создана первая нелокальная компьютерная сеть. Она убедительно продемонстрировала, что сеть с коммутацией соединений (каналов) неприемлема для таких задач.

В 1962 г. в журнале «Коммунист» (№ 12) появилась статья академика АН СССР Александра Александровича Харкевича «Информация и техника». В ней впервые в мире были сформулированы основные принципы создания единой сети связи (ЕСС), предугадана важность цифровых методов передачи и коммутации различных видов информации в цифровой форме. ЕСС, по мнению А.А. Харкевича, должна представлять собой крупнейший инженерный комплекс, объединяющий все существующие сети связи и развивающийся путем планомерного его наращивания в органическом взаимодействии с системой вычислительных, управляющих и справочных центров.

Знаковыми для сетевых технологий стали 1967–1968 гг. В NPL заработала первая ЛВС с пакетной коммутацией, во многом благодаря ее директору Дональду Дэвису. Сеть работала с пиковой скоростью — до 768 кбит/с (в начале 1970-х гг. она объединяла порядка 200 компьютеров со скоростью обмена до 250 кбит/с). В том же 1968-м г. сотрудник шведского отделения компании IBM Олаф Содерблюм разработал сеть Token Ring. МО США одобрило версию первого в мире стандарта на ЛВС — MIL-STD-1553 (протокол обмена данными по общему последовательному каналу посредством манчестерского линейного кода с выделенным контроллером (отечественный аналог — ГОСТ 26765.52–87). Этот стандарт после ряда модификаций до сих пор применяется в бортовых системах.

Но самое главное — в октябре 1967 г. был представлен начальный план сети ARPANET, развитием которой занимался департамент методов обработки

информации IPTO (Information Processing Techniques Office) агентства перспективных исследовательских проектов ARPA (Advanced Research Projects Agency) МО США. В декабре 1968 г. группа во главе с Фрэнком Хартом из компании Bolt, Beranek и Newman (BBN) выиграла конкурс ARPA на создание так называемого интерфейсного процессора сообщений (Interface Message Processor). В 1969 г. в рамках программы ARPANET в Калифорнийском университете в Лос-Анджелесе «отец» пакетной коммутации Леонард Клейнрок построил первый узел ARPANET — прообраз грядущего Интернета. В том же году компания BBN установила в Калифорнийском университете первый интерфейсный процессор сообщений и подключила к нему первый компьютер. Второй узел был образован в Стэнфордском исследовательском институте (SRI). Двумя следующими узлами ARPANET стали Калифорнийский университет в Санта-Барбара и Университет штата Юта. Эмбрион Интернета начал делиться.

В 1970 г. появилась первая пакетная радиосеть передачи данных (через спутник) — знаменитая ALOHA (aloha — приветствие в гавайском диалекте английского языка). Ее разработал и построил Норман Абрамсон (совместно с Франком Куо и Ричародом Биндером) из Гавайского университета. Сеть связывала различные университетские учреждения, разбросанные по отдельным островам Гавайского архипелага. В 1972 г. ALOHA соединили с сетью ARPANET. В ALOHA был реализован принцип подтверждения и повторной посылки пакетов (ARQ), а также механизм множественного доступа к каналу с контролем несущей CSMA. Тогда же начали развиваться проекты создания пакетных радиосетей, в том числе спутниковых.

В октябре 1972 г. известный специалист из компании BBN Роберт Кан на международной конференции по компьютерным коммуникациям впервые публично продемонстрировал работу сети ARPANET. В 1974 г. появляется статья Вирта Серфа (сотрудника Стэнфордского исследовательского института) и Роберта Кана (Cerf V.G., Kahn R.E. A protocol for packet network interconnection // IEEE Trans. Comm. Tech. Vol. COM-22. V. 5. May 1974. P. 627–641), в которой впервые была описана концепция протокола TCP/IP. В том же году компания BBN запустила первую открытую службу пакетной передачи данных (коммерческая версия ARPANET) — известный сегодня любому специалисту Telnet.

В 1973 г. сотрудник исследовательского центра компании Xerox в Пал-Альто Роберт Метклаф, до прихода в Xerox защитивший в MIT докторскую диссертацию в области теории пакетной передачи информации и участвовавший в создании сети ARPANET, представил своему руководству докладную записку, в которой впервые появилось слово Ethernet (эфирная сеть). В том же году Метклаф совместно с Дэвидом Боггсом построил первую Ethernet-ЛВС, связывавшую два компьютера со скоростью 2,944 Мбит/с. В основу технологии Ethernet был положен усовершенствованный принцип CSMA/CD с обнаружением коллизий. Через шесть лет, в 1979 г., при активном участии Р. Метклафа три ведущие в своих областях компании США — Xerox, Intel и Digital Equipment (DEC) — начали процесс стандартизации протокола Ethernet, успешно завершившийся через год. В том же 1979 г. Метклаф при участии DEC основал знаменитую компанию 3COM для выпуска Ethernet-совместимого оборудования.

В 1976 г. CCITT выпустила рекомендацию X.25, которая стала первым и чрезвычайно успешным стандартом сети с пакетной передачей данных по вы-

деленному каналу (Interface between DTE and DCE for Terminal Operations in Packet Mode and Connected to Public Data Networks by Dedicated Circuit). Массовая пакетная коммуникация стала реальностью.

В 1977 г. будущий вице-президент компании Sony Марио Токорои и другой японский ученый Киичироу Тамару предложили метод адаптации технологии Ethernet к передаче данных через радиоканал посредством механизма подтверждений (Acknowledging Ethernet). Эта работа заложила основу будущих беспроводных ЛВС (IEEE 802.11 и IEEE 802.15).

В 1978 г. в Бахрейне телефонная компания Batelco (Bahrain Telephone Company) начала эксплуатацию коммерческой системы беспроводной телефонной связи, которая считается первой в мире реальной системой сотовой связи. Две зоны с 20 каналами в диапазоне 400 МГц обслуживали 250 абонентов. Использовалось оборудование японской компании Matsushita Electric Industrial. В том же году в Чикаго компания AT&T начала испытания сотовой системы Advanced Mobile Phone Service (AMPS), работающей в диапазоне 800 МГц. Сеть из 10 зон охватывала связью 54 тыс. км^2.

В 1977 г. Деннис Хайес основал компанию Hayes Microcomputer Products и выпустил на рынок первый массовый модем Micromodem II для персональных компьютеров (Apple II). Он работал со скоростью 110/300 бит/с и стоил 280 долл. В 1979 г. в Женеве CCITT утверждает первую модемную рекомендацию V.21, определяющую стандартный протокол модуляции на скорости 300 бит/с.

Новый этап начался в 1980 г., когда стек протоколов TCP/IP был принят в качестве военного стандарта США. Годом раньше пакетная радиосеть заработала на военной базе США Форт-Брэгг. В 1983 г. сеть ARPANET была переведена на протокол TCP/IP взамен действовавшего изначально NCP. Из ARPANET, которую вскоре все стали называть Интернетом, выделилась сеть MILNET, обслуживающая оперативные нужды МО США.

События периода 1960-х годов в области сетевых технологий описаны во множестве книг, воспроизводить которые здесь невозможно, да и не нужно. За каждой датой, за каждым событием стоят напряженная работа и выдающиеся достижения специалистов всего мира. В это время сетевые технологии непрерывно развивались в сторону повышения быстродействия и надежности сетей передачи информации, возможности интегрированной передачи данных, голоса и видеоинформации. Так, в области локальных сетей было создано семейство технологий Ethernet-Fast Ethernet-Gigabit Ethernet, обеспечивающих иерархию скоростей 10/100/1000 Мбит/с. В глобальных сетях произошел переход от технологии X.25 к технологии Frame Relay, использованию стека протоколов TCP/IP, ATM и Gigabit Ethernet.

Важно отметить, что и в СССР работало немало выдающихся ученых и специалистов в области систем связи, в том числе и беспроводной. Уже в 1970–1980-х годах проектировались и строились современные сети связи, например, система цифровой телефонной связи «Кавказ-5», многочисленные ведомственные сети связи. Хорошо известны системы «Сирена» (первая в СССР гражданская сеть пакетной коммутации) и «Экспресс» для автоматизации бронирования и продажи авиа- и железнодорожных билетов соответственно. Но, видимо, закрытость как самих работ, так и общества никак не согласовывалась с концепцией открытых сетей. Возможно, именно поэтому изначально созданная на деньги МО США

открытая сеть Интернет завоевала весь мир, породила множество сетевых технологий, стимулировала развитие смежных отраслей, прежде всего разработку соответствующей аппаратуры и элементной базы для нее, т. е. микроэлектронику.

Видимо, именно Интернету мы исторически обязаны тем, что сегодня беспроводные сети получили столь бурное развитие. Их появление было бы невозможно без соответствующей полупроводниковой элементной базы. А она, в свою очередь, не может появиться, если нет массового (многомиллионного) спроса. Историческая заслуга и гениальное провидение тех, кто в 1960-е годы начиная работы по сетям пакетной передачи, в том, что они изначально сумели сформулировать принципы будущей глобальной сети и воплотили их. Тем самым был создан рынок устройств для работы в сети, ставший основой для промышленности и науки в этой области. Не случайно первым директором (с 1962 г.) департамента IPTO в ARPA, т. е. человеком, руководившим финансированием научных исследований в области компьютерных сетей, был психолог из Массачусетского технологического института Джозеф Карл Ликлайдер. Еще в начале 1960-х годов он сумел предвидеть появление глобальной сети взаимосвязанных компьютеров. Ему принадлежит ряд публикаций о концепции «галактической сети» (*Licklider J.C.R.* // On-Line Man Computer Communication, August 1962).

Разумеется, не менее основополагающим для беспроводных сетей стало массовое появление персональных компьютеров и развитие сотовой телефонии, а также стремительное развитие полупроводниковых технологий (создание дешевых сигнальных процессоров и микроконтроллеров, аналоговых СВЧ интегральных схем).

С 1989 года началось развитие стандартов IEEE 802.11 беспроводных локальных сетей. Постепенно они получили широчайшее развитие и стали использоваться на «последней миле» и для создания сетей беспроводного доступа регионального масштаба. Отдельно развивалась линия WLL (беспроводной абонентский доступ) — технология, разработанная компанией AT&T. WLL использовала пакетную передачу голоса и данных со скоростью 128 кбит/с, являясь узкополосной беспроводной системой с временным распределением. Ряд подобных систем был коммерчески доступен в 1990-е годы. Желание совместить широкополосность стандарта IEEE 802.11 и надежность систем операторского класса в лицензионных диапазонах привело к созданию стандарта IEEE 802.16. Изначально он задумывался как технология фиксированного доступа — транспортная сеть распространения информации регионального масштаба. Однако со временем, во многом благодаря усилию международной организации WiMAX-форум, этот стандарт превратился в технологию мобильного широкополосного доступа, т. е. стал ориентированым на конечных пользователей. Таким образом, у технологий сотовой связи появился мощный конкурент. Те не замедлили с адекватным ответом. В результате сегодня мы говорим о зарождении технологий связи четвертого поколения (4G).

1.2. Классификация и технологии беспроводных сетей

Классификация чего бы то ни было — задача неблагодарная, поскольку и критериев классификации можно разработать достаточно много, и реальные объекты могут не укладываться в четкие границы определенного класса, да и по ме-

ре развития устоявшиеся системы классификации могут устаревать. Все это справедливо и для беспроводных сетей передачи информации (БСПИ). Поэтому остановимся на наиболее популярных способах ранжирования различных беспроводных систем. Обычно БСПИ подразделяют по:

- способу обработки первичной информации — на цифровые и аналоговые;
- ширине полосы передачи — на узкополосные, широкополосные и сверхширокополосные;
- локализации абонентов — на подвижные и фиксированные;
- географической протяженности — на персональные, локальные, региональные (городские) и глобальные;
- виду передаваемой информации — на системы передачи речи, видеоинформации и данных.

Вполне справедливы и системы градации на основе используемой технологии (спутниковые сети, атмосферные оптические линии и т. п.), по назначению и др.

Все рассматриваемые в нашей монографии технологии относятся к цифровым беспроводным широкополосным системам. Приведем их отличительные признаки, охарактеризовав и «сопредельные» системы. Термин «беспроводность» определяется легко — отсутствует соединительный провод (оптоволоконный или медный кабель). Также относительно просто определить, цифровая система или нет. К цифровым относят системы, у которых входная аналоговая информация (например, голос, аналоговый телевизионный сигнал и т. п.) первоначально преобразуется в цифровую (дискретную) форму. Однако уже здесь возникает некоторая нечеткость. В самом деле, любой сигнал при передаче через физический канал имеет чисто аналоговый вид, он в принципе не должен быть дискретным (чем дальше форма сигнала от бесконечной синусоиды, тем больше паразитных гармоник и связанных с ними неприятностей), чего добиваются специальными методами. Поэтому термин «цифровая система» говорит только о том, что в ней входящие аналоговые данные оцифрованы и обрабатываются (фильтрация, скремблирование, коммутация) преимущественно цифровыми методами.

Еще сложнее с шириной полосы. Строгого определения тут нет. С технической точки зрения обычно полагают, что если ширина спектральной полосы F, в которой работает система, много меньше центральной частоты этой полосы f_c, то система узкополосная (т. е. $F/f_c \ll 1$). В противном случае система широкополосная. Критерий весьма расплывчат. В области цифровых систем передачи приводят и другие определения широкополосности [5]: например, система широкополосная, если передаточная функция канала в этой полосе существенно изменяется в зависимости от частоты (т. е. передаточная функция в рабочей полосе узкополосной системы практически не зависит от частоты). Очевидно, что определения эти достаточно расплывчаты.

С пользовательской точки зрения широкополосным доступом называют доступ к ресурсам с некой «достаточной» скоростью, причем эта скорость постоянно увеличивается. Еще не так давно к широкополосным относили скорости в 64 кбит/с, а скоро уже мегабитные скорости не будут являться широкополосными в полном смысле. Поэтому под термином «широкополосная система» мы будем понимать такие системы, где проявляются специфические эффекты и свойства,

связанные с широкой рабочей полосой частот, на уровне 1,25–40 МГц и выше. Более строгий критерий едва ли возможен.

Подразделение на мобильные и подвижные системы, казалось бы, столь простое, на самом деле также не является тривиальным. Следует различать собственно возможность мобильности абонентов, предоставляемую технологией, и подразделение на мобильную и фиксированную службы связи, связанное с вопросами частотного распределения и лицензирования. Наиболее характерным примером такой двусмысленности является история появления в России беспроводной телефонной связи стандарта IS-95 (CDMA). Оборудование этого стандарта изначально было разрешено к использованию в нашей стране только для предоставления услуг фиксированной связи. Однако, как известно, IS-95 является стандартом мобильной сотовой связи.

Технологически его никак нельзя «зафиксировать». Аналогичная неопределенность сложилась сейчас и в спутниковой связи. Если же говорить с технической точки зрения, ограничивать мобильность может чувствительность технологии связи к скорости движения абонента, сложность перехода из одной зоны обслуживания в сопредельную без разрыва связи, восприимчивость к кратковременным пропаданиям связи и т. п.

Подразделение по размеру зоны обслуживания также достаточно условно, если рассматривать соседние градации. К персональным сетям (WPAN — wireless personal area network) относят системы с радиусом действия от сантиметров до нескольких метров (до 10–15 м). Основное назначение таких сетей состоит в замещении кабельной системы для связи оборудования (например, компьютера и периферийных устройств). При этом мощность излучения передатчиков, как правило, 1–10 мВт. Локальные сети (WLAN — wireless local area network) подразумевают взаимную удаленность устройств на расстояние до сотен метров и мощности передатчиков порядка 100 мВт. Это сети, предназначенные для объединения устройств в пределах локальной зоны (здания, предприятия и т. п.). Отметим, что на основе стандартов локальных беспроводных сетей вполне успешно строят и сети городского масштаба. Например, в этом качестве используют такие технологии, как DECT и IEEE 802.11.

К сетям городского масштаба (региональным) можно отнести множество различных технологий. Это и наземное теле- и радиовещание, и сотовая связь, и транкинговые системы. Изначально стандарт IEEE 802.16 также задумывался как система региональной (городской) связи. Если же говорить о глобальных беспроводных системах передачи данных, то они представлены спутниковыми системами связи. Однако с учетом того, что, например, практически все сети сотовой телефонии так или иначе связаны друг с другом, все они разрабатываются с учетом возможности взаимодействия, можно (правда, с некоторой натяжкой) говорить и о глобальных сотовых сетях. Аналогична ситуация и с развитием IEEE 802.16 — сети мобильного WiMAX претендуют именно на глобальность.

Особой градацией является подразделение в зависимости от типа передаваемой информации, например, на системы передачи речи (или видеоинформации) и несинхронных данных. С одной стороны, речь — это один из видов информации. После оцифровки поток речевых данных по виду неотличим от потока любой другой информации. Развитие цифровых технологий в различных областях телекоммуникаций (например, в проводной телефонии) давно проде-

монстрировало эффективность цифровых методов обработки, когда и речь, и данные обрабатываются едиными способами. С другой стороны, потребность в информации разного вида уже сделала реальной интеграцию различных информационных сетей (телефония, телевидение, сети передачи цифровых данных, телеметрия) на бытовом уровне. По единому каналу передаются данные самой различной природы. Поэтому можно достаточно уверенно предположить, что недалек тот день, когда вся речевая информация будет обрабатываться исключительно цифровыми методами. Здесь можно было бы остановиться, но возникает важный нюанс. Каждому виду информации свойственны характерные требования при передаче. Человек чувствует задержку передачи речи, когда она превышает 0,25 с. При задержках около 0,5 с восприятие речи для многих становится неприемлемым. Причем дело не только собственно в задержке, но и в неизбежном при дуплексной связи эхо-сигнале, который при таких задержках устранить крайне сложно. С другой стороны, речевая информация малочувствительна к спорадическим помехам и потерям данных. Это означает, что при пакетной передаче речи важно, чтобы задержки распространения сигнала в канале были минимальными, а маршрутизация и восстановление потока данных из пакетов (даже если их последовательность нарушена) происходили в реальном времени. При этом допустима даже потеря отдельных пакетов. Аналогична ситуация и с передачей видеоинформации — задержка между приемом отдельных пакетов (например, MPEG-2) не должна превышать некоего заданного значения, но потеря пакета, как правило, допустима. Совершенно иные требования предъявляются к передаче телеметрической информации, текстовых данных и т. п. Здесь, как правило, не важен режим реального времени (в определенных пределах), но и недопустима потеря данных. Учет этих особенностей может приводить к созданию особых технологий, ориентированных на трансляцию определенных видов информации. Характерным примером было появление технологии Frame Relay — способа пакетной передачи, при котором не происходит проверок прохождения отдельных пакетов (в отличие от традиционных сетей пакетной коммутации X.25 с подтверждением и повторной передачей каждого пакета). В современных мультимедийных сетях для передачи разнородных данных необходимо введение дополнительных механизмов — приоритезации данных, системы обеспечения качества услуг (QoS) и т. п.

Приведенные выше рассуждения показывают, что любое определение, так или иначе ранжирующее БСПИ, не стоит воспринимать буквально и уж тем более не надо удивляться применению той или иной технологии «не по назначению».

1.3. Стандартизация в области телекоммуникаций

Важнейшим аспектом развития современных телекоммуникационных систем является их стандартизация. Стандартизация необходима всем обитателям мира телекоммуникаций, включая производителей электронных компонентов, изготовителей аппаратуры, разработчиков сетей и конечных пользователей. Прежде всего, стандартизация означает массовость производства, что ведет к низким ценам и широкому распространению технологии. Разумеется, выбор и утверждение стандарта — это процесс не только технический, но и политический. Как правило, различные фирмы прорабатывают альтернативные варианты будущей

технологии. От того, какой из них будет утвержден в качестве стандарта, зависят и объемы будущих прибылей. Поэтому, чтобы стандарт действительно стал общепризнанным, стандартизирующая организация должна быть чрезвычайно авторитетной, а сама процедура утверждения — максимально открытой и беспристрастной.

«Головной» организацией в мире в области стандартизации в телекоммуникациях является Международный союз электросвязи (ITU — International Telecommunications Union), работающий под эгидой ООН. После реорганизации 1 марта 1993 г. два его сектора вобрали в себя три важнейшие стандартизирующие организации. В сектор радиосвязи (ITU-R, ITU Radiocommunication Sector) вошли Международный консультативный комитет по радиовещанию (CCIR — International Radio Consultative Committee) и Международный комитет по регистрации радиочастот (IFRB — International Frequency Registration Board). До этого ключевую роль в распределении спектра играли именно эти организации. Они также занимались вопросами спутниковых систем связи, глобальных радиосистем и др. Сектор телекоммуникаций Международного союза электросвязи ITU-T (ITU Telecommunication Standardization Sector) стал преемником Международного консультативного комитета по телеграфии и телефонии (CCITT — Consultative Committee for International Telephone and Telegraphy), долгие годы издававшего свои знаменитые «цветные книги» — сборники стандартов в области телекоммуникаций. Третий сектор ITU — исследовательский сектор ITU-D — был создан на базе организованного в 1989 г. Бюро телекоммуникационных исследований (Telecommunication Development Bureau — TDB).

Надо отметить, что не случайно именно ITU играет главенствующую роль в области международных телекоммуникационных стандартов. История этой организации неразрывно связана с историей телекоммуникационной индустрии. Впервые аббревиатура ITU появилась почти полтора столетия (!) назад. 17 мая 1865 г. в Париже представители 20 государств приняли первую международную телеграфную конвенцию и учредили Международный телеграфный союз — International Telegraph Union (ITU). Напомним, тогда телеграф был единственным видом электросвязи! С 1885 г. ITU занимался и вопросами международной стандартизации телефонии. В 1906 г. под эгидой ITU в Берлине прошла первая международная конференция по вопросам радиотелеграфии, на которой была принята первая конвенция, заложившая основу того, что сегодня называют частотным регулированием.

К середине 1920-х годов были основаны Международные консультативные комитеты в области телефонии (The International Telephone Consultative Committee — CCIF, 1924), телеграфии (The International Telegraph Consultative Committee — CCIT, 1925) и радиовещания (CCIR, 1927). В 1932 г. в Мадриде ITU обрел свое современное название — International Telecommunication Union (официальная дата переименования — 1 января 1934 г.). 15 октября 1947 г. в Атлантик-Сити ITU получил статус агентства только что созданной ООН. В том же году был основан и Международный комитет по регистрации радиочастот IFRB. В 1956 г. CCIT и CCIF слились в единую организацию CCITT. 1 марта 1993 г. завершилась одна из крупнейших реорганизаций, в результате чего ITU обрел свой современный вид. Отметим, что структура ITU как нельзя лучше отражает ситуацию в области современных телекоммуникаций:

тесное переплетение и единение технологий проводных и беспроводных, аналоговых и цифровых.

Огромную роль в области утверждения международных стандартов играет Международная организация по стандартизации ISO (International Organization for Standardization). Это сеть институтов стандартизации 148 различных стран. Само название организации ISO — не сокращение (было бы IOS), а производное от греческого слова isos (равный). Если ITU, как агентство ООН, представляет собой межправительственную организацию, то ISO — неправительственный орган. Среди членов ISO не только правительственные, но и частные организации. Прародителем ISO стала основанная в 1906 г. и успешно действующая до сих пор Международная электротехническая комиссия (МЭК, IEC — International Electrotechnical Commission), занимающаяся вопросами стандартизации в области электротехники и электроники. В 1926 г. была создана ISA (International Federation National Standardizing Associations) — Международная федерация национальных стандартизирующих организаций. С началом Второй мировой войны ISA прекратила существование, но была возрождена под названием ISO на конференции национальных стандартизирующих организаций (25 стран), проходившей с 14 по 26 октября в 1946 г. в Лондоне (официальное начало работы ISO — 23 февраля 1947 г.). Важную роль в рождении ISO сыграл созданный в 1944 г. Координационный комитет стандартов Объединенных Наций (United Nations Standards Coordinating Committee), влившийся в ISO. Сегодня ISO играет важнейшую роль в области стандартизации в телекоммуникациях, работая в плотной кооперации с IEC. Можно сказать, что IEC и ISO поделили сферы влияния: IEC — стандарты в области электроники и электротехники, ISO — все остальное. Они используют единую систему нумерации, и в кодах стандартов зачастую фигурирует название обеих организаций, например ISO/IEC 8802-3.

Отметим, что и ITU, и ISO, и IEC выступают скорее как самые авторитетные утверждающие организации. Сегодня общемировая практика такова, что собственно технической проработкой будущих стандартов занимаются ведущие национальные и межнациональные организации. Прежде всего, это Европейский институт стандартизации в области телекоммуникаций ETSI (European Telecommunications Standards Institute), Европейская конференция почтовых и телекоммуникационных ведомств (CEPT — Conference of European Postal and Telecommunication Administrations), национальные институты стандартизации. В США это ANSI, в Японии — JESA (Japanese Engineering Standards Association), в Великобритании — Министерство почт и телекоммуникаций (MPT — Ministry of Posts and Telecommunications). Весьма значимы различные промышленные ассоциации и объединения, такие, как Ассоциация электронной промышленности США EIA (Electronics Industries Association) и Институт инженеров по электротехнике и электронике IEEE (США). Очевидно, что эти организации сами по себе никаких технологий не разрабатывают, но они организуют и координируют всю работу, необходимую для выбора оптимального варианта из нескольких предложенных, его доработки, документального оформления, утверждения, разрешения конфликтов и т. п. Как правило, для каждого будущего стандарта создается рабочая группа, в которую входят представители всех заинтересованных сторон, которая и занимается проработкой вопроса.

Особняком в перечне стандартизирующих организаций стоит Институт инженеров по электротехнике и электронике — IEEE (Institute of Electrical and

Electronics Engineers). Достаточно сказать, что членами IEEE являются ANSI и ISO. IEEE выпускает свои собственные стандарты, имеющие общемировое значение. Как правило, они затем утверждаются ISO и/или ITU, но это уже формальность.

Безусловно, IEEE — явление американское. Но сама история этой организации показывает, насколько важную роль она играла и играет в развитии телекоммуникаций в мировом масштабе, насколько ее структура отражает реалии этой столь динамичной индустрии. Корни IEEE уходят в позапрошлый век. 13 мая 1884 г. по инициативе 25 наиболее авторитетных специалистов-электротехников США, среди которых были Т. Эдисон, Э. Томсон и Э. Хьюстон, был учрежден Американский институт инженеров-электротехников AIEE (American Institute of Electrical Engineers). В области телекоммуникаций эта организация занималась вопросами проводной электросвязи. Именно AIEE принадлежит заслуга в создании первых стандартов США в данном направлении. В 1907 г. в Бостоне было организовано Общество инженеров по беспроводной телеграфии (Society of Wireless Telegraph Engineers — SWTE). Через два года в Нью-Йорке создали Беспроводной институт (The Wireless Institute — TWI). Однако к 1912 г. обе эти организации оказались в кризисе и решили объединиться. В результате был создан Институт радиоинженеров IRE (Institute of Radio Engineers). Эта организация, ведающая вопросами стандартизации в беспроводной связи, в 1963 г. объединилась с AIEE. Так 1 января 1963 г. родился IEEE. Отметим, что, несмотря на международный статус, IEEE тяготеет к Северо-Американскому континенту. Однако это не мешает ему быть законодателем мировой моды в области телекоммуникаций. В рамках темы настоящей книги особого внимания заслуживает комитет IEEE 802, занимающийся вопросами стандартизации технологий сетей передачи данных. Усилиями этого комитета оформились в виде стандартов такие технологии, как Ethernet, Token Ring, беспроводной Ethernet и др.

В России вопросами стандартизации в качестве головного национального института занимался и продолжает заниматься Госстандарт (в сотрудничестве с отраслевыми институтами — ЦНИИС, ЛОНИИС и др.). К сожалению, работа Госстандарта в области телекоммуникаций в плане участия в международной стандартизационной деятельности не отличается оперативностью. Возможно, в этом есть и положительная сторона, поскольку сегодня отечественные специалисты оперируют международными стандартами — ISO, ITU, IEEE. Это гарантирует, что не появится очередной стандарт, либо не согласующийся с общемировой практикой (как до сих пор действующий в области телефонии ГОСТ 7153-85 с уникальным требованием на сопротивление наборного ключа телефонного аппарата при замыкании шлейфа менее 50 Ом), либо с явно не лучшими техническими решениями (вспомним частотную систему сигнализации «2 из 6» с равномерным шагом между частотами 200 Гц в диапазоне 700–1700 Гц, что требует существенно разнополосных фильтров). Но по мере интеграции с мировым телекоммуникационным сообществом ситуация, надеемся, будет меняться к лучшему.

1.4. Модель взаимодействия открытых систем

Эталонная модель взаимодействия открытых систем (МВОС, OSI — open system interconnection) — это наиболее удачная попытка стандартизировать протоколы

обмена информацией. Она была разработана и утверждена ISO в тесном взаимодействии с CCITT в 1984 г. MBOC не только стала основой для разработки сетевых стандартов, но и явилась хорошей методологической основой для изучения и сравнения сетевых технологий. Несмотря на то что были разработаны и другие модели, большинство разработчиков и поставщиков сетевых продуктов используют терминологию эталонной модели MBOC.

В соответствии с MBOC все протоколы взаимодействия систем подразделяются на семь уровней — физический, канальный (звена данных), сетевой, транспортный, сеансовый, представительский и прикладной. Рассмотрим кратко основные функции перечисленных уровней.

Нижним уровнем иерархии является физический (Physical). Он определяет электрические и механические характеристики подключения к физическим каналам связи, а также процедуры передачи потока битов от одного узла к другому. Иными словами, функции этого уровня — передать поток битов между двумя точками по заданному каналу связи.

Физический уровень предоставляет сервис для канального уровня или уровня звена данных (Data link), отвечающего за передачу данных по каналу связи между двумя точками (узлами сети). К функциям канального уровня в первую очередь относятся упаковка информации в кадры определенной длины, формирование контрольных сумм и проверка содержимого кадров после их передачи, формирование подтверждений о приеме кадров, повторная передача неподтвержденных кадров и т. д.

Сетевой уровень (Network) обеспечивает взаимодействие между узлом и сетью. Он формирует сетевые адреса пакетов, управляет потоками, адресацией, маршрутизацией, организацией и поддержанием транспортных соединений. Единицей информации протоколов сетевого уровня является пакет, поэтому иногда этот уровень называют пакетным.

Транспортный уровень (Transport) предназначен для трансляции потоков данных из одного порта в другой. Под портом понимается конец логического канала сети передачи данных, где фактически завершаются операции транспортировки данных и начинаются вычислительные процессы. На этом уровне происходит прозрачная трансляция данных от передатчика к приемнику через сколь угодно сложную среду передачи — через различные сети посредством разнообразных сетевых и физических технологий. На транспортном уровне устанавливаются и разъединяются транспортные соединения, формируются пакеты, принадлежащие передаваемому в сеансе связи потоку. Транспортный уровень — последний в иерархии MBOC, обеспечивающий транспортный сервис; он освобождает более высокие уровни от организации передачи данных.

Сеансовый уровень (Session) служит для организаци, поддержания и окончания сеансов (логической связи) между прикладными процессами. Сеансы устанавливаются через уровень представления.

Уровень представления (Presentation) необходим для преобразования данных в форму, удобную для прикладной программы. На этом уровне преобразуются форматы данных и команд.

Прикладной уровень (Application) представляет собой процесс обработки информации (прикладные процессы). Он обеспечивает работу прикладной программы так, как если бы обмен данными происходил бы не через сеть передачи данных, а автономно в вычислительной машине.

Отметим, что, несмотря на несомненную полезность МВОС, не существует ни одной коммуникационной системы, структурированной в соответствии со всеми семью уровнями этой модели. И если между физическим и канальным уровнем еще можно провести достаточно четкую границу, то последний уже распадается на два подуровня — контроля доступа к среде передачи (MAC — Medium Access Control) и управления логическим соединением (LLC — Logical Link Control). Однако МВОС внесла определенный порядок в описание процедур взаимодействия телекоммуникационных систем, и хотя бы в этом она сослужила добрую службу.

1.5. Методы доступа к среде передачи в беспроводных сетях

Одна из основных проблем построения беспроводных систем — это решение задачи доступа многих пользователей к ограниченному ресурсу среды передачи. Существует несколько базовых методов множественного доступа (их еще называют методами уплотнения или мультиплексирования), основанных на разделении между станциями таких параметров, как пространство, время, частота и код. Задача множественного доступа — выделить каждому каналу связи пространство, время, частоту и/или код с минимумом взаимных помех и максимальным использованием характеристик передающей среды.

Множественный доступ с пространственным разделением (Space или Spatial Division Multiplexing — SDM) основан на разделении сигналов в пространстве, когда каждое беспроводное устройство может вести передачу данных только в границах одной определенной территории (пространственной области), на которой любому другому устройству запрещено передавать свои сообщения. Самый простой способ пространственного разделения — это ограничение мощности передатчиков.

Еще недавно данный метод считался малоэффективным — до тех пор, пока не получили промышленное развитие системы, обеспечивающие достаточно точную локализацию зон действия отдельных передатчиков. С появлением аппаратуры (и соответствующих стандартов), обеспечивающей адаптивную перестройку мощности передатчиков абонентских и базовых станций, а также систем на основе антенн с перестраиваемой диаграммой направленности, данный метод получил широкое распространение. Характерный пример — системы сотовой телефонной связи, системы с цифровым формированием диаграмм направленности и др.

В схемах *множественного доступа с частотным разделением* (Frequency Division Multiplexing — FDM) каждое устройство работает на строго определенной частоте, благодаря чему несколько устройств могут вести передачу данных на одной территории. Это один из наиболее известных методов, так или иначе используемый в самых современных системах беспроводной связи. Характерный пример схемы FDM — работа нескольких радиостанций на одной территории, но на разных частотах. При этом их рабочие частоты должны быть разделены защитным частотным интервалом, позволяющим исключить взаимные помехи. Эта схема, хотя и позволяет использовать множество устройств на определенной территории, сама по себе приводит к неоправданному расточительству обычно

скудных частотных ресурсов, поскольку требует выделения отдельной частоты для каждого беспроводного устройства.

Более гибким является *множественный доступ с временным разделением* (Time Division Multiplexing — TDM). В данной схеме каналы распределяются по времени, т. е. каждый передатчик транслирует сигнал на одной и той же частоте, но в различные промежутки времени (как правило, циклически повторяющиеся) при строгой синхронизации процесса передачи.

Подобная схема достаточно удобна, так как временные интервалы могут динамично перераспределяться между устройствами сети. Устройствам с большим трафиком назначаются более длительные интервалы, чем устройствам с меньшим объемом трафика.

Однако метод временного уплотнения не может использоваться в чисто аналоговых сетях — даже если исходные данные аналоговые (например, речь), он требует их оцифровки и разбиения на пакеты. Скорость передачи отдельного пакета, как правило, существенно превосходит скорость передачи исходных оцифрованных данных. Характерный пример применения временного уплотнения (в проводных сетях) — это метод передачи телефонного трафика посредством каналов E1. На узловой АТС каждый аналоговый телефонный канал преобразуется в поток данных со скоростью 64 кбит/с (8 разрядов оцифровки × × 8 кГц частоты выборок). Фрагменты по 8 бит из 32 каналов (30 телефонных и 2 служебных) образуют цикл. Длительность каждого цикла — 125 мкс, соответственно скорость передачи данных — (32 × 8 бит)/125 мкс = 2048 кбит/с (т. е. 2 048 000 бит/с). Данный поток транслируется по магистральным каналам и восстанавливается (демультиплексируется) на приемном конце.

Основной недостаток систем с временным уплотнением — это мгновенная потеря информации при срыве синхронизации в канале, например, из-за сильных помех, случайных или преднамеренных. Однако успешный опыт эксплуатации таких знаменитых TDM-систем, как сотовые телефонные сети стандарта GSM, свидетельствует о достаточной надежности механизма временного уплотнения.

Еще один тип множественного доступа — это мультиплексирование *с кодовым разделением* (Code Division Multiplexing — CDM). Первоначально, из-за сложности реализации, данная схема использовалась в военных целях, но со временем прочно заняла свое место в гражданских системах. Именем основанного на CDM механизма разделения каналов (CDMA — CDM Access) даже назван стандарт сотовой телефонной связи IS-95a, а также ряд стандартов третьего поколения сотовых систем связи (cdma2000, W-CDMA и др.). В данной схеме все передатчики передают сигналы на одной и той же частоте, но с разными базовыми кодами.

Принцип кодового уплотнения иллюстрирует ситуация, когда много людей в одной комнате разговаривают на разных языках. При этом каждый человек понимает только один определенный язык. Для каждого речь на непонятном языке будет восприниматься как ничего не значащий шум, лишенный полезной информации. А на фоне этого шума он будет воспринимать поток информации на понятном ему языке.

В схеме CDM каждый передатчик заменяет каждый бит исходного потока данных на CDM-символ — кодовую последовательность длиной в 11, 16, 32, 64 и т. п. бит (их называют чипами). Кодовая последовательность уникальна для

каждого передатчика, причем их подбирают так, чтобы корреляция двух любых CDM-кодов была минимальна (а в ряде случаев — чтобы автокорреляция CDM-кода при фазовом сдвиге была также минимальна). Как правило, если для замены 1 в исходном потоке данных используют некий CDM-код, то для замены 0 применяют тот же код, но инвертированный.

Приемник знает CDM-код передатчика, сигналы которого должен воспринимать. Он постоянно принимает все сигналы, оцифровывает их. Затем в специальном устройстве (корреляторе) производит операцию свертки (умножения с накоплением) входного оцифрованного сигнала с известным ему CDM-кодом и его инверсией. В несколько упрощенном виде это выглядит как операция скалярного произведения вектора входного сигнала и вектора с CDM-кодом. Если сигнал на выходе коррелятора превышает некий установленный пороговый уровень, приемник считает, что принял 1 или 0. Для увеличения вероятности приема передатчик может повторять посылку каждого бита несколько раз. При этом сигналы других передатчиков с другими CDM-кодами приемник воспринимает как аддитивный шум. Более того, благодаря большой избыточности (каждый бит заменяется десятками чипов) мощность принимаемого сигнала может быть сопоставима с интегральной мощностью шума. Похожести CDM-сигналов на случайный (гауссов) шум добиваются, используя CDM-коды, порожденные генератором псевдослучайных последовательностей. Такие кодовые последовательности называют шумоподобными, соответственно модулированные ими сигналы — шумоподобными сигналами (ШПС). Очевидно, что при передаче посредством ШПС спектр исходного сообщения расширяется во много раз. Поэтому данный метод еще называют расширением спектра сигнала посредством прямой последовательности (DSSS — Direct Sequence Spread Spectrum).

Наиболее сильная сторона данного уплотнения заключается в повышенной защищенности и скрытности передачи данных: не зная кода, невозможно получить сигнал, а в ряде случаев — и обнаружить его присутствие. Кроме того, кодовое пространство несравненно более значительно по сравнению с частотной схемой уплотнения, что позволяет без особых проблем присваивать каждому передатчику свой индивидуальный код. Основной же проблемой кодового уплотнения до недавнего времени являлась сложность технической реализации приемников и необходимость обеспечения точной синхронизации передатчика и приемника для гарантированного получения пакета.

Отметим, что уплотнение с кодовым разделением — метод синтетический, т.е. он базируется на частотном либо временном методе уплотнения. В наиболее «чистом» виде метод кодового уплотнения реализуется в случае DSSS. Кроме того, известны и используются методы расширения спектра посредством частотных и временных скачков (соответственно FHSS — Frequency Hoping Spread Spectrum и THSS — Time Hoping Spread Spectrum). В случае расширения спектра посредством частотных скачков (еще его называют методом псевдослучайной перестройки рабочей частоты — ППРИ) в заданном частотном диапазоне F одновременно работает несколько передатчиков, каждый в узкой полосе, во много раз меньшей F. Центральная частота каждого передатчика в ходе работы дискретно изменяется по закону, задаваемому уникальной для него кодовой последовательностью. Приемник знает эту кодовую последовательность и перестраивается по частоте приема синхронно с передатчиком. Кодовые последовательности выбирают так, чтобы минимизировать вероятность одно-

временной работы двух передатчиков. Тем самым обеспечивается определенная защита от прослушивания и помех. Данный метод в ряде случаев оказывается достаточно эффективным и применяется, в частности, в такой популярной сегодня технологии БСПИ, как Bluetooth.

Если метод частотных скачков представляет собой метод частотного уплотнения с изменением частотной полосы, то метод временных скачков аналогичен временному уплотнению, только моменты начала трансляции пакетов передатчика не строго периодичны, а изменяются по псевдослучайному закону. Как правило, кодовая последовательность определяет время отклонения начала трансляции очередного пакета от заданного периода. Подобный механизм, в частности, реализован в системах связи со сверхширокой спектральной полосой компании Time Domain.

Еще одна важная производная методов кодового и частотного уплотнения — механизм мультиплексирования посредством ортогональных несущих (OFDM — Orthogonal Frequency Division Multiplexing). Его суть: весь доступный частотный диапазон разбивается на достаточно много поднесущих (от нескольких сот до тысяч). Одному каналу связи (приемнику и передатчику) назначают для передачи несколько таких несущих, выбранных из всего множества по определенному закону. Передача ведется одновременно по всем поднесущим, т. е. в каждом передатчике исходящий поток данных разбивается на N субпотоков, где N — число поднесущих, назначенных данному передатчику. Распределение поднесущих в ходе работы может динамически изменяться, что делает данный механизм не менее гибким, чем метод временного уплотнения.

До недавнего времени распространение технологии OFDM сдерживала сложность ее аппаратной реализации. Однако с развитием полупроводниковой технологии это уже не является преградой. В результате метод OFDM приобретает все большее распространение, в частности, используется в системах связи таких популярных стандартов, как IEEE 802.11 и DVB, является одним из основных механизмов стандарта широкополосных региональных БСПИ IEEE 802.16-2004. Более того, все наиболее перспективные стандарты БСПИ (IEEE 802.16e, LTE, cdma200 Rev.C) основаны именно на технологии OFDM. И в сетях 4G будет использоваться этот метод множественного доступа.

Как правило, описанные схемы в беспроводных сетях используются в сочетании друг с другом. Например, для мобильных сетей GSM одновременно используются схемы уплотнения SDM, TDM и FDM, в системах стандарта IEEE 802.16 эффективно сочетаются технологии OFDM, CDM, FDM/TDM и SDM.

Рассмотренные выше механизмы — это способы разделения единого ресурса на каналы передачи. Однако эти каналы надо еще назначить конкретным устройствам. Рассмотрим несколько наиболее популярных схем распределения канальных ресурсов на базе технологии TDM (аналогичные механизмы возможны и при других методах уплотнения).

Простейший алгоритм для схемы уплотнения TDM — это *фиксированное распределение временных интервалов* между различными устройствами. Распределением занимается базовая станция (центральное устройство), которая сообщает каждому абонентскому устройству время начала передачи. Подобная схема идеально подходит для беспроводных сетей, которые имеют фиксированную пропускную способность. Однако она не оптимальна в случае нерегулярной

передачи, поскольку во время молчания устройства его интервал не может быть использован другим терминалом. Поэтому число абонентских станций (либо допустимая скорость передачи) принципиально и существенно ограниченно.

Противоположностью данной схемы является механизм полностью случайного доступа или *классическая схема Aloha*. В ней отсутствует какой-либо алгоритм, который позволял бы избежать коллизий (одновременной работы двух передатчиков в одно время на одной частоте). Это означает, что любое устройство может передавать данные в любое время и нет никакой гарантии, что эти данные будут успешно доставлены получателю. Данная схема — один из самых первых механизмов доступа для систем беспроводной связи. Она была разработана в 1970-х годах в Гавайском университете и применялась в сети ALOHANET для беспроводного соединения нескольких станций (университетских зданий, располагавшихся на разных островах Гавайского архипелага). Данная схема хорошо работает в сетях со слабой загрузкой, т.е. в сетях, имеющих малое число устройств или передающих небольшое количество информации в единицу времени. При пуассоновском распределении интенсивности генерации пакетов устройствами максимальная пропускная способность системы достигается уже при 18%-ной загрузке.

Усовершенствованием основной схемы Aloha явился *метод множественного доступа с детектированием несущей* (Carrier Sense Multiple Access — CSMA). Детектирование несущей частоты означает лишь то, что канал прослушивается устройством. Если он занят, т.е. другое устройство передает данные, то передатчик переходит в ждущий режим до того момента, когда канал станет свободным. Этот метод позволяет значительно улучшить пропускную способность системы. Как и в методе случайного доступа, в данной схеме не требуется наличия центрального устройства, т.е. каждое устройство принимает решение о передаче самостоятельно. Поскольку фактически доступ к среде получает та станция, которая первой начала передачу, данный механизм еще называют методом конкурентного доступа.

Существует несколько версий схемы CSMA. При использовании неустойчивой схемы CSMA станции слушают канал и, если канал свободен, немедленно начинают передачу. Если канал занят, станция перед повторным определением состояния канала выжидает случайный промежуток времени, после чего опять слушает канал. Если канал свободен, то терминал передает данные. В p-настойчивых схемах CSMA узлы тоже определяют состояние канала, но данные передаются с вероятностью p. Устройство может отложить передачу до следующего временного интервала с вероятностью $1-p$, т.е. осуществляется дополнительное разделение доступа к среде. В l-настойчивых системах CSMA все станции, которым необходимо передавать данные, одновременно получают доступ к среде, как только она освобождается.

Другой вариацией данного метода является CSMA/CA (CA — Collision Avoidance, с предотвращением конфликтов), использующаяся в беспроводных ЛВС стандарта IEEE 802.11. Здесь после определения незанятости канала время ожидания выбирается случайно в некотором временном промежутке. В спецификации HIPERLAN 1 описана схожая схема — бесприоритетный множественный доступ с исключением (Elimination Yield — Non-Preemptive Multiple Access, EY — NPMA).

Схема с цифровым детектированием (DSMA — Digital Sense Multiple Access) использует схожий с CSMA/CA принцип работы. Этот метод также называют множественным доступом с детектированием подавления (Inhibit Sense Multiple Access — ISMA). Различие заключается в том, что занятость канала определяется не путем прослушивания, а посредством посылки базовой станцией пакета, в котором определяется статус канала. В данной схеме базовая станция должна быть синхронизирована с передатчиками так, чтобы передатчики не передавали данные во время передачи статуса канала. Если канал занят, то станции ждут случайного промежутка времени для последующей передачи. Поскольку несколько станций могут одновременно передать данные, центральная станция посылает пакет с подтверждением о получении пакета данных.

В современных БСПИ, как правило, используют сочетание механизмов централизованного назначения временных интервалов и методов конкурентного доступа. По сути, работа этих систем происходит в два этапа. Первый этап — резервирование ресурсов (временных интервалов) для будущей передачи. На этом этапе все станции заявляют (пытаются заявить) о своих потребностях в канальных ресурсах. На втором этапе происходит непосредственная передача данных в отведенном временном интервале. В этих схемах используется центральный терминал, с помощью которого производится синхронизация передач и осуществляется резервирование. Как правило, механизмы резервирования приводят к увеличению времени задержки получения пакетов при слабой загрузке системы, но при этом обеспечивают ей более высокую пропускную способность.

Примером подобного механизма является *схема множественного доступа с распределением по запросу* (Demand Assigned Multiple Access — DAMA), называемая также схемой Aloha с резервированием. Она, в частности, применяется в спутниковых системах связи. В течение определенного временного интервала, разбитого на мини-интервалы, все станции пытаются зарезервировать для себя будущие временные интервалы для передачи данных. Поскольку на стадии резервирования происходят конфликты, некоторым станциям не удается зарезервировать канальный ресурс. Если станции удалось зарезервировать временной интервал, то ни одна другая станция не сможет в это время осуществлять передачу. Таким образом, базовая станция собирает все успешные запросы (остальные игнорируются) и посылает назад список с указанием прав доступа к последующим временным интервалам. Этому списку подчиняются все станции. Схема DAMA относится к схемам с явным резервированием, когда каждый интервал для передачи резервируется явно.

Схема TDMA с резервированием отличается от предыдущей схемы тем, что этап резервирования происходит не на основании конкурентного доступа, а по обычной фиксированной схеме TDMA. Каждому устройству назначается временной мини-интервал, в течение которого оно сообщает, будет ли передавать данные. Поэтому в начале каждого цикла передачи базовая станция передает пакет, разбитый на N интервалов, в каждом их которых указано, зарезервирован канал или нет. Затем следуют $N \cdot k$ интервалов для данных. Данный метод гарантирует каждой зарезервировавшей канал станции определенную пропускную способность. Остальные станции могут пересылать данные в течение интервалов, которые никто не зарезервировал, но уже на принципах конкурентного доступа и без гарантии доставки пакетов.

Схема с резервированием пакетов (PRMA — Packet Reservation Multiple Access) является примером со скрытым резервированием, поскольку интервалы резервируются неявно. Центральное устройство в начале каждого цикла рассылает список с распределением временных интервалов. Само же резервирование происходит по другой схеме. Представим, что какому-либо устройству необходимо передать данные, но при этом оно не зарезервировало временной интервал. Это устройство регулярно получает список с зарезервированными интервалами. К примеру, в полученном списке указано, что третий, пятый и восьмой интервалы не зарезервированы, т. е. свободны. Устройство случайным образом принимает решение о том, в каком интервале можно попытаться передавать данные. Например, устройство передает сообщение в пятый интервал. Если передача прошла успешно, устройство получает об этом подтверждение. Базовая станция резервирует этот канал для нового устройства и включает его в свой список. Если запрос не дошел до базовой станции, устройство должно попробовать вновь послать данные в один из свободных интервалов.

Литература

1. Вишневский В. М. Беспроводные сети широкополосного доступа к ресурсам Интернета. — Электросвязь, 2000, № 10, с. 9–13.
2. Saunders S., Heywood P., Dornan A., Bruno L., Allen L. Wireless IP: Ready or Not, Here it Comes. — Data Communications, 1999, № 9, p. 42–68.
3. Шахнович И. Современные технологии беспроводной связи. — М.: Техносфера, 2004.
4. Wideband wireless digital communications/ Ed. Andreas F. Molisch. — Prentice Hall PTR, 2001.
5. Столлингс В. Беспроводные линии связи и сети/ Пер. с англ. — М.: Изд. дом «Вильямс», 2003.

ГЛАВА 2

КОДЫ И ИХ ПРИМЕНЕНИЕ В СИСТЕМАХ ПЕРЕДАЧИ ИНФОРМАЦИИ

В данной главе рассматриваются все виды кодирования информации в современных системах передачи дискретной информации. При этом любое кодирование есть преобразование одной цифровой информации в другую как на передающей стороне, так и на приемной. Однако цели у такого преобразования могут быть совершенно разные. Это либо устранение содержащейся в информации избыточности, либо, наоборот, сознательное внесение избыточности для повышения помехоустойчивости, либо шифрация информации.

2.1. Математические основы передачи информации

В последующих двух главах будут описаны основы области знаний, которая имеет несколько названий, иногда обозначающих одно и то же, иногда смежные, иногда пересекающиеся области. Итак, это теория информации, теория передачи информации, теория связи, теория сигналов и т. д. Название не столь важно. Важно то, что фундаментальные математические основы передачи информации интенсивно развивались в течение последних 50 лет прошлого столетия, начиная с работ Шеннона и Котельникова. С задержкой в несколько десятков лет результаты этих исследований стали реализовываться на практике, и сегодня мы присутствуем на том временном этапе, когда любые даже самые фантастические элементы теории передачи информации реализуются в реальных системах. Как правило, реализация этих идей происходит на физическом уровне (PHY) модели взаимодействия открытых систем (OSI) и, иногда, на подуровне управления доступом к каналу (MAC) канального уровня. Именно поэтому сначала рассматриваются сами фундаментальные основы теории передачи информации, а затем международные стандарты, на основе которых строятся реальные системы, и сами реальные системы. При этом основной упор делается на группу стандартов IEEE 802.16 (WiMAX).

Блок-схема любой системы передачи дискретной информации «точка-точка» показана на рис. 2.1. Различным видам кодеров на передающей стороне соответствуют «обратные» декодеры. Кодирование используется для сжатия информация, криптографической защиты, а также повышения помехоустойчивости за счет внесения избыточности на передающей стороне. Выход корректирующего кодера и вход соответствующего декодера образуют вход и выход дискретного канала связи.

В общем случае дискретный канал связи является математическим объектом с K входами и J выходами с определенными переходными вероятностями.

Если детализировать физическое содержание дискретного канала, то окажется, что он состоит из последовательной цепочки «модулятор – непрерывный канал – демодулятор», причем цепочка «модулятор – непрерывный канал» образует полунепрерывный канал. Непрерывный канал является каналом с непрерывным входом и непрерывным выходом, а полунепрерывный канал — с дискретным входом и непрерывным выходом. Позже будет видно, что непрерывный канал используется для синтеза сигнально-кодовых конструкций, а полунепрерывный канал — для декодирования с мягким решением. Важно также понимать, что все рассмотренные каналы являются математическими моделями соответствующих физических каналов связи.

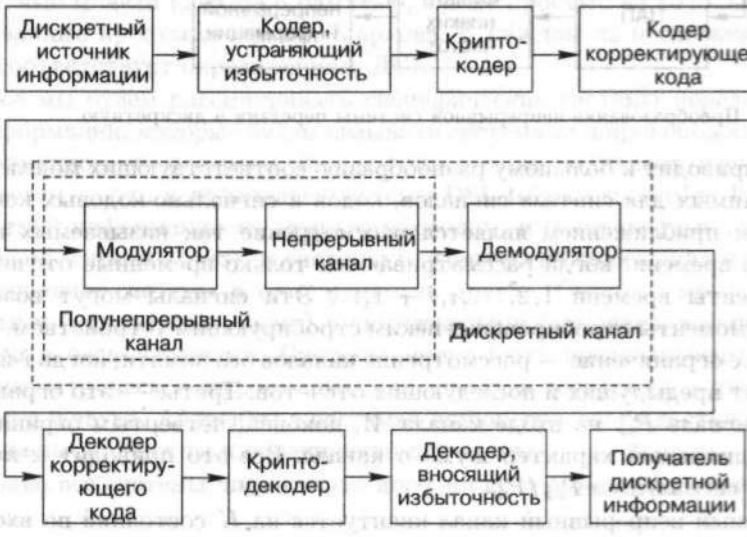


Рис. 2.1. Блок-схема передачи дискретной информации «точка-точка»

Реальная система передачи дискретной информации содержит большое число узлов и функциональных элементов, не отображенных на рис. 2.1, тем не менее, без которых система не функциональна. В первую очередь это устройства синхронизации по несущей, тактовой и блоковой частоте, без которых демодуляция и декодирование сигналов невозможны. Кроме того, это задающие генераторы, системы управления, передатчики, приемники, антенны в случае радиосистем и многие другие элементы. Однако данное рассмотрение мы ограничим узлами, изображенными на рис. 2.1.

Если рассматриваемая система является непрерывной, то при помощи теоремы Котельникова система из непрерывной превращается в дискретную. Цепочка этих преобразований показана на рис. 2.2. Таким образом, дискретные системы передачи покрывают все системы, которые будут рассмотрены ниже.

При рассмотрении необходимо различать реальные физические каналы связи, возникающие при широкополосном доступе, и соответствующие им модели. В зависимости от среды передачи широкополосный доступ может быть проводным и беспроводным. Проводные каналы характеризуются более постоянными параметрами, чем беспроводные. Всем каналам — и проводным, и беспроводным — присущи помехи и мешающие воздействия от других абонентов системы. В слу-

чае беспроводных каналов параметры последних могут существенно меняться во времени. Это может быть вызвано, во-первых, замираниями передаваемого сигнала в результате отражений во время распространения и, во-вторых, эффектом Доплера, возникающим при движении абонентов.

Рис. 2.2. Преобразование непрерывной системы передачи в дискретную

Все это приводит к большому разнообразию соответствующих моделей каналов, необходимых для синтеза сигналов, кодов и сигнально-кодовых конструкций. Первым приближением является рассмотрение так называемых каналов дискретного времени, когда рассматриваются только временные отсчеты сигнала в моменты времени $1, 2, \ldots, i, i+1, \ldots$ Эти сигналы могут возникать, например, в моменты стробирования неким стробирующим устройством. Второе существенное ограничение — рассмотрение каналов без памяти, когда i-й отсчет не зависит от предыдущих и последующих отсчетов. Третье — это ограничение мощности сигнала $P_{вх}$ на входе канала. И, наконец, четвертым ограничением является аддитивный характер шума в канале. Все это приводит к каналу с отношением сигнал/шум $P_{вх}/P_{ш}$.

Если данный непрерывный канал квантуется на K состояний по входу и J состояний по выходу, получается дискретный канал без памяти с K входами и J выходами с соответствующими переходными вероятностями. На входе канала возникает одна из K букв из алфавита $A = \{a_1, a_2, \ldots, a_K\}$, а на выходе одна из J букв из алфавита $B = \{b_1, \ldots, b_J\}$. Переходные вероятности получения на выходе буквы b_j при условии, что на входе была буква a_k, обозначаются как $p(b_j/a_k)$, $k = 1, \ldots, K$, $i = 1, \ldots, J$. Важно понимать, что в общем случае размеры алфавитов по входу и выходу канала не совпадают. Дискретный канал без памяти называется симметричным, если набор переходных вероятностей для одного входа или выхода задает простыми перестановками соответствующие наборы для других входов или выходов.

Самым простым каналом называется двоичный симметричный канал, для которого алфавиты по входу и выходу совпадают и состоят из двух букв $\{0, 1\}$. Переходные вероятности равны $p(0/0) = p(1/1) = 1 - p$, а $p(0/1) = p(1/0) = p$. Тогда p называется вероятностью ошибки, а $1 - p$ — вероятностью правильного приема.

Также очень прост так называемый стирающий канал, для которого имеется два входа $\{0, 1\}$ и три выхода $\{0, 1, *\}$. Символ $*$ называется стиранием или отказом. Переходные вероятности равны $p(0/0) = p(1/1) = 1 - \tau$, $p(0/1) = p(1/0) = 0$, а $p(0/*) = p(1/*) = \tau$. Тогда τ — это вероятность стирания, а $1 - \tau$ — это вероятность правильного приема.

В случае полунепрерывного канала, необходимого при мягком декодировании, выход канала обычно квантуется на несколько значений. В этом случае размер алфавита по выходу канала J существенно превышает алфавит по входу канала K.

Очень большое внимание во второй главе будет уделяться гауссову каналу с межсимвольной интерференцией. Такой канал дискретного времени состоит из последовательного включения линейного фильтра и источника аддитивного гауссова шума.

Наиболее традиционными моделями непрерывных каналов с памятью являются каналы с релеевским или райсовскими замираниями. Часто используют дискретные модели каналов с памятью, самые простые из которых — каналы, образованные простыми цепями Маркова. В каждом из нескольких состояний канал соответствует определенному ДСК.

Ниже мы будем рассматривать специфические системы передачи дискретной информации, которые будем называть системами широкополосного доступа к информационным ресурсам. Под это определение подпадает огромное количество систем — это и проводные системы DSL (digital subscriber line) передачи дискретной информации по медным проводам, и беспроводные системы BWA (broadband wireless access), и системы цифрового радиовещания и телевидения, и спутниковые системы и сети и многие другие. Системы могут быть и дуплексными (двунаправленными) и симплексными (однонаправленными), и проводными и беспроводными. Системы могут быть предназначены для любого вида телеметрической информации, для передачи телефонии, а также трансляции видео и звука. Абоненты в системах широкополосного доступа могут быть неподвижны, подвижны и ограниченно подвижны.

Однако все системы широкополосного доступа объединяет несколько общих свойств.

Все системы широкополосного доступа используют среду передачи (канал связи) на пределе физических возможностей, т. е. пропускной способности, что влечет за собой применение новейших методов синтеза и приема сигналов, кодов и сигнально-кодовых конструкций.

Все системы широкополосного доступа используют эффективные методы доступа к среде передачи и мультиплексирования сообщений в этой среде, что влечет за собой применение новейших методов множественного доступа.

Все системы широкополосного доступа для максимизации своей эффективности используют сжатую информацию, что влечет за собой использование новейших методов кодирования источника.

При передаче непрерывных сообщений используется их оцифровка и дискретная передача. Если на начальном этапе развития таких методов оцифровка сообщения приводила к существенному увеличению полосы сигнала, то теперь оцифрованное сообщение часто занимает полосу в три-четыре раза меньше исходного аналогового.

И, наконец, все системы широкополосного доступа, заменяя собой узкополосные и среднеполосные системы, существенно расширяют пользовательские свойства и приводят к мультисервисности услуг для абонентов.

Теперь можно рассмотреть фундаментальные основы теории передачи информации, позволяющие эффективно строить системы широкополосного доступа.

2.2. Коды, устраняющие избыточность

2.2.1. Введение в теорию кодирования

Рассмотрим простейший пример. Пусть имеется дискретный источник без памяти (ДИБП), на выходе которого может возникать одна из четырех букв с соответствующими вероятностями: $a1$, $P(a1) = 0{,}5$; $a2$, $P(a2) = 0{,}25$; $a3$ и $a4$, $P(a3) = P(a4) = 0{,}125$. Если пронумеровать эти буквы последовательностями из двух бит равной длины, то будет тратиться в среднем два бита на букву, $N_{\text{ср}} = 2$.

Пронумеруем эти буквы другим способом: $a1$ — 0; $a2$ — 10; $a3$ — 110; $a4$ — 111. Тогда средняя длина буквы $N_{\text{ср}} = 1{,}75$, причем все буквы будут однозначно декодироваться, так как никакая более короткая буква не является префиксом (началом) более длинной буквы. Отметим, что энтропия данного источника

$$H = \sum_{i=1}^{i=K} -P(ai)\log 2P(ai)$$

также равна 1,75. Отношение средней длины и длины буквы при равномерном и неравномерном кодировании может трактоваться как выигрыш или коэффициент сжатия.

Даже на таком простейшем примере видно, что сообщения содержат избыточность, устраняя которую можно «сжимать» информацию.

Предположим, что мы имеем дискретный канал, вероятность возникновения ошибки в котором близка к нулю (в идеале = 0). Такой канал называют идеальным каналом или каналом без шума. При наличии идеального канала естественно поставить вопрос о возможности передачи по нему без потерь информации от произвольного дискретного источника A, характеризуемого энтропией $H(A)$, со скоростью, равной пропускной способности канала. Блок-схема такой системы передачи информации показана на рис. 2.3.

Рис. 2.3. Блок-схема системы передачи информации через идеальный канал связи

Кодер осуществляет кодирование сообщений, т. е. каждому дискретному сообщению по определенному правилу ставит в соответствие последовательность символов из алфавита A. При этом по отношению к входу канала выдаваемые кодером символы сами являются дискретными элементами сообщений, статистические свойства которых должны отличаться от статистических свойств сообщений исходного источника. Возможность построения кодера, полностью устраняющего избыточность произвольного исходного источника сообщений, и определяет возможность решения поставленной задачи безошибочной передачи информации со скоростью, равной пропускной способности канала связи. При полном решении этой задачи энтропия как раз и оказывается равной средней длине буквы на выходе кодера.

Степень приближения к точному выполнению равенства зависит от степени уменьшения избыточности источника сообщений. Кодирование, позволяющее устранять избыточность источников сообщений, называется эффективным или статистическим. Коды, получаемые в результате такого кодирования, называются эффективными, статистическими или устраняющими избыточность.

Рассмотрим основные идеи, которые могут быть положены в основу эффективного кодирования, обусловленного двумя причинами:

1) памятью источника;
2) разной вероятностью исходных сообщений.

Универсальным способом уменьшения избыточности, обусловленной памятью источника, является укрупнение элементарных сообщений или букв. При этом кодирование осуществляется длинными блоками. Вероятностные связи между блоками меньше, чем между отдельными элементами сообщений, и чем длиннее блоки, тем больше эти вероятностные связи разрушаются. Смысл укрупнения можно пояснить на примере буквенного текста: если вероятностные связи между соседними буквами в любом языке достаточно сильны, то между словами они значительно слабее, еще меньше — между фразами, еще меньше — между абзацами. Поэтому, применяя кодирование слов, фраз, абзацев, можно практически устранить избыточность, обусловленную вероятностными связями. Однако при этом возрастает задержка передачи сообщений и растет сложность кодирования и декодирования.

Уменьшение избыточности, обусловленной разной вероятностью исходных сообщений, может быть достигнуто применением неравномерных кодов. Основная идея построения таких кодов состоит в том, что наиболее вероятным сообщениям ставятся в соответствие наиболее короткие блоки кодовых символов (кодовые комбинации), а наименее вероятным — более длинные. В силу неравномерности кодов слов и случайного характера сообщения передача без потерь информации с постоянной скоростью возможна лишь при наличии буфера достаточной длины.

2.2.2. Теорема Шеннона для дискретного источника

Предельные возможности статистического кодирования раскрываются в теореме Шеннона для дискретного источника, что является одной из основ теории передачи информации. Эта теорема может быть сформулирована для кодовых слов одинаковой длины следующим образом [1].

Пусть дискретный источник без памяти имеет конечную энтропию $H(U)$. Рассмотрим кодирование последовательностей из L букв источника в последовательности из N кодовых букв, принадлежащих кодовому алфавиту объемом D. С каждой кодовой последовательностью может сопоставляться только одна последовательность источника. Пусть P_e — вероятность появления последовательности источника, которой не соответствует никакая кодовая последовательность. Тогда, если при каком-либо $b > 0$

$$N/L \geq [H(U) + b]/\log D,$$

то величину P_e можно сделать произвольно малой, выбирая значение L достаточно большим.

Обратно, если

$$N/L \le [H(U) - b]/\log D,$$

то P_e становится сколь угодно близкой к L, когда L становится достаточно большим.

Эта теорема имеет достаточно простое и эвристическое обоснование. Можно закодировать буквы произвольного дискретного источника без памяти таким образом, что энтропия кодовых букв будет максимальной.

В случае кодовых слов разной длины теорема имеет следующий вид [1].

Пусть источник сообщений имеет конечный ансамбль U и энтропию $H(U)$. Имеется кодовый алфавит из D символов. Можно так приписать кодовые слова буквам источника, что будет выполняться свойство префикса и средняя длина кодового слова $N_{\text{ср}}$ будет удовлетворять условию

$$N_{\text{ср}} < H(U)/\log D + 1.$$

Более того, для любого однозначно декодируемого множества кодовых слов

$$N_{\text{ср}} \ge H(U)/\log D.$$

Полученный результат позволяет дать следующее толкование энтропии: энтропия источника есть наименьшее количество двоичных символов на сообщение на выходе наилучшего кодера для этого источника при условии, что сообщения могут быть восстановлены по выходу кодера сколь угодно точно.

Естественно, возможно обобщение и на непрерывный случай.

Возможны различные варианты доказательства этой теоремы, иллюстрирующие два возможных подхода к построению эффективных кодов, основанных на использовании равномерного и неравномерного кодирования. При неравномерном кодировании обеспечивается однозначное декодирование всех сообщений.

Предложенный Шенноном метод эффективного кодирования практически совпадает с методом, предложенным другим американским ученым Фано, по которому сообщение длины L, записанное в порядке невозрастания вероятностей, разделяется на две части так, чтобы суммарные вероятности сообщений в каждой части были по возможности равны. Сообщениям первой части приписывается в качестве первого символа 0, сообщениям второй части — 1. Затем каждая из этих частей (если она содержит более одного сообщения) опять делится на две примерно равные части, и в качестве второго символа для первой из них берется 0, а для второй — 1. Этот процесс повторяется до тех пор, пока в каждой из полученных частей не останется по одному сообщению. Именно такой пример и был рассмотрен в начале раздела.

Существуют и другие методы эффективного кодирования. Кодирование по методу Шеннона–Фано, так же как и другими методами, может применяться не только к последовательностям из L-элементных сообщений, но и непосредственно к источникам неравновероятных элементных сообщений. При этом уменьшается выигрыш в эффективности. В том случае, когда левая часть вышеприведенной системы неравенств обращается в равенство, имеем $N_{\text{ср}} = H(U)$. Код, обладающий этим равенством, называется оптимальным. Для того чтобы сообщение источника можно было закодировать двоичным оптимальным кодом, необходимо и достаточно, чтобы все вероятности источника сообщения представляли собой числа, равные целой отрицательной степени числа 2.

Рассуждая аналогичным образом, можно показать, что и в случае кодирования сообщений источника неравномерным кодом с произвольным основанием M оптимальный код может быть получен при условии равенства вероятности всех сообщений целым отрицательным степеням числа M. Если распределение вероятностей кодированного источника не обладает указанным свойством, эффективный код не будет оптимальным и соответствующая ему $N_{\text{ср}}$ будет больше энтропии $H(U)$. Величина $H(U)/N_{\text{ср}}$, характеризующая степень близости неравномерного статистического кода к оптимальному коду, называется эффективностью кода. Таким образом, нижний предел в условии теоремы может быть достигнут лишь при конкретном распределении вероятности источника сообщений. Однако приближение к нему может быть сколь угодно близким при увеличении длины L последовательности кодируемых сообщений. При этом рост эффективности системы передачи информации сопровождается увеличением задержки сообщений.

2.2.3. Применение кодов, устраняющих избыточность

Наибольшее распространение получил способ построения эффективного кода, предложенный Хаффманом.

Рассмотрим его на примере. Пусть задан алфавит из пяти символов A_1–A_5 и их вероятности. В табл. 2.1 наряду с этими исходными данными приведены также результаты кодирования по Хаффману: кодовые слова и их длины N_i, а также средняя длина $N_{\text{ср}}$.

Таблица 2.1. Исходные данные и результаты кодирования по Хаффману

A_i	P_i	Слова	N_i
A_1	0,25	10	2
A_2	0,17	00	2
A_3	0,08	010	3
A_4	0,35	11	2
A_5	0,15	011	3
	$\sum P_i = 1,0$		$N_{\text{ср}} = 2,23$

Сама процедура построения кода Хаффмана проиллюстрирована в табл. 2.2.

Таблица 2.2. Процедура построения кода Хаффмана

A_i	P_i	Шаги объединения				Слова
		1	2	3	4	
A_1	0,35	0,35	0,40	0,60	1,00	11
A_2	0,25	0,25	0,35	0,40		10
A_3	0,17	0,23	0,25			00
A_4	0,15	0,17				011
A_5	0,08					010

На первом этапе символы упорядочивают по убыванию вероятностей, а затем выполняют несколько шагов «объединения», на каждом из которых суммируются вероятности наиболее редко встречающихся символов, и столбец вероятностей пересортировывается.

На втором этапе строится дерево кода, ветви которого отображают в обратном порядке процесс «объединения вероятностей». При построении дерева принимается правило соответствия большей вероятности одному из направлений

ветви (например, «левому») и определенному значению бита кода, например 1. Цепочки битов от «корня» до конца каждой ветви соответствуют кодам исходных символов.

Процедура кодирования сводится к выбору из кодовой таблицы цепочек, соответствующих каждому символу источника. Декодирование предусматривает выделение в битовом потоке кодов символов и их расшифровку в соответствии с табл. 2.2.

Код Хаффмана может быть двухпроходным и однопроходным. Первый строится по результатам подсчета частот (вероятностей) появления различных символов в данном сообщении. Второй использует готовую таблицу кодирования, построенную на основе вероятностей символов в сообщениях похожего типа. Например, кодирование текста на русском языке в первом случае включает его предварительный анализ, подсчет вероятностей символов, построение дерева кода и таблицы кодирования индивидуально для данного сообщения. Во втором случае будет работать готовая таблица, построенная по результатам анализа множества русскоязычных текстов. Двухпроходной код более полно использует возможности сжатия. Однако при этом вместе с сообщением нужно передавать и кодовую таблицу. Однопроходной код не оптимален, однако прост в использовании, поэтому на практике обычно применяют именно его.

Рассмотренные выше основные предельные соотношения привели к широкому применению кодов, устраняющих избыточность во всех телекоммуникационных и вычислительных системах и сетях. Изложенный выше метод кодирования Хаффмана, хотя и является оптимальным, обладает двумя существенными недостатками: большой задержкой буферирования, необходимостью знать или оценить исходные вероятности сообщений и согласованностью только с дискретным источником без памяти.

Здесь для компенсации первых двух недостатков нельзя не упомянуть простой и достаточно эффективный метод кодирования источника с неизвестным распределением вероятностей, известный как сжатие при помощи «стопки книг», или как сжатие сортировкой, или хешированием. Метод был разработан Рябко в 1980 году. Идея метода состоит в следующем: пусть алфавит источника состоит из K символов с номерами $1, 2, \ldots, K$. Кодирующий алгоритм сохраняет последовательность символов, представляющую собой некоторую перестановку символов в последовательности первичного входного алфавита. При поступлении на вход некоторого символа, имеющего в этой переставленной последовательности номер i, кодирующий алгоритм записывает код этого символа (например, монотонный префиксный код). Затем поступивший символ переставляется в начало последовательности и номера всех символов, стоящих перед ним, увеличиваются на 1. Таким образом, наиболее часто встречающиеся символы будут переходить в начало списка и иметь более короткие коды, что, в свою очередь, снизит объем выходного потока при их записи в качестве символов выходного потока.

Кодирование Лемпеля–Зива использует синтаксический метод для динамического источника и борется с третьей проблемой. Очевидно, что посимвольное кодирование, рассмотренное выше, не использует резервы сжатия информации, связанные с повторяемостью цепочек символов, т. е. памятью канала. Наиболее удачным алгоритмом сжатия, основанным на таком подходе, является алгоритм

Лемпеля–Зива, который в разных модификациях используется, в частности, в большинстве программ-архиваторов. Основная идея алгоритма состоит в том, что цепочки символов, уже встреченные ранее, кодируются ссылкой на их «координаты» (номер первого символа и длину) в «словаре», где находится уже обработанная часть сообщения. Более детально основные идеи алгоритма иллюстрирует рис. 2.4.

Рис. 2.4. Иллюстрация идеи алгоритма Лемпеля–Зива

«Сжимаемое» сообщение постепенно входит в буфер источника. Ядро кодера выделяет в буфере блок (цепочку) символов первоначально максимальной длины (обычно порядка 16 символов) и пытается найти совпадающую цепочку в словаре источника. Если это не удается, кодер повторяет поиск для более короткого «урезанного» варианта цепочки. Когда эта цепочка обнаруживается в словаре, в канал передаются ее координаты. Если же поиск не дал результата даже для самого короткого варианта цепочки из двух символов, каждый из них передается по каналу самостоятельно.

На приемной стороне ядро декодера принимает коды и восстанавливает исходное сообщение по собственному словарю. При этом восстановленные цепочки тут же попадают в словарь приемника так, что его содержимое синхронизируется с содержимым словаря источника.

При этом следует помнить, что:

- коды координат цепочки и коды отдельных символов различаются битовыми признаками (например, в первом случае — 1, во втором — 0);
- поскольку цепочки находятся чаще в начале словаря и чаще бывают короткими, дополнительный выигрыш получают за счет статистического кодирования (по Хаффману) их «адресов» и «длин»;
- «канал» — понятие, применимое и к реальному каналу передачи данных, и к файлу, куда данные записываются для хранения. В последнем случае декодер «отрабатывает» при разворачивании сжатого файла;
- при ограниченной длине словаря (обычно от 4 до 16 кбайт) новые поступающие символы и цепочки «вытесняют» прежние (текст как бы «вдвигается» в словарь). Разумеется, вначале, когда словарь не заполнен, эффективность сжатия невысока. Рост объема словаря позволяет повысить степень сжатия, но значительно увеличивается трудоемкость поиска цепочек.

Алгоритм Лемпеля – Зива используется в большинстве популярных программ-архиваторов (в том числе, например, в zip, rar, arj и их windows-версиях). Различие скорости и эффективности операций кодирование – декодирование определяется в основном особенностями программной реализации. Алгоритм Лемпеля – Зива требует большого количества вычислительной работы. Его модификация — алгоритм Лемпеля – Зива – Велча является менее трудоемким, хотя и дает несколько худшие результаты по сжатию.

На похожих принципах и основаны все существующие методы сжатия одномерных цифровых файлов (архиваторы) или одномерных цифровых потоков (например, оцифрованная речь). При оцифровке речи, правда, возникает дополнительная специфика.

Оцифровка звука сводится к дискретизации (взятию отсчетов U_i амплитуды сигнала с шагом по времени Δt или частотой $f_g = 1/\Delta t$), последующему квантованию с шагом ΔU и аналого-цифровому преобразованию.

С учетом особенностей слухового аппарата человека и принимая во внимание теорему Котельникова, стандартизированы следующие уровни частот:

- 44 кГц (для тренированного уха максимальной воспринимаемой считается $f_m = 20$ кГц);
- 22 кГц (в случае, когда восприятие высокочастотных составляющих не критично $f_m = 10$ кГц);
- 8 кГц (для речевого сигнала — с учетом экспериментально установленного порога «слоговой разборчивости» $f_m = 3{,}4$ кГц).

Способ представления непрерывного сигнала последовательностью кодов, которые отражают амплитуду импульсов-отсчетов, называется импульсно-кодовой модуляцией (ИКМ).

Частота дискретизации согласно теореме Найквиста – Котельникова должна быть не меньше удвоенной максимальной частоты спектра сигнала.

Погрешности квантования определяются исходя из емкости кода для максимального уровня сигнала — 8 или 16 бит на отсчет.

Один из классических способов сжатия — так называемая «адаптивная дифференциальная ИКМ» (АДИКМ). Она основана на следующем:

- вместо кода абсолютного значения очередного импульса передают код изменения его амплитуды по отношению к предыдущему (дифференциальному);
- при увеличении скорости передачи сигнала шаг его квантования возрастает, а при уменьшении — уменьшается.

Использование АДИКМ позволяет уменьшить объем передаваемых данных примерно на порядок.

Второй способ сжатия именуется компадированием и использует известную особенность человеческого слухового аппарата, вследствие которой ухо менее критично к погрешностям квантования при более громком звуке.

Механизм компадирования работает следующим образом. Звуковой сигнал перед оцифровкой подается на так называемый «компрессор» КМП — нелинейный преобразователь, «растягивающий» слабые сигналы и «сжимающий» сильные. После передачи зашумленный погрешностями АЦП сигнал подается на вход

«эспандера», который в противоположность компрессору сжимает слабые сигналы и растягивает сильные, восстанавливая первоначальный баланс, при этом погрешности ΔU, попавшие в область слабых сигналов, уменьшаются, а сильных — возрастают.

Это не только позволяет учесть особенности человеческого восприятия звука, но и приводит к уменьшению «среднестатистических» погрешностей квантования — дело в том, что слабые сигналы встречаются чаще.

Стоит добавить, что в современных системах вместо комбинации аналоговый компрессор (эспандер) + АЦП (ЦАП) реально используются АЦП с нелинейной характеристикой.

Наряду с описанными, при сжатии звука используются еще дополнительно некоторые способы, например классическое сжатие.

Однако часто цифровые файлы, или потоки, бывают двумерными, например возникающими в результате оцифровки изображений. В этом случае использование двумерных (строчных и столбцовых) вероятностных связей при сжатии изображений дает дополнительный эффект. В настоящее время в мире существуют три стандарта представления цветного телевизионного видеосигнала: PAL, SECAM и NTSC. Все они преобразуют при оцифровке исходный аналоговый видеосигнал в форму, удобную для передачи по линии связи. Несмотря на различие в методе кодирования видеоинформации, все три системы телевидения представляют исходный телевизионный сигнал в виде сигнала яркости (Y) и двух цветоразностных ($R-Y$) и ($B-Y$) сигналов, а затем каждая по-своему кодирует его.

Один кадр телевизионного изображения представляется в цифровом виде как поле 768×576 точек или, как еще говорят, пикселей — в системе PAL/SECAM и 640×480 — в NTSC. Так как человеческий глаз более чувствителен к перепадам яркости, чем к цвету, важнее точная передача сигнала яркости, при оцифровке видеосигнала соотношение числа отсчетов сигнала яркости и цветоразностных сигналов принято равным 4 : 2 : 2, т. е. частота измерения сигнала Y в два раза выше, чем частота измерения сигналов $R-Y$ и $B-Y$, — одна строка оцифрованного телевизионного видеосигнала содержит 768 измерений яркости и по 384 измерения каждого из цветоразностных сигналов. Соотношение частот оцифровки яркости и цвета также определяет качество изображения, которое впоследствии может быть восстановлено. Это отношение может выражаться соотношениями от 4 : 1 : 1 для формата DV до 4 : 4 : 4 для цифрового формата. Оцифровка по стандарту 4 : 4 : 4 предполагает использование одной и той же частоты дискретизации для яркостного и цветоразностных сигналов. Выбор такого способа оцифровки очень хорошо стыкуется с системами PAL/SECAM и NTSC, разница заключена в количестве строк и кадров в секунду (576/25 для PAL/SECAM и 486/30 для NTSC).

Нетрудно подсчитать, что при использовании 10-битного АЦП одна секунда оцифрованной видеозаписи будет занимать на диске файл размером 27 684 000 байт, а для хранения всего 49 секунд видеоинформации со звуком потребуется более 1 Гбайт свободного дискового пространства (что соответствует скорости передачи более 160 Мбит/с).

Чтобы уменьшить размеры файла или скорости цифровых потоков с видеоинформацией, применяют различные методы сжатия (кодеки). Наиболее распро-

страненными являются два метода сжатия: M-JPEG (неподвижная «картинка») и MPEG (движущееся изображение). Оба метода основаны на так называемом дискретно-косинусном преобразовании изображения (частный случай дискретного преобразования Фурье), разработанном для формата JPEG. Отличие состоит в том, что при сжатии по методу M-JPEG устраняется внутрикадровая избыточность, т.е. такое сжатие устраняет избыточность каждого кадра в отдельности. Сжатие по методу MPEG, кроме устранения внутрикадровой избыточности, устраняет и межкадровую.

Метод MPEG имеет много разновидностей, например MPEG1 и MPEG2. Метод MPEG1 не поддерживает чересстрочной развертки и работает только с половинным разрешением. Для сохранения оцифрованного видеосигнала с полным разрешением нужно использовать формат MPEG2. Следует отметить, что MPEG-сжатие существенно сложнее M-JPEG и требует специального оборудования или программного обеспечения. Существующие сегодня алгоритмы и их реализации достигают коэффициентов сжатия в 30 раз и более без заметного ухудшения качества изображения.

Приведем краткое описание алгоритма JPEG.

В настоящее время широко используется «блочная» версия алгоритма, в которой все изображение разбивается на блоки 8×8 и в дальнейшем эти блоки «огрубляются» таким образом, чтобы код, который их описывает, стал как можно короче (исходное описание каждого такого блока требует $8 \times 8 \times 3 = 192$ байта). Дальнейший алгоритм включает следующие этапы:

- происходит определенная взаимообратимая смена кодирования от RGB к YUV;
- выделяются блоки 8×8 пикселей. При этом каждой из составляющих YUV отвечает матрица коэффициентов P_{ij}. (Например, в матрице U значение $P_{ij} = 0$ означает, что у данной точки отсутствует красная цветовая составляющая, а $P_{ij} = 255$ значит, что она будет максимально яркой.);
- выполняется так называемое «прореживание». Четверки блоков изображения 8×8 объединяются в «макроблоки» 16×16, а затем для цветовых составляющих U и V из соответствующих матриц исключаются все четные строки и столбцы. При этом матрица яркостной составляющей Y остается нетронутой. В итоге количество коэффициентов сокращается вдвое (вместо $4 \times 3 = 12$ блоков остается 6);
- для всех оставшихся блоков изображения выполняется так называемое дискретное косинус-преобразование (ДКП). При этом матрицы $N \times N$ ($N = 8$) коэффициентов P преобразуются в матрицы D в соответствии со следующей процедурой:

$$d_{ij} = \frac{1}{2N} C_i C_j \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} p_{xy} \cos\left[\frac{(2x+1)i\pi}{2N}\right] \cos\left[\frac{(2y+1)i\pi}{2N}\right]$$

Смысл косинус-преобразования заключается в том, что коэффициенты d_{ij} отражают «амплитуду колебаний» яркости пикселей. Например, если все пиксели блока имеют одинаковую яркость, то максимальным будет коэффициент d_{11}, а остальные $d_{ij} = 0$. Чем больше деталей в изображении, тем большими будут значения «удаленных» коэффициентов;

— все коэффициенты масштабируются, т. е. огрубляются; если их поделить, скажем, на 8, то длина кода для каждого из них сократится на 3 бита из 8. Но поскольку «удаленные» коэффициенты (правая нижняя часть матрицы) обычно малы, они просто обнуляются в результате масштабирования. Нужно сказать, что именно на этом этапе информация о деталях изображения необратимо теряется. Регулируя величину делителя, можно задавать соотношение «степень сжатия–качество восстановленного изображения»;

— последний этап — кодирование оставшихся коэффициентов. Он включает несколько шагов: коэффициенты записываются в цепочку в порядке «обхода» (в примере цепочка имеет вид: 16 8 6 1 3 4 0 0 0 1 ...); последовательности нулей в цепочке кодируются методом повторов; длины «нулевых» серий кодируются по Хаффману. В результате подобного кодирования примерно из полутора тысяч бит, описывающих блок, остается обычно несколько десятков — код изображения сжимается в десятки раз.

2.3. Общее понятие о шифровании информации

Рассмотрим классическую модель Шеннона для канала связи с шифрованием. Между источником и приемником существует два канала — открытый и закрытый. В открытом канале передачи $Кн°$ сообщение T от источника $И$ либо может быть перехвачено «злоумышленником» $Зл$, который получает несанкционированный доступ к его содержанию, либо $Зл$ может сообщение исказить, либо $Зл$ может подставить свое сообщение вместо сообщения источника. Первая задача соответствует классическому шифрованию, вторая — работе в условиях преднамеренных помех, третья — электронной подписи. Рассмотрим для простоты первую задачу.

Чтобы избежать перехвата, кодер $К$ шифрует сообщение с помощью функции $Ш = F_{ш}(T_1 K)$, которая использует секретный ключ $К$. Декодер $ДК^{ш}$ выполняет обратную операцию дешифрования $T = F_{ош}(Ш_1 K)$, применяя тот же ключ $К$. Чтобы доставить ключ $К$ получателю сообщений, используется секретный канал $К^{с}$, в котором информация не может быть перехвачена злоумышленником. Разумеется, организовать передачу по секретному каналу намного сложнее, зато, один раз доставив с его помощью секретный ключ, можно затем многократно использовать этот ключ, передавая зашифрованные данные по открытому каналу.

Одним из первых исторических примеров, реализующих эту схему, был так называемый шифр Цезаря, который предполагал замену всех букв сообщения на буквы циклически сдвинутого алфавита. Например, если циклический сдвиг латинского алфавита выполнить на $n = 3$ позиции, то букве a будет соответствовать d, c — e и т. д. В данном случае ключом является значение сдвига n. Зная сам алгоритм подстановки, подобрать ключ не так сложно (перебор по длине алфавита). Между тем теория шифрования исходит из того факта, что сам алгоритм шифрования всегда может быть известен, а неизвестен только ключ. Разумеется, для этого ключ должен быть достаточно длинным, чтобы количество его возможных вариантов было велико, а функция шифрования/дешифрования — достаточно сложной, чтобы затруднить прямой перебор. Например, если вернуться к шифру Цезаря, можно предложить усовершенствованный способ подстановки,

когда исходному алфавиту соответствует измененный, который и является ключом

$$abcde\ldots xyz$$
$$foynl\ldots rem.$$

Теперь количество вариантов ключа стало огромным (порядок 10^{26}). Однако, если использовать сведения о частоте появления отдельных букв, выделять повторяющиеся слова и т. д., перебор становится существенно меньше.

Интересно, что теоретически нераскрываемый шифр все же существует. Чтобы обеспечить абсолютную защиту, необходимо использовать ключ, длина которого не меньше длины сообщения, и применять этот ключ только один раз.

Существуют разные классы шифров с секретным ключом. Блочные шифры предполагают разбиение всего сообщения T на стандартные блоки и шифрование каждого блока

$$Ш_i = F(T_i, K).$$

Такой подход ускоряет решение задачи. Вместе с тем его слабым местом считается повторяемость блоков, которая упрощает несанкционированный доступ.

Примеры шифрования подстановкой рассмотрены на примере шифра Цезаря. Возможны также шифры на основе перестановок, которые дополнительно затрудняют несанкционированный доступ. Идея состоит в записи сообщения в некоторую матрицу по столбцам, а в считывании — по строкам, причем в порядке следования букв в слове-ключе в алфавите. На практике используют сочетания подстановки и перестановки — комбинированные блочные шифры.

Возможны также непрерывные или поточные шифры, в которых сообщение обрабатывается как единое целое. Наиболее распространенный способ — так называемое «гаммирование». «Гаммой» называют битовую последовательность $Г$, которая накладывается на исходное сообщение, представленное в двоичном виде (используется поразрядное суммирование по модулю 2):

$$Ш = T \oplus Г.$$

Сама гамма — это так называемая псевдослучайная последовательность двоичных чисел, получаемая, например, с помощью формулы

$$C_{i+1} = (A + C_i + B) \mod M,$$

где A и B — константы, M — основание модуля. При такой процедуре ключом K является стартовое число C_o. Достоинство такого способа шифрования — его высокое быстродействие. Его слабость в том, что, когда имеется уже зашифрованный текст, по нему легко восстановить участки гаммы, что влечет низкую устойчивость. Способ гаммирования в ряде случаев используют совместно с элементами блочного шифрования.

Наибольшее распространение в прошлом веке получил алгоритм шифрования DES (Data Encrypting Stand), разработанный в 1977 году компанией IBM и принятый как стандарт в США, а затем и во всем мире.

Стандарт DES реализует блочное шифрование данных с использованием 56-битного ключа. Каждый 64-битный блок данных T_i разделяется на «левую» и «правую» половины A_i и B_i. Шифрование включает 16 итераций перестановок и подстановок, в ходе которых к преобразуемым A_i и B_i «подмешиваются» различные фрагменты ключа K (56 бит ключа являются секретными, а еще

8 бит — проверочными символами). Операции шифрования и дешифрования были достаточно удачно оптимизированы по скорости, а надежность шифра уже почти четверть века позволяет ему оставаться стандартом. Хотя сейчас количество вариантов ключа 2^{56} уже считается недостаточно большим (в принципе этот недостаток можно обойти за счет, например, двукратного шифрования на ключах K_1 и K_2, тогда криптостойкость определяется их суммарной длиной).

Однако с появлением и развитием Интернета на повестку дня стала задача передавать секретный ключ по тому же каналу, что и сообщение. Действительно, при сотнях миллионов пользователей Интернета невозможно организовать такое же количество секретных каналов. Тогда были предложены системы с открытыми ключами. Идея таких систем состоит в следующем. У каждого пользователя сети имеется два ключа — открытый *Ко* и закрытый *Кз*. Ключи *Ко* и *Кз* генерируются по определенному алгоритму самим абонентом, но получить один ключ из другого невозможно. Открытые ключи всех абонентов публикуются в сети, а закрытые абоненты хранят у себя.

В случае передачи зашифрованной информации от *А* к *Б* абонент *А* закрывает информацию открытым ключом абонента *Б* — *Коб* по алгоритму $Ш = F_{ш}(T_1 Коб)$. Абонент *Б* раскрывает информацию своим закрытым ключом *Кзб* по алгоритму $T = F_{ош}(Ш_1 Кзб)$.

В случае передачи информации с электронной подписью задача инверсна. Абонент *А* подписывает свою информацию с использованием своего закрытого ключа *Кза* по алгоритму $П = F_{ш}(T_1 Кза)$, а любой другой абонент сети может убедиться в том, что информация подписана *А* при помощи открытого ключа *А–Коа* по алгоритму $T = F_{ош}(П_1 Кзб)$. Система с открытыми ключами является чрезвычайно распространенной.

2.4. Корректирующие коды

2.4.1. Блок-схема системы связи и примеры простейших кодов

Рис. 2.5. Блок-схема системы связи при наличии кодера и декодера корректирующего кода

На рис. 2.5 показана блок-схема системы связи в случае наличия кодера и декодера корректирующего кода, при этом мы считаем, что кодирование, устраняющее избыточность, включено в источник и получатель дискретной информации. Также мы считаем, что канал связи дискретный по входу и по выходу, т. е. модулятор и демодулятор включен в канал. Следует отметить, что мы постоянно будем пользоваться искусственными приемами укрупнения тех или иных блоков теоретико-информационных блок-схем — каналов, источников/получателей информации и т. д.

Рассмотрим два простейших примера корректирующего кода.

Код с проверкой на четность

Список кодовых слов, каждое из которых состоит из информационной и проверочной части, показан в табл. 2.3.

Таблица 2.3. Список кодовых слов

Информационная часть	Проверочная часть
000	0
001	1
010	1
011	0
100	1
101	0
110	0
111	1

Как мы видим, проверочный символ равен сумме по модулю 2 всех трех *информационных. Легко также заметить, что кодовые слова отличаются друг* от друга не менее чем в двух позициях. На три информационных символа кода тратится один проверочный, т.е. относительная скорость кода равна 3/4, а избыточность — 1/4.

Пусть кодовые слова передаются побитно по двоичному симметричному каналу (ДСК), где p — вероятность ошибки, а $(1-p)$ — вероятность правильной передачи. Легко заметить, что, если в кодовом слове исказился один символ или три, эта ситуация может быть обнаружена общей проверкой на четность (определением принадлежности кодового слова к коду). Если же произошло два или четыре искажения (ошибки), слово будет считаться правильным, т.е. произойдет неисправимая ошибка.

Таким образом, возможны три результата декодирования:

A — правильное декодирование (нет ошибок в канале);

$Б$ — обнаружение ошибки (одна или три ошибки в слове);

B — ошибочное декодирование (две или четыре ошибки в слове).

Соответственно, вероятности этих событий вычисляются следующим образом:

$$P_A = (1-p)^4; \qquad P_Б = 4p(1-p)^3 + 4p^3(1-p); \qquad P_B = 6p^2(1-p)^2 + p^4.$$

Если, например, вероятность ошибки в канале ДСК равна 10^{-3}, то вероятность ошибки декодирования равна примерно $6 \cdot 10^{-6}$, а вероятность обнаружения ошибки — $4 \cdot 10^{-3}$. С ухудшением канала (стремлением p к 1/2) вероятность ошибки декодирования P_B растет, но стремится к 7/16, а вероятность обнаружения ошибки $P_Б$ также растет, но стремится к 3/8. Вероятность правильного декодирования P_A падает и стремится к 1/16. Естественно, этот код имеет хоть какой-либо смысл при малых вероятностях ошибки p в канале.

Код-повторение

Список, состоящий всего из двух кодовых слов, показан в табл. 2.4.

Как мы видим, проверочные символы повторяют один информационный. Легко также заметить, что кодовые слова отличаются друг от друга в четырех

позициях. На один информационный символ кода тратится три проверочных, т. е. относительная скорость кода равна 1/4, а избыточность — 3/4.

Таблица 2.4. Список кодовых слов

Информационная часть	Проверочная часть
0	000
1	111

Можно предложить алгоритм декодирования голосованием по большинству одинаковых символов. Тогда правильное декодирование будет происходить при одной или менее ошибок в канале, ошибочное декодирование — при трех или более. При двух ошибках в канале будет обнаруживаться ошибка:

$$P_A = (1-p)^4 + 4p(1-p)^3; \qquad P_Б = 6p^2(1-p)^2; \qquad P_B = 4p^3(1-p) + p^4.$$

Если, например, вероятность ошибки в канале ДСК равна 10^{-3}, то вероятность ошибки декодирования равна примерно 4×10^{-9}, а вероятность обнаружения ошибки — 6×10^{-6}. С ухудшением канала (стремлением p к 1/2) вероятность ошибки декодирования P_B растет, но стремится к 5/16, а вероятность обнаружения ошибки $P_Б$ также растет, но стремится к 3/8. Вероятность правильного декодирования P_A падает и стремится к 5/16. Если сравнивать этот код с предыдущим, видно, что у него существенно выше помехоустойчивость, правда, достигается это ценой существенно большей избыточности. Таким образом, даже на таком простом примере видно, что, как для любой системы связи, существует обмен между помехоустойчивостью и скоростью передачи.

Также можно изменением алгоритма декодирования (порога принятия решения) «перекачивать» вероятности событий A, $Б$ и B. Повысим порог при декодировании второго кода — будем принимать решение, только если 4 символа будут равны 0 или 1. В противном случае всегда будет обнаруживаться ошибка. Тогда вероятности принимают следующий вид:

$$P_A = (1-p)^4; \qquad P_Б = 4p(1-p)^3 + 6p^2(1-p)^2 + 4p^3(1-p); \qquad P_B = p^4.$$

Мы видим, что существенно уменьшилась вероятность ошибки декодирования, правда, ценой увеличения вероятности обнаружения ошибки и уменьшения вероятности правильного декодирования.

Таким образом, на простейших кодах мы проиллюстрировали ключевые свойства теории корректирующего кодирования. Для достижения пропускной способности канала необходимо увеличивать длину кодов. В данном случае код с проверкой на четность будет практически всегда обнаруживать ошибки, а скорость кода повторения будет стремиться к нулю, т. е. ни тот, ни другой код не интересны для практического использования, однако являются хорошей иллюстрацией поведения реальных кодов. Теория корректирующего кодирования, начиная с первых работ Шеннона, как раз и создавала такие легко кодируемые и декодируемые коды на больших длинах, с помощью которых можно стремиться к пропускной способности канала при фиксированной и ненулевой скорости передачи. Для короткого обзора таких методов нам понадобятся некоторые математические определения.

2.4.2. Теорема Шеннона для канала с шумами

Данная теорема является фундаментальным положением теории информации и называется также основной теоремой кодирования Шеннона. Она может быть сформулирована следующим образом.

Если энтропия источника сообщений $H(U)$ меньше пропускной способности канала C ($H(U) < C$), то существует такая система кодирования, которая обеспечивает возможность передачи сообщений источника со сколь угодно малой вероятностью ошибки (или со сколь угодно малой ненадежностью). Если $H(U) > C$, то можно закодировать сообщение таким образом, что ненадежность в единицу времени будет меньше, чем $H(U) - C + e$, где e стремится к нулю (прямая теорема).

Не существует способа кодирования, обеспечивающего ненадежность в единицу времени меньшую, чем $H(U) - C$ (обратная теорема).

В такой формулировке эта теорема была представлена Шенноном [2]. В литературе часто вторая часть прямой теоремы и обратная теорема объединяются в виде обратной теоремы, сформулированной так: если $H(U) > C$, то такого способа кодирования не существует.

Для доказательства первой части прямой теоремы используется множество длинных последовательностей элементарных дискретных сообщений источника длиной T, распадающееся на подмножества высоковероятных или типичных и маловероятных или нетипичных последовательностей. Пусть при некотором ансамбле входных сигналов дискретного канала XY обеспечивается пропускная способность канала

$$C = \max_{Q(0),Q(1),\ldots,Q(K-1)} I(X,Y) =$$

$$= \max_{Q(0),Q(1),\ldots,Q(K-1)} \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} Q(k)P(j) \log \left[P(j/k) \Big/ \sum_{i=0}^{K-1} Q(i)P(j/i) \right],$$

где $Q(0), Q(1), \ldots, Q(K-1)$ — вероятности возникновения K букв на входе канала, $P(j/k)$, $k = 0, 1, \ldots, K-1$; $j = 0, 1, \ldots, J-1$ — переходные вероятности для дискретного канала без памяти с K входами и J выходами, $I(X,Y)$ — средняя взаимная информация между входом X и выходом Y дискретного канала без памяти.

В процессе кодирования каждой типичной последовательности источника ставится в соответствие одна из типичных последовательностей канальных сигналов. Нетипичные последовательности сообщений длительности T (если источник все же выдаст одну из них) не передаются, соглашаясь с тем, что каждая такая последовательность будет принята ошибочно. После выполнения указанного кодирования всеми возможными случайными способами проводится усреднение вероятности ошибок по всему этому большому классу возможных систем кодирования. Это равносильно вычислению вероятности ошибок при случайном связывании типичных последовательностей источника сообщений и канальных сигналов. Далее при помощи усреднения можно оценить среднюю вероятность ошибки.

При любом заданном $e > 0$ можно выбрать столь большое T, что будем иметь среднюю вероятность ошибки меньше e. Но если среднее некоторого множества чисел не больше, чем e, то в этом множестве должно существовать по крайней

мере одно число, меньше e. Тогда среди всех возможных M кодов, обеспечивающих среднее значение вероятности ошибки, обязательно существует хотя бы один, у которого вероятность ошибки не превышает среднее. Таким образом, доказывается первая часть теоремы.

Вторую часть прямой теоремы легко доказать исходя из того, что можно просто передавать C бит в секунду от источника сообщений, совсем пренебрегая остатком создаваемой информации.

Физический смысл эффекта повышения вероятности при увеличении длительности кодируемых сообщений, вытекающий из доказательства прямой теоремы, заключается в том, что с ростом T увеличивается степень усреднения шума, действующего в канале, и, следовательно, уменьшается степень его мешающего воздействия. Кодирование сообщений длительности T способом, предполагаемым при доказательстве теоремы Шеннона, может начаться лишь тогда, когда сообщение целиком поступило на кодирующее устройство. Декодирование же может начаться, когда вся принятая последовательность поступила на декодирующее устройство. Поэтому суммарная задержка сообщений во времени между входом кодера и выходом декодера есть $2T + T_0$, где T_0 — время, затрачиваемое на кодирование, декодирование и прохождение по каналу. При большом T можно принять, что суммарная задержка есть $2T$. Из теоремы также следует принципиальная возможность обмена между вероятностью ошибки, задержкой и скоростью передачи информации. На практике сложность кодирования и декодирования существенно возрастает с ростом T.

2.4.3. Введение в теорию групп, колец и полей

Приведем короткую сводку теории групп, колец и полей в минимальном объеме, *необходимом для чтения литературы по корректирующим кодам.*

Группы

Алгебраическая система $\{G, O\}$, образованная непустым множеством G и операцией O, определенной для любых двух элементов a и b из G, называется группой, если выполнены следующие четыре аксиомы:

А1 (замкнутость).

Для любых двух элементов a и b из G однозначно определен элемент aOb, принадлежащий G.

А2 (ассоциативность).

Для любых трех элементов a, b и c из G выполняется равенство

$$aO(bOc) = (aOb)Oc.$$

А3 (существование единичного элемента).

В G существует элемент e, называемый единичным, такой, что для любого элемента a из G

$$aOe = eOa = a.$$

А4 (существование обратного элемента).

Для любого элемента a из G существует элемент x, называемый обратным к a, такой, что

$$xOa = aOx = e.$$

Группа $\{G, O\}$ называется коммутативной или абелевой, если справедлива следующая аксиома.

А5 (коммутативность).

Для двух произвольных элементов a и b из G

$$aOb = bOa.$$

Если в качестве групповой операции используется $+$, т. е. $\{G, +\}$, то группа называется аддитивной.

Группами является множество целых чисел с операцией обычного сложения и множество комплексных чисел с соответствующим сложением.

Число элементов в группе G называется порядком G. Группа, порядок которой конечен, называется конечной.

Легко доказывается, что единичный элемент в группе единственный и что для каждого элемента группы обратный элемент также единственный.

Пусть $\{G, O\}$ группа, а H — подмножество G, являющееся группой относительно той же групповой операции O. Тогда H называется подгруппой G.

Пусть H подгруппа группы G, а g — произвольный элемент из G. Множество элементов x, таких, что $x = gh$ (для всех h, принадлежащих H), обозначенное ниже через gH, называется левым смежным классом группы G по подгруппе H, порожденным элементом g. Аналогично можно ввести понятие правого смежного класса.

Можно доказать (Лагранж), что порядок любой подгруппы конечной группы является делителем порядка группы.

Тогда любую конечную группу можно однозначно разложить на смежные классы. Если M — порядок конечной группы, N — порядок ее подгруппы, $M = NJ$, то элементы группы могут быть разложены в двумерную таблицу размера J строк на N столбцов, где первая строка есть собственно элементы подгруппы, а остальные строчки есть ее смежные классы.

Более подробно с теорией групп можно ознакомиться в монографиях по алгебраической теории кодирования [3–7].

Кольца

Рассматриваемые ниже кольца и поля в отличие от групп являются алгебраическими системами с двумя операциями.

Пусть R — множество, для любых двух элементов которого определены две операции $+$ и $*$. Алгебраическая система $\{R, +, *\}$ называется кольцом, если выполняются следующие три аксиомы.

Б1. $\{R, +\}$ — коммутативная группа.

Б2. $\{R, *\}$ — полугруппа, т. е. алгебраическая система, для которой выполняются аксиомы А1 и А2.

Б3 (дистрибутивность). Для любых трех элементов a, b и c из R выполняются тождества

$$a(b + c) = ab + ac,$$
$$(a + b)c = ac + bc.$$

Если операция $*$ коммутативна в кольце, то кольцо называется коммутативным. Если в кольце существует единичный элемент относительно операции $*$, то кольцо называется кольцом с единицей.

Целые числа с обычным сложением и умножением являются, например, коммутативным кольцом с единицей.

Множество всех многочленов вида

$$a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n, \qquad a_i \in R,\ 0 \leqslant n$$

с коэффициентами из R является кольцом относительно обычных операций сложения и умножения многочленов. Это кольцо называется кольцом многочленов от неизвестного x над кольцом R и обозначается через $R[x]$.

Подмножество I кольца R называется идеалом кольца R, если оно удовлетворяет следующим двум условиям:

В1. Для любых двух элементов a и b из I элемент $a + b$ принадлежит I.

В2. Для любого элемента x кольца R и любого a из I оба произведения ax и xa лежат в I.

Если I идеал кольца R, то в R следующим образом можно ввести новое бинарное отношение:

$$a \equiv b \,(\mathrm{mod}\, I) \Leftrightarrow a - b \in Ia \equiv b \,(\mathrm{mod}\, I) \Leftrightarrow a - b \in I.$$

Поля

Пусть F — множество, содержащее по крайней мере два элемента, такое, что для любых двух элементов a и b из F определены две операции $+$ и $*$. Алгебраическая система $\{F, +, *\}$ называется полем, если она удовлетворяет следующим аксиомам:

Г1. $\{F, +\}$ — коммутативная группа.

Пусть O — единичный элемент группы $\{F, +\}$, а F^* — множество, полученное из F удалением элемента O.

Г2. $\{F^*, *\}$ — коммутативная группа.

Г3 (дистрибутивность). Для любых трех элементов a, b и c из F справедливы тождества

$$a(b + c) = ab + ac,$$
$$(a + b)c = ac + bc.$$

Множество R действительных чисел и множество C комплексных чисел являются полями относительно операций обычного сложения и умножения.

Поле, состоящее из конечного числа элементов, называется конечным полем, а число его элементов p — порядком поля.

Любое поле имеет следующие свойства:

– для любого элемента поля a выполняется тождество

$$a*0 = 0*a = 0;$$

– для любых двух ненулевых элементов поля a и b выполняется тождество

$$a*b \neq 0;$$

– для любых двух ненулевых элементов поля a и b выполняется тождество

$$-(a*b) = (-a)*b = a*(-b);$$

– если $a*b = a*c$ и $a \neq 0$, то $b = c$;
– по определению поле должно содержать не меньше двух элементов.

Операции в поле из двух элементов показаны в табл. 2.5.

Таблица 2.5. Операции в поле из двух элементов

+	0	1
0	0	1
1	1	0

*	0	1
0	0	0
1	0	1

Если порядок поля задан, то степень свободы при построении поля не так велика. Но если порядок поля фиксирован, конечные поля определяются одинаково с точностью до изоморфизма (отображения одного в другое).

Существует очень простой способ построения некоторых конечных полей. Если порядок поля p — простое число, то операции в поле выполняются по модулю этого простого числа p (mod p). Ниже показаны примеры операций в поле из трех, пяти и семи элементов (табл. 2.6–2.8).

Таблица 2.6. Пример операции в поле из трех элементов

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

*	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

Таблица 2.7. Пример операции в поле из пяти элементов

+	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

*	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Таблица 2.8. Пример операции в поле из семи элементов

+	0	1	2	3	4	5	6
0	0	1	2	3	4	5	6
1	1	2	3	4	5	6	0
2	2	3	4	5	6	0	1
3	3	4	5	6	0	1	2
4	4	5	6	0	1	2	3
5	5	6	0	1	2	3	4
6	6	0	1	2	3	4	5

*	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6
2	0	2	4	6	1	3	5
3	0	3	6	2	5	1	4
4	0	4	1	5	4	6	3
5	0	5	3	1	6	4	2
6	0	6	5	4	3	2	1

Легко заметить по таблицам операций, что выполняются все аксиомы и свойства полей.

Так, в поле из трех элементов «1» и «2» обратны друг другу по сложению, а «2» обратна «2» по умножению.

В поле из пяти элементов по сложению обратны друг другу «1» и «4», а также «2» и «3», а по умножению обратны друг другу «2» и «3», а также «4» и «4».

В поле из семи элементов по сложению обратны друг другу «1» и «6», «2» и «5», «3» и «4», а по умножению — «2» и «4», «5» и «3», «6» и «6».

Если подмножество F_0 поля F само является полем, то F_0 называется подполем F, а F называется расширением поля F_0.

Можно доказать, что любое конечное поле является расширением поля из простого числа элементов p.

В любом конечном поле F порядок p единичного по умножению элемента 1 как элемента аддитивной группы поля F определяется однозначно и называется характеристикой поля F. В поле F характеристики p для любого элемента a поля F p-кратная сумма элемента a всегда равна 0.

В дальнейшем любое конечное поле из q элементов будем называть полем Галуа и обозначать $GF(q)$. По аналогии с полем из простого числа элементов можно строить поля с числом элементов, равным любой целой степени простого числа. Только операции в таком поле выполняются над многочленами по модулю некоторого многочлена, а элемент поля задается коэффициентами многочлена. Ниже для простоты будем рассматривать поля с числом элементов, равным любой целой степени двойки.

Поле $GF(2^a)$, являющееся расширением поля $GF(2)$, содержит 2^a элементов. Все 2^a-1 ненулевые элементы образуют циклическую мультипликативную группу. Это значит, что в поле $GF(2^a)$ существует хотя бы один элемент b, такой, что любой ненулевой элемент $GF(2^a)$ представляет собой некоторую степень элемента b. Элемент b, обладающий этим свойством, называется примитивным элементом поля $GF(2^a)$.

Каждый элемент поля $GF(2^a)$ можно представить в виде слова длины a над полем $GF(2)$ или многочлена $f(x)$ над полем $GF(2)$, т. е. с коэффициентами из $GF(2)$, степень которого меньше a.

В этом случае сложение элементов p, q из поля $GF(2^a)$, т. е. $p+q=r$, выполняется по правилу представляющих их многочленов, $p(x)+q(x)=r(x)$ — у соответствующих многочленов складываются поразрядно коэффициенты при x, причем сложение происходит в поле $GF(2)$.

Умножение элементов в поле $GF(2^a)$, т. е. $pq=r$, выполняется по правилу представляющих эти элементы многочленов по модулю некоторого заданного многочлена $F(x)$ степени a, т. е. $p(x)q(x)=r(x) \mod F(x)$.

Деление одного элемента p поля $GF(2^a)$ на другой элемент q поля $GF(2^a)$ соответствует умножению многочлена $p(x)$ на многочлен $r(x)$, соответствующий элементу r, обратному q, где многочлен $r(x)$ должен удовлетворять условию

$$p(x)q(x)=r(x) \mod F(x).$$

В качестве многочлена $F(x)$ следует выбирать такой многочлен, один из корней которого является примитивным элементом поля $GF(2^a)$.

Этот примитивный элемент b представляется многочленом x или словом вида $(00\ldots10)$, т. е. $b=x=(00\ldots10)$.

Единичный элемент поля $GF(2^a)$ — b^0 записывается в виде $b^0=1=(00\ldots01)$, а все остальные элементы выражаются как степени примитивного элемента b, причем $b(2^a-1)=b^0=1$.

Таким образом, при последовательном вычислении

$$b^2 = bb;\ b^3 = b^2b;\ \ldots;\ b^{m+1} = b^mb;\ \ldots$$

как произведения по модулю многочлена $F(x)$, определяются все элементы поля. Отсюда ясно, что поле $GF(2^a)$ полностью определяется многочленом $F(x)$. С подробным обоснованием структуры и правил построения конечных полей Галуа можно ознакомиться в [3–7]. Мы же ограничимся тремя примерами построения полей $GF(4), GF(8)$ и $GF(16)$ (табл. 2.9–2.11).

Легко убедиться, что приведенных таблиц вполне достаточно, чтобы выполнять все операции в заданном поле. Если надо складывать элементы, удобнее пользоваться их многочленным представлением, если элементы надо перемножать, удобнее пользоваться представлением через степень примитивного элемента. Если надо поделить один элемент на другой, надо умножить делимое на элемент, обратный делителю.

Таблица 2.9. Поле $GF(4)$, многочлен $F(x) = x^2 + x + 1$

Номер элемента поля	Степень примитивного элемента	Соответствующий элементу многочлен	Двоичное представление элемента
1	–	0	0 0
2	b^0	1	0 1
3	b	x	1 0
4	b^2	$x+1$	1 1

Таблица 2.10. Поле $GF(8)$, многочлен $F(x) = x^3 + x + 1$

Номер элемента поля	Степень примитивного элемента	Соответствующий элементу многочлен	Двоичное представление элемента
1	–	0	0 0 0
2	b^0	1	0 0 1
3	b	x	0 1 0
4	b^2	x^2	1 0 0
5	b^3	$x+1$	0 1 1
6	b^4	x^2+x	1 1 0
7	b^5	x^2+x+1	1 1 1
8	b^6	x^2+1	1 0 1

Пример для поля $GF(16)$. Пусть мы хотим сложить 10-й и 11-й элементы. Для этого мы поразрядно складываем их двоичные представления и получаем вектор (1 1 1 1), т. е. 14-й элемент. Если мы хотим эти элементы перемножить, мы пользуемся их степенным представлением и получаем $b^8b^9 = b^{17} = b^2b^{15} = b^21 = b^2$, т. е. 4-й элемент. Если мы хотим поделить 10-й элемент на 11-й, мы находим элемент, обратный 11-му по умножению, b^6 (8-й элемент), потому что $b^6b^9 = b^{15} = 1$. Затем мы умножаем 10-й элемент на 8-й и получаем $b^8b^6 = b^{14}$, т. е. 16-й элемент.

Отметим еще одно интересное свойство поля. Например, рассмотрим поле $GF(8)$. В таблице двоичных представлений элементов рассмотрим поведение любого разряда, например младшего, для всех элементов, кроме нулевого. Получаем вектор (1 0 0 1 0 1 1). Это так называемая последовательность максимальной длины или m-последовательность. В общем случае последовательность длины

$2^a - 1$ генерируется полиномом степени a, что мы и видели при построении поля. Любые ее циклические сдвиги отличаются друг от друга в 2^{a-1} позициях. В данном примере, если рассмотреть все возможные семь циклических сдвигов этой последовательности и добавить нулевое слово, мы получим корректирующий код из восьми слов длины 7, любые два слова которого отличаются в четырех позициях.

Таблица 2.11. Поле $GF(16)$, многочлен $F(x) = x^4 + x + 1$

Номер элемента поля	Степень примитивного элемента	Соответствующий элементу многочлен	Двоичное представление элемента
1	–	0	0 0 0 0
2	$b^0 x^4$	1	0 0 0 1
3	b	x	0 0 1 0
4	b^2	x^2	0 1 0 0
5	b^3	x^3	1 0 0 0
6	b^4	$x+1$	0 0 1 1
7	b^5	x^2+x	0 1 1 0
8	b^6	x^3+x^2	1 1 0 0
9	b^7	x^3+x+1	1 0 1 1
10	b^8	x^2+1	0 1 0 1
11	b^9	x^3+x	1 0 1 0
12	b^{10}	x^2+x+1	0 1 1 1
13	b^{11}	x^3+x^2+x	1 1 1 0
14	b^{12}	x^3+x^2+x+1	1 1 1 1
15	b^{13}	x^3+x^2+1	1 1 0 1
16	b^{14}	x^3+1	1 0 0 1

2.4.4. Введение в пространства Хемминга

Для дальнейшего рассмотрения нам понадобятся некоторые определения. Двоичным пространством Хемминга размерности n называется множество 2^n всех возможных двоичных векторов длины n. Каждый такой вектор ассоциируется с точкой в этом пространстве $\bar{x}_n = (x_1, x_2, \ldots, x_n)$, $x_i \in \{0,1\}$. По аналогии с обычным евклидовым пространством существует n ортогональных измерений (осей), по каждому измерению (оси) возможны только два значения 0 и 1. Набор всех возможных точек пространства Хемминга называется n-мерным кубом. Естественно, возможно обобщение на q-й случай, когда все элементы принимают значения из поля $GF(q)$, но мы в дальнейшем будем рассматривать двоичный случай для простоты и экономии объема.

Весом вектора (или точки) $\bar{x}$ будем называть число ненулевых (для двоичного случая единичных) компонентов вектора и обозначать $\omega(\bar{x})$. Соответственно, расстоянием Хемминга между двумя векторами $\bar{x}$ и $\bar{y}$ $d(\bar{x}, \bar{y})$ будем называть вес вектора z, являющегося результатом поразрядного сложения по модулю 2 двух векторов $\bar{x}$ и $\bar{y}$, т. е. $d(\bar{x}, \bar{y}) = \omega(\bar{z} = \bar{x} \oplus \bar{y})$.

Сферой $Sp(r)$ радиуса r вокруг некоторой точки $\bar{x}$ будем называть множество таких точек $\bar{y}$, расстояние Хемминга которых до точки $\bar{x}$ равно r. Соответственно, шаром $Sh(r)$ радиуса r вокруг некоторой точки $\bar{x}$ будем называть множество

таких точек $\bar{y}$, расстояние Хемминга которых до точки $\bar{x}$ равно или меньше r. Площадью сферы $S_{Sp(r)}$ будем называть число точек, лежащих на этой сфере. Объемом шара $V_{Sh(r)}$ будем называть число точек, содержащихся в этом шаре. Легко подсчитать, что $S_{Sp(r)} = C_n^r$, а

$$V_{Sh(r)} = \sum_{i=0}^{r} C_n^i.$$

В общем случае корректирующим блочным кодом A в двоичном пространстве Хемминга размерности n будем называть набор M точек в этом пространстве, таких, что вокруг каждой точки можно провести сферы радиуса r, не касающиеся друг друга. Легко убедиться, что тогда минимальное расстояние по Хеммингу между кодовыми словами будет удовлетворять оценке $d_{\min} = 2r + 1$. Для простоты рассмотрения будем считать $k = \log_2 M$. Тогда результирующий корректирующий код будем обозначать $(n, k, d_{\min})$.

Пусть все V кодовых слов перенумерованы двоичными векторами из k бит. Эти двоичные векторы будем называть информационными векторами. Процесс кодирования состоит в замене информационного вектора длины k на кодовый вектор длины n. Соответственно, величина $R = k/n$ называется относительной скоростью кода. Далее кодовое слово $\bar{x}$ передается по двоичному симметричному каналу и на выходе канала имеем слово $\bar{y} = \bar{x} \oplus \bar{e}$, где $\bar{e}$ есть вектор ошибок, происшедших в канале.

Пусть декодер устроен следующим образом. Он вычисляет расстояние Хемминга между принятым словом $\bar{y}$ и всеми кодовыми словами $\bar{x} \in A$. В качестве декодированного выбирается то слово $\bar{x}'$, которое имеет наименьшее расстояние по Хеммингу до принятого слова. Соответствующий этому слову информационный вектор выдается получателю. Такое декодирование будем называть по минимуму расстояния.

Можно ограничить принятие решения по декодированию дополнительным сравнением результирующего слова с минимальным расстоянием r. Если оно меньше либо равно r, слово выдается получателю; если больше r, происходит отказ от декодирования — обнаруживается ошибка. Понятно, что такой алгоритм декодирования и такой код гарантированно исправляют r ошибок, происшедших в канале. Если при передаче слово остается внутри шара, соответствующего переданному слову, декодирование происходит правильно. Если принятое слово попадает в другой шар (соответствующий другому кодовому слову) — происходит ошибочное декодирование. Если слово не попадает ни в какой шар, происходит обнаружение ошибки.

Можно искусственно уменьшить радиус сферы, по которому принимается решение о декодировании. Пусть решение принимается по величине $r' < r$. Тогда мы будем говорить, что код и данный алгоритм декодирования гарантированно *исправляют r' ошибок и дополнительно обнаруживают $2r - r'$ ошибок*, причем $d_{\min} = r' + (2r - r') + 1 = 2r + 1$. Другой крайней границей (см. примеры в начале главы) является отсутствие исправления и обнаружение $2r$ ошибок.

В приведенном в начале главы примере в коде-повторении рассматриваются две сферы — вокруг нулевого и единичного слова, а в коде с проверкой на четность кодовыми сферами является половина всех точек пространства Хемминга с четным весом.

Отметим также, что существуют коды, для которых сферы радиуса r занимают все пространство Хемминга — точек между сферами не остается. Это свойство выполняется при соблюдении следующего условия (равенство объемов 2^k шаров радиуса r и объема всего пространства Хемминга)

$$2^k \sum_{i=0}^{r} C_n^i = 2^n.$$

Это условие называется границей Хемминга, а коды — соответствующими этому условию, лежащими на границе Хемминга, или совершенными кодами. В случае декодирования совершенных кодов по расстоянию r вероятность обнаружения ошибки будет равна 0 — будет либо ошибочное, либо правильное декодирование.

2.4.5. Линейные коды

Линейные коды составляют лишь небольшой подкласс блочных кодов, однако благодаря своей красивой структуре они стали основным объектом в теории кодирования. Пусть V_n — векторное пространство размерности n над полем $GF(q)$. Подпространство размерности k пространства V_n называется q-м линейным кодом длины n с k информационными символами или (n, k) кодом.

Если A — линейный код, то подпространство A будем называть линейным подкодом A. При $q = 2$ линейные коды называются групповыми. Пусть $d_{\min}$ — минимальное расстояние линейного кода A. Тогда для любого ненулевого вектора $\bar{v} \in A$ выполняется неравенство $\omega(\bar{v}) = \omega(\bar{v} - \bar{0}) \geqslant d_{\min}$, так как нулевой вектор также принадлежит коду A. Для определения линейного кода удобно пользоваться матрицей G. Пусть $\bar{g}_1, \bar{g}_2, \ldots, \bar{g}_k$ есть некоторый базис (n, k) кода A и пусть матрица G — матрица из k строк и n столбцов, i-я строка которой — базисный вектор $\bar{g}_i$. Матрица G называется порождающей матрицей кода A. Из свойства базиса следует, что любой кодовый вектор может быть представлен как линейная комбинация строк $\bar{g}_1, \bar{g}_2, \ldots, \bar{g}_k$ матрицы G и, наоборот, любая линейная комбинация строк $\bar{g}_1, \bar{g}_2, \ldots, \bar{g}_k$ матрицы G представляет собой кодовый вектор и, более того, различные кодовые комбинации представляют собой различные кодовые векторы. Общее число кодовых векторов равно q^k.

Для задания линейных кодов также используется матрица H, которая определяется следующим образом. Пусть A_0 — множество всех векторов $\bar{y} = (y_1, y_2, \ldots, y_n)$, таких, что для любого $\bar{x} = (x_1, x_2, \ldots, x_n) \in A$ выполняется равенство

$$\bar{y}\bar{x}^T = \bar{0} \quad (\text{т. е.} \sum_{i=1}^{n} y_i x_i = 0),$$

где индекс T означает транспонирование.

Как известно, A_0 является подпространством размерности $(n-k)$ векторного пространства размерности n. Это подпространство называется подпространством, двойственным коду A. Порождающую матрицу H двойственного кода A_0 назовем проверочной матрицей кода A. Проверочная матрица имеет размер $(n-k) \times n$ и, по определению, есть

$$H\bar{x}^T = \bar{0} \Leftrightarrow \bar{x} \in A.$$

В частности, если взять в качестве векторов $\bar{x}$ базисные векторы кода A, то получим

$$HG^T = \bar{0}.$$

Поскольку по определению ранг матрицы H равен $(n-k)$, то H содержит невырожденную подматрицу H' размера $(n-k)\times(n-k)$. Путем перестановки столбцов H и соответствующего изменения нумерации компонент кодовых векторов всегда можно добиться того, чтобы матрица H' была образована последними $(n-k)$ столбцами H. С помощью элементарных операций над строками матрица H всегда может быть приведена к следующему виду:

$$H_0 = [P, I_{n-k}],$$

где P — некоторая матрица размера $(n-k)\times k$, а I_{n-k} — единичная матрица размера $(n-k)\times(n-k)$.

Поскольку матрица H_0 получена элементарными операциями над строками, то

$$H\bar{x}^T = \bar{0} \Leftrightarrow H_0\bar{x}^T = \bar{o},$$

т. е. матрица H_0 также называется проверочной матрицей кода A.

Обозначим через p_{ij} элемент подматрицы P матрицы H_0, стоящий в i-й строке и j-м столбце. Очевидно, что необходимым и достаточным условием принадлежности вектора $\bar{x} = (x_1, x_2, \ldots, x_n)$ коду A является выполнение следующей системы неравенств:

$$x_{r+i} = -\sum_{j=1}^{k} p_{ij}x_j, 1 \leqslant i \leqslant n-k.$$

Таким образом, если компоненты $x_1, x_2, \ldots, x_k \in GF(q)$ заданы, то существует ровно один кодовый вектор, первые k компонентов которого — $x_1, x_2, \ldots, x_k$, и можно вычислить по проверочным соотношениям остальные компоненты этого вектора. Первые k символов называются информационными, а остальные $(n-k)$ — проверочными.

Справедливо следующее утверждение. Пусть проверочная матрица H_0 линейного (n,k) кода A имеет вид $H_0 = [P, I_{n-k}]$. Тогда матрица $G_0 = [I_r - P^T]$ является порождающей матрицей кода A. Справедливо также и обратное утверждение.

Можно доказать следующее утверждение. Минимальный вес линейного (n,k) кода A равен d тогда и только тогда, когда любые $(d-1)$ столбцов проверочной матрицы этого кода линейно независимы, но некоторые d столбцов этой матрицы линейно зависимы.

Следствием этого является неравенство $d \leqslant n-k+1$.

Пусть $\bar{x}$ — переданный кодовый вектор по двоичному симметричному каналу, $\tilde{\bar{x}} = \bar{x} \oplus \bar{e}$ — принятый вектор, а $\bar{e}$ — вектор ошибок. Тогда вектор $\bar{s} = \tilde{\bar{x}}H^T$ называется синдромом. По определению проверочной матрицы

$$\bar{s} = \tilde{\bar{x}}H^T = (\bar{x} \oplus \bar{e})H^T = \bar{x}H^T \oplus \bar{e}H^T = \bar{e}H^T.$$

Векторы $\bar{x}, \bar{y}$ называются сравнимыми по модулю A, если $(\bar{x} - \bar{y}) \in A$; сравнимость векторов определяется следующим образом:

Из определения сравнимости векторов следует, что:

$$\bar{x} \equiv \bar{x} \mod A;$$

$$\text{если } \bar{x} \equiv \bar{y} \mod A, \quad \text{то } \bar{y} \equiv \bar{x} \mod A;$$

$$\text{если } \bar{x} \equiv \bar{y},\ \bar{y} \equiv \bar{z} \mod A, \quad \text{то } \bar{x} \equiv \bar{z} \mod A.$$

В этом случае все пространство Хемминга размерности n можно разбить на непересекающиеся классы таким образом, что любые два вектора из одного класса будут сравнимы по модулю A. Эти классы называются смежными классами. Смежный класс, содержащий вектор $\bar{x}$, обозначается через $(\bar{x})$. Смежный класс $(\bar{0})$ совпадает с кодом A. Каждый смежный класс состоит из q^k векторов, а всего смежных классов q^{n-k}. Для всех векторов из одного смежного класса синдром $\bar{s}$ одинаков. Вектор смежного класса с минимальным весом называется образующим (лидером) смежного класса.

Легко заметить, что для представления порождающей матрицы в таком виде кодирование сводится к умножению информационного слова на порождающую матрицу кода. Также важно заметить, что в случае линейных кодов минимальный ненулевой вес кодового слова совпадает с минимальным кодовым расстоянием.

Для декодирования линейного кода можно построить таблицу взаимно-однозначного соответствия лидеров смежного класса и соответствующих им синдромов. Тогда алгоритм декодирования имеет следующий вид:

1) по принятому вектору вычисляется синдром;
2) по синдрому определяется лидер смежного класса;
3) *в случае двоичного кода лидер смежного класса поразрядно суммируется* по модулю 2 с принятым словом и результат выдается получателю.

Этот алгоритм декодирования полностью соответствует описанному выше алгоритму по расстоянию. Для понимания этого факта заметим, что если мы имеем таблицу смежных классов, где в качестве первой строки записан сам код A, а дальнейшие лидеры смежных классов расположены по росту веса Хемминга, то каждому столбцу соответствуют векторы из шара радиуса r вокруг слова из первой строки вместе с некоторыми векторами вне шара (для несовершенных кодов). Таблица записи смежных классов называется таблицей стандартного расположения кода. В ней, как указывалось, первые $V_{Sh(r)} = \sum_{i=0}^{r} C_n^i$ строк соответствуют исправлению ошибок, а остальные $q^{n-k} - V_{Sh(r)} = \sum_{i=0}^{r} C_n^i$ строк — обнаружению.

Интересно отметить одно важное преимущество линейных кодов. Теперь нет необходимости запоминать кодовую таблицу при кодировании. Для обнаружения ошибок в принятом слове также достаточно умножения на матрицу.

Рассмотрим следующий пример линейного кода. Пусть проверочная матрица H представляется в следующем виде:

$$H = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}.$$

Путем линейных комбинаций приведем ее к систематическому виду (с правой единичной матрицей):

$$H_0 = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

Из проверочной матрицы в систематическом виде H_0 получим порождающую матрицу кода:

$$G_0 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

Получился линейный код (7,4), причем, поскольку любые два столбца проверочной матрицы линейно независимы (разные), минимальное расстояние кода по Хеммингу $d_{min} = 3$ и код имеет параметры (7,4,3). Это так называемый код Хемминга и совершенный код. Код имеет 16 кодовых слов. Ниже в табл. 2.12 показано разбиение пространства Хемминга размерности 7 на смежные классы, где первая строка таблицы есть сам код. Всего имеется 8 смежных классов. Также слева в таблице показан соответствующий синдром.

Таблица 2.12. Разбиение пространства Хемминга размерности 7 на смежные классы

Синдром кода	1-е слово	2-е слово	3-е слово	4-е слово	5-е слово	6-е слово	7-е слово	8-е слово
000	0000000	1000011	0100101	1100110	0010110	1010101	0110011	1110000
011	1000000	0000011	1100101	0100110	1010110	0010101	1110011	0110000
101	0100000	1100011	0000101	1000110	0110110	1110101	0010011	1010000
110	0010000	1010011	0110101	1110110	0000110	1000101	0100011	1100000
111	0001000	1001011	0101101	1101110	0011110	1011101	0111011	1111000
100	0000100	1000111	0100001	1100010	0010010	1010001	0110111	1110100
010	0000010	1000001	0100111	1100100	0010100	1010111	0110001	1110010
001	0000001	1000010	0100100	1100111	0010111	1010100	0110010	1110001
Синдром кода	9-е слово	10-е слово	11-е слово	12-е слово	13-е слово	14-е слово	15-е слово	16-е слово
000	0001111	1001100	0101010	1101001	0011001	1011010	0111100	1111111
011	1001111	0001100	1101010	0101001	1011001	0011010	1111100	0111111
101	0101111	1101100	0001010	1001001	0111001	1111010	0011100	1011111
110	0011111	1011100	0111010	1111001	0001001	1001010	0101100	1101111
111	0000111	1000100	0100010	1100001	0010001	1010010	0110100	1110111
100	0001011	1001000	0101110	1101101	0011101	1011110	0111000	1111011
010	0001101	1001110	0101000	1101010	0011011	1011000	0111110	1111101
010	0001110	1001101	0101011	1101000	0011000	1011011	0111101	1111110

Видно, что код является совершенным. В каждом столбце под соответствующим кодовым словом содержатся все (семь) слова сферы радиуса 1 с центром в этом слове. Синдром кода соответствует тому столбцу проверочной матрицы, на позиции которого произошла ошибка.

Также видно еще одно очень важное свойство линейных кодов — спектр расстояний от нулевого слова до всех других кодовых слов такой же, как и от любого

другого кодового слова. Тогда очень легко подсчитать вероятность ошибки декодирования при условии, что передается нулевое слово и исправляется одиночная ошибка. Ошибочное декодирование произойдет тогда, когда принятое слово попадет в сферу любого другого кодового слова. Для этого достаточно подсчитать вес всех слов во всех шарах кроме первого (правильного).

Для этого кода и алгоритма декодирования вероятность обнаружения ошибки $P_{\text{Б}} = 0$, а вероятность ошибки декодирования вычисляется по формуле

$$P_{\text{В}} = 21p^2(1-p)^5 + 30p^3(1-p)^4 + 35p^4(1-p)^3 + 26p^5(1-p)^2 + 7p^6(1-p) + p^7.$$

Подчеркнем, что это есть точное вычисление вероятности ошибки декодирования в соответствии с таблицей стандартного расположения, т.е. спектром весов кода.

Если модифицировать алгоритм, как в примере в начале главы, и уменьшить радиусы сфер от единицы до нуля, тогда вероятность ошибки декодирования будет равна

$$P_{\text{В}} = 7p^3(1-p)^4 + 7p^4(1-p)^3 + p^7.$$

Вероятность обнаружения ошибки в этом случае равна

$$P_{\text{Б}} = 21p^2(1-p)^5 + 23p^3(1-p)^4 + 28p^4(1-p)^3 + 26p^5(1-p)^2 + 7p^6(1-p).$$

В общем случае двоичные коды Хемминга имеют параметры 2^m-1, 2^m-1-m, 3 для любого целого m. Проверочная матрица любого кода Хемминга состоит из всех различных ненулевых столбцов длины m. В пространстве Хемминга шары радиуса 1 вокруг всех кодовых слов заполняют все пространство Хемминга, т. е. код Хемминга является совершенным и для него выполняется граница Хемминга. В случае кодов Хемминга граница Хемминга принимает следующий вид:

$$2^{2^m-1-m}\sum_{i=0}^{1} C_{2^m-1}^i = 2^{2^m-1} \Rightarrow \sum_{i=0}^{1} C_{2^m-1}^i = 2^m \Rightarrow 1 + 2^m - 1 = 2^m.$$

2.4.6. Циклические коды

Циклические коды являются подклассом линейных кодов, которые не изменяют своих свойств при циклической перестановке компонент кодовых слов. Благодаря этому свойству процедура кодирования и простейших процедур декодирования по сложности линейная, а не квадратичная, как в случае линейных кодов. Все совершенные коды, рассмотренные ниже конкретные классы кодов — Рида–Маллера, Боуза–Чоудхури–Хоквингема, Рида–Соломона — имеют циклические аналоги.

Пусть A представляет собой q-й циклический (n, k) код. Если вместе с каждым кодовым вектором $\bar{a} = (a_0, a_1, \ldots, a_{n-1})$ коду также принадлежит вектор $\bar{a}_{\text{Ц}} = (a_{n-1}, a_0, a_1, \ldots, a_{n-2})$, то код называется циклическим. Кодовый вектор удобно также представлять в качестве многочлена $a(x) = a_0 + a_1x + a_2x^2 + + \cdots + a_{n-1}x^{n-1}$ над полем $GF(q)$. Тогда вектору $\bar{a}_{\text{Ц}} = (a_{n-1}, a_0, a_1, \ldots, a_{n-2})$ соответствует многочлен $xa(x) - a_{n-1}(x^n - 1)$. Первая и последняя компоненты циклического кода рассматриваются как соседние, поэтому степени x^0 и x^n тождественны.

Различные многочлены $a(x)$ и $b(x)$ степени $n-1$ и менее принадлежат различным классам вычетов по модулю $x^n - 1$. Класс вычетов по модулю $x^n - 1$, содержащий многочлен $a(x)$, будем называть $(a(x))$. Циклическому сдвигу $\bar{a}_{\text{Ц}}$

соответствует класс вычетов $(xa(x)) = (x)(a(x))$. Множество классов вычетов с соответствующими операциями по модулю многочлена $x^n - 1$ будем обозначать R_n. Если при представлении кодовых векторов в виде последовательностей необходимо рассматривать циклические перестановки (соответствующие операции над матрицами), то при операциях над R_n достаточно лишь линейных по сложности операций над многочленами.

Пусть I — подмножество R_n, соответствующее коду A, а именно:

$$I = \{a(x) \in R_n \Leftrightarrow \bar{a} \in A\}.$$

Тогда для произвольного многочлена $(b(x)) \in R_n$ выполняется тождество $(b(x)a(x)) \in I$, т. е. I является идеалом в R_n.

Пусть $g(x)$ — ненулевой нормированный по старшей степени многочлен минимальной степени, такой, что $(g(x)) \in I$. Пусть $b(x)$ — произвольный многочлен, принадлежащий идеалу I. Тогда, если $r(x)$ — остаток от деления $b(x)$ на $g(x)$, то

$$b(x) = a(x)g(x) + r(x) \quad \text{и} \quad r(x) = (b(x)) - (a(x))(g(x)).$$

Отсюда следует, что $r(x) \in I$.

Но тогда многочлен $r(x)$ должен быть нулевым, т. е. $b(x)$ делиться на $g(x)$.

Таким образом, практически доказана следующая теорема.

Пусть A — циклический (n, k) код и $I = \{a(x) \in R_n \Leftrightarrow \bar{a} \in A\}$. Тогда существует такой нормированный многочлен $g(x)$ степени $n - k$, делящий $x^n - 1$, что выполнение сравнения $b(x) \equiv 0 \mod g(x)$ — необходимое и достаточное условие того, что $b(x) \in I$. Многочлен $g(x)$ определяется однозначно и называется порождающим многочленом циклического кода A.

Пусть $g(x) = g_0 + g_1 x + \cdots + g_{n-k}x^{n-k}$ — порождающий многочлен циклического кода A. Пусть G — матрица из k строк и n столбцов следующего вида:

$$G = \begin{bmatrix} g_0 & g_1 & . & . & . & g_{n-k} & 0 & . & . & . & 0 \\ 0 & g_0 & g_1 & . & . & . & g_{n-k} & 0 & . & . & 0 \\ . & . & . & . & . & . & . & . & . & . & . \\ 0 & . & . & . & g_0 & g_1 & . & . & . & g_{n-k} & 0 \\ 0 & . & . & . & . & g_0 & g_1 & . & . & . & g_{n-k} \end{bmatrix}.$$

Строки матрицы G являются кодовыми словами, а ранг равен k. Следовательно, это порождающая матрица кода A. Порождающий многочлен $g(x)$ служит кратким представлением матрицы. Многочлен $h(x) = (x^n - 1)/g(x)$ называется проверочным многочленом кода. Этот многочлен порождает код A_0, двойственный к коду A.

Пусть $h(x) = h_0 + h_1 x + \cdots + h_k x^k$. Справедлива следующая теорема.

Двойственный код A_0 циклического (n, k) кода A также является циклическим. Порождающий многочлен кода A_0 — это многочлен вида $h_0^{-1}x^k h(x^{-1})$, где $h(x)$ — проверочный многочлен кода A.

Из теоремы следует, что следующая матрица является порождающей матрицей кода A и проверочной матрицей кода A_0:

$$H = \begin{bmatrix} h_k & h_{k-1} & . & . & . & h_0 & 0 & . & . & . & 0 \\ 0 & h_k & h_{k-1} & . & . & . & h_0 & 0 & . & . & 0 \\ . & . & . & . & . & . & . & . & . & . & . \\ 0 & . & . & 0 & h_k & h_{k-1} & . & . & . & h_1 & 0 \\ 0 & . & . & . & 0 & h_k & h_{k-1} & . & . & . & h_1 \end{bmatrix}.$$

Многочлен $b(x) = x^{n-k}b_0(x) + r(x)$, где $b_0(x)$ — информационный многочлен, является кодовым тогда и только тогда, когда он делится на порождающий многочлен $g(x)$. Понятно, что в качестве многочлена проверочных символов можно взять остаток от деления $x^{n-k}b_0(x)$ на $g(x)$, заменив все коэффициенты на противоположные. Это фактически задает алгоритм кодирования по порождающей матрице кода (порождающему многочлену).

Аналогичным образом можно кодировать по проверочной матрице кода (проверочному многочлену). Естественно, аналогичным образом можно вычислять синдром кода. Более подробно теория циклических кодов изложена в классических монографиях [3–7].

Существуют элементарные операции над кодами. К числу самых важных из них относятся укорочение кода и добавление общей проверки на четность. Укорочение кода за счет выкалывания первых m информационных координат сводит (n, k) код с расстоянием $d_{\min}$ к $(n - m, k - m)$ коду с расстоянием не меньше $d_{\min}$. Это бывает очень важно при согласовании источника сообщений с конкретными параметрами кода. Добавление общей проверки на четность сводит (n, k) код с расстоянием $d_{\min} = 2r + 1$ к $(n + 1, k)$ к коду с расстоянием, на единицу большим.

2.4.7. Наиболее известные классы блоковых кодов

Коды Боуза–Чоудхури–Хоквингема

Коды Боуза–Чоудхури–Хоквингема (БЧХ) составляют один из больших классов кодов, исправляющих независимые ошибки веса t, метод построения которых может быть явно задан. Пусть m — произвольное целое положительное число и n — один из делителей $q^m - 1$. В поле $GF(q^m)$ всегда существует элемент порядка n. Пусть b — один из таких элементов. Заметим, что все элементы b^i $(0 \leqslant i < n)$ различны. Можно доказать [3–6] следующую теорему.

Если порождающий многочлен $g(x)$ циклического кода A длины n имеет своими корнями $b^{l+1}, b^{l+2}, \ldots, b^{l+r}$ $(1 \leqslant r < n)$, где l $(0 \leqslant l < n)$ — некоторое целое число, то минимальное расстояние этого кода не меньше чем $r + 1$.

Расстояние $r + 1$, гарантируемое этой теоремой, называется нижней границей кодового расстояния БЧХ-кодов. Чтобы код гарантированно исправлял t ошибок, надо положить $r = 2t$. Тогда минимальное расстояние кода будет удовлетворять оценке $d_{\min} \geqslant 2t + 1$. Эту границу называют конструктивным расстоянием кода. Реальное расстояние может быть больше. Легко убедиться, что число проверочных символов кода не превосходит величины $2mt$.

В частном случае $q = 2$, $n = 2^m - 1$, $l = 0$, $r = 2t$ БЧХ-коды называются примитивными или БЧХ-кодами в узком смысле. При $t = 1$ — это циклические коды Хемминга.

Рассмотрим алгоритм построения двоичных кодов БЧХ с кодовым расстоянием не менее заданного. Для нечетных $d_{\min} = 2t + 1$ элементы $b, b^2, \ldots, b^{2t}$ являются корнями $g(x)$ и для четных $d_{\min} = 2t + 2$ элементы $1, b, b^2, \ldots, b^{2t}$ являются корнями $g(x)$. В обоих случаях порождающий многочлен $g(x)$ есть произведение некоторых минимальных функций $m_i(x)$ [3–6]. В соответствии со свойствами минимальных функций каждая четная степень элемента b, большая нуля, является корнем минимальной функции для некоторой предшествующей нечетной степени элемента b. Например, b^2, b^4 — корни $m_1(x)$, b^6 — корень $m_3(x)$,

b^{10} — корень $m_5(x)$ и т. д. Следовательно, порождающий многочлен можно получить по формуле:

$$g(x) = \left\{ \begin{array}{l} HOK(m_1(x)m_3(x)\dots m_{2t-1}(x)), d_{\min} = 2t+1, \\ HOK(m_0(x)m_1(x)m_3(x)\dots m_{2t-1}(x)), d_{\min} = 2t+2. \end{array} \right\}.$$

Таким образом, построение БЧХ-кода с кодовым расстоянием не менее заданного состоит в следующем.

Задается конструктивное кодовое расстояние $d_{\min} = 2t+1$ или $d_{\min} = 2t+2$ и определяется соответствующее значение t.

Находятся все минимальные функции $m_{2i-1}(x)$, $i = 1, \dots, t$.

Определяется порождающий многочлен $g(x)$, степень которого равна числу проверочных символов. Число информационных символов находится из соотношения $k = n - r - 2^m - 1 - \deg g(x)$.

Проверяется, не делится ли полученный многочлен $g(x)$ на $m_{2t+1}(x)$. Если нет, построение закончено. Если делится, проверяется многочлен большей степени и т. д. Таким образом, находится реальное расстояние $d_{\min}$.

Коды Рида–Соломона

Коды Рида–Соломона (РС) представляют собой частный случай недвоичных циклических кодов БЧХ. Эти коды, как и двоичные, задаются порождающим многочленом $g(x)$ при $l = 0$ и $m = 1$.

Таким образом, коды имеют длину $n = q - 1$ и порождающий многочлен $g(x)$ с коэффициентами из $GF(q)$ должен содержать в качестве своих корней элементы $b, b^2, b^3, \dots, b^{d_{\min}-1}$, где b — примитивный элемент $GF(q)$. Порождающий многочлен определяется следующим соотношением:

$$g(x) = \prod_{i-1}^{d_{\min}-1} (x + b^i).$$

Таким образом, число проверочных символов равно степени порождающего многочлена $d_{\min} - 1$, а число информационных символов $k = n - d_{\min} + 1$.

Как следует из определения линейных кодов, всегда имеется в коде слово веса 1 (только с одной единицей на информационных позициях). Если предположить, что на всех проверочных позициях такого слова также единицы, вес этого слова равен $n-k+1$. Следовательно, в линейном коде минимальное расстояние не может быть больше величины $n-k+1$. Поэтому РС-коды являются кодами с максимально возможным минимальным расстоянием. Естественно, максимальность расстояния достигается за счет недвоичной структуры кода, т. е. двоичных кодов с аналогичными параметрами n и k не существует.

Декодирование кодов БЧХ и РС

Для кодов БЧХ и РС существуют алгебраические процедуры декодирования, сводящиеся в основном к следующей вычислительной процедуре [7, 8]:

1) по принятому слову вычисляется $d_{\min} - 1$ значений «частичных» синдромов. Если все синдромы равны нулю, принятое слово является кодовым;
2) вычисляется многочлен локаторов стираний;
3) вычисляется многочлен обобщенных проверок;
4) вводится многочлен локаторов ошибок;

5) многочлен локаторов ошибок определяется в результате решения ключевого уравнения. Рекурсивная процедура решения этого уравнения подробно изложена в [7, 8];
6) определяются локаторы ошибок, т. е. те места, где произошли ошибки;
7) определяются значения ошибок.

В общем случае алгоритм позволяет исправлять t ошибок и τ стираний, причем $d_{\min} = t + \tau + 1$, где $\tau \geqslant t$.

Достаточно давно существуют вычислительные и аппаратные реализации декодирования (естественно, и кодирования) БЧХ- и РС-кодов достаточно большой длины.

Коды Рида–Маллера

Коды Рида-Маллера (РМ) существуют в обобщенном аналоге, но мы приведем только упрощенный двоичный случай. Двоичные коды РМ существуют в широкой области скоростей, минимальных расстояний и длин. Они эквивалентны циклическим кодам с общей проверкой на четность. Они являются основой для мажоритарно декодируемых кодов.

Для любых m, $r < m$, существует РМ-код, для которого

$$n = 2^m, \qquad k = 1 + C_m^1 + \cdots + C_m^r,$$
$$n - k = 1 + C_m^1 + \cdots + C_m^{m-r+1}, \qquad d_{\min} = 2^{m-r}.$$

Такой код называется РМ-кодом r-го порядка.

Обозначим векторы

$$\bar{a}_0 = (00000000\ldots00000000),$$
$$\bar{a}_1 = (01010101\ldots01010101),$$
$$\bar{a}_2 = (00110011\ldots00110011),$$
$$\bar{a}_3 = (00001111\ldots00001111),$$
$$\bar{a}_4 = (00000000\ldots11111111) \quad \text{и т. д.}$$

Код РМ порядка r определяется как код, базисом которого являются все векторы $\bar{a}_0, \bar{a}_1, \bar{a}_2, \ldots, \bar{a}_m$ и все векторные произведения r или меньшего числа этих векторов. Легко убедиться, что коды РМ порядка r и $m - r - 1$ двойственны, или дуальны друг другу. Так, например, РМ-код нулевого ($r = 0$) порядка есть не что иное, как код-повторение $(n, 1, n)$. Двойственный ему код $(m-1)$-го порядка есть $(n, n-1, 2)$-код с проверкой на четность. Код первого порядка $(2^m, m+1, 2^{m-1})$ есть код на основе m-последовательности, рассмотренный выше, с общей проверкой на четность. Ему двойственен код $(m-2)$-го порядка $(2^m, 2^m - m - 1, 4)$, являющийся кодом Хемминга с общей проверкой на четность. Все эти коды допускают интересные геометрические интерпретации [5, 6] и различные алгоритмы декодирования.

2.4.8. Итеративные и каскадные коды

С одной стороны, как гласит теорема Шеннона для канала с шумами (см. раздел 4.3.2), необходимо стремиться для повышения пропускной способности к большой длине кода. С другой стороны, при увеличении длины кода существенно возрастает сложность алгоритмов декодирования. Даже для кодов БЧХ она хотя и не экспоненциальная, но, тем не менее, растет от длины как некото-

рая степень полинома. Все это приводит к затруднительности реализации очень длинных кодов.

Для преодоления этой проблемы в самом начале развития теории кодирования Элайесом [9] были предложены так называемые произведения кодов. Начнем с примера. Пусть имеется код с проверкой на четность длины 3: (3,2,2). Словами кода являются все слова четного веса длины 3. Таких слов 4. Это слова: (000), (011), (101), (110). Минимальное расстояние кода равно двум. Будем рассматривать этот код как элемент построения кода-произведения или итеративного кода. Для этого будем поступать следующим образом. Запишем все информационные слова в виде матрицы размером 2×2. Каждую строчку этой матрицы закодируем кодом (3,2,2). Получаем матрицу размером 2×3. Далее, каждый столбец этой матрицы также закодируем кодом (3,2,2). Получаем матрицу размером 3×3.

Тем самым мы получили линейный алгоритм преобразования матрицы 2×2 (информационного слова длины 4) в матрицу размером 3×3 (кодовое слово длины 9), т. е. линейный (9,4) код. Такой код будем называть произведением кодов (3,2,2) или итеративным кодом, в котором первый код (строчный) (3,2,2) является внешним, а второй код (столбцовый) (3,2,2) — внутренним.

Оценим минимальный ненулевой вес кодового слова такого кода. Если кодовое слово будет содержать одну единицу на информационных позициях, то это приведет к трем единицам на проверочных:

$$\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}.$$

Если на информационных позициях две единицы, возможны две ситуации.

У этих двух единиц есть общая строка или столбец. Тогда вес кодового слова также равен четырем:

$$\begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}.$$

У этих двух единиц нет общей строки или столбца. Тогда вес кодового слова еще больше (6):

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

Также легко убедиться, что если на информационных позициях три единицы, то, как минимум, одна приведет к появлению проверочной единицы. Таким образом, в этом случае вес кодового слова будет также не менее четырех. Тем самым мы показали, что вес кодового вектора, а следовательно, и минимальное кодовое расстояние такого кода будет не менее четырех, т. е. мы построили (9,4,4) код. Существует лучший код — код Хемминга с общей проверкой на четность — (8,4,4), однако мы с самого начала не настаивали на оптимальности кодов.

Существует очень простой алгоритм декодирования такого кода. Для этого достаточно выполнить все столбцовые и строчные проверки на четность кода. Если все проверки удовлетворяются (нулевые), то принятое слово кодовое.

Оно и выдается получателю. Если по одной строчной и столбцовой проверке не удовлетворяются, то на пересечении этих проверок исправляется ошибка. Если в столбцах или строках имеется более чем по одной неудовлетворенной проверке, то обнаруживается ошибка. Легко убедиться, что такой алгоритм гарантированно исправляет одиночную ошибку и обнаруживает двойную, т. е. реализует кодовое расстояние такого кода. Очевидно, что все рассуждения остаются в силе для образующих кодов $(n, n-1, 2)$ произвольной длины. Получается результирующий $(n^2, n^2+1-2n, 4)$ код. Аналогичным образом можно строить трех- и четырехмерные коды и коды произвольной размерности. Так, в нашем примере получается (27,8,8) код. Видно, что у таких кодов все кодовые параметры — длина, число информационных символов и минимальное кодовое расстояние — произведение соответствующих параметров образующих кодов.

В общем случае имеется внешний строчный код (n_2, k_2, d_2) и внутренний строчный код (n_1, k_1, d_1). Их произведение является итеративным кодом:

$$(n_1 n_2, k_1 k_2, d_1 d_2).$$

Самый простой алгоритм декодирования такого кода состоит в попеременном построчном и постолбцовом декодировании таких кодов, однако он не реализует кодового расстояния. Для реализации кодового расстояния нужны более сложные алгоритмы.

Следует отметить еще одно очень важное достоинство таких кодов. При кодировании символы записываются в таблицу построчно, а считываются в канал столбцами. Это приводит к так называемому перемежению, если в канале возникают пакетированные ошибки, при декодировании они эффективно исправляются.

Каскадные коды были предложены Форни [10] и являются модификацией итеративных кодов. Исходные коды обозначаются так же, как и в итеративных, а вот внешний является q-м, где символ внешнего кода — информационное слово внутреннего кода, т. е. $q = 2^{k_1}$. Наиболее часто в качестве внешних используются оптимальные недвоичные коды — РС-коды. Алгоритм декодирования состоит в декодировании внутреннего столбцового кода и внешнего РС-кода, причем, как правило, декодирование внутреннего кода согласовано с каналом (носит вероятностный характер), а внешнее декодирование является алгебраическим. Результаты оценки помехоустойчивости такой каскадной пары будут приведены в конце главы при сравнении с турбокодами.

2.4.9. Мягкое декодирование, энергетический выигрыш кодирования — основные определения

Энергетический выигрыш кодирования (ЭВК) G определяет выигрыш по помехоустойчивости при применении корректирующего кода в конкретной системе. Как правило, в системах предполагают применение символов с двухпозиционной фазовой модуляцией ФМ2, использующей противоположные сигналы (см. ниже). В отличие от пропускной способности произвольного канала, задаваемой при стремящейся к нулю вероятности ошибки, ЭВК определяется для фиксированной вероятности ошибки декодирования. В стандартной ситуации ЭВК рассматривается для фиксированной вероятности ошибки в бите, а не в кодовом слове, при этом дополнительное обнаружение ошибки также не учитывает-

ся. Для получения заданного значения вероятности ошибочного приема одного символа P_b в информационной последовательности надо обеспечить на выходе демодулятора приемника некоторое необходимое минимально допустимое отношение сигнал/шум. При передаче информации с корректирующим кодированием уже вместо k информационных символов за заданное время требуется передача n символов с добавлением проверочных за то же время при том же уровне мощности сигналов. При этом придется сокращать длительность символов при передаче в n/k раз, что потребует расширения полосы частот в n/k раз. Исходное заданное значение вероятности ошибки в канале p будет обеспечиваться уже при другом отношении сигнал/шум. Разница отношений сигнал/шум при применении корректирующего кода и без него при ее положительном значении определяет ЭВК, выражаемый в децибелах. Обычно ЭВК растет с уменьшением требуемой вероятности ошибки P_b, причем существует практически в любой системе граничное отношение сигнал/шум, начиная с которого корректирующий код дает выигрыш, т. е. ЭВК положителен.

Быстрая ориентировочная оценка энергетической эффективности кодирования для целей оперативного сравнения кодов производится по асимптотическому энергетическому выигрышу от кодирования (АЭВК) $G_A = 10 \lg R d_{\min}$ (дБ), где $R = k/n$ — относительная скорость кода; $d_{\min}$ — минимальное кодовое расстояние. Величина АЭВК G_A характеризует ЭВК при вероятности $P_b \to 0$ и является верхней границей реального ЭВК. Выигрыш от кодирования может быть использован наиболее эффективным способом, например путем уменьшения мощности передатчиков в системах связи, уменьшения размеров антенн или увеличения скорости передачи.

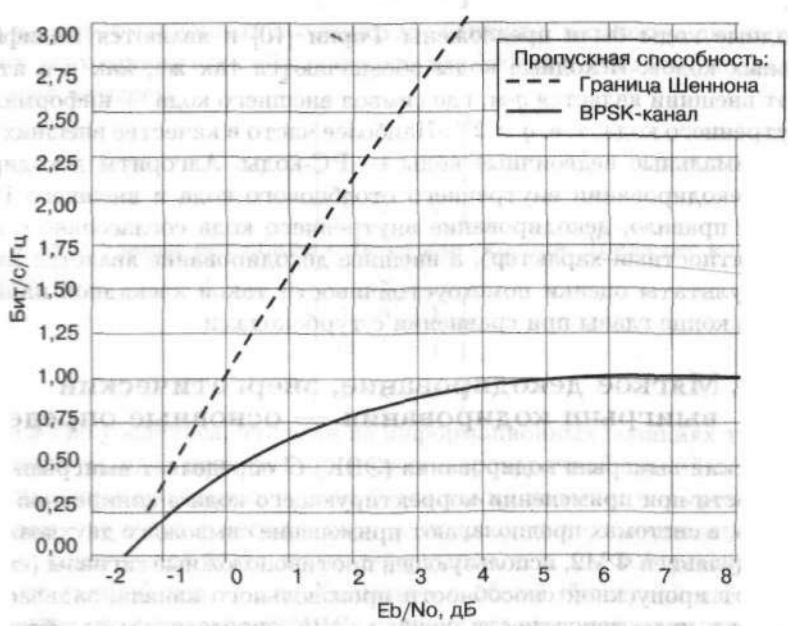


Рис. 2.6. *Сравнение пропускной способности каналов: произвольного с аддитивным гауссовым шумом* (Shannon Bound) и с двоичной модуляцией и корректирующим кодом в зависимости от отношения сигнал/шум на бит (BPSK Capacity)

Следует отметить, что такое определение ЭВК не учитывает частотной эффективности системы — наиболее эффективной может оказаться система с очень большим расширением полосы сигнала. Кроме этого, использование двоичной модуляции и корректирующего кода с относительной скоростью R приводит к существенному ограничению пропускной способности канала связи. Так, на рис. 2.6 показано сравнение пропускной способности канала в случае произвольного канала с аддитивным гауссовым шумом (Shannon Bound) и канала с двоичной модуляцией и корректирующим кодом с относительной скоростью R (BPSK).

На рис. 2.6 в случае двоичной модуляции рассматривается так называемый полунепрерывный канал, в котором в отличие от канала с «жестким» решением не принимаются решения о принятом символе. Точнее говоря, решение о принятом символе может и приниматься, но декодеру сообщается вся информация, содержащаяся в сигнале, соответствующем принятому символу. Такой алгоритм декодирования называется мягким. Обычно декодеру сообщается квантованное напряжение с выхода демодулятора — нескольких бит квантования оказывается достаточным. Мягкое декодирование в пределе для G_A дает дополнительный энергетический выигрыш в 3 дБ.

2.4.10. Низкоплотностные коды

Низкоплотностные коды были предложены Р. Галлагером в 1960 году [11] и долгое время оставались красивой идеей, слабо реализуемой из-за вычислительной сложности. Однако в последнее время низкоплотностные коды становятся одной из достаточно модных процедур кодирования и декодирования. Ниже приводится одна из наиболее простых, но достаточно эффективных процедур реализации низкоплотностных кодов.

Метод построения проверочной матрицы

Декодирование низкоплотностных кодов осуществляется с помощью их проверочной матрицы, которая порождается следующим образом. Пусть J и $K > J$ — это некоторые целые положительные числа, а P — некоторое простое число. Обозначим через E единичную матрицу размером $P \times P$. Рассмотрим исходную проверочную матрицу H_0, которая представляет собой J рядов из K подматриц, каждая из которых получается в результате некоторой циклической перестановки столбцов единичной матрицы E. В дальнейшем совокупность строк проверочной матрицы, соответствующих некоторому ряду подматриц, будем называть полосой. Отметим, что каждый столбец любой полосы содержит ровно одну единицу. Соответственно, каждый столбец матрицы H_0 содержит J единиц, а каждая строка — K единиц.

Полученная исходная матрица H_0 задает линейный код длиной $N = P \times K$ со скоростью передачи $R \approx 1 - J/K$. Приблизительный знак равенства использован потому, что, как правило, построенная таким образом проверочная матрица имеет линейно зависимые строки. Таким образом, число информационных символов оказывается несколько большим, чем $P \times (N - J)$. При $K > P$ можно выбрать циклические перестановки столбцов единичных подматриц таким образом, что матрица H_0 будет задавать J ортогональных проверок для каждого символа кода. Будем говорить, что такая матрица обладает свойством ортогонально-

сти проверок, и будем использовать ее при порождении других матриц путем перестановки столбцов в каждой полосе исходной проверочной матрицы.

Задается некоторое целое положительное число M — количество попыток перестановок столбцов в каждой полосе, которые осуществляются последовательно для каждой полосы. При каждой очередной попытке псевдослучайно выбираются два столбца в данной полосе, которые затем меняются местами. После этого проводится проверка наличия свойства ортогональности проверок. При положительном результате этой проверки перестановка столбцов сохраняется, в противном случае — аннулируется. После проведения всех попыток перестановок во всех полосах матрицы H_0 получается некоторая псевдослучайная проверочная матрица H, обладающая свойством ортогональности проверок. Как показало проведенное моделирование низкоплотностных кодов в канале с белым аддитивным гауссовым шумом, матрица H, как правило, обеспечивает лучшие результаты, чем матрица H_0.

Алгоритм декодирования низкоплотностных кодов

Описываемый ниже алгоритм декодирования низкоплотностных кодов основывается на использовании демодулятора с мягким решением. Таким образом, из демодулятора в декодер должны передаваться два вектора: вектор жестких двоичных решений для каждого символа $(x_1, x_2, \ldots, x_N)$ и вектор надежностей символов $(E_1, E_2, \ldots, E_N)$, где значение надежности i-го символа — величины E_i пропорционально логарифмическому отношению правдоподобия символа x_i. По полученному вектору жестких решений для символов кода вычисляется синдром $(s_1, s_2, \ldots, s_r)$, где r — число строк проверочной матрицы.

Алгоритм декодирования состоит из заданного числа итераций. При каждой итерации выполняется последовательная обработка всех строк проверочной матрицы. Обработка строки заключается в коррекции надежностей входящих в данную строку символов кода в зависимости от значения соответствующей данной строке компоненты синдрома. При этом следует отметить, что в данном алгоритме надежности символов рассматриваются как положительные величины. Если при коррекции надежности некоторого символа его надежность стала отрицательной, жесткое значение символа меняется на противоположное, при этом его надежность меняет знак на положительный.

Опишем теперь детально процедуру обработки строки. Рассмотрим строку с номером j. Перед обработкой данной строки имеются текущие значения векторов жестких решений, надежностей и синдрома. После обработки строки выдаются модифицированные значения этих векторов. Обозначим через $n(j,k)$ позицию k-й единицы в j-й строке. Таким образом, $n(j,1), n(j,2), \ldots, n(j,K)$ — это позиции единиц в строке с номером j. Рассмотрим набор надежностей символов, соответствующих единицам в данной строке, $E_{n(j,1)}, E_{n(j,2)}, \ldots, E_{n(j,K)}$. Далее для простоты будем обозначать эти надежности как $E1, E2, \ldots, EK$.

Обработка каждой строки состоит из двух шагов. На первом после последовательного просмотра набора надежностей находятся позиции двух наименьших надежностей. Пусть $Ek1$ — минимальная надежность, а $Ek2$ — надежность, не превосходящая все остальные, кроме $Ek1$.

Выполнение второго шага зависит от значения компоненты синдрома, соответствующей данной строке. Если $s_j = 0$, то в позициях $k1$ и $k2$ устанавливаются значения надежностей $Ek1 + Ek2$, в остальных позициях значения надежностей

увеличиваются на $Ek1$. Если $s_j = 1$, то в позициях $k1$ и $k2$ устанавливаются значения надежностей $Ek2 - Ek1$, в остальных позициях значения надежностей уменьшаются на $Ek1$. Кроме того, в этом случае изменяется жесткое значение символа, соответствующего позиции $k1$, и производится соответствующая модификация синдрома.

После обработки всех строк проверочной матрицы проводится процедура, необходимость которой не вытекает из математической сущности алгоритма, а обусловлена ограничениями на технические возможности вычислительных устройств. Речь идет о так называемой процедуре нормировки надежностей. Дело в том, что после обработки всех строк проверочной матрицы значения надежностей заметно увеличиваются. Таким образом, после нескольких итераций эти значения могут выйти за рамки допустимых для данного вычислительного устройства. Для того чтобы избежать этого нежелательного эффекта, значения всех надежностей с помощью соответствующего сдвига помещаются в некоторый заданный диапазон.

Оценка сложности алгоритма декодирования

В качестве оценки вычислительной сложности алгоритма будет рассматриваться количество элементарных арифметических операций в алгоритме декодирования. Под элементарными операциями будем понимать сложение, вычитание и сравнение целых положительных чисел, а также сложение по модулю 2 двоичных чисел и операцию сдвига целых чисел на определенное число разрядов. Сложность алгоритма определяется как суммарное количество указанных выше операций.

Для первоначального вычисления синдрома требуется не более JN сложений по модулю 2.

Для выполнения шага 1 при обработке одной строки требуется не более двух $(K - 2)$ сравнений.

Для выполнения шага 2 при обработке одной строки требуются $K + 2$ операции сложения или вычитания, а также не более J сложений по модулю 2 в случае корректировки синдрома.

Для проведения процедуры нормировки надежностей достаточно N операций сравнения и $N + J$ операций сдвига.

Обозначим через $L(N)$ вычислительную сложность одной итерации для последовательного вычислительного устройства (без предварительного вычисления синдрома). Легко видеть, что

$$L(N) \leqslant (2(K-2) + K + 2 + J) + 2N + J.$$

Учитывая, что $r = NJ/K$ и $J/K \approx 1 - R$, получаем

$$L(N) \leqslant NJ(4 - R).$$

Пусть Q — количество итераций алгоритма, а $L(Q, N)$ — вычислительная сложность всего алгоритма при Q итерациях для последовательного вычислительного устройства. Тогда

$$L(Q, N) \leqslant QNJ(4 - R) + NJ = QNJ(4 - R - 1/Q).$$

Обозначим через $L(Q)$ вычислительную сложность всего алгоритма на один кодовый символ при Q итерациях для последовательного вычислительного уст-

ройства. Тогда

$$L(Q) \leqslant QJ(4 - R - 1/Q).$$

Данная оценка вычислительной сложности алгоритма получена для последовательного вычислительного устройства, в котором в каждый момент времени выполняется одна элементарная операция. Далее будет оцениваться вычислительная сложность алгоритма в предположении, что имеется «параллельное» вычислительное устройство, в котором в каждый момент времени может выполняться одновременно любое количество элементарных операций. Обозначим через $L*(N), L*(Q,N)$ и $L*(Q)$ вычислительную сложность одной итерации, сложность всего алгоритма при Q итерациях и сложность на один кодовый символ при Q итерациях соответственно для параллельного вычислительного устройства.

Легко видеть, что для параллельного вычислительного устройства сложность предварительного вычисления синдрома приблизительно равна $J\log_2 N$, а сложность процедуры нормировки — $\log_2 N$.

Поскольку в любой полосе проверочной матрицы каждый столбец содержит ровно одну единицу, все строки в одной полосе могут обрабатываться одновременно. Следовательно, параллельная сложность обработки одной полосы не превышает $3K + J - 2$.

Таким образом,

$$L*(N) \leqslant J(3K + J - 2) + \log_2 N,$$
$$L*(Q,N) \leqslant J\log_2 N + QJ(3K + J - 2) + Q\log_2 N,$$
$$L*(Q) \leqslant ((J + Q)\log_2 N + QJ(3K + J - 2))/N.$$

В качестве примера рассмотрим низкоплотностный код длиной $N = 4112$, $R = 0{,}75$, $J = 4$, $K = 16$. Пусть при декодировании этого кода используется 10 итераций. Тогда с помощью приведенных формул находим, что в случае последовательного вычислительного устройства на один символ приходится 126 элементарных операций, в то время как для параллельного устройства — не более 0,6 операции на один символ.

Хотя определенное выше параллельное вычислительное устройство является практически нереализуемым, поскольку требуется одновременное выполнение неограниченного числа элементарных операций, тем не менее результат, полученный для этого устройства, имеет определенный практический смысл. Дело в том, что требуемая параллельность вычислений при реализации алгоритма декодирования ограничивается количеством строк в одной полосе проверочной матрицы, равным P (в рассмотренном выше примере $P = 257$). Таким образом, для достижения результата параллельного устройства, технически реализуемое устройство должно выполнять одновременную обработку P строк в одной полосе проверочной матрицы.

Оценим теперь объем оперативной памяти, требуемый для реализации алгоритма декодирования. Для хранения векторов жестких решений, надежностей и синдрома требуется N, $2N$ и $(1 - R)N$ байт соответственно (при этом не используется упаковка компонент жестких решений и синдрома). Основной объем оперативной памяти уходит на хранение двух матриц, содержащих позиции единиц в строках и столбцах проверочной матрицы соответственно, на что требуется $4NJ$ байт. Таким образом, общий объем составляет $N(4(J+1) - R)$ байт. Отсюда для рассмотренного выше кода получаем около 80 кбайт.

Результаты моделирования помехоустойчивости

На рис. 2.7–2.9 иллюстрируются результаты моделирования трех низкоплотностных кодов в гауссовом канале с двоичной фазовой модуляцией (табл. 2.13–2.15). Вероятность ошибки в бите P_b обозначена BER, а вероятность обнаружения ошибки в кодовом слове обозначена BLER.

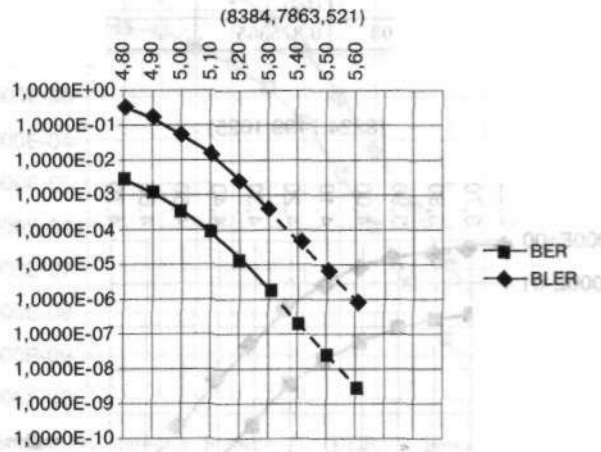


Рис. 2.7. Вероятность ошибки в бите P_b (BER) и вероятность обнаружения ошибки в кодовом слове (BLER) в зависимости от отношения сигнал/шум на бит для низкоплотностного кода в гауссовом канале с двоичной фазовой модуляцией при кодовых параметрах (8384,7863,521) (seed = 1237 shuffles = 4000)

Таблица 2.13. Кодовые параметры (8384,7863,521) (seed = 1237 shuffles = 4000)

E_b/N_0	4,8	4,90	5,00	5,10	5,20
(BER)	2,8838E-03	1,1984E-03	3,4850E-04	1,0058E-04	1,4909E-05
(BLER)	3,4247E-01	1,6949E-01	5,5866E-02	1,7379E-02	2,7813E-03

E_b/N_0	5,30	5,40	5,50	5,6
(BER)	1,7564E-06	2,0692E-07	2,4377E-08	2,8718E-09
(BLER)	3,7139E-04	4,9592E-05	6,6221E-06	8,8425E-07

N	**K**	**R**			
8384	7863	0,937858			
$L_d(Q,N)$	$L_d^*(Q,N)$	$L_c(N)$	$L_c^*(N)$	V_m, кбайт	V_{rom}, кбайт
993384	7819,151	128019,5	7863	159,817	512,0779
$[L_d(Q)]$	$[L_d^*(Q)]$	$[L_c]$	$[L_c^*]$		
118,4857	0,932628	15,2695	0,937858		

Таблица 2.14. Кодовые параметры (8224,7199,1025) (seed = 1237 shuffles = 8000)

E_b/N_0	3,60	3,70	3,80	3,90	4,00	4,10
(BER)	2,4071E-02	2,0366E-02	1,5011E-02	8,8418E-03	4,6465E-03	1,8007E-03
(BLER)	9,9010E-01	9,4340E-01	8,1301E-01	5,5866E-01	3,4364E-01	1,5152E-01

E_b/N_0	4,20	4,30	4,40	4,50	4,60	4,7
(BER)	4,4340E-04	5,0334E-05	6,3057E-06	4,5647E-07	3,3044E-08	2,3920E-09
(BLER)	3,9185E-02	5,0103E-03	7,2245E-04	5,4000E-05	4,0363E-06	3,0169E-07

Таблица 2.14 (окончание)

N	K	R			
8224	7199	0,875365			
$L_d(Q,N)$	$L_d^*(Q,N)$	$L_c(N)$	$L_c^*(N)$	V_m, кбайт	V_{rom}, кбайт
994984	3946,05	230593	7199	157,281	922,3719
$[L_d(Q)]$	$[L_d^*(Q)]$	$[L_c]$	$[L_c^*]$		
120,9854	0,479821	28,03903	0,875365		

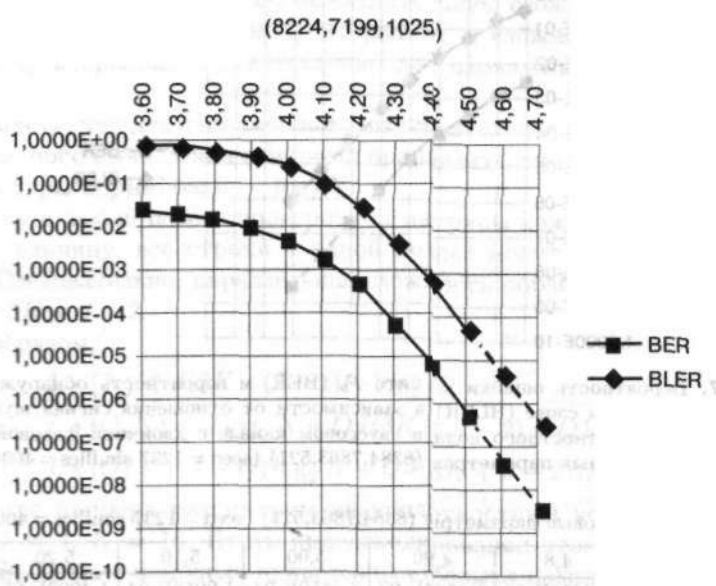


Рис. 2.8. Вероятность ошибки в бите P_b (BER) и вероятность обнаружения ошибки в кодовом слове (BLER) в зависимости от отношения сигнал/шум на бит для низкоплотностного кода в гауссовом канале с двоичной фазовой модуляцией при кодовых параметрах (8224,7199,1025)(seed = 1237 shuffles = 8000)

Таблица 2.15. Кодовые параметры (8336,6255,208[illegible]) (seed = 1237 shuffles = 4000)

E_b/N_0	2,8	2,90	3,00	3,10	3,20	3,30
(BER)	4,5229E-02	3,8264E-02	2,8797E-02	1,[illegible]087E-02	6,1945E-03	1,6505E-03
(BLER)	9,9010E-01	9,9010E-01	9,1743E-01	5,[illegible]945E-01	2,5907E-01	7,9808E-02

E_b/N_0	3,40	3,50	3,60	3,70	3,80	3,9
(BER)	2,4494E-04	1,8341E-05	9,6529E-07	5,[illegible]803E-08	2,6738E-09	1,4072E-10
(BLER)	1,2903E-02	1,0507E-03	5,7460E-05	3,[illegible]423E-06	1,7185E-07	9,3978E-09

N	K	R			
8336	6255	0,75036			
$L_d(Q,N)$	$L_d^*(Q,N)$	$L_c(N)$	$L_c^*(N)$	V_m, кбайт	V_{rom}, кбайт
1050216	2017,661	406770,5	6255	160,465	1627,082
$[L_d(Q)]$	$[L_d^*(Q)]$	$[L_c]$	$[L_c^*]$		
125,9856	0,242042	48,79684	0,75036		

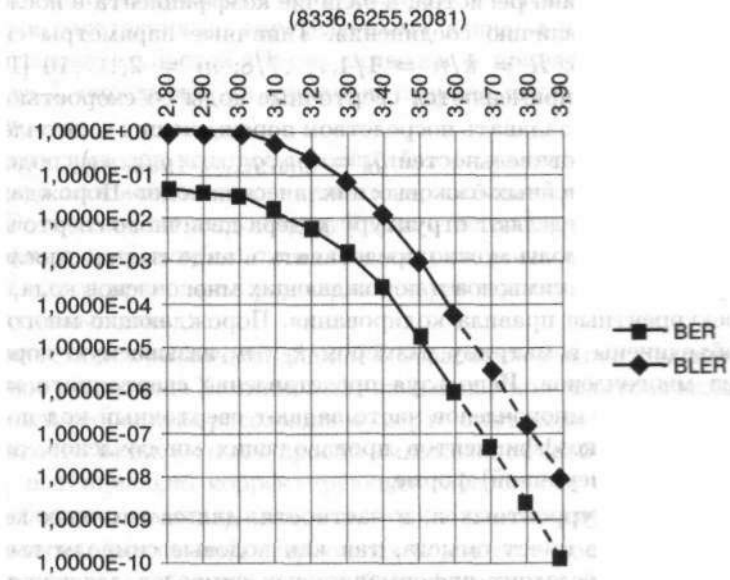


Рис. 2.9. Вероятность ошибки в бите P_b (BER) и вероятность обнаружения ошибки в кодовом слове (BLER) в зависимости от отношения сигнал/шум на бит для низкоплотностного кода в гауссовом канале с двоичной фазовой модуляцией при кодовых параметрах (8336,6255,2081)(seed = 1237 shuffles = 4000)

2.4.11. Сверточные коды

Первый непрерывный рекуррентный код был предложен в 1955 году советским ученым Л. М. Финком [12]. Однако спустя 4 года (в 1959 году) «вновь открытый» рекуррентный код был назван по имени его западного автора кодом Хегельбергера [6]. Если a_k — информационные символы, а b_k — проверочные, то кодовая последовательность имеет вид $a_1, b_1, a_2, b_2, \ldots, a_k, b_k, a_{k+1}, b_{k+1}, \ldots$ Информационные символы определяются передаваемым сообщением, а проверочные формируются по правилу $b_k = a_{k-s} \oplus a_{k+s+1} \pmod 2$, где s — произвольное целое число, называемое шагом кода ($s = 0, 1, 2$). Очевидно, что при ошибочном приеме некоторого проверочного символа b_i кодовое соотношение в принятой последовательности не будет выполнено для $i = k$. В случае же ошибочного приема информационного символа a_i кодовое соотношение не будет выполняться при двух значениях k. В принятой кодовой последовательности для каждого b_k проверяется кодовое соотношение. Если оно оказалось невыполненным при двух значениях k, разделенных величиной $2s + 1$, то информативный элемент a_{k+s+1} должен быть заменен на противоположный. Очевидно, что избыточность такого кода равна $1/2$. Так, если $s = 0$, обеспечивается правильное декодирование, когда между двумя ошибочно принятыми символами имеется не менее трех правильно принятых символов.

Структура сверточного двоичного кодера в общем виде имеет k регистров, куда за один кодовый такт входят k символов, и n сумматоров, связанных с определенными разрядами регистров. Связи j-го сумматора по (mod 2) описываются путем задания j-й порождающей последовательности $\bar{g}_i = (g_{i0}, g_{i1}, \ldots, g_{i(m-1)})$,

$i = 1, \ldots, k$, где m — длина регистра, а наличие коэффициента в последовательности соответствует наличию соединения. Типичные параметры сверточных кодов: $k, n = 1, 2, \ldots, 8$; $R = k/n = 1/4, \ldots, 7/8$; $m = 2, \ldots, 10$ [13]. Наиболее часто на практике применяются сверточные коды со скоростью $R = 1/2$. Сверточный код удобно задавать посредством порождающих многочленов, определяемых видом последовательностей $\bar{g}_i = (g_{i0}, g_{i1}, \ldots, g_{i(m-1)})$, подобно тому, как это делается для линейных блоковых циклических кодов. Порождающие многочлены полностью определяют структуру кодера двоичного сверточного кода. Выходные кодовые символы можно представить в виде свертки последовательности информационных символов и порождающих многочленов кода, задающих линейные рекуррентные правила кодирования. Порождающие многочлены могут быть объединены в матрицу размером $k \times n$, называемую порождающей матрицей из многочленов. Используя представление сверточного кода, с помощью порождающих многочленов часто задают сверточный код посредством последовательностей коэффициентов производящих многочленов, записанных в двоичной (или восьмеричной) форме.

Очевидно, для рекуррентных и, в частности, для сверточных кодов понятие кодового слова не имеет смысла, так как кодовые символы вычисляются по текущему блоку последних информационных символов для каждого такта работы кодера. Поэтому подобные коды называют также цепными или скользящими. Вообще говоря, сверточные коды можно рассматривать как обобщение блочных кодов, так как за каждый такт работы кодера по текущему блоку последних информационных символов формируется определенный блок выходных кодовых символов. Однако блоки выходных кодовых символов сверточного кода, формируемые в следующие друг за другом такты работы кодера, являются функционально зависимыми, тогда как при блоковом кодировании зависимость между соседними блоками кодовых символов (кодовыми словами) отсутствует.

Сверточный кодер как конечный автомат с памятью описывают диаграммой состояний. Диаграмма состояний представляет собой направленный граф, вершины которого отождествляются с возможными состояниями кодера, а ребра, помеченные стрелками, указывают возможные переходы между состояниями. Над каждым из ребер записывают кодовые символы, порождаемые кодером при соответствующем переходе из состояния в состояние.

Определенная последовательность кодируемых информационных символов задает конкретную последовательность смены состояний, и при этом порождаются кодовые символы, записанные над соответствующими ребрами, соединяющими состояния кодера на диаграмме состояний.

Рассматриваемую диаграмму состояний можно развернуть во времени, при этом получается так называемая решетчатая диаграмма. На ней принято, что штриховые линии (ветви) соответствуют переходам, происходящим при приходе информационного символа 1, а сплошные линии (ветви) — информационного символа 0. Из решетчатой диаграммы видно, что ее структура после окончания «переходного процесса» в кодере становится повторяющейся. Важное значение решетчатого представления состоит в том, что с ростом числа входных символов число вершин в решетке не растет, а остается равным 2^{m-1}, где m — число ячеек в регистре сдвига. Решетчатая диаграмма показывает все разрешенные пути, по которым может продвигаться кодер при кодировании.

Сверточные коды являются непрерывными и характеризуются многими минимальными расстояниями, определяемыми длинами начальных сегментов кодовых последовательностей. Число символов в принятой для обработки длине сегмента L определяет на приемной стороне число ячеек в декодирующем устройстве.

Это число символов, которое декодер должен хранить в памяти для обработки принимаемой кодовой последовательности, называется шириной окна декодирования. Если ставится цель обнаружения и исправления как можно большего числа конфигураций ошибок, то в общем случае увеличение ширины окна декодирования всегда приводит к улучшению характеристик, однако в конце концов происходит насыщение.

В соответствии с различной длиной L обрабатываемых в декодере сегментов минимальное расстояние Хемминга для любых пар кодовых слов называется L — минимальным свободным расстоянием сверточного кода и обозначается d_L. Если L достаточно велико, то это просто минимальное свободное расстояние $d_{\min}$. Очевидно, потенциально корректирующая способность сверточного кода тем выше, чем больше его минимальное свободное расстояние.

Практическая реализация сверточных кодов со скоростями $R = k/n$ встречает затруднения, особенно в случае больших скоростей передачи. Упрощение алгоритма обработки может быть получено при выборе кода с $R = 1/n$ и «выкалывании», или удалении некоторых символов в выходной последовательности для получения кода с $R = k/n$. Такие коды называются перфорированными.

В настоящее время используется три основных метода декодирования сверточных кодов: пороговое [14], аналогичное мажоритарному методу декодирования блоковых кодов, последовательное [15, 16] и декодирование по алгоритму Витерби [17]. Наиболее простыми в реализации являются алгоритмы мажоритарного декодирования как блоковых, так и сверточных кодов. Сложность реализации декодеров растет практически пропорционально полной длине кодового ограничения. Декодеры достаточно просты при исправлении ошибок невысокой кратности. Однако дальнейшее увеличение кратности исправляемых ошибок приводит к значительному усложнению схемного построения декодеров, которое не оправдывается возрастанием величины ЭВК. Наибольшую сложность имеют декодеры Витерби, объем вычислений (сложность) которых возрастает экспоненциально с ростом длины кодового ограничения. При использовании алгоритма Витерби увеличение длины кодового ограничения на единицу увеличивает объем декодера более чем вдвое, но дает прирост ЭВК, равный 0,4–0,5 дБ [18]. Метод последовательного декодирования — это метод вероятностного декодирования, при котором число операций, необходимых для декодирования одного символа, является случайной величиной.

При пороговом декодировании вычисляются синдромы, затем эти синдромы или последовательности, полученные посредством линейного преобразования синдромов, подаются на входы порогового элемента, где путем «голосования» (мажоритарный метод) и сравнения его результатов с порогом выносится решение о значении декодируемого символа. Основное достоинство этого метода декодирования — простота реализации. Однако он не полностью реализует потенциальные корректирующие способности сверточного кода. Кроме того, не все сверточные коды могут быть декодированы этим методом. Чтобы сверточный

код допускал декодирование пороговым методом, он должен обладать свойством ортогональности.

При последовательном декодировании число операций, которое должен выполнить декодер, для того чтобы декодировать один символ, изменяется в зависимости от уровня шумов в канале. Число операций при последовательном декодировании является функцией скорости передачи и шумов в канале. При всех скоростях передачи, меньших определенной скорости, число операций при декодировании оказывается небольшим. Последовательный декодер строится по схеме, позволяющей проводить вычисления со средней скоростью, в несколько раз большей скорости передачи символов, и включает в свой состав буферное запоминающее устройство, предназначенное для хранения поступающих данных при повышении уровня шумов в канале. В случае, если число возникших в канале ошибок превысит корректирующую способность кода или переполнится буфер, возникают ошибки декодирования.

Идея алгоритма Витерби состоит в том, что в декодере воспроизводят все возможные пути последовательных изменений состояния сигнала, сопоставляя получаемые при этом кодовые символы с принятыми аналогами по каналу связи, и на основе анализа ошибок между принятыми и требуемыми символами определяют оптимальный путь. Декодирование по методу Витерби, по существу, представляет собой алгоритм поиска наилучшего, максимально правдоподобного пути на графе — решеточной диаграмме кода. При декодировании с мягким решением в качестве оптимального пути выбирают путь с максимальной суммарной метрикой, что соответствует максимальной накопленной доверительной вероятности.

Важным достоинством декодера Витерби является то, что, когда в результате воздействия шумов в канале связи или по другим причинам при декодировании сделана ошибка в выборе пути на решетчатой диаграмме кода, т. е. выбран неправильный путь, за несколько тактов, в течение которых могут происходить ошибки при декодировании, декодер вновь выходит на правильный путь

Эффективность сверточных кодов зависит от выбранного кода и характеристики канала передачи. Итоговую оценку производят по величине коэффициента ошибок при заданном отношении сигнал/шум в канале и типе канала. Поскольку сверточные коды хорошо исправляют одиночные ошибки, но чувствительны к пакетам ошибок, то их обычно применяют вместе с перемежением передаваемых по каналу символов или в качестве внутренних в каскадном коде.

Анализ эффективности сверточных кодов показывает [18], что применение коротких сверточных кодов, декодируемых по алгоритму Витерби с мягким решением, позволяет получить ЭВК порядка 4–6 дБ. Переход к жесткому решению снижает ЭВК примерно на 2 дБ. Квантование выхода демодулятора на четыре уровня снижает ЭВК на 0,7–0,8 дБ, а квантование на восемь уровней — на 0,25 дБ. Обычно ограничиваются квантованием на восемь уровней, используя практически полностью возможности мягкого решения. Также следует отметить, что ЭВК в случае сверточных кодов растет достаточно медленно с уменьшением вероятности ошибки декодирования P_b, что также оправдывает их применение в каскадной паре в качестве внутренних кодов.

2.4.12. Турбокоды

Хотя турбокоды и были введены французским ученым Клодом Берру (Claude Berrou) сравнительно недавно — в 1993 году [19] и, следовательно, представляют собой новый тип кодов для исправления ошибок, они содержат «давно забытые» идеи начального этапа теории кодирования. Первая из этих идей — коды-произведения или итеративные коды, которыми фактически и являются турбокоды. Вторая идея — многоразовое декодирование, когда результат предыдущего декодирования используется на последующих шагах. Вторая идея использовалась, например, Р. Галлагером при декодировании низкоплотностных кодов. Однако эти старые идеи в турбокодах обрели «второе дыхание». Уже в первой работе по турбокодам [19] была практически продемонстрирована возможность получения значения вероятности ошибки в бите $P_b = 10^{-5}$ для ФМ2 в канале с аддитивным белым гауссовым шумом при отношении сигнал/шум на бит всего 0,7 дБ, что лишь на 0,5 дБ больше теоретического предела.

Турбокоды есть и сверточные, и блоковые. Исторически раньше появились сверточные турбокоды, а позже блоковые. Сверточные турбокоды удобнее для малых относительных скоростей кода, а блоковые — для больших.

Сверточные турбокоды

Принцип построения кодера турбокода достаточно прост [20] и показан на рис. 2.10. Из структуры кодера видно, что турбокод представляет собой систематический непрерывный код, в котором проверочные символы генерируются двумя кодерами составных рекурсивных сверточных кодов (РСК), причем информационная последовательность подается в кодер первого РСК (РСК1) непосредственно, а в кодер второго РСК (РСК2) — через устройство псевдослучайного перемежения. Схема «выкалывания» проверочных символов применяется для регулирования общей скорости турбокода.

Рис. 2.10. Структура кодера турбокода

Если при подаче определенной информационной последовательности на вход кодера РСК1 вес его проверочной последовательности оказывается малым, то перемеженная версия этой информационной последовательности, подаваемая на

вход кодера РСК2, с высокой вероятностью приведет к генерации проверочной последовательности большого веса. Если какая-либо комбинация ошибок не может быть исправлена одним РСК, то это почти наверняка будет сделано с помощью другого РСК, и наоборот. Заметьте, что при использовании в составе турбокода нерекурсивной формы сверточных кодов с такой же корректирующей способностью выигрыш от кодирования оказывается намного меньше. Это происходит потому, что вес выходной последовательности сверточных кодов в нерекурсивной форме слабо зависит от вида входной информационной последовательности.

Число составных кодов в турбокоде может быть и больше двух.

Для декодирования турбокодов в настоящее время повсеместно применяется концепция так называемого итеративного декодирования, сущность которой можно раскрыть, рассматривая структуру итеративного декодера турбокода (рис. 2.11).

Рис. 2.11. Структура итеративного декодера турбокода

Итеративный декодер образован последовательным соединением декодеров двух элементарных кодов (РСК1 и РСК2), так называемых декодеров с мягкими входным и выходным сигналами (Soft-In Soft-Out — SISO).

В процессе декодирования турбокода элементарные SISO-декодеры обмениваются друг с другом внешней информацией, с каждой итерацией улучшая окончательное решение в смысле снижения вероятности ошибки в бите в декодированной информационной последовательности (одна итерация включает в себя декодирование РСК1 и РСК2). Однако уже после первой итерации внешняя информация, подаваемая на вход декодера РСК1 по цепи обратной связи, оказывается коррелированной с информацией, получаемой из мягких решений демодулятора для проверочных символов РСК1. Поэтому улучшение окончательного решения с каждой итерацией становится меньше и, таким образом, величина вероятности ошибки на бит, достигаемая декодированием по этому методу, стремится к определенному пределу. Окончательное (жесткое) решение о передаваемых информационных битах принимается после завершения последней итерации декодером РСК2 и подается на его отдельный выход. Методы построения SISO элементарных декодеров практически сводятся к использованию алгоритмов декодирования элементарных кодов, способных вырабатывать мягкие выходные решения о передаваемых информационных символах.

В качестве примера рассмотрим характеристики помехоустойчивости конструкции на основе ФМ2 и турбокода со следующими параметрами: $R = 1/2$, составляющие коды со скоростями 2/3 с 16 состояниями с полиномами в восьмеричном виде 33 и 31. На рис. 2.12 приведены зависимости вероятности ошибки в бите (P_b) от отношения сигнал/шум на бит E_b/N_0 для случая использования указанного выше турбокода совместно с ФМ2 в канале связи с АБГШ и независимыми ошибками (без квантования выходного сигнала демодулятора) для нескольких значений длин информационного пакета K и при двух значениях количества итераций для каждой длины пакета (цифры около каждой кривой) [20]. Из рисунка следует, что ЭВК турбокода существенно увеличивается с ростом длины информационного пакета. Так, при увеличении длины пакета в 32 раза (4096 вместо 128) ЭВК увеличивается более чем на 2 дБ при одинаковом числе итераций (для $P_b = 10^{-5}$). При больших длинах пакета увеличение числа итераций также оказывается энергетически выгодным. *Для $K = 128$ увеличение числа итераций в два раза приводит к возрастанию ЭВК в среднем только на 0,1–0,2 дБ, в то время как при $K = 4096$ это увеличение уже составляет около 0,3 дБ.*

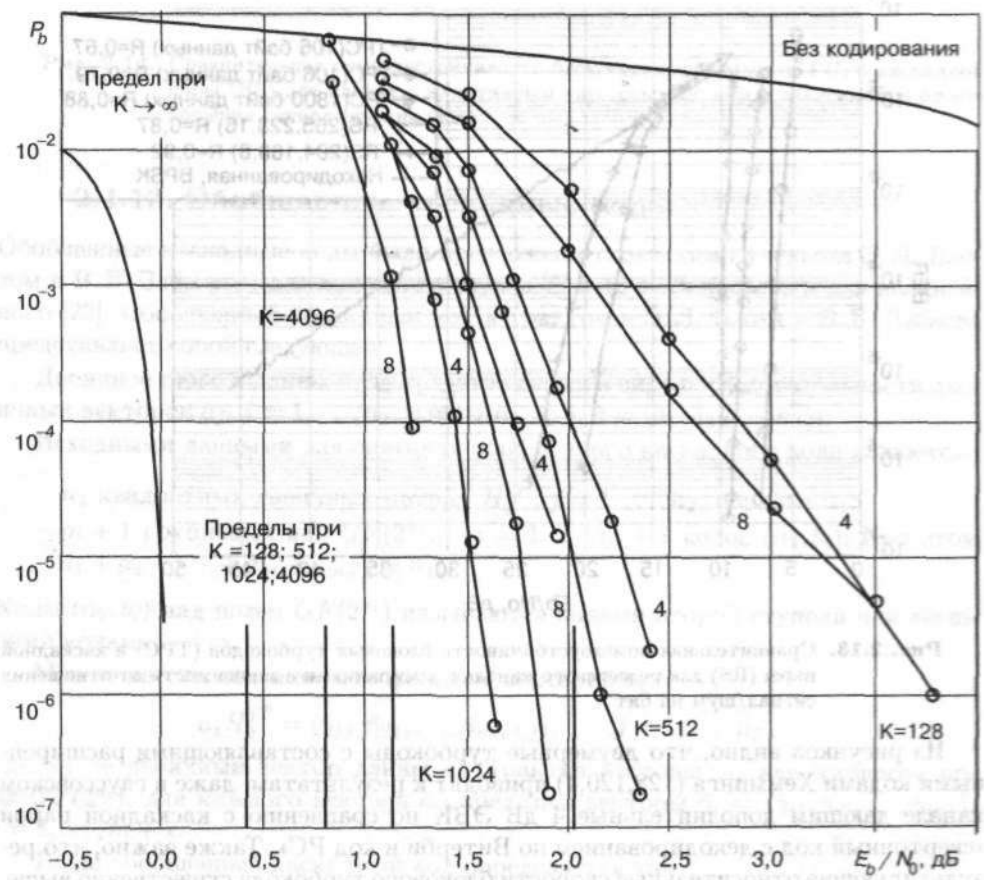


Рис. 2.12. Зависимость вероятности ошибки в бите от отношения сигнал/шум на бит (см. текст)

Блоковые турбокоды

Блоковые турбокоды являются кодами-произведениями или итеративными кодами. В качестве составляющих кодов используются либо простейшие коды с проверкой на четность, либо расширенные коды Хемминга с общей проверкой на четность. Коды могут быть не только двумерными, но и трехмерными. В некоторых случаях возможны комбинации с низкоплотностными кодами. В качестве алгоритмов декодирования составляющих кодов используются специализированные алгоритмы мягкого декодирования, выдающие метрику каждого бита на выход декодера (SISO). Декодирование, как и в случае сверточных турбокодов, осуществляется несколько раз.

На рис. 2.13 и 2.14 показаны результаты сравнительной помехоустойчивости блоковых турбокодов и каскадной пары — внутренний сверточный код и внешний код РС. Результаты взяты с сайта компании AHA (www.aha.com) [21], производителя микросхем для этих кодовых конструкций. На рис. 2.13 показаны результаты для релеевского канала с замираниями, а на рис. 2.14 — для канала с аддитивным гауссовым шумом.

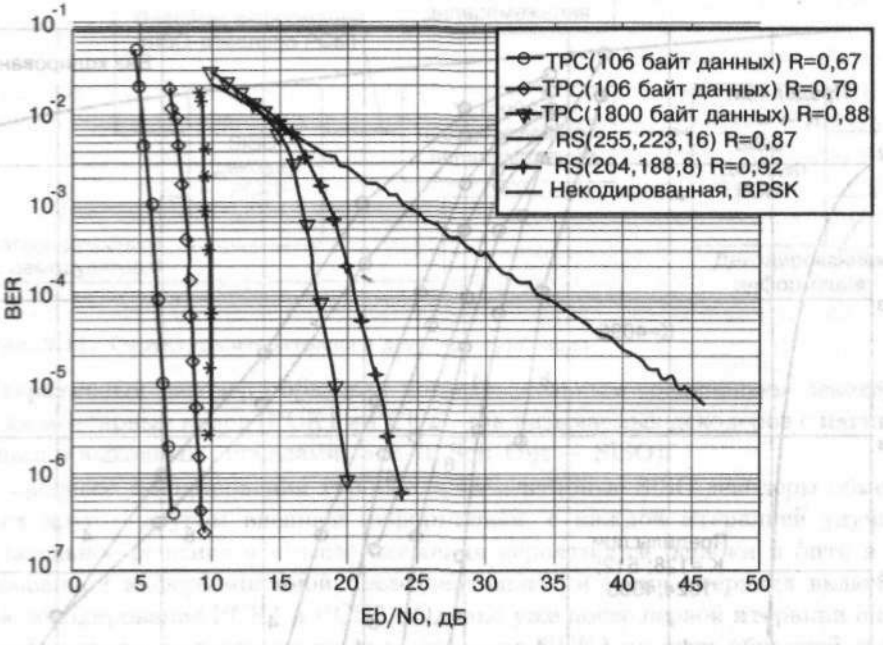


Рис. 2.13. Сравнительная помехоустойчивость блоковых турбокодов (TPC) и каскадной пары (RS) для релеевского канала с замираниями в зависимости от отношения сигнал/шум на бит

Из рисунков видно, что двумерные турбокоды с составляющими расширенными кодами Хемминга (128,120,4) приводят к результатам, даже в гауссовском канале дающим дополнительные 4 дБ ЭВК по сравнению с каскадной парой «сверточный код с декодированием по Витерби и код РС». Также важно, что результирующие относительные скорости блокового турбокода существенно выше, чем у сверточных.

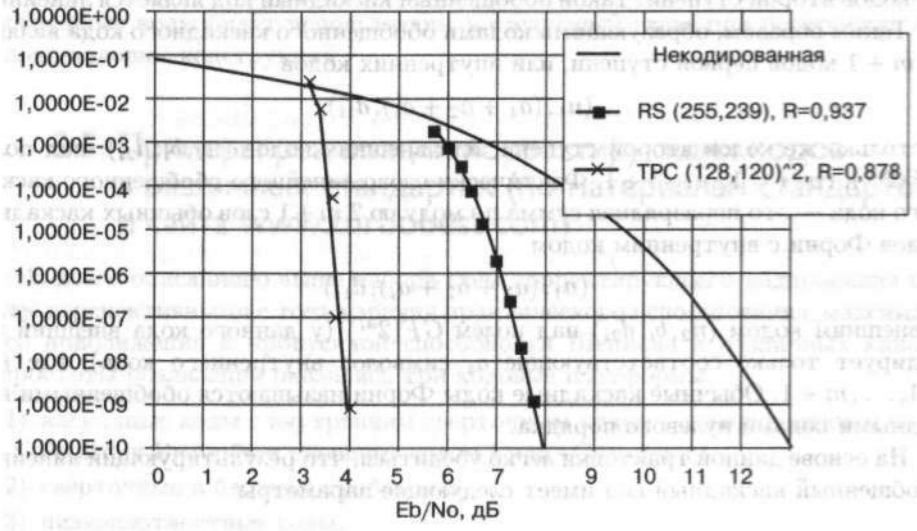


Рис. 2.14. Сравнительная помехоустойчивость блоковых турбокодов (TPC) и каскадной пары (RS) для канала с аддитивным гауссовым шумом в зависимости от отношения сигнал/шум на бит

2.4.13. Обобщенные каскадные коды

Обобщенные каскадные коды были предложены советскими учеными Э. Л. Блохом и В. В. Зябловым для линейного случая [8] и В. А. Зиновьевым для нелинейного [22]. Обобщенный каскадный код в трактовке Э. Л. Блоха и В. В. Зяблова представляет собой следующее.

Двоичное слово $\bar{a}$ длины $n_1 n_2$ представляется в виде последовательности двоичных векторов $\bar{a}_j$, $j = 1, \ldots, n_2$, длины n_1, т. е. $\bar{a} = (\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_{n_2})$.

Исходными данными для описания обобщенного каскадного кода являются:

- n_2 квадратных двоичных матриц $H_0^{(j)}$, $j = 1, \ldots, n_2$, порядка n_1;
- $m+1$ групповых над $GF(2^{a_i})$, $i = 1, \ldots, m+1$ кодов (n_2, b_i). При этом $a_1 + a_2 + \cdots + a_m + a_{m+1} = n_1$.

Коды (n_2, b_i) над полем $GF(2^{a_i})$ называются кодами второй ступени или внешними кодами.

Можно ввести линейное отображение векторов $\bar{a}_j$:

$$\bar{a}_j H_0^{jT} = (\bar{\gamma}_{1j}, \bar{\gamma}_{2j}, \ldots, \bar{\gamma}_{m+1,j}), \qquad j = 1, \ldots, n_2,$$

где $\bar{\gamma}_{ij}$ — двоичный вектор длины a_i. Трактуя векторы $\bar{\gamma}_{ij}$ как элементы поля $GF(2^{a_i})$, для каждого вектора $\bar{a}$ определим векторы $\bar{\gamma}_i = (\bar{\gamma}_{i1}, \bar{\gamma}_{i2}, \ldots, \bar{\gamma}_{in_2})$, $i = 1, \ldots, m+1$.

Тогда обобщенный каскадный код определяется так.

Двоичное слово $\bar{a}$ длины $n = n_1 n_2$ является кодовым словом обобщенного каскадного кода порядка m тогда и только тогда, когда все связанные со словом $\bar{a}$ векторы $\bar{\gamma}_i$, $i = 1, \ldots, m+1$ представляют собой кодовые слова соответствующих

i-х кодов второй ступени. Такой обобщенный каскадный код является линейным.

Таким образом, образующими кодами обобщенного каскадного кода являются $m+1$ кодов первой ступени, или внутренних кодов

$$(n_1, (a_1 + a_2 + a_i), d_{1i}),$$

и столько же кодов второй ступени, или внешних кодов (n_2, b_i, d_{2i}), над полем $GF(2^{a_i})$, где $i = 1, \ldots, m+1$. Фактически слово линейного обобщенного каскадного кода — это поразрядная сумма по модулю 2 $m+1$ слов обычных каскадных кодов Форни с внутренним кодом

$$(n_1, (a_1 + a_2 + a_i), d_{1i})$$

и внешним кодом (n_2, b_i, d_{2i}) над полем $GF(2^{a_i})$ (у данного кода внешний код кодирует только соответствующие a_i символов внутреннего кода), где $i = 1, \ldots, m+1$. Обычные каскадные коды Форни называются обобщенными каскадными кодами нулевого порядка.

На основе данной трактовки легко убедиться, что результирующий линейный обобщенный каскадный код имеет следующие параметры:

$$\left\{ \begin{array}{l} n = n_1 n_2, \\ k = \sum\limits_{i=1}^{m+1} a_i b_i, \\ d_{\min} = \min_{i=1,\ldots,m+1}(d_{1i} d_{2i}) \end{array} \right\}.$$

Существенным требованием при построении и существовании линейного обобщенного каскадного кода с данными параметрами является необходимость вложенности внутренних кодов одного в другой, т. е.

$$(n_1, a_1, d_{1,1}) \subset (n_1, (a_1 + a_2), d_{1,2}) \subset \cdots \subset (n_1, (a_1 + a_2 + \ldots + a_m), d_{1m}) \subset$$
$$\subset (n_1, (a_1 + \ldots + a_m + a_{m+1}), d_{1,m+1}).$$

Естественно, что для минимальных расстояний внутренних кодов выполняется условие $d_{1,1} \geqslant d_{1,2} \geqslant \cdots \geqslant d_{1,m} \geqslant d_{1,m+1}$. Для максимизации минимального расстояния обобщенного каскадного кода естественно выбирать минимальные расстояния внешних кодов d_{2i} таким образом, чтобы выполнялось условие $d_{\min} = d_{11} d_{21} = d_{12} d_{22} = \cdots = d_{1,m+1} d_{2,m+1}$.

Самый простой алгоритм декодирования состоит из $m+1$ шагов. На каждом шаге осуществляется попеременное декодирование i-м внутренним, затем i-м внешним кодами и коррекция исправленных символов в принятом слове. Затем происходит переход к $i-1$ шагу, $i = m+1, \ldots, 1$. Суммарная вероятность ошибки декодирования складывается из $m+1$ соответствующих вероятностей ошибок на каждом шаге. Расстояния внешних кодов d_{2i} выбираются из условия равенства составляющих вероятностей ошибки декодирования. Практика показывает, что это происходит при условии большей защищенности начальных «слоев» декодирования $d_{1,m+1} d_{2,m+1} \geqslant d_{1,m} d_{2,m} \geqslant \cdots \geqslant d_{12} d_{22} \geqslant d_{11} d_{21}$. Переход к обобщенным каскадным кодам от каскадного кода нулевого порядка приводит к дополнительному ЭВК более 1 дБ [23]. В качестве внешних в линейных обобщенных кодах обычно используются РС-коды.

Ключевым при построении обобщенных каскадных кодов m-го порядка является требование наличия системы из $(m+1)$-го вложенного внутреннего кода. При этом, если не требовать линейности от системы внутренних кодов, по-

лучается естественное обобщение на нелинейный случай. Такие нелинейные обобщенные коды будут использованы в следующей главе при построении сигнально-кодовых конструкций.

2.5. Примеры реализации корректирующих кодов в различных стандартах (по материалам стандартов и сайта www.turbobest.com)

Из всего описанного выше набора схем корректирующего кодирования наиболее перспективными с точки зрения практического использования, максимального приближения к пропускной способности Шеннона в различных каналах и простоты реализации оказались три кодовые платформы:

1) каскадные коды с внутренним сверточным кодом с декодированием по алгоритму Витерби и внешним кодом Рида – Соломона;
2) сверточные и блочные турбокоды;
3) низкоплотностные коды.

Во всех новейших стандартах беспроводного доступа применяются те или иные комбинации этих кодовых платформ.

Рассмотрим некоторые примеры.

2.5.1. Схема корректирующего кодирования и декодирования в стандарте IEEE 802.3an

Низкоплотностный кодек (LDPC) используется в сетях 10GBit Ethernet (10GBASE-T) в соответствии со стандартом IEEE 802.3an (опубликован в 2006 году). Стандарт описывает передачу по сетям медным витым линиям категории 6 и 6A (класс E и F) протяженностью от 55 м (без экрана) и до 100 м. Система должна работать с очень маленькой задержкой, поддерживая протокол IP. Полоса сигнала 500 МГц, 128-уровневая модуляция и низкоплотностный код с рабочей вероятностью ошибки 10^{-13}.

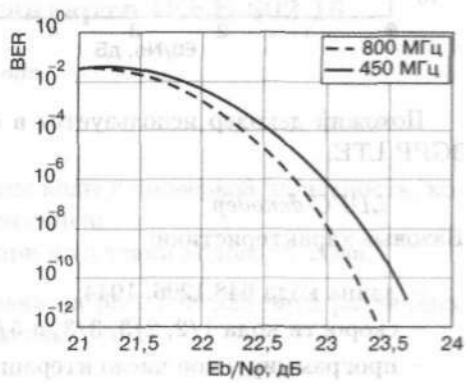


Рис. 2.15. Помехоустойчивость декодера компании Turbobest для двух значений тактовой частоты — 450 и 800 МГц с задержкой 0,32 мкс

Низкоплотностный код построен на алгебраической конструкции с очень низкой «полкой» (10^{-13}). Это (6,32)-размерный RS-LDPC низкоплотностный код на основе кода Рида – Соломона длины 2048 и размерности 1723. Проверочная матрица кода, используемая для декодирования, — это 6×32 блоковая матрица, каждым элементом которой являются перестановки определенной матрицы 64×64.

Поскольку любой алгоритм декодирования LDPC кода многошаговый или «многопопыточный», то помехоустойчивость любого декодера зависит от такто-

вой частоты обработки. На рис. 2.15 приводится помехоустойчивость реализованного компанией Turbobest декодера для двух значений тактовой частоты — 450 и 800 МГц с задержкой 0,32 мкс.

2.5.2. Схемы корректирующего кодирования и декодирования в стандарте IEEE 802.11n

Декодер Витерби сверточного кода с $K = 7$

Декодер рекомендован со скоростью 1/2, остальные скорости получаются перфорированием. Достигается пропускная способность 540 Мбит/с. Характеристики помехоустойчивости показаны на рис. 2.16.

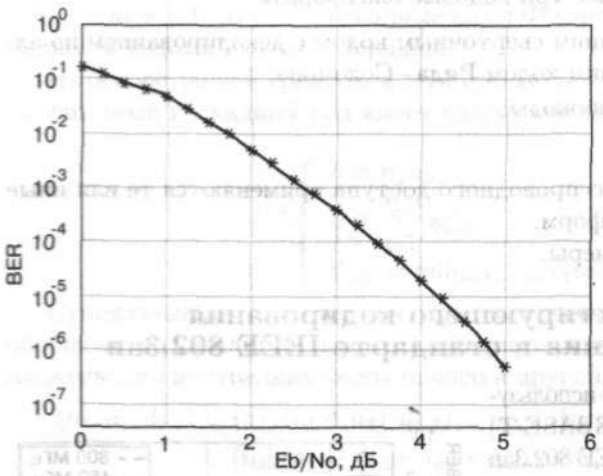


Рис. 2.16. Помехоустойчивость декодера Витерби сверточного кода с K = 7, длина пакета 2048 бит

Похожий декодер используется в стандартах IEEE 802.15, IEEE 802.16 и в 3GPP LTE.

LDPC-декодер

Базовые характеристики:

- длина кода 648,1296, 1944;
- скорости кода 1/2, 2/3, 3/3, и 5/6 для каждой длины кода;
- программируемое число итераций декодирования;
- скорость до 360 Мбит/с;
- остановка декодирования в случае получения кодового слова.

2.5.3. Схемы корректирующего кодирования и декодирования в стандарте IEEE 802.16

LDPC-кодек стандарта IEEE802.16e.

В стандарте предусмотрены четыре кодовые скорости 1/2, 2/3, 3/3 и 5/6 и 19 кодовых длин от 576 до 2304. Вообще, в стандарте также предусмотрены кодовые

схемы с кодеком Витерби и со сверточными турбокодами. Характеристики низкоплотностного кода оказываются наилучшими, и они показаны на рис. 2.17 в случае 50 итераций.

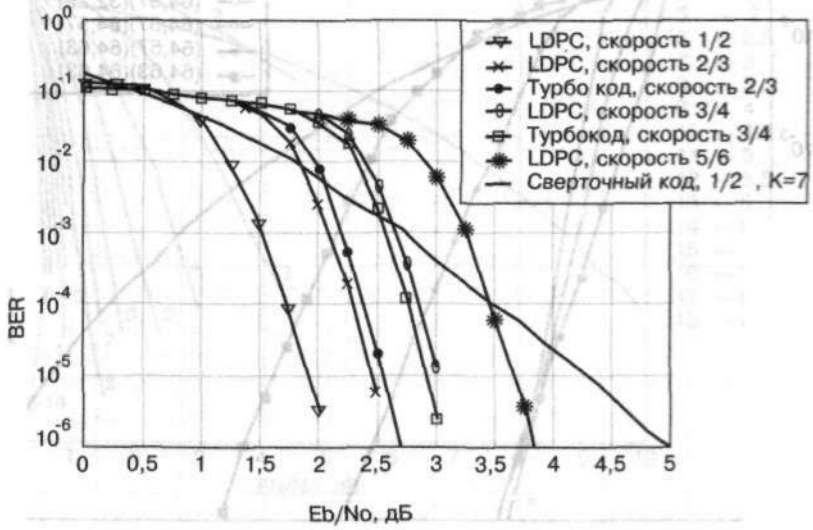


Рис. 2.17. Помехоустойчивость декодеров различных кодов для стандарта IEEE 802.16e. Длина пакета 2304, 50 итераций

2.5.4. Схема корректирующего кодирования и декодирования в стандарте IEEE 802.16

Характеристики блокового турбокода:

- длина кода от 64 бит до 4 кбит;
- 64 возможных кода произведения;
- в качестве составляющих используются коды с проверкой на четность, коды Хемминга и обобщенные коды Хемминга;
- возможна остановка декодирования при получении кодового слова.

Помехоустойчивость декодирования показана на рис. 2.18 для пяти различных кодов в сравнении с некодированной передачей.

2.5.5. Схемы корректирующего кодирования и декодирования в 3GPP LTE

В стандарте 3GPP LTE рекомендованы различные комбинации кодеков Витерби для сверточных кодов и сверточных турбокодов.

Декодер сверточного турбокода:

- $K = 4$, 8 состояний;
- базовая скорость $R = 1/3$, остальные скорости достигаются внешним перфорированием;
- блоковая длина 40–5114.

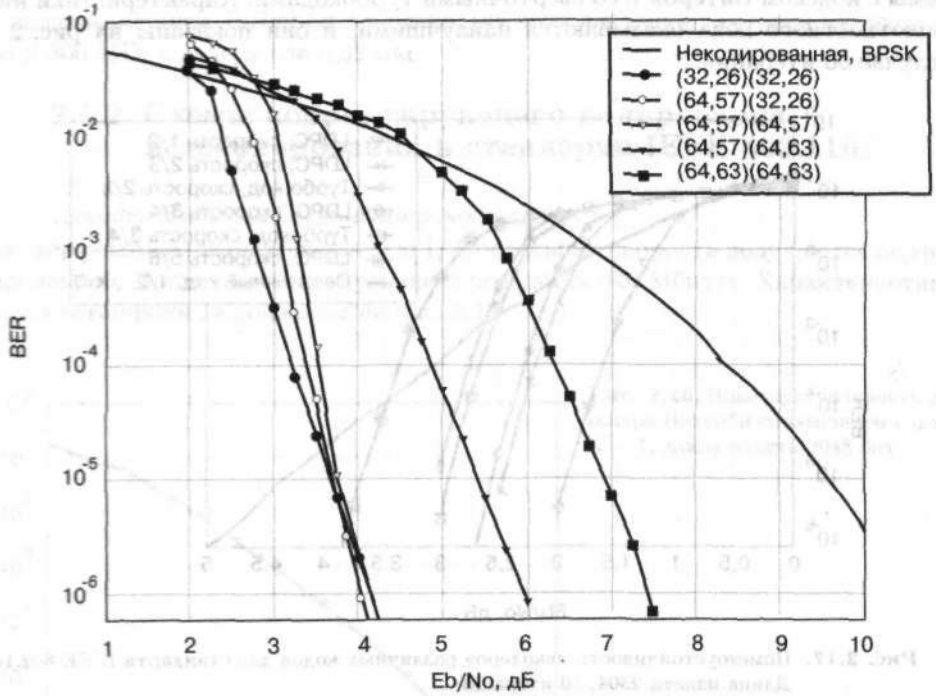


Рис. 2.18. Помехоустойчивость блокового турбокода, используемого в стандарте IEEE 802.16

На рис. 2.19 и 2.20 показаны одна из возможных реализаций декодера и его помехоустойчивость для блоковой длины 5114, относительной скорости кода $R = 1/3$ и при различном среднем на бит количестве итераций декодирования.

Декодеры Витерби

В стандарте рекомендованы различные кодеки Витерби со скоростями $R = 1/2$ и $R = 1/3$ и длинами кодового ограничения $K = 7$ и $K = 9$. На рис. 2.21 проиллюстрирована помехоустойчивость работы декодера для $R = 1/3$ и $K = 9$.

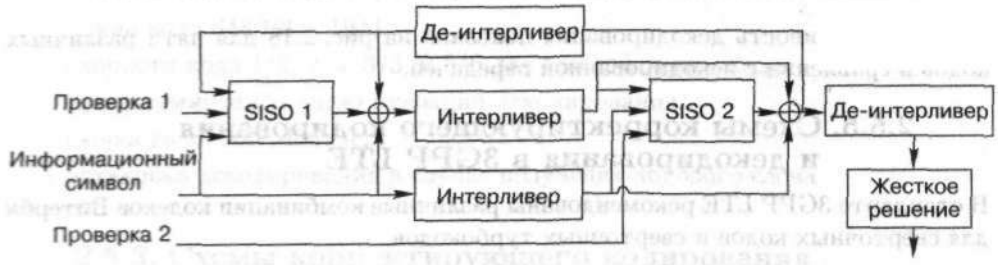


Рис. 2.19. Функциональная схема декодера сверточного турбокода

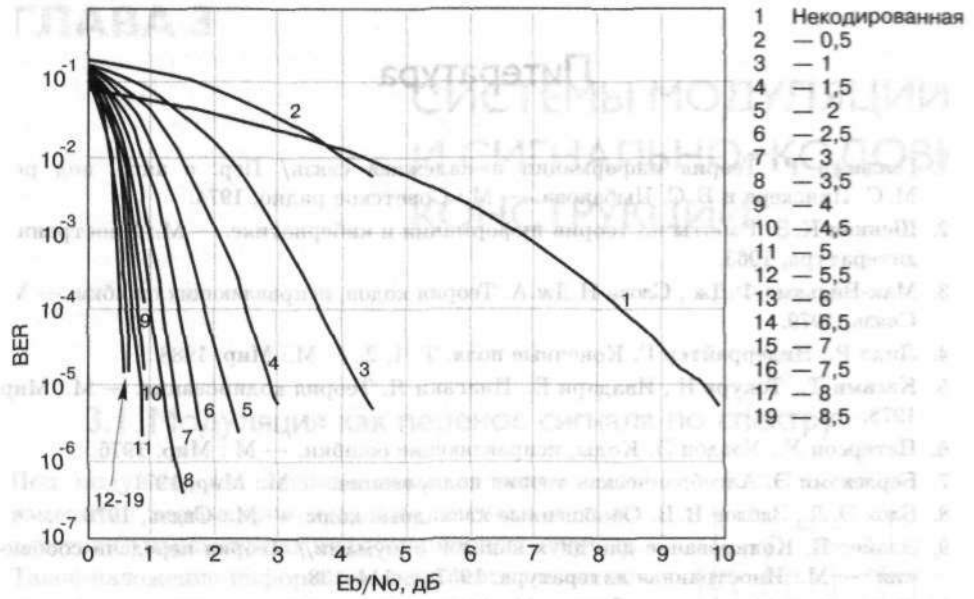


Рис. 2.20. Помехоустойчивость декодера сверточного турбокода

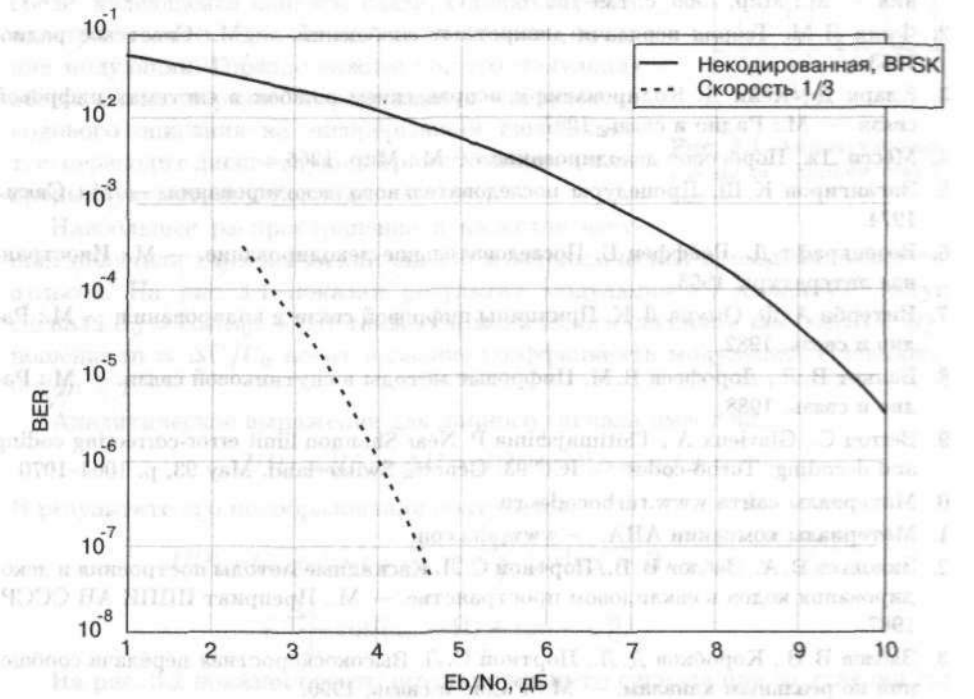


Рис. 2.21. Помехоустойчивость декодера Витерби сверточного кода с $K = 9$, $R = 1/3$

Литература

1. Галлагер Р. Теория информации и надежная связь/ Пер. с англ. под ред. М. С. Пинскера и Б. С. Цыбакова. — М.: Советское радио, 1974.
2. Шеннон К. Э. Работы по теории информации и кибернетике. — М.: Иностранная литература, 1963.
3. Мак-Вильямс Ф.Дж., Слоэн Н.Дж.А. Теория кодов, исправляющих ошибки. — М.: Связь, 1979.
4. Лидл Р., Нидеррайтер Г. Конечные поля. Т. 1, 2. — М.: Мир, 1988.
5. Касами Т., Токура Н., Ивадари Е., Инагаки Я. Теория кодирования. — М.: Мир, 1978.
6. Петерсон У., Уэлдон Э. Коды, исправляющие ошибки. — М.: Мир, 1976.
7. Берлекэмп Э. Алгебраическая теория кодирования. — М.: Мир, 1971.
8. Блох Э. Л., Зяблов В. В. Обобщенные каскадные коды. — М.: Связь, 1976.
9. Элайес П. Кодирование для двух каналов с шумами// Теория передачи сообщений. — М.: Иностранная литература, 1957, с. 114–138.
10. Форни Г. Д. Каскадные коды. — М.: Мир, 1970.
11. Галлагер Р. *Коды с малой плотностью проверок на четность*// Теория кодирования. — М.: Мир, 1966, с. 139–165.
12. Финк Л. М. Теория передачи дискретных сообщений. — М.: Советское радио, 1963.
13. Кларк Д., Кейн Д. Кодирование с исправлением ошибок в системах цифровой связи. — М.: Радио и связь, 1987.
14. Месси Дж. *Пороговое декодирование*. — М.: Мир, 1966.
15. Зигангиров К. Ш. Процедуры последовательного декодирования. — М.: Связь, 1974.
16. Возенкрафт Д., Рейффен Б. Последовательное декодирование. — М.: Иностранная литература, 1963.
17. Витерби А. Д., Омура Д. К. Принципы цифровой связи и кодирования. — М.: Радио и связь, 1982.
18. Банкет В. Л., Дорофеев В. М. Цифровые методы в спутниковой связи. — М.: Радио и связь, 1988.
19. Berrou C., Glavieux A., Thitimajshima P. Near Shannon limit error-correcting coding and decoding: Turbo-codes. — ICC'93, Geneva, Switzerland. May 93, p. 1064–1070.
20. Материалы сайта www.turbocodes.ru.
21. Материалы компании AHA. — www.aha.com.
22. Зиновьев В. А., Зяблов В. В., Портной С. Л. Каскадные методы построения и декодирования кодов в евклидовом пространстве. — М.: Препринт ИППИ АН СССР, 1987.
23. Зяблов В. В., Коробков Д. Л., Портной С. Л. Высокоскоростная передача сообщений по реальным каналам. — М.: Радио и связь, 1990.

ГЛАВА 3

СИСТЕМЫ МОДУЛЯЦИИ И СИГНАЛЬНО-КОДОВЫЕ КОНСТРУКЦИИ

3.1. Модуляция как перенос сигнала по спектру

Под модуляцией в технике связи подразумевают изменение параметров несущего сигнала в соответствии с параметрами информационного сигнала. Такое наложение информационного сигнала на несущий может быть необходимо, например, потому, что последний хуже распространяется в конкретной среде, являющейся каналом связи. Однако перенос спектра — это одна и не самая важная функция модуляции. Гораздо важнее то, что модуляция обеспечивает перенос информации с «цифрового» кодового описания на «непрерывное» сигнальное, т. е. переводит дискретную метрику Хемминга в непрерывную метрику Евклида.

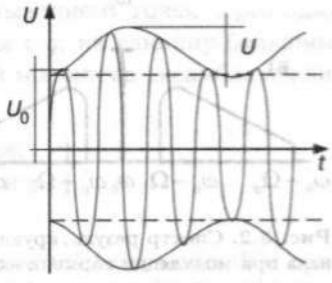


Рис. 3.1. Амплитудно-модулированный гармонический сигнал

Наибóльшее распространение в качестве несущих получили гармонический сигнал и периодическая последовательность импульсов. На рис. 3.1 показан результат модуляции по амплитуде несущего сигнала $U_0 = \cos(\omega_0 t + \varphi_0)$ также гармоническим сигналом $\Delta U \cos(\Omega t + \psi)$. Отношение $m = \Delta U / U_0$ носит название коэффициента модуляции. Очевидно, что $0 \leqslant m \leqslant 1$.

Аналитическое выражение для данного сигнала имеет вид:

$$U(t) = [U_0 + \Delta U \cos(\Omega t + \psi)] \cos(\omega_0 t + \varphi_0).$$

В результате его преобразования получается:

$$U(t) = U_0 \cos(\omega_0 t + \varphi_0) + \frac{\Delta U}{2} \cos[(\omega_0 + \Omega) + (\varphi_0 + \psi)] +$$

$$+ \frac{\Delta U}{2} \cos[(\omega_0 - \Omega) + (\varphi_0 - \psi)].$$

На рис. 3.2 показан спектр результирующего сигнала при модуляции гармоническим сигналом, периодическим сигналом сложной формы и произвольным сигналом. Из рис. 3.2 видно, что спектр модулированного сигнала имеет три составляющие — с несущей частотой ω_0 и так называемыми «боковыми» частотами, отличающимися от несущей на величину $+\Omega$ и $-\Omega$. Поскольку этот

результат распространяется на любую частотную составляющую произвольного модулирующего сигнала, можно легко определить качественный вид спектра гармоники, модулируемой периодическим сигналом сложной формы, апериодическим сигналом произвольной формы со спектром в диапазоне от $\Omega_н$ до $\Omega_в$.

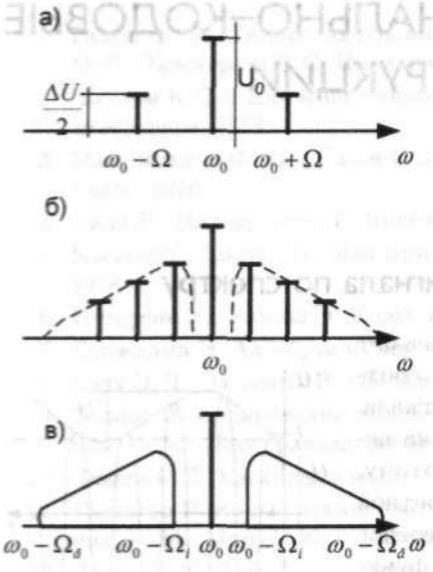


Рис. 3.2. Спектр результирующего сигнала при модуляции гармоническим сигналом (а), периодическим сигналом сложной формы (б) и произвольным сигналом (в)

Из рис. 3.2 следует важный результат — с помощью модуляции спектр модулирующего сигнала можно перенести в любую необходимую область частотного диапазона.

До сих пор речь шла лишь об амплитудной модуляции (АМ) гармонического сигнала. Наряду с амплитудой в качестве изменяемого параметра несущей можно использовать также частоту и фазу (ЧМ и ФМ). Не вдаваясь в подробности анализа спектра ЧМ- и ФМ-сигналов, отметим лишь, что такие сигналы также включают несущую и полосы боковых частот. При этом последние шире, чем при АМ.

Еще один распространенный вариант — модуляция последовательности импульсов. При амплитудно-импульсной модуляции (АИМ), частотно-импульсной (ЧИМ), фазово-импульсной (ФИМ) и широтно-импульсной модуляции (ШИМ) пропорционально амплитуде модулирующего сигнала меняются соответствующие названию параметры несущего сигнала.

В дальнейшем будем считать, что результирующий модулированный сигнал

$$U(t) = A(t)\cos(\omega_0 t + \varphi(t))$$

является узкополосным, т. е. ширина спектра сигнала существенно меньше его несущей частоты $\Delta\omega \ll \omega_0$. Тогда вся информация, имеющаяся в сигнале, заложена в его комплексной огибающей $A(t)e^{j\varphi(t)}$, а это значит, что на так называемой сигнальной плоскости можно показать все изменения сигнала. В каждый момент времени сигнал отображается на сигнальной плоскости точкой с амплитудой и фазой.

3.2. Дискретная модуляция

В случае дискретной модуляции сигнал $U(t) = A(t)\cos(\omega_0 t + \varphi(t))$ в каждый момент времени принимает одно из набора $Q = 2^q$ дискретных значений. Все множество $Q = 2^q$ сигналов называется сигнальным алфавитом. Как и корректирующий код, сигнальный алфавит характеризуется двумя основными параметрами — по скорости и по помехоустойчивости.

По скорости ансамбль сигналов характеризуется скоростью в битах на измерение канала: $R = q/2$, где q — число бит, которыми могут быть перенумерованы

сигналы. По помехоустойчивости ансамбль сигналов характеризуется квадратом нормированного по средней мощности ансамбля евклидова расстояния Δ^2.

Самый простой вид дискретной модуляции — фазовая (ФМ). Все сигналы ФМ лежат на комплексной плоскости на одной окружности через равные промежутки. Легко убедиться, что в этом случае квадрат нормированного по средней мощности ансамбля евклидова расстояния $\Delta^2 = 4\sin^2(\pi/2^q)$. Например, в случае ФМ8 $\Delta^2 = 0{,}576$. Достоинством ФМ является постоянная амплитуда сигнала, что приводит к ее использованию в каналах, где имеется нелинейное преобразование амплитуды передаваемого сигнала.

Однако фазовая модуляция обладает одним существенным недостатком — не используется сигнальная плоскость внутри круга, что приводит к уменьшению Δ^2. Исходя из этого в большинстве реализованных систем используется наиболее общий вид модуляции — амплитудно-фазовая (АФМ). Самая простая и легко реализуемая разновидность АФМ — так называемая квадратурная амплитудная модуляция (КАМ), когда сигнал располагается в узлах равномерной квадратной решетки. В случае четного q построение очевидно. В случае нечетного q ансамбль сигналов КАМ строится «выкалыванием» точек через одну построчно из ансамбля сигналов КАМ с четным q и с q, на единицу большим. В общем случае квадрат нормированного по средней мощности ансамбля евклидова расстояния есть

$$\Delta^2 = \left\{ \begin{array}{ll} 6/(2^q-1), & q\text{— четное}, \\ 6/(2^q-1/2), & q - \text{нечетное} \end{array} \right\}.$$

Так, например, в случае КАМ16 имеем $\Delta^2 = 0{,}4$.

Для большого числа сигналов ансамбля в случае фазовой модуляции ФМ приближенно имеем $\Delta^2 \approx \pi^2/2^{2(q-1)}$, а в случае КАМ $\Delta^2 \approx 3/2^{(q-1)}$. Это значит, что в случае увеличения ансамбля сигналов вдвое помехоустойчивость фазовой модуляции ФМ падает в четыре раза, т. е. на 6 дБ, а квадратурной амплитудной модуляции КАМ — в два раза, т. е. на 3 дБ.

Отдельной проблемой является нумерация сигналов ФМ или КАМ. Одним из наиболее распространенных вариантов является нумерация сигналов кодом Грея, когда ближайшие по евклидову расстоянию сигналы отличаются номерами с расстоянием Хемминга, равным 1. В случае КАМ выбирается код Грея и по горизонтали, и по вертикали. Другим видом нумерации является обычная двоичная нумерация.

3.3. Сигнально-кодовые конструкции (СКК) в гауссовом канале

До сих пор мы оценивали эффективность систем с двоичной модуляцией и кодированием при помощи энергетического выигрыша кодирования (ЭВК). Такая оценка не учитывает частотной эффективности системы связи. Пусть R [бит/с] есть скорость передачи информации конкретной системы связи, $\Delta\omega$ [Гц] есть ширина спектра сигнала, а E_b/N_0 [дБ] есть отношение энергии одного бита к спектральной плотности шума. В соответствии с [1] определим частотную эффективность системы как $\gamma = R/\Delta\omega$ [бит/с/Гц].

Эффективность системы определим как отношение скорости передачи к пропускной способности канала $\eta = R/C$, где C есть пропускная способность, определяемая по формуле Шеннона (см. рис. 2.6):

$$C = \Delta\omega \log_2(1 + P_c/P_ш),$$

где $P_c/P_ш$ — отношение сигнал/шум, вычисляемое по формуле

$$P_c/P_ш = (E_b/N_0)(R/\Delta\omega) = \gamma(E_b/N_0).$$

Тогда эффективность системы может быть определена в следующем виде:

$$\eta = R/(\Delta\omega \log_2(1 + P_c/P_ш)) = \gamma/\log_2(1 + P_c/P_ш).$$

Максимальное значение коэффициент эффективности системы принимает при $\eta_{max} = 1$. Тогда для оптимальной системы (на границе пропускной способности) получим:

$$\gamma = \log_2(1 + \gamma E_b/N_0) \text{ или } 2^\gamma = 1 + \gamma E_b/N_0.$$

Оптимальные обменные соотношения между помехоустойчивостью E_b/N_0 (энергетическая эффективность) и скоростью передачи γ (частотная эффективность), известные как граница Шеннона, окончательно представляются в следующем виде:

$$E_b/N_0 = (2^\gamma - 1)/\gamma.$$

Таким образом, можно оценивать эффективность любой реальной системы передачи как точку в осях E_b/N_0 и γ, ограниченных кривой Шеннона сверху. Следует помнить, что кривая Шеннона показывается для вероятности ошибки, стремящейся к нулю, а точка конкретной системы фиксируется для конкретной вероятности ошибки в бите. Расстояния от конкретной точки системы до кривой Шеннона по осям абсцисс и ординат есть проигрыш пропускной способности по скорости и помехоустойчивости. Также видно, что для увеличения частотной эффективности γ на 1 бит/с/Гц необходимо увеличить энергетическую эффективность на 3 дБ. Из этого следует, что КАМ с ростом числа сигналов ведет себя по отношению к пропускной способности Шеннона одинаково (не приближается и не удаляется), ФМ с ростом числа сигналов отдаляется от пропускной способности — асимптотический угол наклона в два раза больше.

Анализ эффективности систем показывает, что использование многопозиционной модуляции хотя и позволяет повысить скорость передачи по сравнению с двоичной модуляцией, но не позволяет приблизиться к пропускной способности ни по частотной, ни по энергетической эффективности. Использование же корректирующих кодов вместе с двоичной модуляцией позволяет приблизиться к пропускной способности для частотной эффективности, меньшей единицы.

Как же приблизиться к пропускной способности в области с высокой частотной эффективностью? Здесь прослеживается полная аналогия с корректирующими кодами и дискретными каналами. Совершенно понятно, что использования сигналов на двумерной сигнальной плоскости недостаточно для достижения пропускной способности. Необходимо передавать информацию многомерными сигналами, где размерность сигнала $N \to \infty$. При этом ансамбль M сигналов строится таким образом, что все сигналы являются точками в N-мерном пространстве Евклида. Сигналы необходимо выбирать таким образом, чтобы вокруг каждого сигнала можно было бы очертить непересекающиеся N-мерные сферы максимального радиуса.

Тогда, по аналогии с корректирующими кодами ансамбль таких сигналов будем обозначать $(N, K, D^2_{\min})$, где N — размерность сигнала, $K = \log_2 M$ — число бит, переносимых сигналом, $D^2_{\min}$ — нормированный по средней мощности ансамбля сигналов $P_{\text{ср}}$ квадрат минимального расстояния Евклида. Скорость такого ансамбля есть $R = K/N$ [бит/изм.].

В этих обозначениях сигналы ФМ или КАМ из 2^q точек принимают вид $(2, q, \Delta^2)$. В этих обозначениях, хотя и прослеживаются аналогии с корректирующими кодами, есть некоторые различия. В отличие от ненормированного расстояния по Хеммингу $d_{\min}$ квадрат минимального расстояния Евклида $D^2_{\min}$ нормируется по средней мощности ансамбля сигналов $P_{\text{ср}}$. Из этого следует, что при построении многомерных ансамблей сигналов надо стремиться не к росту $D^2_{\min}$, а минимизировать его падение.

Рассмотрим пример, иллюстрирующий выигрыш при увеличении размерности сигнала.

Пусть имеется сигнал ФМ4 с параметрами (2,2,2). Скорость составляет $R = 1$ бит/изм. Все четыре сигнала ФМ4 на сигнальной плоскости отображаются точками на круге через равные промежутки:

$$\bar{x}_2 = (x_1, x_2) = \{(+1,0), (0,+1), (-1,0), (0,-1)\}.$$

Разложим сигнал ФМ4 на два подмножества ФМ2 —

$$\{(+1,0), (-1,0)\} \quad \text{и} \quad \{(0,+1), (0,-1)\}.$$

Будем формировать четырехмерный сигнал $\bar{x}_4 = (x_1, x_2, x_3, x_4)$ как последовательный набор двух сигналов ФМ4, причем первый сигнал выбирается произвольно, а второй берется из того же самого подмножества, что и первый. Всего возможно восемь четырехмерных сигналов следующего вида:

$$\bar{x}_4 = \left\{ \begin{array}{l} (+1,0,+1,0), (+1,0,-1,0), (0,+1,0,+1), (0,+1,0,-1), \\ (-1,0,+1,0), (-1,0,-1,0), (0,-1,0,+1), (0,-1,0,-1) \end{array} \right\}.$$

Видно, что реальный квадрат минимального расстояния Евклида для этого ансамбля по сравнению с ФМ4 вырос в два раза, а нормированный по мощности четырехмерного сигнала остался без изменения. В наших обозначениях параметры этого ансамбля есть (4,3,2), $R = 0{,}75$ бит/изм. Если сравнивать этот ансамбль с ФМ2, то при том же расстоянии Евклида скорость в полтора раза больше. Точки этого ансамбля выбраны в четырехмерном пространстве, что и обеспечивает этот выигрыш.

Таким образом, мы убедились, что задача построения оптимального ансамбля сигналов сводится к задаче оптимальной упаковки шаров в N-мерном пространстве Евклида и нормировании потом по средней мощности сигнала. Последняя задача, хотя и является традиционной математической задачей, мало подходит для данного случая из-за отсутствия неперебoрных алгоритмов декодирования таких конструкций. Исходя из наличия хорошо развитой теории корректирующих кодов, было предложено решать вопросы оптимальных упаковок в пространстве Евклида размерности N при помощи сигнально-кодовых конструкций (СКК). СКК фактически является каскадным или обобщенным каскадным кодом, где в качестве внутреннего кода используется простой ансамбль двумерных сигналов, а в качестве внешнего — корректирующий код [2–5]. Возможно использование как блоковых, так и сверточных корректирующих кодов.

3.4. Описание блоковых СКК в гауссовом канале

Как сказано выше, блоковая СКК является нелинейным каскадным кодом с внутренней системой сигналов и внешним корректирующим кодом. Ввиду различия метрик системы сигналов и корректирующих кодов возможно несколько способов согласования метрик и отсюда — несколько СКК [3].

Конструкция 1 состоит из одного внутреннего ансамбля сигналов и одного внешнего корректирующего кода. Внутренний ансамбль 2^q двумерных сигналов (например, ФМ или КАМ) имеет параметры $(2, q, \Delta^2)$. Внешним является корректирующий код $(N/2, k, d_{\min})$ с элементами из поля $GF(2^q)$. Тогда СКК имеет следующие параметры: N — размерность, $K = q \cdot k$ — число информационных символов, $D^2_{\min} = 2\Delta^2 d_{\min}/N$ — квадрат минимального расстояния Евклида, нормированный по средней мощности.

В данной конструкции нумерация сигналов в ансамбле модуляции не важна, так как q-й корректирующий код не использует тот факт, что разные сигналы находятся друг от друга на разных расстояниях. Прием СКК состоит в последовательном выполнении демодуляции ансамбля сигналов и декодировании корректирующим кодом.

Конструкция 2 также состоит из одного внутреннего ансамбля сигналов и одного внешнего корректирующего кода. Внутренний ансамбль сигналов такой же, как в конструкции 1, а внешний корректирующий код является двоичным с параметрами $(qN/2, k, d_{\min})$. В [3] показано, что параметры конструкции следующие: N — размерность, $K = k$ — число информационных символов, $D^2_{\min} = 2\Delta^2 d_{\min}/N$ — квадрат минимального расстояния Евклида, нормированный по средней мощности.

Для обеспечения указанных параметров существенна нумерация сигналов во внутреннем ансамбле сигналов — она должна быть кодом Грея. Для кода Грея и в случае ФМ, и в случае КАМ выполняется условие, что два двумерных сигнала, отличающиеся в m битах, имеют разницу в расстоянии Евклида не менее $m\Delta^2$. Тогда при доказательстве минимального расстояния СКК можно использовать тот факт, что у двух слов на минимальном расстоянии друг от друга различные биты разбросаны по разным сигналам.

Прием данной СКК также состоит в последовательном выполнении демодуляции ансамбля сигналов и декодировании корректирующим кодом. Отличие состоит в том, что в случае мягкого декодирования внешним кодом вырабатываются специальные метрики для каждого бита в отдельности в одном и том же сигнале.

Конструкция 3 является обобщенным каскадным кодом $(q-1)$-го порядка и состоит из q пар внутренних и внешних кодов. В качестве внутренних кодов используется система вложенных ансамблей сигналов.

В случае ансамблей ФМ имеем вложенную цепочку ансамблей ФМ2 $\subset$ ФМ4 $\subset$ ФМ8 $\subset \cdots \subset$ ФМ2^q. Параметры модуляции на каждом шаге соответствуют описанным выше.

В случае ансамблей КАМ вложенная цепочка ансамблей принимает вид:

$$(\text{КАМ2})' \subset (\text{КАМ4})' \subset (\text{КАМ8})' \subset \cdots \subset \text{КАМ}2^q.$$

Здесь только исходная модуляция является обычной КАМ-модуляцией. Все остальные ансамбли получаются один из другого «прореживанием» точек в два

раза. Таким образом, на каждом шаге прореживания квадрат минимального расстояния Евклида увеличивается ровно в два раза, правда, немного смещается центр ансамбля.

Внешние корректирующие коды имеют параметры $(N/2, k_i, d_i)$, $i = \overline{1,q}$.

Тогда СКК имеет следующие параметры: N — размерность, $K = \sum_{i=1}^{q} k_i$ — число информационных символов, $D^2_{min} = \min_{i=\overline{1,q}}(2\Delta_i^2 d_i/N)$ — квадрат минимального расстояния Евклида, нормированный по средней мощности, а Δ_i^2 — квадрат нормированного расстояния ансамбля сигналов на i-м шаге вложенности.

Прием такой СКК аналогичен декодированию обобщенных каскадных кодов и состоит из q шагов. На i-м шаге происходит демодуляция сигналов в алфавите из 2^{q+1-i} сигналов, декодирование i-м внешним кодом и коррекция ансамблей сигналов по результатам декодирования (выбор соответствующего ансамбля из 2^{q-i} сигналов для следующего шага).

Возможна комбинация этих конструкций, когда в конструкции 3 на i-м шаге происходит разбиение на q_i ансамблей сигналов и используется конструкция 1 или 2.

Приведем пример построения конструкций конкретной размерности.

В качестве ансамбля сигналов используем ансамбль квадратурной амплитудной модуляции КАМ16. Размерность всех сравниваемых конструкций выберем $N = 32$, а квадрат минимального расстояния Евклида, нормированный по средней мощности, $D^2_{min} = 0,2$.

Тогда в конструкции 1 внешним является код РС над полем $GF(16)$ с параметрами (16,9,8). СКК содержит $4 \times 9 = 36$ информационных символов и имеет скорость $R = 1,125$ бит/изм.

В конструкции 2 выберем в качестве внешнего код Рида–Маллера 3-го порядка (64,42,8). СКК содержит 42 информационных символа и имеет скорость $R = 1,3125$ бит/изм.

В конструкции 3 выберем в качестве внешних коды (16,5,8), (16,11,4), (16,15,2) и (16,16,1). СКК содержит $5+11+15+16 = 47$ информационных символов и имеет скорость $R = 1,46875$ бит/изм.

Как видно, для данного примера эффективность конструкций возрастает с номером конструкции. Часто конструкция 2 оказывается лучше конструкции 3. Конструкция 1 обычно им проигрывает. В [2] приводятся подробные примеры построения СКК с различными длинами и числом информационных символов.

3.5. Описание сверточных СКК в гауссовом канале

Для сверточных кодов в принципе возможно построение таких же трех конструкций, как и в случае блоковых кодов. Однако все эти конструкции прежде всего эффективны для кодов регулярной структуры. Наиболее же эффективные и часто употребляемые сверточные коды являются кодами с достаточно коротким кодовым ограничением, строятся методами перебора и декодируются также перебором по алгоритму Витерби. Наверное, благодаря сказанному, наибольшее распространение для сверточных СКК получил отдельный подход, начало которому положила работа Г. Унгербека [6].

Суть подхода в следующем. Сверточный код, как и ранее, выбирается методами перебора из всего класса кодов данной структуры. На ребрах, в отличие от двоичного случая, выбираются либо непосредственно сигналы, либо номера подмножеств в разложении сигналов на достаточно большое число подмножеств. Также ребрам могут соответствовать не двумерные, а четырехмерные сигналы. Таким образом, строится оптимальная сверточная конструкция заданной сложности (с заданным числом узлов решетчатой диаграммы). Достигаются ЭВК порядка 5–6 дБ при том же R, что и в безызбыточной конструкции.

3.6. Модель канала с межсимвольной интерференцией (МСИ)

В реальных частотно-ограниченных каналах связи помимо аддитивного шума возникает межсимвольная интерференция (МСИ), вызванная памятью каналов. Отклик канала на последовательность входных сигналов вызывает взаимное наложение сигналов на выходе канала. Если нормировать по мощности амплитудно-частотную характеристику канала, то можно сказать, что МСИ приводит к значительному изменению расстояний между сигналами на выходе канала и, что особенно важно, к уменьшению минимального расстояния между ними.

При синтезе сигналов и кодов для каналов с МСИ этот эффект, как правило, не учитывается, т. е. в качестве входных сигналов выбираются такие, которые согласованы с идеальным каналом без МСИ. Однако МСИ стремятся учитывать при синтезе оптимального приемника (декодера). Широко известным решением такого рода является алгоритм Витерби и его модификация, учитывающая сверточное кодирование [7, 8].

Рассмотрим подход к кодированию в каналах с МСИ, основанный на синтезе таких сигнально-кодовых конструкций, которые учитывают «деформацию» пространства сигналов при передаче по реальному каналу [3, 9]. Основой этого подхода является показанная ниже возможность преобразования каналов с МСИ в совокупность гауссовых каналов без памяти, т. е. без МСИ, но отличающихся один от другого скалярным коэффициентом передачи или отношением сигнал/шум. В 90-х годах прошлого века данный подход получил название ортогональной частотной модуляции или ортогонального частотного мультиплексирования — Orthogonal Frequency Division Modulation (OFDM).

Широкий класс реальных каналов может быть представлен в виде линейной модели, состоящей из линейного фильтра и сумматора с аддитивным шумом [7, 8]. Вход и выход такого канала связаны выражением

$$\tilde{Z}(t) = \int_0^\infty h(\tau) \mathrm{Z}(t-\tau)\alpha\tau + \gamma(t),$$

где $h(\tau)$ — импульсная реакция канала (отклик на δ-функцию), $\gamma(\tau)$ — аддитивный шум. Импульсная реакция канала, аддитивный шум и входной сигнал $Z(t)$ в общем случае являются комплексными. Импульсная реакция канала $h(t)$ связана с передаточной функцией канала $K(\omega)$ преобразованием Фурье

$$h(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} K(\omega) e^{j\omega t} \alpha\omega.$$

Положим, что канал низкочастотный, т. е. $K(\omega)$ существенно отлична от нуля в интервале, середина которого — $\omega = 0$. Для любых линейных методов модуляции всегда может быть получен низкочастотный эквивалент канала. Канал также предполагается физически возможным, т. е. каузальным: $|h(t)| = 0, \forall t < 0$. Функция $h(t)$ принадлежит пространству функций $L_2(0, \infty)$, из чего вытекает следующее утверждение.

Существует точка на оси времени $t_{\text{гр}}(\varepsilon)$, такая, что

$$\int_{t_{\text{гр}}(\varepsilon)}^{\infty} |h(t)|^2 \alpha t < \varepsilon, \quad \text{где} \quad \varepsilon > 0.$$

Это позволяет аппроксимировать канал системой с конечной импульсной реакцией — в дальнейшем будем считать, что $h(t) = 0, \forall t > t_{\text{гр}}$. Цифровой сигнал на входе канала может быть представлен в виде

$$Z(t) = \sum_q Z_q \delta(t - qT), \qquad q = \overline{-\infty, \infty},$$

где Z_q — комплексный отсчет сигнала на входе канала, $\delta(t)$ — дельта функция, T — тактовый интервал.

Сигнал на выходе канала может быть представлен как

$$\tilde{Z}(t) = \sum_q Z_q h(t - qT) + \gamma(t).$$

Этот сигнал в приемнике дискретизируется по времени с частотой $f_T = 1/T$ в моменты времени $kT+Q$, где Q — начальная фаза стробирования. В результате этой операции получается:

$$\tilde{Z}(kT + Q) = \sum_q Z_q h(kT + Q - qT) + \gamma(kT + Q).$$

Обозначим через $\tilde{Z}_k = \tilde{Z}(kT + Q)$, $k - q = n$, $h_n = h(kT + Q - qT)$, $\gamma_k = \gamma(kT + Q)$. Тогда получаем выражение, связывающее последовательность отсчетов на входе канала Z_k, $k = \overline{-\infty, \infty}$ с последовательностью на его выходе:

$$\tilde{Z}_k = \sum_n h_n Z_{k-n} + \gamma_k.$$

В дальнейшем будем считать, что аддитивный шум является белым и гауссовым (АБГШ). Это значит, что γ_k — комплексная гауссова величина: $\gamma_k = \varsigma_k + j\iota_k$, причем $E(\gamma_k) = 0$,

$$E(|\gamma_k|^2) = \sigma_{\text{ш}}^2, \quad \forall k, \quad E(\gamma_k^* \gamma_e) = 0, \quad \forall k \neq l, \quad E(\varsigma_k \iota_k) = 0, \quad \forall k.$$

Величина $\sigma_{\text{ш}}^2$ представляет собой мощность АБГШ на входе приемника.

Описанная модель представляет собой общепринятую модель канала с межсимвольной интерференцией [8].

Последовательность отсчетов импульсной реакции $\{h_n\}$ будет конечной, поскольку сама импульсная реакция конечна. Пусть $h_o \neq 0$, $h_{L-1} \neq 0$ соответственно первый и последний отсчеты импульсной реакции канала $h(t)$. При этом параметр L определяет память канала, поскольку каждый отсчет сигнала на входе представляет собой линейную комбинацию L переданных отсчетов.

Импульсной характеристикой гауссова канала дискретного времени называется последовательность отсчетов $\{h_n\}_{n=0}^{L-1}$.

Перепишем сигнал на выходе канала с учетом конечности импульсной характеристики:

$$\tilde{Z}_k = \sum_{n}^{L-1} h_n Z_{k-n} + \gamma_k.$$

Именно это выражение и будет использовано в дальнейшем для описания канала, коэффициенты h_n могут быть как постоянными, так и медленно меняться во времени, что соответствует радиоканалам с переменными параметрами.

Временной характеристикой канала дискретного времени является импульсная характеристика $\{h_n\}_{n=0}^{L-1}$, а частотной — периодическая функция вида:

$$K_g(\omega) = \sum_{h=0}^{L-1} h_n e^{-j\omega nT} = \frac{1}{T}\sum_{q} K\left(\omega - q\frac{2\pi}{T}\right).$$

Передаточной функцией в полосе Найквиста называется фрагмент (один период) $K_g(\omega)$ на интервале $\left[-\frac{\pi}{T}, \frac{\pi}{T}\right]$.

Название «полоса Найквиста» связано с критерием Найквиста, определяющим условия отсутствия МСИ.

Для отсутствия ($h_n = 0,\ \forall n \neq 0,\ h_0 = 1$) необходимо и достаточно выполнение:

$$\frac{1}{T}\sum_{q} K\left(\omega - q\frac{2\pi}{T}\right) = 1, \quad \forall\omega.$$

Особенность рассмотренных ниже методов передачи дискретных сообщений в канале с МСИ заключается в том, что отсчеты сигнала поступают на вход канала не непрерывным потоком, а блоками конечной длины (N отсчетов), отделенными один от другого защитными интервалами, обеспечивающими отсутствие интерференции между блоками. В течение защитного интервала сигнал на входе канала тождественно равен нулю.

Если длина (число отсчетов) защитного интервала L_0 больше или равна памяти канала $L-1$, то блоки на выходе канала перекрываются. Таким образом, условие $L_0 \geqslant L$ является условием отсутствия межблочной интерференции. Длина блока на выходе больше, чем на входе, и равна $N_0 = N + L - 1$.

Переход от непрерывной последовательности отсчетов на входе канала к последовательности изолированных блоков позволяет воспользоваться векторным описанием канала с МСИ. Пусть $\overline{Z}_N^T = (Z_0, Z_1, \ldots, Z_{N-1})$ — вектор отсчетов сигнала на входе канала, $\widetilde{\overline{Z}}_{N_0}^T = (\tilde{Z}_0, \tilde{Z}_1, \ldots, \tilde{Z}_{N_0-1})$ — вектор отсчетов сигнала на выходе канала, $\overline{\Delta}_{N_0}^T = (\gamma_0, \gamma_1, \ldots, \gamma_{N_0-1})$ — вектор отсчетов АБГШ.

Тогда вход и выход канала связаны выражением

$$\widetilde{\overline{Z}}_{N_0} = K_{N_0 N}\overline{Z}_N + \overline{\Delta}_{N_0}, \tag{3.1}$$

где $K_{N_0 N}$ — матрица канала ($N_0 \times N$), элементами которой являются компоненты весовой последовательности $\{h_n\}_{n=0}^{L-1}$

$$K_{N_0N} = \begin{bmatrix} h_0 & & & & & \\ h_1 & h_0 & & & & \\ . & h_1 & & & h_0 & \\ . & . & & & h_1 & \\ h_{L-1} & . & & & . & \\ & h_{L-1} & & & . & \\ . & . & & & h_{L-1} & \end{bmatrix}.$$

Столбцы ленточной матрицы K_{N_0N} содержат все компоненты импульсной характеристики и отличаются один от другого сдвигом ненулевых элементов, причем ранг этой матрицы равен N.

Гауссовым каналом с МСИ (ГКМСИ) будем называть такой канал дискретного времени с конечной памятью, вход и выход которого связаны выражением (3.1), а блоки сигнала на выходе канала не перекрываются (не интерферируют): $L_0 \geqslant L-1$.

Импульсную характеристику ГКМСИ и его передаточную функцию в полосе Найквиста будем считать нормированными:

$$\sum_{n=0}^{L-1} |h_n| = \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} |K_g(\omega)|^2 d\omega = 1.$$

Введение защитных интервалов между блоками приводит к тому, что часть времени канал не используется для передачи сообщений. Это уменьшает скорость передачи в N_0/N раз. Относительная скорость определяется выражением:

$$\wp_0 = \frac{N}{N_0} = \frac{1}{1+L_0/N}.$$

Однако с увеличением размера блока N потери в скорости, связанные с введением защитных интервалов, могут быть сделаны сколь угодно малыми.

Все рассмотренные ниже конструкции сигналов для ГКМСИ получены в предположении, что ограничена средняя мощность сигнала на входе канала.

Средней мощностью сигнала на входе ГКМСИ называется величина

$$P_{\text{вх}}(\cdot)^* = \frac{1}{N_0} S_p(\bar{Z}_N \bar{Z}_N^*),$$

где $(\cdot)^*$ означает транспортирование и комплексное сопряжение, $S_p(\cdot)$ — оператор следа.

Допустимая средняя мощность на входе канала $P_1 \geqslant P_{\text{вх}}$ определяет с учетом нормировки импульсной характеристики и передаточной функции канала отношение сигнал/шум $S/N_{\text{ш}} \triangleq P_1/\sigma_{\text{ш}}^2$.

3.7. Преобразование канала с МСИ в параллельные каналы без памяти

Решение задачи построения эффективных СКК для ГКМСИ основано на замене последнего эквивалентным ему каналом, представляющим собой множество параллельных гауссовых каналов без памяти (ГКБП). Ниже рассмотрены два

способа преобразования ГКМСИ в совокупность параллельных ГКБП. Первый из этих способов обеспечивает строгую эквивалентность исходного и преобразованного каналов, однако на практике использован быть не может; второй — связан с некоторыми энергетическими потерями, но существенно менее сложен в реализации [3, 9].

Гауссовым каналом без памяти будем называть канал, вход и выход которого связаны между собой выражением $y = ax + e$, где x, y соответственно вход и выход ГКБП, $e = \varsigma + j\iota$ — не зависящий от x комплексный гауссов шум с независимыми составляющими ς и ι, имеющими одинаковое распределение $N(0, \sigma_ш/\sqrt{2})$, a — некоторая константа.

Преобразование ГКМСИ с помощью сингулярного разложения матрицы канала

Можно доказать [3, 9] следующее утверждение.

ГКМСИ с N-мерными блоками на входе и N_0-мерными блоками на выходе эквивалентен N параллельным ГКБП, выходы и входы которых связаны выражением

$$\tilde{Z}_i = \sqrt{\rho_i}\tilde{Z}_i + e_i, \quad i = \overline{0, N-1},$$

где e_i — независимые по i комплексные гауссовы случайные величины, ρ_i — сингулярные значения матрицы канала K_{N_0N} (собственные значения матрицы $K^*_{N_0N}K_{N_0N}$).

Доказательство осуществляется с помощью представления матрицы ГКМСИ K_{N_0N} в виде разложения по сингулярным значениям:

$$K_{N_0N} = \sum_{j=0}^{N-1} \sqrt{\rho_i}\bar{q}_i\bar{t}_i^* = Q_{N_0N}R_N^{1/2}T_N^*,$$

где Q_{N_0N} — матрица $(N_0 \times N)$: $Q_{N_0N} = (\bar{q}_0, \bar{q}_1, \ldots, \bar{q}_{N-1})$, T_N — унитарная матрица $(N_0 \times N)$: $T_N = (t_0, t_1, \ldots, t_{N-1})$, R_N — диагональная матрица: $R_N = \operatorname{diag}\{\rho_0, \rho_1, \ldots, \rho_N\}$; числа ρ_i, $i = \overline{0, N-1}$ являются сингулярными значениями матрицы K_{N_0N}, они совпадают с собственными значениями матрицы $K^*_{N_0N}K_{N_0N}$ и ненулевыми собственными значениями матрицы $K_{N_0N}K^*_{N_0N}$. Поскольку

$$\operatorname{rank} K_{N_0N} = \operatorname{rank} K^*_{N_0N}K_{N_0N} = \operatorname{rank} K_{N_0N}K^*_{N_0N} = N,$$
$$\rho_i > 0, \qquad \forall_i = \overline{0, N-1}.$$

Последовательность преобразований ГКМСИ в совокупность параллельных независимых ГКБП состоит в следующем. В передатчике вектор сообщения $\bar{Z}_N$ подвергается унитарному преобразованию T_N, и полученный сигнал $\bar{Z}_N$ поступает в канал. В приемнике принятый сигнал $\bar{Z}_{N_0}$ подвергается преобразованию $Q^*_{N_0N}$, матрица которого имеет ортонормированные столбцы $\bar{q}_i$, в результате чего получается сигнал на выходе эквивалентного канала $\bar{\tilde{Z}}_N$, связанный с переданными выражениями:

$$\bar{\tilde{Z}}_N = R_N^{1/2}\bar{Z}_N + \bar{E}_N, \quad \text{где} \quad \bar{\tilde{Z}}_N = Q^*_{N_0N}\bar{Z}_{N_0}.$$

МСИ, действующая в исходном канале, проявляется в преобразованном канале в различии коэффициентов передачи $\sqrt{\rho_i}$ каждого из ГКБП. Заметим, что поскольку матрица $K^*_{N_0N}K_{N_0N}$ — эрмитова, то все ее собственные значения ρ_i — вещественные.

Полученные в результате преобразования ГКБП отличаются один от другого коэффициентом передачи $\sqrt{\rho_i}$ или, что то же самое, отношением сигнал/шум.

Строгий метод преобразования канала с МСИ с использованием матриц Q_{N_0N} и T_N является единственным и дает наилучший из возможных результатов, однако весьма сложен с практической точки зрения по следующим причинам. Во-первых, для его реализации требуется точное знание на передающей и приемной стороне импульсной характеристики канала и учет последней при вычислении матриц Q_{N_0N} и T_N и коэффициентов передачи $\sqrt{\rho_i}$. Учитывая сложность вычисления собственных значений и собственных векторов, это представляется весьма проблематичным. Во-вторых, для данных преобразований неизвестны «быстрые» алгоритмы уменьшения сложности.

Преобразование ГКМСИ с помощью замены апериодической свертки циклической

Выражение, связывающее вход и выход ГКМСИ, соответствует апериодической свертке последовательности отсчетов входного сигнала с импульсной характеристикой $\{h_n\}_{n=0}^{L-1}$ канала [3, 9]. Заменим N-мерный вектор сигнала на входе ГКМСИ N_0-мерным, где $N_0 = N + L_0$, $L_0 \geqslant L - 1$, по правилу $\bar{Z}_{N_0} = P_{N_0N}\bar{Z}_N$, где P_{N_0N} — матрица $(N_0 \times N)$ периодического продолжения $P_{N_0N} = \begin{bmatrix} 0I_{L_0} \\ I_N \end{bmatrix}$, а I_{L_0}, I_N — единичные матрицы соответствующих размеров.

Видно, что L_0 первых компонент вектора $\bar{Z}_{N_0}$ совпадают с L_0 последних компонент вектора $\bar{Z}_N$: $\bar{Z}_{N_0}^T = (Z_{N-L_0}, Z_{N-L_0+\nu}, \ldots, Z_{N-1}, Z_0, Z_1, \ldots, Z_{N-1})$. Тогда на выходе ГКМСИ имеем

$$\bar{\bar{Z}}_{N_0+L-1} = K_{(N_0+L-1)N}\bar{Z}_{N_0} + \bar{\Delta}_{N_0+L-1}.$$

Здесь матрица $K_{(N_0+L-1)N}$ имеет размер $(N_0+L-1)N_0$. Выделим на приеме N компонент вектора $\bar{\bar{Z}}_{N_0+L-1}$, начиная с L_0-й компоненты $\bar{\bar{Z}}_N = S_{(N_0+L-1)N}^T\bar{\bar{Z}}_{N_0+L-1}$, где $S_{(N_0+L-1)N} = \begin{bmatrix} O_{(L_0-1)N} \\ I_N \\ O_{(L-1)N} \end{bmatrix}$. Здесь $O_{(L_0-1)N}$ и $O_{(L-1)N}$ — нулевые матрицы соответствующего размера. Тогда получаем, что сигнал на выходе канала соответствует выражению

$$\bar{\bar{Z}}_N = S_{(N_0+L-1)N}^T K_{(N_0+L-1)N} P_{N_0N}\bar{Z}_N + S_{(N_0+L-1)N}^T\bar{\Delta}_{N_0+L-1},$$

проанализировав которое убеждаемся в том, что матрица

$$C_N = S_{(N_0+L-1)N}^T K_{(N_0+L-1)N} P_{N_0N}$$

— циркулянт $(N \times N)$ вида:

$$C_N = \begin{bmatrix} h_0 & & & h_{L-1} & \dots & h_0 \\ & h_0 & & & & \\ & & \ddots & & & \\ & & & \ddots & & \\ h_{L-1} & & & & & h_{L-1} \\ & h_{L-1} & & & & \\ & & \ddots & & & \\ & & & \ddots & & \\ & & & h_{L-1} & \dots & h_0 \end{bmatrix}.$$

Из этого следует, что вход и выход канала связаны выражением $\tilde{\bar{Z}}_N = C_N \bar{Z}_N + \bar{\Delta}_N$, где $\bar{\Delta}_N = S^T_{(N_0+L-1)N} \bar{\Delta}_{N_0+L-1}$ — выделяемый в приемнике вектор шума.

Таким образом, мы показали, что периодическое продолжение вектора сигнала на передаче $\bar{Z}_N$ на $L_0 \geqslant L-1$ компонент и выделение на приеме N компонент, начиная с L_0, вектора $\tilde{\bar{Z}}_{N_0+L-1}$ преобразует ГКМСИ с апериодической сверткой K_{N_0N} в ГКМСИ с циклической сверткой C_N. Преобразование апериодической свертки в циклическую не увеличивает потерь в скорости передачи, связанных с введением защитных интервалов.

Заметим, что преобразование свертки в циклическую непременно сопряжено с потерей части энергии полезного сигнала в защитных интервалах.

Средняя мощность сигнала P_{L_0}, теряемая при выделении вектора $\tilde{\bar{Z}}_N$ на приеме, определяется величиной $P_{L_0} \leqslant P\frac{L_0}{N}$, где P — средняя мощность сигнала.

Выбором длины блока N доля мощности, теряемой в защитных интервалах, может быть сделана сколь угодно малой.

Такой канал будем называть гауссовым каналом с МСИ и циклической сверткой (ГКМСИ-Ц). Если C_N — циркулянт, то его собственные значения K_i являются компонентами N-точечного ДПФ импульсной характеристики ГКМСИ:

$$K_i = \sum_{n=0}^{L-1} h_n e^{-j\frac{2\pi}{N} i_n}, \quad i = \overline{0, n-1}.$$

Тогда получается, что ГКМСИ-Ц эквивалентен N параллельным ГКБП, входы и выходы которых связаны выражением:

$$\tilde{Z}_i = K_i Z_i + e_i, \quad i = \overline{0, N-1},$$

где e_i — независимые по i комплексные гауссовы случайные величины, K_i — собственные значения циркулянта C_N.

Необходимо еще раз подчеркнуть, что полученные ГКБП неэквивалентны исходному ГКМСИ, а эквивалентны полученному из последнего ГКМСИ-Ц. Отличие ГКМСИ-Ц от ГКМСИ состоит, во-первых, в упомянутой выше потере части энергии полезного сигнала из-за стирания в приемнике сигнала, попадающего в эти интервалы, и, во-вторых, в том, что $\operatorname{rank} C_N \leqslant N$, тогда как $\operatorname{rank} K = N$.

С практической точки зрения второй способ преобразования ГКМСИ в параллельные ГКБП существенно предпочтительнее первого. Причины этого в следующем:

- для реализации преобразования достаточно знать лишь верхнюю границу памяти канала $L_0 \geqslant L-1$, поскольку унитарные преобразования на передаче U_N^* и на приеме U_N являются инвариантными, т. е. от канала независящими;
- сложность преобразования ГКМСИ в совокупность параллельных ГКБП определяется сложностью вычисления ДПФ, для чего могут быть использованы «быстрые» алгоритмы со сложностью $O(\log N)$ операций на один отсчет сигнала, например алгоритм Кули–Тьюки.

Таким образом, применение обратного ДПФ и циклического продолжения вектора сигнала на передаче и выделение неискаженного межблочной интерференцией N-мерного блока на приеме в сочетании с прямым ДПФ на приеме позволяет преобразовать ГКМСИ в совокупность параллельных независимых ГКБП, используя минимальные априорные сведения о канале (L_0) и с минимальной сложностью цифровой обработки.

В заключение целесообразно подчеркнуть связь собственных значений K_i циркулянта C_N с параметрами исходного непрерывного канала и полученного из него канала дискретного времени:

$$K_i = K_g(\omega_i) = \frac{1}{T}\sum_q K(\omega_i - q\frac{2\pi}{T}),$$

где $\omega_i = 2\pi i/NT$, $i = \overline{0, N-1}$, $K(\omega)$ — передаточная функция непрерывного канала, $K_g(\omega)$ — передаточная функция в полосе Найквиста.

3.8. Пропускная способность канала с МСИ

Пропускной способностью ГКМСИ называется максимум взаимной информации между входом и выходом канала по всем распределениям $\bar{Z}_N$, удовлетворяющим ограничению на среднюю мощность сигнала на входе канала:

$$\frac{1}{N_0} S_p E(\bar{Z}_N \bar{Z}_N^*) \leqslant P_1.$$

Если мощность АБГШ на выходе ГКМСИ равна $\sigma_ш^2$, а средняя мощность сигнала на его входе ограничена величиной P_1, то пропускная способность ГКМСИ с матрицей K_{km}, сингулярные значения которой $\rho_0 \geqslant \rho_1 \geqslant \cdots \geqslant \rho_{N-1} > 0$, равна

$$C = C = \wp_0 \frac{1}{N}\sum_{i=0}^{M}\frac{1}{2}\log_2\left[\rho_i \frac{N}{M}\left(\frac{P_1}{\sigma_ш^2}\times\frac{1}{\wp_0} + \frac{1}{N}\sum_{p=0}^{M}\frac{1}{\rho_p}\right)\right],$$

где $M \leqslant N-1$ — наибольшее число, для которого

$$P_M = P_1\frac{N}{M}\times\frac{1}{\wp_0} + \sigma_ш^2\left(\frac{N}{M}\times\frac{1}{N}\sum_{\rho=0}^{M}\frac{1}{\rho_p} - \frac{1}{\rho_M}\right) > 0.$$

В предыдущем разделе было показано, что ГКМСИ может быть преобразован в канал с циклической матрицей ГКМСИ-Ц, отличающийся от исходного тем, что защитные интервалы на передаче заполняются циклическим продолжением блоков, энергия которых на приеме не используется.

Для такого канала справедливо следующее утверждение. Если мощность АБГШ на выходе ГКМСИ-Ц равна $\sigma_ш^2$, а средняя мощность сигнала на его

выходе ограничена величиной P_1, то пропускная способность ГКМСИ-Ц, собственные значения которого $|K_0| \geqslant |K_1| \geqslant \cdots \geqslant |K_{N-1}|$, равна

$$C = \wp_0 \frac{1}{N} \sum_{i=0}^{M} \frac{1}{2} \log_2 \left[|K_i|^2 \frac{N}{M} \left(\frac{P_1}{\sigma_{\text{ш}}^2} + \frac{1}{N} \sum_{P=0}^{M} \frac{1}{|K_P|^2} \right) \right],$$

где $M \leqslant N-1$ — наибольшее число, для которого

$$P_M = P_1 \frac{N}{M} + \sigma_{\text{ш}}^2 \left(\frac{N}{M} \times \frac{1}{N} \sum_{p=0}^{M} \frac{1}{|K_p|^2} - \frac{1}{|K_\mu|^2} \right) > 0.$$

Отличие этих двух формул для P_M состоит в том, что во втором случае отсутствует множитель $1/\wp_0$, учитывающий тот факт, что мощность блока сигнала на входе ГКМСИ может быть увеличена в $1/\wp_0$, так как в защитном интервале мощность на входе равна нулю. В ГКМСИ-Ц мощность сигнала в защитном интервале равна средней мощности сигнала. Здесь еще раз уместно подчеркнуть, что полученные результаты представляют собой частные случаи хорошо известного [7] общего гауссова канала.

При $N \to \infty$ ГКМСИ и ГКМСИ-Ц обращаются в гауссов канал дискретного времени, входом которого является стационарная случайная последовательность $\{Z_K\}_{K \to -\infty}^{\infty}$, а выходом — стационарная случайная последовательность $\left\{\tilde{Z}_K\right\}_{K \to -\infty}^{\infty}$.

3.9. Построение СКК для канала с МСИ и переменными параметрами (OFDM)

В результате описанного выше преобразования ГКМСИ-Ц в совокупность независимых параллельных ГКБП было получено, что вход и выход каждого канала связаны выражением:

$$\tilde{Z}_i = K_i Z_i + e_i, \quad i = \overline{0, N-1}.$$

Если исходный канал дискретного времени имеет существенную неравномерность амплитудно-частотной характеристики в полосе Найквиста, то полученные каналы могут быть весьма различны. Различие ГКБП должно учитываться при построении сигналов и СКК. Возможно очень много разных вариантов преобразования ГКМСИ-Ц в совокупность независимых параллельных ГКБП [9]. Рассмотрим два наиболее интересных из них с точки зрения реализации.

Исследуем возможность преобразования ГКМСИ-Ц в совокупность независимых параллельных одинаковых ГКБП [9].

Пусть, как и ранее, собственные значения матрицы канала упорядочены $|K_0| \geqslant |K_1| \geqslant \cdots \geqslant |K_{N-1}|$. Пусть M_1 таково, что $|K_{M_1-1}| > 0$, $|K_{M_1}| = |K_{M_1+1}| = \cdots = |K_{N-1}| = 0$, $0 < M_1 \leqslant N-1$. Иначе говоря: $\operatorname{rank} C_N = M_1$, где C_N — матрица ГКМСИ-Ц. Положим, что для всех $0 < i \leqslant M_1 - 1$ $Z_i = \frac{1}{|K_i|} \aleph_i$, $\tilde{\aleph}_i = b_i \tilde{Z}_i$, $b_i = e^{-j \arg K_i}$.

Тогда получаем, что $\tilde{\aleph}_i = \aleph_i + \tilde{e}_i$, где $\tilde{e}_i = e^{-j \arg K_i} e_i$.

Поскольку умножение комплексной гауссовой случайной величины с независимыми действительной и мнимой частями e_i на «поворачивающий» множитель

$e^{-j\arg K_i}$ не изменяет статистику шума, то вместо обозначения $\tilde{e}_i$ будем использовать e_i. Имеем:

$$\tilde{\aleph}_i = \aleph_i + e_i, \quad i = \overline{0, M_1 - 1}.$$

Тогда справедливо следующее утверждение.

Предыскажение на передаче $Z_i = \frac{1}{|K_i|}\aleph_i$ и коррекция на приеме $\tilde{\aleph}_i = b_i \tilde{Z}_i$ преобразуют ГКБП, эквивалентные ГКМСИ-Ц, в M_1 независимых одинаковых ГКБП, вход и выход которых связаны выражением $M_1 = \operatorname{rank} C_N$, C_N — матрица ГКМСИ-Ц.

Следующий вариант учитывает различие в каналах, но группирует их по различным «подмножествам». Рассмотрим случай, когда в параллельных ГКБП с предыскажениями используются различные алфавиты КАМ, но с одним и тем же минимальным расстоянием Евклида Δ, не зависящим от номера ГКБП i. Необходимость рассмотрения этого варианта объясняется возможностью построения на его основе эффективных сигналов и сигнально-кодовых конструкций.

Пусть $0 < \mu_1 < \mu_2 < \mu_3 < \cdots < \mu_Q \leqslant M_1$ — некоторое разбиение последовательности номеров ГКБП. Положим, что в ГКБП с номерами от μ_{j-1} до $\mu_j - 1 (\mu_0 = 0)$ используется алфавит КАМ с 2^{q_j} символами, $1 \leqslant j \leqslant Q$, причем $q_1 > q_2 > \cdots > q_Q \geqslant 1$. Это означает, что алфавиты с большим числом точек используются в ГКБП с большим отношением сигнал/шум или, что то же самое, с большими собственными значениями K_i.

Средняя мощность на выходе i-го ГКБП имеет вид:

$$P_{\text{вх}i} = P_{q_j} / |K_i|^2, \quad \mu_{j-1} \leqslant i \leqslant \mu_j - 1,$$

где $P_{q_j} = \Delta^2 \varphi(2^{q_j})$ — средняя мощность сигнала АФМ на входе и выходе ГКБП с номерами μ_{j-1} до $\mu_j - 1$, а

$$\varphi(q) = \left\{ \begin{array}{ll} (2^q - 1/2)/6, & q = 2m - 1, \\ (2^q - 1)/6, & q = 2m, \ m = 1, 2, \ldots \end{array} \right\}.$$

Средняя мощность на входе ГКМСИ-Ц ограничена величиной P_1:

$$\frac{1}{N} \sum_{i=0}^{M_1 - 1} P_{bxi} \leqslant P_1.$$

Подставляя два последних выражения одно в другое, получаем

$$\Delta^2 \sum_{j=1}^{Q} \varphi(2^{q_j}) \frac{1}{N} \sum_{i=\mu_{j-1}}^{M_j - 1} \frac{1}{|K_i|^2} \leqslant P_1.$$

С учетом того, что

$$\frac{1}{N} \sum_{i=\mu_{j-1}}^{\mu_j - 1} \frac{1}{|K_i|^2} = f_\mu(\mu_j) - f_\mu(\mu_{j-1}),$$

где $f_N(M) = (1/N) \sum\limits_{i=0}^{M-1} 1/|K_i|^2$, получаем

$$\Delta^2 \sum_{j=1}^{Q} \varphi(2^{qj}) \left[f_N(\mu_j) - f_N(\mu_{j-1}) \right] \leqslant P_1.$$

Максимальная скорость в каждом ГКБП при фиксированном q_j определяется следующим образом:

$$R(q_j, P_{qj}/\sigma^2_{ш}) = R(q_j, \Delta^2\varphi(2^{qj})/\sigma^2_{ш}).$$

Суммарная скорость в ГКМСИ-Ц задается выражением:

$$R = \wp_0 \frac{1}{N} \sum_{j=1}^{Q} \nu_j R(q_j, \Delta^2\varphi(2^{q_j})/\sigma^2_{ш}),$$

где $\nu_j = \mu_j - \mu_{j-1}$, $\mu_0 = 0$ — число ГКБП с одним и тем же алфавитом КАМ.

Оптимизация рассмотренного варианта по скорости при ограниченной средней мощности сигнала на входе канала свелась к выбору оптимального разбиения параллельных ГКБП на группы с одинаковой скоростью, оптимальному выбору алфавитов КАМ и минимального расстояния Δ в них. Это значит, что справедливо следующее утверждение.

Максимальная скорость, достижимая в ГКМСИ-Ц с предыскажениями и произвольными алфавитами КАМ в каждом из параллельных ГКБП при условии, что минимальное расстояние во всех алфавитах постоянно и равно Δ, задается выражением:

$$R_{\max} = \max_{\Delta>0} \max_{\nu_j=1,2,\ldots} \max_{q_j=1,2,\ldots} \wp_0 \frac{1}{N} \sum_{j=1}^{Q} \mu_j R(q_j, \frac{\Delta^2\varphi(2^{qj})}{\sigma^2_{ш}})$$

при ограничениях, приведенных выше, на допустимую среднюю мощность сигнала на входе ГКМСИ-Ц, а $\nu_j = \mu_j - \mu_{j-1}$, $\mu_0 = 0$, $0 < \mu_1 < \mu_2 < \cdots < \mu_Q \leqslant M_1$ — разбиение множества ГКБП на группы из ν_j параллельных каналов, в каждом из которых используется один и тот же алфавит КАМ со средней мощностью

$$P_{qj} = \Delta^2\varphi(2^{qj}).$$

Тем самым мы получили ступенчатую конструкцию, в которой передаваемый блок в результате виртуальных преобразований на передающей и приемной стороне преобразует систему к N параллельным каналам, из которых μ_1 модулируются КАМ2, μ_2 сигналов модулируются КАМ4, μ_3 сигналов модулируются КАМ8 и т. д. Последние μ_Q сигналов модулируются КАМ2^Q.

Теперь на данную сигнальную конструкцию можно «накладывать» корректирующий код и получать СКК. Операцию наложения кода можно делать независимо на каждый ГКБП, т. е. во временной области. Однако гораздо более привлекательно выглядит построение СКК в частотной области, когда единым блоком СКК «накрываются» все каналы в одном временном блоке. Такой вариант оказывается наиболее предпочтительным, так как он практически не увеличивает задержки. В принципе возможно применение любой из блоковых или сверточных СКК, описанных выше. Необходима лишь достаточно простая модификация СКК, учитывающая разность алфавитов в последовательных символах. В результате получается так называемая ступенчатая СКК, кодируемая и декодируемая в пределах одного блока преобразования сигналов.

В данном разделе совершенно не затронуты вопросы оценки импульсной характеристики канала для точного выполнения OFDM-преобразований на передаче и приеме. Наиболее простой подход, используемый на практике, — отведение некоторого числа равномерно расставленных подканалов, используемых только

для оценки характеристик канала. Особенно важно это в радиоканалах с меняющимися параметрами.

В том или ином виде данная СКК, основанная на параллельных каналах, и является так называемой модуляцией OFDM, входящей практически во все стандарты, описанные ниже. Дальнейшей модификацией этой идеи является ее обобщение на множество абонентов (OFDMA), что нашло свое отражение во всех перспективных стандартах, как например IEEE 802.16e. Разделение абонентов может происходить различными способами: по времени, по частоте или по коду.

Литература

1. Банкет В. Л., Дорофеев В. М. Цифровые методы в спутниковой связи. — М.: Радио и связь, 1988.
2. Зиновьев В. А., Зяблов В. В., Портной С. Л. Каскадные методы построения и декодирования кодов в евклидовом пространстве. — М.: Препринт ИППИ АН СССР, 1987.
3. Зяблов В. В., Коробков Д. Л., Портной С. Л. Высокоскоростная передача сообщений по реальным каналам. — М.: Радио и связь, 1990.
4. Гинзбург В. В. Многомерные сигналы для непрерывного канала. — Проблемы передачи информации, 1984, т. 21, № 1, с. 14–27,
5. Портной С. Л. Характеристики систем модуляции и кодирования с точки зрения каскадных кодов. — Проблемы передачи информации, 1985, т. 20, № 1, с. 28–46.
6. Ungerboeck G. Channel coding with multilevel phase signals. — IEEE Trans. Inform Theory. July, 1999, vol. 45, p. 1456–1467.
7. Галлагер Р. Теория информации и надежная связь/ Пер. с англ. под ред. М. С. Пинскера и Б. С. Цыбакова. — М.: Советское радио, 1974.
8. Витерби А. Д., Омура Д. К. Принципы цифровой связи и кодирования. — М.: Радио и связь, 1982.
9. Зяблов В. В., Коробков Д. Л., Портной С. Л. Каскадные методы кодирования и синтеза сигналов в каналах с межсимвольной интерференцией. — М.: Препринт ИППИ АН СССР, 1988.

ГЛАВА 4

СТАНДАРТЫ ЦИФРОВОГО ВИДЕО- И РАДИОВЕЩАНИЯ

Цифровое телевидение и радиовещание находятся несколько в стороне от темы нашего изложения. Это — исключительно вещательные, однонаправленные сети. Однако они, безусловно, занимают определенную — и очень массовую — нишу в мире широкополосных сетей передачи информации. Нельзя пренебрегать их влиянием на развитие технологий передачи информации поколения 4G. Тем более, что, в конечном итоге, именно телевидение всегда выступало мощнейшим стимулом для повышения пропускной способности всех широкополосных сетей передачи информации. Поэтому очень бегло рассмотрим основные технологии в этой области.

4.1. Цифровое ТВ-вещание

Телевидение — одна из важнейших областей применения технологий беспроводной высокоскоростной передачи данных. Первой ТВ-трансляцией по сети передачи информации считают передачу телевизионного изображения президента США Герберта Гувера по телефонным линиям из Вашингтона в Манхэттен в 1927 году, организованную компанией Bell System.

Телевидение — еще и одна из наиболее «древних» областей электронной техники, по возрасту не уступающая радио. Работы в этом направлении начались еще до появления беспроводной связи. Основой для многих систем раннего, так называемого оптико-механического телевидения был созданный в 1884 году и запатентованный в 1895-м немецким ученым Паулем Нипковым вращающийся перфорированный «развертывающий диск» из селеновых ячеек, раскладывающий изображение на элементы. В 1920-х–1930-х годах во всех странах работы в области телевидения, вплоть до попыток создания сетей вещания и цветного телевидения, были связаны именно с этим устройством. Сегодня мы переживаем очередную технологическую революцию в телевидении — начало массового распространения цифрового телевизионного вещания. Цифровое телевидение (ЦТВ) — это принципиально новые возможности, интерактивность, среда доставки мультимедийного трафика. Поэтому изменение формата ТВ-вещания — не просто сложная техническая задача, это — серьезнейший фактор, действующий в экономическом и социальном плане в общемировом масштабе.

Что дает цифровое телевидение? До массового зрителя практически без искажений доходит сигнал студийного качества, исчезают помехи, характерные для аналогового вещания. Появляется возможность передавать видеоизображение телевидения высокой четкости (ТВЧ, HDTV) с числом строк развертки 720,

1080 (соотношение строк и столбцов 9 : 16) и выше против стандартных сегодня 480–625 строк (формат 3 : 4). Однако увидеть эту высокую четкость можно лишь на экране ТВ-приемника с соответствующими характеристиками кинескопа. Поэтому не менее важно, что ЦТВ позволяет гораздо эффективнее использовать спектральный диапазон — в полосе одного аналогового ТВ-канала можно формировать несколько цифровых, на порядок возрастает число одновременно доступных ТВ-программ. Наконец, развитие средств доставки цифрового телевизионного сигнала — важный шаг к столь интенсивно прорабатываемому сегодня «телевидению по запросу», когда зритель выбирает уже не канал, а непосредственно тот фильм или передачу, которые хочет смотреть.

Выделяют четыре основных механизма передачи ТВ-трафика конечным потребителям — кабельный, спутниковый, наземный (эфирное вещание) и так называемый сотовый. Последний метод реализуют высокочастотные системы с труднопроизносимыми аббревиатурами: MMDS (Multichannel Microwave Distribution System), LMDS (Local Microwave Distribution System), MWS (Multimedia Wireless System), работающие (в России) в диапазонах 2,5–2,7; 27,5–29,5 и 40,5–43,5 ГГц соответственно. Нас системы ЦТВ будут интересовать как пример широкополосных БСПИ, характерная особенность которых — выраженная асимметрия трафика. Стратегический вопрос развития телевидения в государстве — сеть наземного вещания. Остановимся на ней подробнее.

В области массового вещания действуют два основных стандарта передачи сигнала: ATSC (Advanced Television Systems Committee, США) и DVB (Digital Video Broadcasting, Европа). Отдельно стоит Япония со стандартом ISDB (Integrated Services Digital Broadcasting). Противостояние американского и европейского стандартов, по крайней мере для России, уже вопрос прошлого, поскольку выбор в пользу DVB сделан, по всей видимости, окончательно. Эти стандарты аналогичны в том смысле, что оба ориентированы на передачу видео- и аудиоданных, кодированных и компрессированных посредством MPEG-2 [2]. Звук может кодироваться с помощью других алгоритмов, например Dolby AC-3. Поэтому качество картинки мало зависит от выбора одного или другого метода модуляции, если она успешно принята приемником.

В США при выборе стандарта цифрового телевизионного вещания основное внимание уделяли повышению качества изображения. При этом увеличение числа каналов со стандартным качеством изображения не рассматривалось. В Европе при разработке цифровой телевизионной системы DVB учли печальный опыт десятилетней давности, связанный с развертыванием коммерческого телевизионного вещания высокой четкости, в основном аналоговой ТВЧ-системы MAC (Multiplexed analog components). Поэтому основное внимание уделили увеличению числа каналов со стандартной разрешающей способностью. Это позволило начать переход к ЦТВ с выпуска относительно дешевых (тогда — 400–600 долл.) приставок для уже имеющихся телевизионных приемников. Кроме того, европейский стандарт рассматривается как основа для единой телекоммуникационной системы, ориентированной на передачу данных самой различной природы, отсюда требования к низкой вероятности ошибок. В то же время в ATSC отсутствует опция передачи дополнительных данных.

ATSC ориентирован на трансляцию в полосе шириной 6 МГц (стандартная ширина канала американской системы аналогового ТВ NTSC) одного потока

со скоростью 19,28 Мбит/с при наземном вещании и двух таких потоков — в сетях кабельного телевидения. DVB гораздо гибче: в стандартной полосе 8 МГц он обеспечивает выбор скорости в диапазоне от 4,98 до 31,67 Мбит/с (возможна работа с каналами шириной 6 и 7 МГц). Соответственно изменяется и число ТВ-программ в этой полосе — от 16 до 2, причем возможна одновременная трансляция программы с низким разрешением, но высокой надежностью и высоким разрешением при пониженной надежности приема. Уникальная особенность DVB — это мобильность приемника, он может перемещаться со скоростью до 300 км/ч — это поезда, междугородний пассажирский автотранспорт, мобильные службы (скорая помощь, полиция) и т. п.

Однако самое существенное — это надежность доставки сигнала, в чем ATSC уступает DVB. Причиной этого является принятая в ATSC многопозиционная амплитудная модуляция с частично подавленной боковой полосой 8-VSB (vestigal-sideband modulation system for broadcast). В целом 8-VSB можно назвать вершиной технической реализации систем амплитудной модуляции с подавлением боковой полосы. Но по сравнению с DVB в реальных условиях этот механизм проигрывает по надежности, гибкости, требованиям к приемным антеннам. Рассмотрим конкурирующие стандарты подробнее.

4.1.1. Стандарт ATSC

ATSC разрабатывался с учетом того, что какое-то время в одном диапазоне будут транслироваться сигналы как в стандарте ATSC, так и в действующем в США аналоговом стандарте NTSC. Постепенно NTSC должен быть вытеснен, и весь занятый аналоговым вещанием диапазон станет цифровым. В отличие от ATSC, спектр NTSC достаточно неравномерен. Поэтому для борьбы с интерференционными помехами в ATSC предусмотрен режекторный гребенчатый фильтр с подавлением основных составляющих сигнала NTSC (яркость, цвет и звук). Однако данный фильтр вдвое ухудшает для приемника соотношение сигнал/шум, поэтому он включается, только если приемник обнаруживает в эфире NTSC-сигнал.

Очень кратко рассмотрим систему VSB, принятую в ATSC. VSB — это амплитудная модуляция с подавлением нижней боковой полосы и частичным подавлением несущей. Уровней модуляции может быть от 2 (2-VSB) до 16 (16-VSB), они располагаются симметрично относительно 0. Так, амплитуда сигнала в 8-VSB принимает значения, пропорциональные целым числам от −7 до 7 с шагом 2. Для кабельного вещания принята наиболее быстрая система 16-VSB, для наземного — 8T-VSB с предварительным решетчатым (сверточным) кодированием со скоростью 2/3 (T — от trellis, решетка).

На вход VSB-системы поступают пакеты MPEG-2 объемом 188 байт (рис. 4.1). Из них удаляется байт синхронизации, к оставшимся 187 байтам добавляется 20 проверочных байт кода Рида–Соломона. Дополнительно включенный в 8T-VSB решетчатый кодер превращает каждые 2 бита получившегося 207-байтного слова в 3 бита, т. е. в один символ. Очевидно, что скорость передачи данных пропорциональна двоичному логарифму от числа уровней амплитудной модуляции (число бит на символ), однако чем больше уровней, тем ниже помехозащищенность. Частота следования символов в VSB — 10,76 МГц. В результате скорость в системе 16-VSB, где каждым 4 битам соответствует один символ,

в два раза выше (38,6 Мбит/с), чем в 8T-VSB (19,3 Мбит/с), поскольку для передачи одного пакета требуется вдвое меньше символов. Однако и пороговый уровень соотношения сигнал/шум у данных систем соответственно 28,3 и 14,9 дБ.

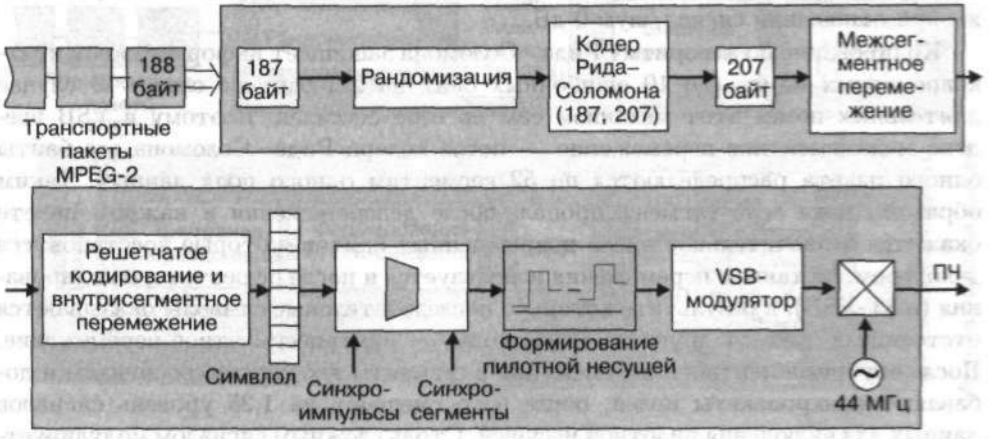


Рис. 4.1. Формирование сигнала в системе 8T-VSB

В итоге в 8T-VSB каждый байт превращается в 4 символа, 207 байт кодированного пакета — в 828 символов. Вместо синхробайта MPEG-2 используется синхросигнал (бросок амплитуды с уровня 5 до уровня −5) длительностью 0,37 мкс, что соответствует 4 символам. Получившиеся $828 + 4 = 832$ символа образуют сегмент. Каждые 312 сегментов данных и дополнительный синхросегмент объединяются в поле данных, два поля данных — в кадр (626 сегментов). Синхросегмент содержит ряд опорных последовательностей для настройки корректоров приемника, информацию о числе уровней модуляции и (только в 8T-VSB) 12 последних символов предшествующего сегмента.

Перед поступлением на вход кодера Рида–Соломона данные входного потока рандомизируют — превращают в псевдослучайные числа. Для этого каждый байт входного потока побитно складывается по модулю 2 с циклической псевдослучайной последовательностью. Генератором этой последовательности служит сдвиговый регистр из 16 триггеров, охваченный набором обратных связей. Во время синхроимпульса первого сегмента данных каждого поля в регистр загружается число $F180_{16}$. Восемь выходов регистра формируют поток байтов псевдослучайной последовательности. В результате сигнал становится практически независимым от изображения и шумоподобным, его спектральная плотность равномерно распределяется по полосе. Кроме повышения спектральной эффективности улучшается и синхронизация передачи: поскольку уровни модуляции симметричны относительно 0, а амплитуды информационных символов достаточно случайны, средний уровень сигнала также близок к 0. На этом фоне легко детектировать периодически повторяющийся синхроимпульс и синхросегмент, даже при отношении сигнал/шум 0 дБ.

Поскольку средний уровень информационных сигналов близок к 0, амплитуда несущей при амплитудной модуляции также была бы нулевой. Но несущая

в качестве пилот-сигнала необходима в данной системе для восстановления синхронизации в приемнике, поэтому вводят смещение — уровень каждого сигнала данных увеличивают на 1,25 единицы. Это соответствует появлению маломощного периодического сигнала несущей, добавляющего лишь 0,3 дБ к общей мощности сигнала. Данного уровня достаточно для детектирования несущей даже при отношении сигнал/шум 0 дБ.

Кодирование по алгоритму Рида–Соломона защищает информацию от кратковременных помех (до 10 ошибочных байт на 207-байтное слово). В случае длительных помех этот механизм сам по себе бессилен. Поэтому в VSB введено межсегментное перемежение — после кодера Рида–Соломона все байты одного пакета распределяются по 52 сегментам одного поля данных. Таким образом, даже если сегмент пропал, после деперемежения в каждом пакете окажется незначительное число поврежденных байтов, которые восстановятся декодером. Механизм перемежения используется и после решетчатого кодирования (в 8T-VSB), в результате которого последовательные символы оказываются отстоящими друг от друга на 12 символов — внутрисегментное перемежение. После внутрисегментного перемежения в сегменты вводят синхросигналы и добавляют синхропакеты полей, после чего смещают на 1,25 уровень сигналов данных для включения пилотной несущей. Столь сложным сигналом модулируется несущая на промежуточной частоте (ПЧ) 44 МГц, которая затем переносится непосредственно в полосу заданного телеканала.

К сожалению, имея лучшие теоретические показатели пороговых соотношений сигнал/шум и энергетической эффективности сигнала (на бит информации) [3], ATSC недостаточно надежен при многолучевом распространении сигналов, что неизбежно в условиях городской застройки. При этом отраженный сигнал, пришедший с задержкой, уже является помехой по отношению к основному. И если уровень отраженного сигнала менее чем на 15 дБ отличается от прямого, приемник теряет работоспособность — появляются помехи. То же самое относится и к приему сигналов от двух передатчиков. Поэтому заявленный выигрыш систем ATSC по отношению к DVB по показателю сигнал/шум в 4–6 дБ может проявиться разве что в лабораторных условиях. Практика же показывает, что в Нью-Йорке передатчик ATSC мощностью 350 кВт не обеспечивает 100% приема в радиусе 10 км, тогда как в Лондоне достаточно передатчика DVB-T мощностью 10 кВт для уверенного приема в радиусе 114 км [4]. Приходится усложнять антенные системы, что, очевидно, пользователей не радует.

4.1.2. Стандарт DVB

DVB, принятый в Европе, — это набор спецификаций, охватывающий кабельное DVB-C (cable), спутниковое DVB-S (sattelite) и наземное DVB-T (terrestrial) вещание. Наиболее сложный алгоритм — в DVB-T [8], поскольку условия работы и требования к передаче при наземном вещании наиболее жесткие. Очень кратко остановимся на наиболее примечательных особенностях DVB.

Предварительная обработка пакетов в DVB в принципе аналогична 8-VSB, хотя механизмы реализации функций различны. На вход кодера поступают транспортные пакеты MPEG-2 по 188 байт (1 синхробайт (всегда 47_{16})+187 байт данных) (рис. 4.2). Прежде всего, они рандомизируются посредством сложения по модулю 2 с двоичной псевдослучайной последовательностью (генератор —

15-разрядный сдвиговый регистр). Генератор инициализируется через каждые восемь пакетов одним и тем же числом ($4B80_{16}$). Синхробайты не рандомизируют, каждый восьмой синхробайт инвертируют.

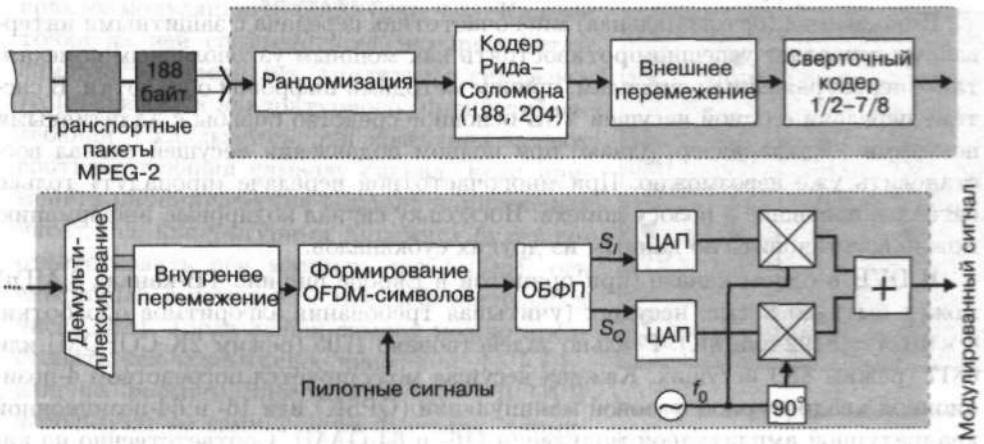


Рис. 4.2. Формирование сигнала в системе DVB-T

После рандомизации пакеты защищают кодом Рида – Соломона, в результате чего к 188 байт добавляются 16 проверочных — возможно исправление до восьми ошибочных байт на кодированный 204-байтный пакет. Затем байты перемешиваются внутри кодированных пакетов, причем так, что синхробайты остаются на своих местах, — это внешнее перемежение. Затем следует внутреннее сверточное кодирование. Его реализует сдвиговый регистр из шести триггеров, превращающий каждый входной бит в два выходных (скорость кодера 1/2). В DVB можно выбирать скорость сверточного кодирования (1/2, 2/3, 3/4, 5/6, 7/8), используя не оба элемента выходных пар, а лишь один. До сих пор функционально все было аналогично системе 8-VSB. Дальше начинаются принципиальные различия, связанные с модуляцией радиосигнала.

В стандарте использована модуляция COFDM (Coded Orthogonal Frequency-Division Multiplexing) — вариант мультиплексирования посредством ортогональных несущих (OFDM) с предварительным кодированием сигнала. Модуляция OFDM подразумевает, что весь диапазон канала вещания разбит на множество ортогональных поднесущих. Ортогональность означает, что усредненное по времени произведение двух несущих равно нулю. Частоты поднесущих задаются как $f_n(t) = \cos 2\pi(f_0 + n/\tau)t$, где f_0 — нижняя частота диапазона; n — номер поднесущей ($n = 0 \ldots N-1$; N — число поднесущих); τ — временной интервал передачи одного символа. Поток данных разбивается на N субпотоков, несущая каждого из которых модулируется с гораздо меньшей скоростью. Разнос несущих по частоте равен $1/\tau$.

Поскольку в отдельном субканале скорость передачи невелика, перед каждым символом можно ввести защитный интервал — временной отрезок до $0{,}25\tau$, в течение которого транслируется фрагмент уже переданного символа (для сохранения ортогональности несущих). Основное назначение защитных интервалов — борьба с межсимвольными помехами, вызванными в том чис-

ле и переотражениями сигналов. Действительно, поскольку скорость символов мала, переотраженный сигнал в приемнике «накладывается» на прямо распространяющийся сигнал в интервале одного символа, а не следующего, попадая в защитный интервал.

Независимая (ортогональная) многочастотная передача с защитными интервалами позволяет успешно противостоять как мощным узкополосным помехам, так и переотраженным сигналам, причем методами цифровой обработки. В системе передачи с одной несущей VSB основное средство борьбы с аддитивными помехами — эквалайзер, однако при полном подавлении несущей сигнал восстановить уже невозможно. При многочастотной передаче «пропадут» только сигналы, попавшие в полосу помехи. Поскольку сигнал кодирован, информацию можно восстановить по данным из других субканалов.

В DVB в одном канале (при принятой в Европе ширине ТВ-канала 8 МГц) может быть до 8 тыс. несущих (учитывая требования алгоритмов обработки, $8 \times 1024 = 8192$ или 8K). Реально задействовано 1705 (режим 2K COFDM) или 6817 (режим 8K) несущих. Каждая несущая модулируется посредством 4-позиционной квадратурной фазовой манипуляции (QPSK) или 16- и 64-позиционной квадратурной амплитудной модуляции (16- и 64-QAM). Соответственно на каждой несущей один модуляционный символ определяет от 2 до 6 бит.

Напомним, что при квадратурной модуляции выходной сигнал формируется сложением двух смещенных друг относительно друга на 90° гармонических колебаний на одной частоте f — синфазного $S_I(t) = A_S \cos \omega t$ и квадратурного $S_Q(t) = -A_Q \sin \omega t$ ($\omega = 2\pi f$). Их сумма — исходное колебание с фазовым сдвигом на угол $\varphi = \operatorname{arctg}(A_Q/A_S)$:

$$A_S \cos \omega t - A_Q \sin \omega t = A \cos(\omega t + \varphi); \qquad A = \sqrt{A_S^2 + A_Q^2}.$$

В соответствии с числом уровней модуляции исходный поток данных разбивается на n субпотоков — по числу битов в модуляционном символе. Для QPSK таких субпотоков два, для 16-QAM — четыре. Демультиплексирование происходит побитно — скажем, при модуляции 64-QAM ($n = 6$) первый бит попадает в первый субпоток, шестой — в шестой, седьмой — снова в первый и т. д. В DVB в каждом субпотоке биты переставляются по определенному правилу (своему для каждого субпотока) в пределах блока в 126 бит — внутреннее перемежение. Параллельные выходы устройств перемежения формируют модуляционный символ: 2-, 4- или 6-разрядный. На одной несущей OFDM передается один символ, поэтому в режиме 8K одновременно транслируется 48 групп по 126 символов — всего $48 \times 126 = 6048$ информационных несущих (или 12 групп по 126 символов на 1512 несущих в режиме 2K). Одновременно передаваемые QAM-символы входят в ODFM-символ. Они распределяются по субканалам OFDM не последовательно, а опять-таки перемежевываются по определенному закону. Поэтому, если OFDM-символ пропадает, его данные можно восстановить, поскольку биты одного кодированного пакета оказываются распределенными по многим OFDM-символам.

Очевидно, что реализовать метод передачи OFDM «в лоб», т. е. использовать несколько тысяч генераторов модулированных поднесущих, весьма проблематично. А на приемной стороне это и вовсе неразрешимая задача. Однако современные методы цифровой обработки сигнала позволяют существенно упростить

ее решение, используя отработанные алгоритмы прямого и обратного быстрого преобразования Фурье (БПФ и ОБПФ). Как это происходит?

Рассмотрим для примера векторную диаграмму модуляции 16-QAM (рис. 4.3). Каждая точка на ней соответствует четырем битам символа и определяет амплитуды синфазного (абсцисса) и квадратурного (ордината) колебаний. Складываясь, эти колебания задают соответствующий символу сигнал. Если применять математический аппарат комплексных чисел, ось квадратурных амплитуд будет соответствовать оси мнимых чисел (Im), ось синфазных амплитуд — оси действительных чисел (Re). Тогда любой символ можно представить как комплексное число $z = A_S + iA_Q$ или, по формуле Эйлера, $z = Ae^{i\varphi}$. В DVB используют не сами z, а их значения, нормированные на среднюю амплитуду суммарного колебания при выбранном виде модуляции. Это необходимо для усреднения амплитуд сигналов при различных режимах модуляции (очевидно, что при QPSK нормирующий множитель $c = 1/\sqrt{2}$, тогда как при 16-QAM $c = 1/\sqrt{10}$).

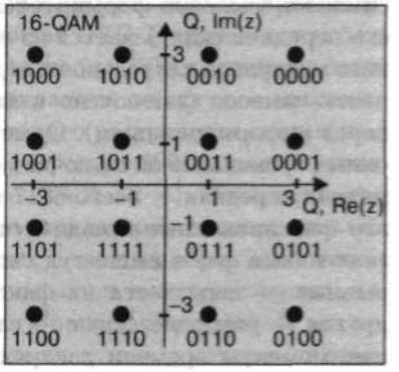


Рис. 4.3. Векторная диаграмма 16-QAM

Запишем в комплексном виде модулированный символом $C_k(t)$ сигнал на k-й несущей (без ее переноса в диапазон вещания): $S_k(t) = C_k(t)e^{i2\pi kt/\tau}$. Учитывая, что сигнал изменяется дискретно, т. е. $t = nT$, где T — длительность дискретного интервала; n — номер отсчета, получим $S_k(nT) = C_k(nT)e^{i2\pi knT/\tau}$. Тогда общий сигнал OFDM-символа

$$S_n = \sum_{k=0}^{N-1} C_k(nT)e^{i2\pi knT/\tau}. \tag{4.1}$$

Математически это аналогично вычислению дискретных значений функции по дискретным значениям амплитуд (C_k) ее гармонических составляющих (с частотами k/τ). Данную процедуру описывает ОПФ:

$$X_n = \sum_{k=0}^{N-1} a_k e^{i2\pi kn/N}, \tag{4.2}$$

где N — число гармонических составляющих, в нашем случае — число несущих. Если длительность интервала дискретизации T выразить как τ/N, выражение (4.1) станет аналогичным ОПФ (4.2). Алгоритмы БПФ, как прямого, так и обратного, достаточно хорошо проработаны, в том числе и с точки зрения их аппаратной реализации. Они наиболее эффективны при N вида 2^m. Поэтому в 8K-COFDM число несущих условно принято равным $2^{13} = 8192$ — просто не все из них используются. Величина $1/T = N/\tau$ — это так называемая системная тактовая частота, для полосы канала в 8 МГц она равна 64/7 МГц. При переходе к другому частотному плану, например с полосой ТВ-канала 7 или 6 МГц, достаточно изменить системную тактовую частоту, сохраняя неизменной всю структуру обработки сигнала (а вместе с ней — основные функциональные устройства). Отметим, что системная тактовая частота одинакова в режимах

2K и 8K, т. е. от числа несущих скорость передачи напрямую не зависит, изменяется только надежность.

Таким образом, посредством ОБПФ из входного массива модуляционных символов численно формируется выходной OFDM-символ. Временной интервал его передачи складывается из собственно времени передачи символа τ и защитного интервала длительностью до $\tau/4$, в течение которого «повторно» передается часть символа (заключено в кавычки, поскольку защитный интервал следует перед информационным). Отметим, что кроме 6048 (в режиме 8K) информационных субканалов он включает еще пилотные сигналы, а также сведения о параметрах передачи — всего 6817 модулированных несущих. Пилотные сигналы — это фиксированные псевдослучайные последовательности с точно известными значениями фаз и амплитуд сигналов. Одна часть пилотных сигналов — непрерывные — передается на фиксированных несущих в каждом OFDM-символе, другая — распределенные — случайным образом (но равномерно) в произвольные моменты времени распределяется по спектральному диапазону передачи. Назначение пилотных сигналов — синхронизация и оценка параметров канала передачи.

Синтезировать OFDM-символы недостаточно, необходимо еще сформировать радиосигнал в заданной частотной области (с нижней частотой f_0). Перенос символа в необходимый диапазон — это его смещение на частоту f_0, что в комплексной форме эквивалентно умножению на комплексное (в виде квадратурных слагаемых) представление несущей f_0. При этом амплитуды перемножаются, а аргументы складываются. Выделяя действительную (синфазную) и мнимую (квадратурную) составляющие $S(n)$ и умножая их соответственно на $\cos(2\pi f_0 t)$ и $-\sin(2\pi f_0 t)$, после суммирования получим полный сигнал одного OFDM-символа.

Описанные механизмы позволяют очень гибко выбирать необходимый режим вещания, а также совмещать два потока пакетов MPEG-2 — с высокой и низкой скоростью. Возможную скорость определяют вид модуляции, скорость сверточного кодирования (R), величина защитного интервала T_3 ($\tau/4$, $\tau/8$, $\tau/16$, $\tau/32$). Учитывая, что при 8K-OFDM $\tau = 896$ мкс, скорость изменяется в пределах от 4,98 Мбит/с (QPSK, R = 1/2, $T_3 = \tau/4$) до 31,67 Мбит/с (64-QAM, R = 7/8, $T_3 = \tau/32$).

Мы чрезвычайно схематично рассмотрели принцип передачи сигнала в DVB-T. Однако сигнал надо еще принять, демультиплексировать и декодировать, что сложнее, чем синтезировать его в передатчике. Для этого дополнительно к алгоритмам передатчика применяют корреляционные детекторы, декодеры с алгоритмом Витерби и т. д. Причем приемное устройство должно быть компактным и недорогим, иначе кто же его купит. Транспортные пакеты MPEG-2 или MPEG-4 также надо декодировать и сформировать ТВ-сигнал — цифровой или аналоговый, в зависимости от типа телевизора. Поэтому ЦТВ-приемник — это достаточно сложный программно-аппаратный комплекс, и только технологические достижения последних лет позволяют делать его недорогим при массовом выпуске.

Дальнейшее развитие стандартов DVB

К 2009 году, вопреки первоначальным планам, ни в одной из стран мира не произошел тотальный переход на сети цифрового ТВ-вещания, хотя некоторые

страны (США, Великобритания и др.) приблизились к этому вплотную. Однако в мире уже продано порядка 244 млн. ТВ-приемников различных типов стандарта DVB (C, S и T). Среди них 107 млн. — стандарта DVB-S и 94 млн. — стандарта DVB-T. В связи со столь массовыми продажами цены на приемные устройства продолжают неуклонно падать, уже сейчас DVB-T-приемники стоят дешевле 30 евро. С другой стороны, в ближайшие годы (до 2012 г.) практически все сколь-нибудь развитые страны планируют прекратить аналоговое вещание. В том числе, это означает и высвобождение частотного ресурса для систем цифрового ТВ. Учитывая все это, консорциум DVB к середине 2008 года выпустил новый стандарт наземного цифрового вещания — DVB-T2 [1]. Пока этот документ имеет статус предварительного стандарта, однако ряд производителей начал выпуск соответствующего ему оборудования, а в Великобритании в июне 2008 года компания BBC приступила к экспериментальным трансляциям по стандарту DVB-T2 (передатчик размещался на юго-западе Лондона). В марте 2009 года в Турине успешно прошел «фестиваль подключений», в котором приняли участие шесть производителей передающего (модулирующего) оборудования DVB-T2 и пять производителей приемного.

Отметим, что DVB-T2 — это стандарт второго поколения. Первый стандарт второго поколения консорциум DVB выпустил в 2003 году — стандарт цифрового спутникового вещания DVB-S2. Но его рассматривать мы не будем, поскольку практически все заложенные в DVB-S2 идеи были воплощены в стандарте DVB-T2.

Стандарт DVB-T2 призван как минимум на 30% улучшить емкость сетей ЦТВ по сравнению с DVB-T, при той же инфраструктуре сети и частотных ресурсах. Однако на практике выигрыш оказывается не ниже 50%. DVB-T2 принципиально отличается как архитектурой системного уровня (MAC-уровня), так и особенностями физического уровня.

На системном уровне принципиальное отличие нового стандарта — это концепция магистральных потоков физического уровня (Physical Layer Pipe — PLP). Если стандарт DVB-T был предназначен исключительно для передачи пакетов MPEG-2, то сеть DVB-T2 способна транслировать самые разные по природе и структуре информационные потоки (рис. 4.4). Система DVB-T2 способна передавать несколько независимых мультимедийных потоков, каждый со своей схемой модуляции, скоростью кодирования и временными интервалами. Возникает относительно сложная кадровая структура как на логическом, так и на физическом уровнях — ничего подобного в DVB-T не было. Соответственно, в системе DVB-T2 появляется новая функция — предварительная обработка входных потоков (рис. 4.5). В целом, общая схема обработки сигналов в системе DVB-2T существенно усложняется (рис. 4.6).

В стандарте различаются три основных типа потоков — транспортный (Transport Stream — TS), обобщенный инкапсулированный (Generic Encapsulated Stream — GSE) и обобщенный непрерывный (Generic Continuous Stream — GCS). Каждый поток представляет собой последовательность пользовательских пакетов (UP — User Packet). Транспортный поток — это последовательность пакетов фиксированной длины (пакеты MPEG-2, 188 байт, первый байт всегда синхробайт со значением 47_{16}). Поток GSE характеризуется пакетами переменной или фиксированной длины, которая указывается в заголовках этих пакетов. Поток

GCS представляет собой непрерывный поток битов. Реально это или последовательность пакетов без указания их длины, или пакеты максимально возможной длины 64 кбит.

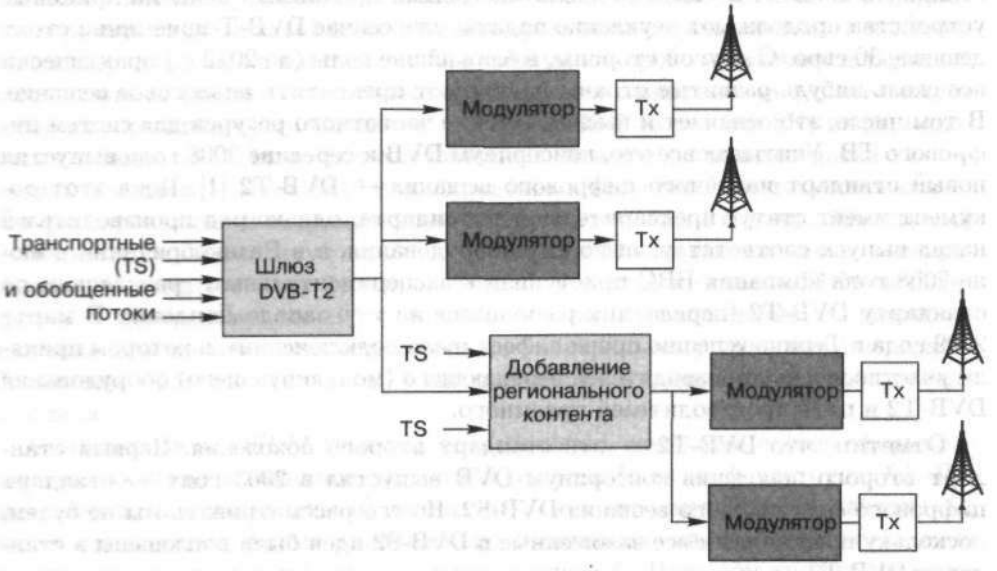


Рис. 4.4. Схема трансляции в сети DVB-T2

Пакеты каждого магистрального потока объединяются в потоковые (Baseband) кадры (BB-кадры) — отдельно для каждого потока (рис. 4.7). BB-кадр содержит BB-заголовок (80 бит), поле данных и поле выравнивания. В последнем можно передавать данные внутриканальной сигнализации. В заголовке пакета содержится информация о типе транспортного потока, размере пользовательского пакета (при необходимости) и всего поля данных, наличии режимов удаления пустых пакетов и дополнительных синхропакетов, используется постоянная/переменная модуляция и т. п. Размер поля данных и выравнивающего поля определяется параметрами сверточного кодера (в сумме не более 53770 бит).

Рис. 4.5. Обобщенная схема обработки передаваемых сигналов в системе DVB-T2

Стандарт DVB-T2 ориентирован на передачу телевизионных потоков, в которых зачастую используются пустые пакеты (для выравнивания скорости потока), разного рода задержки и т. п. для сохранения постоянной скорости потока. Поэтому в DVB-T2 предусмотрены средства удаления этой избыточной инфор-

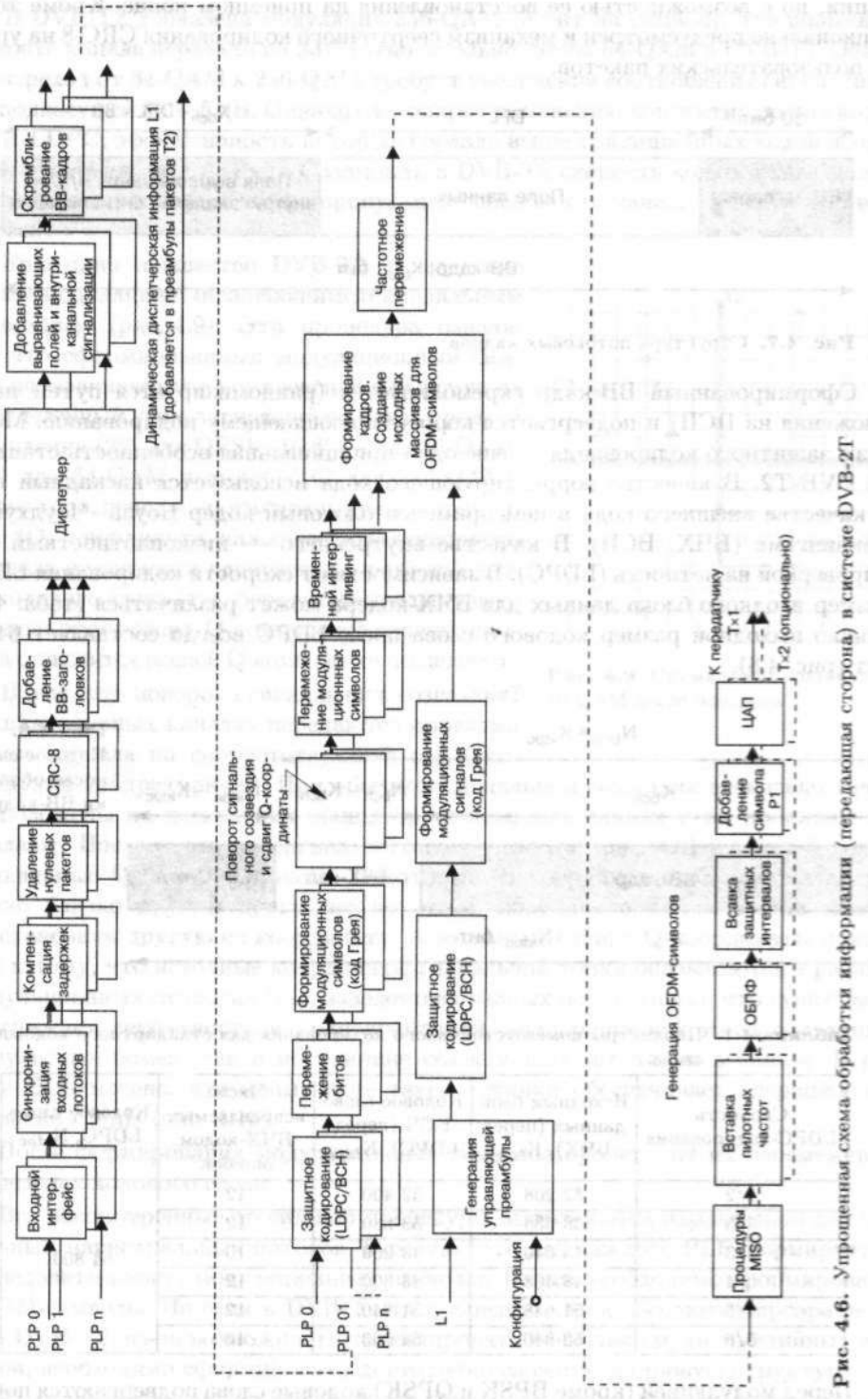


Рис. 4.6. Упрощенная схема обработки информации (передающая сторона) в системе DVB-2T

мации, но с возможностью ее восстановления на приемном конце. Кроме того, опционально предусмотрен и механизм сверточного кодирования CRC-8 на уровне пользовательских пакетов.

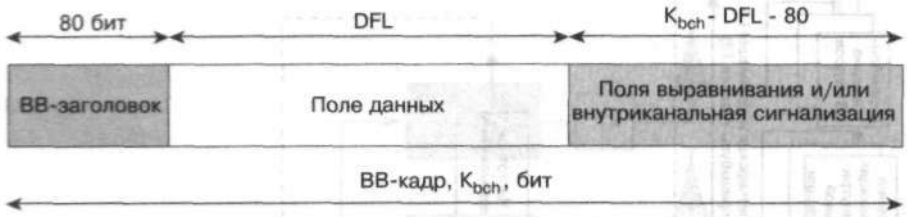


Рис. 4.7. Структура потоковых кадров

Сформированный BB-кадр скремблируется (рандомизируется путем перемножения на ПСП) и подвергается корректирующающему кодированию. Механизм защитного кодирования — еще одна принципальная особенность стандарта DVB-T2. В качестве корректирующего кода используется каскадный код. В качестве внешнего кода в нем применен блоковый кодер Боуза–Чоудхури–Хоквингема (БЧХ, BCH). В качестве внутреннего — низкоплотностный код с проверкой на четность (LDPC). В зависимости от скорости кодирования LDPC размер входного блока данных для БЧХ-кодера может различаться (табл. 4.1), однако выходной размер кодового слова после LDPC всегда составляет 64800 бит (рис. 4.8).

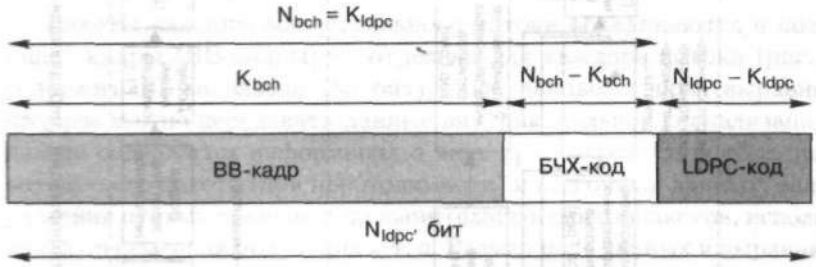


Рис. 4.8. Кодовое слово после обработки BB-кадра

Таблица 4.1. Параметры помехоустойчивого кодирования для стандартного кодового слова

Скорость LDPC-кодирования	Исходный блок данных (перед БЧХ), K_{bch}	Кодовое слово БСЧ (перед LDPC), K_{ldpc}	Число исправляемых БЧХ-кодом ошибок	Кодовое слово LDPC, N_{ldpc}
1/2	32 208	32 400	12	64 800
3/5	38 688	38 880	12	
2/3	43 040	43 200	10	
3/4	48 408	48 600	12	
4/5	51 648	51 840	12	
5/6	53 840	54 000	10	

Перед модуляцией (кроме BPSK и QPSK) кодовые слова подвергаются побитному перемежению и распределяются по модуляционным символам (см. рис. 4.6).

В DVB-T2 добавлена модуляция 256-QAM (8 бит на символ), что повышает емкость канала передачи на 33% (относительно схемы 64-QAM в DVB-T). Обычно переход от 64-QAM к 256-QAM требует увеличения соотношения сигнал/шум на поднесущей на 4–5 дБ. Однако благодаря применению корректирующих кодов BCH-LDPC, эффективность которых гораздо выше традиционных кодов исправления ошибок (в т. ч. Рида-Соломона), в DVB-T2 скорость кодирования может быть намного выше, и общая пропускная способность канала существенно возрастает.

Еще одно новшество DVB-T2 — введение схемы модуляции с «вращающимся» сигнальным созвездием (рис. 4.9). Эта процедура означает, что сформированный модуляционный символ поворачивается в комплексной плоскости на определенный угол, зависящий от числа уровней модуляции (29° для QPSK, 16,8° — для 16-QAM, 8,6° для 64-QAM и arctg(1/16) для 256-QAM). Более того, перед началом вращения квадратурная (Q) координата каждого модуляционного символа циклически сдвигается в рамках одного кодового слова (т. е. берется из предыдущего символа этого слова, Q-компонента первого символа становится равной Q-компоненте последнего).

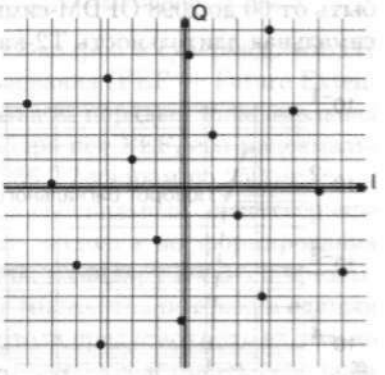


Рис. 4.9. Сигнальное созвездие 16-QAM после поворота

В чем суть поворота сигнального созвездия? В квадратурных каналах передаются проекции точки сигнала на соответствующие оси (синфазную и квадратурную). При обычном сигнальном созвездии несколько точек расположены на нескольких общих ортогональных линиях и их проекции совпадают. После поворота сигнального созвездия (см. рис. 4.9) у каждой точки уникальные Q- и I-координаты. Некоторые из координат оказываются достаточно близко друг к другу, но по одной координате точки всегда можно восстановить другую ее координату. А механизм сдвига Q-координаты приводит к тому, что исходные координаты сигнальной точки оказываются в разных модуляционных символах (т. е. заведомо на разных поднесущих), что существенно снижает вероятность их одновременной деградации как из-за случайных импульсных помех, так и по причине селективных затуханий в канале. В работе [2] отмечено, что применение такой техники обеспечивает операционное усиление 7,6 дБ (рис. 4.10).

После формирования модуляционных символов происходит их перемежение в пределах кодового слова.

Все рассмотренные до сих пор процедуры выполняются параллельно для отдельных магистральных потоков. В результате для каждого PLP формируется последовательность модуляционных символов. Из них необходимо сформировать OFDM-символы. Но если в DVB-T эта процедура была абсолютно прозрачной, то в DVB-T2, из-за возможности транслировать несколько мультимедийных потоков, необходимо сформировать достаточно сложную кадровую структуру.

Кадр физического уровня DVB-T2 (T2-кадр) (рис. 4.11) начинается с преамбулы P1. Это OFDM-символ с модуляцией DBPSK, двумя защитными ин-

тервалами с двух сторон (в сумме 1/2 длительности символа). Он служит для синхронизации, идентификации потока DVB-T2, а также содержит семь информационных бит с начальной информацией о T2-кадре, а именно число номинальных поднесущих в OFDM (1K–32K) и формат передачи следующей за P1 преамбулы P2 (режимы MISO или SISO). Вся остальная информация о T2-кадре (длина, модуляция, скорость кодирования и т. п.) передается в преамбуле P2, которая может занимать несколько OFDM-символов. Далее следует поле данных (информационные OFDM-символы). Замыкает T2-кадр специальный завершающий OFDM-символ. В зависимости от параметров OFDM в T2-кадре может быть от 60 до 2098 OFDM-символов при полосе передачи 8 МГц (табл. 4.2). Максимальная длительность T2-кадра — 250 мс.

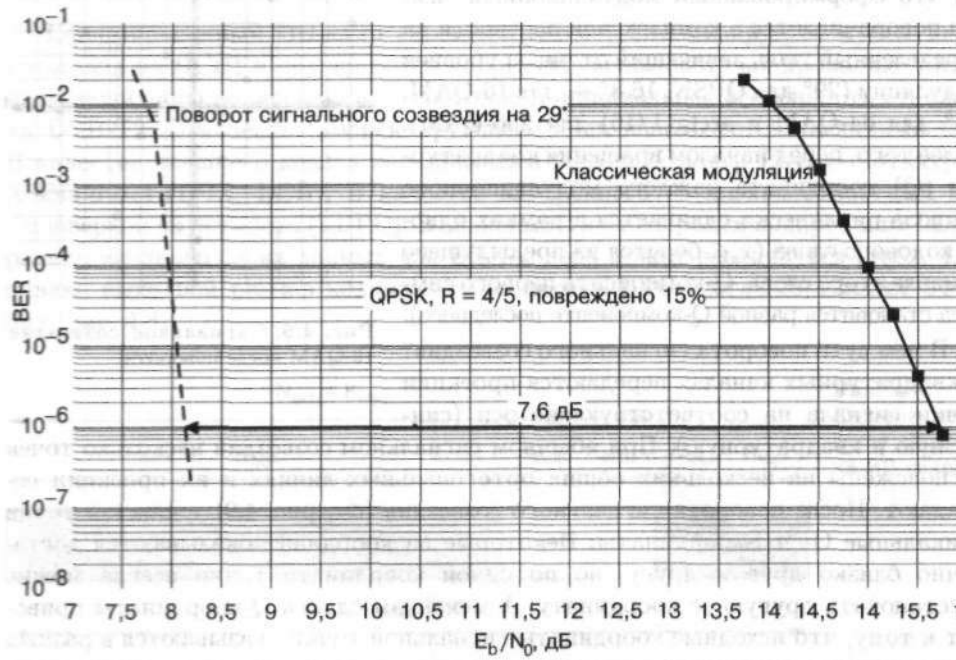


Рис. 4.10. Выигрыш от поворота сигнального созвездия модуляционных символов

Рис. 4.11. Структура кадров в системе DVB-2T

Таблица 4.2. Максимальная длина T2-кадра (OFDM-символов) в полосе 8 МГц

Номинальное число поднесущих	Длительность символа, мс	Защитный интервал						
		1/128	1/32	1/16	19/256	1/8	19/128	1/4
32K	3,584	68	66	64	64	60	60	–
16K	1,792	138	135	131	129	123	121	111
8K	0,896	276	270	262	259	247	242	223
4K	0,448	–	540	524	–	495	–	446
2K	0,224	–	1081	1 049	–	991	–	892
1K	0,112	–	–	2098	–	1 982	–	1 784

T2-кадры объединены в суперкадр. Помимо T2-кадров в суперкадр входят поля, зарезервированные для дальнейшего использования (FEF — Future Extension Frames). Они могут чередоваться в произвольном порядке. Максимальная длительность суперкадра — 128 с. Если в суперкадре нет FEF, его максимальная длительность $T_{SF} = 64$ с, что соответствует 256 T2-кадрам по 250 мс.

Распределением потоков по кадрам занимается специальный диспетчер еще на стадии формирования BB-кадров. Уже тогда, задолго до формирования OFDM-символов, создается сигнальная информация. Стандарт DVB-T2 чрезвычайно гибок с точки зрения мультиплексирования множества потоков в единый трансляционный сигнал. Видимо, разработчики хотели предусмотреть не только все, что знали или могли вообразить, но и то, что даже представить себе пока не могли. Поэтому не будем удивляться столь многообразному и, казалось бы, избыточному набору возможностей.

Формирование OFDM-кадров неразрывно связано с распределением фрагментов различных магистральных потоков как внутри T2-кадра, так и в рамках суперкадра. С этой точки зрения стандарт выделяет три типа потоков PLP — общий, а также потоки данных типа 1 и 2. Общий PLP — это информация, общая для группы из нескольких PLP (например, таблицы программ и сервисов PSI/SI для нескольких транспортных потоков). Потоки PLP типа 1 в T2-кадре не подразделяются на фрагменты — иными словами, в каждом T2-кадре может быть только один фрагмент каждого PLP типа 1. Наконец, потоки типа 2 могут в пределах T2-кадра разделяться на несколько фрагментов (от 2 до 6480), следующих в кадре попеременно (рис. 4.12).

Потоки могут отображаться на T2-кадры по определенным правилам. Например, поток N передается в группах по три смежных T2-кадра, следующих через интервал в один кадр. Более того, перед распределением по T2-кадрам в рамках каждого PLP возможно временное перемежение. Для этого кодовые слова потока PLP после формирования модуляционных символов и их перестановки группируются в так называемые интерливинговые кадры, содержащие динамически изменяющееся целое число кодовых слов. Интерливинговый кадр состоит из одного или нескольких интерливинговых блоков (рис. 4.13). Перемежение символов происходит в пределах всего интерливингового блока. Процедура разбиения на интерливинговые блоки и кадры выполняется на уровне магистральных потоков, с учетом их специфики. Интерливинговые кадры отображаются на кадры физического уровня (T2-кадры) — один в один или один интерливинговый кадр в несколько T2-кадров (см. рис. 4.13). Отметим, что плюс ко всем перечисленным видам перемежения — на уровне бит в кодовых словах, модуляционных симво-

лов, временного интерливинга — используется еще и частотный интерливинг, т. е. перестановка поднесущих в пределах OFDM-символа.

В стандарте DVB-T2 изменения коснулись и структуры OFDM-символов. Увеличено возможное число номинальных поднесущих — помимо 8K добавлены режимы 16K и 32K поднесущих (а также 1K и 4K). Поскольку с увеличением числа поднесущих для OFDM-сигналов спектральная характеристика становится более крутой (рис. 4.14), можно расширить используемый частотный диапазон, не выходя за границы разрешенной спектральной маски. Это позволяет использовать в OFDM-символе больше поднесущих для передачи данных. Такой режим допустимо использовать при 8K, 16K и 32K поднесущих. Эффект от расширенного режима составляет от 1,4% (8K) до 2,1% (32K).

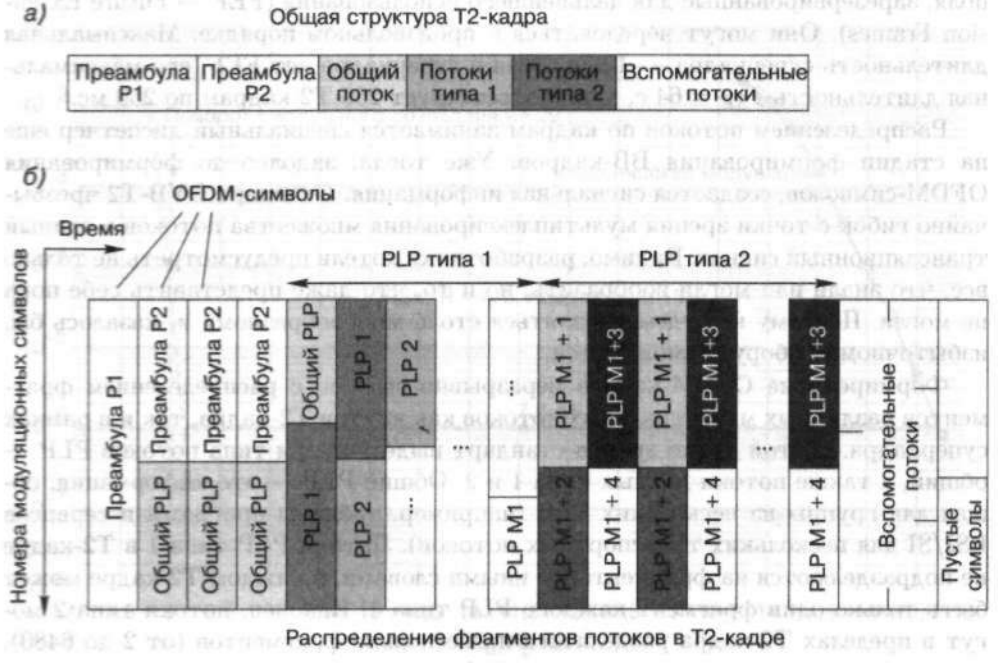


Рис. 4.12. Общая структура T2-кадра *(а)* и распределение фрагментов потоков в T2-кадре *(б)*

Чем больше номинальных поднесущих, тем длительнее может быть OFDM-символ. Это, в свою очередь, позволяет уменьшить защитный интервал до 1/128 (против 1/32 в DVB-T). Использование такого защитного интервала при 32K номинальных поднесущих эквивалентно защитному интервалу 1/32 при 8K поднесущих. Однако пропускная способность при этом возрастает весьма существенно. Всего в DVB-T2 разрешено использовать семь относительных длин защитных интервалов — 1/128, 1/32, 1/16, 19/256, 1/8, 19/128 и 1/4.

В DVB-T2 возможно и более гибкое распределение пилотных поднесущих. Вместо одной фиксированной схемы распределения пилотных частот в DVB-T, в DVB-T2 предусмотрено восемь различных схем их распределения. Выбор варианта зависит от числа номинальных поднесущих и размера защитного интервала. В результате, если в DVB-T распределенные пилотные поднесущие составляли

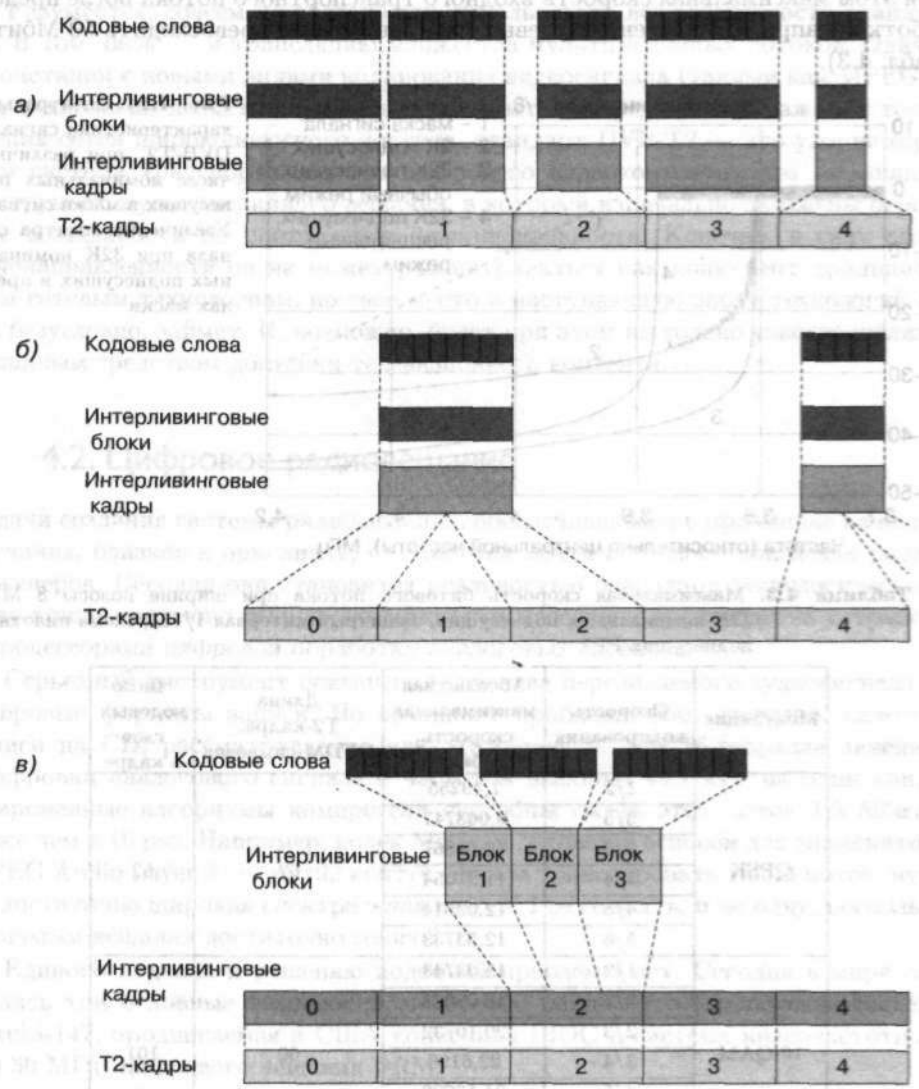


Рис. 4.13. Распределение интерливинговых блоков и кадров различных PLP по T2-кадрам: *а)* — один в один, *б)* — несколько интерливинговых кадров в один T2-кадр (последовательность с пропуском одного кадра), *в)* — три интерливинговых блока в один интерливинговый кадр

8% всех поднесущих, то в DVB-T2 этот показатель может составлять также 1, 2 и 4%.

Еще одна принципиально новая возможность — передача в режиме MISO с использованием схемы Аламоути, т. е. приемник обрабатывает сигнал от двух передающих антенн. Вводятся и дополнительные частотные полосы — 10 и 1,712 МГц (последняя — для мобильных сервисов).

В целом, все эти нововведения позволяют создать очень гибкую и в то же время чрезвычайно эффективную систему трансляции мультимедийных потоков.

При этом максимальная скорость входного транспортного потока после предобработки (например, удаления нулевых пакетов) может превосходить 50 Мбит/с (табл. 4.3).

Рис. 4.14. Спектральная характеристика сигналов DVB-T2 при различном числе номинальных поднесущих и маска сигнала. Увеличения спектра сигнала при 32K номинальных поднесущих в пределах маски

Таблица 4.3. Максимальная скорость битового потока при ширине полосы 8 МГц, 32K номинальных поднесущих, защитный интервал 1/128, схема пилотных поднесущих PP7

Модуляция	Скорость кодирования	Абсолютная максимальная скорость, Мбит/с	Длина T2-кадра, OFDM-символов	Число кодовых слов в кадре
QPSK	1/2	7,49255	62	52
	3/5	9,003747		
	2/3	10,01867		
	3/4	11,27054		
	4/5	12,02614		
	5/6	12,53733		
16-QAM	1/2	15,03743	60	101
	3/5	18,07038		
	2/3	20,10732		
	3/4	22,6198		
	4/5	24,13628		
	5/6	25,16224		
64-QAM	1/2	22,51994	46	116
	3/5	27,06206		
	2/3	30,11257		
	3/4	33,87524		
	4/5	36,1463		
	5/6	37,68277		
256-QAM	1/2	30,08728	68	229
	3/5	36,15568		
	2/3	40,23124		
	3/4	45,25828		
	4/5	48,29248		
	5/6	50,34524		

Разумеется, совсем не обязательно использовать все возможности стандарта. В том числе — и трансляцию множества мультимедийных потоков. Однако в сочетании с новыми видами кодирования видеосигнала (такими как MPEG-4) эта технология является существенным шагом вперед. И что важно с точки зрения сетей широкополосного доступа, стандарт DVB-T2 — это уже не «просто» система транспорта пакетов цифрового видеоконтента. Это — мощный инструмент мультимедийного вещания, в который изначально заложены огромные возможности по расширению функциональности. Конечно, в силу своей однонаправленности он не может рассматриваться как конкурент традиционным сетевым технологиям, но свое место в наступающую эпоху технологий 4G он, безусловно, займет. И, возможно, будет при этом не только узкоспециализированным средством доставки телевизионного контента.

4.2. Цифровое радиовещание

Задачи создания системы радиовещания, обеспечивающей в приемнике качество звучания, близкое к оригиналу, — заветная мечта не одного поколения радиоинженеров. Сегодня она становится реальностью благодаря технологическому чуду конца прошлого века — возможности массового производства устройств с процессорами цифровой обработки аналоговых сигналов.

Серьезный инструмент повышения качества передаваемого аудиосигнала — цифровые форматы записи. Но возникает проблема. Как известно, качество записи на CD, рассматриваемое как эталонное, — это 16-разрядая линейная оцифровка аналогового сигнала с частотой выборки 44,1 кГц на один канал. Современные алгоритмы компрессии способны сжать этот поток 1,5 Мбит/с более чем в 10 раз. Например, кодек Musicam, ставший основой для знаменитого MPEG Audio Layer 2, — до 192 кбит/с. Чтобы транслировать такой поток, нужна достаточно широкая спектральная полоса. Где ее взять, и не одну, поскольку программ вещания достаточно много?

Единого подхода к решению подобной проблемы нет. Сегодня в мире сложились три основные технологии цифрового радио. Это европейская система Eureka-147, продвигаемая в США концепция IBOC и система низкочастотного (до 30 МГц) цифрового вещания DRM.

4.2.1. Система Eureka-147

История европейского стандарта цифрового вещания (Digital Audio Broadcasting — DAB) началась в 1987 году — с официальной даты рождения проекта Eureka-147. Разумеется, работы в области цифрового радио велись и до этого, например, одной из составных частей нового стандарта стала разработка германского института IRT (Institut für Rudfunktechnik) 1981 году. В 1988 году в Женеве на Всемирной административной радиоконференции (WARC — World Administrative Radio Conference) уже демонстрировался первый образец оборудования для мобильного приема. В феврале 1995 года был опубликован европейский стандарт ETS 300 401, его вторая редакция вышла в марте 1997 года. Однако чаще его называют по имени проекта Eureka-147. Впоследствии, в 2000 и 2001 годах, спецификация ETS 300 401 уточнялась. Уже с 1995 года в европейских

странах началось опытное цифровое вещание, с 1997 года — опытная коммерческая трансляция DAB-программ.

Что такое Eureka-147? В его основе [5] два механизма: кодирование аудиосигнала по методу Musicam (MPEG Audio Layer 2, используемое в известных стандартах MPEG 1 и 2) и разделение каналов посредством ортогональных несущих (OFDM), как в уже рассмотренной нами европейской системе цифрового телевидения DVB [6]. Отметим, что с февраля 2007 года действует дополнение DAB+, отличающееся более совершенным аудиокодеком AAC+. Общая схема передачи достаточно стандартна для современных коммуникационных технологий (рис. 4.15): аудиосигнал сжимается MPEG-кодеком, к нему добавляются информационные данные (название программы, курс валюты, прогноз погоды и т. п.) и служебная информация. Вся информация кодируется — сначала выравнивается (рандомизируется) амплитуда сигнала посредством умножения на заданную псевдослучайную последовательность, затем применяется сверточное защитное кодирование, после чего происходит временное перемежение информации. Кодированные таким образом каналы мультиплексируются в один основной сервисный канал. К пакетам этого канала добавляется служебная информация — параметры мультиплексирования, информация о кодировании и т. д. Вместе с пакетами синхронизации вся информация образует кадр передачи, который транслируется посредством нескольких OFDM-символов (терминология стандарта ETS 300 401). OFDM-символ после квадратурной модуляции и ОБПФ представляет собой сигнал, которым можно непосредственно модулировать радионесущую. Полный сигнал эквивалентен низкоскоростной модуляции большого числа (от 1536 до 192) ортогональных несущих. Схема аналогична телевизионному цифровому стандарту DVB.

В полосе 1,536 МГц транслируется цифровой поток порядка 1,5 Мбит/с. Передача возможна в четырех режимах: моно, двухканальном моно, стерео и объединенном стерео (joint stereo). В последнем случае высокочастотные составляющие (выше 2 кГц) стереосигнала передаются в одном потоке, без разделения.

Кодек MPEG Audio Layer 2 предусматривает два режима кодирования аудиосигнала: с частотой выборки 24 и 48 кГц (от 8 до 160 и от 32 до 384 кбит/с). Соответственно с частотой выборки исходный сигнал кодируется в диапазоне 11,3 и 20,3 кГц. Максимальная разрядность выборки 16 бит. В соответствии с алгоритмом весь частотный диапазон делится на 32 одинаковых субканала. В каждом из них разрядность оцифровки сигнала (импульсно-кодовой модуляции — ИКМ) определяется на основе психоакустической модели восприятия, в соответствии с которой восприимчивость уха к различным частотам неодинакова, следовательно, различна и разрядность кодирования. Кроме того, возможен и учет маскирования сигнала в одной полосе сигналом в другой, но это требует достаточно громоздких вычислений. Не углубляясь в механизм кодирования, отметим, что на выходе кодека формируется поток так называемых DAB-аудиокадров, включающих 1152 ИКМ-выборки аудиоинформации, связанные с ней данные и служебную информацию. Общая длительность кадра 24 мс при частоте сканирования 48 кГц. Значения выборок в каждом субканале нормируются с помощью так называемого фактора масштабирования (Scale Factor, SF), одинакового для каждых 12 выборок. Число битов в каждой выборке различно и зависит от номера субканала, выбранной скорости сжатия и частоты сканирования. Число битов в выборке субканала, фактор масштабирования

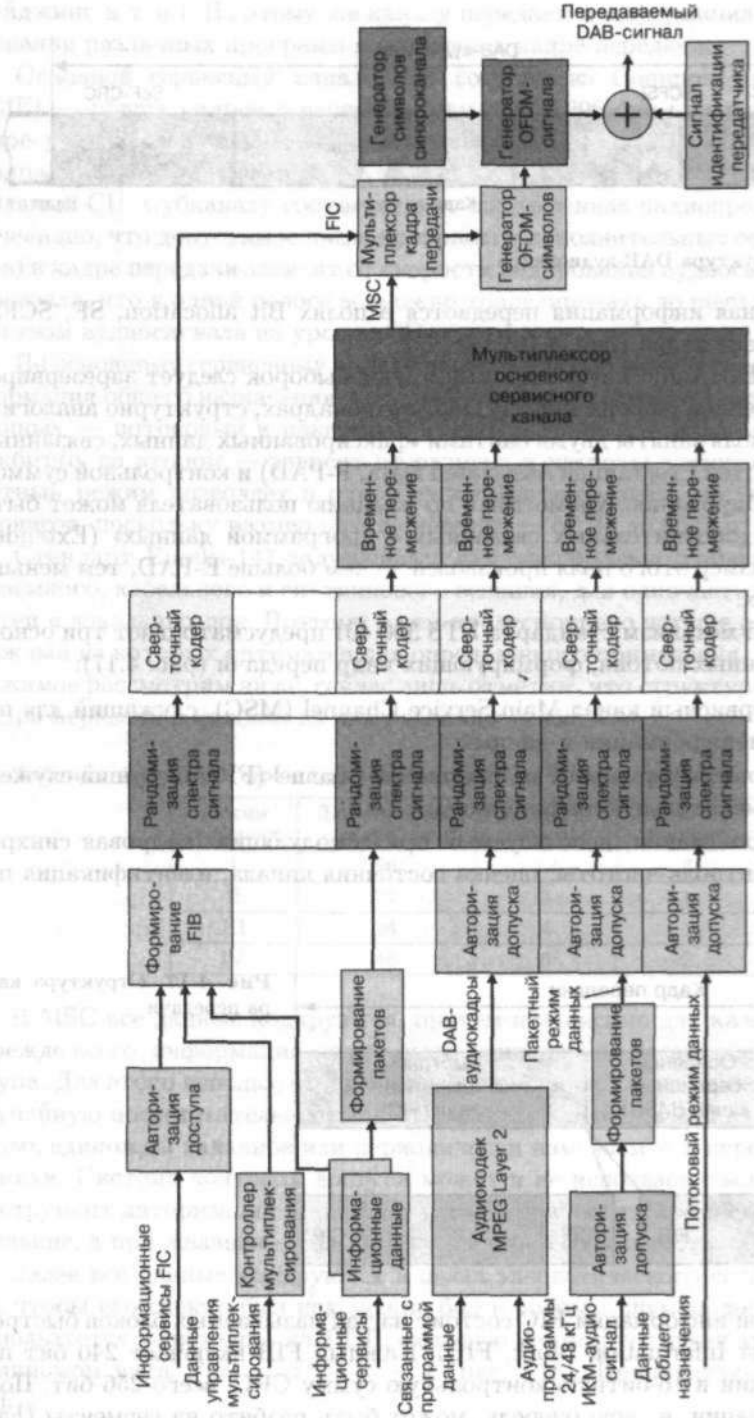


Рис. 4.15. Функциональная схема передающего узла стандарта ETS 300 401 (Eureka-147)

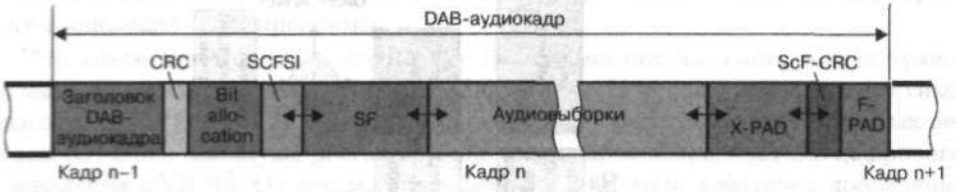


Рис. 4.16. Структура DAB-аудиокадра

и другая служебная информация передается в полях Bit allocation, SF, SCFS, а также в заголовке кадра (рис. 4.16).

В кадрах MPEG Audio Layer 2 за полем аудиовыборок следует зарезервированное поле служебной информации. В DAB-аудиокадрах, структурно аналогичных MPEG, эти поля заняты двумя байтами «фиксированных данных, связанных с программой» (Fixed Programme Associated Data, F-PAD) и контрольной суммой фактора масштабирования. Кроме того, по желанию пользователя может быть добавлено поле «дополнительных связанных с программой данных» (Extended PAD, E-PAD). Размер этого поля произволен — чем больше E-PAD, тем меньше поле аудиоданных.

Транспортный механизм стандарта ETS 300 401 предусматривает три основных информационных потока, формирующих кадр передачи (рис. 4.17):

1) основной сервисный канал Main Service Channel (MSC), служащий для передачи аудиоинформации и данных;
2) канал быстрой информации Fast Information Channel (FIC), несущий служебную и вспомогательную информацию;
3) канал синхронизации, используемый при демодуляции (кадровая синхронизация, контроль частоты, оценка состояния канала, идентификация передатчика).

Рис. 4.17. Структура кадра передачи

Канал быстрой информации FIC состоит из так называемых блоков быстрой информации (Fast Information Block, FIB). Каждый FIB включает 240 бит полезной информации и 16-битную контрольную сумму CRC, всего 256 бит. Поле полезной информации, в свою очередь, может быть разбито на сегменты (Fast Information Groups, FIGs) произвольной длины. Основное назначение FIC — передача многочисленной служебной информации, а также поддержка различных информационных сервисов (индикация времени, радиотекст, меню программ,

пейджинг и т. п.). По этому же каналу передается информация о мультиплексировании различных программ и сервисов в кадре передачи.

Основной сервисный канал MSC состоит из Common Interleaved Frames (CIF) — общих кадров с перемежением, по 55 296 бит каждый. Минимальный адресуемый элемент CIF — Capacity Unit (CU) 64 бит. Всего в CIF 864 CU. CIF подразделяется на субканалы. Каждый из них включает целое число последовательных CU. Субканалу соответствует определенная радиопрограмма (сервис). Очевидно, что допустимое число программ и дополнительных сервисов (субканалов) в кадре передачи зависит от скорости кодирования аудиосигнала. Практика показала, что в одной полосе возможно транслировать до шести программ с качеством аудиосигнала на уровне CD.

По основному сервисному каналу передаются не только аудиоданные, но и информация общего назначения. Стандартом предусмотрено два режима передачи данных — потоковый и пакетный. Скорость данных в первом режиме кратна 8 кбит/с, во втором — зависит от размера и частоты следования пакетов. Пакетный режим позволяет в одном субканале передавать несколько различных сервисов, поскольку размер пакета невелик — от 24 до 96 байт.

Стандарт Eureka-147 задумывался как максимально универсальный — для наземного, кабельного и спутникового вещания, для одночастотных глобальных сетей и локальных зон. Поэтому в нем предусмотрено четыре режима передачи, каждый из которых оптимален для определенного применения. Назначение этих режимов рассмотрим ниже, сейчас лишь отметим, что структура и длительность кадра передачи в каждом из них различны (табл. 4.4).

Таблица 4.4. Структура кадра в зависимости от режима передачи

Режим передачи	Длительность кадра, мс	Число FIB в кадре	Число CIF в кадре
I	96	12	4
II	24	3	1
III	24	4	1
IV	48	6	2

В MSC все данные кодируются, причем независимо для каждого подканала. Прежде всего, информация может быть защищена от несанкционированного доступа. Для этого используется перемножение информации на некоторую псевдослучайную последовательность (ПСП), ключом является 8-байтное контрольное слово, единожды заданное или периодически изменяемое и передаваемое в приемники. Система контроля доступа может и не использоваться. Однако это — инструмент авторизации программ, взимания абонентской платы за их прослушивание, а при желании — за каждый сервис в отдельности.

Далее все данные кодируются в целях энергетического сглаживания сигнала, чтобы его спектр был как можно более ровным, шумоподобным. Для этого используется умножение на ПСП, которую формирует генератор с задающим полиномом вида $P(x) = x^9 + x^5 + 1$ (рис. 4.18). Инициализирующее слово — $1FF_{16}$.

После выравнивающего кодирования начинается сверточное избыточное кодирование, при котором объем информации существенно (на 25–300%) возрастает, а вместе с ним — и степень защиты от ошибок передачи. Схема сверточного

кодера приведена на рис. 4.19. Каждому входному биту соответствуют четыре выходных. Однако используют не обязательно все четыре, поэтому скорость кодирования может варьироваться от 8/9 до 8/32 (число входных/выходных битов). Существенно, что различные информационные поля кодируются с разной скоростью. Кроме того, скорости кодирования аудиоинформации в MSC и данных в FIC различны. В итоге наиболее защищенной оказывается служебная управляющая информация, наименее — поля аудиовыборок, что естественно: потери управляющей информации могут носить фатальный характер. Скорость кодирования в FIC — 1/3, т. е. вместо исходных 256 бит FIB образуются $3 \times 256 = 768$ бит.

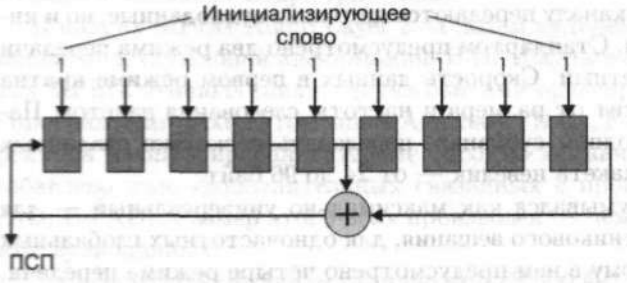


Рис. 4.18. Генератор ПСП для рандомизации спектра сигнала

После сверточного кодирования данные в MSC — и только в нем — подвергаются временному перемежеванию (перемешиванию). В этом канале вся информация структурирована в виде логических кадров — пакетов информации, связанных с исходными блоками данных длительностью 24 мс (DAB-аудиокадры либо блоки данных). Число битов в логическом кадре зависит от стадии и скорости кодирования. Информация из каждых 16 логических кадров перемешивается по определенному закону и распределяется по 16 последовательным пакетам данных. Благодаря этому, если из-за длительной помехи какой-либо пакет пропадет, исчезнет лишь часть информации каждого из 16 логических кадров, которую можно будет восстановить при декодировании.

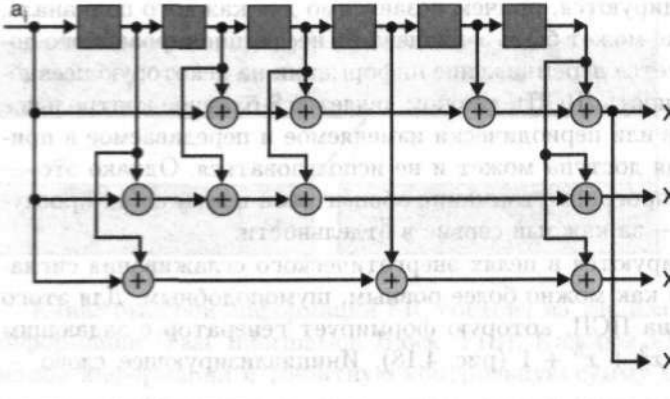


Рис. 4.19. Сверточный кодер

В MSC все процедуры кодирования выполняются независимо для каждого аудиопотока или информационного сервиса. И только после временного перемежевания пакеты мультиплексируются в CIF. Каждому потоку в CIF выделяется

субканал, занимающий целое число 64-битных CU. Информация о схеме мультиплексирования (начальный адрес и размер субканала, скорость кодирования) передается в FIC. После формирования CIF в соответствии с выбранным режимом передачи (см. табл. 4.4) мультиплексор формирует кадр передачи (рис. 4.20).

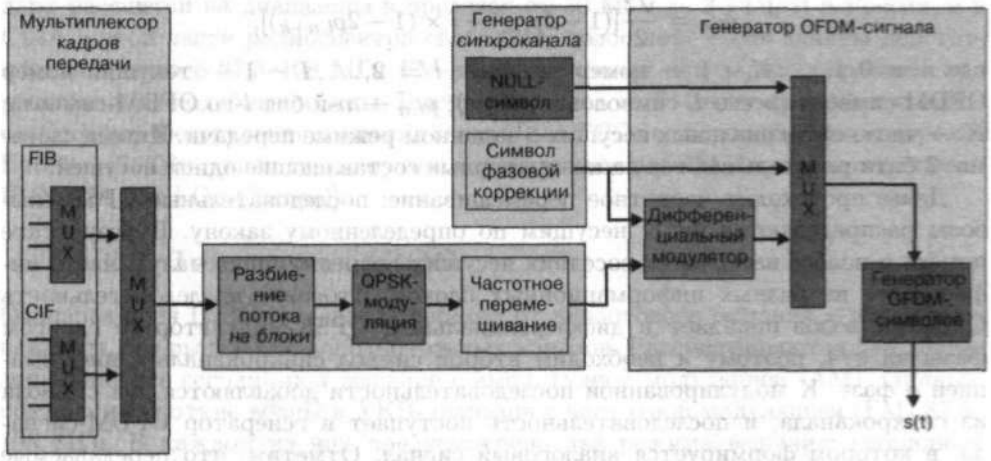


Рис. 4.20. Функциональная схема формирования радиосигнала

Информация кадра передачи транслируется в виде последовательности OFDM-символов. Первые два OFDM-символа кадра принадлежат каналу синхронизации. Это так называемый NULL-символ (уровень сигнала равен 0) и символ, несущий информацию о фазе дифференциальной QPSK-модуляции в следующем OFDM-символе. Информация из FIB и CIF группируется в виде битовой последовательности и разбивается на OFDM-символы. Каждому символу соответствует число битов, вдвое большее, чем число несущих в выбранном режиме передачи (табл. 4.5). Например, для первого режима передачи одному OFDM-символу соответствуют $2 \times 1536 = 3072$ бит. Поскольку размер кодированного FIB — 768 бит, в этом режиме передачи одному OFDM-символу соответствует четыре кодированных FIB, или три OFDM-символа на все 12 пакетов быстрого информационного канала. В основном сервисном канале размер пакета FIC — 55 296 бит, или 18 OFDM-символов на пакет в первом режиме передачи (всего 72 пакета на 4 CIF-кадра передачи).

Таблица 4.5. Параметры режимов передачи Eureka-147

Режим передачи	I	II	III	IV
Длительность кадра, мс	96	24	24	48
Длительность Null-символа, мкс	1297	324	168	648
Длительность защитного интервала, мкс	246	62	31	123
Номинальное максимальное расстояние между передатчиками, км	96	24	12	48
Номинальный частотный диапазон, МГц	До 375	До 1500	До 3000	До 1500
Длительность информационной части символа, мкс	1000	250	125	500
Общая длительность символа, мкс	1246	312	156	623
Число ортогональных несущих	1536	384	192	768

Далее последовательность OFDM-символов кадра, за исключением первых двух из синхроканала, подвергается квадратурной модуляции QPSK. Каждым двум битам двоичного OFDM-символа ставится в соответствие QPSK-символ, который в комплексном виде может быть записан как

$$q_{l,n} = \frac{\sqrt{2}}{2}[(1 - 2p_{l,n}) + i \times (1 - 2p_{l,n+k})],$$

где $n = 0, 1, \ldots, K - 1$ — номер несущей; $l = 2, \ldots, L - 1$ — текущий номер OFDM-символа (всего L символов в кадре); $p_{l,n}$ — n-й бит l-го OFDM-символа; K — число ортогональных несущих в заданном режиме передачи. Иными словами, 2 бита раскладываются на квадратурные составляющие одной несущей.

Далее происходит частотное перемешивание: последовательные QPSK-символы распределяются по K несущим по определенному закону. В результате помеха в полосе нескольких соседних несущих повредит лишь малую часть информации из разных информационных блоков. Итоговая последовательность QPSK-символов попадает в дифференциальный QPSK-модулятор со сдвигом фазы на $\pi/4$, поэтому и необходим второй символ синхроканала с информацией о фазе. К модулированной последовательности добавляются два символа из синхроканала, и последовательность поступает в генератор OFDM-сигнала, в котором формируется аналоговый сигнал. Отметим, что передаваемые OFDM-символы разделены защитным временным интервалом, что избавляет от помех, связанных с многолучевым распространением (переотражениями), неизбежным в городских условиях, — переотраженный сигнал попадает в защитный интервал и не интерферирует со следующим OFDM-символом.

К основному DAB-сигналу может быть добавлен так называемый сигнал-идентификатор передатчика, размещаемый в NULL-символах некоторых кадров передачи. Его основное назначение — передавать информацию о географическом положении передатчика.

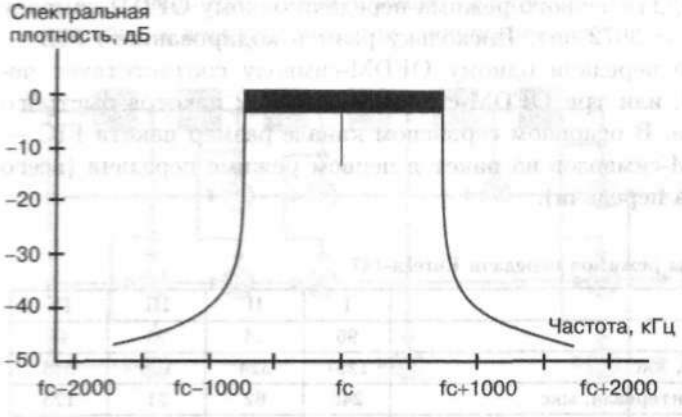


Рис. 4.21. Теоретический спектр DAB-радиосигнала в режиме передачи I

Спектр DAB-сигнала в любом режиме передачи занимает полосу порядка 1,5 МГц (рис. 4.21). Параметры режимов передачи выбраны так, что режим I оптимален для одночастотных наземных сетей вещания в УКВ-диапазоне, режимы II и IV ориентированы для наземной трансляции в УКВ-, а также в L-диапазоне (1452–1492 МГц) в локальных и одночастотных глобальных сетях. Режим III

лучше всего подходит для кабельного и спутникового вещания, а также для гибридных спутниково-наземных сетей.

Несмотря на все достоинства, на пути распространения Eureka-147 встает важнейшая проблема: в каком диапазоне организовывать вещание? Ведь стандарт рассчитан на диапазоны в пределах от 30 МГц до 3 ГГц. И в России, и в США в этой части радиоспектра свободных полос нет — они заняты действующими в полосе 87,5–108 МГц FM-радиостанциями или зарезервированы под военные и им подобные нужды. Быстрого отказа от FM-вещания не предвидится. Поэтому Федеральная комиссия связи США FCC не дала «добро» на внедрение Eureka-147. Там пошли своим путем, развивая систему iDAB на базе технологии IBOC (In Band On Channel).

4.2.2. Технология IBOC

Основная идея IBOC — сделать переход от аналогового вещания к цифровому плавным, без выделения дополнительных каналов. Рассматриваются два диапазона: занятые сегодня под вещание с амплитудной модуляцией (AM) длинные, средние и короткие волны и УКВ-диапазон с частотной модуляцией (FM) 87,5–108 МГц. В каждом из них предусмотрено два режима вещания: гибридный и полностью цифровой. В первом случае происходит одновременное аналоговое и цифровое вещание, во втором — только цифровое.

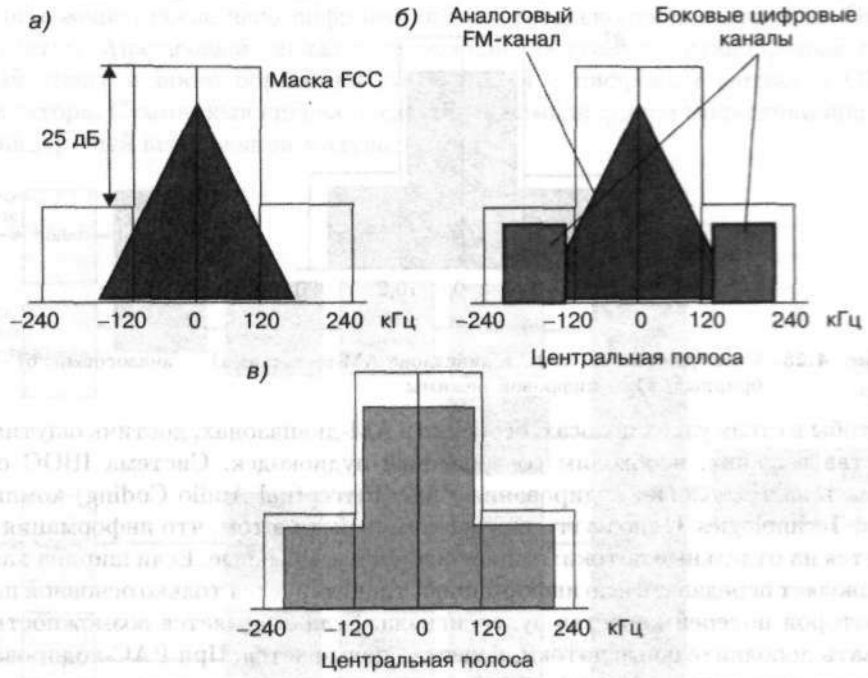


Рис. 4.22. Спектр сигналов IBOC в диапазоне FM-вещания: *а)* — аналоговый; *б)* — гибридный; *в)* — цифровой режимы

В FM-режиме основная полоса каналов занимает около 240 кГц, а две боковые — еще по 120 кГц. По правилам FCC уровень сигнала в боковых полосах

должен быть на 25 дБ ниже, чем в основной (рис. 4.22). В гибридном режиме сигналы цифрового вещания передают в боковых полосах, в цифровом — в основной и боковых. Теоретически возможная скорость цифрового потока в боковых полосах 64 кбит/с, в основной полосе — 96–128 кбит/с.

В АМ-вещании модулируется несущая с частотой 9/10 кГц. Реальная ширина спектральной полосы основного сигнала — порядка 20 кГц, еще по 10 кГц занимают боковые полосы (рис. 4.23). В гибридном режиме полоса цифрового потока перекрывает полосу аналогового вещания, но с меньшим уровнем сигнала. В цифровом режиме сигнал занимает всю полосу в соответствии с разрешенным шаблоном. Скорость передачи в гибридном режиме от 20 до 48 кбит/с, в полностью цифровом — до 64 кбит/с.

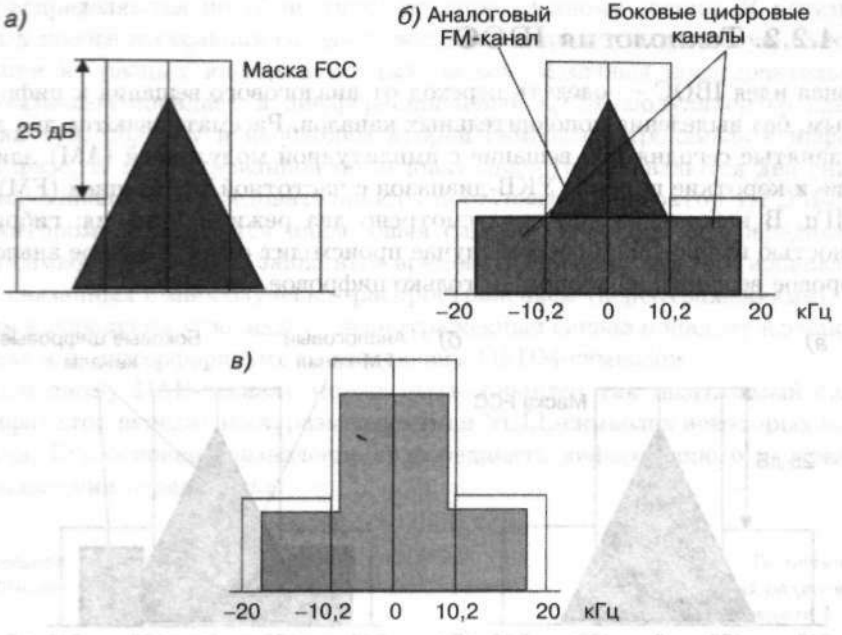


Рис. 4.23. Спектр сигналов IBOC в диапазоне АМ-вещания: *а)* — аналоговый; *б)* — гибридный; *в)* — цифровой режимы

Чтобы в столь узких полосах, особенно в АМ-диапазонах, достичь ощутимого качества вещания, необходим совершенный аудиокодек. Система IBOC основывалась на технологии кодирования PAC (Perceptual Audio Coding) компании Lucent Technologies. Одно из его важнейших свойств в том, что информация разбивается на отдельные потоки: основной и дополнительные. Если ширина канала не позволяет передавать всю информацию, транслируется только основной поток с некоторой потерей качества аудиосигнала. Если появляется возможность передавать дополнительные потоки, качество повышается. При PAC-кодировании при скоростях потока 128–96 кбит/с качество декодированного сигнала мало отличается от оригинала — стереосигнала с CD-диска (16 бит, частота выборки 44,1 кГц). Но такие скорости возможны лишь при полностью цифровом FM-режиме. В АМ-режимах при скорости потока 48 кбит/с качество цифрового вещания примерно соответствует хорошему приему аналогового FM-стерео.

Развитие технологии цифрового вещания в США сегодня сосредоточено в образованной в августе 2000 года корпорации iBiquity Digital. Эта компания родилась в результате совместной двухлетней работы корпораций USADR (USA Digital Radio) и Lucent Digital Radio. USDAR возникла в 1991 году как альянс компаний CBS, Gannett и Westinghouse Electric. С 1998 года она преобразована в независимую компанию. Именно ей принадлежат ключевые патенты на технологию IBOC. Lucent Digital Radio появилась также в 1998 году как подразделение Lucent Technologies в союзе с фирмой Pequot Capital Management. Компания развивала собственную версию DAB на основе технологии аудиокодирования PAC.

Существенная особенность технологии цифрового вещания в США — возможность передавать одну и ту же программу в цифровом и аналоговом виде с некоторым временным сдвигом. Когда цифровой сигнал пропадает, приемник незаметно для слушателя переходит на воспроизведение аналогового сигнала. Это позволяет успешно бороться с так называемым клифф-эффектом, проявляющимся в цифровых вещательных системах в виде периодов неразборчивости сигнала. По мнению специалистов компании iBiquity, переход на полностью цифровое вещание нецелесообразен, пока число IBOC-приемников не составит по меньшей мере 85%.

Система вещания IBOC строится на технологии модуляции с использованием ортогональных несущих COFDM. Структура передатчика приведена на рис. 4.24. Аудиосигнал параллельно транслируется в цифровом и аналоговом виде. В цифровом тракте происходит аудиокодирование, защитное кодирование и перемежение, после чего цифровая последовательность поступает на OFDM-модулятор. Аналоговый сигнал с задержкой поступает в стандартный аналоговый тракт и после обработки смешивается с цифровым сигналом OFDM-модулятора. Суммарный сигнал после фазово-амплитудной модуляции попадает в стандартный передающий модуль.

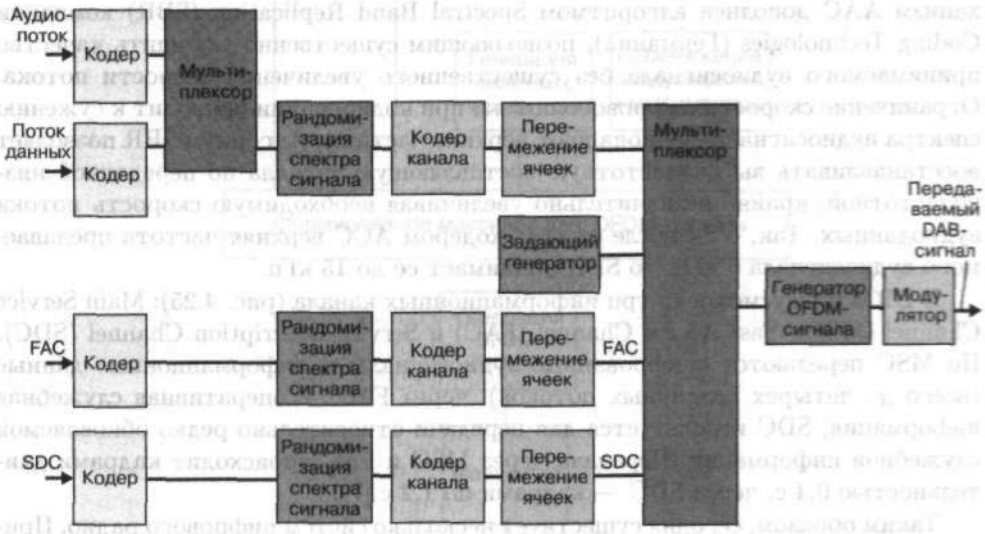


Рис. 4.24. IBOC-передатчик

Основное достоинство американской системы по сравнению с европейской — это более дешевое оборудование, прежде всего приемники. Еще несколько лет назад стоимость DAB-ресивера для Eureka-147 составляла около 1000 долл., в то время как для IBOC-приемников назывались цифры порядка 100 долл. Но ситуация стремительно меняется, сегодня приемники Eureka-147 можно купить чуть больше чем за 30 евро.

4.2.3. Всемирное цифровое радио (DRM)

DRM (Digital Radio Mondiale, всемирное цифровое радио) — это международный консорциум, объединяющий свыше 70 известных исследовательских, производственных и вещательных компаний и организаций всего мира. Он был создан в 1996 году для разработки стандарта цифрового радиовещания в диапазоне ниже 30 МГц. При этом преследуются такие цели, как значительное улучшение качества звука, снижение помех и сохранение зоны обслуживания при значительном (до шести раз) уменьшении мощности передатчика. Поскольку ширина полосы цифрового вещательного сигнала в DRM эквивалентна полосе аналогового, для перехода на цифровое вещание дополнительного спектрального диапазона не требуется. Спецификация DRM признана как ITU (Рекомендация ITU-R BS.1514), так и Европейским институтом стандартов (Спецификация ETSI ES 201 980).

Система DRM также использует OFDM-модуляцию (порядка 200 несущих) с предварительной 16- и 64-позиционной квадратурной амплитудной модуляцией (16- и 64-QAM), защитное сверточное кодирование и временное перемежение информации. Для аудиокодирования принят алгоритм AAC (Advanced Audio Coding) — низкоскоростной алгоритм, применяемый в группе стандартов MPEG-4 (изначально создавался для аудиокодирования в MPEG-2). Кроме того, для кодирования речи предусмотрен низкоскоростной алгоритм CELP. Механизм AAC дополнен алгоритмом Spectral Band Replication (SBR) компании Coding Technologies (Германия), позволяющим существенно улучшить качество принимаемого аудиосигнала без существенного увеличения скорости потока. Ограничение скорости цифрового потока при кодировании приводит к сужению спектра аудиосигнала — пропадают верхние частоты. Алгоритм SBR позволяет восстанавливать высокочастотную составляющую сигнала по переданной низкочастотной, крайне незначительно увеличивая необходимую скорость потока аудиоданных. Так, если после сжатия кодером ACC верхняя частота предаваемого аудиосигнала 6 кГц, то SBR поднимает ее до 15 кГц.

В DRM предусмотрено три информационных канала (рис. 4.25): Main Service Channel (MSC), Fast Access Channel (FAC) и Service Description Channel (SDC). По MSC передаются оцифрованные аудиосигналы и информационные данные (всего до четырех различных потоков), через FAC — оперативная служебная информация, SDC используется для передачи относительно редко обновляемой служебной информации. Передача через MSC и FAC происходит кадрами длительностью 0,4 с, через SDC — кадрами по 1,2 с [7].

Таким образом, сегодня существует несколько систем цифрового радио. Причем, если Eureka-147 ориентирована на широкополосные одночастотные трансляционные сети и в ряде случаев требует пересмотра сложившегося распределения частот, то системы DRM и IBOC изначально предназначены для работы в уже

существующих диапазонах. Недаром наряду с развитием систем Eureka-147 многие европейские производители и радиовещатели (например, BBC) уделяют серьезное внимание DRM. Действительно, традиционное AM-вещание в диапазоне средних и коротких волн умирает, поскольку никого не устраивает по качеству. Цифровое же вещание позволит использовать достоинства распространения волн в этих диапазонах с качеством, конечно, не CD, но уж по крайней мере традиционного FM-вещания.

Рис. 4.25. Общая схема передачи данных DRM

WACR-92 отвела полосу частот 1452–1492 МГц для спутникового цифрового радиовещания (кроме США, России, Белоруссии и Украины). В этой же полосе возможно создание и наземных сетей. Пока эти частоты не могут быть использованы на первичной основе. Система же Eureka-147 сегодня развивается весьма активно и не только в Европе, но и в Австралии, Канаде и Израиле. Свыше 500 млн. человек во всем мире могут принимать более 1000 различных DAB-сервисов. Отметим, что кроме собственно передачи звуковых программ DAB Eureka-147 может обеспечить и различные информационные сервисы, в том числе низкоскоростное мобильное видео (MPEG-2), трансляцию различных картинок (вспомним, что явилось толчком и экономическим фундаментом развития Интернета), различную текстовую новостную и справочную информацию. Все это делает цифровое радиовещание, аналогично цифровому телевизионному вещанию, значимым игроком в мире широкополосной передачи информации — по крайней мере, в области информационных услуг и мобильного видео. Хотя DAB не интерактивна, но разнообразие сервисов DAB в известной мере компенсирует этот недостаток. Причем лицензии на DAB-вещание, например, в Великобритании (порядка 50 тыс. фунтов в 2005 году), несопоставимо дешевле миллиардных 3G-лицензий сотовой связи в тот же период.

Литература

1. Frame structure channel coding and modulation for a second generation digital terrestrial television broadcasting system (DVB-T2). — DVB Document A122, June 2008.
2. Мейтин М. MPEG как предвестник эры интерактивного телевидения. — Электроника: НТБ, 2001, № 4.
3. Варгузин В., Артамнов А. Сравнительные характеристики европейского и американского стандартов цифрового наземного телевидения. — Теле-Спутник, 1999, № 11.
4. Кухарев В. Перспективы развития и преимущества DVB-T в России и странах ближнего зарубежья. — www.nat.ru.
5. ETS 300 401. Radio Broadcasting Systems; Digital Audio Broadcasting (DAB) to mobile, portable and fixed receivers. — European Telecommunications Standards Institute, 2001.
6. Конкурирующие стандарты цифрового телевизионного вещания. — Электроника: НТБ, 2001, № 1, с. 17–19.
7. Draft New Recommendation ITU-R BS. System for Digital Sound Broadcasting in the Broadcasting Bands Below 30 Mhz. — ITU, Document 6/63-E, 25 October 2000.
8. ETSI EN 300 744 V1.6.1 (2009-01). Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital terrestrial television. — European Telecommunications Standards Institute, 2009.

ГЛАВА 5

БЕСПРОВОДНЫЕ ЛОКАЛЬНЫЕ СЕТИ СТАНДАРТОВ IEEE 802.11

5.1. Локальные сети под управлением IEEE 802.11

22 мая 1973 года. Роберт Меткалф, сотрудник исследовательского центра компании Xerox в Пал-Альто, подал своему руководству докладную записку, в которой впервые ввел слово «ethernet» и изложил основные принципы работы новой локальной компьютерной сети, воплощенные позднее в стандарте IEEE 802.3, именуемом сегодня Ethernet'ом. Интересно, выбирая название для новой технологии, автор сознательно остановился на словосочетании ether net — «эфирная сеть», предвидя, что через четверть века Ethernet устремится в эфир?

Беспроводные локальные сети передачи информации (WLAN) развиваются в последние 15 лет невероятно быстро. Простота развертывания таких сетей ограничена только необходимостью оформления разрешительной документации (в тех странах, где это требуется). По пропускной способности они не уступают выделенным медным линиям. Помехоустойчивость, надежность и защищенность современных протоколов передачи сделали WLAN явлением повсеместным, а оборудование для них — массовым продуктом. Отметим, что понятие «локальные сети передачи информации» достаточно условно. Как правило, имеются в виду системы, локализованные в радиусе сотни метров. Однако технологии локальных сетей с успехом применяют и на расстояниях до нескольких десятков километров.

Рынок массовых устройств WLAN достаточно молод. Первые устройства для беспроводных локальных сетей появились в начале–середине 1990-х годов. Но уже в 1999 году объем продаж оборудования WLAN достиг 600–770 млн. долл., а к 2004 году. он составил порядка 2,2–3 млрд. долл. (в месяц продавалось около миллиона адаптеров только стандарта IEEE 802.11b). По экспертным оценкам, к 2003 году. их было установлено свыше 20 млн. (рис. 5.1). Причем стремительно развивалась сама технология передачи и оборудование — скорости выросли от 1–2 до 54 Мбит/с, затем перешагнули и барьер в 100 Мбит/с. С неменьшей стремительностью падала и стоимость оборудования, составляя уже десятки долларов для конечных пользовательских устройств. Несмотря на это, объем продаж ИС для беспроводных сетей возрастет: если в 2002 году он составлял 471 млн. долл., к 2006 году он превысил миллиардную отметку. А в 2008 году составил 3 млрд. долл. — речь идет именно о чипсетах, без учета универсальных компонентов, используемых в устройствах Wi-Fi (например, ИС малошумящих усилителей). Аналитики компании ABI Research ожидают, что в 2011 году будет продано свыше миллиарда чипсетов для Wi-Fi, а к 2013 году объем продаж

превысит 1,6 млрд. штук. А объем всего рынка Wi-Fi в 2008 году составил, по данным агентства In-Stat, 4,83 млрд. долл. — это на 5% больше, чем в 2007 году, невзирая на кризис.

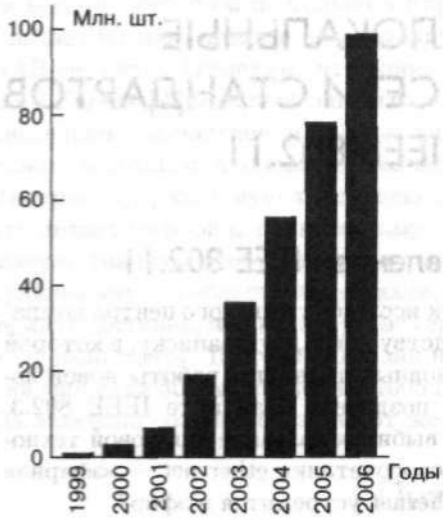


Рис. 5.1. Динамика рынка устройств для сетей IEEE 802.11. По материалам компании IC Insights

Работы над единым стандартом локальных БСПИ начались в 1989 году, когда была организована рабочая группа 11-го комитета IEEE 802. В июле 1997 года в результате работы этой группы был опубликован стандарт IEEE 802.11 «Спецификация физического уровня и уровня контроля доступа к каналу передачи беспроводных локальных сетей» («Wireless LAN Medium Access Control and Physical Layer Specifications»). Он определял архитектуру сети и вытекающие из этого требования к функциям устройств, принципы доступа устройств к каналам связи, формат пакетов, способы аутентификации и защиты данных. Хотя стандарт изначально задумывался как инвариантный по отношению к какому-либо частотному диапазону, на физическом уровне он определял три способа работы: два радиочастотных и оптический. В инфракрасном диапазоне предусматривалась импульсно-позиционная модуляция, в диапазоне 2,400–2,4835 ГГц — режимы модуляции с расширением спектра методом частотных скачков (FHSS) и методом прямой последовательности (DSSS). Скорости обмена устанавливались на уровне 1 и 2 Мбит/с.

Отметим, что устройства, соответствующие исходной спецификации IEEE 802.11, не успели получить развития. Пропускная способность проводных сетей Ethernet сильно возросла, и максимальная скорость передачи 2 Мбит/с, предусмотренная в IEEE 802.11, не удовлетворяла пользователей. Проблему поначалу решило появление стандартов (дополнений) IEEE 802.11b, 802.11a и 802.11g.

Первым стал утвержденный 16 сентября 1999 году стандарт IEEE 802.11b. Он описывал физический и MAC-уровни беспроводных сетей для работы в диапазоне 2,4 ГГц. Стандарт определял работу на скоростях 1 и 2 Мбит/с с модуляцией только методом DSSS. Самое же главное — он предусматривал скорости обмена до 11 Мбит/с (а опционально — и до 33 Мбит/с). Передача данных на скоростях 5,5 и 11 Мбит/с происходит посредством модуляции комплементарных кодовых последовательностей CCK (основной вид модуляции). Кроме того, предусматривалась и работа на скоростях 22 и 33 Мбит/с посредством пакетного бинарного сверточного кодирования (PBCC).

Стандарт IEEE 802.11a, описывающий работу в диапазоне 5 ГГц, был принят одновременно с IEEE 802.11b. В нем использован принципиально иной, чем в IEEE 802.11b, механизм модуляции/мультиплексирования, а именно частотное мультиплексирование посредством ортогональных несущих (OFDM). Данный

метод, в частности, достаточно хорошо зарекомендовал себя в системах цифрового телевизионного вещания DVB. В конце 1999 года были закончены основные работы по созданию европейского 5-ГГц стандарта беспроводных сетей HiperLan2 (HiperLan type 2), который так и не получил массового развития. В июне 2003 года был утвержден высокоскоростной (до 54 Мбит/с) стандарт в диапазоне 2,4 ГГц — IEEE 802.11g.

Сегодня близок к завершению стандарт IEEE 802.11n, описывающий сети со скоростью обмена свыше 100 Мбит/с на основе технологии антенных систем MIMO. Разрабатывается мобильная версия стандарта (IEEE 802.11p) и дополнение, предназначенное для предоставления гарантированного качества связи (QoS), — IEEE 802.11e.

В 2007 году был выпущен обобщенный стандарт IEEE 802.11-2007 [1], в который вошли все стандарты, завершенные к июню 2007 года К ним относятся уже упоминавшиеся стандарты IEEE 802.11a,b,g, а также дополнения IEEE 802.11e/h/i/j.

Стандарт IEEE 802.11 непрерывно совершенствуется и развивается в направлении предоставления пользователям новых сервисов, повышения скорости и качества передачи информации. В 2009 году планируется выпуск целого ряда новых стандартов, работа над которыми началась в 2003–2004 годах. В первую очередь к ним относятся стандарты IEEE 802.11n [2, 3] и IEEE 802.11s [4–6]. Несмотря на то что указанные стандарты находятся в стадии завершения, многие фирмы начали выпуск оборудования, а операторы — эксплуатацию беспроводных сетей на базе черновых версий этих стандартов.

К другим стандартам, утверждение которых ожидают в 2009 году, относятся:

- стандарт IEEE 802.11u [7], описывающий способы взаимодействия сетей, функционирующих под управлением IEEE 802.11, с внешними сетями;
- стандарт IEEE 802.11r [8], регламентирующий процедуры переключения между базовыми станциями для критичных к задержкам приложений, таких как IP-телефония и т. д.;
- стандарт IEEE 802.11p, предназначенный для работы в условиях быстроизменяющихся характеристик среды, в частности, для быстродвижущихся беспроводных устройств [9];
- стандарт IEEE 802.11v, описывающий протоколы управления беспроводной сетью [10];
- стандарт IEEE 802.11w, регламентирующий методы защиты управляющих кадров в беспроводной сети [11];
- стандарт IEEE 802.11z, описывающий протокол обмена данными между станциями напрямую без участия точки доступа [12].

Отметим также стандарт IEEE 802.11k, не вошедший в обобщенный стандарт IEEE 802.11-2007, так как его окончательная версия появилась в конце 2007 года. Указанный стандарт регламентирует механизмы обмена информацией о радиоресурсах, производительности радиоканалов, уровне помех, загрузке каналов и т. д. [13].

Непрерывный рост объемов передаваемой информации, появление новых приложений, таких как передача видеоконтента высокой четкости и др., предъявляют все более высокие требования к пропускной способности беспроводных сетей.

Несмотря на весьма высокие скорости передачи информации в сотовых сетях третьего поколения по технологии LTE и сетях IEEE 802.11n (до 300 Мбит/с) работы по созданию новых технологий в рамках стандарта IEEE 802.11 продолжаются. С 2007 года началась разработка стандарта IEEE 802.11 VNT (Very High Throughput), на базе которого будут реализовываться сверхвысокоскоростные локальные беспроводные сети с номинальной скоростью передачи информации до 500 Мбит/с в частотном диапазоне ниже 6 ГГц. Завершение разработки стандарта планируется в 2012 году.

Высокая пропускная способность сети достигается путем использования технологии MIMO с 8 разнесенными антеннами на приемной и передающей стороне, расширением (по сравнению с другими стандартами IEEE 802.11) полосы частот до 80 МГц путем объединения четырех каналов шириной 20 МГц, использования OFDMA для организации множественного доступа к каналу с разделением по ортогональным частотам, как в протоколе IEEE 802.16. В разрабатываемом стандарте предусматривается совместимость с существующими устройствами, работающими под управлением протоколов IEEE 802.11a/b/g/n.

В Российской Федерации ведется разработка новой технологии и аппаратно-программных средств сверхвысокоскоростных mesh-сетей, функционирующих в частотном диапазоне 60 ГГц [14]. По сравнению с существующими mesh-сетями предлагаемый подход обеспечит не только скорости передачи информации свыше 1000 Мбит/с, но и отсутствие необходимости частотного планирования, работу в дуплексном режиме и т. д.

5.2. Основные принципы IEEE 802.11

Рассмотрим подробнее, что представляет собой стандарт IEEE 802.11 [17] как базовый для всех последующих спецификаций. Как и все стандарты комитета IEEE 802, в документе IEEE 802.11 рассматриваются два нижних уровня модели взаимодействия открытых систем (OSI): физический и канальный (Data Link layer). Причем последний подразделяется на два подуровня. Верхний — Logical Link Control (LLC) — описан в стандарте IEEE 802.2. Стандарт IEEE 802.11 рассматривает только нижний подуровень — Medium Access Control (MAC), т. е. управление доступом к каналу (к среде передачи). Иными словами, на физическом уровне стандарт определяет способ работы со средой передачи, скорость и методы модуляции. На MAC-уровне — принцип, по которому устройства используют (делят) общий канал, способы подключения устройств к точкам доступа и их аутентификации, механизмы защиты данных. Поскольку стандарт IEEE 802.11 разрабатывался как «беспроводной Ethernet», он предусматривает пакетную передачу с 48-битными адресами пакетов, как и любая сеть Ethernet. Комитет IEEE 802 особое внимание уделял совместимости всех своих стандартов, в результате проводные и беспроводные сети IEEE 802 легко сопрягаются друг с другом.

Когда речь заходит о радиотракте, ключевой вопрос — частотный диапазон. IEEE 802.11 привязан к существующим в США и ряде других стран безлицензионным частотным диапазонам. Изначально он был ориентирован на диапазон 2,400–2,4835 ГГц с шириной полосы 83,5 МГц. Определяемая стандартом спектральная маска для одного канала приведена на рис. 5.2 (мощность

отсчитывается относительно пиков функции $\sin(x)/x$). Ширина канала по уровню −30 дБ составляет 22 МГц, следовательно, в полосе 83,5 МГц возможно три неперекрывающихся канала.

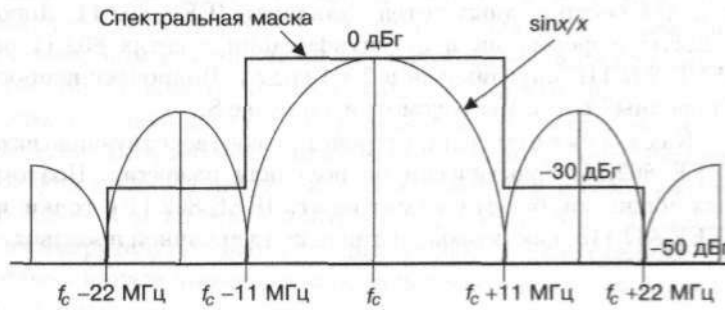


Рис. 5.2. Спектральная маска канала сети 802.11 при модуляции методом DSSS

Стандарт предусматривает два основных способа организации локальной сети: по принципу «равный с равным» (ad-hoc-сеть — рис. 5.3, *а*) и в виде структурированной сети (рис. 5.3, *б*).

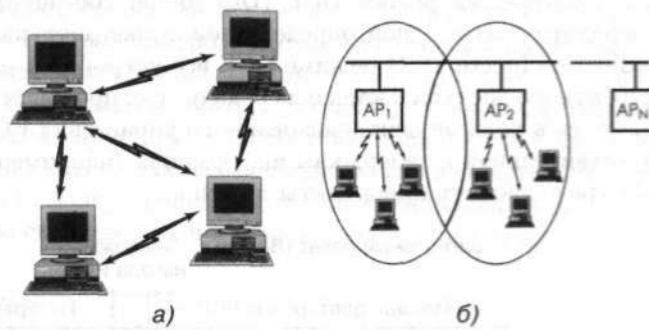


Рис. 5.3. Архитектура сети 802.11: *а)* — ad-hoc-сеть; *б)* — структурированные сети

В первом случае связь устанавливается непосредственно между двумя станциями, и никакого администрирования не предусмотрено. В случае структурированных сетей (а как показала практика, это основной способ построения сетей IEEE 802.11) в их составе появляется дополнительное устройство — точка доступа (AP — Access Point), как правило, стационарная и действующая на фиксированном канале. Связь между устройствами происходит только через AP. Через них же возможен выход во внешние проводные сети. В сети IEEE 802.11 может быть несколько AP, объединенных проводной сетью Ethernet. Фактически такая сеть представляет собой набор базовых станций с перекрывающимися зонами охвата. Стандарт IEEE 802.11 допускает перемещения устройств из зоны одной AP в зону другой (роуминг), тем самым обеспечивая мобильность. Поскольку для мобильных станций важен вопрос ресурса элементов питания, в стандарте заложен специальный протокол управления энергопотреблением — непосредственно при обмене передающее устройство может перевести приемник в режим ожидания.

Важнейшее требование к стандартам беспроводной связи — безопасность передачи данных. В связи с этим на MAC-уровне предусмотрен механизм защиты данных, включающий аутентификацию станций и собственно шифрование

передаваемых данных. Этот механизм должен обеспечивать такой же уровень защиты, как и в обычных сетях Ethernet, поэтому его назвали WEP (Wired Equivalent Privacy — эквивалент проводной конфиденциальности). В дальнейшем были разработаны усовершенствованные алгоритмы и протоколы безопасности для беспроводных сетей стандарта IEEE 802.11. Дополнительные методы защиты информации и аутентификации в сетях 802.11 описаны в стандарте IEEE 802.11i (опубликован в 2004 году.). Подробнее вопросы безопасности беспроводных сетей мы рассмотрим в главе 8.

Как мы уже отмечали, устройства, соответствующие исходной спецификации IEEE 802.11, практически не получили развития. Поэтому далее, без умаления общности, будем рассматривать IEEE 802.11 с точки зрения спецификации IEEE 802.11b, как первой, активно поддержанной производителями аппаратуры.

5.3. MAC-уровень стандарта IEEE 802.11

Стандарт IEEE 802.11 предусматривает два режима управления сетью: когда функции управления распределены между всеми устройствами сети IEEE 802.11 — так называемый режим DCF (Distributed coordination function), и когда они сосредоточены в одной определенной точке доступа — режим PCF (Point coordination function). В режиме DCF все устройства работают по принципу конкурентного доступа к каналу передачи, т. е. приоритетов не существует. Необходимость в режиме централизованного управления PCF возникает при передаче чувствительной к задержкам информации (например, видеопотоков), когда необходимо вводить приоритеты доступа.

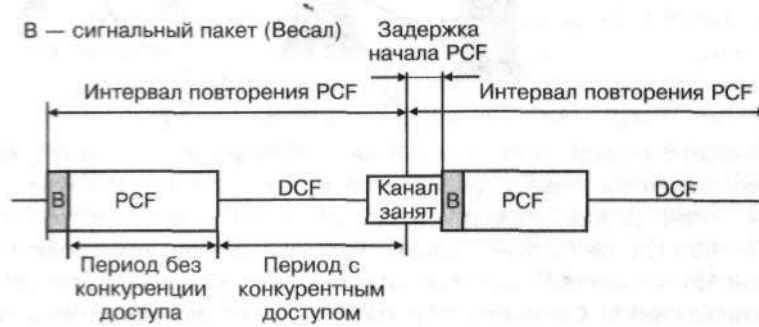


Рис. 5.4. Циклы работы сети в режимах с централизованным (PCF) и распределенным (DCF) управлением

Работа в режиме PCF может происходить только под управлением специальной точки доступа, называемой точкой координации (PC), и только в определенные, периодически повторяющиеся интервалы. Когда сеть переходит в режим PCF, в трафике появляются интервалы, в которых конкурентный доступ отменен, и весь обмен происходит под управлением координирующего устройства (PC) (рис. 5.4). По завершении такого интервала сеть возвращается в режим DCF. Интервалы под управлением PC следуют через строго определенный период, в начале каждого интервала PC выставляет особый сигнальный кадр (Beacon). PC не может передать очередной сигнальный кадр до тех пор, пока

канал не освободится, т. е. очередной «свободный от конкуренции» интервал может начаться с задержкой.

Режим PCF важен для передачи регулярно повторяющейся чувствительной к задержкам информации. Он также эффективен, если сети IEEE 802.11 используются в качестве среды доступа к Интернету (или иным глобальным сетям), т. е. обеспечивают обмен данными между пользователями и централизованным провайдером [16]. Однако основной принцип сетей Ethernet — это все же произвольный конкурентный доступ, что и делает их столь простыми в реализации и эксплуатации. В проводных сетях Ethernet используется механизм множественного доступа к каналу связи с контролем несущей и обнаружением конфликтов (CSMA/CD — Carrier Sense Multiple Acsses with Collision Detection). Станция может начать передачу, только если канал свободен. Если станции обнаруживают, что на одном канале пытаются работать несколько станций, все они прекращают передачу и пытаются возобновить ее через случайный промежуток времени. Таким образом, даже при передаче устройство должно контролировать канал, т. е. работать на прием.

То, что относительно просто при проводной связи, проблематично в беспроводных коммуникациях — затухание сигнала в эфире намного сильнее, чем в проводе. Поэтому возникают две основные проблемы. Во-первых, весьма сложна, если вообще разрешима, задача контроля несущей передающим устройством (когда оно вещает, то собственный сигнал заведомо намного мощнее, чем сигнал удаленного устройства). Во-вторых, возможна ситуация, когда два устройства (A и B) удалены и не слышат друг друга, однако оба попадают в зону охвата третьего устройства C (рис. 5.5) — так называемая проблема скрытых станций. Если оба устройства, A и B, начнут передачу, то они принципиально не смогут обнаружить конфликтную ситуацию и определить, почему пакеты не проходят.

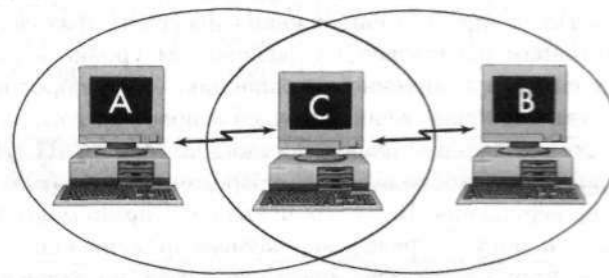


Рис. 5.5. Иллюстрация проблемы скрытых станций

Для устранения подобных проблем в спецификации IEEE 802.11 принят механизм CSMA/CA (Carrier Sense Multiple Acsses with Collision Avoidance) — множественный доступ с контролем несущей и предотвращением коллизий. Перед началом передачи устройство слушает эфир и дожидается, когда канал освободится. Канал считается свободным при условии, что не обнаружено активности в течение определенного промежутка времени — межкадрового интервала (IFS) определенного типа. Если в течение этого промежутка канал оставался свободным, устройство ожидает еще в течение случайного времени отсрочки и, если канал еще не занят, передает пакет. Если пакет предназначен конкретному устройству (не широковещательная или многоадресная передача), то приемник, успешно приняв пакет, посылает передатчику короткий кадр подтверждения

получения ACK (ACKnowledge). Если передатчик не принял ACK, он считает посланный пакет утерянным и повторяет процедуру его передачи.

Примечательно, что, если устройство повторно передает пакет, для определения незанятости канала оно должно использовать увеличенный межпакетный интервал (EIFS). Кроме того, время отсрочки выбирается случайным образом на некотором интервале. При первой попытке передачи этот интервал минимален. При каждой последующей он удваивается до тех пор, пока не достигнет заданного предельного значения. Эти меры приводят к тому, что устройство, успешно передавшее пакет, имеет преимущества в захвате канала (кто ошибается, тот дольше ждет).

Перед первой попыткой получить доступ к каналу устройство загружает длительность случайного интервала отсрочки в специальный счетчик. Его значение декрементируется с заданной частотой, пока канал свободен. Как только счетчик обнулится, устройство может занимать канал. Если до обнуления счетчика канал занимает другое устройство, счет останавливается, сохраняя достигнутое значение. При следующей попытке отсчет начинается с сохраненной величины. В результате неуспевший в прошлый раз получает больше шансов занять канал в следующий. В проводных сетях Ethernet подобного механизма нет.

Однако описанные процедуры доступа не избавляют от проблемы скрытых станций. Для ее преодоления используются два дополнительных кадра: RTS (Request to Send — запрос на передачу) и CTS (Clear to Send — подтверждение готовности). Устройство, желающее отправить пакет данных, передает адресату короткий кадр RTS. Если приемное устройство готово к приему, оно выставляет передающему ответный кадр — CTS. Далее в соответствии с описанной выше процедурой передающее устройство отправляет кадр с данными и дожидается подтверждения ACK.

Стандарт IEEE 802.11 предусматривает два механизма контроля за активностью в канале (обнаружения несущей): физический и виртуальный. Первый механизм реализован на физическом уровне и сводится к определению уровня сигнала в антенне и сравнению его с пороговой величиной. Виртуальный механизм обнаружения несущей основан на том, что в передаваемых кадрах данных, а также в управляющих кадрах ACK и RTS/CTS содержится информация о времени, необходимом для передачи пакета (или группы пакетов) и получения подтверждения. Все устройства сети принимают информацию о текущей передаче и могут определить, сколько времени канал будет занят, т. е. устройство при установлении связи всем сообщает, на какое время оно резервирует канал.

Как мы уже говорили, весь обмен в сетях IEEE 802.11 происходит посредством отдельных кадров (frames). По их структуре особенно четко видно разделение на физический и MAC-уровни. Фактически кадр формируется на MAC-уровне, на физическом уровне к нему добавляется заголовок физического уровня (PLCP). На MAC-уровень пакеты передаются от приложений верхнего уровня. Если их размер превышает максимально допустимый в IEEE 802.11, происходит дефрагментация — большой пакет разбивается на несколько меньших, которые передаются по специальной процедуре.

Кадры MAC-уровня могут быть трех типов: кадры данных, контрольные (ACK, RTS, CTS и т. п.) и кадры управления (например, Beacon). Их структура одинакова (рис. 5.6). Каждый MAC-кадр включает MAC-заголовок, поле дан-

ных (Frame Body) и контрольную сумму CRC. В заголовке передается полная информация о версии протокола стандарта группы IEEE 802.11, типе кадра, системе защиты и т. д. (поле Frame Control); длительности процедуры передачи пакета (Duration/ID), адреса получателя/отправителя (Address 1–4; четыре адресных поля необходимы, если пакеты передаются из подсети одной точки доступа в подсеть другой) и информация о последовательности связанных пакетов (Sequence Control). Поле данных может быть различной длины или вовсе отсутствовать (в контрольных кадрах).

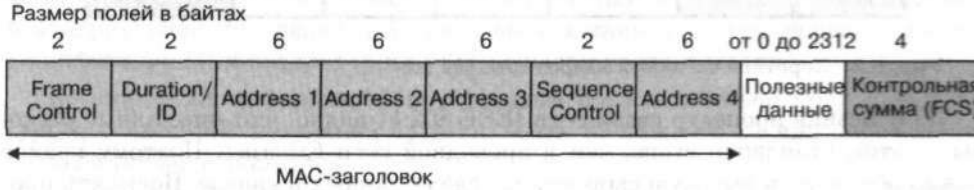


Рис. 5.6. Структура кадров MAC-уровня сети 802.11

5.4. Физический уровень стандарта IEEE 802.11b

На физическом уровне к MAC-кадрам (MPDU) добавляется заголовок физического уровня, состоящий из преамбулы и собственно PLCP-заголовка (рис. 5.7). Преамбула содержит стартовую синхропоследовательность (SYNC) для настройки приемника и 16-битный код начала кадра (SFD) — число $F3A0_{16}$. PLCP-заголовок включает поля SIGNAL (информация о скорости и типе модуляции), SERVICE (дополнительная информация, в том числе о применении высокоскоростных расширений и PBSS-модуляции) и LENGTH (время в микросекундах, необходимое для передачи следующей за заголовком части кадра). Все три поля заголовка защищены 16-битной контрольной суммой CRC.

Рис. 5.7. Структура кадров физического уровня сети 802.11

В стандарте IEEE 802.11b предусмотрено два типа заголовков: длинный и короткий (рис. 5.8). Они отличаются длиной синхропоследовательности (128 и 56 бит), способом ее генерации, а также тем, что символ начала кадра в коротком заголовке передается в обратном порядке. Кроме того, если все поля длинного заголовка передаются со скоростью 1 Мбит/с, то при коротком заголовке преамбула транслируется на скорости 1 Мбит/с, другие поля заголовка — со скоростью 2 Мбит/с. Остальную часть кадра можно передавать на любой из

допустимых стандартом скоростей передачи, указанных в полях SIGNAL и SERVICE. Короткие заголовки физического уровня предусмотрены спецификацией IEEE 802.1b для увеличения пропускной способности сети.

Рис. 5.8. Короткий заголовок кадров сети 802.11b

Из описания процедур связи сети IEEE 802.11 видно, что «накладные расходы» в этом стандарте выше, чем в проводной сети Ethernet. Поэтому крайне важно обеспечить высокую скорость передачи данных в канале. Повысить пропускную способность канала с заданной шириной полосы частот можно, разрабатывая и применяя более совершенные методы модуляции. По этому пути пошла группа разработчиков IEEE 802.11b.

Изначально стандарт IEEE 802.11 предусматривал работу в режиме DSSS с использованием так называемой баркеровской последовательности (Barker) длиной 11 бит: $B_1 = 10110111000$. Каждый информационный бит замещается своим произведением по модулю 2 (операция «исключающее ИЛИ») с данной последовательностью, т. е. каждая информационная единица заменяется на B_1, каждый ноль — на инверсию B_1. В результате бит заменяется последовательностью 11 чипов. Далее сигнал кодируется посредством дифференциальной двух- или четырехпозиционной фазовой модуляции (DBPSK или DQPSK, один или два чипа на символ соответственно). При частоте модуляции несущей 11 МГц общая скорость составляет в зависимости от типа модуляции 1 и 2 Мбит/с.

Стандарт IEEE 802.11b дополнительно предусматривает скорости передачи 11 и 5,5 Мбит/с. Для этого используется так называемая CCK-модуляция (Complementary Code Keying — кодирование комплементарным кодом). В основу данного метода легли работы специалистов компаний Intersil (ранее — Harris Semiconductor) и отчасти Agere Systems (полупроводниковое подразделение, выделенное из Lucent Technologies). Первым, по-видимому, изложил идею бинарных комплементарных кодов Марсель Голей (M. J. E. Golay), более полувека назад описав их принцип и методы генерации [26]. Суть комплементарного кодирования состоит в использовании двух последовательностей a и b, каждая из k элементов, с автокорреляционными функциями:

$$A(i) = \sum_{j=1}^{k-i} a_j a_{j+i} \quad \text{и} \quad B(i) = \sum_{j=1}^{k-i} b_j b_{j+i},$$

где $i = 0, \ldots, k-1$ — возможный сдвиг. Последовательности a и b комплементарны, если $A(i) + B(i) = 0$ при любом $i \neq 0$, и $A(0) + B(0) = 2k$. Физическая интерпретация автокорреляции со сдвигом — одновременный прием прямо распространяющегося сигнала и сигнала с фазовой задержкой на i элементов (чипов). Иными словами, если система связи, использующая комплементарное кодирование, работает в условиях многопутевого распространения сигналов,

то в идеале межсимвольная интерференция (вызванная наложением сигналов с задержками распространения) должна отсутствовать, поскольку сумма их автокорреляционных функций равна нулю.

Метод CCK использует дифференциальную квадратурную модуляцию (DQPSK) в радиотракте. CCK-модуляция строится на выделении из последовательного информационного потока групп по 8 бит (d_0–d_7). Эти 8 бит определяют информационный символ **C** из восьми комплексных чипов $\mathbf{C} = [c_0, \ldots, c_7]$. Используя представление Эйлера, можно записать: $c_i = e^{j\psi_i} = \cos\psi_i + j\sin\psi_i$, где ψ_i — фаза чипа; $j = \sqrt{-1}$. Чипы являются комплексными, поскольку применительно к квадратурной модуляции действительная и мнимая составляющие относятся к синфазному (I) и квадратурному (Q) каналам соответственно.

Восемь чипов информационного символа последовательно модулируют несущую с частотой модуляции 11 МГц. Фаза ψ_i каждого из них определяется в соответствии с формулой:

$$\mathbf{C} = [c_0, \ldots, c_7] = [e^{\phi_1+\phi_2+\phi_3+\phi_4}; e^{\phi_1+\phi_3+\phi_4}; e^{\phi_1+\phi_2+\phi_4}; e^{-(\phi_1+\phi_4)}; e^{\phi_1+\phi_2+\phi_3}; e^{\phi_1+\phi_3}; e^{-(\phi_1+\phi_2)}; e^{\phi_1}]. \tag{5.1}$$

Иначе эту формулу можно записать для вектора Ψ фаз чипов символа **C**:

$$\begin{aligned}\Psi &= [\psi_0, \ldots, \psi_7] = \\ &= [\phi_1 + \phi_2 + \phi_3 + \phi_4;\ \phi_1 + \phi_3 + \phi_4;\ \phi_1 + \phi_2 + \phi_4; \\ &- (\phi_1 + \phi_4);\ \phi_1 + \phi_2 + \phi_3;\ \phi_1 + \phi_3;\ -(\phi_1 + \phi_2);\ \phi_1].\end{aligned} \tag{5.2}$$

Элемент ϕ_1 входит в каждый чип символа, т.е. изменяет фазу всего символа. Фазовая модуляция сигнала называется дифференциальной, поскольку значение ϕ_1 текущего символа определяется относительно значения ϕ_1 предшествовавшего символа. Изменение фазы ϕ_1 задается в зависимости от значений двухразрядных двоичных чисел (дибитов) (d_0d_1). При этом для четных символов установлено следующее соответствие (d_0d_1) → $\Delta\phi_1$: (00) → 0; (01) → $\pi/2$; (11) → π; (10) → $-\pi/2$. Для нечетных символов значение ϕ_1 дополнительно смещают на π.

Как мы уже отмечали, в стандарте IEEE 802.11b посредством CCK кодируют только MAC-кадр. Поэтому первый CCK-символ — тот, что следует сразу за символами заголовка физического уровня. Он имеет номер «0», т.е. является четным. Фаза последнего QPSK-символа заголовка кадра является опорной для определения ϕ_1 первого символа информационного поля — к нему добавляется $\Delta\phi_1$. Остальные три параметра имеют абсолютные значения. Для скорости 11 Мбит/с они определяются как $\phi_2 = (d_2d_3)\cdot\pi/2$; $\phi_3 = (d_4d_5)\cdot\pi/2$; $\phi_4 = (d_6d_7)\cdot\pi/2$, где ($d_id_{i+1}$) — двухразрядные двоичные числа.

Для скорости 5,5 Мбит/с также используется DQPSK-модуляция, но CCK-символ (из восьми чипов) определяют не восемь, а четыре информационных бита (d_0–d_3), поэтому и скорость вдвое ниже. Параметр ϕ_1 вычисляется так же, как и для 11 Мбит/с, остальные иначе: $\phi_2 = d_2 \cdot \pi + \pi/2$; $\phi_3 = 0$; $\phi_4 = d_3\pi$.

На приемной стороне, при условии синхронного приема, декодирующее устройство восстанавливает значения ϕ_1, ϕ_2, ϕ_3, ϕ_4, а по ним — и значения информационных битов. В простейшем случае обработка сводится к следующим вычислениям:

$$\phi_2 = \arg\{r_1 r_2^* + r_3 r_4^* + r_5 r_6^* + r_7 r_8^*\},$$
$$\phi_3 = \arg\{r_1 r_3^* + r_2 r_4^* + r_5 r_7^* + r_6 r_8^*\},$$
$$\phi_4 = \arg\{r_1 r_5^* + r_2 r_6^* + r_3 r_7^* + r_4 r_8^*\},$$
$$\phi_1 = \arg\{r_4 e^{-j\phi_4} + r_6 e^{-j\phi_3} + r_7 e^{-j\phi_2} + r_8\},$$

где $r = [r_1, \ldots, r_8]$ — принятая 8-битная последовательность CCK-символов. Знак * означает комплексное сопряжение, т. е. если $r_i = a + jb$, то $r_i^* = a - jb$. Функция $\arg(r) = \text{arctg}[Im(r)/Re(r)] = \text{arctg}(b/a)$.

В чем достоинство CCK-модуляции? Как видно из формулы (5.2), фазы чипов символа (вектора) **C** определяются на основе последовательностей Уолша–Адамара. Если записать вектор $\mathbf{\Phi} = [\phi_1, \phi_2, \phi_3, \phi_4]$, то для скорости 11 Мбит/с вектор $\mathbf{\Psi} = (\mathbf{\Phi} \times \mathbf{WH} + \mathbf{B})$, где **WH** — верхняя половина матрицы Уолша–Адамара:

$$\mathbf{WH} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Вектор $\mathbf{B} = [0, 0, 0, \pi, 0, 0, \pi, 0]$, физически соответствующий повороту фазы чипов c_3 и c_6 на 180° (знаки «минус» в формуле (5.1)), необходим для улучшения корреляционных свойств кодовых последовательностей. Последовательности Уолша–Адамара хорошо изучены, обладают отличными автокорреляционными свойствами. Что немаловажно, каждая такая последовательность мало коррелирует сама с собой при фазовом сдвиге — очень полезное свойство при борьбе с переотраженными сигналами. Нетрудно заметить, что теоретическое усиление обработки при CCK-модуляции 3 дБ (в два раза), поскольку без кодирования комплементарным полифазным кодом сигнал, модулированный QPSK с частотой 11 Мсимвол/с, может транслировать 22 Мбит/с. Как видно, CCK-модуляция представляет собой вид блокового кода, а потому достаточно проста при аппаратной реализации. Совокупность этих свойств и обеспечила CCK место в стандарте IEEE 802.11b в качестве обязательного вида модуляции.

На практике важно не только операционное усиление. Существенную роль играет и равномерность распределения символов в фазовом пространстве (расстояние Хэмминга) — символы должны как можно дальше отстоять друг от друга, чтобы минимизировать ошибки их детектирования. И с этой точки зрения CCK-модуляция не выглядит оптимальной, ее реальное операционное усиление не превышает 2 дБ. Поэтому изначально прорабатывался другой способ модуляции — пакетное бинарное сверточное кодирование PBCC (Packet Binary Convolutional Coding). Этот метод вошел в стандарт IEEE 802.11b опционально, т. е. как необязательная опция. Механизм PBCC (рис. 5.9) был предложен специалистами фирмы Alantro Communications, в 2000 году вошедшей в состав компании Texas Instruments. PBCC позволяет добиваться в сетях IEEE 802.11b пропускной способности 5,5; 11 и 22 Мбит/с.

Как следует из названия, метод основан на сверточном кодировании. Для скоростей 5,5 и 11 Мбит/с поток информационных битов поступает в 6-разрядный сдвиговый регистр с сумматорами (рис. 5.10, *а*). В начальный момент времени все триггеры сдвигового регистра инициализируют нулем. В результате

каждый исходный бит d заменяется двумя битами кодовой последовательности (c_0, c_1). При скорости 11 Мбит/с c_0 и c_1 задают один символ четырехпозиционной QPSK-модуляции. Для скорости 5,5 Мбит/с используют двухпозиционную BPSK-модуляцию, последовательно передавая кодовые биты c_0 и c_1. Если же нужна скорость 22 Мбит/с, схема кодирования усложняется (рис. 5.10, *б*): три кодовых бита $(c_0$–$c_2)$ определяют один символ в восьмипозиционной 8-PSK-модуляции.

Рис. 5.9. Общая схема PBCC-модуляции

После формирования PSK-символов происходит скремблирование. В зависимости от сигнала s (см. рис. 5.9) символ остается без изменений ($s = 0$) либо его фаза увеличивается на $\pi/2$ ($s = 1$). Значение s определяет 256-битная циклически повторяющаяся последовательность S. Она формируется на основе начального вектора $U = 338B_{16}$, в котором равное число нулей и единиц. S представляет собой 16 последовательных векторов U_i $(i = 0, \ldots, 15)$, причем каждый вектор U_i циклически сдвинут влево по сравнению с U_{i-1} на $3i$ разряда.

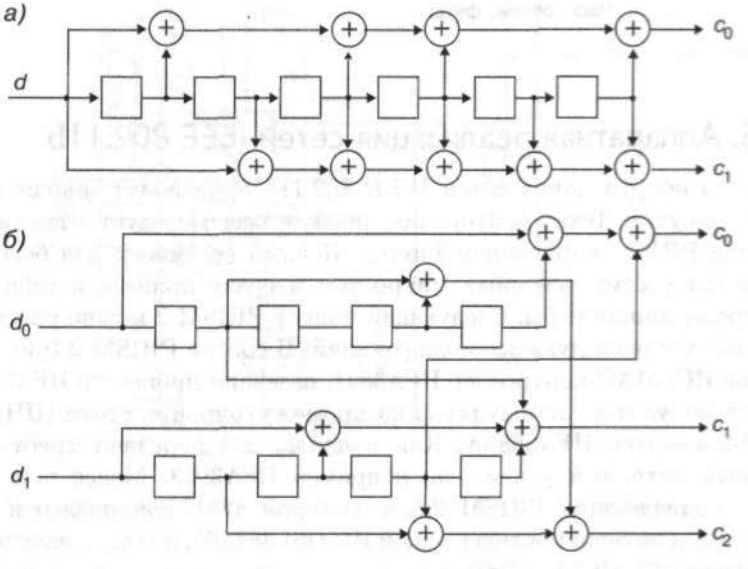


Рис. 5.10. Схема сверточного BCC-кодирования: *а)* — для скоростей 5,5 и 11 Мбит/с; *б)* — для скорости 22 Мбит/с

У 6-разрядного сдвигового регистра, применяемого в PBCC для скоростей 11 и 5,5 Мбит/с, 64 возможных выходных состояния. Так что при модуляции PBCC информационные биты в фазовом пространстве оказываются гораздо дальше

друг от друга, чем при CCK-модуляции. Поэтому PBCC и позволяет при одних и тех же соотношениях сигнал/шум и уровне ошибок вести передачу с большей скоростью, чем в случае CKK (рис. 5.11). Однако плата за более эффективное кодирование — сложность аппаратной реализации данного алгоритма. Видимо, не случайно продвигать PBCC-модуляцию стала компания Texas Instruments — ведущий производитель процессоров цифровой обработки сигнала.

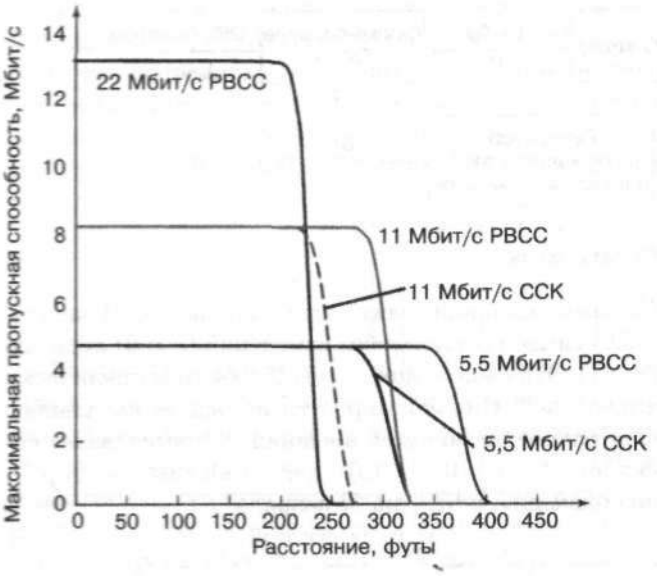


Рис. 5.11. Дальность связи при различных способах модуляции и фиксированном затухании в канале (по материалам компании Texas Instruments о применении ИС ACX100)

5.5. Аппаратная реализация сетей IEEE 802.11b

Чипсеты для оборудования сетей IEEE 802.11b производят многие компании. Кроме упомянутой Texas Instruments, прежде всего следует отметить линейку чипсетов PRISM корпорации Intersil. Первый ее чипсет для беспроводных ЛВС включал восемь основных микросхем и сразу привлек к себе внимание разработчиков аппаратуры. Следующий чипсет PRISM 2 можно рассматривать как базовый для последующих модификаций. В состав PRISM 2 (рис. 5.12) входят четыре ИС: MAC-контроллер HFA3841, baseband-процессор HFA3861B, квадратурный модулятор/демодулятор на промежуточной частоте (ПЧ) HFA3783 и ВЧ/ПЧ-конвертер HFA3683A. Как правило, для передачи требуется высокочастотный антенный усилитель, например HFA3963. Менее чем через год появилась модификация PRISM 2,5, в которой MAC-контроллер и baseband-процессор были интегрированы в одной ИС (ISL3873B), поддерживающей к тому же интерфейсы PCMCIA и USB.

Следующим шагом стало создание к лету 2001 года чипсета PRISM 3 на основе фирменной архитектуры прямого преобразования частоты ZIF (Zero IF, нулевая ПЧ) без использования ПЧ.

Отметим и объявленный чипсет компании Philips Semiconductors, первая ИС которого — однокристальный трансивер SA2400 для диапазона 2,4 ГГц —

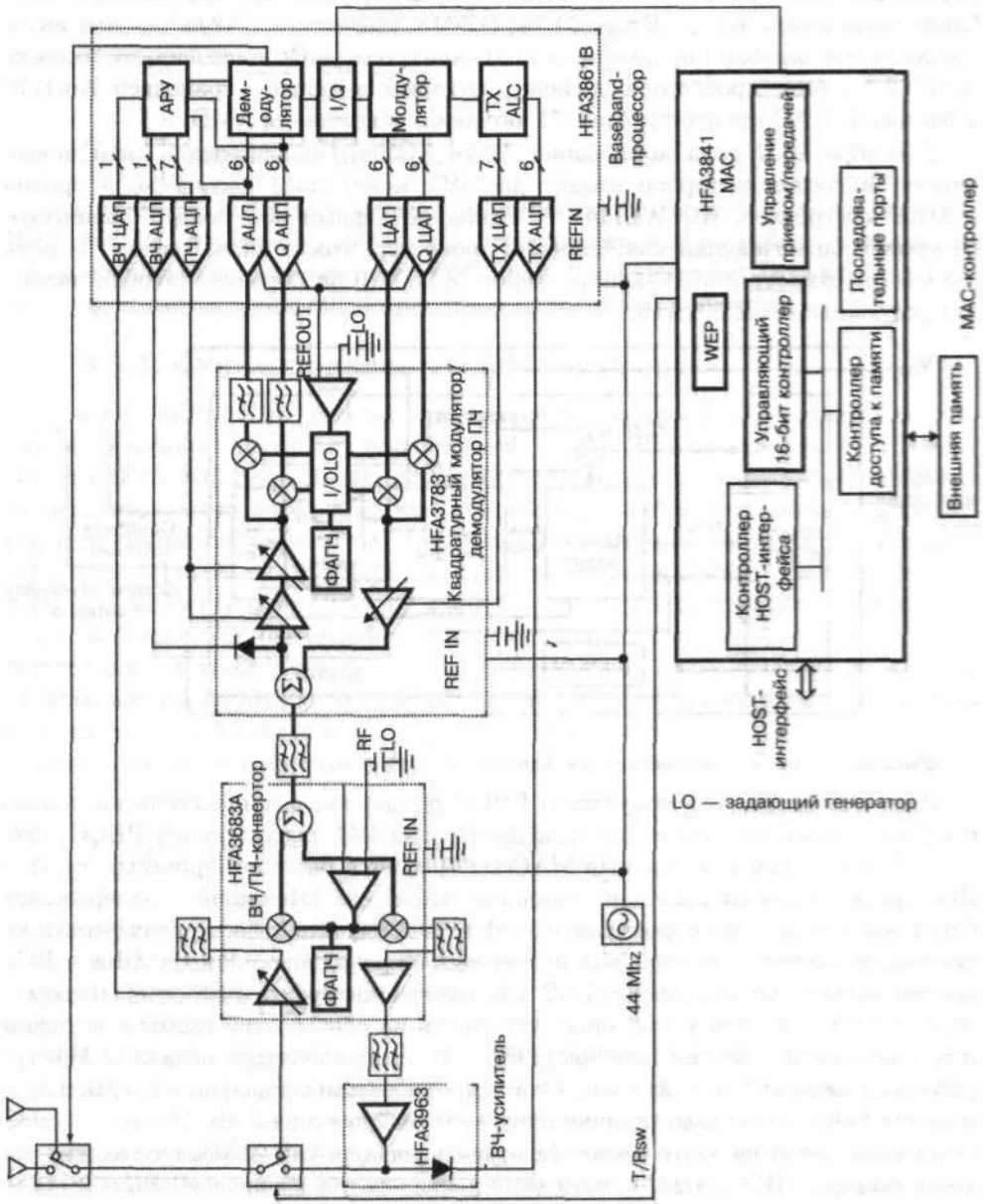


Рис. 5.12. Структурная схема устройства на базе чипсета PRISM 2

была анонсирована в июле 2001 года. Данная ИС также построена по принципу прямого преобразования (без ПЧ) и содержит блоки приемника, передатчика, квадратурного модулятора/демодулятора, задающего генератора на основе ГУН и fractional-N-синтезатора, а также выходного малошумящего усилителя. В составе ИС — полностью интегрированные полосовые канальные фильтры, устройство автоматического контроля усиления, трехпроводная шина

управления основными блоками устройства, интерфейс для сопряжения с baseband-процессором и т. д. Вторая микросхема чипсета — SA2440 — это интегрированные baseband-процессор и MAC-контроллер. Об аналогичном чипсете Am1772 на базе процессора Alchemy, который включает трансивер Am1770 и baseband/MAC-процессор Am1771, объявила и компания AMD.

В ноябре 2002 года корпорация Agere Systems анонсировала свой новый чипсет, в состав которого входят две ИС: контроллер физического уровня и MAC-контроллер. ИС WL1141 — однокристальный контроллер физического уровня, включающий как baseband-процессор, так и аналоговую ВЧ-часть (рис. 5.13). Вместе с MAC-контроллером WL60010 он образует полный чипсет для устройств IEEE 802.11b.

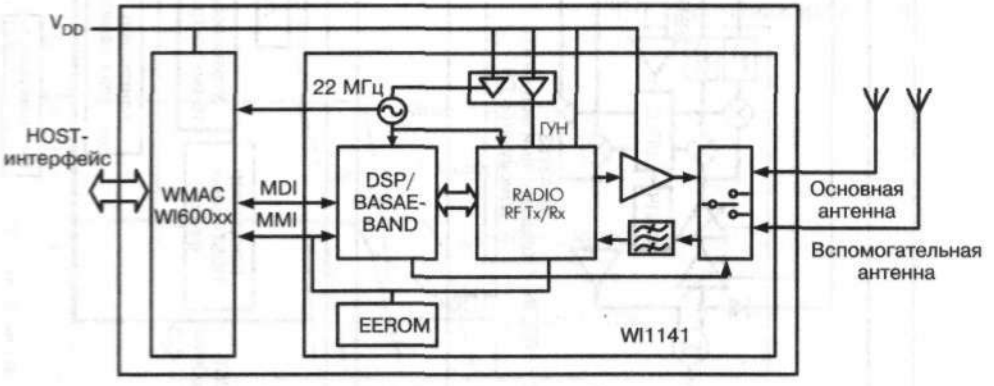


Рис. 5.13. Чипсет компании Agere Systems

Надо отметить, что модуляцию PBCC поддерживают практически только продукты компании Texas Instruments. Первой ИС, реализующей PBCC, стала ACX100 — однокристальный MAC-контроллер с baseband-процессором. Эта ИС, кроме обязательных для стандарта IEEE 802.11b опций, поддерживает PBCC-модуляцию со скоростями 5,5; 11 и 22 Мбит/с. Среди других особенностей микросхемы — аппаратная поддержка 32-разрядных шин CardBus и PCI, последовательного интерфейса USB 1.1, а также интерфейса Ethernet. Интересно отметить, что при испытаниях устройств на основе этой схемы при одном и том же уровне пакетных ошибок ($PER = 10^{-2}$) и скорости передачи 11 Мбит/с работа в режиме CCK была возможной при отношении сигнал/шум 8,5 дБ, а применение PBCC позволяло ухудшить это соотношение до 4,5 дБ.

Однако, несмотря на все усилия Texas Instruments, режим 22 Мбит/с так и не вошел в стандарт IEEE 802.11b, хотя даже появилось обозначение «IEEE 802.11b+» и устройства на основе ACX100 начали производить такие известные компании, как D-Link и NDC. Не получили распространения и сети со скоростью 22 Мбит/с. Видимо, тут играют роль два обстоятельства. Прежде всего, процедура стандартизации — это конкуренция за получение значительных финансовых прибылей в виде лицензионных отчислений тому, чья технология стала стандартной. В данном случае столкнулись интересы двух гигантов полупроводниковой индустрии — компаний Intersil и Texas Instruments. Поддерживать одновременно два вида модуляции сложно в финансовом и техническом отношении, поэтому и был принят CCK-вариант.

С другой стороны, скорости порядка 22 Мбит/с стали мало интересны для потребителей, так как стандарты IEEE 802.11a и g открыли перед ними новые перспективы — до 54 Мбит/с.

5.6. Стандарт IEEE 802.11a

Стандарт IEEE 802.11a появился практически одновременно со стандартом IEEE 802.11b, в сентябре 1999 года. Он ориентирован на работу в диапазоне 5 ГГц и основан на технологии OFDM (Orthogonal Frequency Division Multiplexing — мультиплексирование посредством ортогональных несущих).

5.6.1. Формирование OFDM-символов

Технология OFDM означает, что единовременно информация передается по многим поднесущим частотам, образующим канал. Ширина канала — 20 МГц. В сетях IEEE 802.11a в 20-МГц канале используются 52 поднесущие, однако их номинальное число выбирается из соображений удобства преобразования Фурье и принимается равным 64. Таким образом, интервал между поднесущими $\Delta f = 20$ МГц/64 = 312,5 кГц, а сами поднесущие можно представить как $f_k(t) = a_k \sin[2\pi(f_0 + k\Delta f)t + \phi_k]$, где $k = -26, \ldots, 26$. Центральная поднесущая f_0 не используется (ее амплитуда всегда равна нулю). Поднесущие модулируются посредством квадратурной амлитудно-фазовой модуляции: 2-, 4- 16- и 64-позиционной BPSK, QPSK, 16-QAM и 64-QAM соответственно. Сигнал удобно представлять в комплексной форме $C = a\cos x + b\sin x = Ae^{jx}$, где $j = \sqrt{-1}$. Соответственно, суммарный сигнал на всех поднесущих можно записать как

$$s(t) = \sum_{k=-26}^{26} C_k e^{j2\pi(f_0 + k\Delta f)t}. \qquad (5.3)$$

Здесь C_k — комплексная амплитуда k-й поднесущей, мнимая и действительная составляющие которой соответствуют квадратурному (Q) и синфазному (I) каналам квадратурной модуляции. Значения комплексных составляющих выбираются в соответствии с диаграммами Грея (рис. 5.14), исходя из потока информационных бит.

OFDM-символ представляет собой совокупность всех поднесущих на дискретном интервале длительностью $T_F = 1/\Delta f = 3{,}2$ мкс. Информационная емкость OFDM-символа определяется типом модуляции информационных поднесущих и их числом. Из 52 поднесущих в стандарте IEEE 802.11a для передачи данных используются 48, остальные 4 поднесущие — пилотные. Следовательно, емкость OFDM-символа составляет $48 \times N_b$, где N_b — число бит в одном модуляционном символе (на одной поднесущей), равное двоичному логарифму от числа позиций модуляции. Таким образом, OFDM-символ содержит от 96 до 288 бит.

Отметим, OFDM-модуляция обладает мощным средством борьбы с межсимвольной интерференцией, проявляющейся в том, что из-за множественных переотражений в приемник одновременно поступают два смежных символа — прямо распространяющийся и «запоздавший». Это ведет к потере символов. В случае OFDM-модуляции, которая допускает небольшую скорость передачи данных на одной поднесущей, в каждый OFDM-символ добавляется защитный

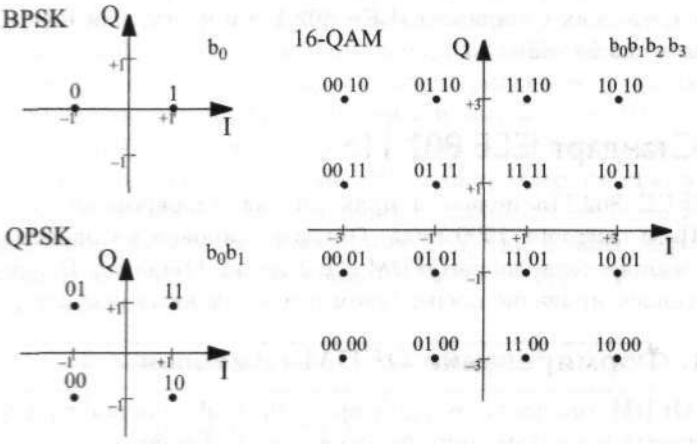


Рис. 5.14. Представление модуляционных символов (диаграммы Грея) для BPSK, QPSK и 16-QAM

интервал GI. В стандарте IEEE 802.11a его длительность равна $T_F/4 = 0{,}8$ мкс. Следовательно, длительность всего символа $T_S = T_{GI} + T_F = 4$ мкс. Защитный интервал транслируется в начале OFDM-символа и представляет собой копию его последних 0,8 мкс. В результате отраженный и пришедший с задержкой символ попадает в защитный интервал и не повреждает прямо распространяющийся символ.

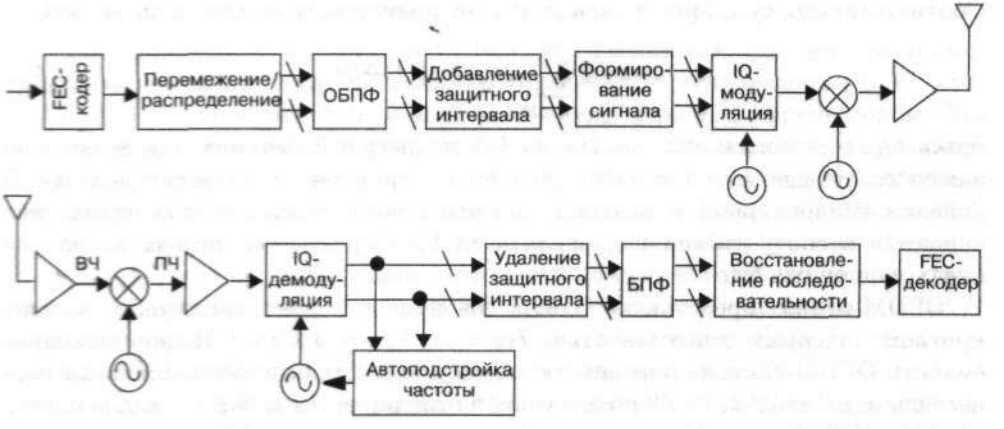


Рис. 5.15. Функциональная схема трактов приема/передачи стандарта IEEE 802.11a

Рассмотрим процедуры формирования выходного сигнала в стандарте IEEE 802.11a (рис. 5.15). Входной поток данных (бит) прежде всего подвергается скремблированию (в данном случае рандомизации) посредством перемножения на псевдослучайную последовательность (ПСП) с циклом повторения 127. Ее формирует генератор с задающим полиномом $G(x) = x^7 + x^4 + 1$ и начальным значением 1111111. При передаче конкретного пакета вектор инициализации генератора ПСП может быть произвольным, но должен принадлежать ПСП.

Приемник восстанавливает его, поскольку известно, что последние 7 бит поля данных (младшие биты поля SERVICE заголовка, см. далее) перед скремблированием всегда равны нулю.

После скремблирования поток данных поступает на сверточный кодер (FEC). Исходя из выбранной скорости передачи данных, скорость кодирования может составлять 1/2, 2/3 и 3/4. Напомним, скорость кодирования — это отношение числа бит в пакете до и после кодера (скорость кодирования $r = 1/2$ означает, что каждый входной бит после кодирования превращается в два бита). Поскольку у кодера два выхода, каждому входному биту x_i соответствует пара бит (y_i, z_i). Значения скорости кодирования, отличные от 1/2, получаются путем исключения из выходной последовательности отдельных значений y_i или z_i (процедура выкалывания).

Далее поток кодированных битов подвергается перемежению (интерливингу) — изменяется порядок битов в последовательности в рамках OFDM-символа. Вся последовательность кодированных битов разбивается на блоки, длина которых равна числу битов в OFDM-символе (N_{CBPS}) при выбранной скорости передачи. В пределах блока биты нумеруются от 0 до $N_{CBPS} - 1$. Затем происходит двухстадийная перестановка. Цель первого этапа — добиться, чтобы смежные биты кодовой последовательности оказались на несмежных поднесущих. Первый этап перемежения эквивалентен тому, что данные последовательно по строкам записываются в таблицу из 16 строк и $N_{CBPS}/16$ столбцов, а затем последовательно считываются по столбцам (т. е. считываются в порядке записи, но из транспонированной таблицы).

После второго этапа перестановки смежные биты оказываются попеременно в старших и младших разрядах групп, определяющих модуляционный символ квадратурной модуляции (см. рис. 5.14). Это делается для того, чтобы соседние биты не оказались в младших разрядах, надежность передачи которых наиболее низка. Математически процедура перемежения выражается двумя уравнениями, в которых k — номер бита в кодированной последовательности, i — его номер после первого этапа стадии перестановок, j — после второго (окончательный):

$$i = (N_{CBPS}/16) \cdot (k \mod 16) + \text{floor}(k/16);$$
$$j = s \cdot \text{floor}(i/s) + (i + N_{CBPS} - \text{floor}[16 \cdot i/N_{CBPS}]) \mod s;$$
$$s = \max(N_{BPSC}/2, 1),$$

где N_{BPSC} — число битов на поднесущую. Функция $x \mod n$ — это остаток от x/n, значение функции $\text{floor}(x)$ равно наибольшему целому числу, не превышающему x.

После интерливинга последовательность битов разбивается на группы по числу позиций выбранной квадратурной модуляции (1; 2; 4 или 6) и в соответствии с диаграммами Грея определяют значения синфазной (младшие биты) и квадратурной (старшие биты) составляющих комплексных амплитуд. Полученные из диаграмм Грея значения амплитуд умножаются на нормировочный коэффициент $1, 1/\sqrt{2}, 1/\sqrt{10}, 1/\sqrt{42}$ для BPSK, QPSK, 16-QAM и 64-QAM соответственно. В результате получаются значения комплексных амплитуд C_k, используемых в формуле (5.3). Различные сочетания скоростей кодирования и методов модуляции поднесущих и определяют набор скоростей передачи данных в стандарте IEEE 802.11a (табл. 5.1).

Таблица 5.1. Параметры OFDM-символов для различных скоростей передачи данных в стандарте IEEE 802.11a

Скорость передачи данных, Мбит/с	Модуляция	Скорость кодирования	Кодированных битов на несущую (N_{BPSC})	Кодированных битов в OFDM-символе (N_{CBPS})	Информационных битов в OFDM-символе (N_{DBPS})
6	BPSK	1/2	1	48	24
9	BPSK	3/4	1	48	36
12	QPSK	1/2	2	96	48
18	QPSK	3/4	2	96	72
24	16-QAM	1/2	4	192	96
36	16-QAM	3/4	4	192	144
48	64-QAM	2/3	6	288	192
54	64-QAM	3/4	6	288	216

Сформированный OFDM-символ подвергается обратному быстрому преобразованию Фурье (ОБПФ), в результате чего формируются выходные синфазный и квадратурный сигналы. К ним добавляется защитный интервал, после чего происходит окончательное формирование аналогового сигнала. Дальнейшая обработка стандартна — квадратурный модулятор, гетеродин для переноса сигнала в заданную область (если это необходимо) и выходные усилители. Если используется гетеродин, в формуле (5.3) $f_0 = 0$. В приемнике преобразования выполняются в обратном порядке (см. рис. 5.15).

5.6.2. Структура пакетов физического уровня

Пакеты MAC-уровня в IEEE 802.11a стандартны: сначала следует MAC-заголовок, содержащий адреса приемников и передатчиков, а также служебную информацию, далее — собственно данные (поле данных), за ними — контрольная сумма (CRC). Сформированный пакет MAC-уровня (MPDU) встраивается в пакет физического уровня. Здесь и начинаются различия с базовым документом IEEE 802.11.

Рис. 5.16. Структура пакетов физического уровня стандарта IEEE 802.11a

На физическом уровне кадр представляет собой последовательность в составе преамбулы, заголовка (PLCP-заголовок) и поля данных, за которым следуют

так называемые хвостовые биты (Tail, равны нулю и обозначают конец поля) и заполняющие биты (Pad), предназначенные для выравнивания длины пакета (рис. 5.16). Преамбула содержит 12 OFDM-символов. Все поля заголовка, кроме SIGNAL, передаются посредством одного OFDM-символа, причем с наименьшей из возможных скоростей (номинальное значение 6 Мбит/с). Оставшаяся часть заголовка и поле данных транслируются с любой заданной скоростью из списка возможных (см. табл. 5.1).

Собственно преамбула (рис. 5.17) включает десять коротких настроечных последовательностей (t)) и две длинные последовательности (T). OFDM-символы коротких последовательностей формируются на основе лишь 12 поднесущих, при этом применяется четырехпозиционная QPSK-модуляция. Длительность короткой настроечной последовательности — 0,8 мкс, защитных интервалов между ними нет. Короткие настроечные последовательности предназначены для автоматической настройки усилителей сигнала (АРУ), а также временной и частотной синхронизации.

Рис. 5.17. Последовательность передаваемых данных в стандарте IEEE 802.11a

Две длинные настроечные последовательности следуют за короткими с промежутком в два защитных интервала GI = 0,8 мкс. Каждой из них соответствуют OFDM-символы, включающие 53 поднесущие, в том числе центральную f_0. Поднесущие модулируются посредством двухпозиционной BPSK, длительность символов — 3,2 мкс, защитных интервалов нет. Длинные последовательности предназначены для оценки канала и точной частотной подстройки приемников. Таким образом, длительность трансляции преамбулы составляет 16 мкс.

За преамбулой следует PLCP-заголовок физического пакета. Он состоит из двух фрагментов — SIGNAL и SERVICE. Фрагмент SIGNAL всегда занимает один OFDM-символ и транслируется посредством BPSK-модуляции со скоростью кодирования 1/2, т.е. максимально надежно. Он не скремблируется. В SIGNAL передается информация о скорости передачи поля данных (поле RATE) и длине пакета (LENGTH). Для надежности используется бит контроля четности (Parity). Шесть последних бит (Tail), всегда равных нулю, обозначают конец фрагмента SIGNAL.

Фрагмент SERVICE (16 бит) формально принадлежит заголовку, но входит в поле данных и передается с выбранной для передачи данных скоростью. Используются только младшие 7 бит поля SERVICE (для инициализации генератора ПСП в приемнике), перед скремблированием они всегда равны нулю. Остальные 9 бит в стандарте IEEE 802.11a не задействованы.

Поле данных завершают 6 разделительных нулевых бит (Tail). Они добавляются после скремблирования и служат как дополнительное средство контроля

ошибок, поскольку в приемнике после сверточного декодера также должны оказаться равными нулю. Кроме того, в конце пакета добавляются специальные биты заполнения Pad (равны нулю), так чтобы общая длина поля данных (включая SERVICE) оказалась кратной числу битов в OFDM-символе при выбранной скорости передачи данных.

Аппаратная поддержка

Диапазон 5,1–5,9 ГГц хорош тем, что там гораздо проще найти широкую полосу для системы связи. В США для безлицензионной работы в этом диапазоне выделены полосы 5,15–5,35 и 5,725–5,825 ГГц — всего 300 МГц по сравнению с 83 МГц в диапазоне 2,4 ГГц. Вместо трех неперекрывающихся каналов в диапазоне 2,4 ГГц для сетей IEEE 802.11b только в нижнем поддиапазоне 5,15–5,35 ГГц имеются восемь неперекрывающихся каналов (рис. 5.18). Аналогичная ситуация в Европе и в России (однако в нашей стране отсутствуют безлицензионные диапазоны) — в более высокочастотной области места больше.

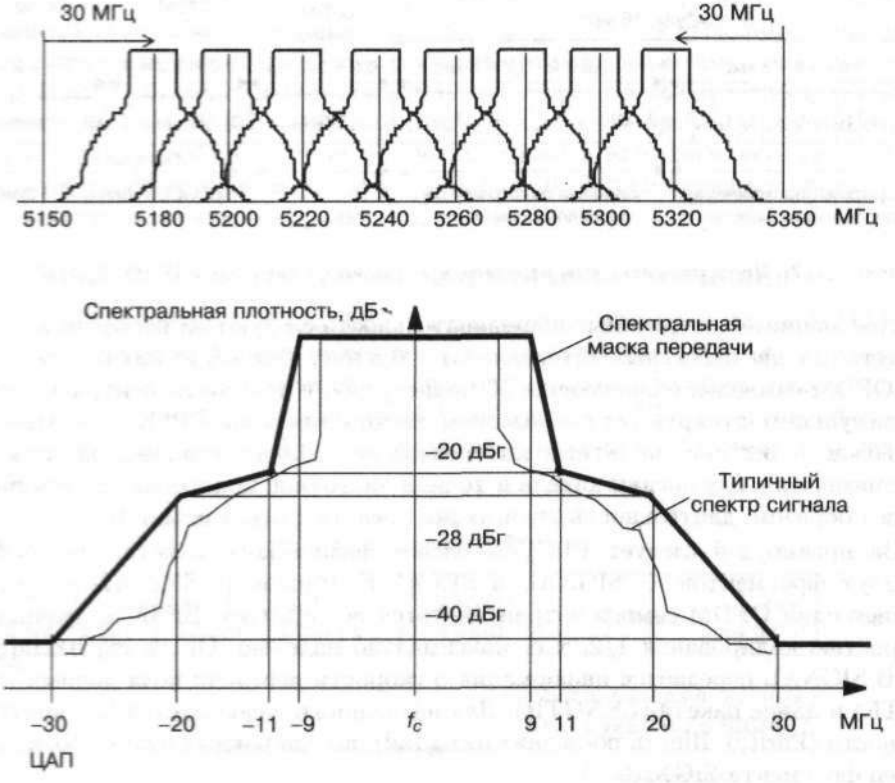


Рис. 5.18. Спектральная маска и распределение неперекрывающихся каналов в диапазоне 5,15–5,35 ГГц для стандарта IEEE 802.11a. Мощность (дБг) определяется относительно пиков функции $\sin(x)/x$

Кроме того, диапазон 2,4 ГГц перегружен различными системами связи — тут и беспроводные телефоны, и устройства Bluetooth (IEEE 802.15.1), и многочисленное оборудование стандарта IEEE 802.11b. Взаимных помех избежать тяжело. Сравнительные испытания убедительно показывают, что в одних и тех

же условиях устройства IEEE 802.11a по скорости обмена превосходят оборудование IEEE 802.11b (рис. 5.19). И до недавнего времени казалось, что будущее принадлежит сетям стандарта IEEE 802.11a. Однако возник ряд вопросов.

Прежде всего, как быть с уже существующими сетями (и оборудованием) в диапазоне 2,4 ГГц? Как обеспечить столь необходимую всем обратную совместимость? С этой проблемой производители справились, разработав двухдиапазонные чипсеты. Характерный пример — компания Atheros, создавшая комплект из трех ИС AR5001X Combo. В его составе ИС baseband-процессора и MAC-контроллера AR5211 и две аналоговые ИС трансиверов — на 2,4 и на 5 ГГц (AR2111 и AR5111 соответственно). Аналогичное решение предложено и компанией Intersil — чипсет PRISM Duette из двух ИС (рис. 5.19): baseband/MAC-процессор ISL3890 и однокристальный двухдиапазонный трансивер ISL3690.

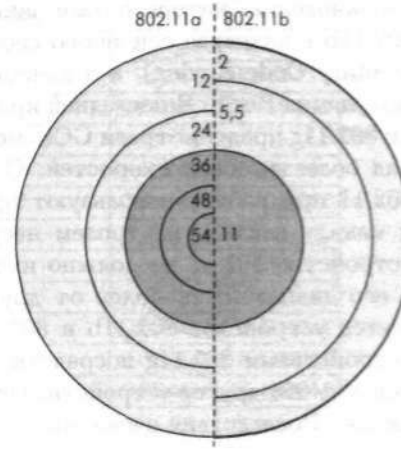


Рис. 5.19. Соотношение скоростей передачи и радиусов действия устройств стандартов 802.11a и 802.11b (по материалам компании Proxim)

Однако к моменту, когда данные решения стали технологически возможны и рентабельны, в Европе был разработан свой стандарт 5 ГГц — HiperLan2 (правда, так и не нашедший поддержки со стороны производителей аппаратуры). Кроме того, работы по ускорению стандарта IEEE 802.11b в диапазоне 2,4 ГГц привели к появлению новой версии — IEEE 802.11g, предусматривающей скорости до 54 Мбит/с. В частности, упомянутые выше чипсеты AR5001X Combo и PRISM Duette в диапазоне 2,4 ГГц поддерживают не только IEEE 802.11b, но и IEEE 802.11g.

5.7. Стандарт IEEE 802.11g

Работы над будущей спецификацией IEEE 802.11g начались в марте 2000 года, когда была сформирована исследовательская группа по изучению возможности увеличения скорости передачи данных свыше 20 Мбит/с в диапазоне 2,4 ГГц. В ноябре 2000 года. эта группа приобрела статус штатной группы разработчиков и получила обозначение G. Через полтора года, рассмотрев несколько альтернативных подходов, специалисты исследовательской группы G предложили использовать применяющуюся в стандарте IEEE 802.11a систему кодирования с мультиплексированием посредством ортогональных несущих OFDM. В качестве опциональных возможностей новый стандарт IEEE 802.11g предусматривал использование таких схем модуляции, как CCK-OFDM и PBCC.

Новая спецификация по сути представляет собой перенесение схемы модуляции OFDM, прекрасно зарекомендовавшей себя в 802.11a, из диапазона 5 ГГц в область 2,4 ГГц. Это возможно, поскольку в стандартах 802.11 ширина одного канала в диапазоне 2,4 и 5 ГГц схожа — 22 МГц по уровню −30 и −20 дБ соответственно. Правда, по уровню −28 дБ маска канала в IEEE 802.11a до-

пускает спектральную полосу шириной 40 МГц, что может создать проблемы, безусловно, преодолимые.

Одним из основных требований к спецификации 802.11g была обратная совместимость с устройствами 802.11b. Это требование привело к очередному столкновению интересов компаний Intersil и TI. Действительно, в стандарте 802.11b в качестве основного способа модуляции принята схема CCK (Complementary Code Keying), а в качестве дополнительной возможности допускается модуляция PBSS. В последней крайне заинтересована компания TI. Разработчики 802.11g предусмотрели CCK-модуляцию для скоростей до 11 Мбит/с и OFDM для более высоких скоростей. С этим были согласны все. Но сети стандарта 802.11 при работе используют принцип CSMA/CA — множественный доступ к каналу связи с контролем несущей и предотвращением коллизий. Ни одно устройство 802.11 не должно начинать передачу, пока не убедится, что эфир в его диапазоне свободен от других устройств. Если в зоне слышимости окажутся устройства 802.11b и 802.11g, причем обмен будет происходить между устройствами 802.11g посредством OFDM, то оборудование 802.11b просто не поймет, что другие устройства сети ведут передачу, и попытается начать трансляцию. Последствия очевидны.

Рис. 5.20. Кадры IEEE 802.11g в различных режимах модуляции

Чтобы подобную ситуацию не допустить, предусмотрена возможность работы в смешанном режиме — CCK-OFDM. Информация в сетях 802.11 передается кадрами. Каждый информационный кадр включает два основных поля: преамбулу с заголовком и информационное поле (рис. 5.20). Преамбула содержит синхропоследовательность и код начала кадра, заголовок — служебную информацию, в том числе о типе модуляции, скорости и продолжительности передачи кадра. В режиме CCK-OFDM преамбула и заголовок модулируются методом CCK (реально — путем прямого расширения спектра DSSS посредством последовательности Баркера, поэтому в стандарте 802.11g этот режим именуется DSSS-OFDM), а информационное поле — методом OFDM. Таким образом, все устройства 802.11b, постоянно «прослушивающие» эфир, принимают заголовки кадров и узнают, сколько времени будет транслироваться кадр 802.11g. В этот период они «молчат». Естественно, пропускная способность сети падает, поскольку скорость передачи преамбулы и заголовка — 1 Мбит/с.

Видимо, данный подход не устраивал лагерь сторонников технологии PBSS, и для достижения компромисса в стандарт 802.11g в качестве дополнительной возможности ввели, так же как и в 802.11b, необязательный режим — PBSS, в котором заголовок и преамбула передаются так же, как и при CCK, а информационное поле модулируется по схеме PBSS и передается на скорости 22 или 33 Мбит/с. В результате устройства стандарта 802.11g должны оказаться совместимыми со всеми модификациями оборудования 802.11b и не создавать взаимных помех. Диапазон поддерживаемых им скоростей отражен в табл. 5.2, зависимость скорости от типа модуляции — на рис. 5.21.

Таблица 5.2. Возможные скорости и тип модуляции в спецификации IEEE 802.11g

Скорость, Мбит/с	Тип модуляции	
	Обязательно	Допустимо
1	DSSS (последовательность Баркера)	
2	DSSS (последовательность Баркера)	
5,5	CCK	PBCC
6	OFDM	CCK-OFDM
9		OFDM, CCK-OFDM
11	CCK	PBCC
12	OFDM	CCK-OFDM
18		OFDM, CCK-OFDM
22		PBCC
24	OFDM	CCK-OFDM
33		PBCC
36		OFDM, CCK-OFDM
48		OFDM, CCK-OFDM
54		OFDM, CCK-OFDM

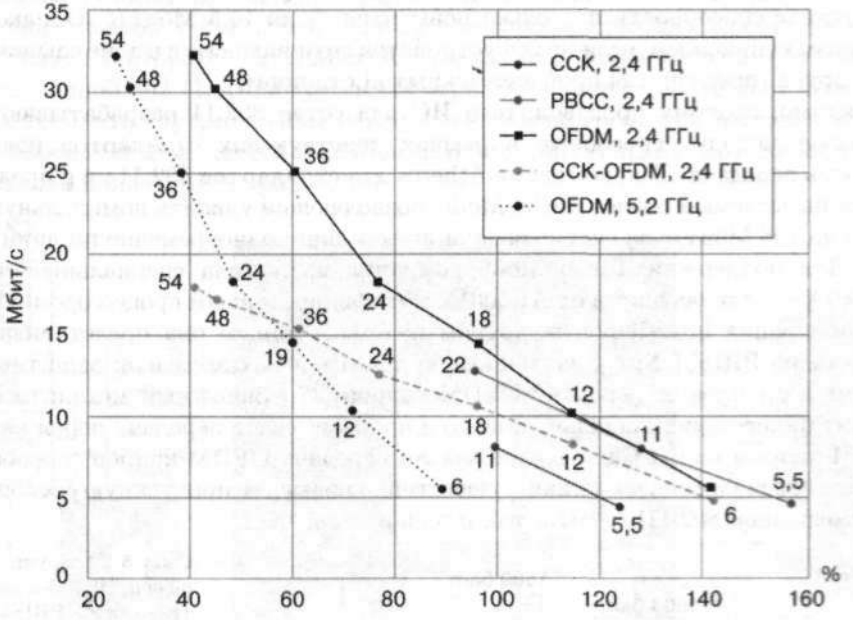


Рис. 5.21. Зависимость скорости передачи от расстояния для различных технологий передачи. Расстояние приведено в процентах, 100% — дальность передачи с модуляцией CCK на скорости 11 Мбит/с

Очевидно, что устройствам стандарта IEEE 802.11g достаточно долго придется работать в одних сетях с оборудованием 802.11b. Также очевидно, что производители в массе своей не будут поддерживать режимы CCK-OFDM и PBSS в силу их необязательности, ведь почти все решает цена устройства. Поэтому одна из основных проблем нового стандарта — как обеспечить бесконфликтную работу смешанных сетей 802.11b/g.

Основной принцип работы в сетях 802.11 — «слушать, прежде чем вещать». Но устройства 802.11b не способны услышать устройства 802.11g в OFDM-режиме. Ситуация аналогична проблеме скрытых станций: два устройства удалены настолько, что не слышат друг друга и пытаются обратиться к третьему, которое находится в зоне слышимости обоих. Для предотвращения конфликтов в подобной ситуации в 802.11 введен защитный механизм, предусматривающий перед началом информационного обмена передачу короткого кадра «запрос на передачу» (RTS) и получение кадра подтверждения «можно передавать» (CTS). Механизм RTS/CTS применим и к смешанным сетям 802.11b/g. Естественно, эти кадры должны транслироваться в режиме CCK, который обязаны понимать все устройства. Однако защитный механизм существенно снижает пропускную способность сети. Так, при физической скорости 54 Мбит/с потолок пропускной способности гомогенной сети 802.11g (с учетом всей служебной и управляющей информации) около 32 Мбит/с, а реальные показатели оборудования — на уровне 24 Мбит/с. Если же сеть смешанная, то защитный механизм RTS/CTS понизит пропускную способность до 12 Мбит/с. Это практически вдвое превышает пропускную способность однородной сети 802.11b ($\sim$ 6 Мбит/с), но ведь всегда хочется большего. Поэтому вместо механизма RTS/CTS можно использовать только кадры CTS, предшествующие каждому OFDM-кадру. В результате пропускная способность несколько повысится — до 14,5 Мбит/с. Однако этот механизм неприемлем, если не все устройства сети находятся в зоне слышимости друг друга (пресловутая проблема «скрытой станции»).

Видимо, поэтому производители ИС для сетей 802.11 разрабатывают специальные методы, способные в рамках действующих стандартов повысить скорость передачи. Так, компания Atheros для стандартов 802.11a и g предложила так называемый режим Turbo Mode, позволяющий удвоить номинальную скорость до 108 Мбит/с за счет передачи информации одновременно по двум каналам. Для поддержки Turbo Mode компания выпустила специальный чипсет AR5001X+, отличающийся от AR5001X модифицированным процессором AR5212.

Корпорация Intersil пошла другим путем. В апреле она представила свою технологию PRISM Nitro, включающую два основных элемента: защитный механизм и групповую передачу OFDM-кадров [27]. Защитный механизм не содержит ничего принципиально нового и подразумевает передачу перед каждым OFDM-кадром кадра CTS. Групповая же передача OFDM-кадров способна, по мнению специалистов компании, существенно повысить пропускную способность как смешанной 802.11b/g сети, так и однородной.

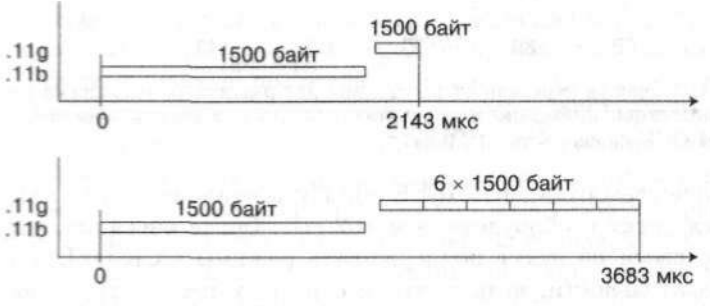


Рис. 5.22. Групповая передача OFDM-пакетов по технологии PRISM Nitro в смешанной сети

В случае смешанной сети предлагается каждому устройству предоставлять в трафике примерно равный временной интервал. Действительно, для передачи

ССК-кадра со скоростью 11 Мбит/с требуется примерно столько же времени, сколько для передачи шести OFDM-кадров со скоростью 54 Мбит/с (с учетом всех накладных расходов). Если устройства 802.11b и g поочередно передают одинаковый объем информации, на передачу, например, пакетов с информационным полем 1500 байт двум устройствам потребуется 2143 мкс (рис. 5.22).

Рис. 5.23. Групповая передача OFDM-пакетов по технологии PRISM Nitro в однородной сети

Если же каждому устройству для трансляции выделить равные временные интервалы, устройство 802.11g передаст шесть пакетов (9000 байт) — всего 10 500 байт за 3683 мкс. В первом случае пропускная способность сети составит 11,2 Мбит/с, во втором — 22,8 Мбит/с: выигрыш — более чем в два раза. В случае однородной 802.11g-сети групповая передача пакетов также дает выигрыш за счет того, что внутри группы между пакетами не требуется выставлять кадр CTS и выжидать межкадровый интервал. Необходим только короткий кадр подтверждения приема ACK (рис. 5.23).

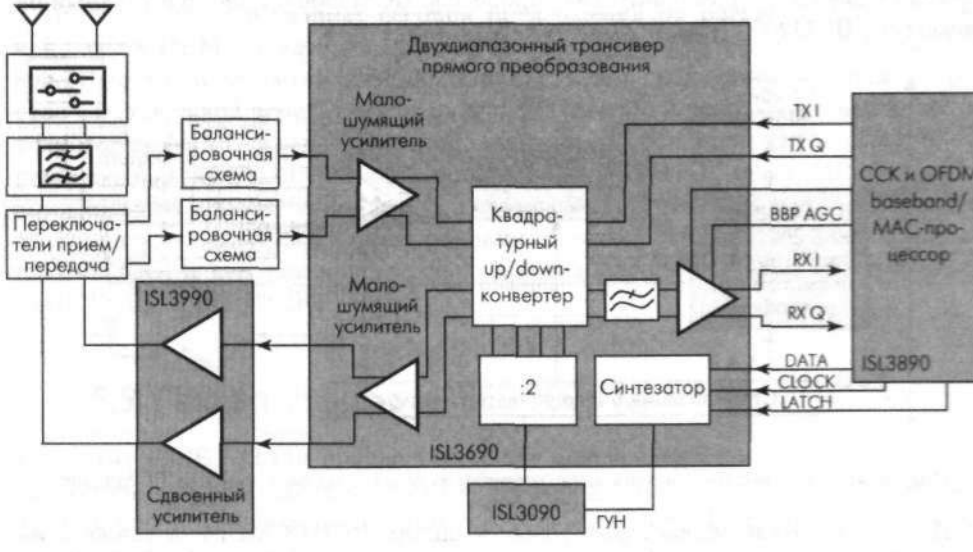


Рис. 5.24. Чипсет PRISM Duette

Технология PRISM Nitro реализуется на уровне системного программного обеспечения. Она специально разрабатывалась для применения совместно с чипсетами PRISM Duette и GT (рис. 5.24). Ее создатели утверждают, что

она полностью соответствует требованиям спецификации IEEE 802.11g. PRISM Nitro была, в частности, использована в маршрутизаторе CONNECT2AIR и сетевых картах компании Fujitsu Siemens Computers. Отметим, что эта компания первой применила новейший тогда процессор фирмы Intersil для точек доступа/маршрутизаторов беспроводных сетей ISL3893. Этот процессор, построенный на базе микропроцессорного ядра ARM9, реализует функции как сетевого процессора беспроводной сети, так и сети Ethernet. Он предназначен для работы совместно с чипсетами PRISM GT и PRISM Duette.

5.8. Аппаратная поддержка IEEE 802.11g

Устройства IEEE 802.11g с 2002 года производят такие компании, как Buffalo Technologies, Linksys (вошла в Cisco Systems), D-Link, Apple. Позднее к ним присоединились фирмы Netgear, Belkin, Actiontec, Proxim и многие другие. Такую возможность им предоставили производители наборов микросхем для 802.11g (прежде всего компании Intersil, Atheros Communications, Broadcom). Лидером в данной области выступает компания Intersil, продавшая в 2002 года компонентов для WLAN на 106 млн. долл. Однако наиболее динамичная корпорация в этой области — Broadcom: первый свой чипсет для 802.11b она представила в июле 2002 года, а в ноябре уже объявила о серийном производстве чипсетов 802.11g. К марту 2003 года Broadcom продала 1,3 млн. чипсетов.

Первым набором микросхем для устройств IEEE 802.11g компании Broadcom стал чипсет для абонентских устройств BCM94306. Он входит в состав большого семейства продуктов компании для беспроводных сетей AirForce. Чипсет включает две ИС — baseband/MAC-процессор BCM4306 и однокристальный радиомодуль BCM2050 (рис. 5.25).

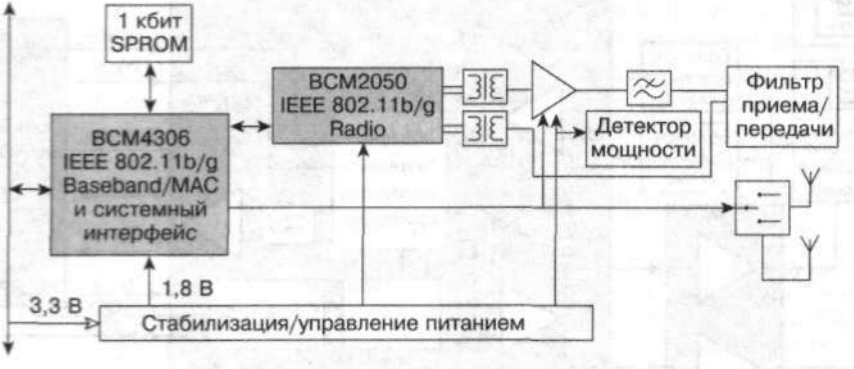


Рис. 5.25. Построение абонентского устройства 802.11g на базе чипсета BCM94306

Процессор поддерживает интерфейс к шинам PCI/PCMCIA, а также асинхронный последовательный интерфейс для коммутации с микросхемами, обеспечивающими связь посредством протоколов Bluetooth и GPRS (у компании Broadcom есть для этого однокристальные решения). Отметим, что в семейство AirForce входит широкий спектр продуктов — как абонентские устройства (рис. 5.26) различного исполнения (CardBus, Mini PCI, USB и др.), так и точки

доступа/маршрутизаторы с поддержкой высокоскоростных проводных интерфейсов [29].

Практически одновременно с компанией Broadcom к производству микросхем для устройств 802.11g приступили корпорации Intersil и Atheros Communications. Intersil выпустила чипсеты PRISM Duette (802.11a/b/g) (см. рис. 5.24) и PRISM GT (802.11b/g). Компания Atheros представила на рынок чипсет AR5001X Combo с поддержкой 802.11a/b/g. В марте 2002 года это было первое промышленное решение «три в одном» (рис. 5.27), поскольку PRISM Duette компания Intersil анонсировала в октябре 2002 года.

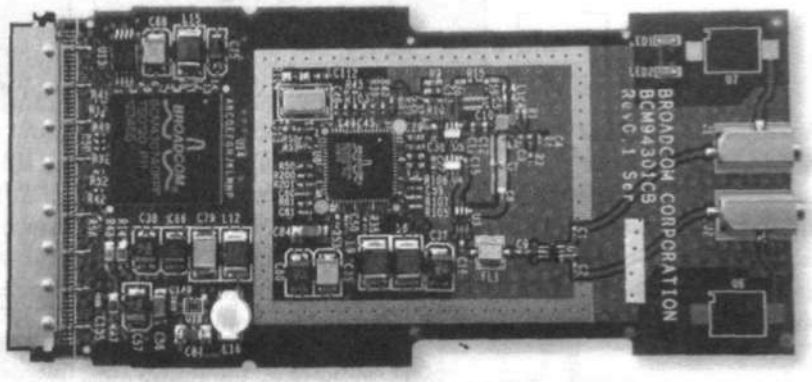


Рис. 5.26. *Сетевая абонентская карта BCM94306CB стандарта 802.11b/g в формате* CardBus компании Broadcom

На базе многочисленных сегодня предложений от разработчиков элементной базы и OEM-модулей создано множество устройств — как для сетей IEEE 802.11b/g, так и для сетей IEEE 802.11a/b/g. Вскоре появились серийные устройства с поддержкой агрегатной скорости 108 Мбит/с и использованием антенной технологии MIMO — предвестники грядущего стандарта IEEE 802.11n. Характерный пример — сетевой адаптер DWL-G650M Wireless MIMO Cardbus компании D-Link. Он строится на основе чипсета AGN300 фирмы Airgo Networks, в состав которого входят однокристальный MAC- и baseband-процессор (AGN303BB) и два однокристальных двухдиапазонных (2,4/5 ГГц) трансивера (AGN301RF и AGN302RF).

5.9. Проект стандарта IEEE 802.11n

Стандарт IEEE 802.11n предназначен для дальнейшего расширения диапазона скоростей передачи данных — до 100 Мбит/с и выше. В целом же он основывается на рассмотренном выше стандарте (дополнении) IEEE 802.11a, поскольку именно в нем описана технология OFDM. Увеличение скорости передачи данных в проекте IEEE 802.11n базируется на двух физических принципах — удвоении полосы пропускания канала, с 20 до 40 МГц, и введении дополнительных антенных каналов приема-передачи (технология многоканальных антенных систем MIMO).

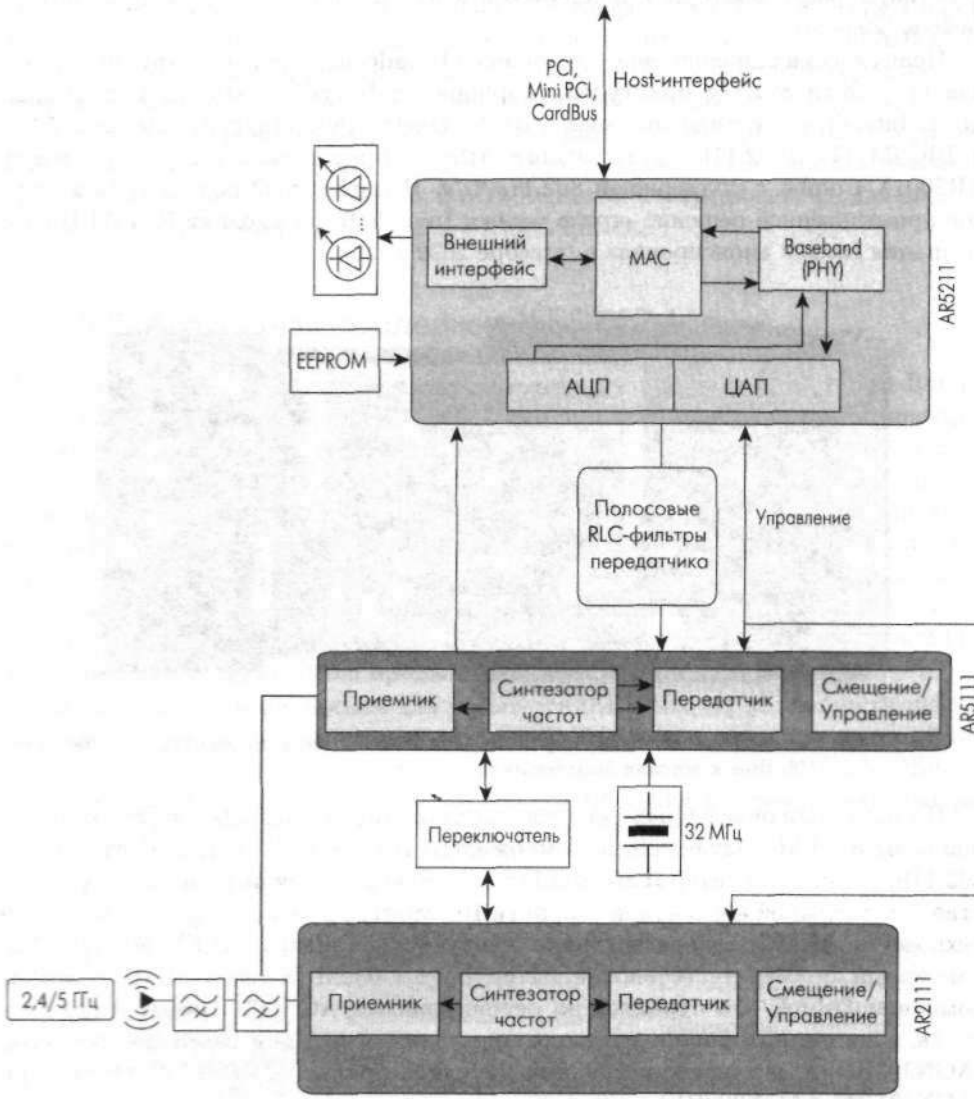


Рис. 5.27. Чипсет AR5001X Combo с поддержкой 802.11a/b/g компании Atheros

История развития проекта IEEE 802.11n — это история умения договариваться. Работы по созданию будущего стандарта IEEE 802.11n официально начались 11 сентября 2002 года, через год образовалась целевая группа TGn. Весной 2005 года ситуация с выбором окончательного варианта зашла в тупик. Входящие в TGn компании фактически разбились на два примерно равных лагеря — WWiSE (World-Wide Spectrum Efficiency) и TGn Sync (и еще небольшая группа MITMOT). Долгое время ни один из альтернативных вариантов не мог набрать необходимые для его утверждения 75% голосов (заметим, открытого голосования). Однако уже к осени 2005 года противоборствующим кланам удалось договориться и создать объединенное предложение, которое учло интересы

всех. В результате в январе 2006 года на регулярной конференции на Гавайях были утверждены устраивающие всех документы [30, 31] (описания физического и MAC-уровней нового стандарта), которые вошли в предварительный (draft) стандарт IEEE 802.11n. Казалось, что в 2007 году долгожданный стандарт будет утвержден. Но не тут-то было. Страсти разгорелись с новой силой, и лишь к 2008 году ситуация снова начала стабилизироваться. Вполне вероятно, что к концу 2009 года этот стандарт будет утвержден. Но производители окончательного его согласования дожидаться не стали, приступив к выпуску элементной базы и оборудования. Поэтому правомерно рассматривать основные принципы проекта IEEE 802.11n, основываясь на предварительных документах.

5.10. Отличия физического уровня

5.10.1. Каналы и режимы передачи

Проект IEEE 802.11n допускает как стандартные каналы с шириной полосы 20 МГц (как и всех предшествующих стандартов IEEE 802.11, включая IEEE 802.11a), так и расширенные до 40 МГц. Поскольку каналы шириной 40 МГц приемлемы не для всех стран, противореча национальной политике распределения частотных ресурсов, то их применение — это опциональная (необязательная) возможность. Собственно, введение обязательной поддержки 40-МГц каналов и было одним из камней преткновения в противостоянии групп WWiSE и TGn Sync.

Проект IEEE 802.11 предусматривает поддержку как традиционных режимов передачи (как в IEEE 802.11a), так и режимов с высокой пропускной способностью (HT — High Throughput). В традиционных (L — Legacy) режимах число поднесущих не изменено. В HT-режимах оно увеличено: в 20-МГц канале их 56, из них 52 — информационные и 4 пилотные. Только из-за этого скорость возрастает на 8%. Еще один фактор увеличения пропускной способности — повышение допустимой скорости кодирования до 5/6 (т. е. каждые 5 бит исходной последовательности превращаются в 6 бит кодированной). Опционально предусмотрена возможность двукратного сокращения длительности защитных интервалов GI в OFDM-символах — с 0,8 до 0,4 мкс. В результате скорость возрастает до 65 и 72,2 (опционально) Мбит/с.

Режим 20 МГц — обязательный, для него установлен базовый набор скоростей (табл. 5.3). В 40-МГц каналах поднесущих 114, из них 108 информационных и 6 пилотных. Один лишь этот фактор увеличивает пропускную способность канала на 125%.

Отметим, что традиционный режим также позволяет увеличивать (удваивать) скорость передачи данных. Однако это происходит при непосредственном удвоении полосы пропускания (40 МГц) — фактически передача ведется на двух смежных каналах IEEE 802.11a, используются $52 \times 2 = 104$ поднесущие (из 128 номинальных поднесущих не задействованы частоты с индексами от -5 по 5).

Таблица 5.3. Базовый набор скоростей проекта IEEE 802.11n

Модуляция	Скорость кодирования	Число битов на поднесущую	Кодированных битов на символ	Информационных битов на символ	Скорость передачи данных	
					GI = 0,8 мкс	GI = 0,4 мкс
BPSK	1/2	1	52	26	6,5	7,2
QPSK	1/2	2	104	52	13,0	14,4
QPSK	3/4	2	104	78	19,5	21,7
16-QAM	1/2	4	208	104	26,0	28,9
16-QAM	3/4	4	208	156	39,0	43,3
64-QAM	2/3	6	312	208	52,0	57,8
64-QAM	3/4	6	312	234	58,5	65,0
64-QAM	5/6	6	312	260	65,0	72,2

5.10.2. Формирование сигналов MIMO-OFDM

Основное отличие стандарта IEEE 802.11n от его предшественников — появление нескольких антенных каналов в приемнике и передатчике. Обязательный режим подразумевает поддержку двух антенных каналов оборудованием точек доступа (AP) и одного канала — пользовательскими (терминальными) станциями. Всего и у AP, и у терминальной станции может быть до четырех антенных каналов приема-передачи.

Поскольку проектом IEEE 802.11n предусмотрена технология MIMO, изменяется структурная схема передающего и приемного устройств (рис. 5.28). Данные после скремблирования поступают на сверточный кодер (как и в IEEE 802.11a). Если скорость передачи данных превышает 300 Мбит/с, используются два сверточных кодера. В отличие от IEEE 802.11a, поддерживается скорость кодирования 5/6. Кроме того, опционально вместо двоичного сверточного кодера (BCC) предусмотрено применение блокового кодера LDPCC (Low density parity check codes).

Дальше различия более серьезные. Кодированный поток битов разбивается на так называемые пространственные (пространственно-разделенные) потоки (spatial streams). Число таких потоков N_{SS} не может быть меньше, чем число антенных каналов в передатчике N_{TX}. Предусмотренный алгоритм распределения по потокам прост — входная последовательность разбивается на группы длиной

$$S = \sum_{i_{SS}=0}^{N_{SS}-1} s(i_{SS}),$$

где $s(i_{SS}) = \max(1, N_{BPSC}(i_{SS})/2)$ — число битов, определяющих действительную или мнимую составляющую комплексного модуляционного символа на одной поднесущей, $N_{BPSC}(i_{SS})$ — число битов на поднесущую в одном OFDM-символе. Если в каждом пространственном потоке используется одинаковая схема мультиплексирования, то $S = N_{SS} \cdot \max(1, N_{BPSC}/2)$.

Группы длиной S последовательно (по кругу) распределяются между N_{SS} пространственных потоков. Далее в каждом потоке происходит перемежение би-

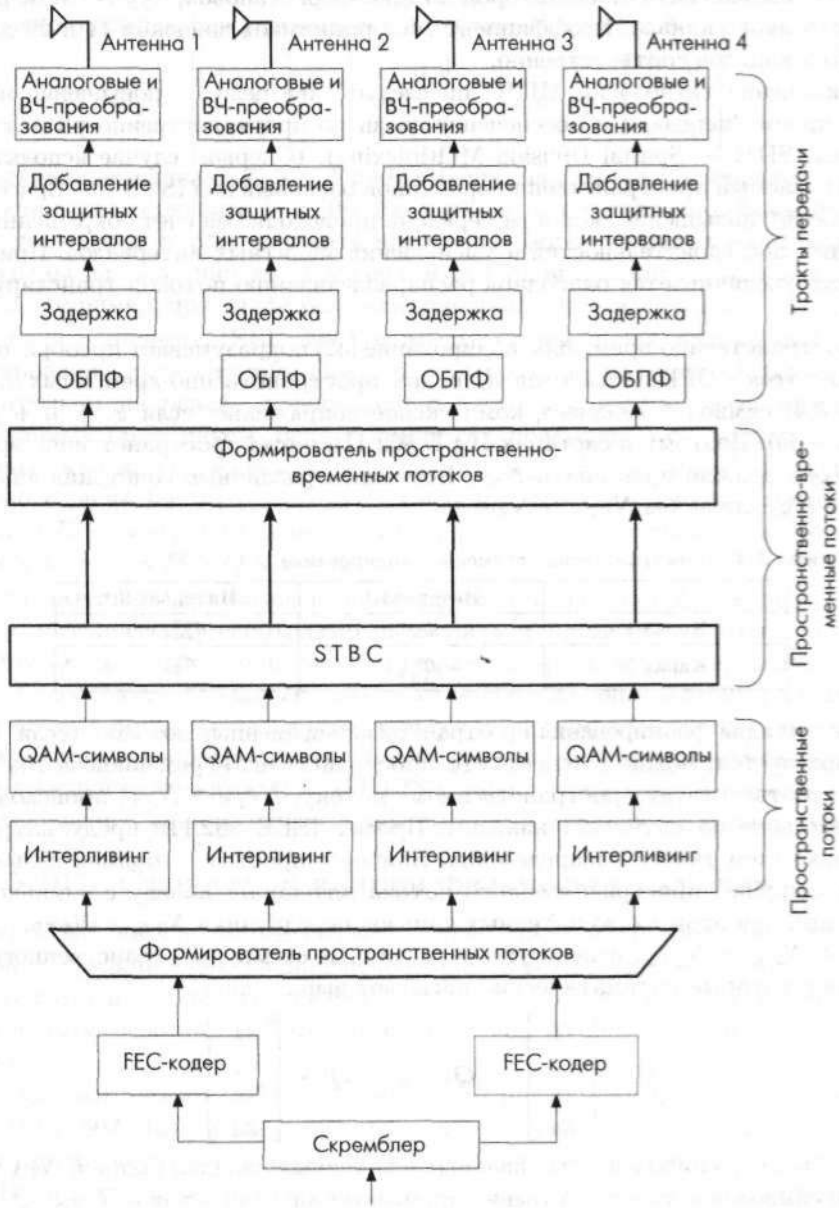


Рис. 5.28. Структура аппаратуры передачи стандарта IEEE 802.11n

тов, аналогично IEEE 802.11a (на первой стадии для перестановок применяется таблица размером 13 строк на $4N_{BPSC}$ столбцов для 20-МГц канала и $18 \times 4N_{BPSC}$ — для 40-МГц канала). Если использовано более одного пространственного потока, появляется третья стадия перемежения — частотная ротация. Она описывается выражением:

$$r = \{j - [2i_{SS} \bmod 3 + 3 \cdot \text{floor}(i_{SS}/3)] \cdot N_{rot} \cdot N_{BPSC}\} \cdot N_{CBPS},$$

где j — индекс бита после второй стадии перестановок; i_{SS} — номер пространственного канала. Коэффициент N_{rot} принимает значения 11 и 29 для 20- и 40-МГц каналов соответственно.

Применение технологии MIMO преследует две цели — повышение надежности приема/передачи и обеспечение связи по пространственно разделенным каналам (SDM — Spatial Division Multiplexing). В первом случае используется так называемый пространственно-временной блоковый код (STBC — Space Time Block Code), повышение скорости передачи происходит за счет сокращения проверочных последовательностей и уменьшения защитных интервалов. При SDM скорость увеличивается благодаря распараллеливанию потоков транслируемых данных.

Пространственно-временное кодирование [32] подразумевает преобразование одного потока OFDM-символов d_i в два пространственно-временных потока (табл. 5.4, символ * означает комплексное сопряжение: если $x = a + jb$, то $x^* = a - jb$). Поэтому в системах IEEE 802.11n число пространственно-временных N_{STS} должно превышать N_{SS}. Возможны различные сочетания значений N_{STS} и N_{SS}, лишь бы $N_{STS} > N_{SS}$.

Таблица 5.4. Пространственно-временное кодирование

	Интервал 1	Интервал 2
Канал 1	d_{2n}	d_{2n+1}
Канал 2	$-d^*_{2n+1}$	d^*_{2n}

После стадии формирования пространственно-временных потоков (если STBC не используется, можно считать, что пространственно-временной поток однозначно соответствует пространственному потоку, $N_{STS} = N_{SS}$) происходит их распределение по антенным каналам. Проект IEEE 802.11n предусматривает несколько схем такого распределения. Наиболее простая — прямое назначение (direct mapping) пространственного потока антенному каналу с одинаковыми номерами, при этом число антенных каналов передатчика $N_{TX} = N_{STS}$.

Если $N_{TX} > N_{STS}$, используются различные схемы пространственного расширения, которые математически описывает выражение:

$$\boldsymbol{r} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{bmatrix} = [\boldsymbol{Q}] \times \boldsymbol{x} = [\boldsymbol{Q}] \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix},$$

где $\boldsymbol{r}$ — вектор символов, передаваемых в антенных каналах (длиной N_{TX}); $\boldsymbol{x}$ — вектор символов в пространственно-временных каналах длиной N_{STS}; $\boldsymbol{Q}$ — матрица с N_{TX} строками и N_{STS} столбцами. В схеме прямого назначения $\boldsymbol{Q}$ — диагональная единичная матрица.

В случае двух пространственно-временных потоков и трех антенных трактов первый поток может передаваться через первый и третий антенные тракты, т.е. $\boldsymbol{Q} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$, при $N_{TX} = 3$ и $N_{STS} = 1$ $\boldsymbol{Q} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Отметим, что столбцы в матрицах $\boldsymbol{Q}$ должны быть ортогональными (например, на основе матриц Адамара).

Предусмотрен и режим, когда матрица $\boldsymbol{Q}$ принимает различные значения для каждой поднесущей k: в рассмотренном случае $N_{TX} = 3$ и $N_{STS} = 2$:

$$\boldsymbol{Q}_k = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{или} \quad \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

В антенных каналах 2, 3 и 4 информация передается с временной задержкой (относительно канала 1) в 400, 200 и 600 нс соответственно. Такой фазовый сдвиг необходим, чтобы избежать непроизвольного формирования стабильной диаграммы направленности передающей антенной системы, а это может произойти, например, при синфазном или противофазном излучении.

Однако в проекте IEEE 802.11n предусмотрен и так называемый режим формирования луча (beamforming). В этом режиме матрица $\boldsymbol{Q}$ формируется передатчиком на основе информации о состоянии канала связи между ним и выбранным приемником. Вектор принятого приемником совокупного по всем антенным каналам сигнала можно записать как $\boldsymbol{y} = [y_1, \ldots, y_{RX}]^T$. Вектор передаваемого сигнала $\boldsymbol{r} = [r_1, \ldots, r_{TX}]^T$. Тогда $\boldsymbol{y} = \boldsymbol{H} \times \boldsymbol{r} + \boldsymbol{n}$, где $\boldsymbol{n}$ — вектор шума в канале; $\boldsymbol{H}$ — матрица (размерности N_{RX}, N_{TX}) состояния канала. Поскольку $\boldsymbol{r} = \boldsymbol{Q} \times \boldsymbol{x}$, то $\boldsymbol{y} = \boldsymbol{H} \times \boldsymbol{Q} \times \boldsymbol{x} + \boldsymbol{n}$.

Следовательно, для компенсации задержек и затуханий в канале, описываемых матрицей $\boldsymbol{H}$, необходимо вычислить и применить соответствующую матрицу $\boldsymbol{Q}$. Для этого приемник вычисляет и транслирует передатчику либо матрицу состояния канала $\boldsymbol{H}$, либо уже подготовленные матрицы $\boldsymbol{Q}$. В качестве тестовой последовательности (заранее известной передатчику) выступают преамбулы пакетов физического уровня (точнее — длинные подстрочные последовательности режима HT, HT-LTF) либо специальные тестовые пакеты «прослушивания» канала, включая пакеты прослушивания без поля данных.

5.10.3. Структура кадров физического уровня

Проект IEEE 802.11n на физическом уровне предусматривает три структуры кадров — традиционную (совпадающую с IEEE 802.11a/g), смешанную и так называемое Зеленое поле (Green Field) (рис. 5.29). Эти структуры соответствуют одноименным режимам работы. Последние два из них называют скоростными (HT).

Смешанный режим (MM — mixed mode) предназначен для совместимости сетей IEEE 802.11a/g и IEEE 802.11n — традиционные устройства распознают присутствие своих высокоскоростных собратьев и определяют режим их передачи (например, длительность захвата канала), что позволяет предотвратить коллизии, особенно в случае скрытых станций. Для этого в структуре пакетов типа MM присутствуют как традиционные поля — короткая и длинная подстроечные последовательности (L-STF и L-LTF) и заголовок SIGNAL (L-SIG), так и поля, необходимые для HT-режима. К последним относятся управляющее поле HT-SIGNAL (HT-SIG), а также короткая и длинные подстроечные последовательности (HT-STF и HT-LTF). Временная задержка между каналами добавляется именно перед HT-полями.

Режим Зеленого поля (GF) (точнее — чистого, незагрязненного поля) позволяет передавать данные наиболее эффективно. В GF-кадрах присутствуют

только HT-поля, и традиционные приемники IEEE 802.11 такие посылки не воспримут. Поэтому данный режим опционален и может использоваться, только если все станции сети поддерживают IEEE 802.11n.

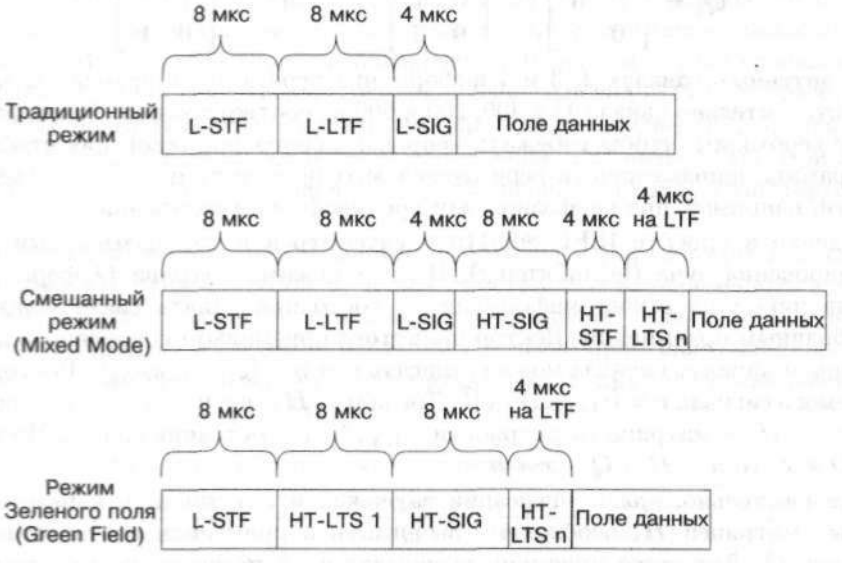


Рис. 5.29. Пакеты физического уровня стандарта IEEE 802.11n

Управляющее поле HT-SIGNAL состоит из двух частей по 24 бит каждая. Эти части кодируются (со скоростью 1/2), подвергаются перемежению и прочей обработке, как и в случае IEEE 802.11a. Поднесущие модулируются посредством BPSK, только модуляционные символы смещены на 90° относительно традиционного случая (приведенного на рис. 5.14), т. е. на диаграмме Грея значения битов 0 и 1 соответствуют -1 и 1 на квадратурной (мнимой) оси $\boldsymbol{Q}$. В результате поле HT-SIGNAL занимает два OFDM-символа.

Информация, передаваемая в HT-SIGNAL, — номер схемы кодирования/модуляции (MCS), признак ширины канала (20/40 МГц), длина поля данных, признак оценки канала smoothing (интегральная или на каждой поднесущей отдельно), признак пакетов прослушивания канала (sounding), признак объединения MAC-пакетов (aggregation), номер схемы STBC-кодирования, указатель типа кодирования (BCC/LDPCC), признак короткого защитного интервала в OFDM-символах (short GI), число дополнительных длинных подстроечных последовательностей (N HT-LTF), контрольная сумма CRC и разграничительные 6 бит (Tail) (рис. 5.30).

Короткая подстроечная последовательность аналогична традиционной (с поправкой на число используемых поднесущих). Длинные подстроечные последовательности служат для оценки каналов передачи (для каждого антенного тракта), поэтому их число не может быть меньше, чем число прстранственно-временных потоков NSTS. Кроме того, HT-LTF — это механизм прослушивания канала, поэтому в кадре может быть больше HT-LTF, чем необходимо для передачи поля данных именно этого кадра (т. е. больше, чем в данный момент используется пространственно-временных каналов). Эти дополнительные

последовательности предназначены для оценки каналов, которые передатчик предполагает использовать. Физически вышесказанное означает, что если данные в кадре передаются, например, через антенные тракты 1 и 2, то в них используют длинные подстроечные последовательности HT-LTF1 и HT-LTF2, а одновременно в незадействованных антенных трактах 3 и 4 могут транслироваться последовательности HT-LTF3 и HT-LTF4 для оценки соответствующих каналов. Длительность HT-LTF — 4 мкс, в режиме GF первая HT-LTF вдвое длиннее.

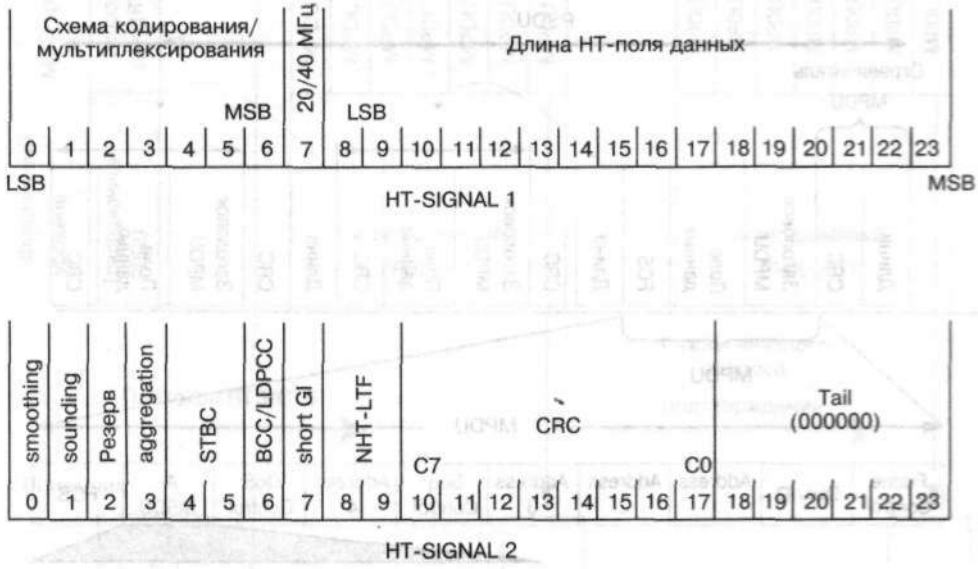


Рис. 5.30. Поле SIGNAL заголовка пакета физического уровня стандарта IEEE 802.11n

В результате применения дополнительных антенных трактов, а также 20- и 40-МГц каналов число возможных скоростей передачи и соответствующих им схем модуляции/кодирования существенно возрастает по отношению к базовому набору (см. табл. 5.3). Причем возможны как симметричные схемы модуляции (одинаковые в каждом пространственном потоке), так и несимметричные (вид модуляции в каждом потоке различен). Проект описывает по 32 симметричные схемы для 20- и 40-МГц каналов (для HT-режимов) и по 44 асимметричные схемы. В результате для 20-МГц каналов предусмотрены скорости до 288,9 и 600 Мбит/с — для 40-МГц каналов.

5.10.4. Особенности MAC-уровня

Повышение скорости передачи возможно не только за счет изменений на физическом уровне. MAC-уровень в проекте IEEE 802.11n также модернизирован. Одна из важнейших особенностей IEEE 802.11n — возможность объединения (агрегирования) нескольких MAC-пакетов (MPDU) в один пакет физического уровня PSDU (A-MPDU) (рис. 5.31). У такого объединенного пакета может быть только один адрес получателя. При этом повышается скорость передачи данных, поскольку сокращается удельный объем служебной информации (заго-

ловки и преамбулы пакетов физического уровня). Объединение пакетов — это обязательное требование, и только те MPDU, которые требуют индивидуального подтверждения приема (Ack), могут передаваться как отдельные пакеты. Более того, один MPDU может агрегировать несколько пакетов MSDU (MAC service data unit), принадлежащих различным сервисным потокам (различным приложениям) и обладающих различными требованиями к качеству предоставления услуг (QoS), лишь бы у них был единый адрес приемника. В результате формируется объединенный MAC-пакет A-MSDU (см. рис. 5.31).

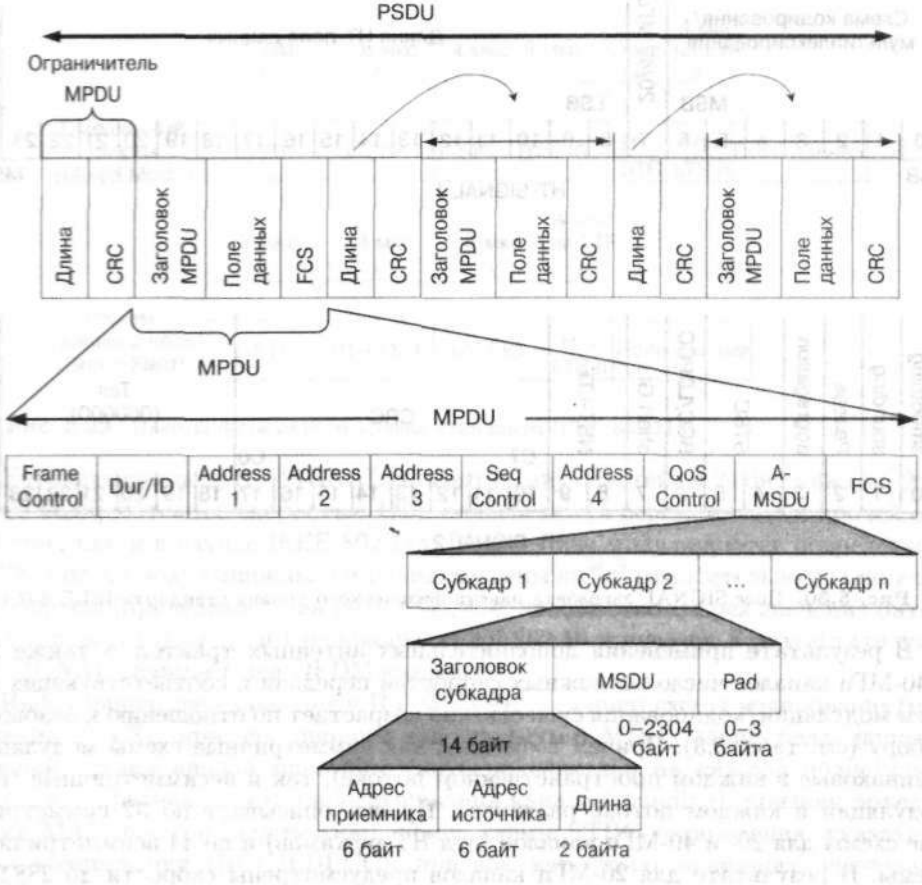


Рис. 5.31. Интегрированные заголовки MAC-уровня стандарта IEEE 802.11n

Для подтверждения приема пакетов физического уровня PSDU используется специальный пакет блочного подтверждения (Block Acknowledge — BA), описанный в стандарте IEEE 802.11e (посвящен обеспечению QoS). Однако в IEEE 802.11n применяют «сжатые» BA — в этом пакете предусмотренное стандартом IEEE 802.11e поле подтверждения размером 128 байт сокращено до 8 байт, причем каждый бит в этом поле подтверждает прием отдельного MSDU (рис. 5.32). Более того, необходимый запрос подтверждения (отдельный пакет BAR — BA request) может не использоваться, вместо этого достаточно задать определенный тип политики подтверждений. По утверждениям разработчиков, приведенный на

рис. 5.32 механизм обмена (объединенный пакет и блок подтверждения) более чем в 2,5 раза эффективнее, чем традиционный «данные–подтверждение».

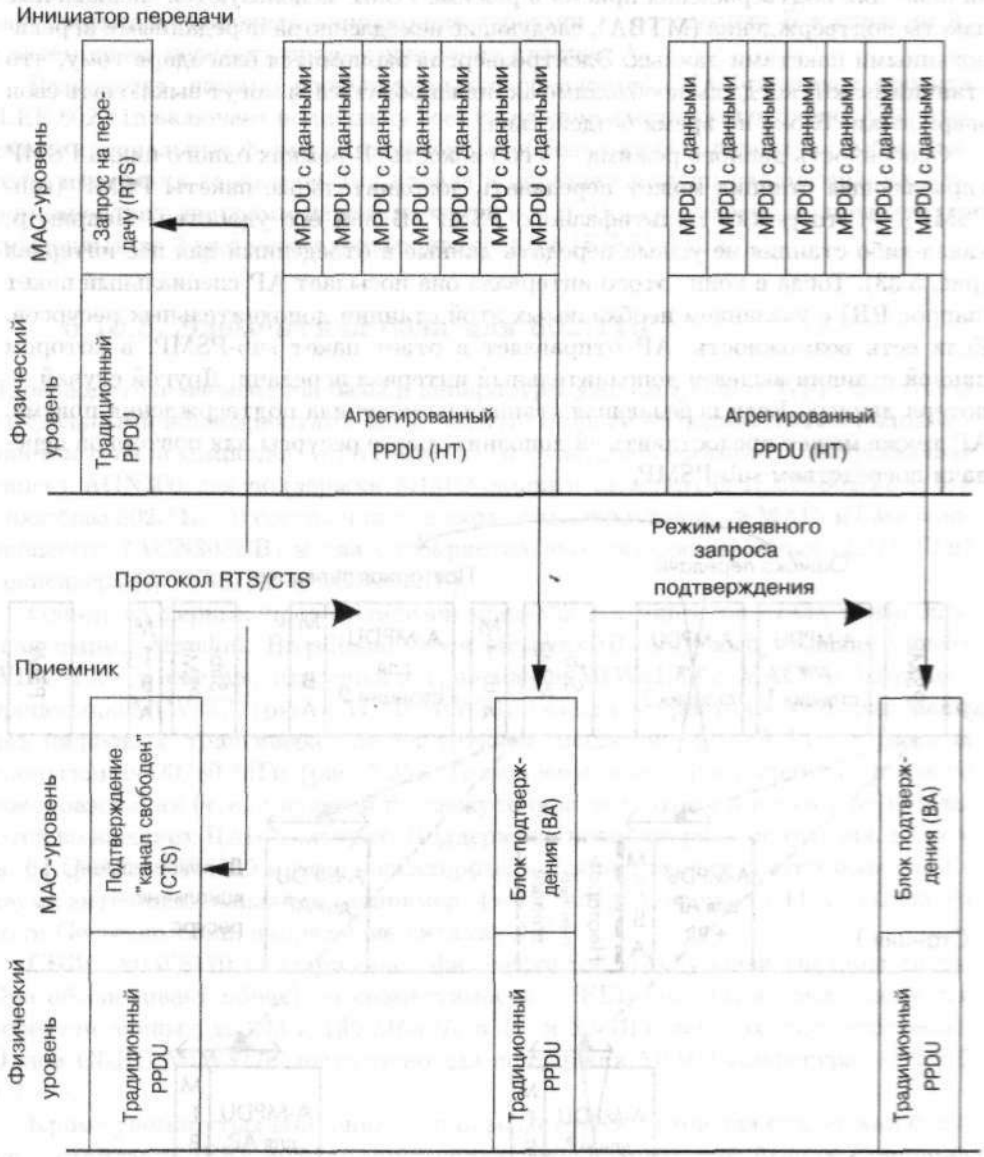


Рис. 5.32. Передача посредством агрегированных пакетов и блоков подтверждений

Важная особенность проекта IEEE 802.11n — режим PSMP (Power Save Multi-Poll) — энергосберегающий множественный опрос. Он похож на стандартный режим централизованного распределения ресурсов PCF — управляющая станция AP транслирует специальный управляющий пакет PSMP, в котором для каждой станции (из тех, которым разрешен режим PSMP) назначается время и длительность приема и передачи (нисходящего и восходящего соеди-

нений) в интервале обслуживания (цикле повторения приема-передачи). Причем длительность интервала обслуживания для каждой станции может быть различной. Для подтверждения приема в режиме PSMP используются специальные пакеты подтверждения (MTBA), следующие немедленно за переданными агрегированными пакетами данных. Электроэнергия экономится благодаря тому, что станции знают, когда им необходимо активизироваться, и могут выключать свои энергоемкие блоки на время бездействия.

Особенность данного режима — его гибкость. В рамках одного цикла PSMP управляющая станция может передавать дополнительные пакеты PSMP (sub-PSMP) — это режим мультифазного PSMP. В чем его удобство? Например, какая-либо станция не успела передать данные в отведенный для нее интервал (рис. 5.33). Тогда в конце этого интервала она посылает AP специальный пакет (запрос RR) с указанием необходимых этой станции дополнительных ресурсов. Если есть возможность, AP отправляет в ответ пакет sub-PSMP, в котором данной станции выделен дополнительный интервал передачи. Другой случай — потеря данных. Если передающая станция не получила подтверждения приема, AP также может предоставить ей дополнительные ресурсы для повторной передачи посредством sub-PSMP.

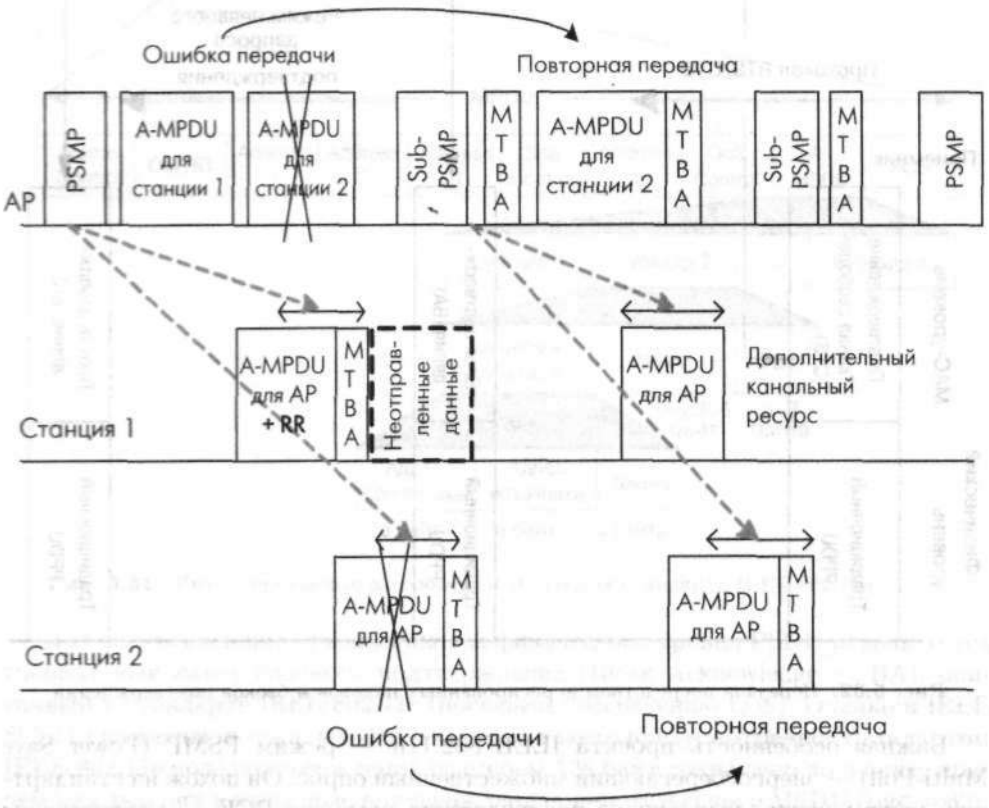


Рис. 5.33. Выделение дополнительного окна передачи для досылки информации и повторной отправки данных в режиме мультифазового PSMP

Не менее интересная особенность IEEE 802.11n — режим изменения направления обмена соединения. Пусть станция А передает данные станции Б. Оставаясь в рамках того же соединения, т. е. без дополнительных процедур доступа к каналу, возможно изменить направление передачи — от станции Б к станции А, а затем вновь передать права трансляции станции А.

Разумеется, помимо перечисленных особенностей, MAC-протокол проекта IEEE 802.11n включает поддержку всех изменений физического уровня. Он определяет специальные форматы пакетов калибровки канала передачи, выбора антенн, измерения характеристик канала и передачи матриц параметров канала, формирования диаграмм направленности и т. п. Предусмотрено применение сокращенных до 2 мкс межкадровых интервалов (RIFS) и т. д.

5.10.5. Элементная база для 802.11n

Производители элементной базы и аппаратуры уже выпускают Wi-Fi-продукты с отдельными возможностями IEEE 802.11n. Едва ли не первой в этом направлении выступила компания Airgo Networks (www.airgonetworks.com), выпускающая чипсет AGN300 для поддержки MIMO-опции в стандартах IEEE 802.11 a/b/g (прообраз 802.11n). В состав чипсета входят однокристальный MAC- и baseband-процессор (AGN303BB) и два однокристальных двухдиапазонных (2,4/5 ГГц) трансивера (AGN301RF и AGN302RF).

Одним из первых производителей пред-802.11n чипсетов стала и израильская фирма Metalink Broadband (www.MetalinkBB.com), выпустившая чипсет WLANPlus в составе сдвоенного трансивера MtW8150 и MAC- и baseband-процессора MtW8170 (рис. 5.34). В MtW8150 на одном кристалле интегрированы два цифровых трансивера для частотного диапазона 4,9–5,9 ГГц с полосой пропускания 20/40 МГц (рис. 5.35). Трансиверы используют принцип прямого преобразования (т. е. с нулевой промежуточной частотой), а потому не нуждаются во внешних ПАВ-фильтрах. Поддержана квадратурная модуляция, вплоть до 64-QAM. MtW8150 можно каскадировать для поддержки систем более чем с двумя антенными каналами (например, 4×4). Чип размером 11×11 мм выполнен по Si-Ge-технологии, напряжение питания 3 В.

СБИС MtW8170 поддерживает физический и MAC-уровни системы связи. Она обеспечивает обратную совместимость с IEEE 802.11a, а также скорости передачи данных до 243 и 135 Мбит/с в 40- и 20-МГц каналах соответственно. Одной СБИС MtW8170 достаточно для поддержки MIMO-конфигураций 2×2 и 2×3.

Кроме упомянутых компаний, к производству чипсетов для поддержки будущего стандарта IEEE 802.11n приступили такие фирмы, как Atheros Communications (чипсет AR5008) и Broadcom (чипсет INTENSI-FI). В состав последнего входит MAC- и baseband-процессор BCM4321 и трансивер BCM2055 (рис. 5.36), предназначенные для работы в диапазонах 2,4–2,5 и 4,9–5,85 ГГц.

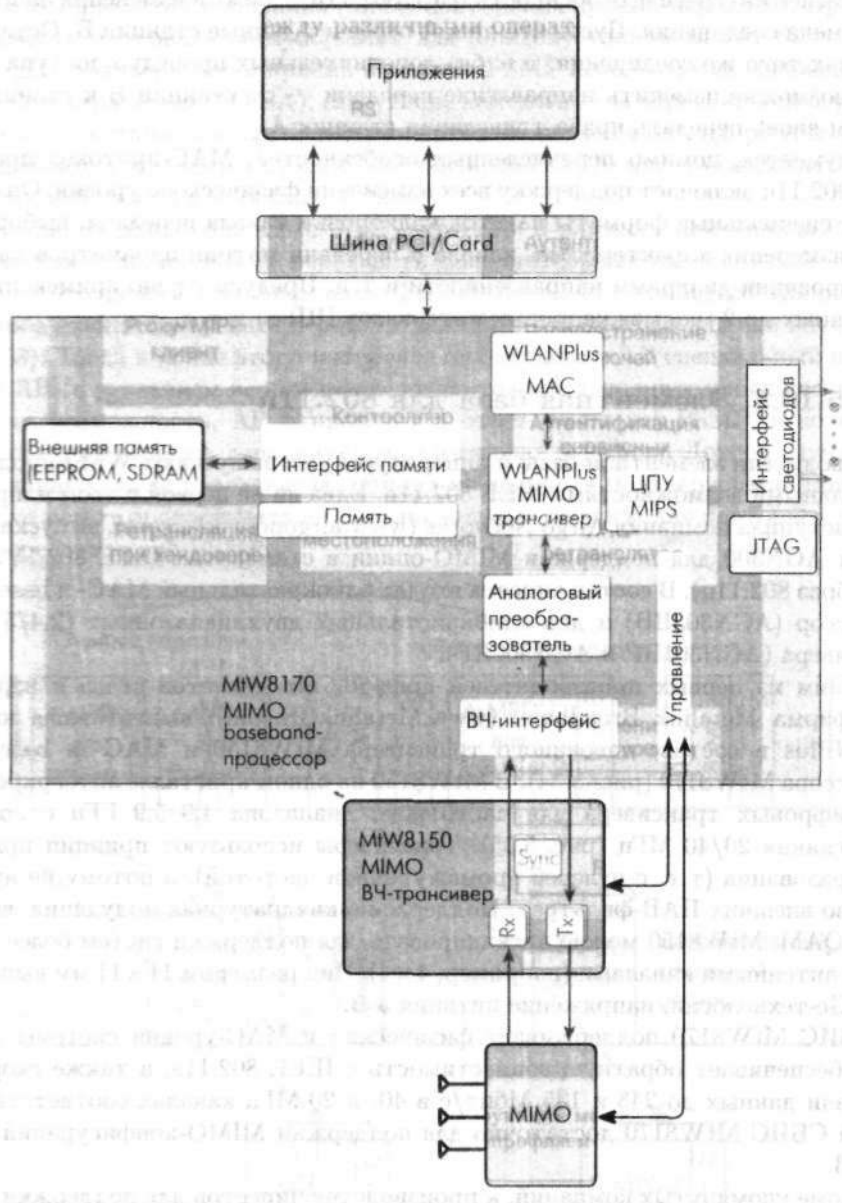


Рис. 5.34. Чипсет WLANPlus компании Metalink

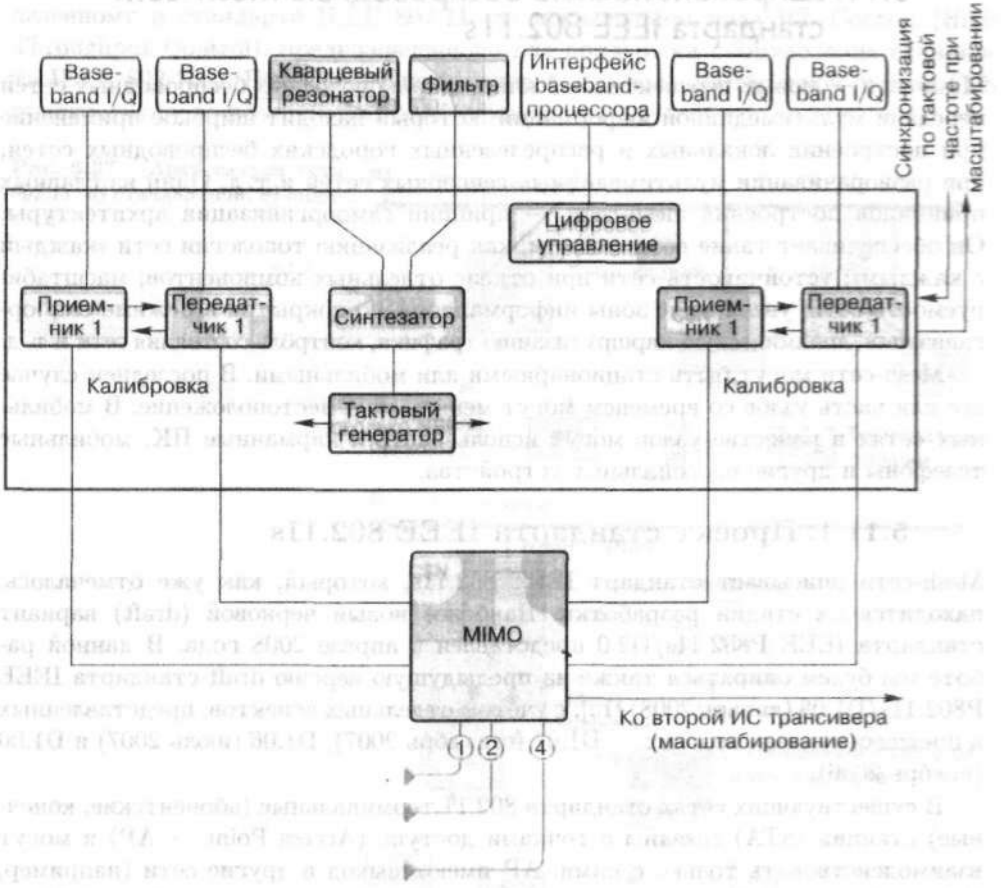


Рис. 5.35. Сдвоенный трансивер прямого преобразования MtW8150 компании Metalink

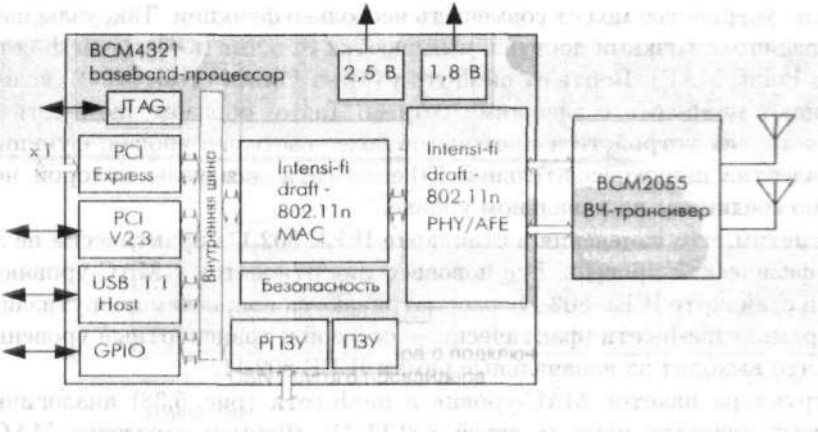


Рис. 5.36. Чипсет INTENSI-FI (трансивер и baseband-процессор) компании Broadcom

5.11. Широкополосные беспроводные mesh-сети стандарта IEEE 802.11s

Mesh-сети — новый перспективный класс широкополосных беспроводных сетей передачи мультимедийной информации, который находит широкое применение при построении локальных и распределенных городских беспроводных сетей, при разворачивании мультимедийных сенсорных сетей и т. д. Один из главных принципов построения mesh-сети — принцип самоорганизации архитектуры. Он обеспечивает такие возможности, как реализацию топологии сети «каждый с каждым»; устойчивость сети при отказе отдельных компонентов; масштабируемость сети, увеличение зоны информационного покрытия в режиме самоорганизации, динамическую маршрутизацию трафика, контроль состояния сети и т. д.

Mesh-сети могут быть стационарными или мобильными. В последнем случае все или часть узлов со временем могут менять свое местоположение. В мобильных сетях в качестве узлов могут использоваться карманные ПК, мобильные телефоны и другие персональные устройства.

5.11.1. Проект стандарта IEEE 802.11s

Mesh-сети описывает стандарт IEEE 802.11s, который, как уже отмечалось, находится на стадии разработки. Наиболее новый черновой (draft) вариант стандарта IEEE P802.11s/D2.0 представлен в апреле 2008 года. В данной работе мы будем опираться также на предыдущую версию draft-стандарта IEEE P802.11s/D1.08 (январь 2008) [15], с учетом отдельных аспектов, представленных в предшествующих версиях — D1.07 (сентябрь 2007), D1.06 (июль 2007) и D1.00 (ноябрь 2006).

В существующих сетях стандарта 802.11 терминальные (абонентские, конечные) станции (STA) связаны с точками доступа (Access Point — AP) и могут взаимодействовать только с ними. AP имеют выход в другие сети (например, Ethernet), но не могут обмениваться информацией друг с другом (рис. 5.37, *а*). В mesh-сети, помимо терминальных станций и точек доступа, присутствуют особые устройства — узлы mesh-сети (Mesh Point — MP), способные взаимодействовать друг с другом и поддерживать mesh-службы (рис. 5.37, *б*).

Одно устройство может совмещать несколько функций. Так, узлы mesh-сети, совмещенные с точками доступа, называются точками доступа mesh-сети (Mesh Access Point, MAP). Порталы mesh-сети (Mesh Point Portal, MPP), являясь MP, соединяют mesh-сеть с внешними сетями. Таким образом, mesh-сеть с точки зрения других устройств и протоколов более высокого уровня, функционально эквивалентна широковещательной Ethernet-сети, все узлы которой непосредственно соединены на канальном уровне.

Отметим, что изменения в стандарте IEEE 802.11s практически не затрагивают физический уровень. Все нововведения относятся к MAC-уровню. Кроме того, в стандарте IEEE 802.11s рассматриваются вопросы маршрутизации пакетов в рамках mesh-сети (фактически — сетевой и транспортный уровень модели OSI), что выходит за изначальные рамки IEEE 802.11.

Структура пакетов MAC-уровня в mesh-сети (рис. 5.38) аналогична стандартному формату пакетов сетей 802.11 [1]. Формат заголовка MAC-пакета в mesh-сети полностью соответствует MAC-заголовку пакета данных, опре-

деленному в стандарте IEEE 802.11 (за исключением поля HT Control (High Throughput Control), предназначенного для поддержки оборудования стандарта IEEE 802.11n). Первые три поля заголовка и поле контрольной суммы FCS присутствуют во всех пакетах MAC-уровня.

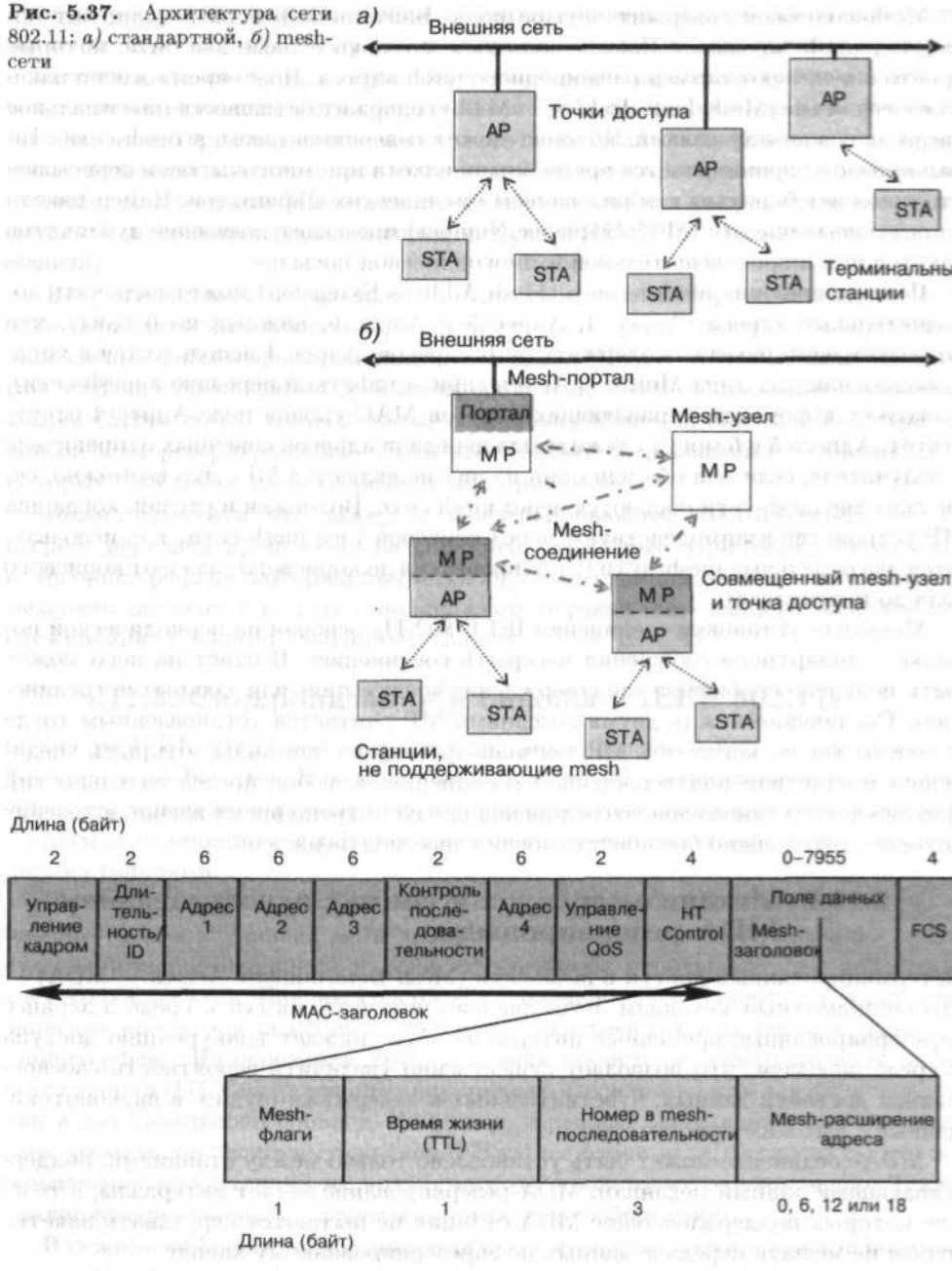


Рис. 5.37. Архитектура сети 802.11: *а)* стандартной, *б)* mesh-сети

Рис. 5.38. Формат MAC-кадра с mesh-заголовком

Отличие MAC-пакетов IEEE 802.11s заключается в наличии mesh-заголовка в начале поля данных. Этот заголовок присутствует в пакетах данных тогда и только тогда, когда они передаются от mesh-узла к mesh-узлу по установленному между ними соединению, он также присоединяется к одному из типов (Multihop Action) управляющих пакетов.

Mesh-заголовок содержит четыре поля. Байт mesh-флагов регулируют обработку mesh-заголовка. Пока используются только первые два бита, которые просто определяют размер расширенного mesh-адреса. Поле «время жизни пакета в mesh-сети» (Mesh Time To Live — MTL) содержит оставшееся максимальное число шагов между узлами, которое может совершить пакет в mesh-сети. Таким образом ограничивается время жизни пакета при многошаговой пересылке, что помогает бороться с образованием циклических маршрутов. Номер пакета в последовательности (Mesh Sequence Number) пресекает появление дубликатов пакетов при широковещательной и многоадресной посылке.

Поле расширения mesh-адреса (Mesh Address Extension) может включать дополнительные адреса (Адрес 4, Адрес 5 и Адрес 6, каждый по 6 байт), что позволяет mesh-пакетам содержать до 6 адресов. Адрес 4 используется в управляющих пакетах типа Multihop Action (при эстафетной передаче в mesh-сети), поскольку в формате управляющих пакетов MAC-уровня поле Адрес 4 отсутствует. Адреса 5 и 6 могут служить для передачи адресов конечных отправителя и получателя, если они оба или один из них не являются MP. Это возможно, если узлы вне mesh-сети общаются через mesh-сеть. Возможен и случай, когда два MP-устройства взаимодействуют через корневой узел mesh-сети, т. е. используются два отдельных mesh-пути (от отправителя до корневого узла и от корневого узла до получателя).

Механизм установки соединений IEEE 802.11s основан на периодической посылке стандартного сообщения «открыть соединение». В ответ на него может быть получено сообщение «подтверждение соединения» или «закрытие соединения». Соединение между двумя соседними MP считается установленным тогда и только тогда, когда оба MP послали друг другу команды «открыть соединение» и ответили подтверждением соединения (в любой последовательности). Для каждого установленного соединения предусмотрено время жизни, в течение которого оно должно быть использовано либо подтверждено.

5.11.2. Механизм доступа к среде с использованием MDA-резервирования

Детерминированный доступ в mesh-сети (Mesh Deterministic Access — MDA) — это опциональный механизм, позволяющий получать доступ к среде в заранее зарезервированные временные интервалы. Это снижает конкуренцию доступа к среде передачи, что позволяет существенно увеличить вероятность своевременной доставки данных, чувствительных к задержкам (аудио- и видеопотоки, данные с высоким приоритетом и т. п.).

MDA-соединение может быть установлено только между станциями, поддерживающими данный механизм. MDA-резервирование задает интервалы, в течение которых поддерживающие MDA-станции не пытаются передавать пакеты, чтобы не мешать передаче данных по зарезервированному каналу.

Создание MDA-соединения инициируется узлом-источником, а принимается или отклоняется узлом-адресатом данных. Для установления MDA-соединения

устройство выбирает интервалы времени, не занятые другими MDA-соединениями, о которых ему известно. Если резервирование нового соединения вызывает превышение допустимого числа MDA-соединений для соседей данного узла, то устройство отказывается от создания соединения.

При установлении MDA-соединения узел направляет соответствующий запрос узлу-адресату, указав, в какие моменты времени и какой длины интервалы он хочет использовать. Получатель запроса выполняет аналогичные проверки о допустимости создания соединения с запрошенными параметрами и шлет ответ — положительный или отрицательный. Если выбранные узлом-источником временные интервалы пересекаются с другими MDA-соединениями, о которых известно получателю, он может в ответе предложить альтернативные интервалы времени. Для разрыва MDA-соединения получатель или отправитель могут послать специальный информационный элемент (MDAOP Set Teardown information element).

Все устройства, которые знают о существовании MDA-резервирований, обязаны периодически сообщать о них своим соседям (рекламировать) посредством либо специальных информационных элементов (MDAOP Advertisements information element), включенных в биконы, либо используя специальные служебные кадры (MDA action frame). Узлы, поддерживающие механизм MDA, хранят список всех резервирований, о которых узнают из рекламных сообщений и в которых участвуют сами (передают или принимают).

Важно отметить, что даже при зарезервированном MDA-интервале доступ к среде передачи происходит на конкурентной основе. При этом учитывается категория трафика, который пытается передать станция, для чего используется механизм доступа к каналу с поддержкой дифференцированного качества обслуживания (Enhanced Distributed Channel Access, EDCA).

5.11.3. Синхронизация и биконы в IEEE 802.11s

Стандарт IEEE 802.11 поддерживает два режима работы беспроводных сетей: hot spot и ad hoc [16]. В режиме hot spot одна из станций работает в качестве точки доступа, и данные могут передаваться только между точкой доступа и другими станциями сети. В режиме ad hoc передача возможна между любыми двумя станциями.

В режиме hot spot точка доступа регулярно рассылает специальные кадры — биконы (beacon), главная цель которых заключается в синхронизации часов станций и информировании о сервисах и режимах работы, которые поддерживает точка доступа. Биконы содержат специальное поле Timestamp, в котором записано время, когда первый бит бикона оказывается переданным через радиоинтерфейс. На основании этого значения происходит синхронизация часов всех станций [17]. Синхронизация внутренних часов важна как для физического, так и для канального уровней. Например, в режиме модуляции с расширением спектра методов частотных скачков (FHSS) необходимо гарантировать, что переключение всех станций на новую частоту происходит одновременно. Также синхронизация важна для работы режима энергосбережения.

В режиме ad hoc биконы выполняют ту же функцию, что и в режиме hot spot. Но процесс передачи бикона является распределенным, т.е. в нем участвуют все станции. Станция, которая организует сеть ad hoc, задает серию момен-

тов времени, которые называют ожидаемым временем передачи бикона (Target Beacon Transmission Time, TBTT). Последовательные моменты TBTT отделены друг от друга равными интервалами времени — бикон-интервалами. В каждый момент TBTT начинается так называемое ATIM-окно (Announcement Traffic Indication Message — сообщение уведомления о трафике), во время которого могут быть переданы только биконы или ATIM-кадры (используются механизмом энергосбережения), в то время как трансляция других пакетов запрещена для снижения вероятности коллизии.

Передача бикона основана на том же механизме конкурентного доступа с контролем несущей, что и при передаче данных. В момент TBTT каждая из станций замораживает счетчик времени отсрочки передачи данных и инициализирует таймер передачи бикона случайно выбранным числом слотов (единица дискретного времени в сети 802.11), равномерно распределенным в интервале от нуля до некоей константы $(2 \cdot aCWmin)$. Если среда передачи не занята в течение слота, станция уменьшает значение таймера на единицу. Если одна из станций начинает передачу, другие станции замораживают свои таймеры на время передачи плюс интервал времени DIFS. Если происходит коллизия, т. е. более одной станции передают одновременно, то вместо времени DIFS используется более длинный интервал EIFS. Станция начинает передачу бикона в момент, когда значение ее таймера становится равным нулю. При получении бикона от любой из станций все остальные станции отменяют передачу своих биконов.

Алгоритм посылки биконов, остающийся неизменным в течение 10 лет существования стандарта IEEE 802.11, используется и в стандарте IEEE 802.11s. Именно этот механизм поддерживает глобальную синхронизацию сети, когда все устройства работают по единому времени, привязанному к ожидаемому времени передачи бикона.

Узлы mesh-сети MP могут, но не обязаны поддерживать глобальную синхронизацию в сети. Соответственно, они подразделяются на синхронные и асинхронные MP. Асинхронные MP передают биконы подобно точкам доступа в сетях hot spot. При этом каждая станция поддерживает независимо от других станций серию моментов TBTT и не подводит свои часы при получении биконов. Синхронные MP стараются поддерживать общее для всех время Mesh TSF.

Синхронные MP передают биконы по тому же алгоритму, что и в сетях ad hoc, за исключением следующего аспекта. Если MP получило бикон от соседнего устройства mesh-сети, оно может отменить запланированную передачу собственного бикона, но не обязано это делать, как в ad hoc-сети. В mesh-сети одного бикона от случайно выбранного MP может оказаться недостаточно.

По сравнению с сетями ad hoc, mesh-сети поддерживают дополнительные mesh-сервисы, и биконы ответственны за их поддержку. Например, механизм детерминированного доступа MDA использует биконы для передачи в них специального информационного элемента MDAOP Advertisements с рекламой MDA-резервирований. Этот и другие дополнительные информационные элементы делают биконы в mesh-сети более индивидуальными, по сравнению с биконами в сетях ad hoc, которые разнятся только значением временной метки (поля бикон-кадра, описывающие возможные режимы работы станции, не меняются в течение всего времени существования сети ad hoc). Потому важно, чтобы каждое MP отправляло свой бикон как можно чаще.

В дополнение к алгоритму рассылки биконов, используемого в сетях hot spot и ad hoc, в первой версии 802.11s/D1.00 было введено понятие распространителя биконов (точки биконов) — Beacon Broadcaster (BB). Когда BB выбран, оставшиеся MP биконы не передают. Роль BB периодически передается от одного MP другому. Однако в mesh-сети некоторые станции скрыты друг от друга, что приводит к появлению нескольких BB, и задача ротации BB становится слишком сложной. В связи с этим в более поздних версиях стандарта mesh-сетей использование BB исключено.

В текущей версии IEEE 802.11s [15] сделан еще один шаг в сторону от использования принципа глобальной синхронизации mesh-сети. Дело в том, что в mesh-сети глобальная синхронизация требует больших издержек: размер ATIM-окна должен быть увеличен по сравнению с сетями ad hoc, чтобы уместить возможно большее число биконов MP. Поэтому вместо поддержки глобальной синхронизации MP могут лишь поддерживать синхронизацию попарно. При этом MP рассылают биконы независимо, без привязки к единому времени TBTT и единому ATIM-окну. Издержки сети при этом могут уменьшиться, но, по-видимому, качество ее работы снизится: без глобальной синхронизации сети трудно защитить биконы от коллизий с данными, а значит, нельзя обеспечить качество обслуживания и эффективную работу режима сохранения энергии.

Работа над дополнением к стандарту IEEE 802.11s еще не завершена. Пока не ясно, какая парадигма синхронизации будет принята в mesh-сети. Но все же отметим, что глобальная синхронизация mesh-сети и алгоритм рассылки биконов, похожий на алгоритм в сетях ad hoc, позволят обеспечить качество обслуживания (QoS) в рамках всей mesh-сети, а также позволят применять эффективные методы энергосбережения, все более востребованные на рынке телефонов, коммуникаторов и т. п. В сети без глобальной синхронизации обеспечение QoS представляется чрезвычайно трудной, если вообще разрешимой задачей. Поэтому сегодня основное внимание уделено изучению алгоритма рассылки биконов синхронными MP, как они описаны в первой завершенной версии стандарта IEEE 802.11s/D1.00 [18].

5.11.4. Энергосбережение в IEEE 802.11s

Режим энергосбережения в mesh-сетях является опциональным. Так, MAP-узлы всегда активны, поскольку в любой момент к ним могут обратиться устройства, не поддерживающие 802.11s и соответствующий режим энергосбережения. Однако для устройств с автономным питанием (разного рода датчики, ноутбуки, телефоны и т. п.) сбережение энергии — актуальная задача.

Узлы сети обязаны сообщать о своей способности поддерживать спящий (энергосберегающий) режим. Для этого используется информационное поле возможностей (capability information field) в биконах и в ответах на пробные пакеты. В этом же поле сообщается, что узел находится в режиме энергосбережения либо имеет связь с узлом, который пребывает в этом режиме. Если устройство, желающее работать в режиме энергосбережения, видит, что его сосед не поддерживает эту возможность, то оно может либо не устанавливать соединения с таким устройством, либо установить его, но отказаться от перехода в режим энергосбережения. Узел не может переходить из активного режима в режим энергосбережения (и обратно), пока не проинформирует все устройства, с ко-

торыми у него установлено соединение, о своем желании переключиться. Для информирования соседей о смене режима энергосбережения используются пустые пакеты данных (null-data frame).

Узел в спящем режиме периодически просыпается, чтобы получить биконы от своих соседей либо послать свои. Узел просыпается по крайней мере один раз за так называемый DTIM-интервал (delivery traffic indication message — сообщение о наличии пакетов для станции) и остается активным в промежутке времени окна ATIM (Announcement Traffic Indication Message — окно для сообщений о трафике). Все узлы mesh-сети, поддерживающие режим энергосбережения, откладывают посылку пакетов, предназначенных для устройств в спящем режиме (в том числе широковещательных и многоадресных) и отправляют их только в назначенный промежуток времени. О наличии этих пакетов узел-источник сообщает в сообщении Mesh TIM в биконе или в передаваемом ATIM-кадре, следующем за DTIM-биконом. Устройства, находящиеся в режиме энергосбережения, слушают такие сообщения о наличии для них данных, и, если обнаруживают их, остаются активными после ATIM-окна. Если узел получил широковещательный или многоадресный пакет, то он остается активным до тех пор, пока не получит пакета, в котором поле о наличии данных (More Data field) говорит о том, что адресованных ему данных более не осталось, либо Mesh TIM элемент с той же информацией.

Спящие узлы могут проснуться в любой момент времени, если у них в очереди оказывается пакет на передачу. В этом случае такой узел остается бодрствовать, по крайней мере, до следующего момента времени TBTT.

Режим энергосбережения отличается для синхронных и асинхронных MP. Так, асинхронные MP используют свои собственные значения ATIM и DTIM, а все узлы, с которыми они установили соединение, сохраняют эти параметры для дальнейшей работы. Синхронные же MP, присоединяясь к сети, используют общие ATIM и DTIM значения, которые они получают в биконах от соседей, в этом случае все спящие устройства в сети будут просыпаться одновременно.

5.11.5. Маршрутизация в широкополосных беспроводных mesh-сетях стандарта IEEE 802.11s

Вопросы, связанные с маршрутизацией в mesh-сети, играют существенную роль — им посвящено более 20% стандарта IEEE 802.11s. Такое большое внимание к маршрутизации связано со сложной топологией, высокой мобильностью, большим количеством устройств и прочими особенностями mesh-сетей, которые делают нахождение и поддержание кратчайшего пути до адресата важнейшим механизмом, необходимым для эффективной работы сети.

Для выбора оптимальных маршрутов в сети используются различные критерии (метрики). Метрики могут включать в себя такую информацию, как длина пути (количество шагов), надежность, задержка, пропускная способность, загрузка, стоимость передачи трафика и так далее. Наиболее распространенной метрикой является длина пути. Некоторые протоколы маршрутизации позволяют администратору сети присвоить каналу (путь длиной в один шаг) произвольную длину. При этом длина пути — это сумма длин каналов, через которые пролегает путь от источника (отправителя) к адресату (получателю). Другие

протоколы определяют число шагов — сколько сетевых устройств (например, маршрутизаторов) должен пройти пакет на своем пути к получателю.

Еще один критерий выбора оптимальных маршрутов — надежность. Под метрикой «надежность» обычно подразумевается доля потерь пакетов в каждом из каналов. Некоторые каналы разрываются чаще, чем другие. Или восстанавливаются проще или быстрее после ошибки в работе сети. Любые факторы надежности могут учитываться при получении численного значения данной метрики.

Другая популярная метрика — задержка, т. е. время, необходимое для доставки пакета от отправителя к получателю. Задержка зависит от многих факторов, включая пропускную способность каналов, очереди в портах устройств на пути пакета, загрузку сети во всех промежуточных каналах, а также физическое расстояние, которое нужно преодолеть. Поскольку задержка зависит от ряда важных факторов, это распространенная и полезная метрика.

Пропускная способность также часто используется как критерий выбора пути. Под ней подразумевается объем данных, который может быть передан по сети в единицу времени.

Метрика, напрямую связанная с пропускной способностью, — это загрузка, которая отражает степень занятости сетевых ресурсов, таких как каналы и маршрутизаторы. Загрузку можно вычислить различными способами, включая загрузку процессора и число обрабатываемых или передаваемых в секунду пакетов. Следует отметить, что постоянный анализ этих показателей сам по себе может потребовать значительные ресурсы сетевого оборудования.

Существенно отличается от перечисленных выше критериев метрика «стоимость». Некоторые компании в целях экономии предпочитают использовать пути через собственные каналы, а не более высокопроизводительные, но платные каналы других операторов.

Метрика отдельных каналов может быть статической (задаваемой администратором сети) и динамической. Пример первой — метрика стоимости. Динамическая метрика может измеряться по уровню сигнала, задержке пакетов и множеству других параметров. Причем она может определяться как пассивно, без дополнительных служебных пакетов, так и использовать специальные «пробные» пакеты для сбора статистики по каждому каналу (задержки, потери и пр.).

Стандарт IEEE 802.11s требует, чтобы все устройства поддерживали метрику времени передачи в канале (Airtime Link Metric). Эта обязательная метрика необходима для совместимости устройств.

Метрика времени передачи в канале задается формулой $c_a = (O + \frac{B_t}{r})/(1 - e_f)$, где O и B_t — константы, определенные стандартом для различных физических реализаций (802.11a, 802.11b): B_t — число битов в тестовом пакете (8192), O — накладные расходы доступа к каналу, которые включают в себя заголовки пакетов, кадры протоколов доступа и т. д.; r — скорость передачи данных в канале (Мбит/с); e_f — вероятность возникновения ошибки (измеряется экспериментально на пакетах длиной B_t). Эта метрика представляет собой оценку времени передачи (в секундах) пробного пакета длиной B_t с учетом возможных ретрансляций при потерях в канале. Способ определения параметров r и e_f в стандарте не приводится, однако можно предположить, что для этого должна использоваться периодическая рассылка пробных пакетов длиной $B_t = 8192$ бит.

В основе метода выбора пути для передачи данных в стандарте IEEE 802.11s лежит механизм профилей. Этот механизм обеспечивает совместимость устройств от разных производителей, которые при этом могут поддерживать как стандартизованные механизмы, так и собственные. Профиль — это запись вида ⟨Идентификатор профиля⟩ ⟨Идентификатор протокола маршрутизации⟩ ⟨Идентификатор метрики протокола маршрутизации⟩. Устройство может поддерживать несколько профилей работы, но единовременно лишь один из них может быть активным. Обязательный для реализации профиль использует протокол HWMP и метрику времени передачи Airtime Link Metric.

Производители вольны реализовать собственные алгоритмы маршрутизации и метрики к ним, а также определять дополнительные проприетарные профили. Поэтому вопрос эффективности механизмов маршрутизации и метрик является одним из важнейших для разработчиков mesh-устройств.

Гибридный протокол маршрутизации HWMP (Hybrid Wireless Mesh Protocol) использует стандартный набор служебных пакетов, правил их создания и обработки, наподобие хорошо известного протокола дистанционно-векторной маршрутизации по запросу (Ad Hoc On Demand Distance Vector, AODV) [19]. Однако HWMP адаптирован для работы с адресами MAC-уровня и метриками путей. Гибридным он назван потому, что объединяет в себе два режима построения путей, которые могут быть использованы как по отдельности, так и одновременно в одной сети:

- реактивный режим — построение маршрутных таблиц в узлах mesh-сети непосредственно перед передачей данных (по запросу);
- проактивный режим — регулярная процедура обновления информации в маршрутных таблицах узлов всей сети. Процедуру инициирует корневой узел, в результате на сети строится граф (дерево) путей с вершиной в корневом узле.

В реактивном режиме HWMP узел отправляет широковещательный PREQ-пакет запроса пути (Path Request). Пути выбираются на основании метрики, для распространения информации о которой служит специальное поле в служебных пакетах запроса пути. Этот пакет распространяется через соседние узлы по всей сети, пока не будет достигнут узел-адресат. По мере продвижения от узла к узлу модифицируется поле метрики пути от текущего узла до отправителя. В итоге формируется полная метрика пути получатель–отправитель. Узел-адресат отправляет инициатору пакет подтверждения PREP (Path Reply), содержащий итоговое значение метрики пути «инициатор–получатель». Приняв его, узел-инициатор получает информацию об установленном пути.

Очевидно, что в mesh-сети широковещательные пакеты запроса проходят до получателя по множеству путей через различные узлы. При этом они могут начать передаваться по замкнутым маршрутам (циклам), не единожды проходя через какой-либо узел. Чтобы избежать такой ситуации, используется порядковый номер запроса. В стандарте IEEE 802.11s он именуется порядковым номером назначения (Destination Sequence Number), что вносит невероятную путаницу. Кроме DSN, стандарт оперирует понятием DSN инициатора (поиска пути) — Originator's DSN (OSN). Именно этот параметр и служит порядковым номером при рассылке пакетов поиска пути. Каждое mesh-устройство имеет

собственный DSN. Перед началом процедуры поиска пути DSN инициатора увеличивается на 1 и записывается в поле Originator's DSN пакета запроса PREQ. Кроме того, в пакете содержится адрес инициатора (адрес начала пути). Все узлы mesh-сети хранят информацию о каждом узле mesh-сети. Такая информация распространяется в служебных пакетах (в полях «адрес отправителя», «метрика пути», «порядковый номер»).

Узел, получив пакет PREQ, сравнивает значение OSN с ранее сохраненным значением для этого же отправителя (если таковое имеется). Узел принимает, обрабатывает и ретранслирует пакет PREQ, только если текущий OSN в пакете больше ранее сохраненного или они равны, но метрика пути ранее полученного пакета хуже, чем у вновь полученного (т. е. повторного приема и ретрансляции одного и того же пакета быть не может). В реактивном режиме пакеты подтверждения PREP может отправлять не только узел назначения, но и все промежуточные узлы, успешно принявшие пакет запроса PREQ (если в пакете установлены соответствующие флаги).

Помимо полей метрики пути, по мере прохождения от узла к узлу в пакете может изменяться значение поля «время жизни» (Time to Live, TTL) — число промежуточных узлов, которое может пройти данный пакет. Если этот параметр используется, он декрементируется в каждом узле следования. Когда $TTL = 0$, обработка и трансляция пакета прекращается.

Проактивный режим отличается от реактивного тем, что в сети назначается корневой узел (узлы). Этот узел периодически рассылает пакеты PREQ, которые распространяются по всей сети. Все узлы сети, принявшие проактивный PREQ, сохраняют адрес узла-отправителя (через который лежит путь к корневому узлу), широковещательно транслируют PREQ с измененными полями (поля метрики и TTL) и отправляют PREP корневому узлу (либо не отправляют, в зависимости от установок).

Помимо описанных методов выбора пути на основе пакетов PREQ/PREP, стандарт предусматривает процедуру на основе пакетов оповещения о корневом узле RANN (Root Announcement). Но этот метод принципиально не отличается от уже рассмотренного.

Возможны ситуации одновременного использования реактивного и проактивного режимов HWMP. Например, в сети штатно используется проактивный режим протокола HWMP, но какой-либо узел использует метод выбора пути по запросу для установления прямого соединения с другим заданным узлом.

Протокол маршрутизации HWMP обязателен для всех устройств стандарта IEEE 802.11s как протокол по умолчанию.

Помимо протокола маршрутизации HWMP, ранние версии IEEE 802.11s предполагали использование *стандарта RA-OLSR (Radio Aware OLSR)* — модификации оптимизированного протокола маршрутизации по состоянию канала OLSR (Optimized Link State Routing). OLSR — это описанный в документе IETF RFC 3626 [20] проактивный протокол маршрутизации для мобильных ad hoc-сетей. Он поддерживает маршрутные таблицы в узлах сети при помощи регулярных процедур обновления маршрутной информации в сети. Протокол эффективен для больших и плотных мобильных сетей.

OLSR основан на понятии многоточечной эстафеты MPR (MultiPoint Relay). Каждый узел сети m выбирает несколько узлов из числа своих соседей (т. е. из уз-

лов, с которыми у него установлено соединение). В итоге в сети формируется набор узлов MPR(m). Причем он формируется так, что все узлы, находящиеся в сфере с радиусом 2 шага от узла m (соседи соседей), имеют симметричные каналы с MPR(m). Это означает, что узлы MPR связаны со всеми узлами в сфере с радиусом 2 шага. MPR выбираются каждый раз, когда обнаруживается изменение в сфере с радиусом 1 или 2.

Каждый узел сети хранит свою таблицу маршрутизации, которую формирует на основании информации о топологии сети. Она распространяется по всей сети посредством служебных пакетов выбора маршрута Topology Control (TC). Причем только MPR-узлы участвуют в пересылке TC-пакетов, остальные узлы принимают и обрабатывают такие пакеты, но не пересылают их дальше.

Для каждого MPR формируется список соседних узлов, выбравших его в качестве MPR — список MPR Selectors (MPRS). Информация о MPRS передается в специальных HELLO-пакетах, которые передаются только между двумя соседними узлами. В сеть (в TC-пакетах) передается только информация о состоянии соединений между MPR и его MPRSs. Данный механизм позволяет существенно снизить число передач служебных пакетов по сравнению с лавинной рассылкой [20, 21].

В протоколе OLSR служебные сообщения содержат последовательные номера (аналог DSN в HWMP), которые увеличиваются в последующих сообщениях. Таким образом, получатель контрольного сообщения может при необходимости с легкостью определить, какая информация является более новой, даже если сообщения пришли в обратном порядке.

OLSR разработан как совершенно распределенный протокол и не зависит от каких-либо корневых узлов. Кроме того, каждый узел шлет контрольные пакеты периодически, поэтому протокол устойчив в случае потери части этих сообщений, что довольно часто случается с широковещательными пакетами в беспроводных сетях.

Протокол RA-OLSR практически совпадает с оригинальным OLSR [20]. По сравнению с протоколом OLSR, модификация RA-OLSR сводится к управлению энергопотреблением и отсутствием фиксированной процедуры выбора узлов MPR (ее может задавать производитель устройства). Он присутствовал в ранних вариантах стандарта как опциональный, однако в версии D1.07 от него отказались. Поводом для этого послужил тот факт, что RA-OLSR дублирует функциональность HWMP, являющегося и про- и реактивным протоколом, а стандарт допускает применение любых других протоколов маршрутизации [22]. Кроме того, RA-OLSR вызвал множество замечаний, в основном указывающих на неточности его описания в тексте стандарта. Также оказался слишком большим размер самого описания этого протокола (48 из 246 страниц документа IEEE 802.11s были посвящены RA-OLSR) [22]. Таким образом, отказ от RA-OLSR не связан с его недостатками [23, 24]: решено оставить вопрос выбора любого альтернативного механизма маршрутизации производителям оборудования.

Ниже приведен сравнительный анализ протоколов маршрутизации достаточно сложной сети (рис. 5.39) посредством компьютерного моделирования на языке GPSS World (General Purpose Simulation System).

Рассматривалась сеть без шумов, работающая по протоколу IEEE 802.11a с дополнением IEEE 802.11s на скорости 54 Мбит/с. Практически все потоки

данных в этой модели направляются от конечных узлов (1–20) в Интернет через шлюзы (узлы 41–42), в обратном направлении передается примерно 2% от общего числа пакетов.

Рис. 5.39. Топология анализируемой mesh-сети

Для оценки производительности протоколов маршрутизации сравнивались пропускная способность, средние длины путей (в шагах) от конечных узлов до ближайшего шлюза (табл. 5.5) и другие параметры. На рисунке видны кратчайшие пути от узлов до шлюзов: три шага для узлов 1–20, два шага для узлов 21–34 и один шаг для узлов 35–40.

Таблица 5.5. Сравнение протоколов маршрутизации при стандартной метрике времени передачи

Протокол маршрутизации	RA-OLSR	HWMP
Средняя длина пути, шагов	3,3	4,5
Пропускная способность, Мбит/с	15,3	10,6
Отношение посылок пакетов к доставленным пакетам	5,7	6,3
Время доставки, мс	30	36

Видно, что в случае HWMP пути становятся в полтора раза длиннее, а это, в свою очередь, существенно отражается на остальных параметрах сети. Так, пропускная способность сети тоже снижается практически в 1,5 раза по сравнению с RA-OLSR. Число посылок пакетов данных в сети на каждый доставленный пакет возрастает на 10%, а время доставки пакетов — на 20%. Это объясняется тем, что с ростом длины путей увеличивается число коллизий, а с ними — и число повторных посылок.

Однако в приведенном эксперименте, кроме метода маршрутизации, сыграл свою роль и недостаток метода измерения метрики времени передачи, предложенной в стандарте IEEE 802.11s. Для измерения метрики использовалась

групповая отправка пробных пакетов, которая увеличивала загрузку сети, а полученная таким образом метрика случайным образом менялась от измерения к измерению, что влекло выбор неоптимальных маршрутов.

Чтобы выяснить характеристики именно метода маршрутизации, аналогичный эксперимент проводился с использованием простейшей метрики — количества шагов до узла (табл. 5.6). В этом случае алгоритм RA-OLSR выбирает практически идеальные пути, а HWMP, хотя и дает заметно лучшие результаты по сравнению с предыдущим экспериментом, тем не менее остается на 16% хуже по этому параметру. Как и в случае с предложенной в стандарте метрикой, это оказывает большое влияние на пропускную способность сети. Время доставки пакетов и отношение числа посылок пакетов данных к числу доставленных пакетов в случае HWMP становится даже лучше, чем для RA-OLSR. Но объясняется это лишь тем, что алгоритм RA-OLSR обеспечивает почти в 1,5 раза большую загрузку сети (пропускная способность), что существенно отражается на данных параметрах.

Таблица 5.6. Сравнение протоколов маршрутизации, если метрика — число шагов

Протокол маршрутизации	RA-OLSR	HWMP
Средняя длина пути, шагов	3,0	3,5
Пропускная способность, Мбит/с	16,7	11,8
Отношение посылок пакетов к доставленным пакетам	5,2	5,1
Время доставки, мс	30	25

Возникающие проблемы с HWMP объяснить несложно. Ведь данный протокол предельно прост и хранит минимум информации. Так, ему известен только один путь до каждого из узлов mesh-сети. Каждый вновь прибывший от данного отправителя PREQ-пакет, если его DSN больше предыдущего или метрика лучше, считается пришедшим по единственно верному пути. Если же PREQ-пакет, шедший по более короткому пути, был потерян (а для широковещательных пакетов это явление довольно частое), то путь автоматически становится длиннее, чем он есть на самом деле. В случае с RA-OLSR таких проблем не возникает, так как узлы знают всю (или почти всю) топологию сети, да и путь через узел исчезает только при многократном неполучении информации о нем.

Таким образом, хотя в стандарте и остался только один протокол маршрутизации и одна метрика, они требуют серьезной доработки. В случае, если его недостатки не будут исправлены, производители устройств будут вынуждены сами выбирать оптимальные методы маршрутизации и метрики. И очевидным кандидатом на эту роль, как видно из приведенного исследования, выступает протокол маршрутизации RA-OLSR.

5.11.6. Реализация mesh-сетей на базе стандарта IEEE 802.11.s

Несмотря на то что работа над стандартом 802.11s еще не завершена, многие ведущие производители телекоммуникационного оборудования, программного обеспечения уже представили свои версии реализации mesh-сетей и оборудования для них и собственные протоколы маршрутизации.

Одно из наиболее полных и законченных решений предложила компания Cisco Systems. Она представила беспроводную платформу Cisco Aironet 1520 Se-

ries, включающую в себя точку доступа mesh-сети внешнего исполнения Cisco Aironet 1522, на базе которой и строится mesh-сеть, при этом используя закрытый фирменный протокол маршрутизации Adaptive Wireless Path Protocol (AWPP). Логика протокола скрыта, однако по косвенным данным можно предположить, что этот протокол базируются на одной из версий протокола HWMP, работающего в проактивном режиме. Управлением и мониторингом сети занимается специальное устройство — контроллер беспроводной сети Cisco Wireless LAN Controller. Компания рекомендует использовать в mesh-сетях контроллеры серии 4400. Этот контроллер также может служить центром безопасности сети, поскольку включает в себя RADIUS-сервер и поддерживает ряд других служебных сервисов. Контроллер и устройства сети обмениваются между собой служебной информацией по протоколу управления Lightweight Access Point Protocol (LWAPP). Открытая версия этого протокола редактируется и обсуждается на сайте открытого международного сообщества IETF (Internet Engineering Task Force).

Компания Tropos Networks также представила свое решение в области маршрутизации в mesh-сетях. Яркий пример внедрения ее разработок — сеть Google WiFi, объединяющая свыше 400 маршрутизаторов в опорной сети, охватывающая более 30 км^2 и 15 тыс. домов для обслуживания 25 тыс. пользователей. Данного результата удалось достичь благодаря разработке и использованию протокола Predictive Wireless Routing Protocol (PWRP), способного работать в больших сетях без потери пропускной способности. PWRP является закрытым проприетарным протоколом, поэтому точных данных о его работе нет. Однако из официальных документов разработчика следует, что данный протокол — полностью распределенный и в первую очередь ориентирован на обеспечение связи клиент-сервер, которая динамически оптимизируется и с легкостью масштабируется при расширении сети. Про метрику, используемую в протоколе, известно лишь то, что она основана на измерении действительной производительности беспроводной сети.

Довольно много информации о маршрутизации в своих сетях представила корпорация Microsoft. Компания разработала реактивный протокол маршрутизации, основанный на алгоритме динамической маршрутизации источника DSR (Dynamic Source Routing). Он очень похож на протокол Ad Hoc On Demand Distance Vector (т. е. на HWMP), с той разницей, что для маршрутизации от источника до адресата используется маршрутная таблица источника, а не промежуточных узлов.

Компания Microsoft предложила и протокол маршрутизации источника по качеству канала (Link Quality Source Routing, LQSR), который является адаптацией DSR на виртуальный 2,5-уровень модели OSI. Введение промежуточного уровня предпринято компанией, чтобы сделать протокол прозрачным для более высокого уровня, но при этом обеспечить его корректную работу при переходе между проводной и беспроводной сетями. К предложенному протоколу прилагается пять различных метрик: количество шагов; время на получение ответа (Round Trip Time, RTT); время на посылку пробного пакета от источника до адресата и обратно (Packet Pair); ожидаемое время передачи (Expected Transmission Time, ETT) и взвешенное совокупное ожидаемое время передачи (Weighted Cumulative ETTs, WCETT). Помимо основного протокола LQSR, есть версия многоинтерфейсного LQSR (MR-LQSR — Multi-Radio Link Quality Source

Routing), которая, согласно экспериментам, дает существенный прирост производительности сети, узлы которой поддерживают несколько интерфейсов.

Специально для mesh-сетей в Голландском институте беспроводной и мобильной связи (Twente Institute for Wireless and Mobile Communications) разработан протокол Forwarding LAyer for MEshing (FLAME). Он работает на виртуальном втором с половиной уровне модели OSI, аналогично протоколу LQSR. Это наделяет FLAME теми же преимуществами, что и LQSR, т.е. прозрачностью с точки зрения протоколов верхних уровней и независимостью от среды передачи данных. Однако в отличие от LQSR, протокол FLAME не использует никаких метрик (первый пришедший от узла пакет считается пришедшим по кратчайшему пути, которым и используется в дальнейшем), любой полученный пакет является основанием для обновления информации о его источнике. При этом в таблицу маршрутизации заносится интерфейс и соседний узел, через которые пролегает путь к источнику пакета. Для этого в сети под управлением FLAME ко всем передаваемым пакетам добавляется FLAME-заголовок.

Группа OLPC team (известный проект One Laptop per Child — каждому ребенку по лэптопу) предложила упрощенную версию протокола HWMP. Неоспоримые преимущества этого решения — открытость проекта и исходных кодов и его поддержка крупными компаниями.

Примечательно и решение компании Nortel — точка доступа Wireless Access Point 7220. Именно на его основе построена московская беспроводная сеть Golden Wi-Fi, которая в 2007 году была признана крупнейшей городской сетью Wi-Fi в мире. Для мониторинга и управления сетью в данном решении используется специальный графический пользовательский интерфейс ENMS, который базируется на протоколе SNMP.

Компания Firetide анонсировала точки доступа mesh-сети HotPoint серии 4000. Эти устройства осуществляют полностью прозрачный переход между существующей проводной и беспроводной mesh-сетями.

Свое решение для mesh-сетей представила и широко известная фирма Proxim. В серии устройств ORiNOCO Wi-Fi Mesh Series разработан специальный протокол ORiNOCO Mesh Control Protocol (OMCP), позволяющий использовать один и тот же беспроводной интерфейс как для формирования транспортной mesh-сети, так и для организации доступа пользователей к беспроводной сети.

5.12. Анализ информационной безопасности беспроводных сетей стандарта IEEE 802.11

5.12.1. Методы защиты информации в спецификации IEEE 802.11 и их уязвимости

Для беспроводных сетей вопрос безопасности стоит гораздо острее, чем для обычных проводных сетей, так как весь обмен трафиком в сети производится в радиоканале и для его перехвата достаточно недорогого стандартного оборудования. Разработчики стандартов Wi-Fi это понимали и сделали все возможное, чтобы обеспечить уровень безопасности, по крайней мере, не ниже чем в проводных сетях.

Протокол безопасности WEP

Первая спецификация IEEE 802.11 (1997 год) не имела какой-либо защиты, кроме сокрытия идентификатора беспроводной сети SSID (service set identifier), который необходимо знать для подключения к сети. Однако идентификатор SSID передается в открытом виде, и его перехват не является сложной задачей. Следует отметить, что большинство точек доступа использует broadcast SSID как настройку по умолчанию, т. е. передает идентификатор сети в эфир в открытом виде.

В последующей версии IEEE 802.11 1999 был введен протокол безопасности WEP (Wired Equivalent Privacy — безопасность, эквивалентная проводной). Исторически он начал использоваться в устройствах IEEE 802.11b.

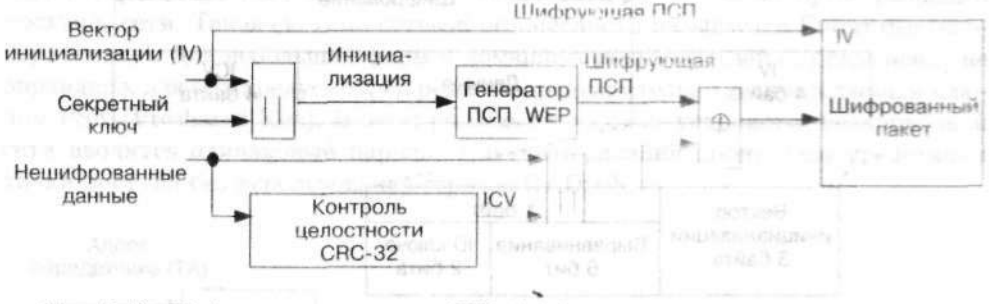


Рис. 5.40. Шифрование в протоколе WEP (RC4)

Алгоритм WEP основан на использовании четырех общих для одной сети секретных ключей (паролей пользователя) длиной 40 бит. Само шифрование происходит по алгоритму RC4 компании RSA Security. Алгоритм использует перемножение блоков исходных данных на псевдослучайную последовательность такой же длины, что и блок шифруемых данных (соответствует кадру MAC-уровня) (рис. 5.40). Генератор псевдослучайной последовательности инициализируется 64-разрядным числом (ключем), состоящим из 24-разрядного вектора инициализации (IV — initialization vector) и 40-разрядного секретного ключа. Существенно, что если секретный ключ известен устройствам сети и неизменен, то вектор IV может изменяться от пакета к пакету. Для защиты от несанкционированного изменения передаваемой информации как для шифрованный пакет защищается 32-разрядной контрольной суммой CRC-32, ее значение передавалось в параметре ICV (integrity check value). Таким образом, при шифровании к передаваемым данным добавляется 8 байт: 4 для ICV, 3 для IV, и еще 1 байт содержит информацию о номере используемого секретного ключа (одного из четырех) (рис. 5.41). Отметим, что ключ может быть не только 64, но и 128 бит. В последнем случае под пароль отводится не 40, а 104 бита.

Алгоритм RC4 является симметричным, т. е. для шифрования и дешифрования служит один и тот же ключ. Это обеспечивает высокую скорость работы, но низкую криптостойкость. Алгоритм WEP обладает чрезвычайно низкой криптостойкостью не только в силу своей симметричности. Ключ длиной в 64 бита (а реально — 40 бит, что остается за вычетом 24 бит вектора инициализации IV) подбирается методом полного перебора за несколько секунд. Для подбора 128-битного ключа потребуется достаточно большое время, но с появлением так называемой FMS-атаки (по первым буквам фамилий изобретателей —

Fluher, Martin, Shamir) необходимость в лобовом переборе отпала. Атака FMS использует слабые места в алгоритме распределения ключей RC4, благодаря чему для взлома достаточно было собрать около 6 млн. пакетов. Для несильно загруженных сетей это достаточно много, и на атаку могло уйти от нескольких часов до нескольких суток, но благодаря стараниям людей из лаборатории Dasb0den Labs число требуемых для взлома пакетов сократилось до 500 тысяч. Вскоре после публикации статьи об уязвимости RC4 начали появляться первые утилиты под Linux и FreeBSD, предназначенные специально для взлома WEP. В новом оборудовании уязвимость (ее суть — неудачный механизм генерации IV) была устранена, а для ранее выпущенного производители создали патчи драйверов.

Рис. 5.41. Пакет после WEP-шифрования

В августе 2004 года хакер KoreK написал новый статистический криптоанализатор, который позволял взламывать 40- и 104-битные ключи, используя 200 и 500 тыс. пакетов соответственно. Созданный им алгоритм был использован в утилите aircrack, обновленная версия которой и сегодня является основным инструментом для взлома ключей WEP.

Протокол WEP также выполняет функцию аутентификации, однако ее нельзя назвать полноценной, так как она основана на задаваемом пользователем пароле и на предположении, что его знают лишь легитимные отправитель и получатель. Эту проблему пытались решить за счет использования списка MAC-адресов допустимых клиентских устройств (фильтрация по MAC). Но MAC-адреса передаются в открытом виде, они могут быть перехвачены и использованы для подмены MAC-адреса устройства нарушителя.

Второй проблемой аутентификации в протоколе WEP является его односторонность, т. е. аутентифицируются только клиентские устройства, которые не могут определить легитимность точки доступа. Это позволяет использовать атаки с фальшивыми точками доступа.

Проблему могло решить существенное удлинение ключа шифрования, что привело бы к неоправданному снижению пропускной способности сети. Поэтому разработчики не стали модифицировать этот алгоритм, в более поздних реализациях просто увеличив ключ до 256 бит и приступив к разработке нового стандарта безопасности WPA.

Стандарт WPA

Стандарт WPA (Wi-Fi Protected Access) представляет собой подмножество спецификаций из принятого в 2004 году стандарта IEEE 802.11i. Весь же набор

спецификаций стандарта IEEE 802.11i Wi-Fi Alliance называется WPA2. Спецификации WPA были призваны дать пользователям альтернативу для WEP и стали переходной ступенью между ним и новым стандартом IEEE 802.11i, разработка которого сильно затянулась. Структуру WPA можно представить в виде формулы WPA = IEEE 802.1X + EAP + TKIP + MIC, т. е. WPA является суммой нескольких элементов, рассмотренных ниже.

Протоколы IEEE 802.1X и EAP (Extensible Authentication Protocol — наращиваемый протокол аутентификации) обеспечивают механизм аутентификации пользователей, которые должны предъявить мандат или свидетельство для доступа в сеть. В больших корпоративных сетях для аутентификации часто используют сервер RADIUS. В иерархии сети он находится выше точки доступа и содержит базу данных со списком пользователей, которым разрешен доступ к сети. Такая система сетевой безопасности называется Enterprise (корпоративная). Для небольших фирм и домашних пользователей ее применение не оправдано, для них предусмотрен режим с предварительно распределяемым ключом PSK (Preshared Key). В этом режиме на каждом устройстве беспроводной сети вводится одинаковый пароль, и аутентификация происходит средствами точки доступа без использования сервера RADIUS.

Рис. 5.42. Шифрование по протоколу TKIP

Протокол TKIP (Temporal Key Integrity Protocol — протокол временной целостности ключа) выполняет функции обеспечения конфиденциальности и целостности данных. Функционально TKIP является расширением WEP (рис. 5.42). Аналогично WEP, он использует алгоритм шифрования RC4, но более эффективный механизм управления ключами. Протокол TKIP генерирует новый секретный ключ для каждого передаваемого пакета данных, и один статический ключ WEP заменяется на, примерно 500 миллиардов возможных ключей, которые могут использоваться для шифрования данного пакета данных. Изменен и сам механизм генерации ключа. Он получается из трех компонентов: базового ключа длиной в 128 бит (TK), номера передаваемого пакета (TSC) и MAC-адреса устройства-передатчика (TA). Также в TKIP используется 48-разрядный вектор

инициализации, чтобы избежать повторного использования IV, на котором основана выше описанная атака FMS.

Алгоритм TKIP использует сквозной счетчик пакетов (TSC) длиной 48 бит. Он постоянно увеличивается, сбрасываясь в 1 только при генерации нового ключа. Младшие 16 бит TSC включаются в новый IV (рис. 5.43). Таким образом формируется механизм противостояния так называемым атакам с воспроизведением.

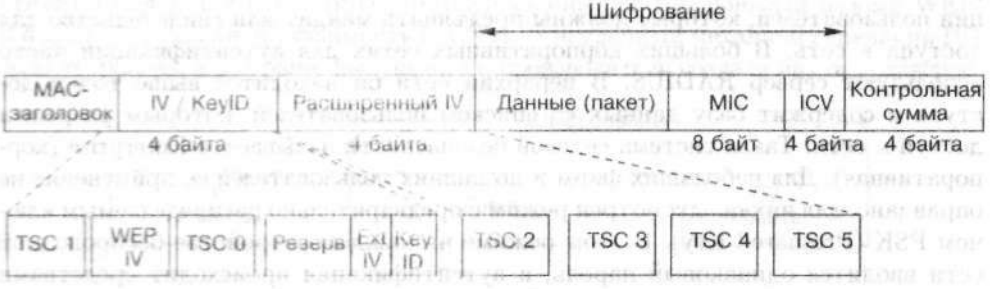


Рис. 5.43. Пакет после шифрования по TKIP

Атаки с воспроизведением используют специальные инструменты, которые внедряют новый трафик для ускорения взлома или даже сканируют порты беспроводных хостов. Такую атаку в рамках WEP остановить невозможно, поскольку стандарт ничего не говорит о том, как должен выбираться IV. В большинстве случаев выбор производится (псевдо?) случайно. Напротив, в TKIP IV увеличиваются последовательно (вместе с TSC), причем те пакеты, где порядок IV (TSC) нарушен, отбрасываются. Это смягчает остроту проблемы воспроизведения трафика, но лишь вступает в некоторое противоречие с одним усовершенствованием, направленным на повышение качества обслуживания, которое было предложено группой по разработке стандарта IEEE 802.11e.

Дело в том, что подтверждать каждый принятый кадр, как описано в алгоритме CSMA/CA, неэффективно. Поэтому было предложено улучшение, названное групповым подтверждением (burst-ACK). Смысл его в том, чтобы подтверждать не каждый отдельный кадр, а группу из 16 кадров. Если один из 16 посланных кадров не дошел до получателя, то применяется селективное подтверждение (аналогично селективному ACK в опциях TCP), требующее повторно передать только потерянный кадр, а не все 16. Разумеется, из-за порядкового счетчика TKIP ошибочно переданный кадр будет отвергнут, если уже получены кадры с большими номерами IV. Чтобы решить эту проблему, в протоколе TKIP применяется окно воспроизведения (replay window), в котором хранятся последние 16 принятых IV, и проверяется, присутствует ли среди них кадр-дубль. Если это так и если этот кадр не был получен ранее, он принимается.

Не менее важен механизм MIC (Message Integrity Check или Michael). Он обеспечивает проверку целостности сообщений вместо очень простого и небезопасного вектора проверки целостности ICV в WEP. Он также строится на основе CRC-32, но длина MIC — 64 бита (два 32-разрядных слова), посредством чего препятствует изменению содержимого передаваемых пакетов. В отличие от ICV,

механизм MIC использует мощную хэш-функцию, которую применяют отправитель и получатель, а затем сравнивают результат. Если он не совпадает, то данные считаются ложными и пакет отбрасывается. Более того, в алгоритм заложены меры противодействия атакам. Если приемник обнаружит две ошибки в MIC в период не более 60 с, соединение будет разорвано и восстановлено не ранее чем через 60 с со сменой всех ключей.

Тем не менее, несмотря на все меры, в ноябре 2009 года была опубликована работа [33], в которой описывался алгоритм атаки на протокол TKIP. Повысить степень криптозащиты призван стандарт IEEE 802.11i (WPA2).

Несмотря на все нововведения, WPA несовместимость с ранее выпущенным оборудованием спецификаций IEEE 802.11a/b/g, и для его использования в большинстве случаев достаточно обновить драйвер и сменить прошивку программы процессора устройства. Для использования же нового стандарта IEEE 802.11i (WPA2 по обозначению комитета Wi-Fi Alliance) необходимо новое оборудование, так как в нем на смену шифрованию RC4 пришел стандарт AES.

5.12.2. Архитектура стандарта IEEE 802.11i

Принятый в июне 2004 года, стандарт IEEE 802.11i возможен сходен с WPA, однако предлагает более высокий уровень безопасности. В IEEE 802.11i определена концепция надежно защищенной сети — Robust Security Network (RSN). Стандарт использует протокол CCMP (Counter Mode CBC MAC Protocol) на основе блокового шифра стандарта AES (Advanced Encryption Standard). Для протокола CCMP алгоритм AES играет ту же роль, что и RC4 для протокола TKIP. Оба протокола работают с одним и тем же механизмом управления ключами. Как и TKIP, CCMP использует 48-битные IV и несколько измененный алгоритм MIC. Благодаря использованию стойкого шифра AES отпала необходимость в генерации пакетных ключей (новый ключ для каждого пакета), и теперь один ключ, создаваемый при каждой ассоциации клиента с сервером, используется для шифрования трафика и для генерации контрольной суммы. В целом, в архитектуре IEEE 802.11i можно выделить два «рубежа»: улучшенные протоколы шифрования и протокол контроля доступа на базе портов (IEEE 802.1X).

IEEE 802.11i предусматривает наличие трех участников процесса аутентификации. Это сервер аутентификации (Authentication Server, AS), точка доступа (Access Point, AP) и рабочая станция (Station, STA). В процессе шифрования данных участвуют только AP и STA (AS не используется).

Стандарт предусматривает двустороннюю аутентификацию (в отличие от WEP, где аутентифицируется только рабочая станция, но не точка доступа). При этом, местами принятия решения о разрешении доступа являются STA и AS, а местами исполнения этого решения — STA и AP.

Для работы по стандарту IEEE 802.11i создается иерархия ключей, включающая Master Key (MK), Pairwise Master Key (PMK), Pairwise Transient Key (PTK), а также групповые ключи (GTK), служащие для защиты широковещательного сетевого трафика:

MK — симметричный ключ, воплощающий решение STA и AS о взаимной аутентификации. Для каждой сессии создается новый ключ MK.

PMK — обновляемый симметричный ключ, владение которым означает разрешение (авторизацию) на доступ к среде передачи данных в течение данной

сессии. PMK создается на основе MK. Для каждой пары STA и AP в каждой сессии создается новый ключ PMK.

PTK — коллекция операционных ключей, которые используются для привязки PMK к данным STA и AP, распространения GTK и для шифрования данных. Выделяется пять фаз работы IEEE 802.11i:

- В *фазе обнаружения* STA находит AP, с которой может установить связь и получает от нее параметры безопасности, используемые в данной сети. Таким образом STA узнает идентификатор сети (SSID) и методы аутентификации. Затем STA выбирает метод аутентификации и между STA и AP устанавливается соединение.
- В *фазе аутентификации IEEE 802.1X* выполняется взаимная аутентификация STA и AS, создаются MK и PMK. В данной фазе STA и AP блокируют весь трафик, кроме трафика IEEE 802.1X.
- В третьей фазе AS перемещает PMK на точку доступа. Теперь STA и AP владеют действительными ключами PMK.
- Четвертая фаза — *управление ключами IEEE 802.1X*. В этой фазе происходит генерация, привязка и верификация ключа PTK.
- Пятая фаза — *шифрование и передача данных*. Для шифрования используется соответствующая часть PTK.

Процесс аутентификации и доставки ключей определяется стандартом IEEE 802.1X. Он предоставляет возможность использовать в беспроводных сетях традиционные серверы аутентификации. Спецификация IEEE 802.11i не определяет тип сервера аутентификации, но де-факто стандартным является использование сервера RADIUS (Remote Authentication Dial-In User Server).

Стандартом IEEE 802.11i предусмотрен режим Pre-Shared Key (PSK), который позволяет обойтись без сервера доступа. При использовании этого режима на STA и на AP вручную вводится Pre-Shared Key, который используется в качестве PMK. Дальше генерация PTK происходит описанным выше порядком. Режим PSK может использоваться в небольших сетях, где нецелесообразно устанавливать AS, а также при работе в режиме ad-hoc.

Подробнее рассмотрим перечисленные нами основные элементы стандарта WAP2.

Протокол IEEE 802.1X

Первоначально стандарт IEEE 802.1X [34] задумывался для обеспечения аутентификации пользователей на уровне 2 в коммутируемых проводных локальных сетях. В беспроводных локальных сетях стандарт IEEE 802.1X имеет дополнительную функцию: динамическое распределение ключей. Для ее поддержки генерируется два набора ключей. Первый набор состоит из сеансовых (или попарных) ключей, уникальных для каждого соединения (ассоциации) между клиентским хостом и точкой доступа. Сеансовые ключи обеспечивают приватность канала и решают проблему «одного ключа WEP для всех». Второй набор состоит из групповых ключей. Групповые ключи разделяются всеми хостами в одной соте сети IEEE 802.11 и применяются для шифрования трафика, вещаемого на группу. Длина сеансовых и попарных ключей составляет 128 бит. Попарные ключи порождаются из главного попарного ключа (Pairwise Master

Key — PMK) длиной 256 бит. PMK выдается RADIUS-сервером каждому устройству сети.

Аналогично, групповые ключи порождаются из главного группового ключа (Groupwise Master Key — GMK). В ходе процедуры порождения PMK и GMK используются в сочетании с четырьмя ключами квитирования EAPOL, которые в совокупности называются попарным временным ключом. Зачастую нецелесообразно применять RADIUS-сервер с базой данных конечных пользователей. В таком случае для генерирования сеансовых ключей используется только предварительно распределенный ключ PMK (вводится вручную).

Поскольку в локальных сетях IEEE 802.11 нет физических портов, то ассоциация между беспроводным клиентским устройством и точкой доступа считается сетевым портом доступа. Беспроводной клиент рассматривается как претендент, а точка доступа — как аутентификатор. Таким образом, в терминологии стандарта IEEE 802.1X точка доступа играет роль коммутатора в проводных сетях Ethernet. Очевидно, что проводной сегмент сети, к которому подключена точка доступа, нуждается в сервере аутентификации. Его функции обычно выполняет RADIUS-сервер, интегрированный с той или иной базой данных пользователей, в качестве которой может выступать стандартный RADIUS, LDAP, NDS или Windows Active Directory. Коммерческие беспроводные шлюзы высокого класса могут реализовывать как функции сервера аутентификации, так и аутентификатора. То же относится и к программным шлюзам на базе Linux, которые могут поддержать стандарт IEEE 802.1X с помощью HostAP и установленного RADIUS-сервера.

В стандарте IEEE 802.1X аутентификация пользователей на уровне 2 выполняется по протоколу EAP (Extensible Authentication Protocol) [35], который был разработан Группой по проблемам проектирования Интернета (IETF). Протокол EAP — это замена метода CHAP, который применяется в протоколе «точка-точка» (PPP) [36], он предназначен для использования в локальных сетях (EAP over LAN). Спецификация EAP over LAN (EAPOL) определяет, как кадры EAP инкапсулируются в кадры сетей стандартов IEEE 802.3, IEEE 802.5 и IEEE 802.10.

Рис. 5.44. Обмен кадрами EAP

Принцип выполнения аутентификации (рис. 5.44): после того как канал установлен, аутентификатор посылает начальный запрос идентификации (Identity Request), за которым следует один или несколько запросов о предоставлении информации для аутентификации. Претендент посылает ответ на каждый запрос. Аутентификатор завершает процесс аутентификации отправкой пакета об успехе или неудаче аутентификации. Отметим, что структура сообщения EAP

аналогична структуре пакета RADIUS. Детальная информация о типах пакетов EAP приведена в RFC 3748 [35].

Существует несколько вариантов протоколов EAP, разработанных с участием различных компаний-производителей:

- EAP-MD5 — это обязательный уровень EAP, который должен присутствовать во всех реализациях стандарта IEEE 802.1X, именно он был разработан первым. Функционально он дублирует протокол CHAP. Мы не рекомендуем пользоваться протоколом EAP-MD5 по трем причинам. EAP-MD5 не поддерживает динамическое распределение ключей. Он уязвим для атаки «человек посередине» с применением фальшивой точки доступа и для атаки на сервер аутентификации, так как аутентифицируются только клиенты. В ходе аутентификации противник может подслушать запрос и зашифрованный ответ, после чего провести атаку с известным открытым или шифр-текстом.
- EAP-TLS (Transport Layer Security, RFC 2716) поддерживает взаимную аутентификацию на базе сертификатов. EAP-TLS основан на протоколе SSLv3 и требует наличия удостоверяющего центра.
- EAP-LEAP (Lightweight EAP или EAP-Cisco Wireless) — это запатентованный компанией Cisco вариант EAP, реализованный в точках доступа и беспроводных клиентских картах Cisco Aironet. LEAP был первой (и на протяжении длительного времени единственной) схемой аутентификации в стандарте IEEE 802.1X, основанной на паролях. Поэтому LEAP приобрел большую популярность и поддерживается в сервере Free-RADIUS, несмотря на то, что это фирменное решение. В основе LEAP лежит прямой обмен запрос-свертка пароля. Сервер аутентификации посылает клиенту запрос, а тот должен вернуть пароль, предварительно выполнив его свертку со строкой запроса. Будучи основанным на применении паролей, EAP-LEAP аутентифицирует пользователя, а не устройство. В то же время очевидна уязвимость этого варианта для атак методом полного перебора и по словарю, не характерная для методов аутентификации с применением сертификатов.

К числу менее расцространенных реализаций EAP относятся PEAP (Protected EAP, неутвержденный стандарт IETF) и EAP-TTLS (Tunneled Transport Layer Security EAP), разработанный компанией Certicom and Funk Software. Эти варианты достаточно развиты и поддержаны производителями, в частности Microsoft и Cisco.

Для работы EAP-TTLS требуется, чтобы был сертифицирован только сервер аутентификации, а у претендента сертификата может и не быть, так что процедура развертывания упрощается. EAP-TTLS поддерживает также ряд устаревших методов аутентификации, в том числе PAP, CHAP, MS-CHAP, MS-CHAPv2 и EAP-MD5. Чтобы обеспечить безопасность при использовании этих методов, EAP-TTLS создает зашифрованный по протоколу TLS туннель, внутри которого эти протоколы и работают. Примером практической реализации EAP-TTLS может служить программное обеспечение для управления доступом в беспроводную сеть Odyssey от компании Funk Software (Windows XP/2000/98/Me).

Такое разнообразие вносит дополнительные проблемы совместимости. В результате выбор подходящего оборудования и программного обеспечения для беспроводной сети становится нетривиальной задачей.

Протокол шифрования CCMP

Рассмотренный нами в рамках WAP протокол TKIP не обязателен для реализации в окончательной версии стандарта IEEE 802.11i, но он обратно совместим со старым WEP и не требует полного обновления беспроводного оборудования. Напротив, протокол CCMP обязателен для совместимости со стандартом IEEE 802.11i. В нем применяется шифр Advanced Security Standard (AES или шифр Rijndael) в режиме счетчика со сцеплением блоков шифр-текста и кодом аутентификации сообщения (CBC-MAC) (рис. 5.45). Режим счетчика (CCM) был создан специально для стандарта IEEE 802.11i, но позже был представлен комитету NIST для универсального применения совместно с шифром AES.

Рис. 5.45. Шифрование по протоколу CCMP

В стандарте IEEE 802.11i определено, что размер ключа AES равен 128 бит. Как и в TKIP, в CCMP используются 48-разрядные IV (здесь они называются номерами пакетов, PN) (рис. 5.46) и несколько видоизмененный алгоритм MIC. В CCMP функции порождения пакетных ключей не реализованы, поскольку сильный шифр AES делает их излишними. В этом протоколе один и тот же

Рис. 5.46. Пакет после шифрования по CCMP

ключ, создаваемый отдельно для каждой ассоциации, применяется как для шифрования данных, так и для генерирования контрольной суммы. Контрольная сумма длиной 8 октетов, применяемая для гарантии целостности сообщения, считается гораздо более эффективной, чем вычисляемая алгоритмом Michael в протоколе TKIP.

Таблица 5.7. Возможности протоколов шифрования, используемых в беспроводных сетях

Протокол	Open System	WEP	WPA-PSK	WPA-EAP	WPA2-PSK	WPA2-EAP
Алгоритм шифрования	RC4	RC4	RC4	RC4	AES (CTR)	AES (CTR)
Аутентификация	Нет	Разделяемый ключ	Разделяемый ключ	IEEE 802.1X	Разделяемый ключ	IEEE 802.1X
Длина ключа, бит	64 или 128	64 или 128	128 (шифров.), 64 (аутент.)	128 (шифров.), 64 (аутент.)	128	128
Повторяемость ключа	24-битный IV	24-битный IV	48-битный TSC	48-битный TSC	48-битный PN	48-битный PN
Целостность данных	CRC-32	CRC-32	Michael	Michael	AES (CBC-MAC)	AES (CBC-MAC)
Целостность заголовка	Нет	Нет	Michael	Michael	AES (CBC-MAC)	AES (CBC-MAC)
Управление ключами	Статические для всей сети			На основе EAP	Статич. для всей сети	На основе EAP

Отметим, что уже созданы микросхемы с аппаратной реализацией AES, что снижает вычислительную нагрузку на аппаратуру сетевых устройств. Это, а также выход на рынок продуктов, поддерживающих протокол CCMP, влечет полный пересмотр архитектуры оборудования для сетей IEEE 802.11. Кроме того, еще остается несколько вопросов, не решенных в стандарте IEEE 802.11i. В частности, речь идет о безопасности независимых сетей, быстрой передаче пользователя от одной точки доступа к другой, а также о процедурах прекращения сеанса и отсоединения.

Модель AAA. Протокол RADIUS

AAA (Authentication, Authorization, Accounting — аутентификация, авторизация и учет) считается краеугольным камнем службы удаленной аутентификации пользователей RADIUS (Remote Authentication Dial-In User Service). Она предназначена для управления доступом к компьютерным ресурсам, проведения определенных политик, анализа использования ресурсов и предоставления информации, необходимой для взимания платы за пользование ими. Все эти процедуры жизненно важны для эффективного и рационального управления сетью.

В основе модели AAA лежат три понятия — аутентификация, авторизация и учет. *Аутентификация* — это способ идентификации пользователя путем за-

проса его «верительных грамот» и проверки их правильности. Предполагается, что у каждого пользователя имеется уникальный набор характеристик для получения доступа. Сервер, удовлетворяющий требованиям AAA, сравнивает эти характеристики с теми, которые хранятся в базе данных. Если «верительные грамоты» совпадают, то пользователю предоставляется доступ к запрошенным сетевым ресурсам, в противном случае в доступе будет отказано. По завершении процедуры аутентификации следует *авторизация*, т. е. принятие решения о том, разрешено ли пользователю выполнять определенные задачи и пользоваться теми или иными сетевыми ресурсами. Обычно авторизация производится одновременно с аутентификацией; если разрешение получено, то пользователь будет иметь доступ к ресурсам. Таким образом, авторизация — необходимая составная часть разумной политики администрирования. Последняя составляющая модели AAA — это *учет*. Под учетом понимается процесс измерения и протоколирования используемых сетевых ресурсов. Сюда включается мониторинг и фиксация событий для различных целей, в том числе биллинга, анализа трендов, учета потребления ресурсов, планирования вычислительных мощностей и текущего сопровождения.

Несмотря на то, что протокол RADIUS [37] был разработан еще до оформления модели AAA, он может служить хорошим примером ее практической реализации. Протокол RADIUS широко применяется во многих сетях. Его можно определить как протокол безопасности, в котором для аутентификации удаленных пользователей используется модель клиент-сервер. Реализуется он в виде серии запросов и ответов, которые клиент передает от сервера доступа к сети (Network Access Server — NAS) до конечного пользователя. Протокол RADIUS был разработан в ответ на настоятельную необходимость иметь некий метод аутентификации, авторизации и учета действий пользователей, которым необходим доступ к разнородным вычислительным ресурсам.

Основные особенности протокола RADIUS:

- Модель клиент-сервер. Сервер NAS выступает в роли клиента RADIUS. Этот клиент отвечает за доставку информации о пользователе RADIUS-серверу и выполнении тех или иных действий в зависимости от полученного ответа. RADIUS-серверы отвечают за прием запросов на установление соединения, аутентификацию пользователей и возврат всех деталей конфигурации клиенту, который будет предоставлять пользователю определенные сервисы. Кроме того, RADIUS-сервер может выступать в виде прокси-клиента для других RADIUS-серверов или иных серверов аутентификации.
- Сетевая безопасность. Во время аутентификации пользователя обмен данными между клиентом и сервером шифруется с помощью общего секретного кода, который никогда не передается по сети в открытом виде. Пароли пользователей клиент передает RADIUS-серверу также в зашифрованном виде, чтобы исключить возможность прослушивания.
- Гибкие механизмы аутентификации. RADIUS-сервер допускает различные методы аутентификации пользователей. Получив имя пользователя и его пароль, он может поддержать процедуры аутентификации по протоколу PAP, CHAP или ОС UNIX, а также поискать информацию в иных хранилищах, например PAM, LDAP, SQL и т. д..
- Расширяемый протокол. Все данные передаются в виде троек переменной длины: атрибут-длина-значение. Можно добавлять новые значения атрибу-

тов, не нарушая корректность работы существующей реализации, за счет чего протокол становится более гибким и динамичным, способным к расширению.

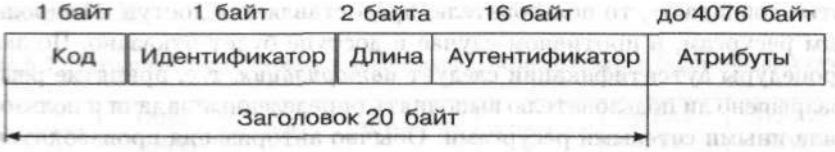


Рис. 5.47. Структура пакета протокола RADIUS

Пакеты протокола RADIUS (рис. 5.47) инкапсулированы в поток данных протокола UDP. Поле кода определяет тип пакета. Получив пакет с некорректным значением кода, сервер игнорирует его без каких-либо уведомлений. Идентификатор позволяет клиенту RADIUS сопоставить полученный от сервера ответ с ранее посланным запросом. В поле «длина» указывается длина пакета сообщения в байтах (от 20 до 4096), включая заголовок. Аутентификатор используется для аутентификации и верификации ответа от RADIUS-сервера, а также как механизм сокрытия паролей. В этом поле могут передаваться значения двух типов: идентификаторы запроса и ответа.

Аутентификатор запроса может встречаться в пакетах типа Access (Доступ) и Accounting Request (Запрос учетной информации), его значение должно быть случайно и уникально. Ответ передается в пакетах типа Access-Accept (Доступ разрешен), Access-Reject (Доступ запрещен) и Access-Challenge (Запрос). Аутентификатор ответа должен содержать значение хеш-функции (по алгоритму MD5 [38]), вычисленное по значениям полей кода, идентификатора, длины, аутентификатора, атрибутов и по общему секретному коду.

В поле атрубутов передаются различные характеристики службы, обычно для анонсирования конкретных предлагаемых или запрашиваемых возможностей.

Уязвимости протокола RADIUS

Известен целый ряд слабостей RADIUS, причиной которых является как сам протокол, так и неудачная реализация клиентов. Перечислим некоторые уязвимости, далеко не исчерпав весь список проблем протокола. Отметим, что сам по себе протокол UDP позволяет подделывать пакеты. Атаки можно отнести к следующим категориям:

- подбор «верительных грамот» пользователя методом полного перебора;
- DoS-атака;
- повтор сеанса;
- внедрение поддельных пакетов.

Различные виды атак многократно описаны в литературе, отметим лишь некоторые наиболее общие типы, основываясь на данных работ и источников [43–48].

Атака на аутентификатор ответа

Аутентификатор ответа (Response Authenticate) — это, по существу, MD5-свертка [38]. Если злоумышленник сумел перехватить корректную последовательность

пакетов типа Access-Request, Access-Accept или Access-Reject, то он может, уже не находясь в сети, попробовать вскрыть общий секретный код методом полного перебора, поскольку остальные параметры, на основе которых вычисляется свертка, известны (код + идентификатор + длина + аутентификатор + атрибуты). Затем возможно повторять эту свертку при каждой попытке угадать общий секретный код.

Атака на общий секретный код на основе атрибута Password

Мандат, содержащий пару имя-пароль, является защищенным, но противник может получить информацию об общем секретном коде, если будет следить за попытками аутентификации. Если взломщик предпримет попытку аутентифицироваться с известным паролем, а затем перехватит отправляемый в результате пакет Accept-Request, то он сможет сравнить защищенную часть атрибута User-Password с паролем, который он сообщил клиенту ранее. Поскольку аутентификатор ответа известен (его можно увидеть в пакете Accept-Request), то противник получает возможность провести атаку методом полного перебора на общий секретный код, уже не находясь в сети.

Атака на пароль пользователя

Эта атака аналогична предыдущей: зная общий секретный код, противник может пробовать различные пароли путем модификации и воспроизведения пакетов типа Access-Request. Если сервер не ограничивает число безуспешных попыток аутентификации одного пользователя, то атакующий сумеет выполнить полный перебор всех паролей, пока не отыщет правильный. Следует отметить, что применение стойкой схемы аутентификации в пакете Access-Request сделает такую атаку почти невозможной.

Атака на аутентификатор запроса

Безопасность пакета в протоколе RADIUS зависит от значения аутентификатора запроса. Оно должно быть уникальным и непредсказуемым. Однако в спецификациях протокола генерированию этого поля не уделено должного внимания, поэтому существует много реализаций, в которых алгоритм оставляет желать лучшего. Если клиент пользуется генератором псевдослучайных чисел с коротким периодом, то желаемый уровень безопасности протокола не будет достигнут.

Атака воспроизведением ответов сервера

Противник может создать базу данных с аутентификаторами запросов, идентификаторами и соответствующими им ответами сервера, если будет периодически прослушивать и перехватывать трафик между клиентом и сервером. Увидев запрос с уже встречавшимся ранее аутентификатором, противник замаскирует себя под сервер и повторит наблюдавшийся ранее ответ. Кроме того, можно воспроизвести похожий на легитимный ответ сервера типа Access-Accept и тем самым аутентифицироваться, не представив корректных «верительных грамот».

Атака на общий секретный код

Стандарт протокола RADIUS допускает использование одного и того же общего секретного кода многими клиентами. Это небезопасно, так как позволяет некорректно реализованным клиентам скомпрометировать сразу много машин. Рекомендуется задавать разные секретные коды для каждого клиента, причем

выбирать слова, отсутствующие в словаре, чтобы их невозможно было предсказать.

5.12.3. Обеспечение конфиденциальности и целостности данных с использованием VPN

Виртуальная частная сеть VPN — это технология сетевого доступа, позволяющая на основе телекоммуникационной инфраструктуры общего пользования, например, Интернетом, сформировать защищенные каналы обмена информацией между отдельными сетями и/или пользователями. Поскольку беспроводные сети IEEE 802.11 легко доступны для случайного или злонамеренного прослушивания, то именно в них развертывание и обслуживание VPN приобретает особую важность, если необходимо обеспечить высокий уровень защиты информации. Стандарт IEEE 802.11i снизил необходимость развертывания VPN, но она полностью не отпадает в сетях, где безопасность информации играет решающую роль. Кроме того, в реализациях стандарта IEEE 802.11i обнаружилось немало проблем с безопасностью. Можно утверждать, что с течением времени будут разработаны новые атаки против этого стандарта. Да и для обеспечения безопасности особо секретных данных нельзя полагаться на какой-то один механизм или на защиту лишь одного уровня сети. В случае двухточечных каналов проще и экономичнее развернуть VPN, покрывающую две сети, чем реализовывать защиту на базе стандарта IEEE 802.11i, включающую RADIUS-сервер и базу данных о пользователях.

В названии VPN слово «виртуальный» подразумевает мирное сосуществование в одном сегменте сети двух различных сетей, не создающих помех друг другу, будь то сети IP, IPX и DDP в одной локальной сети или трафик IP, IPSec и L2TP в Интернете. Слово «частная» означает признание того факта, что весь обмен данными и вообще наличие какой-то сети понятны лишь концевым точкам канала и никому больше. Это равным образом относится к секретности и аутентичности передаваемых данных.

VPN строится на двух механизмах — туннелировании и шифровании. Суть туннелирования — исходные передаваемые пакеты вставляются (инкапсулируются) в пакеты транспортной сети, через которую реализуется передача. При этом форматы исходных пакетов не имеют никакого значения, что допускает передачу через, например, IP-сеть любой не-IP информации. Разумеется, возможен и вариант передачи IP-в-IP. В этом случае каждая дейтаграмма «как есть» вставляется в другую дейтаграмму в начале туннеля и передается по сети. В узле-получателе (конец туннеля), определяемом заголовком внешней дейтаграммы, происходит выделение (декапсуляция) внутренней дейтаграммы и ее обработка в соответствии с ее заголовком. Дополнительно, инкапсулируемый пакет может шифроваться любым способом, причем не только поле данных, но и служебные заголовки. Возможно использование методов аутентификации (т. е. получатель должен удостовериться в легитимности источника пакета перед его обработкой). Для каждого из этих механизмов существуют свои протоколы, широко используемые в современных сетях. Но для отправителя исходных пакетов и их конечного получателя все эти механизмы скрыты, а среда передачи выглядит абсолютно прозрачной, как будто они находятся в одной локальной сети.

Распространено представление о том, что VPN обязательно должна шифровать все проходящие через нее данные, но в общем случае это не так. VPN отвечает трем условиям: конфиденциальность, целостность и доступность. Следует отметить, что никакая VPN не является устойчивой к DoS- или DDoS-атакам (Отказ в обслуживании или Распределенный отказ в обслуживании). Также VPN не защищает от несанкционированного доступа на физическом уровне просто в силу своей виртуальной природы и зависимости от нижележащих протоколов.

Основное достоинство связи через VPN — это сокращение расходов на построение каналов связи между удаленными точками. На данный момент альтернативой VPN служат выделенные линии или внедрение сервера удаленного доступа. Выделенные линии обычно организуются для критически важных приложений, которым требуется гарантированная пропускная способность, тогда как передача данных по сетям общего пользования представляется ненадежной, а их доступность в любой момент времени не может быть гарантирована. Создание беспроводного двухточечного канала — это еще одна недорогая альтернатива, но в свете атак, рассмотренных выше, такое решение нельзя считать достаточно безопасным. Традиционные для сетей IEEE 802.11a/b/g механизмы аутентификации и шифрования сами по себе не в состоянии обеспечить необходимый уровень защиты от опытного взломщика. Но если развертывание протокола IEEE 802.1X и RADIUS-сервера слишком дорого для двухточечных беспроводных мостов, то большинство имеющихся на рынке сетевых устройств могут поддержать VPN, обеспечивающую примерно такой же уровень защиты.

Топология VPN может быть самой разной, но основные — это «сеть-сеть» и «хост-сеть» [39]. Топология «сеть-сеть», очевидно, связывает две удаленные локальные сети (рис. 5.48).

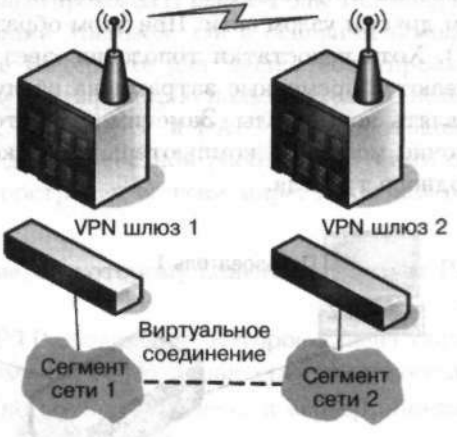


Рис. 5.48. VPN между разнесенными сетями

Топология хост-сеть

Топология «хост-сеть» означает, что сначала клиент устанавливает соединение с сетью общего пользования (с Интернетом), а через нее — с VPN-шлюзом нужной его сети (например, сети своей организации) (рис. 5.49). После успешной аутентификации создается туннель поверх сети общего пользования.

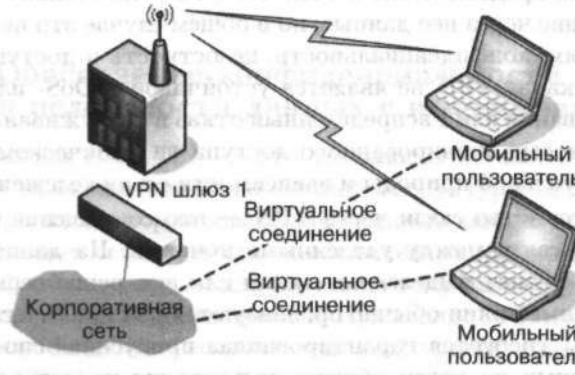


Рис. 5.49. Топология VPN «хост-сеть»

Топология типа «звезда» — самая распространенная для VPN. Центральный VPN-концентратор организует туннели со всеми удаленными клиентами (рис. 5.50). Масштабируемость и общая производительность такой сети ограничена пропускной способностью и вычислительной мощностью VPN-концентратора. Последнее особенно актуально, поскольку сначала нужно расшифровать принятые данные, а затем снова зашифровать перед отправкой. Но в сети «звезда» проще выполнять конфигурирование, обслуживание, контроль доступа и учет. Но при выходе концентратора из строя перестанет работать вся сеть. Звездная топология применима в сетях с каналами «точка-многоточка», но она менее безопасна, чем топология «хост-сеть» поскольку позволяет беспроводным хостам взаимодействовать между собой (через концентратор).

Полносвязная топология предполагает, что каждый узел напрямую соединен туннелем с любым другим узлом сети. При этом образуется «переплетение» соединений (рис. 5.51). Хотя недостатки топологии «звезда»и устраняются, но существенно увеличиваются временные затраты на обслуживание сети и становится трудно добавлять новые узлы. Заметим еще, что оконечные клиенты должны быть достаточно мощными компьютерами, поскольку им приходится поддерживать более одного туннеля.

Рис. 5.50. Топология «звезда»

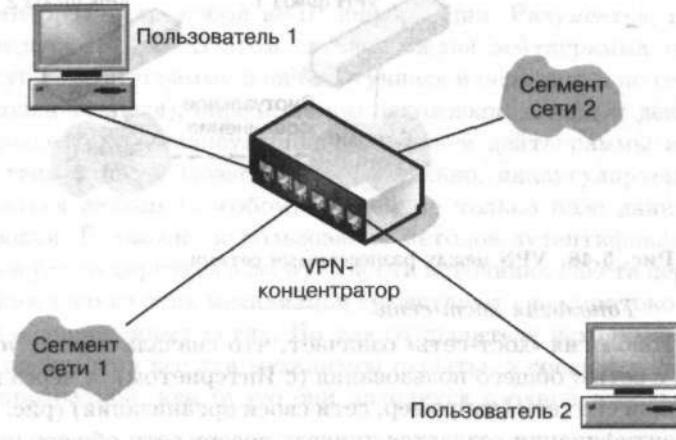
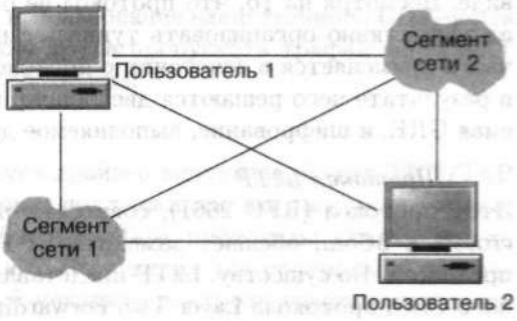


Рис. 5.51. Полносвязная топология



Распространенные туннельные протоколы и VPN

Протокол IPsec

IPSec — наиболее широко признанный и стандартизованный из всех протоколов VPN. IPSec — это набор различных открытых алгоритмов, работающих поверх стека IP. Он предоставляет службы аутентификации и шифрования данных на третьем уровне модели OSI и может быть реализован на любом устройстве, которое работает по протоколу IP. В отличие от многих других схем шифрования, которые защищают конкретный протокол верхнего уровня, IPSec может защитить весь IP-трафик. Он применяется также в сочетании с туннельными протоколами на уровне 2 для шифрования и аутентификации трафика, передаваемого по протоколам, отличным от IP.

Протокол PPTP

Двухточечный туннельный протокол (Point-to-Point Tunneling Protocol — PPTP) — это технология компании Microsoft. Он в основном используется для создания безопасных каналов в Интернете для машин под упроавлением ОС Windows. PPTP обеспечивает аутентификацию пользователей с помощью таких протоколов, как MS-CHAP, CHAP, SPAP и PAP. Этот протокол недостаточно гибок и мало приспособлен для совместной работы с другими протоколами VPN, но прост и широко распространен во всем мире. Протокол определяет следующие типы коммуникаций:

- PPTP-соединение, по которому клиент организует PPP-канал с провайдером;
- управляющее PPTP-соединение, которое клиент организует с VPN-сервером и по которому согласует характеристики туннеля;
- PPTP-туннель, по которому клиент и сервер обмениваются зашифрованными данными.

Этот протокол имеет известные уязвимости и использует относительно слабые шифры MD4 и DES.

Протокол GRE

Протокол Generic Routing Encapsulation (GRE) разработан компанией Cisco и применяется для туннелирования трафика между различными частными сетями, включая не-IP трафик, который нельзя пропустить по сети в неизменном

виде. Несмотря на то, что протокол не осуществляет шифрования, он позволяет эффективно организовать туннель с низкими накладными расходами. GRE часто применяется в сочетании с протоколами шифрования на сетевом уровне, в результате чего решаются две задачи: инкапсуляция не-IP трафика, реализуемая GRE, и шифрование, выполняемое другим протоколом, например, IPSec.

Протокол L2TP

Этот протокол (RFC 2661), совместно разработанный компаниями Cisco, Microsoft и 3Com, обещает заменить PPTP в качестве основного туннельного протокола. По существу, L2TP представляет собой комбинацию PPTP и созданного Cisco протокола Layer Two Forwarding (L2F). Протокол L2TP применяется для туннелирования PPP-трафика поверх IP-сети. Для установления соединения по коммутируемой линии в нем используется PPP с аутентификацией по протоколу PAP или CHAP. Но в отличие от PPTP, L2TP определяет свой собственный туннельный протокол. Поскольку L2TP работает на уровне 2, то через туннель можно пропускать и не-IP трафик. L2TP совместим с любым канальным протоколом, например ATM, Frame Relay или IEEE 802.11. Сам по себе протокол не содержит средств шифрования, но может быть использован в сочетании с другими протоколами или механизмами шифрования на прикладном уровне.

Помимо стандартных протоколов VPN существуют и специализированные варианты. Рассмотрим некоторые из них.

Протокол cIPe

Разработчики утверждают, что cIPe обеспечивает почти такой же уровень безопасности, как IPSec. Протокол работает на уровне IP и позволяет туннелировать протоколы более высоких уровней (например, ICMP, TCP, UDP). Принцип работы напоминает PPP, но cIPe инкапсулирует передаваемые IP-пакеты в UDP-дейтаграммы. При разработке cIPe была поставлена цель создать облегченный протокол, в котором для шифрования данных применяются достаточно стойкие криптографические алгоритмы Blowfish и IDEA, но при этом простой для установки и обслуживания и в то же время несколько более производительный, чем IPSec. В cIPe используется единственный UDP-порт для организации туннеля, трафик без труда проходит через механизм преобразования сетевых адресов NAT (network address translation) и межсетевой экран с запоминанием состояния. Существуют бесплатные реализации cIPe как для UNIX, так и для Windows. К сожалению, были выявлены многочисленные недостатки, допущенные при проектировании cIPe, которые, вероятно, не будут исправлены до выхода следующей версии протокола.

Пакет OpenVPN

OpenVPN — это еще одно открытое решение, по своей функциональности аналогичное cIPe. Пакет легко инсталлируется и конфигурируется. Он работает на большинстве UNIX-подобных систем, в которых есть драйверы виртуальной сети TUN/TAP. OpenVPN имеет модульную структуру. Все криптографические функции реализованы посредством библиотеки OpenSSL, в том числе и самые современные шифры, к примеру AES с 256-битным ключом. Следовательно, протокол в полной мере поддерживает реализованные в OpenSSL механизм PKI для аутентификации сеансов, протокол TLS для обмена ключами, не зависящий от шифра интерфейс EVP для шифрования данных и коды HMAC для аутенти-

фикации данных. Как и в случае cIPe, использование единственного UDP-порта для инкапсуляции туннеля позволяет без труда пропускать трафик через NAT и межсетевые экраны с запоминанием состояния.

Пакет VTun

VTun — это решение, которое использует драйвер виртуальной сети TUN/TAP для туннелирования IP-трафика. Протокол поддерживает все распространенные протоколы уровня 3, в том числе IPX и AppleTalk, протоколы для работы по последовательным линиям связи PPP и SLIP, а также все программы, работающие с конвейерами UNIX. Встроенный механизм контроля трафика позволяет ограничивать входную и выходную скорость работы туннеля, что отличает это решение от всех остальных. С точки зрения конфиденциальности, VTun не претендует на звание самого безопасного протокола, основные усилия при разработке были направлены на быстродействие, стабильность и удобство эксплуатации. Тем не менее, используется алгоритм шифрования Blowfish с 128-битным ключом для шифрования данных и MD5 для генерирования 128-разрядных сверток. Версии для Windows не существует, применение ограничено UNIX-подобными ОС, которые поддерживают драйвер TUN/TAP.

Протокол IPSec

Протокол IPSec — это набор открытых стандартов, разрабатываемых под эгидой IETF. Набор протоколов IPSec состоит из трех основных частей, которые определяют два режима его работы (AH и ESP):

- AH (Authentication Header — заголовок аутентификации) обеспечивает аутентификацию источника данных, целостность и защиту от воспроизведения;
- ESP (Encapsulating Security Payload — инкапсуляция зашифрованных данных, обеспечивает аутентификацию источника данных, целостность, защиту от воспроизведения, конфиденциальность данных и до некоторой степени скрытность управления потоком;
- IKE (Internet Key Exchange — схема обмена ключами через Интернет) предоставляет средства согласования криптографического алгоритма и отвечает за распределение ключей, используемых в AH и ESP.

В режиме аутентификации к стандартной дейтаграмме IP добавляется специальный заголовок аутентификации AH (рис. 5.52) [40]. Поле «следующий заголовок» определяет тип данных после HA. Параметр безопасности SCI определяет тип профиля безопасности для данного пакета. Порядковый номер пакета служит для предотвращения атак повторением. В поле данных аутентификации содержится информация в соответствии с выбранной схемой обеспечения безопасности. Для проверки аутентичности вычисляется свертка в соответствии с механизмом HMAC (с MD5 или SHA-1).

Режим инкапсуляции зашифрованных данных ESP [41] намного сложнее режима HA. Этот механизм обеспечивает не только аутентификацию, но и целостность и конфиденциальность передаваемых пакетов. Помимо собственно шифрования данных, к исходной дейтаграмме добавляются поля «Заголовок ESP» и «Трейлер ESP» (рис. 5.53).

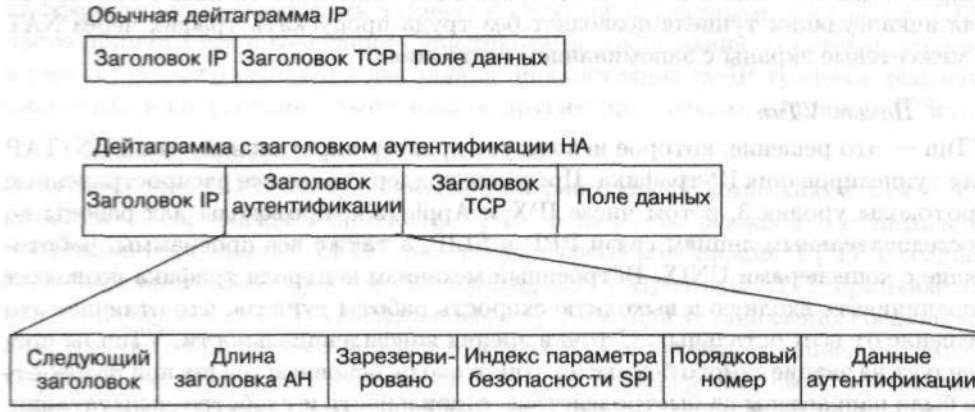


Рис. 5.52. Формат пакета с заголовком аутентификации НА, протокол IPSec

Оба режима — АН и ESP — полагаются на механизм согласования сторонами параметров безопасного соединения (профиля безопасности — security association, SA) по алгоритму IKE. В SA хранятся параметры, о которых договорились оба участника обмена по сети VPN. К ним относятся криптографические ключи и время их жизни, используемые криптографические алгоритмы и режим работы IPSec.

Рис. 5.53. Формат пакета с инкапсулированными шифрованными данными ESP, протокол IPSec

Для каждого режима работы нужно два SA: для входящего и исходящего трафиков. Эти два набора параметров, описывающих данные, отправляемые и получаемые хостом, называются парой SA (SA bundle). В каждом SA явно указывается протокол АН или ESP, IP-адрес получателя для исходящего соединения или IP-адрес отправителя для входящего соединения и 32-разрядный уникальный

идентификатор SPI. Еще одна важная характеристика SA — время жизни. Этот параметр определяет интервал времени, по истечении которого следует провести повторное согласование или считать SA недействительным. Время жизни задается либо как число обработанных байтов, либо как временной интервал; по достижении любого из этих порогов начинается повторное согласование SA. Существует два значения порога времени жизни SA: жесткий и мягкий. В случае если достигнут мягкий порог, то SA согласуется заново, по достижении же жесткого порога SA удаляется из памяти хоста.

Каждый хост-участник хранит SA в базе данных параметров безопасности (SA Database — SAD). Для работы IPSec необходима база данных политик безопасности (Security Policy Database — SPD), в которой хранятся сведения о политиках, применяемых к трафику. SPD содержит набор правил, которые, в свою очередь, состоят из селекторов, несущих информацию о типах выполняемых действий. Когда приходит пакет, по базе данных SPD определяется дальнейшее действие над ним: отбросить, пропустить дальше или передать для обработки протоколу IPSec. В отличие от SPD, база данных SAD хранит только необходимые параметры соединения.

Чтобы определить необходимые действия над пакетом, из его заголовка извлекаются три поля, которые сопоставляются с информацией, хранящейся в SAD (протокол IPSec, IP-адрес и SPI). Если соответствие найдено, то далее параметры сравниваются с полями AH или ESP. Если соответствие не найдено, пакет отбрасывается.

Традиционно в IPSec использовался шифр DES или 3DES. Шифр DES считается слабым и может быть вскрыт за несколько дней или даже часов, поэтому его использование не рекомендуется. Шифр 3DES гораздо более стоек, но требует большого объема вычислений и медленно работает на маломощных устройствах, таких как точки доступа и карманные компьютеры. Возможны и другие шифры, например Rijndael.

Добавление новых заголовков к IP-пакету после инкапсуляции ведет к увеличению размера пакета, т. е. к накладным расходам на организацию туннеля. В случае протокола ESP пакет может вырасти на 300 байт, что негативно сказывается на производительности. В протоколе IPSec предпринята попытка решения этой проблемы посредством встроенного протокола сжатия IP-пакетов (IPComp), в котором обычно применяются алгоритмы DEFLATE или LZS.DEFLATE. Сжатие выполняется до модификации в соответствии с IPSec и до фрагментации. Обычно сжатие случайных или уже сжатых данных неэффективно, более того, иногда применение избыточного сжатия приводит даже к увеличению размера IP-пакета. Применение протокола IPComp должно быть согласовано обеими сторонами с помощью механизма IKE. Следует отметить, что IPComp достаточно гибок, он позволяет выборочно применять сжатие только к конкретному протоколу транспортного уровня или лишь на одном конце соединения.

Протокол обмена и управления ключами в IPSec (IPSec Key Exchange and Management Protocol — ISAKMP) [42] входит в набор протоколов IPSec и определяет процедуры согласования, создания, модификации и удаления SA, а также форматы соответствующих пакетов. Он спроектирован так, чтобы не зависеть ни от какого конкретного метода обмена ключами или генерации ключей,

криптографического алгоритма или механизма аутентификации. ISAKMP описывает лишь общий каркас в довольно абстрактных терминах.

Internet Key Exchange (IKE) — это протокол общего назначения для обмена информацией, относящейся к безопасности. Он предоставляет вспомогательный сервис для аутентификации узлов в протоколе IPSec, согласования параметров безопасности (SA) и ключей алгоритмов шифрования [42].

В IPSec есть внутренний механизм, который позволяет посылать по протоколу IKE сообщение с извещением об удалении, когда одна из сторон уничтожает SA. К сожалению, хост обычно не посылает такое извещение, например, из-за неожиданной остановки по причине сбоя электропитания или в беспроводных сетях — из-за выхода из зоны покрытия. Для решения этой проблемы предусмотрен механизм обнаружения неработающего хоста (Dead Peer Discovery — DPD). Идея в том, что сообщение с извещением посылается раньше данных, если период неактивности продлился дольше заранее установленного порогового значения. Работающий хост на поступившее извещение должен ответить своим собственным.

Отметим, что в беспроводных сетях вероятна ситуация, когда клиент соединяется по беспроводному каналу и получает новый IP-адрес от DHCP-сервера, который изменяется время от времени. Поскольку один из IP-адресов динамический, его нельзя использовать для идентификации данной стороны. Для аутентификации таких хостов необходимо применять другие методы, например сертификаты X.509.

Литература

1. IEEE Std 802.11-2007. IEEE Standard for Information Technology — Telecommunications and information exchange between systems — Local and metropolitan area networks — Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. — Revision of IEEE Std 802.11-1999. IEEE, 2007.
2. IEEE P802.11n/D6.0. Draft Standard for Information Technology — Telecommunications and information exchange between systems — Local and metropolitan area networks — Specific requirements — Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Amendment 5: Enhancements for Higher Throughput / IEEE Standards Activities Department. IEEE, 2008.
3. Шахнович И. В. Современные технологии беспроводной связи. — М.: Техносфера, 2006.
4. IEEE P802.11s/D2.0. Draft Standard for Information Technology — Telecommunications and information exchange between systems — Local and metropolitan area networks — Specific requirements — Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment: Mesh Networking / IEEE Standards Activities Department. IEEE, 2008.
5. Вишневский В. М., Лаконцев Д. В., Сафонов А. А., Шпилев С. А. Mesh-сети. В ожидании стандарта IEEE 802.11s. — Электроника: НТБ, 2008, № 3, с. 98–106.
6. Вишневский В. М., Лаконцев Д. В., Сафонов А. А., Шпилев С. А. Маршрутизация в широкополосных беспроводных mesh-сетях стандарта IEEE 802.11s. — Электроника: НТБ, 2008, № 6, с. 64–69.

7. IEEE P802.11u/D3.0. Draft Amendment to Standard for Information Technology — Telecommunications and information exchange between systems — Local and metropolitan area networks — Specific requirements — Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Amendment 7: Interworking with External Networks / IEEE Standards Activities Department. IEEE, 2008.
8. IEEE P802.11r/D9.0. Draft Standard for Information Technology — Telecommunications and information exchange between systems — Local and metropolitan area networks — Specific requirements — Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment 2: Fast BSS Transition / IEEE Standards Activities Department. IEEE, 2008.
9. IEEE P802.11p/D4.0. Draft Standard for Information Technology — Telecommunications and information exchange between systems — Local and metropolitan area networks — Specific requirements — Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment 8: Wireless Access in Vehicular Environments / IEEE Standards Activities Department. IEEE, 2008.
10. IEEE P802.11v/D3.0. Draft Standard for Information Technology — Telecommunications and information exchange between systems — Local and metropolitan area networks — Specific requirements — Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment 8: Wireless Network Management / IEEE Standards Activities Department. IEEE, 2008.
11. IEEE P802.11w/D6.0. Draft Standard for Information Technology — Telecommunications and information exchange between systems — Local and metropolitan area networks — Specific requirements — Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment 4: Protected Management Frames / IEEE Standards Activities Department. IEEE, 2008.
12. IEEE P802.11z/D2.0. Draft Standard for Information Technology — Telecommunications and information exchange between systems — Local and metropolitan area networks — Specific requirements — Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment 7: Direct Link Setup / IEEE Standards Activities Department. IEEE, 2008.
13. IEEE P802.11k/D13.0. Draft Standard for Information Technology — Telecommunications and information exchange between systems — Local and metropolitan area networks — Specific requirements — Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment 1: Radio Resource Measurement of Wireless LANs / IEEE Standards Activities Department. IEEE, 2008.
14. Заявка на патент «Сверхскоростные беспроводные MESH-сети», 2009, Авторы: Вишневский В. М., Фролов С. В.
15. IEEE P802.11s/D1.08. Amendment: Mesh Networking. — IEEE, January 2008.
16. Вишневский В. М., Ляхов А. И., Портной С. Л., Шахнович И. В. Широкополосные беспроводные сети передачи информации. — М.: Техносфера, 2005.
17. IEEE Std 802.11, 1999 Edition. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. — IEEE, August 1999.
18. IEEE P802.11s/D1.00. Amendment: Mesh Networking. — IEEE, November 2006.
19. Perkins C., Belding-Royer E., Das S. Ad hoc On-Demand Distance Vector (AODV) Routing. — IETF RFC 3561, July 2003.
20. Clausen T., Jacquet P. Optimized Link State Routing Protocol (OLSR). — IETF RFC 3626, October 2003.
21. Qayyum, Laouiti A., Viennot L. Multipoint relaying technique for flooding broad-cast messages in mobile wireless networks. — HICSS: Hawai Int. Conference on System Sciences, January 2002.

22. Reconsidering RA-OLSR — IEEE P802.11-07.2547r2, September 2007.
23. Vishnevsky V. M., Gorodov P. V., Shpilev S. A. Performance analisys of RA-OLSR in IEEE 802.11s mesh networks. — International Workshop. Proc. Of Distributed Computer and Communication Networks (DCCN-2007), 2007, vol. 1.
24. Шпилев С. А. Проактивная маршрутизация в IEEE 802.11s mesh-сетях. — Третья всероссийская молодежная научная конференция по проблемам управления. — ВМКПУ-2008, 2008.
25. Вишневский В. М., Гузаков Н. Н., Лаконцев Д. В. Система «Рапира» — базис для отечественных широкополосных беспроводных сетей. — Электроника: НТБ, 2005, № 1, с. 30–34.
26. Golay M. J. E. Complementary series. — IRE Trans, 1961, IT-7, p. 82–87.
27. Warren J., Sargologos N. PRISM NITRO Introduction. — Intersil. 2003.
28. Arensman R. Intel, 11g fuel WLAN boom. — Electronic Business, 2003, № 5(1).
29. AirForce-BR100-R-11.14.02. — Broadcom, 2002.
30. Joint Proposal: High throughput extension to the IEEE 802.11 Standard: PHY. IEEE 802.11-05/1102r4.
31. Joint Proposal: High throughput extension to the IEEE 802.11 Standard: MAC. IEEE 802.11-05/1095r5.
32. Слюсар В. И. Системы MIMO: принципы построения и обработка сигналов — Электроника, НТБ, 2005, № 8, с. 52–58.
33. Martin Beck, Erik Tews. Practical attacks against WEP and WPA. — http://dl.aircrack-ng.org/breakingwepandwpa.pdf, November 8, 2008.
34. IEEE Std IEEE 802.1X-2004 (Revision of IEEE Std IEEE 802.1X-2001). Port-Based Network Access Control. — IEEE, 13 December 2004.
35. RFC 3748. EAP Key Management Framework. — IETF, August 2008.
36. RFC 1661. The Point-to-Point Protocol (PPP). — IETF, July 1994.
37. RFC 2138. Remote Authentication Dial In User Service (RADIUS). — IETF, April 1997.
38. RFC 1321. The MD5 Message-Digest Algorithm. — IETF, April 1992.
39. Пролетарский А. В., Баскаков И. В., Федотов Р. А. и др. Беспроводные сети Wi-Fi/Курс лекций. — www.intuit.ru/department/network/wifi.
40. RFC 2402. IP Authentication Header. — IETF, November 1998.
41. RFC 2406. IP Encapsulating Security Payload. — IETF, November 1998.
42. RFC 2407. IP Security Domain of Interpretation. — IETF, November 1998.
43. Joshua Wright. Layer 2 Analysis of WLAN Discovery Applications for Intrusion Detection. — www.rootsecure.net/content/downloads/pdf/wlan_ids.pdf, 8.11.2002.
44. Joshua Wright. Detecting Wireless LAN MAC Address Spoofing. — www.rootsecure.net/content/downloads/pdf/wlan_macspooof_detection.pdf, 21.01.2003.
45. www.nwifi.ru.
46. Владимиров А. А., Гавриленко К. В. Проблемы безопасности беспроводных сетей стандартов 802.11a/b/g. — www.arhont.com/digitalAssets/197_moscow_wireless.pdf.
47. Владимиров А. А. Актуальные вектора эффективных атак на сетях стандарта 802.11a/б/г. — http://2006.edu-it.ru/docs/4/04_16_Vladimirov.doc.
48. www.securitylab.ru.

ГЛАВА 6

МОБИЛЬНЫЕ СОТОВЫЕ ТЕХНОЛОГИИ

Сотовая связь — одно из революционных достижений в области беспроводных сетей, ставшее обыденным за последние десять лет. Роль этой технологии в 1990-х годах столь же велика, как бум персональных компьютеров в 1980-х. Мобильный телефон превратился в привычный предмет обихода, по стоимости приближающийся к обычному телефонному аппарату (а по распространенности уже превзошедший число телефонных аппаратов фиксированной связи). Широкие возможности сетей мобильной связи не могут не привлекать внимания разработчиков различного рода систем — мониторинговых, охранных, коммуникационных и т. д. Смена же поколений сотовой связи столь стремительна, что достижимые технические возможности ощутимо опережают реальную потребность пользователей, это ярко демонстрирует проблема с внедрением сотовых сетей третьего поколения (3G). А все уже говорят о 4G.

Рост сотовых сетей связи носит поистине «взрывной» характер. К февралю 2009 года число подключенных мобильных телефонов (активных SIM-карт) в мире первысило 4 млрд. Хотя число абонентов несколько меньше, поскольку у многих больше одного телефона, показатель впечатляет. Только в 2008 году, по данным IDC, было продано 1,18 млрд. мобильных телефонов, доход производителей оборудования мобильной связи первысил 700 млрд. долл. А ведь еще в 2001 году в мире было только 600 млн. абонентов сотовых сетей (около 430 млн. абонентов сетей GSM, 65 млн. приверженцев CDMA, 47,1 млн. пользователей DAMPS, еще около 58 млн. абонентов сетей других стандартов [23]). В России к марту 2009 года насчитывалось 190,8 млн. зарегистрированных абонентов сотовых сетей, показатель проникновения (число SIM-карт на число жителей) составил 131,4%. При этом, по данным аналитиков www.mforum.ru, в стране 112,4 млн. действительно активных SIM-карт (т. е. по которым хоть раз в месяц совершалась какая-либо операция) и 96,4 млн. собственно активных абонентов. Примечательно, что еще на 1 января 2001 года в России было порядка 3,4 млн. абонентов — годовой рост составил тогда 152% [24]! К марту 2004 года их число выросло более чем в 10 раз, достигнув 42,37 млн. абонентов. А за 2005 год базы сотовых операторов России (по числу SIM-карт) выросли на 51,9 млн. абонентов и достигли 125,8 млн. [1]. Для сравнения — абонентов проводной телефонной связи в РФ в 2005 году было чуть больше 41 млн.

Очевидно, что в подобной ситуации потребность в едином стандарте или по крайней мере в совместимых стандартах весьма остра. К тому же сотовая связь все активнее вторгается в сферу передачи данных — это и электронная почта, и доступ к ресурсам Интернета, и обмен видеоинформацией и т. п. Но для подобных услуг были необходимы иные скорости обмена (т. е. превосходящие

стандартные для сетей второго поколения (2G) 9,6 или 14,4 кбит/с) и новые принципы организации соединений.

Еще в 1992 году Международный союз электросвязи (ITU) инициировал работы над стандартом «международной подвижной радиосвязи» IMT-2000. Предполагалось, что к 2000 году появится спецификация сетей подвижной связи, действующая в диапазоне около 2000 МГц и со скоростью передачи данных порядка 2 Мбит/с. Одно из требований — возможность точного определения местоположения мобильных терминалов. Причем речь шла о единой системе телекоммуникаций, объединяющей спутниковые, мобильные, фиксированные виды связи. Надеждам этим сбыться не было суждено, однако некая определенность со стандартами сетей третьего поколения наступила. Но прежде чем говорить о 3G, вспомним историю основных стандартов сотовой связи.

Системы первого поколения (1G-first generation) были развернуты в середине 1980-годов (первые коммерческие сети — в конце 1970-х: 1978 год, Бахрейн, 1979 год — Япония). Системы 1G поддерживали только передачу голоса, обладали аналоговым радиотрактом и охватывали территорию отдельных стран, являясь несовместимыми друг с другом. Цифровые мобильные системы второго поколения 2G появились в конце 90-годов прошлого столетия. Они обеспечивали не только качественную передачу речи, но и низкоскоростную передачу данных (до 14,4 Мбит/с). Обладая совместимостью, мобильные сети 2G охватили все страны мира. В наиболее известной системе 2G — глобальной системе мобильной связи GSM (Global System for mobile communication) — миллиардный абонент был зарегистрирован в 2004 году. На американском континенте и ряде азиатских стран широкое распространение получила другая система второго поколения — cdmaOne (IS-95), базирующаяся на технологии CDMA.

Важным этапом на пути развития мобильных сетей явилась разработка системы пакетной радиосвязи общего пользования GPRS (General Paket Radio Service). В отличие от GSM, где речь и данные передаются по коммутируемым каналам, GPRS обеспечивает пользователям возможность получать и отправлять данные с большей скоростью (до 50 кбит/с), используя технологию коммутации пакетов. GPRS относят к системам 2,5G, подчеркивая промежуточное положение между 2G и 3G.

Системы третьего поколения мобильной связи 3G реализуются на базе новой радиотехнологии, обеспечивающей высокую скорость передачи мультимедийной информации и беспроводный доступ в Интернет, не уступающий сервису провайдеров стационарной сети Интернета. В Европе для систем 3G используют термин UMTS — универсальная мобильная телекоммуникационная система (Universal Mobile Telecommunication System). Внедряется и система третьего поколения мобильной связи cdma2000, представляющая собой дальнейшее развития стандарта IS-95 cdmaOne. Таким образом, системы мобильной связи 3G развиваются по двум направлениям — UMTS и cdma2000. В рамках UMTS обеспечивается приемственность GSM и GPRS; разрабатываются технологии повышения пропускной способности нисходящего (к абоненту) и восходящего (к базовой станции) направлений передачи информации: технологии HSDPA и HSUPA, соответственно.

В 2009 году в рамках UMTS завершена разработка первых версий новейшей технологии Super 3G или Long Term Evolution (LTE), которую уже позициониру-

ют как систему 3,9G. Уже разработаны требования по дальнейшему развитию технологии LTE (LTE Advanced, стандарт 3GPP Release 10). В указанных требованиях предусмотрено, что в нисходящем радиоканале будет обеспечена максимальная скорость передачи информации до 1000 Мбит/с, а в восходящем — до 500 Мбит/с. По сути, это требования к стандарту сетей четвертого поколения 4G.

Второе направление сетей мобильной связи, базирующееся на технологии CDMA и американских стандартах IS-95a и IS-95b, обычно объединяют под единым названием — сети cdmaOne. Их дальнейшим развитием явилось появление стандарта cdma2000, разрабатываемого партнерским объединением 3G PP2 (3G Partnership Project 2). Разработка стандарта cdma2000 ведется поэтапно: cdma2000 1xRTT (или просто 1x) и cdma2000 EV-DO. Стандарт EV-DO (Evolution — Data Optimized), являющийся развитием сетей 1xRTT, был разработан компанией Qualcomm в 1999 году. К настоящему времени выпущено несколько ревизий протоколов EV-DO: Rev.A, RevB и RevC. Стандарт EV-DO RevC (UMB — Ultra Mobile Broadband) относят уже к сетям 4G, так как сети этого класса обеспечивают пиковую скорость до 300 Мбит/с в нисходящем направлении и до 100 Мбит/с в восходящем.

Технологии 3G широко используются операторами сотовой связи во всем мире, включая и Россию. В 2005 году ОАО «Московская Сотовая Связь» под торговой маркой «Скай Линк» развернуло сети по технологии cdma2000 1x EV-DO в частотном диапазоне 450 МГц в Москве и Санкт-Петербурге. К 2009 году указанные сети уже функционировали на территории 31 субъекта Российской Федерации. Сотовый оператор «Мегафон» в 2008 году в Санкт-Петербурге начал коммерческую эксплуатацию сети из 45 базовых станций на основе технологии UMTS/HSPA. МТС предоставляет услуги широкополосного мобильного доступа в Интернет на базе технологии 3G в восьми крупных городах России.

6.1. Аналоговые стандарты сотовой связи

Первые стандарты сотовой связи появились в конце 70-х годов прошлого века и носили в основном региональный характер. Первая коммерческая сеть сотовой телефонной связи заработала в 1978 году в Бахрейне (телефонная компания Batelco) на основе оборудования японской компании Matsushita Electric Industrial. Две зоны с 20 каналами в диапазоне 400 МГц обслуживали 250 абонентов. В том же году в Чикаго компания AT&T начала испытания сотовой системы Advanced Mobile Phone Service (AMPS), работающей в диапазоне 800 МГц. Сеть из 10 зон охватывала связью 54 тыс. км^2. К середине 1980-х годов насчитывалось уже девять стандартов сотовой связи первого поколения [25].

Первые коммерческие сети сотовой связи в Европе заработали в странах Скандинавии под управлением стандарта NMT-450 (Nordic Mobile Telephone System) в 1981 году. Это был совместный продукт Дании, Норвегии, Финляндии и Швеции. Система работала в диапазонах 453–457,5 МГц (восходящий канал) и 463–467,5 МГц (нисходящий канал). Система разделения каналов была частотной, с разносом каналов на 25 (20) кГц и разносом дуплексных каналов на 10 МГц. Передача речи в каналах — чисто аналоговая посредством фазовой модуляции. Служебные сообщения в системе NMT были цифровыми и передавались посред-

ством модуляции FFSK (fast frequency shift keying) — единице соответствовал тон 1200 Гц, нулю — 1800 Гц. Скорость такой передачи составляла 1200 бит/с. Через пять лет появились и варианты систем NMT для 900-МГц диапазона.

Пожалуй, европейская система NMT-450/900 была наиболее распространенной, а в России даже была выбрана в качестве федерального стандарта. Остальные системы, начало эксплуатации которых приходится на 1985–1987 годы, гораздо менее известны. Не каждый специалист сегодня вспомнит такие стандарты, как C-450 (Германия), TACS/ETACS (Великобритания), RTMS-101H (Италия) и Radiocom-200 (Франция). Гораздо более известен первый стандарт сотовой связи, действовавший в США, — AMPS.

Первая сеть стандарта AMPS (Advanced Mobile Phone System) заработала 13 октября 1983 года в Чикаго. Родившись на Американском континенте, этот стандарт распространился по всему миру, попав и в Россию. AMPS рассчитан на диапазон 824–840 и 869–894 МГц, каналы — дуплексные с разносом на 45 МГц, ширина канала 30 кГц. AMPS интересен тем, что на его основе родился первый цифровой стандарт второго поколения — Digital AMPS (DAMPS). Произошло это в 1988 году.

DAMPS — цифровой стандарт второго поколения, действующий в том же диапазоне, что и предшественник. Ширина канала в DAMPS — те же 30 кГц, но применено временное разделение каналов — циклически повторяющиеся кадры с тремя временными интервалами. Речевой кодек — VCELP, 8 кбит/с. Поскольку стандарт американский, его сетевая инфраструктура — ANSI-41. Размер соты — до 20 км. Ряд абонентских устройств поддерживали одновременно DAMPS и AMPS, что было даже удобно, когда удаленность от базовой станции превышала фиксированный в цифровом стандарте радиус соты, поскольку при аналоговой связи это приводило лишь к ухудшению сигнала, но не к разрыву соединения. Однако эра стандартов первого–второго поколений завершалась. Поэтому, в отличие от аналогового прародителя, DAMPS особого распространения в мире получить не успел.

6.2. Глобальная система мобильной связи GSM

Безусловный лидер по распространенности на мировом рынке — стандарт GSM. Его история началась в 1982 году, когда Европейская конференция администраций почты и телеграфа (CEPT) создала рабочую группу GSM (Group Special Mobile) для разработки общеевропейской системы подвижной сотовой связи. В 1989 году работы по GSM перешли под эгиду Европейского института стандартизации электросвязи (ETSI), и в 1990 году были опубликованы спецификации первой фазы стандарта. В 1993 году в 22 странах мира действовало 36 сетей GSM. К 1995 году насчитывалось около 5 млн. абонентов, стандарт стал общемировым и расшифровывался уже как Global System for Mobile Communications. За последующие шесть лет число абонентов возросло в 84 раза, что составляет порядка 70% пользователей сотовой связи во всем мире.

GSM действует в диапазонах 900 и 1800 МГц (в США — 1900 МГц). В Европе и России в диапазоне 900 МГц мобильный телефон передает (восходящий канал) в полосе 890–915 МГц, принимает (нисходящий канал) в интервале 935–960 МГц (для GSM-1800 — 1710–1785 и 1805–1880 МГц, соответственно). Весь

диапазон делится на частотные каналы по 200 кГц — в GSM-900 всего 124 канала (124 восходящих и 124 нисходящих), разнос между восходящим и нисходящим каналами — 45/95 МГц (в диапазонах 900/1800 МГц, соответственно). Базовая станция поддерживает от 1 до 16 частотных каналов. Таким образом, в GSM реализован частотный метод дуплексирования каналов (FDD).

Что касается доступа к среде передачи, в GSM использован принцип временного разделения канала — TDMA. Частотные каналы разбиты на кадры по 8 временных интервалов (канальные интервалы) длительностью по 577 мкс. Каждому физическому каналу соответствует один определенный временной интервал на определенной частоте. Таким образом, мобильный терминал (МТ) передает базовой станции (БС) информацию в течение 577 мкс каждые 4615 мкс. БС связывается с МТ точно так же, но на три временных интервала раньше МТ (и на частоте на 45 МГц выше), чтобы разнести во времени прием и передачу. Это существенно упрощает аппаратуру МТ.

Временные интервалы в GSM бывают пяти типов — нормальный, подстройки частоты, синхронизации, установочный и доступа. Структура нормального временного интервала показана на рис. 6.1. Полезная информация передается двумя блоками по 57 бит. Между ними расположена тренировочная последовательность в 26 бит, ограниченная одноразрядными указателями РВ (Pointer Bit). Интервалы ВВ (Border Bit) длиной 3 бита ограничивают всю передаваемую последовательность. После трансляции всех 148 бит канального интервала передатчик «молчит» в течение защитного интервала ST (Shield Time) длительностью 30,44 мкс, что по времени эквивалентно передаче 8,25 бит.

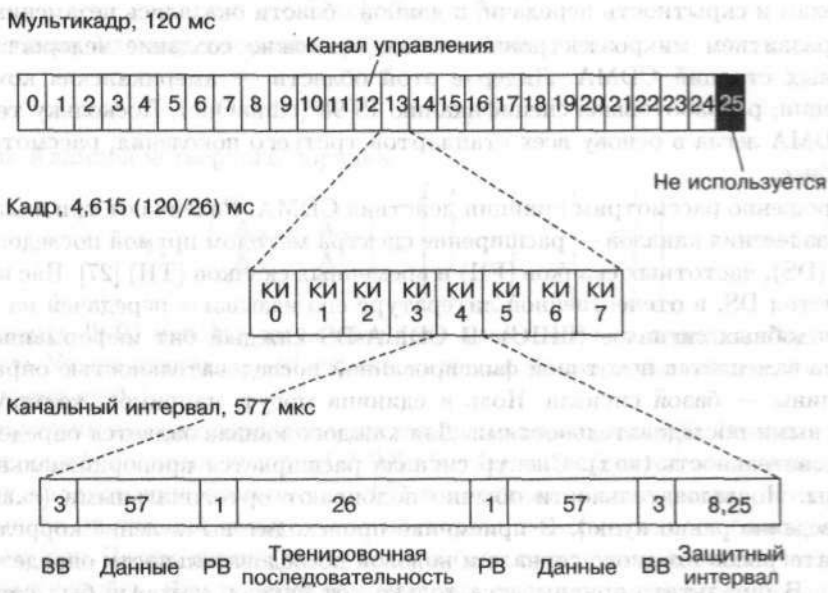


Рис. 6.1. Временное разделение каналов в GSM

Каждые 26 кадров объединены в мультикадр продолжительностью 120 мс. В мультикадре каждый 13-й кадр зарезервирован для канала управления, а в течение каждого 26-го кадра вся система «молчит».

Отметим, что в GSM использован принцип медленных частотных скачков — прием/передача нового кадра может происходить на новой несущей частоте. При этом сохраняется дуплексный разнос в 45 МГц. Начальное значение несущей и последовательность изменения назначаются мобильному терминалу при установлении связи. Модуляция сигнала — двоичная гауссова с минимальным частотным сдвигом GMSK (один бит на символ).

Радиус соты в GSM — до 35 км — ограничен возрастающей временной задержкой распространения сигнала, к которой чувствительна технология TDMA. Сетевая инфраструктура GSM/MAP основана на системе сигнализации ОКС7 (SS7) [26]. Для кодирования речи применен кодек VCELP на основе алгоритма RPE-LTP (Regular Pulse Excitation-Long Term Prediction) со скоростью 13 кбит/с. Скорость передачи данных — до 9,6 кбит/с (по стандартной схеме).

6.3. Стандарт CDMA (cdmaOne)

CDMA расшифровывается как множественный доступ с кодовым разделением каналов (Code-Division Multiple Access). Сама по себе технология не нова: первая в СССР работа на эту тему — «Основы теории линейной селекции» Д.В. Агеева — была опубликована в сборнике ЛЭИС в 1935 году. Значительно продвинули технологию работы К. Шеннона. До определенного момента CDMA находил применение только в военной и специальной технике из-за сложности аппаратуры для обработки сигналов. Зато такие свойства технологии, как высокая стойкость к помехам и скрытность передачи, в данной области оказались незаменимыми.

С развитием микроэлектроники стало возможно создание недорогих портативных станций CDMA. Лидер в этой области — американская компания Qualcomm, разработавшая спецификацию IS-95 (cdmaOne). Поскольку технология CDMA легла в основу всех стандартов третьего поколения, рассмотрим ее подробнее.

Упрощенно рассмотрим принцип действия CDMA. Различают три вида кодового разделения каналов — расширение спектра методом прямой последовательности (DS), частотных скачков (FH) и временных скачков (TH) [27]. Нас интересует метод DS, в отечественной литературе его называют передачей на основе шумоподобных сигналов (ШПС). В CDMA-DS каждый бит информационного сигнала заменяется некоторой фиксированной последовательностью определенной длины — базой сигнала. Ноль и единица могут, например, кодироваться инверсными последовательностями. Для каждого канала задается определенная последовательность (код). Спектр сигнала расширяется пропорционально длине базы. Последовательности обычно подбирают ортогональными (скалярное произведение равно нулю). В приемнике происходит вычисление корреляционных интегралов входного сигнала и кодовой последовательности определенного канала. В результате принимается только тот сигнал, который был расширен посредством заданной кодовой последовательности (корреляционная функция выше порогового значения). Все остальные сигналы воспринимаются как шум. Таким образом, в одной полосе могут работать несколько приемопередатчиков, не мешая друг другу. Благодаря широкополосности сигнала снижается его мощность, причем при очень длинной базе — ниже уровня белого шума.

Сильно возрастает помехоустойчивость, а с ней и качество связи — узкополосная помеха не повлияет на широкополосный сигнал. Кодовая последовательность одновременно является и элементом криптозащиты. Что особенно привлекательно для операторов сотовой связи — упрощается проблема частотного планирования, поскольку все станции работают в одной полосе. Все эти свойства и предопределили успех CDMA.

Естественно, принцип взаимодействия базовой и мобильной станций в стандарте IS-95 гораздо сложнее. Рассмотрим его немного подробнее.

Сети IS-95 занимают практически тот же частотный диапазон, что и сети AMPS: 824–840 и 869–894 МГц. Нисходящий канал (от БС к МТ) всегда на 45 МГц выше восходящего. Ширина канала — 1,25 МГц. Существует и более высокочастотная версия в диапазонах 1890–1930 и 1950–1990 МГц. Там дуплексный разнос — 80 МГц. Ниже мы рассмотрим работу в диапазоне до 900 МГц — в более высокочастотной версии все аналогично, только скорость передачи данных в 1,5 раза выше: до 14,4 кбит/с.

Нисходящий канал содержит 64 логических канала. Логические каналы формируются за счет расширения спектра сигнала последовательностями Уолша (Walsh). Каждая из этих последовательностей представляет собой одну из 64 строк матрицы Адамара (Hadamar). Основное их свойство в том, что все строки матрицы (и их инверсия) взаимно ортогональны.

Способ построения матрицы Адамара прост. Матрица первого порядка $\mathbf{A_1} = [1]$. Матрица $\mathbf{A_{2n}}$ образуется по схеме

$$\mathbf{A_{2n}} = \begin{bmatrix} \mathbf{A_n} & \mathbf{A_n} \\ \mathbf{A_n} & -\mathbf{A_n} \end{bmatrix}.$$

Так, матрица Адамара второго порядка имеет вид

$$\mathbf{A_2} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix};$$

матрица Адамара четвертого порядка

$$\mathbf{A_4} = \begin{bmatrix} \mathbf{A_2} & \mathbf{A_2} \\ \mathbf{A_2} & -\mathbf{A_2} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

В стандарте IS-95 используются матрицы Адамара 64-го порядка. Последовательность Уолша отличается от строки матрицы Адамара только тем, что в ней -1 заменена на 0.

Рассмотрим процесс передачи в нисходящем канале (рис. 6.2). Входной поток (данные, оцифрованный голос) (1,2–9,6 кбит/с) подвергается защитному сверточному кодированию с скоростью 1/2 и попадает в повторитель, который в зависимости от условий связи может повторять передачу одного блока данных до восьми раз. Затем данные поступают в блок перемежения, защищающий от групповых ошибок. Фактически это матрица, которую информационные биты заполняют по строкам, а выводятся по столбцам.

Далее поток перемножается с 42-разрядным числом, так называемой маской длинного кода, фактически — идентификационным номером мобильной станции (речь идет о канале передачи трафика, в каналах другого типа маска может формироваться иначе). Это элемент дополнительной криптозащиты.

Наконец, поток расширяется посредством последовательностей Уолша (каждый бит перемножается на 64-разрядную последовательность). Каждому из 64 каналов соответствует определенная последовательность. Первая последовательность Уолша закреплена за пилотным каналом.

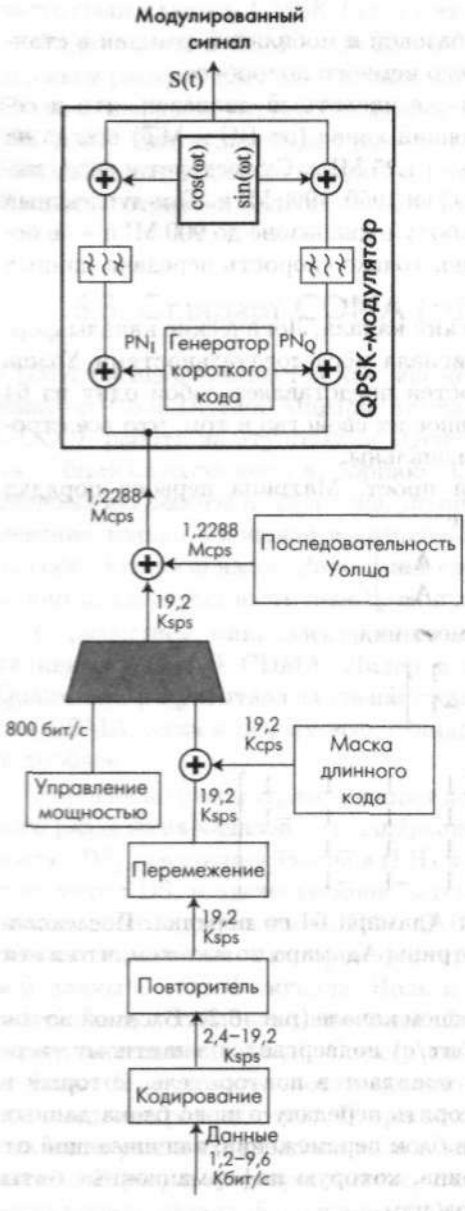


Рис. 6.2. Схема передачи в нисходящем канале cdmaOne

После расширения последовательностями Уолша скорость потока становится 1,2288 Мбит/с (если быть точным — не бит, а чипов, поскольку бит — понятие информационное, а элементы модулированных последовательностей называют чипами). В результате каждому информационному биту исходного потока соответствует 128 чипов выходной последовательности. Выигрыш в отношении сигнал/шум для расширенного и исходного сигнала составляет $10 \lg 128 = 21$ дБ. Если принять, что на входе приемника допустимо соотношение сигнал/шум в 3 дБ, то передачу теоретически можно вести при уровне сигнала на 18 дБ ниже уровня интерференционных помех.

Прежде чем попасть на модулятор, сигнал дополнительно расширяется псевдослучайной последовательностью с так называемым коротким кодом (период — $2^{15} - 1$) и раскладывается на квадратурные составляющие. Несущая модулируется методом четырехпозиционной фазовой манипуляции — QPSK. Поскольку последовательности Уолша взаимно ортогональны, интерференционные помехи между каналами одной БС практически отсутствуют. Передача ведется пакетами длительностью 20 мс.

Восходящий канал делится на $2^{42} - 1$ логических каналов. Каждой мобильной станции присвоен свой уникальный логический канал на основе 42-битного идентификационного номера. Спектр сигнала в обратном канале расширяется на основе так называемых m-последовательностей (длинного кода). Их генерация происходит в 42-разрядном сдвиговом регистре с обратными связями (рис. 6.3).

Если начальные значения во всех m узлах регистра не равны 0, генератор на его основе будет выдавать периодическую псевдослучайную последовательность

длиной $2^m - 1$. Отличительная особенность m-последовательностей — сумма по модулю 2 (операция «исключающее ИЛИ») m-последовательности с той же последовательностью, смещенной по фазе, дает ту же самую последовательность, но с другим фазовым сдвигом $M(t/T) \oplus M(t/T + \varphi_1) = M(t/T + \varphi_2)$. Благодаря этому свойству с помощью идентификационного номера — маски длинного кода — можно задавать начальную фазу последовательности (рис. 6.4). Таким образом, вся система использует один вид псевдослучайной последовательности (ПСП) с очень большим периодом повторения, а селекция логических каналов происходит за счет выбора ее фазы. ПСП обладают всеми свойствами случайных последовательностей, с высокой автокорреляцией при совпадении фаз.

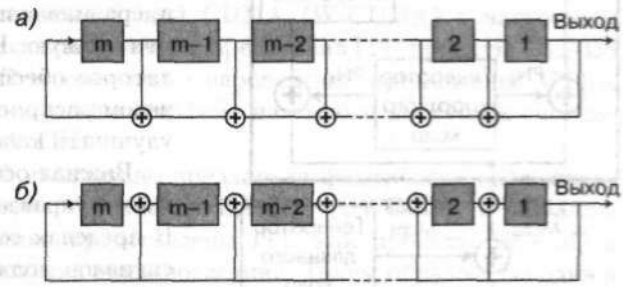


Рис. 6.3. Генерация длинного кода (m-последовательности): схема Фибоначчи (*а*) и схема Галоиса (*б*)

Передача в восходящем канале во многом аналогична передаче в нисходящем (рис. 6.5). Входная информация после сверточного кодирования (со скоростью кодирования 1/3), повторителя и блока перемежения попадает в блок ортогональной модуляции, где каждая группа из 6 бит заменяется соответствующей 64-разрядной последовательностью Уолша. Далее поток последовательностей Уолша перемножается на ПСП (длинный код). При этом каждый элемент последовательности Уолша преобразуется в четыре элемента ПСП. Затем поток квадратизируется посредством так называемого короткого кода с периодом $2^{15} - 1$. Короткий код необходим для первичной синхронизации МТ с БС.

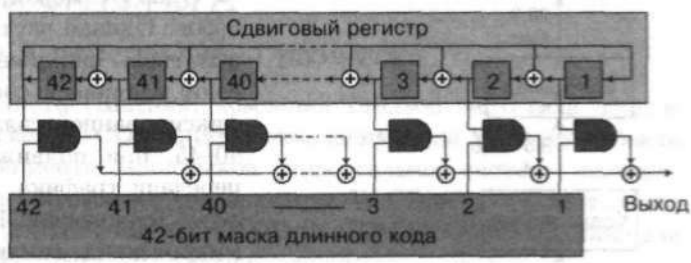


Рис. 6.4. Генератор длинного кода с заданным фазовым сдвигом в cdmaOne

Получающиеся в итоге две последовательности имеют период повторения порядка 2^{57}, что при скорости цифрового потока 1,2288 Мчип/с эквивалентно 3700 годам. Они используются для модуляции несущей. В восходящем канале применяется квадратурная фазовая модуляция со сдвигом O-QPSK (см. рис. 6.5), каждому символу соответствуют два бита. Этот вид модуляции позволяет снизить требования к линейности усилителей передающего тракта МТ. В результате всех преобразований каждый бит исходного сообщения заменяется 256 элементами транслируемой последовательности.

Рис. 6.5. Схема передачи в восходящем канале cdmaOne

Прием сигналов происходит в обратном порядке. Для выделения «своего» сигнала используют цифровые корреляторы, вычисляющие корреляционную функцию с заданной последовательностью Уолша (в МТ) либо с m-последовательностью (БС) в заданной фазе. МТ обычно содержит несколько корреляторов для работы одновременно с несколькими базовыми станциями. Это важно при переходе из соты в соту, когда терминал принимает сигналы от различных БС и, сравнивая их качество, выбирает предпочтительную. Кроме того, несколько корреляторов обеспечивают прием при многолучевом распространении сигнала, что может улучшить качество связи.

Важная особенность стандарта IS-95 — гибкое управление мощностью излучения МТ. В пределах соты уровни принимаемых БС сигналов должны быть одинаковыми независимо от удаления МТ. Для этого мощность МТ регулируется по специальному алгоритму в диапазоне порядка 80 дБ с шагом 1 дБ каждые 1,25 мс. Кроме того, в IS-95 скорость работы голосового кодека не постоянна, как в GSM, а может меняться в зависимости от интенсивности речи от 8 до 1,2 кбит/с. Эти особенности позволяют очень гибко регулировать загрузку в сети, не загружая соту избыточной информацией.

Одна БС может поддерживать до 64 каналов. Однако часть из них — служебные: пилотный, синхронизации, вызова. Оказывают влияние и соседние БС. Однако при фиксированной связи БС поддерживает до 40–45, при подвижной — до 25 каналов передачи трафика. И все это на одной частоте! Технология CDMA требует точной, до микросекунд, синхронизации БС. Для этого используют сигналы глобальной системы позиционирования GPS. Радиус соты — до 20 км, сетевая инфраструктура — ANSI-41.

6.4. Третье поколение сотовой связи

6.4.1. Основные технологии третьего поколения

Основной недостаток систем мобильной связи второго поколения — низкая скорость передачи данных — 9,6–14,4 кбит/с. В рамках же IMT-2000 стояла задача достичь в сетях 3G скорости потока до 2 Мбит/с для малоподвижных абонентов и до 384 кбит/с — для мобильных. В мире сформировались два глобальных партнерских объединения, формирующих стандарты 3G, — 3GPP и 3GPP2 (3G Partnership Project). В первое вошли ETSI (Европа), подкомитет P1 телекоммуникационного комитета ANSI (США), ARIB и TTC (Япония), SWTS (Китай) и TTA (Южная Корея). Участники 3GPP сумели согласовать особенности своих подходов к технологии широкополосной CDMA (W-CDMA) с частотным (FDD) и временным (TDD) дуплексированием, представив ITU проекты IMT-DS и IMT-TC, соответственно. В основу легло европейское предложение UTRA (UMTS Terrestrial Radio Access, радиоинтерфейс наземного доступа к системе UMTS) — UTRA FDD и UTRA TDD.

Члены объединения 3GPP2 изначально предлагали фактически эволюционный путь — варианты развития технологий DAMPS (UWC-136) и cdmaOne (cdma2000). Данные предложения представлены ITU как проекты IMT-SC и IMT-MC. Реальное развитие получил лишь последний. Таким образом, сегодня в мире развиваются две ветви технологий 3G — W-CDMA и cdma2000. Даже название отражает их принципиальную схожесть. Да и их путь развития во многом аналогичен. Однако сети W-CDMA получили более широкое распространение (как наследники GSM), поэтому несколько подробнее остановимся именно на этой технологии.

Отметим, что в качестве одного из стандартов IMT-2000 была утверждена технология микросотовых сетей DECT (проект IMT-FT). В 2007 году полноправным членом IMT-2000 стала технология WiMAX.

6.5. Технология UMTS/HSPA

6.5.1. История и перспективы развития

В 1996 году в городе Чиста (Швеция) компания Ericsson запустила первую опытную сеть с технологией W-CDMA. Эта технология легла в основу проекта наземного мобильного сегмента европейской универсальной системы телекоммуникаций UMTS. Было предложено два варианта W-CDMA — с частотным и временным дуплексом (FDD W-CDMA и TDD W-CDMA) соответственно для парного (2110–2170 и 1920–1980 МГц) и непарного спектров частот.

Технология основывается на расширении спектра методом прямой последовательности в полосе 5 МГц на канал. Изначально определенная скорость потока чипов 4,096 Мчип/с для согласования с другими стандартами была снижена до 3,84 Мчип/с. Таким образом, система может поддерживать требуемые 2 Мбит/с для малоподвижных абонентов и 384 кбит/с — для мобильных. Предусмотрена возможность применения интеллектуальных антенных систем (Smart-антенн с цифровым формированием диаграммы направленности). Принципы технологии FDD W-CDMA во многом аналогичны cdmaOne (конечно, W-CDMA гораздо

сложнее). Одно из принципиальных отличий — сеть на базе FDD W-CDMA может быть асинхронной (возможен и синхронный режим).

W-CDMA (UMTS) изначально разрабатывалась как замена сетей GSM с возможностью плавного перехода. Поэтому ее сетевая инфраструктура совместима с MAP/GSM. Кроме того, она ориентирована на глобальные сети с пакетной коммутацией (IP, X.25). Операторы могут создавать «островки» W-CDMA в особо густонаселенных районах, постепенно расширяя их. Поэтому все абонентские терминалы для W-CDMA в Европе поддерживают GSM. Однако первая коммерческая сеть W-CDMA начала действовать в Японии (оператор — компания NTT DoCoMo) в 2002 году, где телефонов GSM никогда не было. Для японских операторов W-CDMA привлекательна из-за ее высокой абонентской емкости.

Спецификации 3GPP включают в себя описания радиоинтерфейса, базовой сети, а также структуры сервисов. Работа над стандартами 3GPP ведется поэтапно, релизами.

Release 98 (утвержден в 1998 году) и более ранние релизы описывают сети GSM, GPRS и EDGE.

Release 99 (I квартал 2000 года) описывает первые сети универсальной системы мобильной связи UMTS, использующие беспроводный интерфейс с широкополосным множественным доступом W-CDMA. Наземная сеть радиодоступа UMTS (UTRAN) может работать в двух дуплексных режимах: FDD и TDD — в парных и непарных диапазонах, соответственно, что позволяет эффективно использовать спектр с учетом специфики выделения частот в различных регионах. Сигнал подвергается процедуре расширения спектра методом прямой последовательности с чиповой скоростью 3,84 Мчип/с и передается по каналу шириной 5 МГц. Кадр продолжительностью 10 мс делится на 15 таймслотов. Стандарт UMTS Rel'99 обеспечивает скорость передачи данных до 384 кбит/с (пиковая скорость — до 1920 кбит/с). В японском варианте W-CDMA — сети FOMA компании NTT DoCoMo используется чиповая скорость 4,096 Мчип/с и 16 таймслотов на кадр. Именно FOMA, запущенная в эксплуатацию в 2001 году, стала первой сетью стандарта W-CDMA. UMTS совместима с сетями GSM на уровне базовой сети.

Release 4 (II квартал 2001 года). Внесены изменения в стандарт, в частности, представлена полностью-IP базовая сеть.

Release 5 (I квартал 2002 года). Добавлена технология HSDPA (условно относится к поколению 3,5G) — пакетный сервис в прямом (восходящем) канале сети W-CDMA. Скорость передачи данных в прямом канале — до 14,4 Мбит/с, в обратном — 2,0 Мбит/с. Первая сеть HSDPA введена в эксплуатацию в мае 2006 года (SK Telecom, Южная Корея). К 2008 году в мире было свыше 100 коммерческих сетей HSDPA, а в середине 2009 — свыше 300. Но лишь немногие обеспечивали пиковую скорость 14,4 Мбит/с в нисходящем канале: большинство поддерживают 1,8; 3,6 или 7,2 Мбит/с.

Release 6 (IV квартал 2004 года). Добавлена технология HSUPA (условно относится к поколению 3,75G) — сервис в обратном канале сети W-CDMA, увеличивающий пиковую скорость передачи данных до 5,76 Мбит/с. Введение в коммерческую эксплуатацию сетей HSUPA началось в 2007 году. Совместно с HSDPA эти технологии стали называться HSPA.

Release 7 (2007 год). В рамках этого релиза проведена работа над уменьшением задержек, развитием технологий HSPA (HSPA+ или HSPA Evolution) и EDGE

(EDGE Evolution); теоретическая пиковая скорость возросла до 42 Мбит/с в нисходящем и до 11 Мбит/с в восходящем каналах.

Работы по развитию технологии UMTS непрерывно продолжаются. В 2009 году в рамках UMTS завершена разработка первых версий новейшей технологии Super 3G или Long Term Evolution (LTE), которую позиционируют как системы 3,9G. Уже разработаны требования по дальнейшему развитию технологии LTE (LTE Advanced).

Кратко рассмотрим архитектуру и радиоинтерфейс UMTS; более подробное описание можно найти в стандартах 3GPP [6–9] и в работах [10–12].

6.5.2. Архитектура сети UMTS/HSPA

Прежде чем переходить к описанию архитектуры UMTS рассмотрим структурные схемы построения сетей GSM и GPRS. Структурные элементы сети GSM (рис. 6.6): мобильная станция (Mobile Station, MS), подсистема базовых станций (base station subsystem, BSS), подсистема сети и коммутации (Network and Switching Subsystem, NSS).

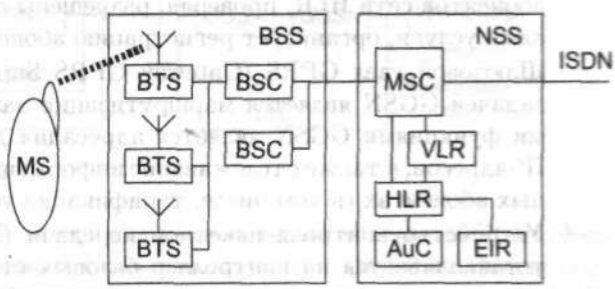


Рис. 6.6. Структурная схема сети GSM

BSS выполняет необходимые функции для контроля радиосоединений с MS, кодирование и декодирование голоса, адаптацию скорости передачи по направлению в беспроводную сеть и из нее. Она включает:

- Базовая станция (Base Transceiver Station, BTS) включает в себя оборудование, необходимое для передачи радиосигнала в пределах географической области, называемой сотой.
- Контроллер базовых станций (Base Station Controller, BSC) резервирует радиочастоты и управляет процессом переключения между BTS, когда мобильная станция перемещается из одной соты в другую.

NSS связывает беспроводную и проводную сети. В нее входят:

- Центр коммутации (Mobile Switching Center, MSC) — коммутатор, который устанавливает соединения с другими MSC и BSC. MCS образуют проводную опорную сеть и могут коммутировать звонки во внешнюю коммутируемую телекоммуникационную сеть (Public Switched Telecommunications Network, PSTN).
- Реестр идентификации оборудования (Equipment Identity Register, EIR) — база данных идентификационных номеров (international mobile equipment identities, IMEI) всех MS сети. Обеспечивает некоторые функции безопасности (например, блокирование звонков).

- Центр авторизации (Authentication Center, AuC), реализует процедуры установления подлинности абонента.
- Реестр собственных абонентов (Home Location Register, HLR).
- Реестр перемещений (Visitor Location Register, VLR) — база данных, в которой временно хранится информация об активных MS на географической территории, к которой относится VLR.

Сети GSM способны передавать как голосовой трафик, так и данные, обеспечивая работу в режимах с коммутацией каналов и с коммутацией пакетов — GPRS.

Система GPRS разрабатывалась на основе сетей GSM для обеспечения более высокой скорости передачи данных. Эта технология является переходной (2,5G) на пути развития от сетей поколения 2G (GSM) к сетям 3G (UMTS).

По сравнению с GSM в сеть GPRS были добавлены следующие структурные элементы (рис. 6.7):

- Узел поддержки GPRS (Serving GPRS Support Node, SGSN) — контролирует доставку пакетных данных, взаимодействует с реестром собственных абонентов сети HLR, проверяя, разрешены ли запрашиваемые пользователями услуги, организует регистрацию абонентов в зоне действия сети.
- Шлюзовой узел GPRS (Gateway GPRS Support Node, GGSN). Основной задачей GGSN является маршрутизация данных через SGSN. Вторичными функциями GGSN является адресация данных, динамическая выдача IP-адресов, а также отслеживание информации о внешних сетях и собственных абонентах (в том числе, тарификация услуг).
- Устройство контроля пакетной передачи (Packet Control Unit, PCU) — устанавливается на контроллер базовых станций BSC и отвечает за направление трафика данных непосредственно от BSC к SGSN.

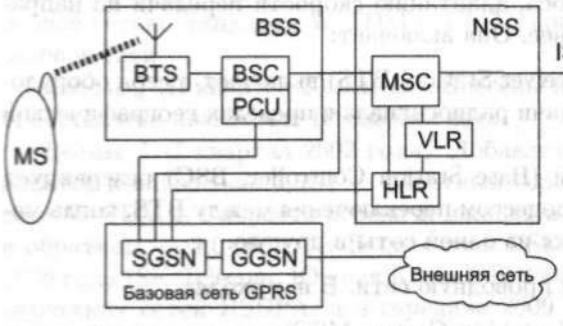


Рис. 6.7. Структурная схема сети GPRS

Для реализации технологии GPRS в существующих сетях GSM помимо описанного выше оборудования требуется обновление программного обеспечения на BTS, BSC, базах данных (HLR, VLR) и новый тип конечного оборудования (обратно совместимое с системой GSM).

Главное отличие UMTS от сетей GSM/GPRS состоит в интерфейсе беспроводной передачи. Для организации нового метода доступа к радиоинтерфейсу требуется новая сеть беспроводного доступа — UTRAN (UMTS Terrestrial Radio Access Network). Для работы с UTRAN в опорную сеть требуется внести лишь

незначительные изменения — а именно контроллеры сети радиодоступа RNC и новые базовые станции Node B. Дуплекс в W-CDMA возможен как по частоте (FDD), так и по времени (TDD).

С функциональной точки зрения, элементы сети (рис. 6.8) объединяются в наземную сеть радиодоступа UTRAN и в базовую сеть (Core Network, CN). Последняя осуществляет переключение и маршрутизацию вызовов, а также подключение к внешним сетям. Кроме того, в состав сети входит оборудование пользователя (UE).

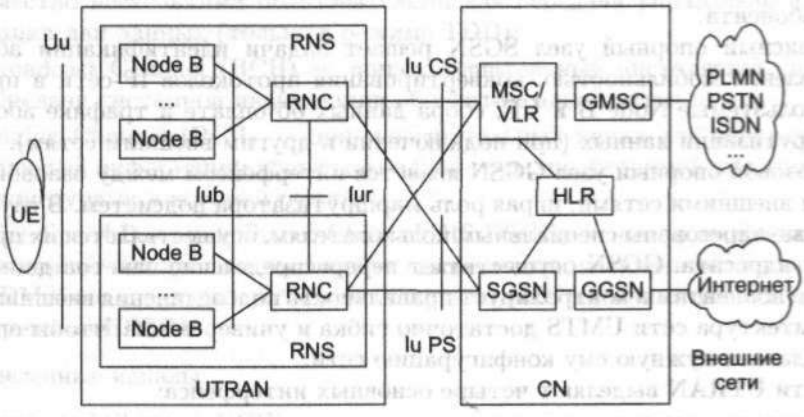


Рис. 6.8. Архитектура UMTS

Сеть UTRAN состоит из совокупности радиосетевых подсистем (Radio Network Subsystem, RNS), соединенных с базовой сетью (см. рис. 6.8). Подсистема RNS включает контроллер RNC и один или несколько базовых станций (Node B). Основной функцией Node B является реализация радиоинтерфейса (обработка радиосигнала, модуляция/демодуляция с расширением/сжатием спектра сигнала, кодирование/декодирование и др.), в том числе, выполнение некоторых операций по распределению радиоресурсов сети (управление мощностью излучения, осуществление хэндовера). Базовая станция может поддерживать режимы FDD, TDD или смешанный (dual-mode) режим работы.

Контроллер сети радиодоступа RNC управляет базовыми станциями и взаимодействует с центром коммутации сети MSC/VLR. Основными функциями RNC являются управление распределением радиоканалов, контроль соединений, регулирование их очередности, удаленная динамическая коммутация, а также контроль за распределением абонентской нагрузки. Контроллеры ведущих мировых производителей телекоммуникационного оборудования строятся, как правило, на базе ATM-коммутатора, расширенного блоками управления радиоканалами.

Мобильный центр коммутации сети MSC/VLR является центральным элементом сети. Он может обслуживать большую группу Node B и обеспечивает все виды соединений, в которых нуждается в процессе работы подвижная абонентская станция. MSC/VLR осуществляет обмен внутри сети UMTS, соединяя между собой различные сетевые элементы, в частности, элементы подсистемы RNS. MSC/VLR обеспечивают соединение с другими MSC, в частности, с

зональными GMSC и другими службами. Совмещенная база данных перемещения абонентов (VLR) содержит копию списка подключенных услуг связи для гостевых абонентов, а также точную информацию о местоположении абонентской станции в рамках обслуживающей системы. Зональный центр коммутации (GMSC) осуществляет коммутацию между сетью UMTS и внешними CS-сетями.

База данных местоположения абонентов HLR представляет собой справочную базу данных о постоянно прописанных в сети абонентах. В ней содержатся опознавательные номера и адреса, а также параметры подлинности абонентов, состав услуг связи, специальная информация о маршрутизации и данные о роуминге абонента.

Сервисный опорный узел SGSN решает задачи идентификации абонента и управления мобильностью, конвертирования протоколов IP-сети в протоколы, используемые Node B и UE, сбора данных об оплате и трафике абонентов и маршрутизации данных (при подключении к другим внешним сетям).

Шлюзовой опорный узел GGSN является интерфейсом между базовой сетью GPRS и внешними сетями, играя роль маршрутизатора подсистем. В случае, если данные адресованы специальным пользователям, осуществляется их проверка и поиск адресата. GGSN осуществляет перераспределение пакетов данных мобильным абонентам и контролирует правильность подсоединения внешних сетей.

Архитектура сети UMTS достаточно гибка и универсальна, чтобы оператор мог создавать нужную ему конфигурацию сети.

В сети UTRAN выделяют четыре основных интерфейса:

- Iu — интерфейс между RNC и базовой сетью;
- Uu — интерфейс между абонентским оборудованием и Node B;
- Iur — интерфейс между контроллерами RNC;
- Iub — интерфейс между Node B и контроллерами RNC.

Протоколы интерфейсов Uu и Iu делятся на две структурные части: протоколы плоскости пользователя (транспортные сервисы, обеспечивают передачу данных через слой доступа) и протоколы плоскости управления (обеспечивают управление транспортными службами радиодоступа и соединением между оборудованием пользователя и сетью, включая запросы на обслуживание, управление ресурсами передачи, хэндовер и т. д.). Также существует прозрачный механизм передачи сообщений уровней, непосредственно не связанных с предоставлением доступа.

6.5.3. Радиоинтерфейс UMTS/HSPA

Радиоинтерфейс UTRAN описан на трех уровнях модели OSI: на физическом (L1), канальном (L2) и сетевом (L3). Передача данных структурирована по так называемым транспортным каналам. Одновременно может транслироваться несколько транспортных каналов.

Транспортные каналы делятся на две группы: общие и выделенные. В выделенных каналах пользователь однозначно определяется физическим каналом, т.е. кодом и частотой для FDD или кодом, временным слотом и частотой для TDD. Общие транспортные каналы:

- Random Access Channel (RACH) — восходящий канал с конкурентным доступом для передачи относительно небольших объемов данных, например для начального доступа;

- Forward Access Channel (FACH) — общий нисходящий канал без управления мощностью внутреннего цикла, для передачи относительно небольших объемов данных. Дополнительно, по FACH передаются широковещательные и мультивещательные данные;
- Downlink Shared Channel (DSCH) — нисходящий канал, использующийся совместно несколькими пользователями для передачи управляющей информации или данных (только в режиме TDD);
- Uplink Shared Channel (USCH) — восходящий канал, использующийся совместно несколькими пользователями для передачи управляющей информации или данных (только в режиме TDD);
- Broadcast Channel (BCH) — широковещательный нисходящий канал для передачи системной информации абонентам одной соты;
- Paging Channel (PCH) — широковещательный канал для передачи управляющей информации абонентам одной соты, необходимой для управления процедурами в ждущем режиме;
- High Speed Downlink Shared Channel (HS-DSCH) — нисходящий канал, разделяемый между пользователями посредством назначения персональных CDMA-кодов из набора кодов, закрепленного за каналом. Используется в режиме HDSPA.

Выделенные каналы:

- Dedicated Channel (DCH) — канал, выделенный одному пользователю в восходящем или нисходящем соединении;
- Enhanced Dedicated Channel (E-DCH) — канал, выделенный одному абоненту для восходящего соединения. Используется в режиме HUSPA.

Для каждого транспортного канала определен *транспортный формат* (Transport Format) или множество транспортных форматов (Transport Format Set) — комбинация параметров кодирования, перемежения, битовой скорости и соответствия физическим каналам.

Для пользовательского оборудования (UE) определено два режима работы: ждущий и активный. После включения питания АС находится в ждущем режиме до тех пор, пока не инициирует соединение. UTRAN не может адресоваться к конкретным пользователям, находящимся в ждущем режиме, только к группам: например, ко всем пользователям в соте или всем пользователям, отслеживающим определенные пейджинговые события. После установления соединения UE переходит в активный режим. Устанавливается соединение между UE и обслуживающим RNC, UE выделяется временный радиосетевой идентификатор (U-RNTI). U-RNTI используется для идентификации UE в сети UTRAN, а также в общих транспортных каналах. По окончании соединения UE возвращается в ждущий режим.

6.5.4. Физический уровень радиоинтерфейса UMTS/HSPA

UTRA может работать в двух дуплексных режимах: FDD и TDD — в парных и непарных диапазонах, соответственно, что позволяет эффективно использовать спектр с учетом специфики выделения частот в различных регионах.

UTRA TDD работает в трех режимах: 7,68 Мчип/с; 3,84 Мчип/с и 1,28 Мчип/с. В режиме TDD в дополнение к DS-CDMA используется также метод множественного доступа с временным разделением (TDMA, Time Division Multiple Access). Такая схема часто обозначается как TDMA/CDMA.

Множественный доступ к среде осуществляется методом кодового разделения с прямым расширением спектра (Direct-Sequence Code Division Multiple Access, DS-CDMA). Ширина полосы составляет 5 МГц для FDD и режима 3,84 Мчип/с TDD (эту технологию называют широкополосным CDMA), 1,6 МГц — для режима 1,28 Мчип/с TDD (что называют узкополосным CDMA), 10 МГц для режима 7,68 Мчип/с. Однако подавляющее большинство сетей W-CDMA используют FDD в парных диапазонах частот с полосой 5 МГц, поэтому в дальнейшем сосредоточимся на этом режиме.

Способ доступа к среде передачи в W-CDMA — комбинированный, представляющий совокупность методов TDMA и CDMA. Одновременно в радиоканале может транслироваться несколько физических каналов, соответствующих транспортным каналам. Каждый канал различается своим канальным кодом и фактором расширения. Кроме того, во временной области передача сигналов W-CDMA ведется радиокадрами длительностью 10 мс. Радиокадр — это область, внутри которой происходит распределение канальных ресурсов (каждому физическому каналу присваиваются коды, тайм-слоты, назначается вид защитного кодирования/модуляции и т. п.).

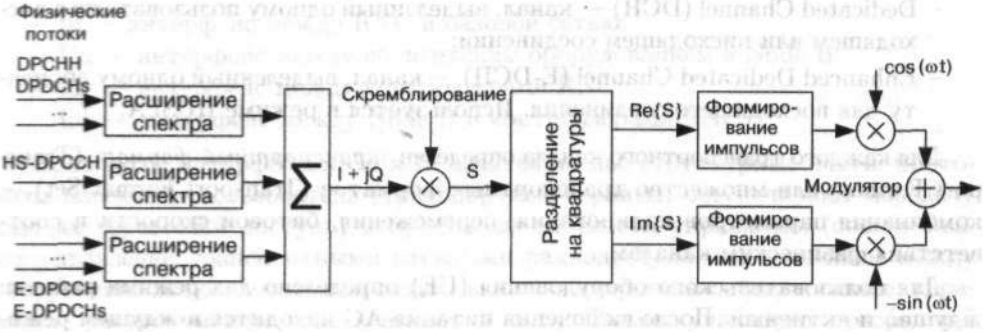


Рис. 6.9. Формирование мультиплицированного радиосигнала в восходящем канале

Общая схема передачи данных в W-CDMA похожа на традиционные CDMA-системы. Потоки данных (в терминологии W-CDMA — физические каналы) подвергаются защитному кодированию и интерливингу. Затем они перемножаются на канальный CDMA-код (формируется последовательность чипов) и скремблируются, умножаются на весовые коэффициенты и мультиплексируются в единый поток (рис. 6.9). После чего происходит модуляция и формирование радиосигнала в квадратурном модуляторе. Разумеется, сама процедура гораздо сложнее, она несколько различается для отдельных физических каналов и для восходящего и нисходящего трактов, но принцип един. В результате формируется мультиплицированный сигнал, объединяющий различные физические каналы (разделенные на основе CDMA и TDMA). Каждый тип каналов передачи данных сопровождается своим каналом управления.

Каждый физический канал характеризуется общей константой — скоростью передачи чипов (элементов после кодового расширения исходной информаци-

онной последовательности) в течение радиокадра, 3,84 Мчип/с. Для каждого физического канала радиокадр делится на 15 канальных интервалов (тайм-слотов). В каждом тайм-слоте 2560 чипов. Информационная скорость в каждом физическом канале зависит от параметра расширения SF. В режиме FDD для восходящих каналов он изменяется от 256 до 2, для нисходящих — от 512 до 4. Физический смысл SF — сколько чипов соответствуют одному модуляционному символу. Поскольку скорость чипов в каждом физическом канале постоянна, скорость следования модуляционных символов в канале зависит только от SF и составляет 3840/SF [кСимволов/с].

Для традиционной W-CDMA (до релиза 4) в восходящих каналах была возможна модуляция BPSK, а позднее — 4PAM (четырехуровневая импульсно-амплитудная модуляция), 1 и 2 бита на символ, соответственно. В нисходящих каналах предусматривалась QPSK, 2 бита на символ. Соответственно, максимальные скорости в физических каналах не могут превосходить 1,92 Мбит/с в нисходящем канале (SF = 4) и 960 кбит/с в восходящем (SF = 4, BPSK). Но очевидно, что такие скорости достижимы лишь теоретически, на практике пропускная способность сетей W-CDMA не превышала 384 кбит/с в нисходящем канале и 128 кбит/с — в восходящем.

Для решения этой проблемы создали технологию HSPDA — высокоскоростная пакетная передача в нисходящем канале. Был введен дополнительный канал HS-PDSCH передачи данных и связанные с ним управляющие каналы. Появилась модуляция 16-QAM (а чуть позднее – и 64-QAM). Канал HS-PDSCH обладает рядом особенностей. Он обеспечивает временное мультиплексирование пакетов данных для различных абонентских станций. Для этого весь радиокадр разбивается на 5 субкадров по 2 мс (три тайм-слота на субкадр). И уже в пределах субкадра происходит распределение канальных ресурсов, но что еще важнее — управление мощностью передатчика и адаптивное изменение схемы кодирования и модуляции. Для этого предусмотрены такие механизмы, как система адаптивной модуляции-кодирования, гибридная система повторного запроса передачи (HARQ) и быстрая диспетчеризация потоков данных. Параметр расширения в HS-PDSCH фиксирован и равен 16, соответственно, скорость модуляционных символов в каждом HS-PDSCH равна 240 кСимволов/с. При модуляции 64-QAM скорость потока данных составляет 1440 кбит/с (960 и 480 кбит/с для модуляций 16-QAM и QPSK соответственно). Под скоростью мы понимаем поток данных после всех защитных кодеков, непосредственно перед процедурой расширения (т. е. перед канальным кодированием/скремблированием). Поскольку система W-CDMA предусматривает применение турбокода с минимальной скоростью кодирования 3/4, реальная скорость полезных данных гораздо ниже. Кроме того, передача канала HS-PDSCH сопровождается передачей канала управления и другой служебной информации.

Относительно низкая скорость символов позволяет передавать несколько HS-PDSCH одновременно в пределах одного радиокадра, используя различные канальные коды CDMA. Возможна одновременная трансляция 5, 10 и 15 таких каналов. При этом скорость передачи может превышать 10 Мбит/с, теоретически достигая значения 14,4 Мбит/с (10 каналов одновременно).

Следующим шагом стало повышение скорости передачи данных в восходящем канале. Эту технологию часто именуют HSUPA (высокоскоростная пакетная

передача данных в восходящем канале). Аналогично нисходящему каналу, был введен новый выделенный высокоскоростной канал E-DCH и связанные с ним управляющие каналы. Сначала в канале стала возможной модуляция QPSK, позднее — 16-QAM. Скорость в одном физическом канале E-DCH при модуляции QPSK и параметре расширения 2 стала достигать 3,84 Мбит/с. При одновременной передаче нескольких E-DCH были достигнуты пиковые скорости 5,76 Мбит/с, а с модуляцией 16-QAM — и 11,5 Мбит/с. Сети, поддерживающие HSUPA, также поддерживают HSDPA. Поэтому обобщенно данные технологии стали называть HSPA, а их развитие — технологиями HSPA+.

Отметим, в мае 2009 года в мире, по данным Ассоциации GSM (hspa.gsmworld.com), насчитывалось свыше 300 действующих сетей HSPA всех типов, развернутых в 107 странах мира. Из них 245 сетей HSDPA, 53 сети HSUPA и 4 сети HSPA+. В нисходящем канале в четырех сетях заявлена скорость 21 Мбит/с, 14,4 Мбит/с — в 11 сетях, 7,2 Мбит/с — в 60 сетях. Но в основном скорости в нисходящем канале составляют 1,8 и 3,6 Мбит/с — 76 и 78 сетей, соответственно.

Для случаев, когда спектральный диапазон ограничен — нет возможности выделять частоты под парные каналы 5 МГц, — проработана версия W-CDMA TDD с временным дуплексированием каналов. При этом весь временной диапазон представляет последовательность равных канальных интервалов. В течение каждого из них в каждом из логических каналов (с кодовым разделением) происходит передача в одном направлении — от БС или от МТ. Таким образом, в определенные промежутки все каналы либо восходящие, либо нисходящие. Соотношение и последовательность восходящих/нисходящих канальных интервалов может гибко изменяться в зависимости от интенсивности трафика в обе стороны. Это крайне важно для многих приложений с асимметричной передачей данных (например, доступ в Интернет). По сравнению с FDD W-CDMA сети с TDD должны быть синхронными, в остальном же их параметры практически совпадают.

6.6. Технология TD-SCDMA

Развитием метода W-CDMA TDD стала система TD-SCDMA, созданная совместно компанией Siemens и китайской Академией телекоммуникационных технологий (China Academy of Telecommunications Technology — CATT). Это стандарт физического уровня беспроводных сетей 3G, одобренный ITU и объединением стандартизирующих организаций 3GPP как часть пула стандартов UMTS. TD-SCDMA (технология CDMA с одной несущей и временным дуплексированием) ориентирована для работы в зонах с высоким дефицитом частотного ресурса — именно такова ситуация в КНР, связанная с высочайшей плотностью населения (в несколько раз выше, чем в густонаселенной Европе).

Сама технология доступа представляет собой комбинацию трех механизмов: временного разделения дуплексных каналов (TDD), временного мультиплексирования каналов (TDMA) и кодового мультиплексирования каналов (CDMA). Обмен происходит циклически повторяющимися кадрами длительностью 5 мс, разделенными на семь временных интервалов (тайм-слотов). Кроме того, в каждом тайм-слоте возможно формирование до 16 CDMA-каналов на основе 16 кодовых последовательностей. Важнейшая особенность — предусмотрена воз-

можность гибкого распределения тайм-слотов исходя из фактически передаваемого трафика. Например, в асимметричных приложениях (доступ в Интернет) для восходящего канала можно выделить один тайм-слот, для нисходящего — остальные шесть.

Ширина одной полосы TD-SCDMA 1,6 МГц. Скорость передачи модуляционных символов 1,28 Мчип/с. Это, вместе с переменным числом тайм-слотов во фрейме, назначенных одному соединению, позволяет добиваться скорости передачи данных в широчайшем диапазоне: от 1,2 кбит/с до 2 Мбит/с. Заявленная дальность передачи — 40 км, допустимая максимальная скорость движения мобильного абонента — не менее 120 км/ч.

Важнейшее достоинство TD-SCDMA — эффективное использование спектра. В технологиях с частотным разносом восходящего/нисходящего каналов на одно соединение всегда выделяются две частотные полосы. И проблема не только в том, что эти две полосы с разносом в 45–220 МГц еще надо найти. При асимметричной передаче (а именно таковы многие мультимедийные приложения) частотный ресурс одного из каналов в большой степени фактически пропадает, поскольку для соединения назначается два частотно-разнесенных канала: приемный и передающий. В случае TD-SCDMA такого не происходит, поскольку частотная полоса одна и соотношение входящего/нисходящего трафика можно гибко варьировать.

Не менее важно, что разработчики TD-SCDMA предусмотрели ее гибкую интеграцию с GSM-сетями, а также мягкий переход к W-CDMA-сетям благодаря поддержке сигнализации и протоколов верхних уровней как GSM, так и W-CDMA. Более того, первые телефоны стандарта TD-SCDMA были двухмодовыми, на основе GSM-чипсета с дополнительной СБИС поддержки TD-SCDMA.

6.7. Технология cdma2000

6.7.1. История и перспективы развития

Семейство протоколов cdma2000 разрабатывается группой 3GPP2 (The Third Generation Partnership Project 2) в рамках проекта IMT-2000. Для Северной Америки cdma2000 представляет собой основную технологию сотовых беспроводных сетей третьего поколения. Она явилась развитием своей предшественницы системы cdmaOne, в которой использовались радиоинтерфес стандарта IS-95 и базовая сеть стандарта ANSI-41 для передачи сигнализации между базовыми станциями и контроллерами базовых станций.

Для доступа к среде используется схема множественного доступа с кодовым разделением (CDMA). Реализация протокола может быть выполнена в широком диапазоне частот; технически она не ограничена диапазоном системы IMT-2000.

Разработка стандарта ведется поэтапно (рис. 6.10):

- cdma2000 1xRTT (или просто 1x, он же IS-2000);
- cdma2000 EV-DO (несколько ревизий: Rev.0, Rev.A, Rev.B, Rev.C).

Рассматривался вариант cdma2000 DC, технически аналогичный W-CDMA, поэтому работы над ним были прекращены.

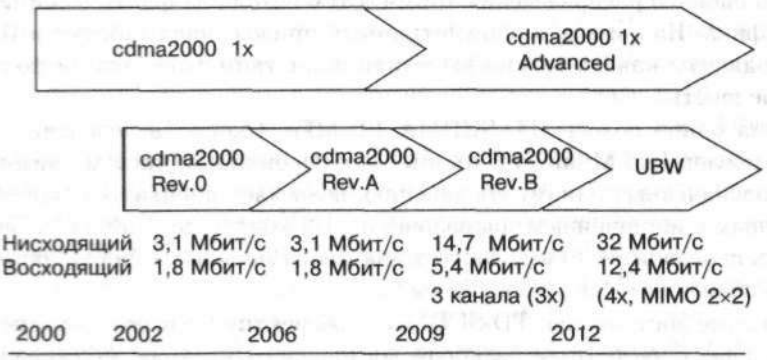


Рис. 6.10. Развитие стандартов cdma2000

В cdma2000 1x используется передача в паре дуплексных радиоканалов шириной по 1,25 МГц (FDD). Пиковая скорость составляет 144 кбит/с (иногда указывается 153 кбит/с), средняя — 50–100 кбит/с. В стандарте стало возможным использовать CDMA-коды различной длины (т. е. различный параметр расширения, от 4 до 256) — решение, аналогичное W-CDMA (или наоборот, в W-CDMA решение аналогично cdma2000). Хотя официально 1xRTT квалифицируется как протокол семейства 3G, из-за невысокой скорости передачи его относят к поколению 2,5G. По сравнению со своим предшественником, стандартом IS-95, 1xRTT практически удвоил емкость сети передачи голосового трафика. Отметим, что развитие линии стандартов cdma2000 1x продолжается, но именно как сервис, ориентированный на голосовые соединения. Цель совершенствования этого направления — повышение емкости каналов передачи, в том числе за счет применения новых вокодеров.

Другая ветвь технологий cdma2000 — cdma2000 EV-DO (сокращение от Evolution — Data Optimized). Они предназначены только для передачи данных, чисто голосовые каналы не предусмотрены (что не исключает методов VoIP). Было намерение создать совмещенную технологию EV-DV, в которой в одном радиоканале совмещались бы возможности передачи голоса (cdma2000 1x) и данных. Однако с марта 2005 года компания Qualcomm приостановила развитие этого направления.

Первый вариант стандарта EV-DO (Rev.0) был представлен компанией Qualcomm в 1999 году. Изменения в основном затронули нисходящий (прямой) канал. В основу легла технология увеличения скорости HDR (High Data Rate). Основная идея новой технологии — гибкое увеличение числа положений вектора сигнала (символов) при фазовой модуляции, следовательно, увеличение числа битов на символ. В стандартной модуляции с CDMA (QPSK) каждый символ определяет два бита. Если увеличить число возможных символов до восьми (8PSK) и 16 (16-QAM), скорость передачи возрастет, соответственно, в полтора и два раза. Очевидно, что чем хуже условия связи, тем меньше символов может распознать приемник (т. е. больший фазовый сдвиг способен распознать его детектор). Поэтому в зависимости от зашумленности эфира вид модуляции меняется.

Для управления параметрами кодирования и модуляции вводится понятие временного кадра длительностью 26,6 мс, разбитого на 16 канальных интервалов

по 1,67 мс. Пакеты данных в нисходящем канале передаются, используя всю ширину канала — т. е. одному пользователю предоставляется весь канал, реализуя тем самым принцип TDMA. При этом для передачи используется 16 CDMA-каналов. Изменение адресата и скорости передачи возможно только в следующем канальном интервале.

Технология cdma2000 EV-DO Rev.0 обеспечивает пиковую скорость передачи данных в прямом соединении 2,4 Мбит/с в полосе 1,25 МГц (среднюю скорость 400–700 кбит/с), в обратном соединении 144 кбит/с (50–100 кбит/с). В целом, технология аналогична HSDPA. Используется тот же механизм сочетания CDMA и TDMA, только несколько различаются чиповые скорости (в cdma2000 они кратны 1,2288 Мчип/с) для обеспечения обратной совместимости с IS-95. Кроме того, в отличие от HSDPA, одновременно может передаваться только один пакет. Коммерческая эксплуатация сетей этого стандарта началась в январе 2002 года (SK Telecom, Корея).

В целом, по сравнению с cdma2000 1x были сделаны значительные изменения, включая следующие:

- в схему мультиплексирования в прямом направлении добавлен метод TDM, чтобы максимизировать эффективность передачи одному пользователю;
- в прямом соединении изменена схема управления мощностью;
- введено адаптивное кодирование (турбокоды) / модуляция;
- добавлена гибридная схема повторной передачи (HARQ);
- схема быстрого планирования в прямом соединении;
- схема передачи обслуживания (handoff) заменена с мягкой на более спектрально эффективную «виртуальную» мягкую.

Следующим шагом стала система cdma2000 EV-DO Rev.A (апрель 2004 года). Развертывание этих сетей началось в 2006 году (KDDI, Япония и Sprint, США). Функционально cdma2000 EV-DO Rev.A аналогична технологии HSUPA (W-CDMA релиз 6, март 2006 года). По сравнению с Rev.0, в новой технологии в 4–10 раз возросла скорость передачи данных в обратном соединении — до 1,8 Мбит/с пиковая, средняя скорость 300 — 400 кбит/с. Увеличилась скорость и в прямом канале, до 3,1 Мбит/с (в среднем 450 — 800 кбит/с). В нисходящем канале изменения по сравнению с EV-DO Rev.0 включают в себя меньшие размеры пакетов и мультиплексирование пакетов от разных пользователей на MAC-уровне. В обратном канале — поддержка HARQ, AMC, меньший размер кадра (6,67 мс). Благодаря этим улучшениям в системе EV-DO Rev.A могут поддерживаться также голосовые сервисы (например, VoIP и игры).

В начале 2006 года был опубликован стандарт cdma2000 EV-DO Rev.B. В нем поддерживается объединение нескольких 1,25-МГц частотных каналов (до 15), а пиковая скорость в нисходящем соединении составляет 4,9 Мбит/с при ширине канала 1,25 МГц, 1,8 Мбит/с — в восходящем соединении. Помимо объединения каналов, стандарт предусматривает модуляцию 64-QAM в нисходящем канале. Так, во время тестирования при объединении 15 каналов была достигнута скорость передачи 73 Мбит/с в нисходящем канале и 27 Мбит/с в восходящем. Однако в реальных сетях выделение полосы частот 20 МГц одному пользователю будет, скорее всего, невозможно. Но 5-МГц полосы (аналогично W-CDMA) вполне доступны, это достигается объединением трех частотных полос (cdma2000

EV-DO Rev.B (3x)). Коммерческая эксплуатация сетей cdma2000 EV-DO Rev.B еще не началась.

Стандарт EV-DO Rev.C (UMB, Ultra Mobile Broadband) относят уже к сетям поколения 4G. Она обеспечит пиковую скорость свыше 275 Мбит/с в нисходящем канале, 75 Мбит/с в восходящем. Технология будет использовать метод множественного доступа OFDMA, каналы шириной 1,25–20 МГц, FDD, переключение без прерывания соединения с сетями стандартов семейства cdma2000 и технологии MIMO. Коммерческая доступность систем этого стандарта планируется не ранее конца 2010 года. Системы UMB функционально и технически подобны будущим системам LTE, разрабатываемым в рамках 3GPP.

Отметим, что несмотря на то, что, в основном, коммерческая эксплуатация систем cdma2000 осуществляется в Северной Америке, эта система также популярна и во многих других странах (особенно в азиатском регионе). В Российской Федерации, начиная с 2006 года, оператор «Скай Линк» ввел в эксплуатацию систему cdma2000 в частотном диапазоне 450 МГц. Общая лицензионная зона компании «Скай Линк» охватывает 71 субъект Федерации. Уже происходит переход на коммерческую эксплуатацию системы EV-DO Rev.A, что позволит существенно повысить скорость передачи данных.

6.7.2. Архитектура сети cdma2000

Семейство стандартов cdma2000 включает в себя описание базового беспроводного интерфейса, спецификацию минимальной производительности, а также описание сервисов. Система cdma2000 совместима с более ранними системами CDMA. Это означает, что мобильная станция cdma2000 может работать как в сетях cdma2000, так и в сетях IS-95. Для обеспечения совместимости стандарт определяет параметры системы, процедуры установления и обработки соединения, а также систему сигнализации. Описаны уровни мощности излучения мобильной станции — для управления взаимным влиянием абонентского оборудования.

В архитектуре сети cdma2000 (рис. 6.11), в отличие от систем IS-95A, добавлены или подверглись модификации элементы:

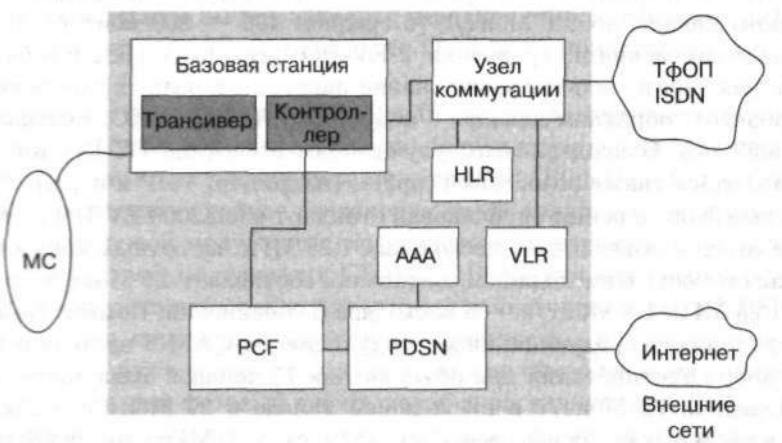


Рис. 6.11. Упрощенная схема сети cdma2000

- PCF (Packet Control Function) — центр управления пакетной передачей;
- PDSN (Packet Data Service Node) — узел, предоставляющий сервисы передачи пакетных данных;
- AAA (Accounting, Authentication, Authorization) — центр авторизации, аутентификации и учета.

Схема сети включает следующие структурные составляющие:

- мобильная станция;
- базовая станция, состоит из приемопередатчика (Base Transceiver System, BTS) и контроллера базовой станции (Base Station Controller, BSC);
- узел коммутации;
- HLR (Home Location Register) — домашний реестр местоположения абонентов;
- VLR (Visitor Location Register) — реестр местоположения гостевых абонентов.

Сравнивая схему сети cdma2000 с рис. 6.8 можно видеть, что базовая сеть пакетной передачи данных использует узел обслуживания пакетных данных PDSN вместо служебного узла поддержки GPRS-SGSN в базовой сети UMTS. В отличие от системы UMTS, где управление мобильностью осуществляется при участии домашнего реестра абонентов HLR, в системе cdma2000 для целей управления мобильностью используется расширенная версия протокола Mobile IP. Учитывая, что новые поколения мобильных сотовых сетей неизбежно будут ориентированы, в основном, на приложения и услуги мобильного Интернета, наличие расширенной версии Mobile IP (соответственно совместимость с протоколом IP) является несомненным достоинством cdma2000. В заключение отметим, что системы UMTS и cdma2000 обеспечивают схожее подключение к телефонной сети общего пользования и к сетям пакетной передачи данных. Более подробное описание архитектуры и протоколов системы cdma2000 можно найти в документах 3G PP2 [2–5].

6.8. Технология и архитектура сетей LTE

Развитие беспроводной связи сопровождается непрерывной сменой технологий. Объем пакетных данных в сетях сотовой связи третьего поколения (3G) уже превышает объем голосового трафика (рис. 6.12), что связано с внедрением технологий HSPA [13]. Это позволяет использовать их для оказания услуг голосовой

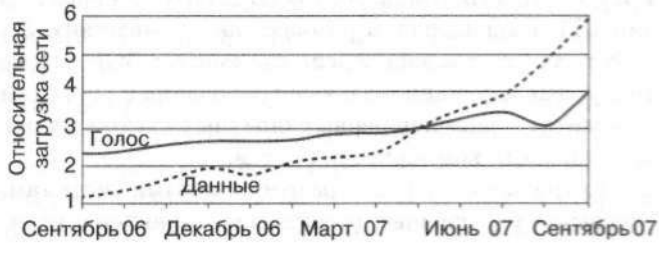


Рис. 6.12. Соотношение объемов трафика голоса и данных в сетях W-CDMA [15]

связи, передачи мультимедийной информации и т. п. В связи с этим само понятие сетей следующего, четвертого, поколения (4G) неразрывно связано (если не синонимично) с созданием универсальных мобильных мультимедийных сетей передачи информации. Сегодня две группы технологий явно нацелены на оказание универсальных услуг связи. Это WiMAX (как развитие линии IEEE 802) и технологии сотовой связи поколений «супер 3G». Причем каждая из них занимает свою нишу на обширном рынке беспроводной связи [28].

Однако требования конечных пользователей к предоставляемым услугам (рис. 6.13) постоянно повышаются. Мобильные сети должны использоваться не только для сотовой связи, но и для передачи видео, мобильного ТВ, музыки и работы с Интернетом с высокими скоростями и качеством передачи. Именно с этой целью в рамках проекта сотрудничества в создании сетей третьего поколения 3GPP (3G Partnership Project) была начата разработка технологии LTE.

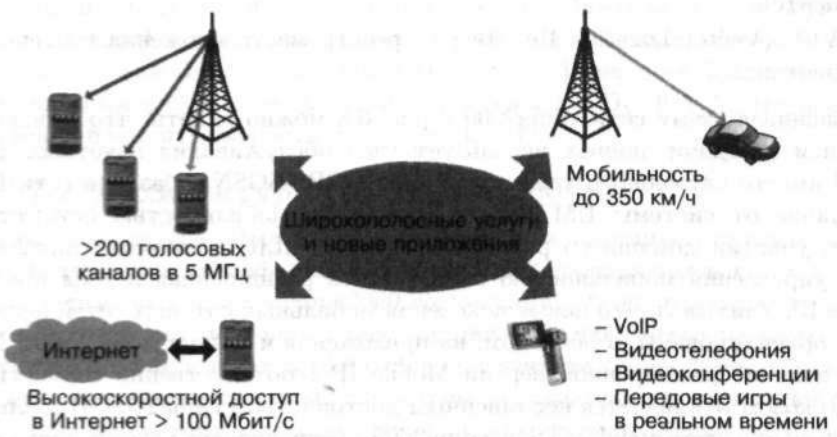


Рис. 6.13. Широкополосные услуги и новые приложения, стимулирующие эволюцию систем 3G

6.8.1. Развитие технологии LTE

Разработка технологии LTE как стандарта официально началась в конце 2004 года (рис. 6.14). Основной целью исследований на начальном этапе был выбор технологии физического уровня, которая смогла бы обеспечить высокую скорость передачи данных. В качестве основных были предложены два варианта: развитие существующего радиоинтерфейса W-CDMA (используемого в HSPA) и создание нового на основе технологии OFDM. В результате проведенных исследований единственной подходящей технологией оказалась OFDM, и в мае 2006 года в 3GPP была создана первая спецификация на радиоинтерфейс Evolved UMTS Terrestrial Radio Access (E-UTRA). Первые, предварительные спецификации LTE создавались в рамках так называемого 3GPP Release 7. А в декабре 2008 года утверждена версия стандартов 3GPP (Release 8), фиксирующая архитектурные и функциональные требования к системам LTE. В середине 2009 года ожидается появление первых опытных систем на основе LTE, а в 2010–2011 годы — первых коммерческих сетей.

По сравнению с ранее разработанными системами 3G, радиоинтерфейс LTE обеспечит улучшенные технические характеристики. В частности, в LTE шири-

на полосы пропускания может варьироваться от 1,4 до 20 МГц (по более ранним источникам — от 1,25 МГц), что позволит удовлетворить потребностям разных операторов связи, обладающих различными полосами пропускания. При этом оборудование LTE должно одновременно поддерживать не менее 200 активных соединений (т. е. 200 телефонных звонков) на каждую 5-МГц ячейку. Также ожидается, что LTE улучшит эффективность использования радиочастотного спектра, т. е. возрастет объем данных, передаваемых в заданном диапазоне частот. LTE позволит достичь внушительных агрегатных скоростей передачи данных — до 50 Мбит/с для восходящего соединения (от абонента до базовой станции) и до 100 Мбит/с для нисходящего соединения (от базовой станции к абоненту) (в полосе 20 МГц). При этом должна обеспечиваться поддержка соединений для абонентов, движущихся со скоростью до 350 км/ч. Зона покрытия одной БС — до 30 км в штатном режиме, но возможна работа с ячейками радиусом более 100 км. Поддерживаются многоантенные системы MIMO.

Рис. 6.14. Основные этапы развития технологии LTE

Радиоинтерфейс LTE позиционируется в качестве решения, на которое операторы будут постепенно переходить с нынешних систем стандартов 3GPP и 3GPP2 [16–19], а его разработка является важным этапом в процессе перехода к сетям четвертого поколения 4G. Фактически спецификация LTE уже содержит большую часть функций, изначально предназначавшихся для систем 4G, поэтому ее иногда именуют «технологией 3,9G».

Но развитие технологии LTE продолжается. Уже разрабатываются спецификации следующего поколения, так называемые LTE-Advanced. И конца этому процессу не видно.

6.8.2. Принципы построения радиоинтерфейса по технологии LTE

LTE базируется на трех основных технологиях: мультиплексирование посредством ортогональных несущих OFDM (Orthogonal Frequency-Division Multiplexing), многоантенные системы MIMO (Multiple Input Multiple Output) и эволюционная системная архитектура сети (System Architecture Evolution).

Принципиально, что дуплексное разделение каналов может быть как частотным (FDD), так и временным (TDD). Это позволяет операторам очень гибко использовать частотный ресурс. Такое решение открывает путь на рынок тем компаниям, которые не обладают спаренными частотами. С другой стороны, поддержка FDD очень удобна для традиционных сотовых операторов, поскольку

у них спаренные частоты есть «по определению» — так организованы практически все существующие системы сотовой связи. Сама же по себе система FDD существенно более эффективна в плане использования частотного ресурса, чем TDD, — в ней меньше накладных расходов (служебных полей, интервалов и т. п.).

Обмен между базовой станцией (БС) и мобильной станцией (МС) строится по принципу циклически повторяющихся кадров (в терминологии LTE — радиокадр) [20]. Длительность радиокадра — 10 мс. Все временные параметры в спецификации LTE привязаны к минимальному временному кванту $T_s = 1/(2048 \cdot \Delta f)$, где Δf — шаг между поднесущими, стандартно — 15 кГц. Таким образом, длительность радиокадра — $307\,200T_s$. Сам же квант времени соответствует тактовой частоте 30,72 МГц, что кратно стандартной в 3G-системах (W-CDMA с полосой канала 5 МГц) частоте обработки 3,84 МГц ($8 \times 3{,}84 = 30{,}72$).

Стандарт LTE предусматривает два типа радиокадров. Тип 1 предназначен для частотного дуплексирования — как для полного дуплекса, так и для полудуплекса. Такой кадр состоит из 20 слотов (длительностью 0,5 мс), нумеруемых от 0 до 19. Два смежных слота образуют субкадр (рис. 6.15). При полнодуплексном режиме радиокадры в восходящем и нисходящем каналах передаются параллельно, но с оговоренным в стандарте временным сдвигом.

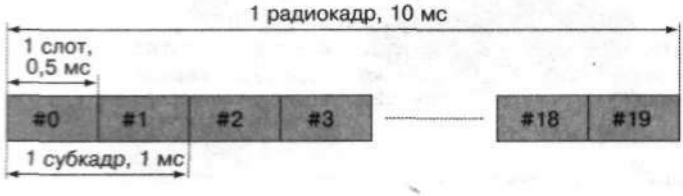


Рис. 6.15. Структура кадра LTE при частотном разделении дуплексных каналов

Радиокадр типа 2 (рис. 6.16) предназначен только для временного дуплексирования. Он состоит из двух полукадров длительностью по 5 мс. Каждый полукадр включает 5 субкадров длительностью 1 мс. Стандарт предусматривает два цикла временного дуплексирования — 5 и 10 мс. В первом случае 1-й и 6-й субкадры идентичны и содержат служебные поля DwPTS, UpPTS и защитный интервал GP. При 10-мс цикле TDD 6-й субкадр используется для передачи данных в нисходящем канале. Субкадры 0 и 5, а также поле DwPTS всегда относятся к нисходящему каналу, а субкадр 2 и поле UpPTS — к восходящему. Распределение остальных субкадров определяется табл. 6.1. Возможно несколько вариантов длительности полей DwPTS, UpPTS и GP, но их сумма всегда равна 1 мс.

Таблица 6.1. Распределение субкадров в радиокадре типа 2

Конфигурация	Цикл TDD, мс	Номер субкадра									
		0	1	2	3	4	5	6	7	8	9
0	5	D	S	U	U	U	D	S	U	U	U
1	5	D	S	U	U	D	D	S	U	U	D
2	5	D	S	U	D	D	D	S	U	D	D
3	10	D	S	U	U	U	D	D	D	D	D
4	10	D	S	U	U	D	D	D	D	D	D
5	10	D	S	U	D	D	D	D	D	D	D
6	5	D	S	U	U	U	D	S	U	U	D

D – нисходящий канал, U — восходящий, S — субкадр со специальными полями

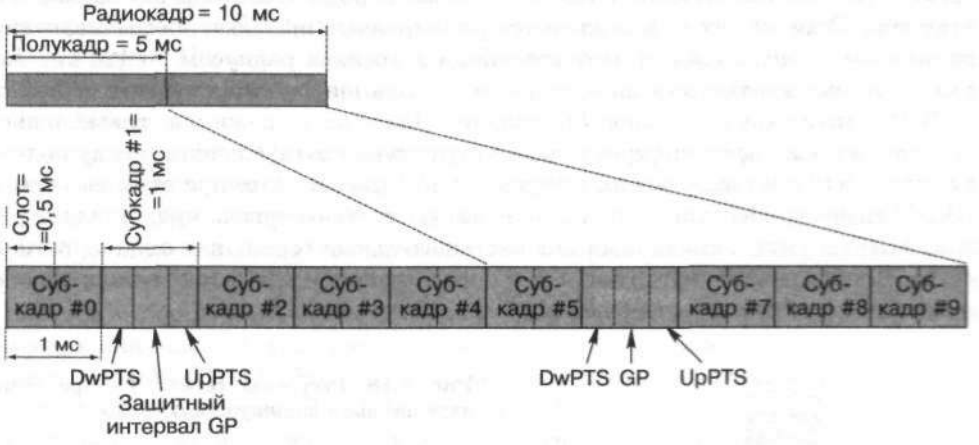


Рис. 6.16. Структура кадра LTE при временном разделении дуплексных каналов

Как уже отмечалось, в LTE используется модуляция OFDM, хорошо исследованная в системах DVB, Wi-Fi и WiMAX [13]. Напомним, технология OFDM предполагает передачу широкополосного сигнала посредством независимой модуляции узкополосных поднесущих вида $S_k(t) = a_k \cdot \sin[2\pi(f_0 + k\Delta f)]$, расположенных с определенным шагом по частоте Δf. Один OFDM-символ содержит набор модулированных поднесущих. Во временной области OFDM-символ включает поле данных (полезная информация) и так называемый циклический префикс CP (Cyclic Prefix) — повторно передаваемый фрагмент конца предыдущего символа (рис. 6.17). Назначение префикса — борьба с межсимвольной интерференцией в приемнике вследствие многолучевого распространения сигнала. Отраженный сигнал, приходящий с задержкой, попадает в зону префикса и не накладывается на полезный сигнал. В LTE принят стандартный шаг между поднесущими $\Delta f = 15$ кГц, что соответствует длительности OFDM-символа 66,7 мкс.

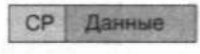


Рис. 6.17. OFDM-символ с циклическим префиксом

Каждому абонентскому устройству (АУ) в каждом слоте назначается определенный диапазон канальных ресурсов в частотно-временной области (рис. 6.18) — ресурсная сетка. Ячейка ресурсной сетки — так называемый ресурсный элемент — соответствует одной поднесущей в частотной области и одному OFDM-символу — во временной. Ресурсные элементы образуют ресурсный блок — минимальную информационную единицу в канале. Ресурсный блок занимает 12 поднесущих (т. е. 180 кГц) и семь или шесть OFDM-символов, в зависимости от типа циклического префикса (табл. 6.2) — так, чтобы общая длительность слота составляла 0,5 мс. Число ресурсных блоков NRB в ресурсной сетке зависит от ширины полосы канала и составляет от шести до 110 (ширина частотных полос восходящего/нисходящего каналов в LTE — от 1,4 до 20 МГц). Ресурсный блок — это минимальный ресурсный элемент, выделяемый абонентскому устройству планировщиком базовой станции. О распределении ресурсов в каждом слоте базовая станция сообщает в специальном управляющем канале.

Длительность префикса 4,7 мкс позволяет бороться с задержкой отраженного сигнала, прошедшего путь на 1,4 км больше, чем прямо распространяющийся

сигнал. Для систем сотовой связи в условиях города этого обычно вполне достаточно. Если же нет — используется расширенный префикс, обеспечивающий подавление межсимвольной интерференции в ячейках радиусом до 120 км. Такие огромные ячейки полезны для разного рода широковещательных сервисов (MBMS), таких как мобильное ТВ-вещание. Для этих же режимов (только в нисходящем канале) предусмотрена особая структура слота, с шагом между поднесущими 7,5 кГц и циклическим префиксом 33,4 мкс. В слоте при этом всего три OFDM-символа. Особый случай широковещательного сервиса представляет режим MBSFN (мультимедийный широковещательный сервис для одночастотной сети). В этом режиме несколько БС в определенной MBSFN-зоне одновременно и синхронно транслируют общий широковещательный сигнал.

Рис. 6.18. Ресурсная сетка LTE при стандартном шаге поднесущих $\Delta f = 15$ кГц

Таблица 6.2. Физический префикс в нисходящем канале при $\Delta f = 15$ кГц

Тип префикса	Длина префикса		Длина слота, OFDM-символов
	T_s	мкс	
Стандартный:			
первый символ слота	160	5,2	7
остальные 6 символов слота	144	4,7	
Расширенный	512	16,7	6

Каждая поднесущая модулируется посредством 4-, 16- и 64-позиционной квадратурной фазово-амплитудной модуляции (QPSK, 16-QAM или 64-QAM). Соответственно, один символ на одной поднесущей содержит 2, 4 или 6 бит. При

стандартном префиксе символьная скорость составит 14000 символов/с, что соответствует, при FDD-дуплексе, агрегатной скорости от 28 до 84 кбит/с на поднесущую. Сигнал с полосой 20 МГц содержит 100 ресурсных блоков или 1200 поднесущих, что дает общую агрегатную скорость в канале от 33,6 до 100,8 Мбит/с.

Спецификации LTE определяют несколько фиксированных значений для ширины восходящего и нисходящего каналов между БС и АС (в сетях E-UTRA) (табл. 6.3). Поскольку в OFDM используется быстрое преобразование Фурье (БПФ), число формальных поднесущих для упрощения процедур цифровой обработки сигнала должно быть кратно $N = 2^n$ (т. е. 128, 256, ..., 2048). При этом частота выборок должна составлять $F_s = \Delta f \cdot N$. При заданных в стандарте значениях она оказывается кратной 3,84 МГц — стандартной частоте выборок в технологии W-CDMA. Это очень удобно для создания многомодовых устройств, поддерживающих как W-CDMA, так и LTE. Разумеется, при формировании сигнала амплитуды «лишних» поднесущих (включая центральную поднесущую канала) считаются равными нулю.

Таблица 6.3. Параметры канала передачи между БС и АУ

Ширина канала, МГц	1,4	3	5	10	15	20
Число ресурсных блоков	6	15	25	50	75	100
Число поднесущих	72	180	300	600	900	1200
Число номинальных несущих для БПФ	128	256	512	1024	1536	2048
Тактовая частота для БПФ, МГц	1,92	3,84	7,68	15,36	23,04	30,72

6.8.3. Нисходящий канал

В нисходящем и восходящем каналах применение технологии OFDM различно. В нисходящем канале эта технология используется не только для передачи сигнала, но и для организации множественного доступа (OFDMA) — т. е. для мультиплексирования абонентских каналов.

Помимо описанного физического структурного блока вводится понятие логического структурного блока. По числу ресурсных элементов они эквивалентны, однако возможно два варианта отображения ресурсных элементов физического блока в логический — один в один и распределенно. В последнем случае элементы логического ресурсного блока оказываются распределенными по всей доступной ресурсной сетке.

В отличие от пакетных сетей, в LTE нет физической преамбулы, которая необходима для синхронизации и оценки смещения несущей. Вместо этого в каждый ресурсный блок добавляются специальные опорные и синхронизирующие сигналы. Опорные сигналы могут быть трех видов — опорный сигнал, характеризующий ячейку (Cell-specific), сигнал, связанный с конкретным абонентским устройством, и сигнал для специального широковещательного мультимедийного сервиса MBSFN. Опорный сигнал служит для непосредственного определения условий в канале передачи (поскольку приемнику известно его месторасположение и исходная форма). На основе этих измерений можно определить реакцию канала для остальных поднесущих и с помощью интерполяции восстановить их исходную форму.

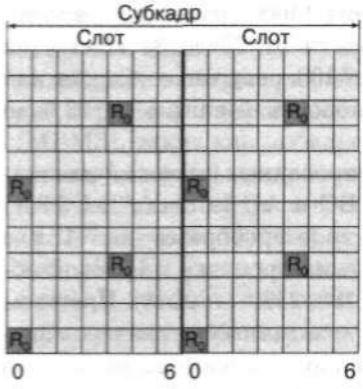


Рис. 6.19. Расположение опорного сигнала (cell-specific) в нисходящем канале ресурсной сетки LTE в случае работы с одной антенной

Опорный cell-specific-сигнал должен присутствовать в каждом субкадре нисходящего канала (кроме случаев MBSFN-передачи). Форма сигнала определяется на основе псевдослучайной последовательности Голда (вариант m-последовательности), при инициализации которой используется идентификационный номер ячейки БС (Cell ID). Такой опорный сигнал равномерно распределен по ресурсным элементам (рис. 6.19). Так, при стандартной длине префикса он транслируется в 0-м и 4-м OFDM-символе, при расширенном CP — во время 0-го и 3-го OFDM-символа. В частотной области опорные сигналы передаются через каждые шесть поднесущих, причем смещение определяется идентификатором ячейки, взятым по модулю 6.

Помимо опорных сигналов, в нисходящем канале транслируются и синхронизирующие сигналы. Синхронизирующие сигналы также однозначно определяют Cell ID. В LTE принята иерархическая структура идентификации ячейки, как и в предшествующей ей технологии W-CDMA. Предполагается, что на физическом уровне доступно 504 Cell ID. Они разбиты на 168 ID-групп, по три идентификатора в каждой. Номер группы N_1 (0–167) и номер идентификатора в ней N_2 (0–2) однозначно определяют ID ячейки. Используется два синхросигнала — первичный и вторичный. Первичный синхросигнал представляет собой 62-элементную последовательность в частотном плане, задаваемую последовательностью Задова-Чу на основе идентификатора N_2. Такая последовательность из 62 поднесущих, распределенных по ресурсной сетке симметрично относительно ее центральной частоты, передается в радиокадре типа 1 в последнем OFDM-символе слотов 0 и 10 (субкадры 0 и 5). В радиокадре типа 2 для передачи первичного синхросигнала используется третий OFDM-символ субкадров 1 и 6. Вторичный синхросигнал генерируется на основе номера ID-группы N_1. Он передается в слотах 0 и 10 радиокадра типа 1 (пятый OFDM-символ при стандартном CP) и в слотах 1 и 11 радиокадра типа 2 (шестой OFDM-символ при стандартном CP).

Формирование сигнала в нисходящем канале достаточно стандартно для современных систем цифровой передачи информации (рис. 6.20). Оно включает процедуры канального кодирования, скремблирования, формирования модуляционных символов, их распределения по антенным портам и ресурсным элементам и синтеза OFDM-символов. Канальное кодирование подразумевает вычисление контрольных сумм (CRC-24) для блоков данных, поступающих с MAC-уровня. Затем блоки с контрольными суммами обрабатываются посредством кодера со скоростью кодирования 1/3. В LTE предусмотрено применение либо сверточного кода, либо турбокода. Кодированная последовательность после перемежения (интерливинга) поступает в скремблер (для входной последовательности $\{x(i)\}$ выполняется процедура вида $d_{scr}(i) = x(i) + c(i)$, где $c(i)$ — определенная скремблирующая последовательность). Затем формируются комплексные модуляционные символы (QPSK, 16- и 64-QAM) и распределяются

по ресурсным элементам. Далее происходит синтез OFDM-символов, их последовательность поступает в модулятор, формирующий выходной ВЧ-сигнал в заданном частотном диапазоне. На стороне приема все процедуры выполняются в обратном порядке.

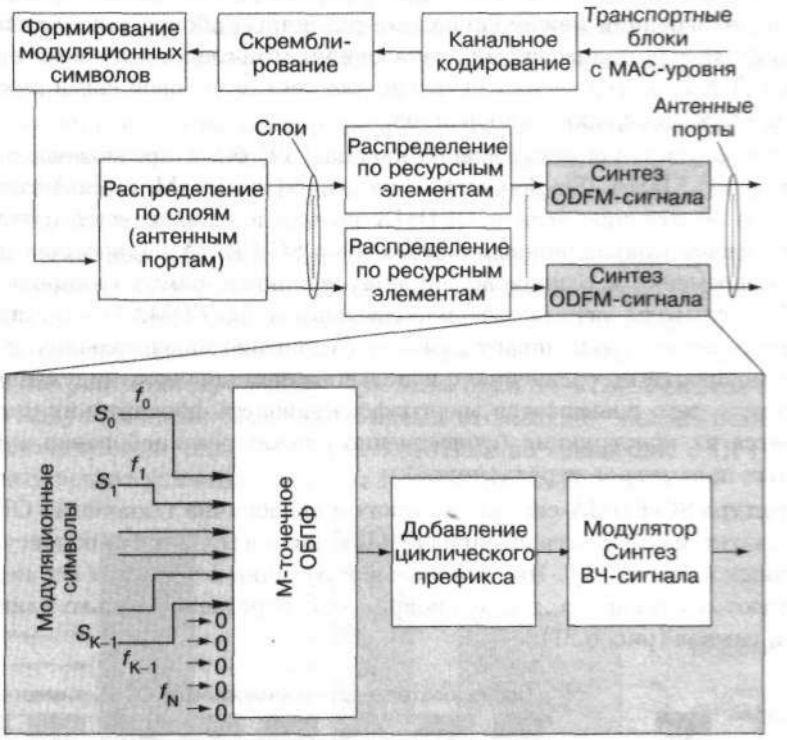


Рис. 6.20. Схема формирования сигнала в нисходящем канале

6.8.4. Восходящий канал

Применение OFDM в сочетании с циклическим префиксом делает связь устойчивой к временной дисперсии параметров радиоканала, в результате на приемной стороне становится не нужным сложный эквалайзер. Это очень полезно для организации нисходящего канала, поскольку упрощается обработка сигнала приемником, что снижает стоимость терминального устройства и потребляемую им мощность.

В восходящем канале допустимая мощность излучения значительно ниже, чем в нисходящем. Поэтому первичным становится энергетическая эффективность метода передачи информации с целью увеличения зоны покрытия, снижения стоимости терминального устройства и потребляемой им мощности.

Основной недостаток технологии OFDMA — высокое соотношение пиковой и средней мощности сигнала (PAR). Это связано с тем, что во временной области спектр OFDM-сигнала становится аналогичным гауссову шуму, характеризующемуся высоким PAR. Кроме того, сама по себе технология OFDMA,

с учетом необходимости минимизировать шаг между поднесущими и сокращать относительную длительность CP, предъявляет очень высокие требования к формированию композитного сигнала. Мало того, что частотные рассогласования между передатчиком и приемником и фазовый шум в принимаемом сигнале могут привести к межсимвольной интерференции на отдельных поднесущих (т. е. к интерференции между сигналами различных абонентских каналов). При малом шаге между поднесущими к аналогичным последствиям может привести и эффект Доплера, что очень актуально для систем сотовой связи, предполагающих высокую мобильность абонентов.

В связи с этим для восходящего канала LTE была предложена новая технология — SC-FDMA (Single-Carrier Frequency-Division Multiple Access). Принципиальное ее отличие: если в OFDMA на каждой поднесущей одновременно передается свой модуляционный символ, то в SC-FDMA поднесущие модулируются одновременно и одинаково, но модуляционные символы короче. То есть в OFDMA символы передаются параллельно, в SC-FDMA — последовательно. Такое решение обеспечивает меньшее отношение максимального и среднего уровней мощности по сравнению с использованием обычной модуляции OFDM, в результате чего повышается энергоэффективность абонентских устройств и упрощается их конструкция (существенно снижаются требования к точности частотных параметров передатчиков).

Структура SC-FDMA-сигнала во многом аналогична технологии OFDM. Так же используется композитный сигнал — модуляция множества поднесущих, расположенных с шагом Δf. Принципиальное отличие в том, что все поднесущие модулируются одинаково, т. е. единовременно передается только один модуляционный символ (рис. 6.21).

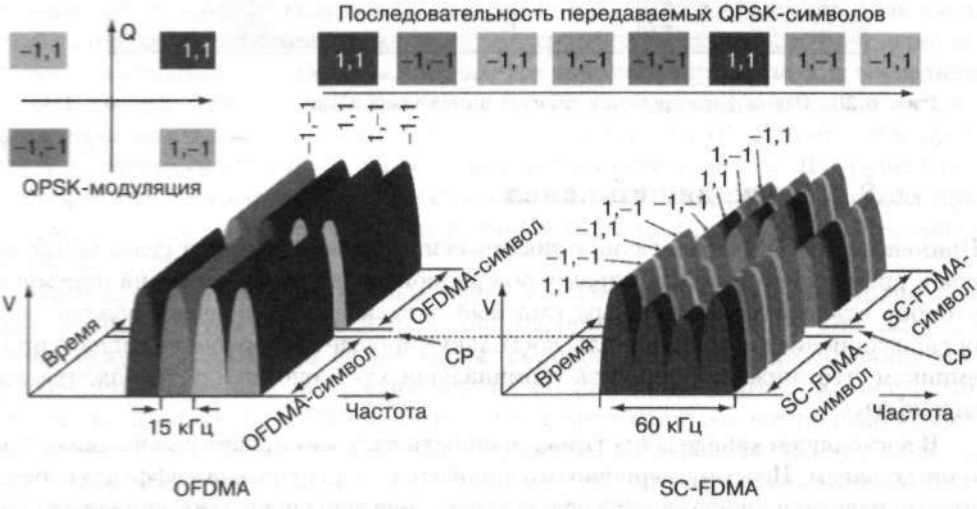


Рис. 6.21. Различие между OFDMA и SC-FDMA при передаче последовательности QPSK-символов

При этом ресурсная сетка полностью аналогична нисходящему каналу. Так же каждый физический ресурсный блок, соответствующий слоту, занимает 12 поднесущих с шагом $\Delta f = 15$ кГц в частотной области (всего 180 кГц)

и 0,5 мс — во временной. Ресурсному блоку соответствуют 7 SC-FDMA-символов при стандартном циклическом префиксе и 6 — при расширенном. Длительность SC-FDMA-символа (без префикса) равна длительности OFDMA-символа и составляет 66,7 мкс (длительности соответствующих циклических префиксов также равны). В сетке может быть от 6 до 110 ресурсных блоков, но их число должно быть кратно 2, 3 или 5, что связано с процедурой дискретного Фурье-преобразования. Еще одна особенность — поддержка модуляции 64-QAM в АУ опциональна.

Каждому абоненту сети для передачи данных от базовой станции с помощью функции планирования на определенное время выделяется определенное число ресурсных блоков. Расписание передается абонентам по служебным каналам в нисходящем радиоканале.

Однако если при OFDMA один модуляционный символ (QPSK, 16- или 64-QAM) соответствует OFDM-символу на одной поднесущей (15 кГц, 66,7 мкс), то при SC-OFDMA ситуация иная. В частотном плане ширина модуляционного символа оказывается равной всей доступной полосе частот (он передается на всех поднесущих одновременно). При этом один SC-FDMA-символ содержит несколько модуляционных символов (в идеале столько же, сколько поднесущих), но в соответствующее число раз более коротких по сравнению с OFDMA, что полностью отвечает условиям теоремы Котельникова-Шеннона.

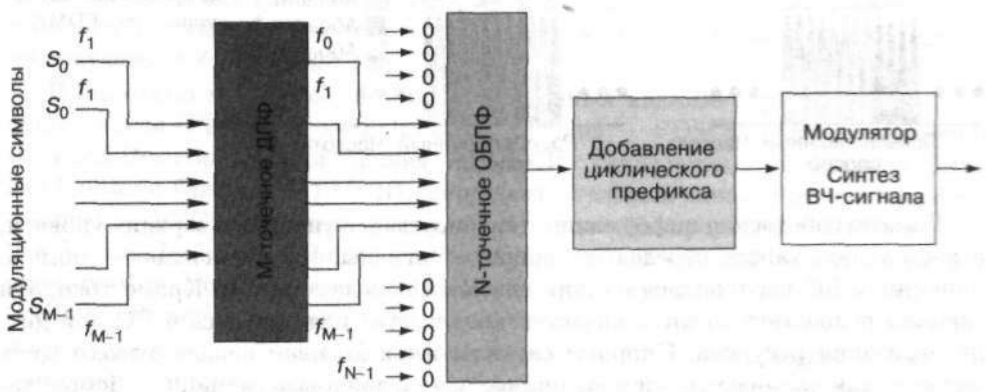


Рис. 6.22. Особенность формирования выходного сигнала в случае SC-FDMA

Сама процедура формирования SC-FDMA-сигнала отличается от схемы OFDMA. После канального кодирования, скремблирования и формирования модуляционных символов они группируются в блоки по M символов — субсимволов SC-FDMA (рис. 6.22). Очевидно, что непосредственно отнести их на поднесущие с шагом 15 кГц невозможно — требуется в N раз более высокая частота, где N — это число доступных для передачи поднесущих. Поэтому, сформировав группы по M модуляционных символов ($M < N$), их подвергают M-точечному дискретному Фурье-преобразованию (ДПФ), т. е. формируют аналоговый сигнал. А уже затем с помощью стандартной процедуры обратного N-точечного Фурье-преобразования синтезируют сигнал, соответствующий независимой модуляции каждой поднесущей, добавляют циклический префикс и генерируют выходной ВЧ-сигнал. В результате такого подхода передатчик

и приемник OFDMA- и SC-FDMA-сигналов имеют схожую функциональную структуру (см. рис. 6.20 и 6.22).

Отметим, что АУ может использовать как фиксированный частотный диапазон (используются смежные ресурсные блоки, т. е. смежные поднесущие), так и распределенный — так называемый режим скачкообразной перестройки частоты (FH). В последнем случае для каждого слота восходящего канала используется новый ресурсный блок из доступной ресурсной сетки. Параметры перестройки частоты задаются сетевым оборудованием и сообщаются как при инициализации абонентской станции в сети, так и по ходу работы в канале управления. В случае распределенного способа информация от каждого абонента расположена во всем спектре сигнала (рис. 6.23), поэтому данный способ устойчив к частотно-избирательному замиранию. С другой стороны, при локализованном способе распределения возможно определить полосу, в которой для данного абонента достигается максимальная устойчивость канала к замираниям. Поскольку области замирания сигнала для всех абонентов различны, то можно достичь общую максимальную эффективность использования радиоканала. Однако это требует непрерывного сканирования частотной характеристики канала для каждого устройства и организации функции диспетчеризации.

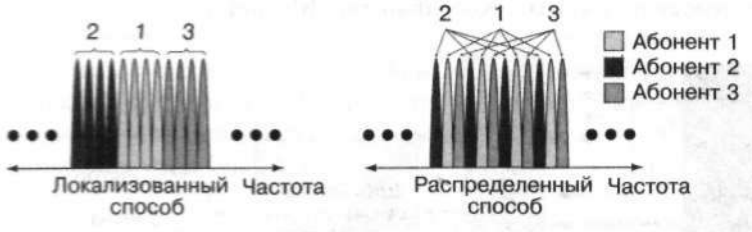


Рис. 6.23. Способы распределения поднесущих в SC-FDMA

Помимо собственно информации, генерируемой функциями верхних уровней, в восходящем канале передаются опорные сигналы. Их назначение — помочь приемнику БС настроиться на определенный передатчик АУ. Кроме того, эти сигналы позволяют оценить качество канала, что используется в БС при диспетчеризации ресурсов. Опорные сигналы в восходящем канале бывают двух видов — так называемые «демодулированные» и зондовые (sounding). Демодулированные опорные сигналы аналогичны опорным сигналам нисходящего канала. Они передаются на постоянной основе. Так, в общем информационном канале последовательность демодулированного опорного сигнала передается в четвертом SC-FDMA-символе каждого слота при стандартном CP. Зондовые сигналы апериодичны. Их основное назначение — дать БС возможность оценить качество канала, если передача еще не ведется.

6.8.5. Информационные потоки

До сих пор мы говорили о способе формирования физического канала обмена между абонентскими и базовыми станциями. Однако как в восходящем, так и в нисходящем каналах передаются различные типы информационных потоков.

В восходящем канале их три — канал общего пользования назначения (PUSCH), управляющий канал (PUCCH) и канал произвольного доступа (PURCH). Назначение первого очевидно — передача информации пользователей.

Управляющий канал содержит такую информацию, как индикатор качества канала, сообщения подтверждения доставки (ACK/NACK) и запрос на получение расписания (о доступных ресурсах). Канал общего пользования и управляющий канал никогда не транслируются одновременно одним АУ. Для передачи управляющего канала используется один ресурсный блок в каждом из слотов одного субкадра. В зависимости от формата PUCCH возможны четыре варианта его расположения на ресурсной сетке (рис. 6.24), определяемые переменной m.

Канал произвольного доступа служит для запроса начальной инициализации в сети, при хэндовере, при выходе из режима ожидания в активный режим и т.п. Абонентской станции назначается интервал в ресурсной сетке (номер физического ресурсного блока и номер субкадра), в течение которого она передает специальный пакет — преамбулу произвольного доступа. Преамбула генерируется на основе последовательностей Задова-Чу с нулевой зоной корреляции, всего определено 64 различных преамбулы на одну ячейку. БС, приняв запрос доступа, отвечает в том же самом канале произвольного доступа (но уже нисходящем) подтверждением. Если подтверждение не получено, АУ повторяет запрос.

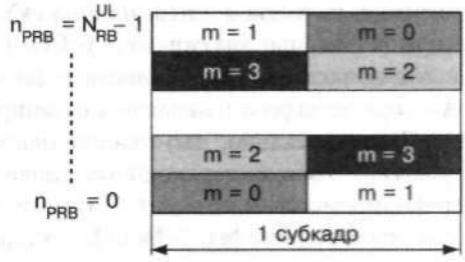


Рис. 6.24. Варианты расположения канала управления PUCCH в нисходящем канале

В нисходящем направлении информационных каналов гораздо больше. Это общий канал (Physical Downlink Shared Channel — PDSCH); канал управления (Physical Downlink Control Channel — PDCCH); канал групповой передачи (Physical Multicast Channel — PMCH); широковещательный канал (Physical Broadcast Channel — PBCH); индикаторный канал управления форматом (Physical Control Format Indicator Channel — PCFICH) и индикаторный канал гибридной процедуры повторного запроса (HARQ) Physical Hybrid ARQ Indicator Channel (PHICH). Назначение общего канала очевидно — передача данных конкретным абонентским устройствам. В канале управления PDCCH передаются таблицы с назначением канальных ресурсов абонентским устройствам — как в нисходящем, так и в восходящем каналах. В канале PCFICH, который передается в каждом субкадре, указываются номера OFDM-символов, которые используются для трансляции сообщений канала управления PDCCH. Канал PHICH предназначен для подтверждения доставки данных в восходящем канале. Назначение каналов групповой передачи и широкого вещания также очевидно. Отметим особенность широковещательного канала — каждый его блок с верхних уровней протокола транслируется в четырех субкадрах, следующих с жестко фиксированным интервалом в 40 мс. Это исключает необходимость в дополнительных указателях на расположение этих субкадров.

6.8.6. Многоантенные системы

Как и все современные технологии беспроводной связи, в LTE поддерживаются многоантенные системы (MIMO). Учитывая ориентацию этой технологии на максимально простые абонентские устройства, техника MIMO в LTE

максимально упрощена. Стандарт рассматривает MIMO-схемы с одной, двумя и четырьмя передающими и приемными антеннами в различных сочетаниях. В MIMO-системах есть два основных вида передачи — пространственное мультиплексирование и диверсифицированная передача. Первый режим означает, что каждый антенный канал транслирует независимый информационный поток. При этом сами каналы должны быть некоррелированными. Возможно два вида пространственно-мультиплексированной передачи — для одного АУ (SU-MIMO) и для группы АУ (MU-MIMO). В первом случае БС передает несколько независимых потоков данных одному АУ. При этом в АУ должно быть по крайней мере не меньше антенн, чем у БС. В MU-MIMO ресурсные элементы с одинаковыми частотно-временными параметрами должны приниматься различными АУ (при этом речь о цифровом формировании диаграммы направленности не идет).

Принципиально, что одновременно по всем антенным каналам может передаваться только два кодовых слова (т. е. только два логически независимых информационных потока). Поэтому, несмотря на четыре возможных антенных канала, в режиме MU-MIMO БС в одном частотно-временном диапазоне способна работать только с двумя АУ.

Диверсифицированная передача означает, что несколько антенных каналов используются для передачи одного потока данных. Эта техника предназначена для борьбы с замираниями в радиоканале и направлена только на улучшение качества передачи в канале. На скорость передачи она влияет опосредованно, через повышение качества канала.

В восходящем канале возможна схема пространственного мультиплексирования множества абонентов MU-MIMO. Несколько АУ, каждое с одной антенной, могут использовать одинаковые частотно-временные ресурсы, но за счет декоррелированных антенных каналов БС работает со всеми ними одновременно.

6.8.7. Механизм диспетчеризации и повторные передачи

Под диспетчеризацией понимается процесс распределения сетевых ресурсов между пользователями. Цель диспетчеризации — сбалансировать качество связи и общую производительность системы. В LTE предусмотрена динамическая и статическая диспетчеризация. Динамическая диспетчеризация распределяет ресурсы в зависимости от текущего состояния канала связи. Она обеспечивает передачу данных на повышенных скоростях (за счет модуляции более высокого порядка, уменьшения степени кодировки каналов, передачи дополнительных потоков данных и меньшего числа повторных передач), задействуя для этого временные и частотные ресурсы с относительно хорошими условиями связи. Таким образом, для передачи любого конкретного объема информации требуется меньше времени.

Для трафика сервисов, пересылающих пакеты с небольшой полезной нагрузкой и через одинаковые промежутки времени (например, IP-TV), объем служебной информации, необходимой для динамической диспетчеризации, может превысить объем полезных данных. Для таких случаев в LTE предусмотрена статическая диспетчеризация.

Для надежной передачи информации в технологии LTE реализована ставшая традиционной система повторной передачи Hybrid Automatic Repeat Request

(HARQ). Особенность ее реализации в LTE в том, что одновременно может поддерживаться несколько (до восьми) HARQ-процессов. Если данные (субкадр), связанные с HARQ-процессом, пришли успешно, приемник отправляет сообщение об успешном приеме/неприеме данных (ACK/NACK). В случае отсутствия подтверждения или сообщения NACK происходит повторная передача. В нисходящем канале расположение и параметры (тип сигнально-кодовой конструкции) повторно передаваемого субкадра сообщаются дополнительно, в канале управления — так называемая адаптивная передача, когда БС выбирает оптимальный ресурс для ретрансляции. В восходящем канале, если АУ не получило сообщения ACK, оно должно повторить передачу. БС может сообщить АУ параметры субкадра для повторной передачи. Если же по каналу управления такого сообщения не поступило, АУ повторяет передачу субкадра с точно такими же параметрами, как и у исходного субкадра, прием которого не был подтвержден — неадаптивная ретрансляция. Повторная передача происходит через заданное в спецификации LTE число субкадров (от 4 до 9), которое зависит от типа дуплексирования, типа радиокадра, схемы распределения каналов в случае TDD и номера неверно принятого субкадра.

6.8.8. Сетевая архитектура SAE

Для технологии LTE консорциум 3GPP предложил новую сетевую инфраструктуру (SAE — System Architecture Evolution). Цель и сущность концепции SAE — эффективная поддержка широкого коммерческого использования любых услуг на базе IP и обеспечение непрерывного обслуживания абонента при его перемещении между сетями беспроводного доступа, которые не обязательно соответствуют стандартам 3GPP (GSM, UMTS, W-CDMA и т. д.) (рис. 6.25) [21].

В сети с архитектурой SAE могут применяться узлы только двух типов — базовые станции (evolved Node B, eNodeB) и шлюзы доступа (Access Gateway, AGW). Уменьшение числа типов узлов позволит операторам снизить расходы как на развертывание сетей LTE/SAE, так и на их последующую эксплуатацию. Ядро сети SAE включает в себя четыре ключевых компонента:

- модуль управления мобильностью (Mobility Management Entity, MME) обеспечивает хранение служебной информации об абоненте и управление ею, авторизацию терминальных устройств в наземных сетях мобильной связи и общее управление мобильностью;
- модуль управления абонентом (User Plane Entity, UPE) отвечает за установление нисходящего соединения, шифрование данных, маршрутизацию и пересылку пакетов;
- 3GPP-якорь играет роль шлюза между сетями 2G/3G и LTE;
- SAE-якорь используется для поддержки непрерывности сервиса при перемещении абонента между сетями, как соответствующими спецификациям 3GPP, так и не соответствующими (WLAN и т. п.).

Последние два компонента представляют собой совершенно новые элементы архитектуры ядра сети мобильной связи (Evolved Packet Core) и обязаны своим появлением требованию поддержки мобильности при перемещении абонента между сетями разных типов.

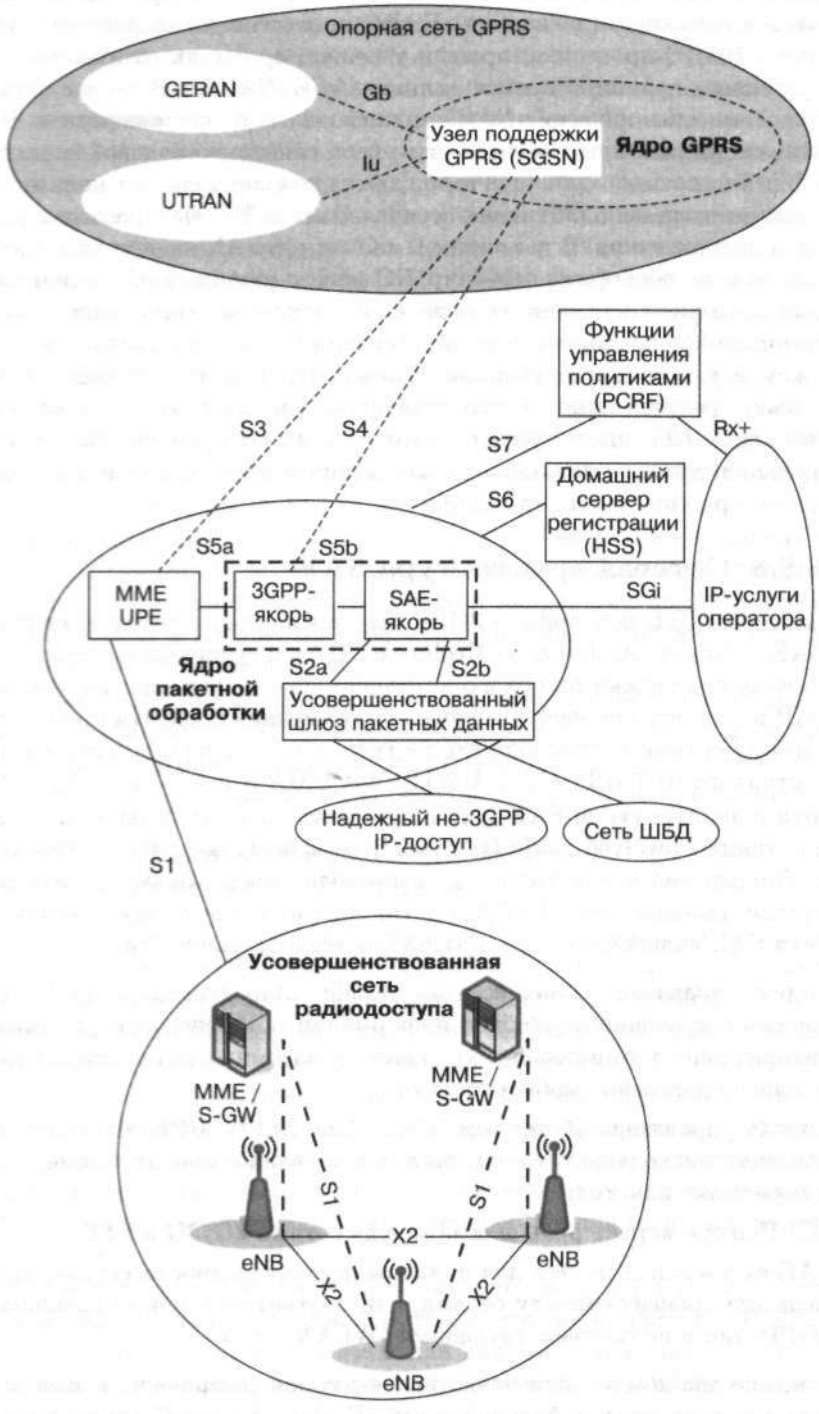


Рис. 6.25. Основные компоненты архитектуры SAE [21]

Функциональные элементы можно по-разному распределять среди аппаратуры сети. Например, 3GPP-якорь допустимо (но не обязательно) располагать вместе с модулем управления абонентом. Аналогично, модули MME и UPE можно совмещать либо реализовывать в разных узлах сети.

Важная особенность SAE — пользовательские данные могут пересылаться между базовыми станциями непосредственно, причем как с помощью проводной, так и беспроводной связи (интерфейс X2). Это особенно важно при хэндовере, для быстрого бесшовного переключения пользователя между БС. Разумеется, допустимо передавать данные между БС и через шлюзы транспортной IP-сети. Возможность непосредственной беспроводной передачи данных между БС фактически означает, что в архитектуре SAE заложена функциональность mesh-сети.

Значительное внимание в документах 3GPP Release 8 уделено обеспечению качества сервиса, выбору сети и использованию идентификационных данных. Появление многомодовых терминалов, предназначенных, например, для работы в сетях Wi-Fi и сотовой связи, позволяет обслуживать абонентов с применением разных вариантов доступа. В этой связи в SAE предусмотрены механизмы выбора наиболее удобной инфраструктуры для предоставления услуг, необходимых абоненту.

Как отмечают разработчики SAE, предложенные ими архитектурные изменения позволят значительно уменьшить задержки передачи данных, которые особенно критичны для таких приложений, как VoIP или онлайновые интерактивные игры. В частности, задержки радиосети при передаче данных пользователя не должны превышать 10 мс (5 мс для коротких IP-пакетов при небольшой сетевой нагрузке). Эти значения, по крайней мере, на 50% лучше аналогичных показателей наиболее совершенных сетей 3G.

6.8.9. Дальнейшие пути развития LTE

Не дожидаясь окончания работ над стандартом 3GPP Release 8, многие ведущие производители телекоммуникационного оборудования уже представили свои первые опытные образцы устройств, поддерживающих LTE. Так, в феврале 2007 года компания Ericsson впервые в мире продемонстрировала работу оборудования LTE со скоростью передачи 144 Мбит/с. В сентябре 2007 года компания NTT Docomo представила оборудование LTE со скоростью передачи 200 Мбит/с и потребляемой мощностью менее 100 мВт. В апреле 2008 года корпорации LG и Nortel продемонстрировали передачу данных по технологии LTE с пропускной способностью 50 Мбит/с при скорости мобильных абонентов 110 км/ч. 18 сентября 2008 года. мобильный оператор T-Mobile и Nortel Networks объявили о достижении скоростей передачи 170 Мбит/с для нисходящего соединения и 50 Мбит/с для восходящего. Испытания проводились в машине на средней скорости 67 км/ч в радиусе действия трех базовых станций.

Дальнейшее развитие технологии LTE будет продолжаться в рамках работ над новым стандартом 3GPP Release 10 (LTE Advanced). Сегодня уже сформулированы основные требования, которым должен будет удовлетворять LTE Advanced [22]. По сути, это требования к стандарту мобильных сетей четвертого поколения (4G):

- максимальная скорость передачи данных в нисходящем радиоканале до 1 Гбит/с, в восходящем — до 500 Мбит/с (средняя пропускная способность на одного абонента — в три раза выше, чем в LTE);

- полоса пропускания в нисходящем радиоканале — 70 МГц, в восходящем — 40 МГц;
- максимальная эффективность использования спектра в нисходящем радиоканале — 30 бит/с/Гц, в восходящем — 15 бит/с/Гц (втрое выше, чем в LTE);
- полная совместимость и взаимодействие с LTE и другими 3GPP системами.

Для решения этих задач предполагается использовать более широкие радиоканалы (до 100 МГц), ассиметричное разделение полос пропускания между восходящим и нисходящим каналами в случае частотного дуплекса; более совершенные системы кодирования и исправления ошибок; гибридную технологию OFDMA и SC-FDMA для восходящего канала, а также передовые решения в области антенных систем (MIMO).

Очевидно, что в столь кратком обзоре мы лишь затронули только основные особенности технологии LTE и только относящиеся к физическому уровню модели взаимодействия открытых систем. Отметим лишь, что эта технология сегодня находится в стадии бурного развития, ежемесячно происходят изменения, в том числе — и в самих стандартах. В самих спецификациях LTE еще хватает незаполненных мест, ошибок, неточностей и неопределенностей. Явно следует ожидать появления новых документов в области сетевой архитектуры. Но безусловное достоинство технологии LTE — ее открытость. Любой специалист может найти ответы на интересующие его вопросы в многочисленных стандартизирующих документах 3GPP, выложенных на сайте http://www.3gpp.org/ftp.

Литература

1. Аверин С. Обзор рынка сотовой связи РФ. Итоги 4Q05 и 2005 года. — www.sotovik.ru.
2. 3GPP2 C.S0001-D v2.0. Introduction to cdma2000 Standards for Spread Spectrum Systems. 10.2005.
3. 3GPP2 S.R0005-B v2.0. Network Reference Model for cdma2000 Spread Spectrum Systems — Revision B. 06.2007.
4. 3GPP2 S.R0134-0 v1.0. Evolution of Ultra Mobile Broadband — System Requirements Document. 02.2009.
5. 3GPP2 C.S0084-000-A v1.0. Overview for Ultra Mobile Broadband (UMB) Air Interface Specification. 12.2008.
6. 3GPP TS 25.201 V6.2.0 Physical layer — General description.
7. 3GPP TS 25.213 V6.5.0 Spreading and modulation (FDD).
8. 3GPP TS 25.301 V6.4.0 Radio Interface Protocol Architecture.
9. 3GPP TS 25.401 V6.9.0 UTRAN Overall Description.
10. Кааранен Х. и др. Сети UMTS. Архитектура, мобильность, сервисы. — М.: Техносфера, 2007, 460 с.
11. Ле Бодик Г. Технологии и службы мобильной передачи данных SMS, EMS и MMS. — М.: Техносфера, 2008, 543 с.
12. Джамалипур А. Беспроводной мобильный интернет: архитектура, протоколы и сервисы. — М.: Техносфера, 2009, 494 с.
13. Вишневский В. М., Ляхов А. И., Портной С. Л., Шахнович И. В. Широкополосные беспроводные сети передачи информации. — М.: Техносфера, 2005.

14. Andrews J. G., Ghosh A., Muhamed R. Fundamentals of WiMAX. — Prentice Hall, 2008.
15. Long Term Evolution (LTE): an introduction. White Paper. — Ericsson, October 2007.
16. Jim Zyren. Overview of the 3GPP Long Term Evolution Physical Layer. White Paper. — www.freescale.com.
17. Moray Rumney. 3GPP LTE: Introducing Single-Carrier FDMA. — Agilent Measurement Journal, 2008, № 4, p. 18–27.
18. Дальман Э., Фурускар А., Ядинг И. Радиоинтерфейс LTE в деталях. — Сети и Системы связи, 2008, № 9.
19. Tomislav Blaich. Evolution of radio access network in 3G mobile systems. — Revija 19.2006.2, p. 54–68.
20. GPP TS 36.211. Physical Channels and Modulation (Release 8). — 3GPP, 12.2008.
21. 3GPP TR 23.882. 3GPP System Architecture Evolution: Report on Technical Options and Conclusions (Release 8). — 3GPP, 09.2008.
22. 3GPP TR 36.913 V8.0.0. Requirements for Further Advancements for E-UTRA (LTE-Advanced), Release 8. — 3GPP, 06.2008.
23. www.sotovik.ru/library/statistika_world.htm.
24. www.sotovik.ru/ratings/market20010101.htm.
25. Громаков Ю. А. Стандарты и системы подвижной радиосвязи. — М.: Мобильные ТелеСистемы — Эко-Трендз, 1997.
26. Гольдштейн Б. С. Сигнализация в сетях связи. Т. 1. — М.: Радио и связь, 1998.
27. Невдяев Л. CDMA: технологии доступа. — Сети, 2000, № 6.
28. Вишневский В., Красилов А., Шахнович И. Технология сотовой связи LTE — почти 4G. — Электроника: НТБ, 2009, № 1, с. 62–72.

ГЛАВА 7

СТАНДАРТ ШИРОКОПОЛОСНОГО ДОСТУПА IEEE 802.16

7.1. Предыстория стандарта IEEE 802.16

7.1.1. Системы MMDS и LMDS/MVDS

Системы беспроводного широкополосного доступа регионального (городского) масштаба своим массовым развитием изначально обязаны телевидению, которое требовало все новых и более высокоскоростных средств доставки ТВ-трафика зрителям. Кабельные системы со своими задачами справлялись не везде, и в 1970 году. Федеральная комиссия связи (FCC) США представила, по-видимому, первую беспроводную многоточечную систему распределения MDS (Multipoint Distribution System). Она предусматривала вещание в радиусе порядка 30 миль от передатчика в диапазоне 2,1–2,7 ГГц. Эксперимент не удался — стоимость оборудования оказалась неконкурентоспособной с кабельным и спутниковым ТВ.

Через четверть века на смену MDS пришли новые системы MMDS (Multichannel (или Microwave) Multipoint Distribution Service) — многоканальные (микроволновые) многоточечные распределительные системы. Они позволяют работать в диапазоне 2,5–2,7 ГГц в зоне радиусом до 35 миль (50–60 км — фактически зона прямой видимости). При этом мощность передатчика базовой станции, как правило, не превышает 100 Вт (до 1 кВт) — сравните с передатчиками 10–50 кВт на ретрансляторах обычных аналоговых телевизионных сетей. В оборудовании MMDS используется квадратурная амплитудная модуляция — от четырехпозиционной QPSK (2 бита на модуляционный символ) до 256-позиционной QAM (8 бит на символ). Поэтому скорости передачи данных в MMDS могут достигать очень больших значений при соответствующей ширине канала. Так, в стандартном телевизионном канале 8 МГц возможна скорость до 64 Мбит/с, но лишь теоретически. Отметим, что концепция систем MMDS разрабатывалась для трансляции телевизионных программ, поэтому она принципиально однонаправленная. В последние годы стали появляться сообщения об оборудовании для двунаправленной работы — производители стали добавлять обратный канал, как правило, с гораздо меньшей (в 4–8 раз) полосой пропускания, что позволяет использовать системы MMDS и для обмена данными (доступ в Интернет, интерактивное ТВ, мосты между локальными сетями и т. д.).

Однако сети MMDS так и не стали массовыми — к 2000 году в США они насчитывали около 20 тыс. абонентов [10]. Причина очевидна: оборудование

слишком дорого по сравнению с альтернативными решениями (кабельное телевидение, xDSL-модемы и т. д.). Следующим шагом в области систем широкополосного доступа стало появление концепции Local Multipoint Distribution Service (LMDS) — локальная многоточечная распределительная система. Ее испытания прошли в 1992 году в Нью-Йорке. В 1998 году FCC объявила о начале лицензирования LMDS. Первой реально действующей LMDS-системой стала сотовая телевизионная сеть компании Cellular Vision в Нью-Йорке в районе Брайтон-Бич (этот район Бруклина не был охвачен сетью кабельного телевидения).

Изначально LMDS предназначалась для работы в диапазоне 27,5–29,5 ГГц, однако рабочий диапазон зависит от распределения частот в конкретном регионе. В Европе появилась аналогичная система MVDS (Multipoint Video Distribution Systems), ориентированная на диапазон 40,5–42,5 ГГц (диапазон, выделенный в Европе для аналогового ТВ-вещания).

Системы LMDS/MVDS называют сотовым телевидением, поскольку радиус действия каждого ретранслятора невелик — порядка 3–8 км. В системах используют относительно маломощные передатчики — не более десятков ватт в групповых передатчиках (до 100–300 мВт на канал). Кроме того, в миллиметровом диапазоне затухание радиоволн весьма велико. Но, с другой стороны, сигналы в этом диапазоне отражаются от препятствий с малыми потерями, что можно эффективно использовать в условиях городской застройки, работая на переотраженных сигналах.

Системы LMDS/MVDS используют те же методы модуляции (QPSK, QAM) и частотные планы (19,5—39 МГц), что и системы спутникового ТВ-вещания. Диапазон их работы определяется наличием свободного частотного ресурса и может быть, например, 10, 24, 31, 38 ГГц. При общей ширине полосы 2 ГГц эти системы позволяют передавать от 96 до 128 аналоговых ТВ-каналов. Сотовая структура сетей LMDS/MVDS открывает широкие возможности для частотного планирования, включая такие механизмы, как различная поляризация сигналов, применение направленных (секторных) антенн, использование одних и тех же каналов в разных сотах и т. д. Важно отметить, что в современных трансляционных системах твердотельные выходные усилители усиливают сигнал только в активных каналах, а не во всей полосе сразу. Именно это и обусловливает относительно невысокую (сотни милливатт на канал) мощность излучения групповых передатчиков. Для сравнения — в Нью-Йорке при внедрении системы LMDS в районе Брайтон-Бич использовались выходные усилители на основе ламп бегущей волны мощностью свыше 100 Вт, усиливающие сразу весь рабочий диапазон. Сейчас там 17 базовых станций обеспечивают 48 аналоговых ТВ-каналов (NTSC) [11].

С 2000 года системы LMDS стали двунаправленными, что открыло перед ними широчайшие возможности для миграции из области ТВ-вещания в зону интерактивных приложений, важнейшее из которых Интернет, а также интерактивное телевидение (Video on Demand). Операторы соответствующих сервисов не замедлили этим воспользоваться. Однако широкому внедрению систем широкополосного доступа мешало отсутствие единого стандарта — аппаратура различных производителей оказывалась несовместимой, специализированная элементная база не могла стать массовой, соответственно цены оставались высокими. Названия MMDS и LMDS/MVDS фактически обозначали только тип

сервиса и самые основные функциональные возможности, а не методы практической реализации. Единой технической концепции (методы модуляции, механизмы доступа к каналам и т. п.) не было. Развитие этих систем резко остановилось, многие известные производители (например, корпорации Nortel, ADC) вообще объявили о прекращении выпуска соответствующего оборудования. Решением возникшей проблемы стал стандарт широкополосных региональных сетей IEEE 802.16.

7.1.2. Появление стандарта широкополосного доступа IEEE 802.16-2004

В августе 1998 года по инициативе Национальной испытательной лаборатории беспроводных электронных систем Национального института стандартов и технологии США (National Wireless Electronics Systems Testbed of the U.S. National Institute of Standards and Technology) комитет 802 IEEE организовал рабочую группу 802.16. С июля 1999 года группа приступила к регулярной работе над новым стандартом широкополосных городских (региональных) сетей передачи данных (MAN — Metropolian Access Network) с фиксированным доступом. Соответственно новый стандарт получил название WirelessMAN. Изначально деятельность велась в трех направлениях — разработка стандартов для диапазонов 10–66 ГГц (первоначально обозначался 802.16.1) и 2–11 ГГц (802.16.3), а также стандарта, регламентирующего совместную работу различных систем широкополосного беспроводного вещания (802.16.2).

Уже в декабре 2001 года был утвержден стандарт IEEE 802.16 «Air Interface for Fixed Broadband Wireless Access Systems» — «Воздушный интерфейс для фиксированных систем с широкополосным беспроводным доступом» (официально опубликован 8 апреля 2002 года). Он описывал общие принципы построения систем широкополосного беспроводного доступа и сосредоточивался на диапазоне 10–66 ГГц. Отметим, что 15 января 2003 года был опубликован документ IEEE 802.16c — поправки и дополнения к IEEE 802.16, касающиеся работы в диапазоне 10–66 ГГц. 10 сентября 2001 года увидел свет стандарт IEEE 802.16.2 «Coexistence of Fixed Broadband Wireless Access Systems» («Сосуществование фиксированных систем широкополосного беспроводного доступа»). Над более низкочастотным диапазоном работы продолжались чуть дольше — стандарт IEEE 802.16a «Medium Access Control Modifications and Additional Physical Layer Specifications for 2–11 GHz» («Модификации управления доступа к среде передачи и дополнительные спецификации физического уровня для диапазона 2–11 ГГц»), регламентирующий работу в диапазоне 2–11 ГГц, был утвержден 29 января 2003 года, а 1 апреля опубликован.

Разумеется, работать с тремя документами вместо одного неудобно. Кроме того, сразу же после публикации этих стандартов стали появляться многочисленные исправления и дополнения, только что выявленные на практике. Труд учитывать все эти поправки взяла на себя рабочая группа 802.16d IEEE. Непосредственно к работе по созданию единого документа с учетом всех поправок она приступила 11 сентября 2003 года (рабочее называние draft-версии того периода — 802.16REVd). Почти через год, 24 июня 2004 года, был официально утвержден новый стандарт — 802.16-2004, заменяющий собой документы 802.16-2001, 802.16c-2002 и 802.16a-2003.

В Европе Институт стандартизации в области телекоммуникаций ETSI принял континентальный эквивалент стандарта IEEE 802.16, именуемый ETSI HIPERMAN. Отметим, что HIPERMAN — это подмножество IEEE 802.16, он предусматривает работу в диапазоне 2–11 ГГц и только в одном из режимов — WirelessMAN-OFDM (см. далее).

Часто используют коммерческое имя стандарта IEEE 802.16 — WiMax (Worldwide Interoperability for Microwave Access). Произошло оно от названия международной организации WiMax Forum (www.wimaxforum.org), в которую входит ряд ведущих коммуникационных и полупроводниковых компаний (Airspan Networks, Alvarion Ltd, Aperto Networks, Fujitsu Microelectronics America, Intel, OFDM Forum, Proxim Corporation, Wi-LAN Inc и др.). Однако следует помнить, что на самом деле WiMax, равно как и европейский HIPERMAN, рассматривает только режим WirelessMAN-OFDM.

WiMax Forum был организован 11 апреля 2003 года. Его целью является содействие разработке беспроводного оборудования для доступа к широкополосным сетям, скорейшее развертывание сетей во всем мире и сертификация оборудования IEEE 802.16, а также подготовка спецификаций, призванных обеспечить совместимость оборудования разных производителей. Одна из целей WiMax — дальнейшее разделение труда на рынке производителей беспроводного оборудования. Поставщики элементной базы, такие, как Intel и Fujitsu, будут разрабатывать ее для всех производителей оборудования, а производители оборудования смогут сконцентрировать свои усилия на оборудовании со стандартной элементной базой. По данным аналитиков, члены WiMax Forum представляют собой более 75% рынка производителей оборудования широкополосного беспроводного доступа. WiMax Forum начал сертификацию оборудования в июле 2005 года. В качестве сертификационной лаборатории уже выбрана испанская компания Cetecom. По условиям WiMax сертификация должна проводиться с проверкой совместимости оборудования трех производителей с одинаковыми профилями.

7.2. Общие принципы IEEE 802.16–2004

Разработанный IEEE стандарт IEEE 802.16-2004 [4] представляет собой рассчитанную на внедрение в городских распределенных (региональных) беспроводных сетях (WirelessMAN) технологию беспроводного широкополосного доступа операторского класса. В последнем — его основное отличие от группы стандартов IEEE 802.11, ориентированных на работу в безлицензионном диапазоне.

При создании стандарта были существенно изменены основополагающие принципы, заложенные в беспроводные системы на предыдущих этапах. Первостепенное значение приобрело оптимальное использование спектрального ресурса радиоканала при любых соотношениях «скорость – помехоустойчивость», а также необходимость обеспечивать заданный уровень качества обслуживания (QoS) любому абоненту сети.

Стандарт IEEE 802.16-2004 описывает принципы построения сетей регионального масштаба в диапазонах до 66 ГГц, точнее, их физический и MAC-уровни (радиоинтерфейсы, методы модуляции и доступа к каналам, системы управления потоками, структуры передаваемых данных, механизмы сопряжения про-

токолов передачи данных верхних уровней (прежде всего ATM и IP) с протоколами физического уровня IEEE 802.16 и др.). Стандарт предусматривает пять режимов организации работы сети (табл. 7.1). Только один из них — WirelessMAN-SC — предназначен для диапазона 10–66 ГГц. Он ориентирован на магистральные сети («точка-точка», «точка-многоточка»), работающие в режиме прямой видимости (так как затухание столь высокочастотных сигналов при отражении очень велико), с типичными скоростями потока данных (bit stream) 120 Мбит/с и шириной канала порядка 25 МГц.

Таблица 7.1. Основные режимы в стандарте IEEE 802.16-2004

Режим	Частотный диапазон	Опции	Метод дуплексирования
WirelessMAN-SC	10–66 ГГц		TDD/FDD
WirelessMAN-SCa	< 11 ГГц	AAS/ARQ/STC/256-QAM	TDD/FDD
WirelessMAN-OFDM	< 11 ГГц	AAS/ARQ/STC/Mesh	TDD/FDD
WirelessMAN-OFDMA	< 11 ГГц	AAS/ARQ/STC	TDD/FDD
WirelessHUMAN	< 11 ГГц*	DFS/AAS/ARQ/Mesh/STC	TDD

* Безлицензионный диапазон (в США и Европе).

Остальные режимы разработаны для диапазонов менее 11 ГГц. Один из них — WirelessMAN-SCa — это «низкочастотная» вариация WirelessMAN-SC (с рядом дополнительных механизмов). Два других режима — WirelessMAN-OFDM и WirelessMAN-OFDMA — это совсем новые методы, ранее входившие в утвержденный в 2003 году. стандарт IEEE 802.16a, но с тех пор претерпевшие ряд изменений.

В стандарте IEEE 802.16-2004 предусмотрен и режим работы в безлицензионном (в США) диапазоне WirelessHUMAN (High-speed Unlicensed Metropolitan Area Network). Фактически речь идет об адаптации методов WirelessMAN-OFDM и WirelessMAN-OFDMA для работы в диапазоне 5–6 ГГц. Основные отличия WirelessHUMAN — это использование только временного дуплексирования, режим динамического распределения частот (DFS — dynamic frequency selection) и механизм сквозной нумерации частотных каналов. Однако, поскольку в России безлицензионных диапазонов в гигагерцевой области нет и ничего подобного нам не грозит, подробно останавливаться на данном режиме не будем.

Отметим, что все режимы диапазона ниже 11 ГГц отличает три характерные детали — механизмы автоматического запроса повторной передачи (ARQ — automatic repeat request), поддержка работы с адаптивными антенными системами (AAS — adaptive antenna system) и пространственно-временное кодирование (STC — space time coding). Кроме того, помимо централизованной архитектуры «точка-многоточка», в режиме WirelessMAN-OFDM предусмотрена поддержка архитектуры mesh-сети («сетки» — децентрализованной сети взаимодействующих друг с другом систем, узлы которой не только обеспечивают доступ к среде передачи, но и поддерживают ретрансляцию трафика). Примечательно, что если в документе IEEE 802.16a речь шла о диапазоне 2–11 ГГц, то в стандарте 802.16-2004 нижняя граница так четко не оговаривается (упоминается, «как правило, не ниже 1 ГГц»). Хотя по информации ряда аналитиков известно, что

всерьез рассматривается возможность использования диапазона 700 МГц для нового стандарта для мобильных абонентов IEEE 802.16e.

7.3. MAC-уровень стандарта IEEE 802.16

7.3.1. Структура MAC-уровня

Задачи непосредственной доставки потоков данных между БС и абонентскими станциями решаются на физическом уровне стандарта IEEE 802.16. Функции же, связанные с формированием структур этих данных, а также управлением работой системы IEEE 802.16, реализуются на MAC-уровне. Оборудование стандарта IEEE 802.16 призвано формировать транспортную среду для различных приложений (сервисов), поэтому первая задача, решаемая в IEEE 802.16 — это механизм поддержки разнообразных сервисов верхнего уровня. Разработчики стандарта стремились создать единый для всех приложений протокол MAC-уровня, независимо от особенностей физического канала. Это существенно упрощает связь терминалов конечных пользователей с городской сетью передачи данных — физически среды передачи в разных фрагментах WMAN могут быть различны, но структура данных едина. В одном канале могут работать (не единовременно) сотни различных терминалов еще большего числа конечных пользователей. Этим пользователям необходимы самые разные сервисы (приложения) — потоки голоса и данных с временным разделением, соединения по протоколу IP, пакетная передача речи через IP (VoIP) и т. п. Более того, заданное качество услуг (QoS) каждого отдельного сервиса не должно изменяться при работе через сети IEEE 802.16. Алгоритмы и механизмы доступа MAC-уровня должны уверенно решать все эти задачи.

Структурно MAC-уровень стандарта IEEE 802.16 подразделяется на три подуровня (рис. 7.1) — подуровень преобразования сервиса CS (Convergence Sublayer), основной подуровень CPS (Common Part Sublayer) и подуровень защиты PS (Privacy Sublayer). На подуровне защиты реализуются функции криптозащиты данных и механизмы аутентификации/предотвращения несанкционированного доступа. Для этого предусмотрены два основных компонента — набор алгоритмов криптозащиты и протокол управления ключом шифрования. Ключ каждой абонентской станции (АС) базовая станция (БС) может передавать в процессе авторизации, используя схему работы «клиент (АС) — сервер (БС)».

На подуровне преобразования сервиса происходит трансформация потоков данных протоколов верхних уровней для передачи через сети IEEE 802.16. Для каждого типа приложений верхних уровней стандарт предусматривает свой механизм преобразования. На сегодня описаны и вошли в спецификацию IEEE 802.16 два основных типа сервисных потоков — ATM и пакетная передача. Последняя подразумевает достаточно широкий набор различных протоколов — IP, Ethernet (IEEE 802.3), виртуальные ЛВС (VLAN, IEEE 802.1Q-1998). Цель работы на CS-подуровне — оптимизация передаваемых потоков данных каждого приложения верхнего уровня с учетом их специфики. Поэтому важнейшая задача, решаемая на данном подуровне, — задача классификации пакетов/ячеек. От ее результатов зависит и оптимизация передаваемых потоков, и выделение полосы пропускания для каждого из них.

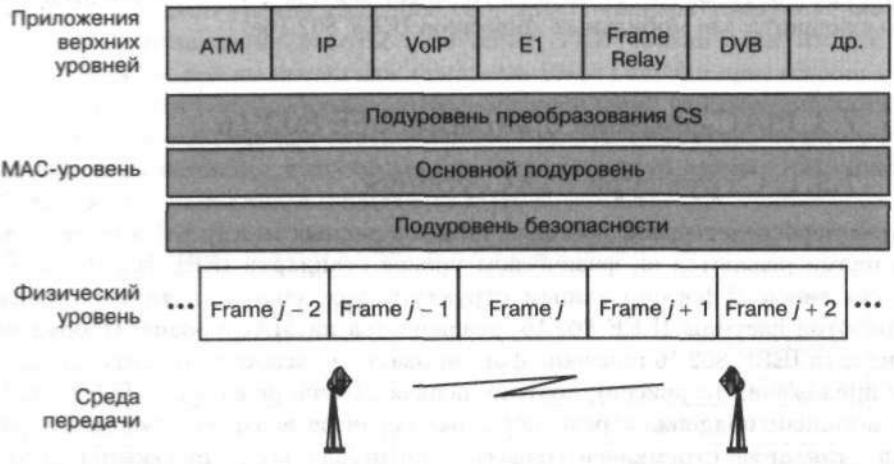


Рис. 7.1. MAC-уровень в стандарте IEEE 802.16

Для оптимизации транслируемых потоков предусмотрен специальный механизм удаления повторяющихся фрагментов заголовков PHS (Payload Header Suppression). Действительно, и в ATM, и в пакетном режиме данные передаются отдельными порциями — ячейками и пакетами. Каждая такая порция данных состоит, в общем случае, из заголовка и поля данных — фиксированных размеров для ячеек ATM (5 и 48 байт соответственно) и достаточно произвольных при пакетной передаче. Во многих случаях заголовки пакетов и ячеек содержат повторяющуюся информацию, излишнюю при трансляции посредством протокола IEEE 802.16. Механизм PHS позволяет избавиться от передачи избыточной информации: на передающем конце пакеты приложений в соответствии с определенными правилами преобразуются в структуры данных MAC-уровня IEEE 802.16, на приемном — восстанавливаются.

7.3.2. Соединения и сервисные потоки

Ключевой момент в стандарте IEEE 802.16 — это понятие сервисного потока и связанные с ним понятия «соединение» и «идентификатор соединения» (CID). Поскольку система IEEE 802.16 — лишь транспортная среда, ее инфраструктура фактически формирует коммуникационные каналы для потоков данных различных приложений верхних уровней (сервисов), таких, как передача видеоданных, ATM-потоки, IP-потоки, передача телефонных мультиплексированных пакетов типа E1 и т. д. Каждое из таких приложений обладает своими требованиями к скорости передачи, надежности (качеству обслуживания, QoS), криптозащите и т.д.

Сервисным потоком в стандарте IEEE 802.16 называется поток данных, связанный с определенным приложением. Сервисный поток характеризуется набором требований к каналу передачи информации — времени задержки символов, уровню флуктуаций задержек (джиттеру) и гарантированной пропускной способности. Каждому сервисному потоку в сети присваивается идентификатор SFID (32 разряда), основываясь на котором БС (а в ряде случаев и АС) определяют необходимые параметры соединения, связанного с конкретным данным сервисным потоком. Для общей стандартизации работы в сети используется

понятие сервисного класса — устойчивого набора параметров для стандартных приложений, например для трансляции телефонных каналов E1. Параметры сервисного потока можно задать, просто указав его принадлежность к определенному сервисному классу.

В терминологии IEEE 802.16 соединение — это установление логической связи на MAC-уровнях на передающей и приемной стороне для передачи сервисного потока. Каждому соединению присваивается 16-разрядный идентификатор CID, с которым однозначно связаны тип и характеристики соединения. В частности, по запросу предоставления/изменения полосы пропускания со стороны AC базовая станция сразу понимает, с каким сервисным потоком имеет дело и какие условия передачи ему нужно обеспечить. Так, при начальной инициализации в сети каждой AC назначается три CID для служебных сообщений трех уровней. Принципиально, что одна AC может устанавливать множество различных соединений с различными CID. Характерен пример, когда связь крупного офиса с телекоммуникационным узлом организована через систему IEEE 802.16. В этом случае одна AC в офисе может поддерживать совершенно разные приложения — телефонию, телевидение, доступ в Интернет и в распределенную корпоративную сеть и т. д. Каждое из этих приложений предъявляет свои требования к QoS и скорости передачи, которые нужно удовлетворить. Посредством CID базовая станция узнает, с чем имеет дело, и предоставляет необходимый ресурс.

7.3.3. Пакеты MAC-уровня

Весь поток данных в сетях IEEE 802.16 — это поток пакетов. На основном подуровне MAC формируются пакеты данных (MAC PDU — MAC Protocol Data Unit, блоки данных MAC-уровня), которые затем передаются на физический уровень, инкапсулируются в физические пакеты и транслируются через канал связи. Пакет MAC PDU (далее — PDU) включает заголовок и поле данных (его может и не быть), за которым может следовать контрольная сумма CRC (рис. 7.2). Заголовок PDU занимает 6 байт и может быть двух типов — общий и заголовок запроса полосы пропускания. Общий заголовок используется в пакетах, у которых присутствует поле данных. В общем заголовке указывается идентификатор соединения (CID), тип и контрольная сумма заголовка, а также приводится информация о поле данных (табл. 7.2).

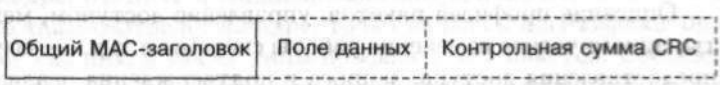


Рис. 7.2. Пакет MAC-уровня IEEE 802.16

Заголовок запроса полосы (также 6 байт) применяется, когда AC просит у БС выделить или увеличить ей полосу пропускания в восходящем канале. При этом в заголовке указывается CID и размер требуемой полосы (в байтах, без учета заголовков физических пакетов). Поля данных после заголовков запроса полосы нет.

Поле данных может содержать:

- подзаголовки MAC;
- управляющие сообщения;

— собственно данные приложений верхних уровней, преобразованные на CS-подуровне.

Таблица 7.2. Структура общего заголовка MAC PDU (от старшего к младшим битам)

Поле	Длина, бит
Тип заголовка = 0	1
Признак цифрования поля данных	1
Тип подзаголовков	6
Не используется	1
Признак наличия CRC	1
Индекс ключа шифрования	2
Не используется	1
Длина пакета, включая заголовок (байт)	11
Идентификатор соединения CID	16
Контрольная сумма заголовка (задающий полином $g(D) = D^8 + D^2 + D + 1$)	8

MAC-подзаголовки могут быть пяти типов: упаковки, фрагментации, управления предоставлением канала, а также подзаголовки mesh-сети и подзаголовок канала быстрой обратной связи Fast Feedback.

Подзаголовок упаковки используется, если в поле данных одного PDU содержится несколько пакетов верхних уровней; *подзаголовок фрагментирования* — если, напротив, один пакет верхнего уровня разбит на несколько PDU.

Подзаголовок управления предоставлением доступа предназначен для того, чтобы АС сообщала БС изменение своих потребностей в полосе пропускания (число байтов в восходящем канале для определенного соединения, сообщение о переполнении выходной очереди в АС, требование регулярного опроса со стороны БС для выяснения необходимой полосы).

Подзаголовок Fast Feedback служит для назначения конкретной АС интервала для быстрого ответа на запрос БС (например, в ходе процедур измерения характеристик канала).

Управляющие сообщения — это основной механизм управления системой IEEE 802.16. Всего зарезервировано 256 типов управляющих сообщений, из них используется 48. Формат управляющих сообщений прост — поле типа сообщения (1 байт) и поле данных (параметров) произвольной длины.

Описание профилей пакетов, управление доступом, механизмы криптозащиты, динамическое изменение работы системы и т. д. — все функции управления, предоставления доступа, запроса и подтверждения реализуются через управляющие сообщения. Рассмотренные ниже карты входящего/нисходящего каналов (UL-/DL-MAP) также являются управляющими сообщениями.

7.3.4. Общая структура кадров IEEE 802.16

Для понимания принципов управления соединениями и потоками данных в сетях IEEE 802.16 рассмотрим общие принципы передачи в физическом канале.

Передача данных на физическом уровне происходит посредством непрерывной последовательности кадров фиксированной длительности. Кадр состоит из двух субкадров — для нисходящего потока (от базовой станции к АС) и для

восходящего (от АС к БС). Дуплексный механизм предусматривает как частотное (FDD — frequency division duplex), так и временное (TDD — time division duplex) разделение восходящего и нисходящего субкадров.

При временном дуплексировании каналов кадры передаются в одном частотном диапазоне, сначала нисходящий (DL), затем восходящий (UL). При частотном дуплексировании восходящий и нисходящий субкадры транслируются одновременно, но с частотным разносом.

DL-субкадр начинается с синхронизирующей последовательности (преамбулы), за которой следует управляющая секция с набором широковещательных (предназначенных всем) служебных сообщений. За ними передаются пакеты физического уровня, содержащие как служебные сообщения, так и данные для различных АС.

Пакеты в нисходящем субкадре следуют друг за другом без интервалов. Чтобы абонентские станции могли отличить один пакет от другого, в управляющей секции передаются карты нисходящего (DL-MAP) и восходящего (UL-MAP) каналов. В карте нисходящего канала указана длительность и номер кадра, идентификационный номер базовой станции, номер последнего переданного дескриптора нисходящего канала (в котором описаны профили пакетов текущего кадра), а также точка начала и тип профиля каждого пакета. Точка начала отсчитывается в так называемых физических слотах, размер которых различен для каждого из режимов IEEE 802.16. Например, в режимах SC и SCa физический слот равен четырем модуляционным символам.

Профиль пакета — это список его параметров, включая метод модуляции, тип FEC-кодирования (с параметрами схем кодирования), а также диапазон значения отношения сигнал/шум в приемном канале конкретной станции, при котором данный профиль может применяться. Каждому профилю пакета в нисходящем/восходящем канале присваивается идентификационный код DIUC/UIUC (Downlink/Uplink Interval Usage Code), который и используется в карте нисходящего канала при распределении ресурсов. Список профилей в виде специальных управляющих сообщений — дескрипторов нисходящего и восходящего каналов (DCD/UCD) — транслируется базовой станцией с периодом не менее 10 с.

В восходящем субкадре для каждой передающей АС базовая станция резервирует специальные временные интервалы (или частотно-временные для режимов OFDMA и частично OFDM) — тайм-слоты. Информация о распределении тайм-слотов между АС записывается в карте восходящего канала UL-MAP, транслируемого в каждом кадре. UL-MAP функционально аналогична DL-MAP — в ней сообщается, сколько тайм-слотов в субкадре, точка начала и идентификатор соединения для каждого из них, а также типы профилей всех пакетов. Сообщение UL-MAP текущего кадра может относиться как к данному кадру, так и к последующему.

Кроме назначенных БС тайм-слотов для определенных АС в UL-канале предусмотрены интервалы конкурентного доступа, в течение которых АС может передать сообщение для первичной регистрации в сети или для запроса канала/изменения полосы пропускания предоставленного канала.

7.3.5. Принцип предоставления канальных ресурсов

Основной принцип предоставления доступа к каналу в стандарте IEEE 802.16 — это доступ по запросу Demand Assigned Multiple Access (DAMA). Ни одна АС

не может ничего передавать, кроме запросов на регистрацию и предоставление канала, пока БС не разрешит ей этого, т. е. отведет временной интервал в восходящем канале и укажет его расположение в карте UL-MAP. Абонентская станция может как запрашивать определенный размер полосы в канале, так и просить об изменении уже предоставленного ей канального ресурса.

Стандарт IEEE 802.16 предусматривает два режима предоставления доступа — для каждого отдельного соединения и для всех соединений определенной АС. Очевидно, что первый механизм обеспечивает большую гибкость, однако второй существенно сокращает объем служебных сообщений и требует меньшей производительности от аппаратуры.

Запросы полосы могут быть как спорадическими для БС, так и планированными. В первом случае запросы реализуются посредством пакетов, состоящих из заголовка запроса, передаваемых на конкурентной основе в специально выделенном для них интервале восходящего канала. Поскольку эти запросы спонтанны, в данных интервалах возможны коллизии, вызванные одновременной работой передатчиков двух и более АС.

Принцип борьбы с коллизиями аналогичен используемому в стандарте IEEE 802.11: после того как АС решила, что ей нужно зарегистрироваться/запросить канал, она не начинает трансляцию в первом же предназначенном для этого интервале. В АС есть генератор случайных чисел (ГСЧ), выбирающий значения из некоего диапазона от 0 до $2^n - 1$. Так, если $n = 4$, ГСЧ выбирает числа в диапазоне 0–15, например 11. Далее АС отсчитывает 11 интервалов, предназначенных для регистрации/запроса канала, и только в 12-м выходит в эфир. Если передача прошла успешно и БС приняла запрос, она в определенный период ответит специальным сообщением. В противном случае АС считает попытку неудачной и повторяет процедуру, только интервал выбора для ГСЧ удваивается. Такая последовательность действий продолжается до тех пор, пока не будет получен ответ от БС. Максимальный размер диапазона возможных значений ГСЧ ограничен — при его достижении он вновь принимает минимальное значение.

Процедура плановых запросов полосы в восходящем канале называется опросом (polling) — БС как бы запрашивает у АС их потребности в канальных ресурсах. Реально это означает, что базовая станция предоставляет конкретной АС интервал для передачи запроса о предоставлении/изменении полосы, т. е. никакой конкуренции уже нет.

Опрос может быть в «реальном времени» — интервалы для запроса предоставляются АС с тем же периодом, с каким у нее может возникнуть потребность в изменении условий доступа (например, в каждом кадре). Этот режим удобен для приложений, когда пакеты данных следуют с фиксированным периодом, но их размер не стабилен (например, видео-MPEG). Другой вариант опроса — вне «реального времени». В этом случае БС предоставляет АС интервал для запроса также периодически, но период этот существенно больше, например, 1 с. Характерное приложение, для которого эффективен этот механизм, — FTP-протокол.

Для приложений, у которых периодичность и размер пакетов фиксированы (например, в телефонии шина E1), предусмотрен механизм доступа к каналу без требования (Unsolicited Grant Service — UGS). В этом случае БС предоставляет АС для передачи данных интервалы фиксированного размера с заданным периодом, соответствующим скорости потока данных.

Если в ходе работы АС нужно изменить условия доступа, она делает это посредством специального MAC-подзаголовка управления предоставлением канала. В этом подзаголовке есть специальный флаг «опроси меня», установив который АС просит у БС интервал для запроса новой полосы. Кроме того, в этом подзаголовке АС может непосредственно запросить у БС дополнительную прибавку (в байтах) к уже предоставленной полосе для конкретного соединения (так называемый PiggyBack Request). В подзаголовке предоставления канала также есть специальный бит индикации переполнения входного буфера передатчика АС, что приводит к потере данных (slip). БС может отреагировать на появление этого сигнала, например, увеличив полосу для данной АС.

7.3.6. Подтверждение приема (ARQ) и быстрая обратная связь

Пожалуй, наиболее существенные особенности MAC-уровня для диапазона ниже 11 ГГЦ — это процедура подтверждения приема пакетов и их повторной отправки (ARQ), а также функция быстрой организации канала обратной связи. Механизм ARQ достаточно хорошо известен и основан на способности корректирующих кодов обнаруживать ошибки передачи. Кратко напомним суть этого механизма. Каждый пакет при передаче кодируется линейным циклическим корректирующим кодом с достаточным числом проверочных символов. На приемной стороне после декодирования проверяется безошибочность принятого кодового слова. Для этого вычисляется так называемый синдром — вектор, равный произведению вектора принятого сигнала и присущей коду матрицы проверочных уравнений. Если синдром нулевой — слово кодовое (т. е. не содержит ошибок), если ненулевой — некодовое. Даже в очень плохом канале с вероятностью ошибки, стремящейся к 0,5, вероятность ошибки декодирования можно сделать сколь угодно малой, выбирая достаточное число проверочных символов, правда, ценой существенного снижения скорости передачи.

Получив пакет, приемник обязан отправить передатчику квитанцию подтверждения (значение синдрома, ноль или нет). Все пакеты нумеруются, поэтому приемник сразу обнаруживает сбой в получении определенного пакета и сообщает об этом передатчику, который повторно отправляет пакет, принятый с неустранимыми ошибками. Режим ARQ — опциональный и назначается конкретному соединению. Причем в рамках одного соединения весь трафик может передаваться либо с применением ARQ, либо без него, смешение недопустимо. Очевидно, что режим ARQ требует эффективного быстрого обратного канала для подтверждения приема. Для этого предназначено специальное сообщение, о наличии которого в пакете сигнализирует специальный признак в основном MAC-заголовке пакета.

Быстрого и гарантированного ответа со стороны АС, помимо ARQ, требуют и ряд других механизмов стандарта. Быстрая обратная связь необходима для периодического измерения параметров канала, запросов физических параметров каналов, коррекции мощности передатчиков и др. Особое значение эти процедуры приобретают при работе с адаптивными многолучевыми антенными системами, где важно определить физические параметры канала связи в каждом луче (секторе). Данные процедуры заключаются в отправке специальных сообщений и получении данных об условиях их прохождения через канал.

В разных режимах (SCa, OFDM, OFDMA) названия и тип подобных сообщений несколько различны, однако сущность от этого не меняется. Поэтому в стандарте IEEE 802.16 предусмотрен набор процедур для обеспечения быстрой обратной связи.

Например, для приоритетной передачи таких данных абонентскими станциями в режиме OFDMA предусмотрен специальный тип сообщений быстрой обратной связи FAST FEEDBACK. Для них выделяется специальная область (канал) в восходящем субкадре. Сами сообщения обратной связи представляют собой 4-разрядные двоичные числа, расположение которых в канале обратной связи для каждой АС базовая станция задает в специальном подзаголовке FAST FEEDBACK. В этом же подзаголовке указывается тип (назначение) сообщения — измерение мощности в нисходящем канале, коэффициенты антенн БС в режиме AAS и др. Предусмотрены и специальные меры для успешной доставки этих 4-разрядных сообщений с 24-кратной избыточностью: каждое сообщение (4 бита) занимает один OFDMA-слот (3 символа × 1 субканал, 6 фрагментов, 48 информационных несущих — см. раздел 11.6), т. е. 4-битной последовательности поставлен в соответствие фиксированный набор всех 48 информационных несущих в слоте. Модуляция в FAST FEEDBACK-канале — QPSK, 2 бита на символ, соответственно на каждый информационный бит сообщения приходится 24 транслируемых бита. Данная избыточность в определенной мере гарантирует достоверный прием сообщения.

7.4. Физический уровень стандарта IEEE 802.16. Режим WirelessMAN-SC

На физическом уровне стандарт IEEE 802.16 предусматривает три принципиально различных метода передачи данных — метод модуляции одной несущей (SC — single carrier, в диапазоне ниже 11 ГГц — SCa), метод модуляции посредством ортогональных несущих OFDM (orthogonal frequency division multiplexing) и метод мультиплексирования (множественного доступа) посредством ортогональных несущих OFDMA (orthogonal frequency division multiple access) (см. табл. 7.1). Отметим, что режим SCa отличается от своего более высокочастотного собрата SC прежде всего методами помехоустойчивого кодирования и модуляции (допускается 256-уровневая квадратурная модуляция 256-QAM).

Метод WirelessMAN-SC стандарта IEEE 802.16 описывает работу в диапазоне 10–66 ГГц сетей с архитектурой «точка-многоточка» (из центра — многим). Это двунаправленная система, т. е. предусмотрены нисходящий и восходящий потоки. При этом каналы подразумеваются широкополосные (до 25–28 МГц), а скорости передачи — высокие (например, 120 Мбит/с).

7.4.1. Канальное кодирование

Тракт обработки данных и формирования выходного сигнала для передачи через радиоканал в стандарте IEEE 802.16 достаточно обычен для современных телекоммуникационных протоколов (рис. 7.3) и практически одинаков для восходящих и нисходящих соединений.

Рис. 7.3. Тракт преобразования сигнала в режиме WirelessMAN-SC

Входной поток данных скремблируется — подвергается рандомизации, т. е. умножению на псевдослучайную последовательность (ПСП), получаемую в 15-разрядном сдвиговом регистре (рис. 7.4) (задающий полином ПСП — $c(x) = x^{15}+x^{14}+1$, начальное значение — $4A80_{16}$). Далее скремблированные данные защищают посредством помехоустойчивых кодов (FEC-кодирование). При этом можно использовать одну из четырех схем кодирования: код Рида – Соломона с символами из поля Галуа GF(256), каскадный код с внешним кодом Рида – Соломона и внутренним сверточным кодом с кодовым ограничением $K = 7$ (скорость кодирования — 2/3) с декодированием по алгоритму Витерби, каскадный код с внешним кодом Рида – Соломона и внутренним кодом с проверкой на четность (8, 6, 2) и блоковый турбокод. Размер кодируемого информационного блока и число избыточных байтов не фиксированы — эти параметры можно задавать в зависимости от условий среды передачи и требований QoS. Так, для кода Рида – Соломона размер исходного блока данных может быть от 6 до 255 байт, а число избыточных байтов — до 32 (всего до 255 байт). Первые два алгоритма кодирования обязательны для всех устройств стандарта, остальные два алгоритма — опциональны.

Рис. 7.4. Генерация ПСП для рандомизации данных

Метод WirelessMAN-SC, как и следует из его названия, предусматривает схему с модуляцией одной несущей в каждом частотном канале. Допускается три типа квадратурной амплитудной модуляции: 4-позиционная QPSK и 16-позиционная 16-QAM (обязательны для всех устройств), а также 64-QAM (опцио-

нально). Кодированные блоки данных преобразуются в модуляционные символы (каждые 2/4/6 бит определяют один символ QPSK/16-QAM /64-QAM) в соответствии с приведенными в стандарте таблицами — каждой группе из 2/4/6 бит ставится в соответствие синфазная (I) и квадратурная (Q) координаты. Далее последовательность дискретных значений в каналах I и Q преобразуется посредством так называемого синусквадратного фильтра (square-root raised cosine filter) в непрерывные (сглаженные) сигналы. Передаточная функция синусквадратного фильтра записывается как

$$H(f) = \begin{cases} 1, & |f| < f_N(1-\alpha), \\ \sqrt{\frac{1}{2} + \frac{1}{2}\sin\left[\frac{\pi}{2f_N}\left(\frac{f_N - |f|}{\alpha}\right)\right]}, & f_N(1-\alpha) \leqslant |f| \leqslant f_N(1+\alpha), \\ 0, & |f| > f_N(1+\alpha), \end{cases}$$

где α — коэффициент избирательности (по стандарту IEEE 802.16 $\alpha = 0{,}25$); f_N — частота Найквиста, равная половине частоты дискретизации.

Фильтрованные потоки $I(t)$ и $Q(t)$ поступают непосредственно в квадратурный модулятор, где формируется выходной сигнал

$$S(t) = I(t)\cos(2\pi f_c t) - Q(t)\sin(2\pi f_c t),$$

где f_c — несущая частота. Далее сигнал усиливается и передается в эфир. На приемной стороне все происходит в обратном порядке. В результате в зависимости от ширины канала и метода модуляции формируется достаточно широкий набор скоростей потока данных (табл. 7.3).

Таблица 7.3. Скорость физического потока данных в зависимости от вида модуляции и ширины канала

Ширина канала, МГц	Скорость символов, Мбод	*Скорость физического потока данных, Мбит/с*		
		QPSK	16-QAM	64-QAM
20	16	32	64	96
25	20	40	80	120
28	22,4	44,8	89,6	134,4

7.4.2. Структура кадров

Передача данных на физическом уровне происходит посредством непрерывной последовательности кадров. Каждый кадр имеет фиксированную длительность — 0,5; 1 и 2 мс, поэтому его информационная емкость зависит от символьной скорости и метода модуляции. Кадр состоит из преамбулы (синхропоследовательности длиной 32 QPSK-символа), управляющей секции и последовательности пакетов с данными (рис. 7.5). Управляющая секция может содержать карты DL/UL-MAP и дескрипторы нисходящего/восходящего каналов. Сообщения этой секции всегда передаются посредством QPSK.

Поскольку определяемая стандартом IEEE 802.16 система двунаправленная, необходим дуплексный механизм. Он предусматривает как частотное (FDD — frequency division duplex), так и временное (TDD — time division duplex) разделение восходящего и нисходящего каналов.

При временном дуплексировании каналов кадр делится на нисходящий и восходящий субкадры (их соотношение в кадре может гибко изменяться в процессе работы в зависимости от необходимой полосы пропускания для нисходящих и восходящих каналов), разделенные специальным интервалом (рис. 7.5, *а*). При частотном дуплексировании восходящий и нисходящий каналы транслируются каждый на своей несущей (рис. 7.5, *б*).

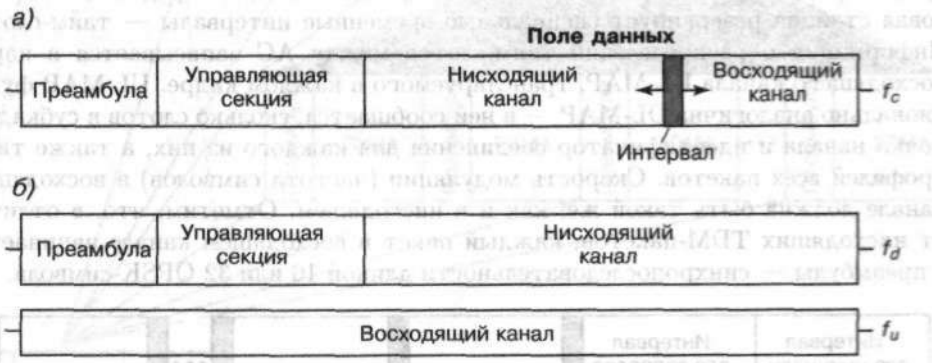


Рис. 7.5. Структура кадра в стандарте IEEE 802.16 для систем с временным *(а)* и частотным *(б)* дуплексированием каналов

В нисходящем канале информация от базовой станции передается в виде последовательности пакетов (метод временного мультиплексирования TDM — time division multiplex) (рис. 7.6). Для каждого пакета можно задавать метод модуляции и схему кодирования данных, т. е. выбирать между скоростью и надежностью передачи. Данные о параметрах пакета, его длине, моменте начала передачи, а также о его принадлежности к определенному соединению (соответственно об адресации определенной АС) содержатся в карте нисходящего канала DL-MAP. Точка начала отсчитывается в так называемых физических слотах (один физический слот равен четырем модуляционным символам).

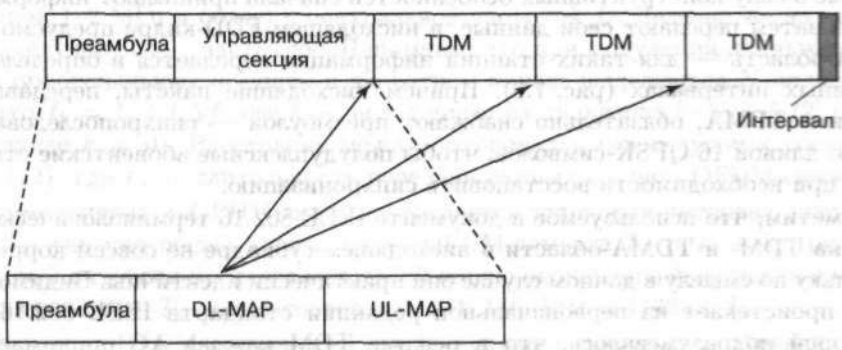


Рис. 7.6. Структура нисходящего канала

В нисходящем субкадре пакеты выстраиваются в очередь в зависимости от типа их кодирования и модуляции так, что самые помехозащищенные передаются первыми (управляющая секция всегда передается посредством QPSK-модуляции). Если этого не сделать, абонентские станции с плохими условиями при-

ема, которым предназначаются наиболее защищенные пакеты, могут потерять синхронизацию в ожидании своей порции информации. Пакеты в нисходящем субкадре следуют друг за другом без интервалов и предваряющих их заголовков и идентифицируются абонентскими станциями на основе информации в DL-MAP.

Абонентские станции получают доступ к среде передачи посредством механизма временного разделения каналов (TDMA — time division multiple access) (рис. 7.7). Для этого в восходящем субкадре для каждой передающей АС базовая станция резервирует специальные временные интервалы — тайм-слоты. Информация о распределении тайм-слотов между АС записывается в карте восходящего канала UL-MAP, транслируемого в каждом кадре. UL-MAP функционально аналогична DL-MAP — в ней сообщается, сколько слотов в субкадре, точка начала и идентификатор соединения для каждого из них, а также типы профилей всех пакетов. Скорость модуляции (частота символов) в восходящем канале должна быть такой же, как и в нисходящем. Отметим, что, в отличие от нисходящих TDM-пакетов, каждый пакет в восходящем канале начинается с преамбулы — синхропоследовательности длиной 16 или 32 QPSK-символа.

Рис. 7.7. Структура восходящего канала

В восходящем канале также предусмотрены интервалы конкурентного доступа для первичной регистрации в сети или для запроса канала/изменения полосы пропускания канала.

Стандарт IEEE 802.16 допускает применение как дуплексных, так и полудуплексных абонентских станций. Последние не способны одновременно принимать и передавать информацию. В режиме FDD для полудуплексных АС, которые в силу конструктивных особенностей сначала принимают информацию и лишь затем передают свои данные, в нисходящем FDD-кадре предусмотрена TDMA-область — для таких станций информация передается в определенных временных интервалах (рис. 7.8). Причем нисходящие пакеты, передаваемые в режиме TDMA, обязательно снабжают преамбулой — синхропоследовательностью длиной 16 QPSK-символов, чтобы полудуплексные абонентские станции могли при необходимости восстановить синхронизацию.

Отметим, что используемое в документе IEEE 802.16 терминологическое деление на TDM- и TDMA-области в нисходящем субкадре не совсем корректно, поскольку по смыслу в данном случае они практически идентичны. Видимо, проблема проистекает из первоначальной редакции стандарта IEEE 802.16-2001, в которой подразумевалось, что в режиме TDM каждая АС принимает все пакеты, декодирует их заголовки и по MAC-адресам распознает «свои». В новой редакции указано, что распределение пакетов в DL-субкадре приведено в DL-MAP (хотя не исключается и «старая» схема). Единственное различие между TDM и TDMA — если при TDM пакеты следуют один за другим, без интервалов, то в зоне TDMA между отдельными пакетами с преамбулами возможны незаполненные интервалы, в течение которых передатчик БС «молчит».

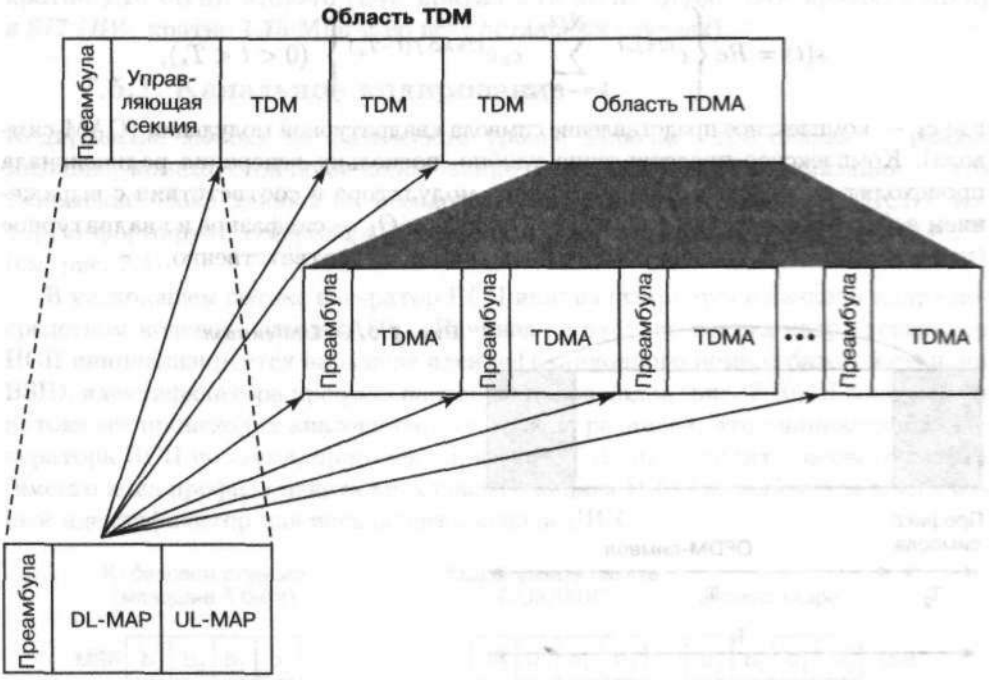


Рис. 7.8. Нисходящий канал в случае FDD при работе с полудуплексными абонентскими станциями

7.5. Режим WirelessMAN-OFDM

Режим OFDM — это метод модуляции потока данных в одном частотном канале (шириной 1–2 МГц и более) с центральной частотой f_c. Деление же на каналы, как и в случае SC, — частотное. Напомним, что при модуляции данных посредством ортогональных несущих в частотном канале выделяются N поднесущих так, что $f_k = f_c + k\Delta f$, где k — целое число из диапазона $[-N/2, N/2]$ (в данном случае $k \neq 0$). Расстояние между ортогональными несущими составляет $\Delta f = 1/T_b$, где T_b — длительность передачи данных в одном OFDM-символе .

Кроме данных, в OFDM-символе передается защитный интервал длительностью T_g, так что общая длительность OFDM-символа $T_s = T_b + T_g$ (рис. 7.9). Защитный интервал представляет собой копию оконечного фрагмента символа. Его длительность T_g может составлять 1/4, 1/8, 1/16 и 1/32 от T_b.

Модуляция OFDM основана на двух основных принципах: разбиение одного канала с переменными параметрами на параллельные гауссовы каналы с различными отношениями сигнал/шум и точное измерение характеристик канала.

В соответствии с первым принципом OFDM, каждая поднесущая модулируется независимо посредством квадратурной амплитудной модуляции. Общий сигнал вычисляется посредством обратного быстрого преобразования Фурье (ОБПФ) как

$$s(t) = Re\left\{ e^{i2\pi f_c t} \sum_{k=-N/2}^{N/2} c_k e^{i2\pi k \Delta f (t-T_g)} \right\} \ (0 < t < T_s),$$

где c_k — комплексное представление символа квадратурной модуляции (QAM-символа). Комплексное представление удобно, поскольку генерация радиосигнала происходит посредством квадратурного модулятора в соответствии с выражением $s_k(t) = I_k \cos(2\pi f)B - Q_k \sin(2\pi f)$, где I_k и Q_k — синфазное и квадратурное (целое и мнимое) значения комплексного символа, соответственно.

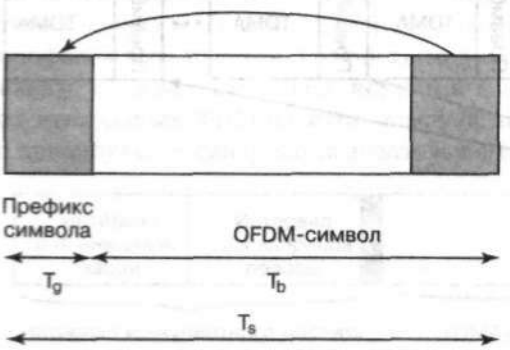


Рис. 7.9. OFDM-символ

Для работы алгоритмов БПФ/ОБПФ удобно, чтобы число точек соответствовало 2^m. Поэтому число несущих выбирают равным минимальному числу $N_{FFT} = 2^m$, превосходящему N. В режиме OFDM стандарта IEEE 802.16 $N = 200$, соответственно $N_{FFT} = 256$. Из них 55 ($k = -128 \ldots -101$ и $101 \ldots 127$) образуют защитный интервал на границах частотного диапазона канала. Центральная частота канала ($k = 0$) и частоты защитных интервалов не используются (т. е. амплитуды соответствующих им сигналов равны нулю). Оставшиеся 200 несущих — информационные.

В соответствии со вторым принципом OFDM, для точного определения параметров канала необходимы так называемые пилотные несущие частоты, метод модуляции и передаваемый сигнал в которых хорошо известен всем станциям в сети. В методе OFDM предусмотрено использование восьми пилотных частот (с индексами ±88, ±63, ±38, ±13). Остальные 192 несущие распределены между 16 логическими подканалами по 12 несущих в каждом, причем в одном подканале частоты расположены не подряд. Например, подканал 1 составляют несущие с индексами −100, −99, −98, −37, −36, −35, 1, 2, 3, 64, 65, 66. Деление на подканалы необходимо, поскольку в режиме WirelessMAN-OFDM предусмотрена (опционально) возможность работы не во всех 16, а в одном, двух, четырех и восьми подканалах — некий прообраз схемы множественного доступа OFDMA. Для этого каждый подканал и каждая группа подканалов имеют свой индекс (от 0 до 31).

Длительность полезной части T_b OFDM-символа зависит от ширины полосы канала BW и системной тактовой частоты (частоты дискретизации): $F_s = N_{FFT}/T_b$. Соотношение $F_s/BW = n$ нормируется и в зависимости от ширины полосы канала принимает значения 86/75 (BW кратно 1,5 МГц), 144/125 (BW

кратно 1,25 МГц), 316/275 (*BW* кратно 2,75 МГц), 57/50 (*BW* кратно 2 МГц) и 8/7 (*BW* кратно 1,75 МГц и во всех остальных случаях).

7.5.1. Канальное кодирование

Кодирование данных на физическом уровне включает три стадии — рандомизацию, помехоустойчивое кодирование и перемежение. Рандомизация — это умножение блока данных на псевдослучайную последовательность (ПСП), которую формирует генератор ПСП с задающим полиномом вида $1 + x^{14} + x^{15}$ (см. рис. 7.4).

В нисходящем потоке генератор ПСП инициализируется с началом кадра посредством кодового слова $4A80_{16}$. Начиная со второго пакета кадра генератор ПСП инициализируется на основе идентификационного номера базовой станции BSID, идентификатора профиля пакета и номера кадра (рис. 7.10). В восходящем потоке все происходит аналогично, с той лишь разницей, что инициализация генератора ПСП по схеме, приведенной на рис. 7.10, происходит с первого пакета (вместо кода профиля пакета нисходящего канала DIUC используется аналогичный идентификатор для восходящего канала UIUC).

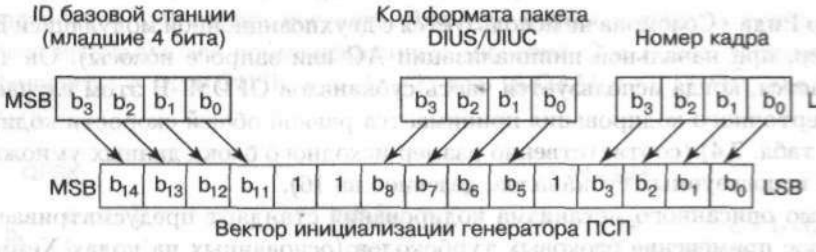


Рис. 7.10. Формирование вектора инициализации генератора ПСП для рандомизации нисходящего потока OFDM

Кодирование данных предполагает кодирование каскадным кодом: внешним кодом Рида–Соломона и внутренним сверточным кодом. Алгоритм кодирования Рида–Соломона, используемый в данном стандарте, строится над полем Галуа GF (256). В базовом виде он оперирует блоками исходных данных по 239 байт, формируя из них кодированный блок размером 255 байт (добавляя 16 проверочных байт). Такой код способен исправить до 8 поврежденных байт или обнаружить до 15 поврежденных или стертых внутренним кодом байт. Поскольку реально используются блоки данных меньшей длины K, перед ними добавляются $(239 - K)$ нулевых байт (так называемое укорочение линейного кода). После кодирования эти байты удаляются. Если необходимо сократить число проверочных слов так, чтобы уменьшить число восстанавливаемых байтов T, используются только 2T первых проверочных байт. Обязательные для поддержки в IEEE 802.16 варианты каскадного кодирования приведены в табл. 7.4.

После кодера Рида–Соломона данные поступают в сверточный кодер (рис. 7.11) с порождающими последовательностями (генераторами кода) $G_1 = 171_8$ (для выхода X) и $G_2 = 133_8$ (для Y) — так называемый стандартный код NASA [12]. Его базовая скорость кодирования — 1/2, т. е. из каждого входного бита он формирует пару кодированных битов X и Y. Кодовое ограничение этого кода $K = 7$. Упуская («выкалывая» или перфорируя) из последовательности пар элементы X_i

или Y_i, можно получать различные скорости кодирования. Так, скорости 2/3 соответствует последовательность (X_1 Y_1 Y_2), скорости 3/4 — (X_1 Y_1 Y_2 X_3), 5/6 — (X_1 Y_1 Y_2 X_3 Y_4 X_5).

Таблица 7.4. Обязательные схемы кодирования/модуляции в режиме OFDM

Модуляция	Блок данных до кодирования, байт	Кодер Рида – Соломона	Скорость кодирования сверточного кодера	Суммарная скорость кодирования	Блок данных после кодирования, байт
BPSK	12	(12,12,0)	1/2	1/2	24
QPSK	24	(32,24,4)	2/3	1/2	48
QPSK	36	(40,36,2)	5/6	3/4	48
16-QAM	48	(64,48,8)	2/3	1/2	96
16-QAM	72	(80,72,4)	5/6	3/4	96
64-QAM	96	(108,96,6)	3/4	2/3	144
64-QAM	108	(120,108,6)	5/6	3/4	144

Кодер Рида – Соломона не используется с двухпозиционной модуляцией BPSK (например, при начальной инициализации АС или запросе полосы). Он также пропускается, когда используется часть субканалов OFDM. В этом случае скорость сверточного кодирования принимается равной общей скорости кодирования (см. табл. 7.4) (соответственно размер исходного блока данных умножается на число используемых субканалов, деленное на 16).

Помимо описанного механизма кодирования стандарт предусматривает опциональное применение блоковых турбокодов (основанных на кодах Хемминга и контроле четности) и сверточных турбокодов.

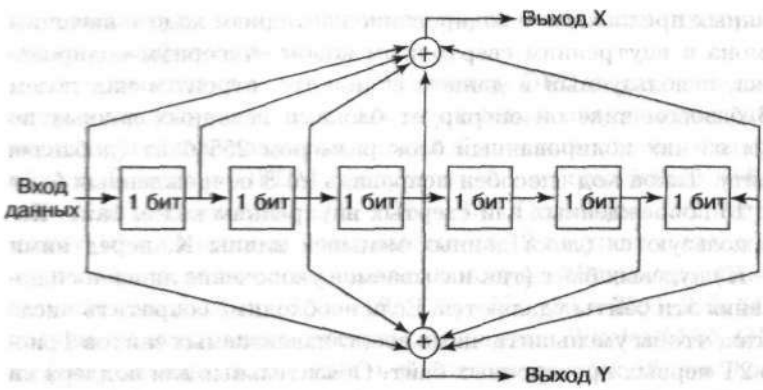


Рис. 7.11. Схема сверточного кодера

После кодирования следует процедура перемежения (интерливинга) — перемешивания битов в пределах блока кодированных данных, соответствующего OFDM-символу. Эта операция проводится в две стадии. Цель первой — сделать так, чтобы смежные биты оказались разнесенными по несмежным несущим. На второй стадии смежные биты оказываются разнесенными в разные половины последовательности. Все это делается для того, чтобы при групповых

(пакетируемых) ошибках в символе повреждались не смежные биты, которые легко восстановить при декодировании. Перемежение реализуется в соответствии с формулами:

$$m_k = (N_{cbps}/12) \cdot (k \mod 12) + \text{floor}(k/12);$$
$$j_k = s \cdot \text{floor}(m_k/s) + [m_k + N_{cbps} - \text{floor}(12m_k/N_{cbps})] \mod s; \quad (7.1)$$
$$k = 0, \ldots, N_{cbps} - 1,$$

где m_k и j_k — номер исходного k-го бита после первой и второй стадий перемежения, соответственно; N_{cbps} — число кодированных бит в OFDM-символе (при заданном числе субканалов); s — 1/2 числа битов на несущую (2/4/6 бит для QPSK/16-QAM/64-QAM, соответственно, для BPSK $s = 1$). Функция $\text{floor}(x)$ — это наибольшее целое число, не превосходящее x; функция $(x \mod r)$ — остаток от x/r.

После перемежения начинается стадия модуляции. Исходя из выбранной схемы модуляции (BPSK/QPSK/16-QAM/64-QAM), блок представляется в виде последовательности групп битов, соответствующих модуляционным символам (по 1/2 /4/6 бит). Каждой группе ставятся в соответствие значения Q и I из векторных диаграмм Грея (рис. 7.12), которые затем используются при непосредственной модуляции несущей. Для усреднения амплитуд квадратурных символов используются нормализованные значения Q и I, т. е. умноженные на коэффициенты c (для QPSK $c = 1/\sqrt{2}$, для 16-QAM $c = 1/\sqrt{10}$, для 64-QAM $c = 1/\sqrt{42}$).

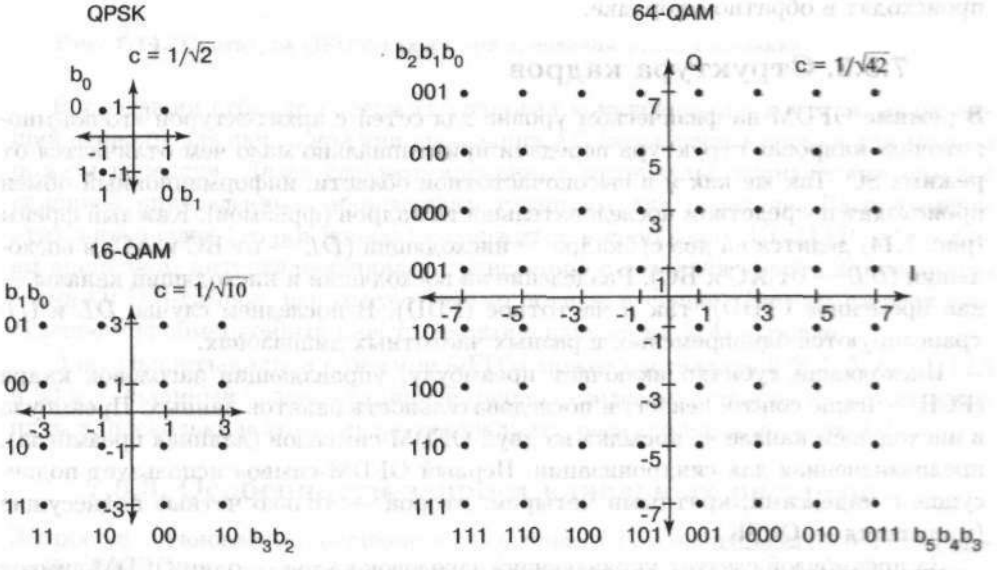


Рис. 7.12. Векторные диаграммы Грея (представление модуляционных символов) для QPSK, 16-QAM и 64-QAM

Пилотные несущие модулируются посредством BPSK. Значения сигналов на этих несущих определяются на основании ПСП w_k с задающим полиномом $x^{11} + x^9 + 1$, причем в нисходящем субкадре k — номер символа относительно начала кадра, в восходящем — номер символа относительно начала пакета

(рис. 7.13). Инициализирующие слова генератора ПСП для нисходящего и восходящего потоков различны ($8FF_{16}$ и 555_{16}, соответственно). Собственно значения BPSK-символов вычисляются как $c_{-88} = c_{-38} = c_{63} = c_{88} = 1 - 2w_k$; $c_{-63} = c_{-13} = c_{13} = c_{38} = 1 - 2\overline{w_k}$ в нисходящем канале и $c_{-88} = c_{-38} = c_{13} = c_{38} = c_{63} = c_{88} = 1 - 2w_k$; $c_{-63} = c_{-13} = 1 - 2\overline{w_k}$ — в восходящем. В результате получается так называемая ступенчатая конструкция OFDM.

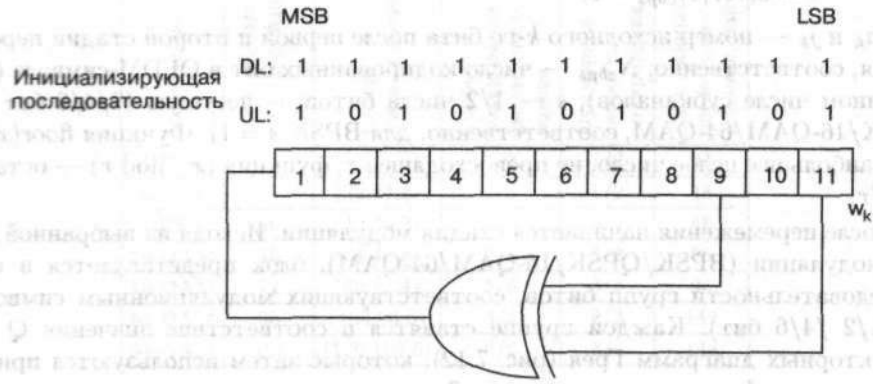


Рис. 7.13. Генерация модулирующей последовательности для пилотных несущих

После определения модуляционных символов посредством ОБПФ вычисляется сам радиосигнал и передается в передатчик. При приеме все процедуры происходят в обратном порядке.

7.5.2. Структура кадров

В режиме OFDM на физическом уровне для сетей с архитектурой «точка-многоточка» кадровая структура передачи принципиально мало чем отличается от режима SC. Так же как и в высокочастотной области, информационный обмен происходит посредством последовательности кадров (фреймов). Каждый фрейм (рис. 7.14) делится на два субкадра — нисходящий (DL — от БС к АС) и восходящий (UL — от АС к БС). Разделение на восходящий и нисходящий каналы — как временное (TDD), так и частотное (FDD). В последнем случае DL и UL транслируются одновременно, в разных частотных диапазонах.

Нисходящий субкадр включает преамбулу, управляющий заголовок кадра (FCH — frame control header) и последовательность пакетов данных. Преамбула в нисходящем канале — посылка из двух OFDM-символов (длинная преамбула), предназначенная для синхронизации. Первый OFDM-символ использует поднесущие с индексами, кратными четырем, второй — только четные поднесущие (модуляция — QPSK).

За преамбулой следует управляющий заголовок кадра — один OFDM-символ с модуляцией BPSK и стандартной схемой кодирования (скорость кодирования — 1/2). Он содержит так называемый префикс кадра нисходящего канала (DLFP — Downlink Frame Prefix), который описывает профиль и длину первого (или нескольких начальных) пакета в DL-субкадре.

В первый пакет входят широковещательные сообщения (предназначенные всем АС) — карты расположения пакетов DL-MAP, UL-MAP, дескрипторы

нисходящего/восходящего каналов DCD/UCD, другая служебная информация. Каждый пакет обладает своим профилем (схема кодирования, модуляция и т. д.) и передается посредством целого числа OFDM-символов. Точки начала и профили всех пакетов, помимо первого, содержатся в DL-MAP.

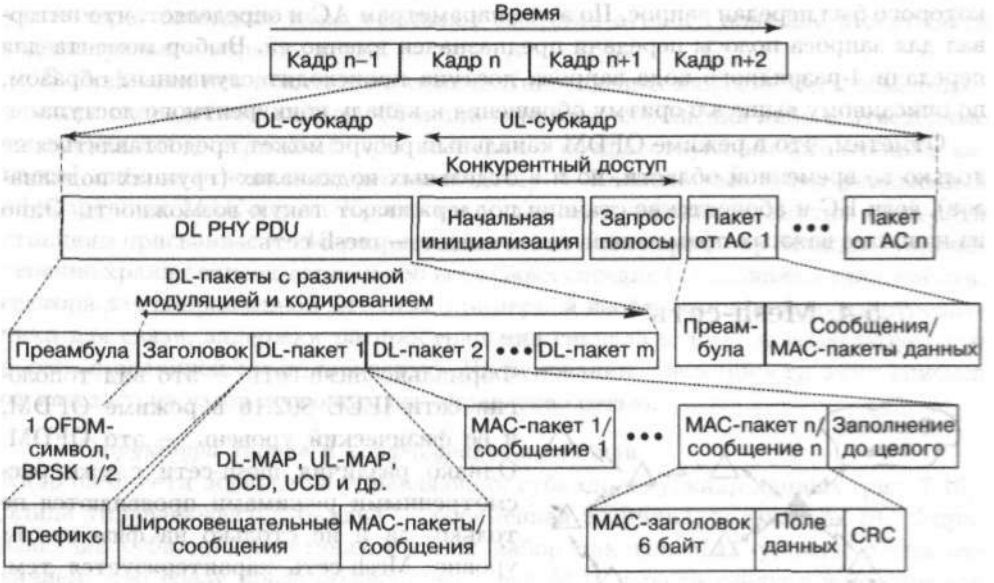


Рис. 7.14. Структура OFDM-кадров при временном дуплексировании

Восходящий субкадр содержит интервал конкурентного доступа, включающий периоды для начальной инициализации АС (вхождение в сеть) и для запроса полосы передачи. Далее следуют временные интервалы, назначенные базовой станцией определенным абонентским станциям для передачи. Распределение этих интервалов (точки начала) содержится в сообщении UL-MAP. АС в своем временном интервале начинает трансляцию с передачи короткой преамбулы (один OFDM-символ, использует только четные несущие). За ним следует собственно информационный пакет, сформированный на MAC-уровне.

Длительность кадров в режиме OFDM может составлять 2,5; 4; 5; 8; 10; 12,5 и 20 мс. Заданный базовой станцией, период следования кадров не может изменяться, поскольку в этом случае потребуется ресинхронизация всех АС.

7.5.3. Особенности запроса канальных ресурсов

Запрос на установление соединения не отличается от общепринятого в стандарте IEEE 802.16, за исключением дополнительного режима «концентрированного» запроса (Region-Focused). Он предназначен только для станций, способных работать с отдельными субканалами. В этом режиме в интервалах конкурентного доступа (заданных в UL-MAP) АС может передать короткий 4-разрядный код на одном из 48 субканалов, каждый из которых включает четыре несущие. Всего предусмотрено восемь кодов. Таблица кодов и подканалов приведена в тексте стандарта IEEE 802.16. Код и номера канала АС выбирает случайным образом.

Получив кодовое сообщение, БС предоставляет АС интервал для передачи «обычного» запроса на предоставление доступа (специального заголовка запроса), если это возможно. Однако, в отличие от других механизмов, БС в UL-MAP не указывает идентификатор запросившей ее станции, а приводит номера кода запроса, подканала, а также порядковый номер интервала доступа, в течение которого был передан запрос. По этим параметрам АС и определяет, что интервал для запроса полосы передачи предназначен именно ей. Выбор момента для передачи 4-разрядного кода запроса доступа происходит случайным образом, по описанному выше алгоритму обращения к каналу конкурентного доступа.

Отметим, что в режиме OFDM канальный ресурс может предоставляться не только во временной области, но и в отдельных подканалах (группах подканалов), если БС и абонентские станции поддерживают такую возможность. Одно из наиболее важных применений такой опции — mesh-сеть.

7.5.4. Mesh-сеть

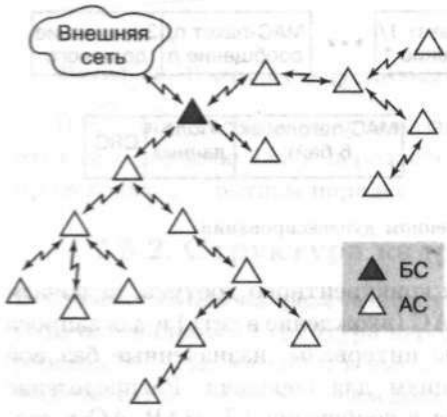


Рис. 7.15. Пример mesh-сети

Формально mesh-сеть — это вид топологии сети IEEE 802.16 в режиме OFDM, и ее физический уровень — это OFDM. Однако различия mesh-сети с уже рассмотренными режимами проявляются не только, да и не столько на физическом уровне. Mesh-сеть характеризуется тем, что ее узлы обеспечивают сквозную передачу (ретрансляцию) трафика. Поэтому основное отличие mesh-сети в том, что если в рассматриваемой до сих пор архитектуре «точка-многоточка» АС может общаться только с БС, то в mesh-сети возможно взаимодействие непосредственно между АС. Соответственно mesh-сети вошли в IEEE 802.16 вовсе не в целях создания одноранговых локальных сетей — для этого есть стандарты группы IEEE 802.11. Причина в ином — необходим инструмент построения широкополосной сети, в которой трафик может передаваться по цепочке из нескольких станций, ликвидируя тем самым проблемы передачи при отсутствии прямой видимости. Поэтому и все механизмы управления, в принципе позволяющие построить децентрализованную распределенную сеть, ориентированы все же на древовидную архитектуру (корневой узел — БС) с выделенной базовой станцией и доминирующими потоками БС–АС.

Топология сети

В mesh-сети все станции (узлы) формально равноправны. Однако практически всегда присутствует один узел, через который происходит обмен трафика mesh-сети с внешним окружением (рис. 7.15). Такой узел называют базовой станцией mesh-сети, именно на него возлагается часть необходимых для управления mesh-сетью функций. При этом управление доступом может происходить либо

на основе механизма распределенного управления, либо централизованным способом под управлением БС. Возможна и комбинация этих методов.

Базовое понятие в mesh-сети — соседи. Под соседями определенного узла понимают все узлы, которые могут устанавливать с ним непосредственное соединение. Все они образуют соседское окружение. Узлы, связанные с заданным узлом через соседские узлы, называют соседями второго порядка. Могут быть соседи третьего порядка и т. д.

В mesh-сети нет понятия восходящих/нисходящих каналов. Весь обмен происходит посредством кадров. Станции передают сообщения либо в отведенные им временные интервалы (в соответствии с предшествующим назначением каналов), либо на основе конкурентного доступа. Каждый узел имеет уникальный 48-разрядный MAC-адрес. Кроме того, для идентификации внутри mesh-сети станциям присваивается 16-разрядный сетевой идентификатор. Каждый узел постоянно хранит список данных обо всех своих соседях (с указанием удаленности, сектора для направленной антенны, примерной необходимой мощности передатчика для связи, задержки распространения сигнала и т. п.) и транслирует его в сеть с заданной периодичностью. На основании совокупности этих списков от каждого из узлов и происходит управление сетью.

Структура кадров и конфигурирование сети

Кадр mesh-сети делится на управляющий субкадр и субкадр данных (рис. 7.16). Длина управляющего субкадра — переменная величина, задаваемая БС. Управляющий субкадр представляет собой набор пакетов MAC-уровня, с тем отличием, что сразу после общего заголовка MAC-пакета следует подзаголовок mesh-сети. Управляющий субкадр в зависимости от реализуемой функции может быть двух типов — управления сетью (network control) и управления очередностью доступа к каналам связи (schedule control). В субкадрах управления всегда используется модуляция QPSK со скоростью кодирования 1/2.

Субкадры управления включают интервалы для подключения к сети новых устройств (Network entry — «сетевой вход») и следующие за ними сообщения «конфигурация сети». Сообщения типа «конфигурация сети» содержат всю необходимую информацию о составе сети. Они же реализуют процедуры управления. Эти сообщения генерирует каждый узел и транслирует по сети через свое соседское окружение. Среди передаваемой информации — списки соседей каждого узла, идентификационный номер БС и число ее соседей, номер логического канала для передачи графика доступа к каналам, удаленность узла (ранг соседства) от БС и т. д. Посредством таких сообщений с заданной периодичностью транслируется дескриптор сети — таблица, полностью описывающая текущие параметры сети. Она содержит такие параметры, как длительность кадров, длина управляющего субкадра, число интервалов для сообщений децентрализованного распределения ресурсов, периодичность следования субпакетов распределения ресурсов, профили пакетов, тип кодирования, соответствие логических каналов физическим и т. п. Дескриптор сети передается от базовой станции ее соседскому окружению, от него — узлам со следующим рангом соседства и т. д. Периодичность передачи дескриптора сети нормирована.

«Сетевой вход» — это интервал, в течение которого новый узел может послать сообщение (NENT) о своем намерении подключиться к сети (аналог интервала конкурентного доступа в сети «точка-многоточка»). Перед этим он

должен принять сообщение о конфигурации сети, выбрать узел для подключения, синхронизироваться с ним и лишь затем отправлять запрос. В ответ узел либо откажет в доступе, либо назначит новому узлу сетевой идентификатор, канал и временной интервал для проведения процедур аутентификации.

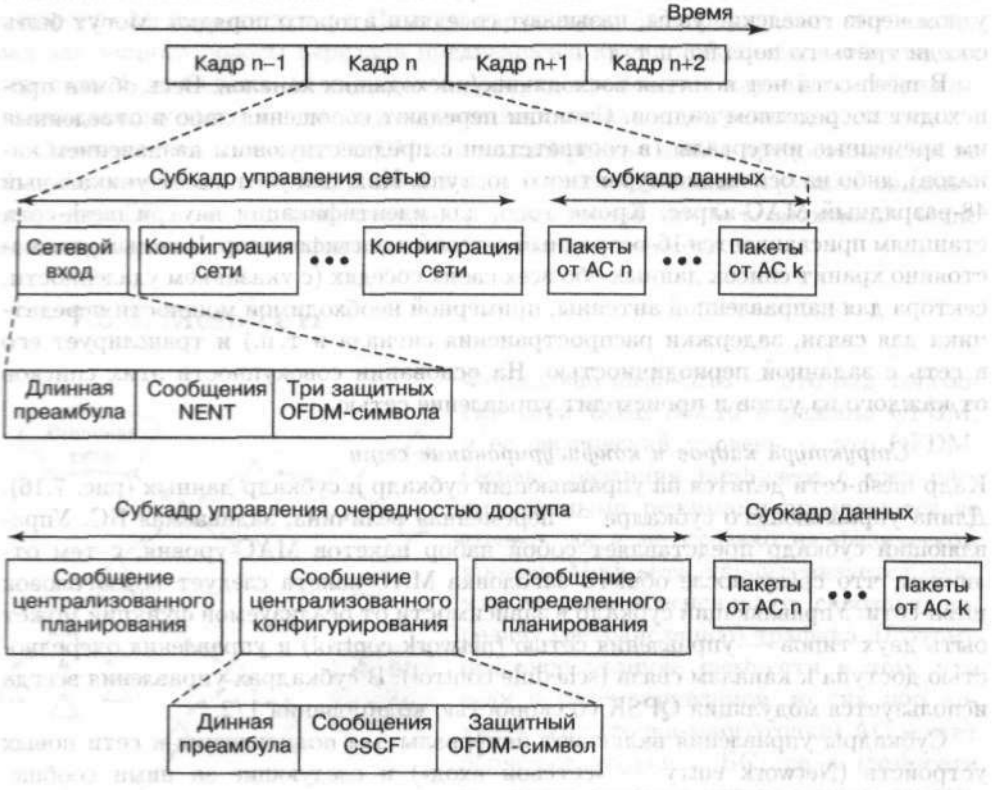


Рис. 7.16. Структура кадра mesh-сети

Методы управления канальными ресурсами

Распределение канальных ресурсов в mesh-сети может быть централизованным и децентрализованным (распределенным). В свою очередь, децентрализованное распределение бывает координированным с БС и некоординированным.

Децентрализованное распределение ресурсов подразумевает, что распределение происходит в пределах одной группы соседей (т. е. между станциями, способными непосредственно связываться друг с другом). При координированном децентрализованном распределении узлы обмениваются между собой специальными сообщениями управления распределением (distributed scheduling — DSCH). DSCH-сообщения — это запросы на получение ресурса и ответные сообщения с предоставлением (подтверждением) свободного ресурса (временного интервала в субкадре данных). Ресурс предоставляется соседом под конкретное соединение.

Координированность заключается в том, что период выдачи таких сообщений каждой станцией определен и известен ее соседям. Координированные

DSCH-сообщения передаются в субкадрах управления очередностью доступа в оговоренных в сетевом дескрипторе интервалах. Некоординированные DSCH-сообщения передаются в субкадре данных.

Централизованное распределение ресурсов подразумевает древовидную топологию сети с БС в вершине. Оно реализовано посредством двух типов сообщений — централизованного конфигурирования CSCF и централизованного планирования CSCH. Эти управляющие сообщения размещаются в начале субкадра управления графиком доступа. Используя сообщения централизованного планирования CSCH, каждый узел определяет потребность в трафике своих дочерних узлов (т.е. трафик которых от/к БС проходит через данный узел) и сообщает свою потребность вышестоящему узлу, вплоть до БС. Проанализировав потребность, БС рассылает сообщение CSCH, информируя каждый узел о выделенной ему полосе пропускания (в бит/с) в восходящем и нисходящем направлениях. Исходя из этих данных, каждый узел уже сам запрашивает (или назначает) расположение пакетов в субкадре данных у (для) своих соседских узлов посредством сообщений децентрализованного планирования DSCH.

Сообщения централизованного конфигурирования CSCF формируются БС и транслируются по сети для информирования всех ее узлов о текущем состоянии. CSCF включает такую информацию, как число доступных логических каналов и их перечень, перечень узлов в сети с указанием числа дочерних узлов для каждого из них, а также профили восходящих/нисходящих пакетов для каждого дочернего узла.

7.6. Режим OFDMA

Режим OFDMA с точки зрения формирования модуляционных символов аналогичен OFDM. Различие проявляется в принципе разделения каналов. Один логический OFDMA-канал образован фиксированным набором несущих, как правило, распределенных по всему доступному диапазону частот физического канала. В упрощенном виде этот механизм опционально используется в режиме OFDM — вспомним разбиение канала на 16 подканалов. Ширина физического канала (BW) не нормирована (в стандарте говорится «не менее 1 МГц»), но в реальных применениях вряд ли окажутся эффективными каналы менее 5 МГц.

7.6.1. Особенности формирования символов и канального кодирования

В OFDMA несущих значительно больше — 2048, соответственно и число подканалов становится достаточным для организации работы сети: в разных режимах их от 32 до 70, по 24 или 48 информационных несущих в каждом. Системная тактовая частота — 8/7 BW.

Метод формирования, структура OFDM-символов и механизм канального кодирования в OFDMA схожи с описанным для OFDM. Канальное кодирование включает рандомизацию, помехоустойчивое кодирование, перемежение и модуляцию. Метод рандомизации также практически идентичен OFDM (различны только способы формирования инициализирующего вектора генератора ПСП).

Помехоустойчивое кодирование в OFDMA в качестве обязательного предусматривает только сверточный кодер — такой же, как в OFDM, и с тем же

набором скоростей кодирования. Кодера Рида – Соломона нет. Опционально предусмотрено применение блочных и сверточных турбокодов. Метод перемежения также идентичен с OFDM — надо лишь в соответствующих формулах (7.1) заменить 12 (для OFDM) на 16.

Схемы модуляции несущих полностью совпадают с OFDM, с той лишь разницей, что предусмотренный набор включает только QPSK и 16-QAM со скоростями кодирования 1/2 и 3/4, а также опционально 64-QAM со скоростями кодирования 1/2, 2/3 и 3/4. Однако в OFDMA после формирования символов квадратурной модуляции и усреднения их амплитуд (нормировке на параметр c) последовательность символов на каждой несущей умножается на бинарную ПСП w_k, схема генератора которого идентична приведенной на рис. 7.13. Каждая k-я несущая умножается на значение $1-2w_k$ (т. е. если $w_k = 0$, то $1-2w_k = 1$ и символ не изменяется; если $w_k = 1$, символ умножается на -1). Символы пилотных несущих модулируются методом BPSK, их значения вычисляются как $c_k = 1 - 2w_k$. Поскольку мощность сигналов пилотных несущих в нисходящем канале (опционально — и в восходящем) должна быть на 2,5 дБ выше средней мощности информационных несущих, значение c_k дополнительно умножается на 4/3.

7.6.2. Структура кадров, методы распределения несущих

Структура кадров (рис. 7.17) в OFDMA схожа со всеми рассмотренными режимами в том, что сохраняется подразделение на восходящий и нисходящий субкадры, как временное, так и частотное. Длительность кадра может составлять 2; 2,5; 4; 5; 8; 10; 12,5 и 20 мс. Кадр — это последовательность OFDMA-символов. Каждый OFDMA-символ включает набор подканалов. Пакеты данных могут передаваться одновременно, на различных OFDMA-подканалах.

Для описания структуры кадра в OFDMA используется понятие слота — минимального ресурса для передачи данных. Слот занимает один подканал и от одного до трех последовательных OFDMA-символов. В нисходящем субкадре длительность слота — один символ в режиме FUSC, два — в режиме PUSC; в восходящем субкадре длительность слота всегда равна трем OFDMA-символам.

Подканал — это набор несущих частот (как и в OFDM). Распределение несущих по подканалам, равно как и число несущих на один подканал, зависит от направления передачи и метода распределения несущих. Стандарт IEEE 802.16 описывает несколько способов распределения несущих как в нисходящем канале, так и в восходящем. Принципиально они подразделяются на FUSC (full usage of the subchannels) — полное использование подканалов передатчиком БС и PUSC (partial usage of subchannels) — использование групп подканалов (сегментов), т. е. не всего доступного диапазона. Какие именно подканалы используются в режиме PUSC, однозначно определяют номера сегментов.

В методах PUSC и FUSC (и их вариациях) одному субканалу присваиваются несущие, равномерно распределенные по всему доступному физическому каналу. Используется и другой подход — применение в подканалах набора соседних поднесущих частот. Он реализован в методе AMC (Advanced modulation and coding), предназначенном для работы с адаптивными антенными системами.

В методе AMC используется также 2048 несущих. Из них 160 нижних и 159 верхних образуют защитные интервалы, центральная частота не использует-

ся. Оставшиеся несущие последовательно разбиваются на 192 группы (группа именуется bin) по 9 несущих в каждой. Центральная (пятая) частота в каждой группе — пилотная. Четыре смежные группы образуют полосу. AMC-подканал использует либо одну группу в шести последовательных по времени OFDM-символах, либо две группы в трех OFDM-символах. Структура AMC-подканалов в восходящем и нисходящем субкадрах одинакова. При этом перестановки поднесущих происходят в пределах субканала (опциональный механизм перестановок соседних поднесущих).

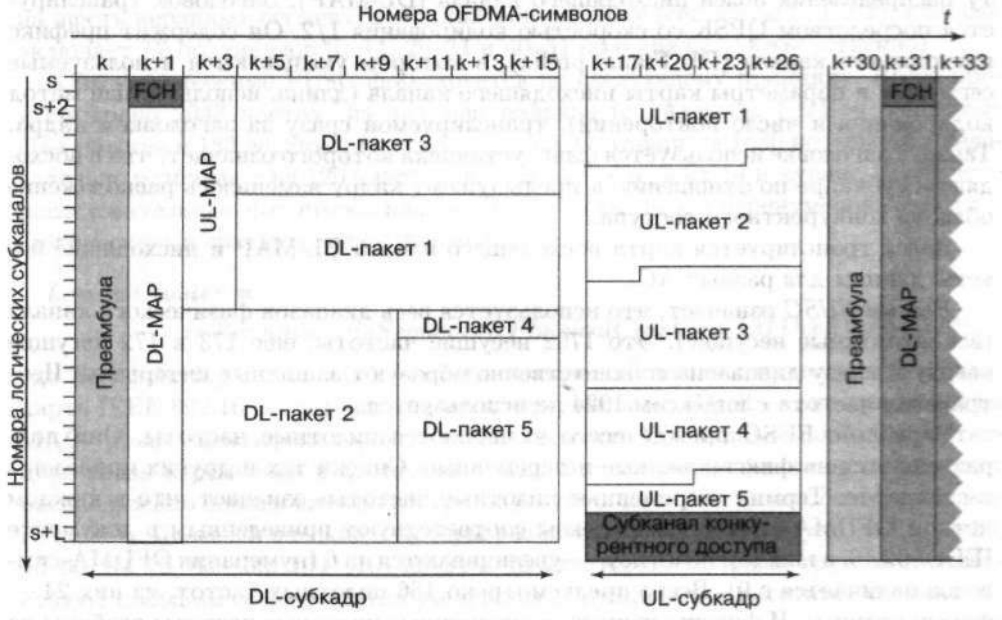


Рис. 7.17. Структура OFDMA-кадров

Отметим, что в пределах одного субкадра возможно использование различных механизмов распределения несущих по подканалам — FUSC, PUSC, AMC и т. д. Границы соответствующих зон (именуемых в стандарте «зонами перестановки» — permutation zone) определены в картах субкадров (рис. 7.18).

Рис. 7.18. Совмещение различных «зон перестановки» в OFDMA-кадре

7.6.3. Нисходящий OFDMA-канал

В нисходящем канале (см. рис. 7.17) первый символ — это преамбула. Несущие преамбулы модулируются посредством BPSK специальным псевдослучайным кодом, зависящим от используемого сегмента (в режиме PUSC) и переменной IDcell, задаваемой на MAC-уровне. *Одному сегменту с номером* $n = [0 \dots 2]$ соответствует набор несущих с шагом 3 и начальным сдвигом n. Распознав тип преамбулы, АС сразу определяет значение переменной IDcell и режим работы БС.

За преамбулой следуют два символа, передающие заголовок кадра FCH и карту распределения полей нисходящего канала (DL-MAP). Заголовок транслируется посредством QPSK со скоростью кодирования 1/2. Он содержит префикс нисходящего канала (DL Frame prefix), в котором указываются используемые сегменты и параметры карты нисходящего канала (длина, используемый метод кодирования и число повторений), транслируемой сразу за заголовком кадра. Также в заголовке используется флаг, установка которого означает, что в восходящем субкадре по отношению к предыдущему кадру изменилось расположение области конкурентного доступа.

Далее транслируется карта восходящего канала UL-MAP и нисходящие пакеты данных для разных АС.

Режим FUSC означает, что используется весь диапазон физического канала (все возможные несущие). Это 1702 несущие частоты, еще 173 и 172 несущие вверху и внизу диапазона соответственно образуют защитные интервалы. Центральная частота с индексом 1024 не используется.

В режиме FUSC прежде всего назначаются пилотные частоты. Они подразделяются на фиксированные и переменные. Списки тех и других приведены в стандарте. Термин «переменные пилотные частоты» означает, что в каждом четном OFDMA-символе их индексы соответствуют приведенным в документе IEEE 802.16, в каждом нечетном — увеличиваются на 6 (нумерация OFDMA-символов начинается с 0). Всего предусмотрено 166 пилотных частот, из них 24 — фиксированные. И фиксированные, и переменные пилотные частоты разбиты на два набора, одинаковые по объему. Это разбиение имеет значение только при работе с адаптивными антенными системами в режиме пространственно-временного кодирования (STC).

После назначения пилотных частот оставшиеся 1536 несущих предназначены для передачи данных. Они подразделяются на $N_{subchannels}B = 32$ подканала по $N_{subcarriers} = 48$ несущих в каждом. Назначение информационных несущих подканалам происходит в соответствии с формулой:

$$subcarrier(k,s) = N_{subchannels}n_k + \{P[(s+n_k) \bmod N_{subchannels}] + \\ + IDcell\} \bmod N_{subchannels}, \quad (7.2)$$

где $subcarrier(k,s)$ — индекс несущей k в подканале s,

$$s = [0 \dots N_{subchannels} - 1], \quad k = [0 \dots N_{subcarriers} - 1],$$
$$n_k = (k + 13s) \bmod N_{subcarriers}.$$

IDCell — определяемый на MAC-уровне идентификатор отдельного сегмента БС (задаваемая базовой станцией целая переменная в диапазоне 0–31). $P(x)$ означает x-й элемент последовательности перестановок $\{P\}$, приведенной в стандарте ($P = \{3, 18, 2, 8, 16, 10, 11, 15, 26, 22, 6, 9, 27, 20, 25, 1, 29, 7, 21, 5, 28, 31, 23, 17, 4, 24, 0, 13, 12, 19, 14, 30\}$). Операция $x \mod k$ означает остаток от x/k.

Очевидно, что перед применением приведенной формулы информационные несущие должны быть перенумерованы так, чтобы их индексы укладывались в диапазон 0–1535 (последнее значение соответствует физическому индексу 1702), т. е. пронумерованы подряд, без учета пилотных частот. Поскольку в четных и нечетных символах расположение пилотных частот различно, распределение информационных несущих для них также нужно вычислять независимо.

В *режиме PUSC* весь доступный диапазон подканалов делится на ряд независимых сегментов. Для работы используется часть из них, но не менее 12 подканалов. Всего в режиме PUSC может быть 60 подканалов. Они разбиты на шесть неравных сегментов. Из них три базовых (сегменты 0, 1 и 2), каждый включает 12 подканалов (0–11, 20–31 и 40–51 подканалы, соответственно).

Символ в режиме PUSC формируется по следующему принципу. Всего предусмотрено 2048 частот, из них центральная (с индексом 1024) и защитные (184 нижние и 183 верхние) не используются. Оставшиеся 1680 несущих последовательно разбивают на 120 кластеров, каждый содержит 14 несущих. После чего последовательные физически кластеры (PhysicalCluster) перенумеровываются в «логические» в соответствии с формулой:

$$LogicalCluster = \\ = RenumberingSequence((PhysicalCluster + 13IDcell) \bmod 120),$$

где $RenumberingSequence(x)$ — соответствующий элемент приведенной в стандарте IEEE 802.16 последовательности перестановок; IDcell — определяемый на MAC-уровне идентификатор отдельного сегмента БС (задаваемая базовой станцией целая переменная в диапазоне 0–31). Эта операция фактически означает распределение последовательных групп несущих по всему диапазону физического канала. Далее логические кластеры разбиваются на шесть групп (0–23, 24–39, 40–63, 64–79, 80–103, 104–119) по 24 и 16 кластеров. Большие группы соответствуют большим сегментам (по умолчанию, группа 0 соответствует сегменту 0, группа 2 — сегменту 1, группа 4 — сегменту 2). В каждом кластере определяются пилотные несущие: для четных символов это 5-я и 9-я несущие, для нечетных — 1-я и 13-я (рис. 7.19).

Рис. 7.19. Структура кластера

После того как пилотные несущие определены, информационные несущие в сегменте нумеруются подряд, без учета пилотных частот. Затем несущие распределяются по подканалам: в соответствии с формулой (7.2) каждому подканалу назначаются по 24 несущих. Распределение происходит в рамках отдельных сегментов, объединяющих 8 или 12 подканалов. Поэтому в формуле (7.2) используются значения $N_{subchannels} = 12$ или 8, $N_{subcarriers} = 24$, а также специальные перестановочные последовательности P_{12} и P_8 для сегментов из 12 и 8 подканалов, приведенных в стандарте.

В стандарте предусмотрены и опциональные методы распределения несущих, в частности, так называемый optional FUSC, принципиально не отличающийся от рассмотренного.

7.6.4. Восходящий канал

Восходящий субкадр следует непосредственно за нисходящим через интервал TTG. Он содержит пакеты от абонентских станций и интервал для запроса доступа/инициализации. Минимальный размер одного сообщения в восходящем подкадре (слот) — три OFDMA-символа в одном подканале. Это привело к появлению в документе IEEE 802.16 термина «фрагмент» (мозаичный элемент, tile).

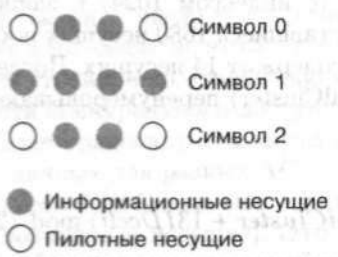


Рис. 7.20. Структура «фрагмента» восходящего канала

«Фрагмент» представляет собой совокупность трех символов и четырех несущих, в котором положения пилотных частот жестко определены (рис. 7.20).

В восходящем канале весь частотный диапазон канала (1680 несущих) разбивается на 420 последовательных «фрагментов», по четыре несущие в каждом. Предусмотрено 70 подканалов. Каждый из них включает шесть «фрагментов», т. е. 24 несущие на символ в одном подканале. Распределение «фрагментов» по подканалам происходит следующим образом. Все 420 «фрагментов» разбиваются на шесть групп по 70 «фрагментов». В каждый подканал включается по одному «фрагменту» из каждой группы в соответствии с уравнением:

$$Tile(n,s) = 70n + \{P[(n+s) \mod 70] + UL\_IDcell\} \mod 70,$$

где $Tile(n,s)$ — «фрагмент» n подканала s; $n = [0 \ldots 5]$; $s = [0 \ldots 69]$, $P(x)$ — перестановочная последовательность; UL_IDcell — переменная в диапазоне 0–69, задаваемая БС на MAC-уровне. В результате каждому подканалу в каждом символе назначается свой набор несущих.

После распределения по подканалам происходит нумерация информационных несущих в каждом слоте — всего их 48 в трех символах. Информационные частоты нумеруются в подканале начиная с наименьшей несущей «фрагмента» с наименьшим индексом — сначала в первом символе, затем во втором и третьем. После чего информационные несущие в каждом слоте перенумеровываются в соответствии с формулой $subcarrier(n,s)B = (n + 13s) \mod 48$, где s — номер подканала; $n = [0 \ldots 47]$ (т. е. происходит циклический сдвиг нумерации информационных несущих на $13s$ в каждом подканале s).

Отметим, что в тексте документа IEEE 802.16 происходит подмена терминов: подканалом в восходящем субкадре авторы текста IEEE 802.16 называют именно слот, информационную структуру размером 24 несущие на три символа. И когда в документе написано, что в субканале 48 информационных несущих, следует помнить, что с точки зрения правильной терминологии речь идет не о субканале,

а о слоте. Реальных же несущих (т. е. физических частот) в субканале всего 24. Умножая их на три (число OFDMA-символов в слоте) и вычитая 24 пилотные несущие, как раз и получим 48 информационных несущих.

Опционально в восходящем канале предусмотрен режим, в котором во «фрагменте» одна пилотная частота (рис. 7.21), шесть «фрагментов» на подканал, всего 96 подканалов (1728 используемых частот).

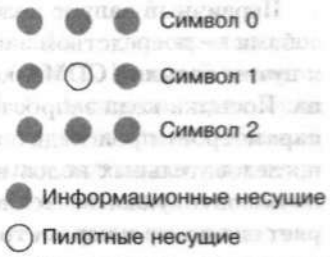


Рис. 7.21. Структура «фрагмента» восходящего канала в опциональном режиме

7.6.5. Запрос полосы и регистрация в сети

Механизмы запроса начальной инициализации в сети и первичного запроса полосы пропускания в режиме OFDMA схожи и принципиально отличаются от других режимов. Для этих запросов в OFDMA используется специально выделенный канал. Он назначается БС и состоит из шести последовательных подканалов, индексы которых приведены в UL-MAP. Запрос представляет собой 144-разрядный CDMA-код, передаваемый посредством BPSK, т. е. 1 бит на несущую. В результате для передачи такого кода достаточно шести подканалов (24 информационные несущие в каждом). Сам код формируется в генераторе ПСП — 15-разрядном сдвиговом регистре с задающим полиномом $1 + X^1 + X^4 + X^7 + X^{15}$. Старшие шесть разрядов вектора инициализации генератора ПСП равны переменной UL_IDcell, остальные девять — константа. Номер кода определяется начальной точкой (т. е. числом тактов генератора ПСП после инициализации) — всего предусмотрено 256 кодов. Причем БС использует только часть из всех возможных кодов — сначала N кодов начальной инициализации, за ними следуют M кодов периодического определения параметров АС, далее L кодов запроса полосы. Для каждой БС задается точка начала этой последовательности кодов.

Начальная инициализация происходит так: АС, приняв дескриптор восходящего канала и UL-MAP, определяет набор CDMA-кодов и посылает в отведенном интервале случайно выбранный код из N возможных. Один и тот же код транслируется в двух последовательных OFDMA-символах. Если продолжительность канала конкурентного доступа составляет более одного слота, АС может отправить CDMA-код в четырех последовательных символах, причем коды должны быть смежными (т. е. следовать в ПСП один за другим).

Успешно приняв и распознав CDMA-код (а это может и не произойти, поскольку в интервале конкурентного доступа возможны коллизии при одновременной работе передатчиков нескольких АС), базовая станция не знает, от какой АС пришел запрос. Поэтому в ответ в UL-MAP следующего кадра она указывает номер принятого CDMA-кода, субканал и символ, в котором код был отправлен.

Так АС определяет, что ее запрос принят и следующее за UL-MAP широковещательное сообщение с указанием диапазона для запроса (номера символа, подканала и длительности) предназначено именно ей. В этом сообщении БС передает необходимые параметры для процесса инициализации в сети (включая идентификатор соединения CID, присвоенный MAC-адрес, набор физических параметров и др.). Далее в указанный в UL-MAP интервал АС приступает к штатной процедуре регистрации в сети.

Первичный запрос полосы в методе OFDMA может происходить двумя способами — посредством заголовков запроса полосы, как и в остальных режимах, и путем посылки CDMA-кода запроса полосы в интервале конкурентного доступа. Посылка кода запроса полосы (равно как и кода периодического измерения параметров) происходит в одном OFDMA-символе. Возможна и посылка трех последовательных кодов в трех символах (какой из вариантов необходимо использовать, указывается в UL-MAP). Приняв CDMA-код, БС в UL-MAP повторяет его номер и параметры, а также сообщает интервал для отправки заголовка запроса полосы уже обычным способом.

7.7. Поддержка адаптивных антенных систем

Важнейшая особенность стандарта IEEE 802.16, принципиально отличающая его, скажем, от стандартов IEEE 802 a/b/g, — это наличие встроенных средств поддержки адаптивных антенных систем (AAS). Подробно подобные системы описаны в монографии [3]. Разумеется, применение AAS — не обязательное требование стандарта. AAS — это системы с секторными направленными антеннами (метод формирования диаграмм направленности антенн в стандарте не оговаривается), т. е. антенные системы с несколькими антенными элементами. Применение AAS существенно увеличивает потенциальную емкость сети стандарта IEEE 802.16, поскольку в разных секторах БС возможна работа в одних и тех же каналах (частотных и OFDMA). Кроме того, направленные антенны позволяют существенно уменьшать общую излучаемую мощность. В результате снижается и межканальная интерференция. Не менее важно применение многоэлементных антенных систем для улучшения прохождения сигналов в каналах с замираниями, так называемых методов пространственно-временного кодирования (разнесения) STC.

Поддержка ASS в спецификации IEEE 802.16 означает модификацию протоколов на физическом и MAC-уровнях, наличие специальных управляющих и контролирующих сообщений для работы с адаптивными антеннами.

7.7.1. Работа с направленными AAS

Стандарт допускает в рамках одного кадра транслировать как ненаправленный, так и направленный (посредством AAS) трафик (рис. 7.22). Для разграничения зон не-AAS и AAS-трафика используются специальные сообщения. Принцип работы с AAS в режимах OFDM и OFDMA, равно как и в SCa, достаточно схож. Наиболее полно он описан в стандарте для случая OFDMA, поэтому остановимся именно на нем. В режиме AAS возможно два механизма назначения канальных ресурсов — сканирование карт (Diversity-Map Scan) и прямая сигнализация (Direct Signaling)

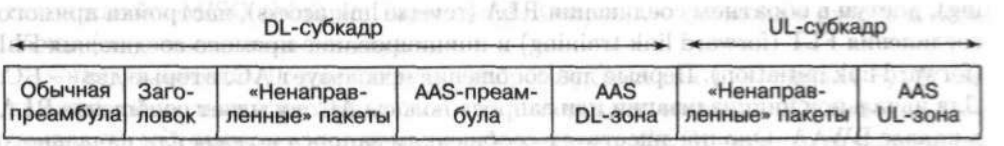


Рис. 7.22. Структура кадров с зоной AAS

Механизм Diversity-Map Scan

В режиме OFDMA предусмотрено два метода работы с AAS — с распределенными несущими (FUSC, PUSC) и с соседними несущими (AMC). Каждый из методов в начале AAS-зоны предусматривает передачу OFDMA-символа преамбулы AAS-зоны и заголовка с префиксом AAS-зоны. Для передачи этих сообщений в AAS-зоне нисходящего субкадра выделены специальные подканалы (два старших для FUSC/PUSC и четвертый с начала и четвертый с конца подканалы в AMC). Сообщения в этих подканалах могут повторяться несколько раз, с тем что, если используется не широковещательная трансляция, а передача с переключением лучей, сообщения с префиксом дошли бы до всех АС. В префиксе указывается код луча антенны, тип и размеры преамбулы ASS-зоны (в восходящем и нисходящем каналах), область для начальной инициализации/запросов полосы, а также области в кадре для каждого AAS-соединения. Префикс, как и в штатном режиме, передается посредством QPSK со скоростью кодирования 1/2 и двукратным повтором (в пределах одного символа). Основное назначение префикса — сообщить АС о том, как будут переданы карты UL/DL-каналов для разделенных по направлениям лучей групп пользователей (очевидно, что распределение канальных ресурсов может происходить независимо в каждом луче).

Для работы в режиме AMC-AAS кадры могут объединяться в суперкадр длительностью не менее 20 обычных кадров. В суперкадр входит по крайней мере один широковещательный кадр, содержащий дескрипторы и карты DL/UL-каналов. Смысл такого объединения — обеспечить минимум управляющих сообщений для группы кадров.

Memod Direct Signaling

Метод прямой сигнализации (Direct Signaling Method) использует механизм последовательного распределения несущих AMC. Прямая сигнализация требует точной пространственной селекции каналов, но позволяет увеличить емкость системы связи.

Особенность метода — в каждом кадре в AAS-зоне выделяется от одного до четырех каналов доступа/распределения ресурсов (BWAA — bandwidth allocation/access). Каждый BWAA-канал состоит из двух субканалов, расположенных в верхней и нижней частях диапазона симметрично относительно центральной частоты (если BWAA-канал один, то он включает самый верхний и самый нижний подканалы). В этом канале передаются префикс нисходящего субкадра (для режима Direct Signaling Method), карты UL-MAP и DL-MAP для каждой из пространственно разделенных АС или групп АС. Благодаря точной пространственной настройке AAS данный метод позволяет в одном кадре передавать сообщения множеству пользователей.

В методе прямой сигнализации предусмотрены четыре специальных кодовых сообщения — настройка (обучение) обратного соединения RLT (reverse link train-

ing), доступ в обратном соединении RLA (reverse link access), настройка прямого соединения FLT (forward link training) и инициирование прямого соединения FLI (forward link initiation). Первые два сообщения использует АС, вторые два — БС. Для начальной инициализации или запроса полосы АС посылает сообщение RLA в канале BWAA. Оно предшествует сообщениям запроса полосы или начального доступа и используется БС для точной настройки своей антенной системы на данную АС. В ответ БС передает сообщение FLI — уникальный код для каждой АС (БС может сама инициировать соединение, послав FLI). FLI транслируется в подканале, выделенном для данной АС. Каждая абонентская станция сканирует все подканалы и, обнаружив по кодовой последовательности адресованное ей сообщение начальной инициализации, отправляет в ответ в том же самом канале (в отведенном для нее временном интервале) последовательность RLT, предназначенную для точной настройки антенн БС на АС в данном подканале. В результате, выполнив все необходимые подстройки, БС и АС устанавливают соединение, в течение которого происходит обмен данными. Причем пакетам данных предшествуют настроечные последовательности FLT (со стороны БС) и RLT (со стороны АС).

7.7.2. Пространственно-временное кодирование

Еще одна важная особенность применения многоэлементных антенных систем — это возможность использовать пространственно-временное разнесение передающих каналов (Space-Time Coding, STC) для улучшения прохождения радиосигналов. Идея метода — разнести, пространственно и во времени, источник одного и того же сигнала, т. е. несколько изменить условия его прохождения. Вероятность безошибочного приема такого сигнала (после соответствующей первичной обработки в приемнике) существенно возрастает.

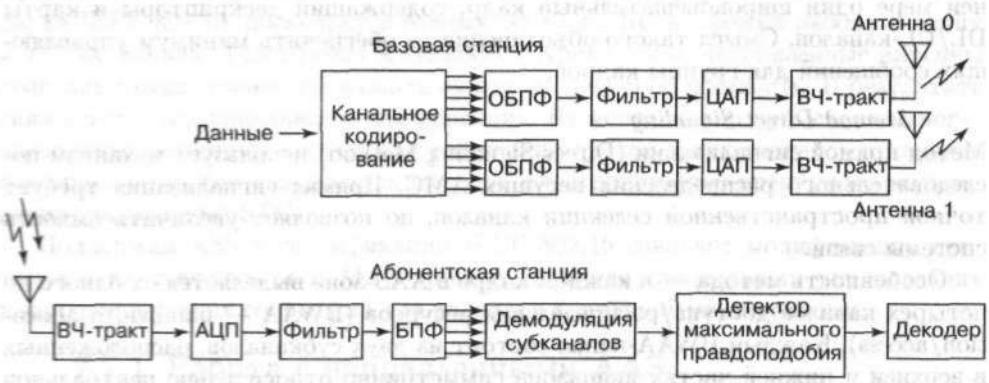


Рис. 7.23. Метод пространственно-временного кодирования по схеме MISO

В стандарте IEEE 802.16 используется схема прстранственно-временного разнесения, предложенная Аламоути [13]. Суть метода проста — выходной поток символов разбивается на два (например, четные и нечетные символы), формируемые параллельно (рис. 7.23). В передатчике используется два антенных канала, действующих параллельно и использующих общий тактовый генератор (что обеспечивает синхронность). Таким образом, реализуется так называемая

схема канала MISO (Multiple Input Single Output) — несколько входов и один выход (по отношению к каналу).

Сначала антенна 0 транслирует символ S_0, антенна 1 — символ S_1. В следующий символьный интервал антенна 0 передает символ $-S_1^*$, антенна 1 — символ S_0^* (S^* означает комплексное дополнение к S). Приемник работает с одной антенной и в каждом символьном интервале принимает сигналы r_0 и r_1. Зная передаточные характеристики каналов (h_0 и h_1), в приемнике можно восстановить переданные сигналы S_0 и S_1 согласно формулам (разумеется, вычисленные значения являются некоторым приближением к исходным значениям S_0 и S_1):

$$S_0 = h_0^* r_0 + h_1 r_1^*; \qquad S_1 = h_1^* r_0 - h_0 r_1^*.$$

С точки зрения протоколов физического уровня, применение STC не требует особых действий. Зона, транслируемая посредством STC, помечается в DL-MAP каждого кадра.

Особенности STC в методе OFDMA

В методе OFDMA предусмотрен ряд особенностей реализации STC. Прежде всего, при формировании OFDMA-символов набор пилотных частот зависит от четности символа и номера антенного канала. Кроме того, в OFDMA в дополнение к пространственно-временному разнесению используется и частотное посредством частотных скачков (frequency hopping diversity coding — FHDC). Суть данного механизма: допустим, несущие в субканале X модулируются сигнальным вектором S_0, в субканале $X+1$ — вектором S_1. Именно такой сигнал передает антенна 0. Антенна 1 транслирует сигнал, в котором несущие подканала X модулируются вектором $-S_1^*$, несущие подканала $X+1$ — вектором S_0^*. Восстановление в приемнике происходит аналогично рассмотренному варианту STC, только вместо передаточных характеристик двух антенных каналов используются характеристики, связанные с подканалами X и $X+1$ (т.е. с наборами несущих этих подканалов). Под принятыми сигналами r_0 и r_1 понимают принятые сигналы в подканалах X и $X+1$ соответственно. Из них восстанавливают S_0 и S_1.

Очевидно, что данную методику можно перенести на пары субканалов, т.е. все подканалы OFDMA-символа разбиваются на смежные пары (X, $X+1$; Y, $Y+1$; ...). В антенне 0 они передаются без изменений, в антенне 1 в каждой паре происходит описанное преобразование.

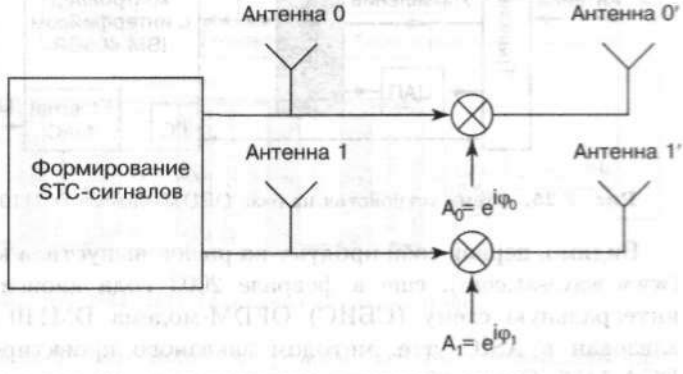


Рис. 7.24. Схема STC с четырьмя передающими антеннами

Все изложенные схемы преобразования можно описать матрицей:

$$\mathbf{A} = \begin{bmatrix} S_0 & -S_1^* \\ S_1 & S_0^* \end{bmatrix}.$$

Однако возможна и упрощенная схема: $\mathbf{B} = \begin{bmatrix} S_0 \\ S_1 \end{bmatrix}$, обеспечивающая, однако, двукратный выигрыш в скорости. Вид матрицы преобразования задается базовой станцией в картах соответствующих каналов.

Метод OFDMA допускает применение STC/FHDC не только в нисходящем, но и в восходящем канале. Кроме того, возможно применение STC на базе не только двух, но и четырех антенных элементов. В последнем случае помимо базовых антенн 0 и 1 (рис. 7.24) добавляются антенны 0′ и 1′, сигнал в которых смещен по фазе (например, сигнал в антенне 0′ $S_0' = S_0 \cdot e^{i\varphi_0}$).

7.8. Интегральная элементная база для устройств стандарта IEEE 802.16

Спецификация IEEE 802.16 предоставляет создателям аппаратуры достаточно широкие возможности, не оговаривая при этом конкретные способы реализации предусмотренных стандартом алгоритмов и механизмов. Фактически IEEE 802.16, как и положено стандарту, описывает самые общие правила игры, следуя которым возможно производить совместимую аппаратуру. И производители достаточно быстро отреагировали на появление нового тогда стандарта. Среди них такие компании, как Intel, Nokia, Analog Devices, Atheros Communications, Fujitsu Microelectronics America и многие другие. Рассмотрим лишь некоторые, самые первые решения в области элементной базы для реализации систем стандарта IEEE 802.16-2004 в низкочастотной области (менее 11 ГГц).

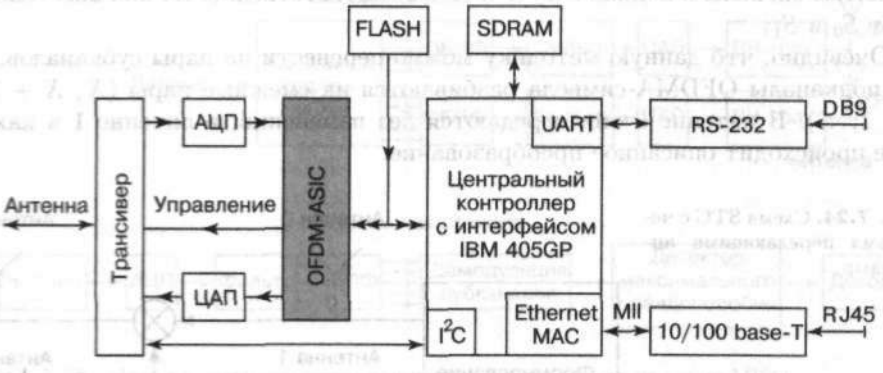


Рис. 7.25. Схема устройства на базе OFDM-модема DM110 компании Wavesat

Видимо, первой свой продукт на рынок выпустила канадская фирма Wavesat (www.wavesat.com), еще в феврале 2003 года анонсировавшая сверхбольшую интегральную схему (СБИС) OFDM-модема DM110 (рис. 7.25). Он был реализован в ASIC, т. е. методом заказного проектирования, в корпусе типа BGA-1156. Схема обладала возможностями, несколько отличающимися от тре-

бований IEEE 802.16 (в чем-то избыточными). Так, поддерживалась модуляция 4-, 16- и 64-QAM, 1024-точечное БПФ, временное и частотное дуплексирование, формирование OFDM-пакетов с защитным кодированием. Допустимая ширина канала — от 3 до 7 МГц, скорость передачи данных — до 35 Мбит/с. Напряжение питания ядра и периферии — 2,5 и 3 В, соответственно.

В декабре 2004 года фирма Wavesat объявила о начале продаж своего нового однокристального OFDM-модема — СБИС DM256. СБИС реализована в корпусе PQFP-208 и принципиально отличается от предшественницы. DM256 оснащена ЦАП и АЦП (10 разрядов). Входной/выходной интерфейс реализован в виде как квадратурных составляющих (I и Q), так и модулированного сигнала на промежуточной частоте 10 МГц. Поддерживается модуляция 2/4/16/64-QAM. В микросхеме реализованы разработанные компанией механизмы временной и частотной синхронизации, поддерживается временное и частотное разделение каналов, в последнем случае — дуплексный и полудуплексный режимы. Ширина канала — 1,75; 3; 7 и 10 МГц, длительность защитного интервала — от 1/4 до 1/32 от длительности OFDM-символа. На аппаратном уровне поддерживается кодек Рида–Соломона и декодер Витерби. Для построения оборудования на базе DM256 дополнительно необходим лишь ВЧ-трансмиттер и контроллер MAC-уровня. СБИС может использоваться как в БС, так и в абонентском оборудовании. Важно отметить, что DM256 поддерживает требования стандарта широкополосного доступа для мобильных приложений IEEE 802.16e (в режиме OFDM).

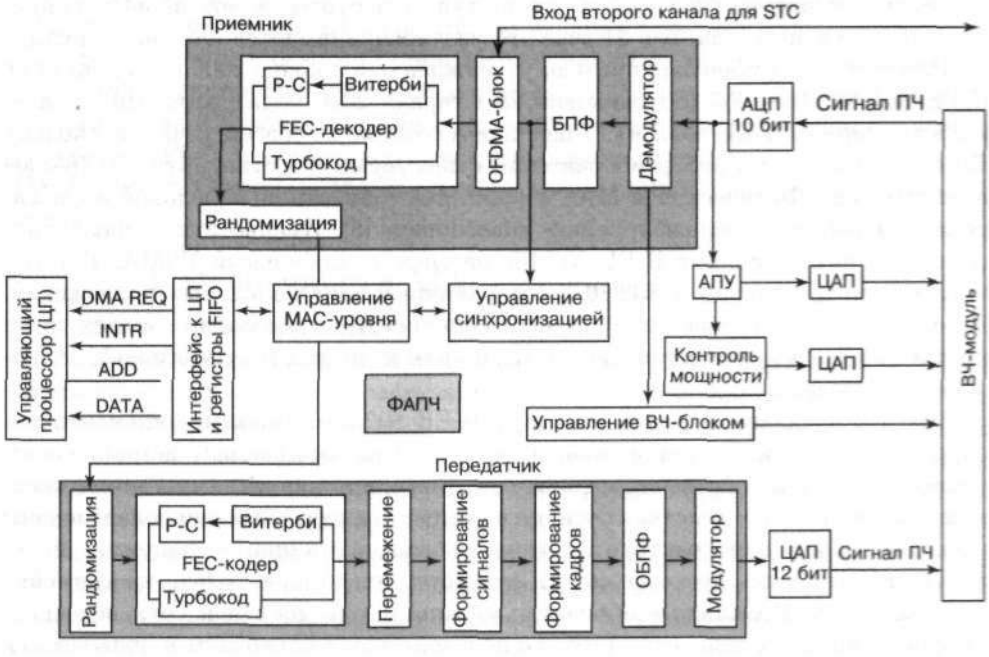


Рис. 7.26. Структура СБИС RN-2234 OFDMA-модема компании Runcom

Если компания Wavesat сосредоточилась на режиме OFDM, то израильская фирма Runcom Technologies (www.runcom.com) выпустила СБИС RN-2234 — мо-

дем с поддержкой режима OFDMA стандарта IEEE 802.16 (рис. 7.26). Данная СБИС реализует схемы модуляции QPSK, 16- и 64-QAM. Поддерживаются как турбокоды, так и кодек Рида – Соломона (Р–С). Модем включает АЦП (10 бит) и ЦАП (12 бит). При ширине канала 14 МГц СБИС RN-2234 обеспечивает скорость передачи данных до 56 Мбит/с. На аппаратном уровне поддержана возможность работы с AAS, включая режим STC в нисходящем канале. СБИС производится по КМОП-технологии уровня 0,18 мкм в корпусе 304 LF BGA ($19 \times 19 \times 1{,}5$ мм). Напряжение питания ядра/периферии — 1,8/3,3 В. Максимальная потребляемая мощность — 2 Вт.

7.9. Широкополосный мобильный доступ под управлением стандарта IEEE 802.16e

В 2004 году появился стандарт беспроводного широкополосного доступа IEEE 802.16-2004 [4]. Тогда казалось, что мир вот-вот шагнет в новую реальность, где пользователю практически в любой точке Земли будут доступны средства высокоскоростного информационного обмена — от передачи данных до телефонной связи и телевидения. Но вскоре выяснилось, что стандарт — это еще не все. Необходимо выделение частотного ресурса, построение инфраструктуры сетей, немалые усилия по интеграции уже существующих служб, в конце концов — привлечение к новой технологии создателей контента для конечных пользователей. Однако главная проблема заключалась в том, что стандарт IEEE 802.16-2004 был ориентирован на фиксированный доступ — местоположение абонента определялось раз и навсегда. А в 21 веке это уже воспринимается как анахронизм.

Изменить ситуацию был призван утвержденный в конце 2005 года документ IEEE Std 802.16e-2005 (опубликован 28 февраля 2006 года), который называют «стандартом IEEE 802.16e». Такое наименование не совсем точно, поскольку IEEE 802.16e — это набор исправлений существующего стандарта 802.16-2004 и дополнения «Физический и MAC-уровни для совместной мобильной и фиксированной работы в лицензируемых диапазонах» [5]. Именно эти «дополнения» (из-за которых стандарт IEEE 802.16e называют «мобильный WiMAX») и открывают путь стандарту 802.16 в безграничный мир мобильных приложений. В результате он становится серьезным конкурентом технологий сотовой связи третьего и последующих поколений, равно как и других перспективных технологий беспроводного доступа.

Что же предлагает новый стандарт IEEE 802.16e? Понятие «мобильность» относят к двум категориям абонентов — к так называемым номадическим («кочующим») и к собственно подвижным. Номадические абоненты могут перемещаться в пределах действия сети, но в момент сеансов связи они локализованы (находятся в зоне одного и того же сегмента базовой станции) — например, пользователи ноутбуков, которые могут включить их дома, в офисе, на скамейке в парке и т. п. Подвижные абоненты должны иметь доступ к сети непосредственно в процессе движения (тот же пользователь с ноутбуком в движущемся автомобиле). Если для номадических абонентов важна быстрая регистрация в любой точке сети (в идеале — сети любого провайдера), то обеспечить подлинную подвижность гораздо сложнее. Прежде всего, необходимы процедуры передачи абонента от одной базовой станции (БС) к другой (или между различ-

ными сегментами одной БС) так, чтобы сам абонент этого не ощущал. Это — функции так называемой эстафетной передачи (хэндовер).

Кроме того, мобильность абонентов диктует совершенно иные требования к управлению ресурсами сети и к возможности их оперативного перераспределения. Ужесточаются и требования к вторичному использованию частотного ресурса сети. Именно поэтому в новой редакции стандарта значительное внимание уделено возможности пропорционального уменьшения частотной полосы канала, а также технологиям многоканальных антенных систем (MIMO). Для мобильных устройств очень важно снизить энергопотребление, чему способствуют специальные режимы и процедуры нового стандарта.

Помимо собственно мобильности особое внимание IEEE 802.16e уделяет проблемам качества предоставляемых услуг (QoS). Ведь IEEE 802.16 рассматривается как стандарт для предоставления услуг операторского класса, в том числе — и для мобильных абонентов. Поэтому вопрос QoS для этой технологии играет первостепенную роль.

Кроме того, мобильность автоматически подразумевает усложнение сетевой архитектуры. Если при фиксированном доступе абонентская станция общается с единственной назначенной ей БС, то мобильная абонентская станция (МС) должна знать свое окружение, общаться одновременно с несколькими БС, переключаться с одной на другую и т. п. Эти требования обусловили появление в стандарте IEEE 802.16e понятий «сервисной БС» и «соседней БС». Сервисная БС для определенной МС — это базовая станция, на которой МС последний раз выполнила процедуру регистрации, при начальном вхождении в сеть или при хэндовере. С сервисной БС абонентская станция работает в обычном режиме. Соседняя БС — это базовая станция, отличная от сервисной, трансляцию с которой (нисходящий поток) способна принять МС.

Изменения и дополнения в документе 802.15e затронули функции физического уровня (Phy), а также подуровня управления контролем доступа к каналу (Media Access Control — MAC) сетевого уровня MBOC. Рассмотрим эти изменения.

7.9.1. Особенности MAC-уровня

Качество обслуживания

На MAC-уровне нововведения и изменения связаны с QoS. Понятие «соединение» заменено на «транспортное соединение». Сервисный поток (со всеми его свойствами) определяется не для всей сети, а только для обмена между конкретной парой БС–АС. Особо отмечено, что каждому сервисному потоку с идентификатором SFID ставится в соответствие единственное транспортное соединение с уникальным идентификатором CID.

Поскольку мобильность предполагает миграцию абонента между различными сетями, вводится понятие «глобальный сервисный класс». От существовавшего понятия сервисного класса его отличает то, что имя глобального сервисного класса остается единым и постоянным для всех БС, и никакая отдельная БС не может его изменить. Таким образом, глобальный сервисный класс — это инструмент управления QoS в рамках глобальной сети и/или объединения нескольких сетей. Имя глобального сервисного класса представляет собой набор из восьми параметров (плюс один резервный) длиной 32 бита (табл. 7.5).

Таблица 7.5. Формат имени глобального сервисного класса

Позиция	Название	Размер, бит
I	Признак восходящего/нисходящего потока	1
S	Максимальная скорость непрерывного трафика (1200–1921000 бит/с)	6
T	Признак привилегированного трафика	1
B	Максимальный размер пакета (1200–1921000 бит)	6
R	Минимальная резервированная скорость	6
L	Максимальная задержка (1 мс – 10 с)	6
S	Признак фиксированной/переменной длины пакетов	1
P	Признак возможности передать мобильной станции пейджинговое сообщение в режиме ожидания	1
R	Резерв	4

Помимо сервисных классов новый стандарт вводит понятие типов служб доставки данных. В отличие от сервисных классов, тип службы доставки не подразумевает присвоения параметрам соединения каких-либо значений, а лишь обозначает список нормируемых для каждой службы параметров. Названия служб ассоциируются с типом планирования запросов на предоставление ресурсов, более того, у восходящих соединений их названия совпадают. Всего предусмотрено пять типов служб доставки:

- доставка без требования (Unsolicited Grant Service — UGS);
- доставка в реальном времени с переменной скоростью (RT-VR);
- доставка в реальном времени с переменной скоростью и расширенными возможностями (ERT-VR);
- доставка вне реального времени с переменной скоростью (NRT-VR);
- доставка по мере возможности (Best Efforts — BE).

Служба доставки без требования UGS предполагает, что оговоренные ресурсы предоставляются на периодической основе. Она предназначена для приложений реального времени, транслирующих данные с известной фиксированной скоростью. Причем размеры MAC-пакетов могут быть различными. Для UGS нормируются такие параметры, как толерантность к джиттеру, размер блоков данных (если они фиксированы), минимальная гарантированная скорость передачи, максимальная задержка и интервал между сеансами передачи.

Служба RT-VR рассчитана на приложения реального времени, которые требуют передачи данных с гарантированными скоростью и временем задержки. Эта служба предоставляется по запросу, для чего вводится параметр — период запросов. БС регулярно (в соответствии с периодом запросов) выделяет в восходящем канале специальный интервал для запроса дополнительного канального ресурса от конкретной МС. То есть приложению гарантируется не сам требуемый ресурс, а возможность его запросить.

Служба NRT-VR, как и следует из ее названия, необходима для передачи данных с заданной скоростью, но с произвольной задержкой. Для этой службы нормируется минимальная гарантированная скорость передачи данных. Вероятность предоставления запрошенного ресурса зависит от приоритета трафика — от 0 (низший) до 7 (высший). Причем, в отличие от службы реального времени, запрос производится на конкурентной основе.

Служба BE подразумевает остаточный принцип предоставления ресурса. В ней определятся только приоритет трафика.

Служба реального времени с расширенными возможностями ERT-VR — это комбинация служб UGS и RT-VR. Типичные ее задачи — высокоприоритетные приложения, требующие гарантированных значений скорости передачи и времени задержки, но характеризующиеся переменной скоростью — например, IP-телефония. Служба ERT-VR, как и UGS, предоставляется без запроса (по расписанию, через заданный интервал), но использует параметр «приоритет трафика».

Спящий режим и ожидание

Для экономии энергии в мобильных станциях, стандарт IEEE 802.16e предусматривает два режима энергосбережения — спящий режим и режим ожидания.

Спящий режим (sleep mode) означает, что МС остается зарегистрированной в сети (в списке абонентов какой-либо БС), но недоступна для приема/передачи. Этот режим не только снижает энергопотребление МС, но и экономит канальные ресурсы базовой станции — «спящую» МС не надо обслуживать. Спящий режим обязателен для поддержки каждой БС и опционален для МС.

Спящий режим МС представляет собой последовательность интервалов сна и прослушивания. Параметры спящего режима связаны с типами соединений (и служб доставки данных), поддерживаемыми МС. Инициация спящего режима, его прекращение, а также определение типа и параметров происходят только по командам и под управлением БС. МС может запросить у БС переход в спящий режим посредством специального сообщения. Переход в этот режим происходит после получения от БС ответного сообщения с подтверждением.

Совокупность параметров спящего режима образует класс энергосбережения. Этот класс присваивается одному или группе соединений, имеющих сходные параметры запроса ресурса. Например, соединения типа BE или NRT-VR могут принадлежать одному классу энергосбережения. В то же время, два соединения типа UGS с различными требованиями к канальным ресурсам принадлежат к различным классам. Если МС поддерживает только соединения, не связанные ни с каким классом энергосбережения, такая МС постоянно активна.

Каждая МС может одновременно поддерживать соединения с несколькими классами энергосбережения. Совокупность окон сна/прослушивания различных классов образует для каждой МС интервалы доступности/недоступности (рис. 7.27). В интервале недоступности БС не может связываться с МС. Поэтому в этот период в мобильной станции могут отключаться функциональные блоки приема и обработки данных от БС. Если БС получает адресованные МС данные, когда последняя находится в интервале недоступности, БС может либо буферизировать их для последующей передачи, либо сбросить.

Выход из спящего режима возможен по специальным управляющим сообщениям от БС, по определенным событиям (например, детектирование порогового значения мощности сигнала от БС) или по расписанию. В спящем режиме допустимы определенные периодические процедуры, например ранжирование (определение условий в канале связи и соответствующих параметров сигнала — времени задержки, мощности излучения и т. п.).

Стандарт определяет три типа классов энергосбережения.

Классы энергосбережения типа I рекомендованы для соединений BE и NRT-VR. Особенность этого типа — постоянно удваивающаяся длительность окна сна

SleepWin, но до определенного предела $B \cdot 2^E$:

$$SleepWin(i+1) = \min(2 \cdot SleepWin(i),\ B \cdot 2^E).$$

Параметры начального значения окна сна $SleepWin(0)$, а также B, E, длительность окна прослушивания и номер фрейма, с которого начинается окно сна, базовая станция передает, когда задает класс энергосбережения. Также сообщаются специальные флаги событий, по которым возможно внеочередное «пробуждение».

Рис. 7.27. Пример работы MC с двумя классами энергосбережения

В активном режиме энергосбережения типа I MC не способна передавать полезные данные или делать запросы на предоставление дополнительных канальных ресурсов. Однако во время интервалов прослушивания возможен прием всего нисходящего трафика, как в обычном режиме.

Классы энергосбережения типа II предназначены для соединений UGS и RT-VR. Для этого типа классов энергосбережения размер окон сна и прослушивания фиксирован и постоянен. Данные параметры, а также номер фрейма, с которого начинается последовательность окон сна/прослушивания, передаются базовой станцией. В отличие от типа I, в окне прослушивания MC может как принимать, так и передавать полезные данные (в рамках соединений, связанных этим типом классов энергосбережения).

Классы энергосбережения типа III служат для многоадресных соединений или для таких служебных процедур, как периодическое ранжирование, динамическое изменение услуг и т. п. Этот тип подразумевает наличие всего одного окна сна (задается параметрами B и E как $B \cdot 2^E$). По его завершении MC автоматически выходит из спящего режима.

Например, если БС знает периодичность поступления данных для многоадресной рассылки, то на весь этот период (когда данные отсутствуют) БС может назначить MC спящий режим типа III. Данные передаются после его завершения, а затем БС снова активирует спящий режим. Аналогичный подход применим и для прочих периодических процедур.

Режим ожидания (пейджинг) (idle mode) — опциональная функция. В стандарте IEEE 802.16e этот механизм делает MC периодически доступными для широковещательного нисходящего трафика без регистрации в какой-либо кон-

кретной БС. Например, при перемещении МС по географически протяженному району со множеством БС мобильная станция может не передавать информацию и оставаться пассивной, пока БС не вызовет ее специальным сообщением о поступлении трафика в адрес данной МС. Достоинство такого режима — МС не нужно тратить ресурсы и энергию на поддержку процедур хэндовера, сканирования БС через дискретные промежутки времени и т. п. Ждущий режим предоставляет простой механизм для предупреждения МС о направляемом ей нисходящем трафике, причем без дополнительной активности со стороны МС. Поскольку ждущий режим — это фактически система передачи коротких сообщений для МС, местоположение которой точно не известно, его еще называют пейджингом.

Для реализации ждущего режима все БС логически подразделяются на пейджинговые группы (рис. 7.28), обеспечивающие радиопокрытие определенного региона. Пейджинговая группа должна быть достаточно обширной, чтобы большинство МС как можно дольше оставались в ее пределах, и достаточно малой, чтобы перекрытие таких групп было эффективным (не излишним).

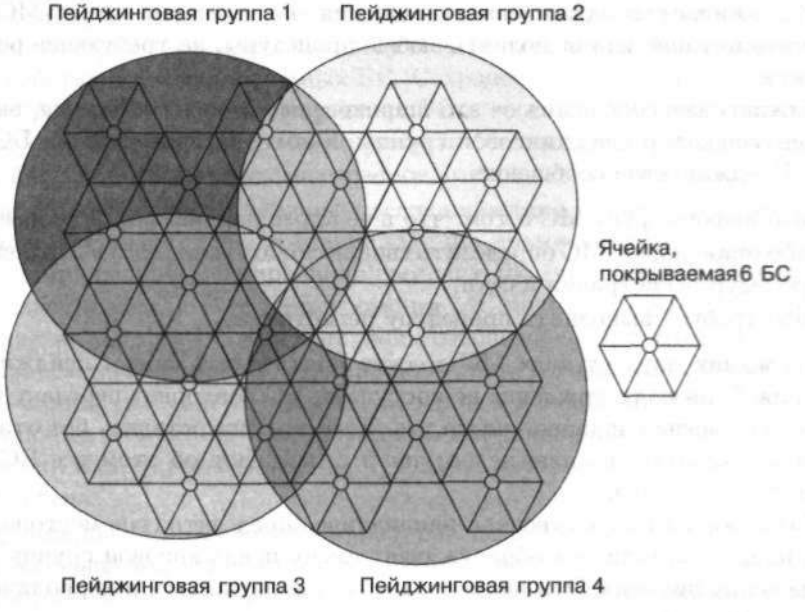


Рис. 7.28. Пейджинговые группы

Переход в ждущий режим и работа в нем включает несколько шагов: инициацию ждущего режима, выбор ячейки, синхронизацию с БС времени пейджинговых сообщений, установление интервалов для прослушивания (для МС) и передачи (для БС) пейджинговых сообщений, собственно передачу пейджинговых сообщений и прекращение режима пейджинга. Инициация режима возможна после того, как МС передаст специальное сообщение о прекращении регистрации на определенной БС и переходе в ждущий режим. БС также может послать МС команду о переходе в ждущий режим. После обмена подтверждениями МС переходит в ждущий режим. Однако МС может потребовать у БС (на которой

была зарегистрирована) или у контроллера пейджинга (БС с дополнительными функциями) хранить ее сетевые настройки в течение некоторого времени. Это необходимо для быстрого перехода из ждущего режима к нормальной работе. Если по истечении определенного времени БС не обнаруживает в своей зоне МС или мобильная станция остается пассивной (в ждущем режиме), то БС удаляет такую МС из списка зарегистрированных станций.

При переходе в режим ожидания МС должна выбрать БС с наилучшими условиями в радиоканале и зарегистрироваться в ней. После чего МС необходимо синхронизироваться с выбранной БС, принять от нее управляющие сообщения (дескриптор нисходящего канала DCD и карту нисходящего канала DL-MAP), извлечь из этих сообщений размер и номер текущего фрейма и определить время до следующего пейджингового интервала. Пейджинговый интервал определяется длительностью (числом кадров), периодом повторения C, а также смещением P_offset. Номер начального фрейма пейджингового интервала в каждом цикле N_f должен удовлетворять условию $N_f \mod C = P\_offset$, где оператор $x \mod y$ — остаток от деления x на y.

Пейджинговый интервал — это период прослушивания, в течение которого МС ожидает пейджинговые сообщения. В остальное время МС может выключить питание или выполнять любые процедуры, не требующие регистрации в сети.

Пейджинговые сообщения — это широковещательные сообщения, включающие идентификатор пейджинговой группы, к которой принадлежит БС-отправитель. Пейджинговое сообщение:

- либо информирует МС о том, что в ее адрес направляются данные,
- либо принуждает МС определить свое местоположение без выполнения всех процедур регистрации в сети,
- либо требует выполнить процедуру регистрации.

В последних двух случаях МС должна подтвердить прием пейджингового сообщения. Если подтверждение не поступает, БС повторяет передачу пейджиногового сообщения заданное число раз. Если все они остались без ответа, БС считает мобильную станцию недоступной и извещает об этом все БС данной пейджинговой группы.

МС способна и по собственной инициативе определять свое местоположение в сети, например, если она обнаруживает смену пейджинговой группы или после включения питания. Эта процедура может выполняться и периодически, по специальному таймеру.

Все функции поддержки режима пейджинга в сети реализует специальная базовая станция – контроллер пейджинга. Она рассылает всем БС специальные сообщения со списками МС, находящихся в режиме ожидания, сохраняет в течение определенного интервала сведения о МС для ее быстрого возвращения в режим нормальной работы, а также извещает БС о повторной регистрации МС в другой точке сети.

Хэндовер

Хэндовер (эстафетная передача, переход МС от одной БС к другой) необходим в ряде ситуаций. Наиболее типичные — при движении МС условия в радиоканале с текущей БС стали ниже допустимых и/или хуже, чем с соседней БС. Либо

текущая БС не обеспечивает заданные параметры QoS и ее необходимо сменить на соседнюю. Однако прежде чем МС сможет переключиться от одной БС к другой, ей нужно найти и выбрать новую базовую станцию. Чтобы упростить МС поиск соседних БС, каждая базовая станция, поддерживающая мобильность, периодически рассылает специальное сообщение — анонс соседей (Neighbor Advertisement). В этом сообщении перечислены все соседние БС и их профили, включая режимы работы (например, OFDM или OFDMA, размер БПФ, ширина полосы, номер канала, частотная литера, эквивалентная мощность в антенне и т. п.), особенности процедуры хэндовера для каждой БС, вид поддерживаемых классов сервиса и т. д.

Сервисная БС по запросу от МС назначает ей специальные временные интервалы, в течение которых МС анализирует обстановку на предмет поиска соседних БС для хэндовера. Такие интервалы называют интервалами сканирования. Интервалы сканирования могут перемежаться с интервалами нормальной работы. При запросе интервалов сканирования МС может перечислить базовые станции (например, из списка, сообщенного сервисной БС), условия связи с которыми она будет анализировать.

При сканировании возможна процедура ассоциирования. Это опциональная функция, которая позволяет МС получить и сохранить физические параметры и свойства QoS опрашиваемых БС. Сохранение этих параметров упрощает и ускоряет хэндовер. Причем сервисная БС в сообщении о предоставлении интервалов сканирования сама может указать базовые станции, с которыми МС должна ассоциироваться (так называемое прямое ассоциирование).

Стандарт определяет три уровня ассоциирования. Уровень 0 — это обычное (некоординированное) ранжирование. МС во время интервалов сканирования выполняет процедуру ранжирования соседних БС (указанных в запросе на интервал сканирования). При ранжировании МС посылает специальное тестовое сообщение, на которое БС отвечает своим сообщением. При этом МС определяет отношение сигнал/шум в канале, мощность принятого сигнала, время задержки и т. п. Запросы базовым станциям посылаются в заданные интервалы опроса, но на конкурентной основе (как при начальном ранжировании при установлении соединения). Поэтому нет гарантии, что такой запрос дойдет до конкретной БС с первого раза.

Координированное ассоциирование (уровень 1) подразумевает участие в сканировании сервисной базовой станции. Этот уровень может быть реализован как по запросу от МС, так и назначен самой БС. Сервисная БС запрашивает соседние БС об удобном для них времени проведения ранжирования. В ответ соседние БС передают сервисной БС уникальный CDMA-код для запроса ранжирования (для режима OFDMA) и условия передачи запроса (фактически, номер фрейма, в котором конкретная БС передаст карту восходящего канала UL-MAP, где указан нужный момент начала запроса). Эти условия и CDMA-коды сервисная БС сообщает МС, и та уже сама в заданное время общается с соседними БС.

Ассоциирование с уведомлением по сети (уровень 2) похоже на координированное ассоциирование. Но в отличие от уровня 1, МС при ранжировании достаточно передать только CDMA-код и не нужно дожидаться ответа от соседней БС. Все соседние БС, получив запрос, передают информацию о физических параметрах (таких, как параметры для коррекции времени отправки сообще-

ния, частоты, уровня мощности передатчика МС, доступные уровни QoS и др.) сервисной БС по магистральной межстанционной сети. Сервисная БС собирает эти данные и передает их МС в одном сообщении.

Сканирование и ассоциирование — это процедуры, предшествующие хэндоверу. Они позволяют сформировать список соседних БС и выбрать из них одну для предстоящего переключения. Собственно процедура хэндовера включает несколько стадий:

- выбор ячейки (на основе непосредственного сканирования или ассоциирования);
- решение о начале хэндовера и инициация этой процедуры;
- синхронизация с выбранной БС;
- установление соединения (регистрация);
- разрыв соединения с предыдущей сервисной БС.

Завершение хэндовера МС подтверждает специальным сообщением. МС может прервать процедуру хэндовера в любой момент до отправки финального сообщения.

Важно отметить, что решение о начале процедуры хэндовера способна принять МС, сервисная БС или система управления сетью. Базовая станция может инициировать хэндовер на основе оценки необходимых МС канальных ресурсов и QoS. Эти требования МС соотносятся с возможностями сервисной и соседних БС. Если собственных ресурсов не хватает, а у соседней БС они есть, сервисная БС заставляет МС переключиться на соседнюю БС. Сервисная БС знает о возможностях своих соседей благодаря информационному обмену по магистральной межстанционной сети. Причем БС может предоставить МС такую информацию в ответ на ее запрос о хэндовере.

При установлении соединения с выбранной БС выполняется либо процедура начальной регистрации в сети, либо сокращенная процедура регистрации при хэндовере. Упростить регистрацию можно благодаря тому, что сервисная БС по магистральной сети передает выбранной БС информацию о МС, устанавливающей с ней соединение. При этом удается опустить такие этапы начальной регистрации, как аутентификация, определение ключа для криптозащиты трафика, сам запрос на регистрацию и др.

Мы описали базовую процедуру хэндовера. Однако стандарт предусматривает две опциональные функции — так называемый макро-диверсифицированный хэндовер (Macro diversity handover — MDHO) и быстрое переключение между базовыми станциями (Fast Base Station Switching — FBSS). Все БС, поддерживающие режимы MDHO и FBSS, должны синхронизироваться на основе единого опорного источника и работать в одном частотном диапазоне.

В режиме MDHO МС может одновременно работать с несколькими БС (диверсификация). Каждая из них должна передавать МС одинаковые пакеты. Для этого формируется список БС, способных поддерживать такой обмен с заданной МС, — лист диверсификации (Diversity Set). У каждой БС из одного списка есть вся информация о МС, которую БС получает при начальной регистрации МС (включая аутентификацию). То есть, если МС зарегистрирована на одной БС, она автоматически регистрируется на всех БС, включенных в список диверсификации.

Среди БС из списка диверсификации одна станция назначается анкерной. Для обмена с несколькими БС мобильная станция должна постоянно принимать управляющую информацию (как минимум, карты нисходящего/восходящего каналов и управляющие заголовки). В режиме MDHO это возможно двумя способами.

Первый метод — МС следит только за управляющей информацией анкерной БС. В этом случае сообщения DL- и UL-MAP анкерной БС должны содержать данные о расположении пакетов, адресованных данной МС от всех БС из листа диверсификации.

Второй метод подразумевает, что МС контролирует управляющую информацию всех активных БС из листа диверсификации. Но тогда в картах DL/UL-MAP каждой из них должна быть информация о расположении пакетов других БС. Базовые станции, поддерживающие режим MDHO и входящие в один лист диверсификации, должны работать с общим набором идентификаторов соединений CID. Отметим, что описанный базовый хэндовер — это частный случай MDHO, когда в лист диверсификации входит всего одна БС.

Режим FBSS отличается от MDHO тем, что МС может работать одновременно только с одной БС из списка диверсификации — с анкерной БС. Анкерной станцией может быть любая БС из списка диверсификации. Смена анкерных станций в режиме FBSS (переключение) происходит очень быстро, поскольку для этого не требуется каких-либо процедур синхронизации и регистрации. Для смены анкерной БС мобильная станция использует либо специальное управляющее сообщение, либо короткое информационное сообщение по специальному каналу быстрой обратной связи.

Групповая и широковещательная передача

Дополнения 802.16e впервые вводят в стандарт IEEE 802.16 понятия групповой (многоточечной, multicast) и широковещательной (broadcast) передачи. Отличия между ними незначительны: в первом случае информация адресована группе МС, во втором — всем МС. Поэтому в стандарте эти понятия практически не разделены, используется единый термин Multicast and Broadcast Service (MBS) — услуга групповой и широковещательной передачи. Сеть может предоставлять сервис MBS в двух режимах — вещание с одной БС или с нескольких БС (групповой MBS). Мобильные станции должны поддерживать оба этих режима. Механизм автоматического повтора отправки не дошедших пакетов (ARQ) при широковещательной передаче не предусмотрен.

Транспортным потокам MBS с заранее заданными свойствами (трафика и QoS) присваивают идентификаторы сервисных потоков и соответствующие CID. МС распознает MBS по этим идентификаторам. Очевидно, что сообщения групповой передачи некоторые МС могут игнорировать (если у них данный CID не определен), в то время как CID широковещательных сообщений должны поддерживать все МС. Важно, что MBS-сообщения МС могут принимать даже в спящем режиме или в режиме ожидания.

В режиме группового MBS несколько БС образуют специальную группу, всем БС которой присваивается идентификатор зоны MBS. В пределах этой зоны БС и МС должны пользоваться едиными CID для работы в режимах MBS. Все БС этой группы синхронно транслируют MBS-сообщения на одном частотном канале и с одинаковыми идентификаторами. Такой режим повышает вероятность приема широковещательных сообщений.

Для того чтобы получать групповые сообщения, МС должна зарегистрироваться как адресат такого сервиса. Но, в отличие от режима широкого вещания с одной БС, при групповом MBS-режиме регистрация на БС, от которой МС получает широковещательное сообщение, не требуется. Более того, МС может быть не зарегистрированной ни в одной БС группы. Достаточно, чтобы она была зарегистрирована в сети как получатель данного сервиса. Информация о МС предается базовым станциям посредством сетевой инфраструктуры. Очевидно, что при этом собственно процедура регистрации получателя MBS-сообщений реализуется на сетевом уровне, т. е. вне рамок стандарта IEEE 802.16, — это уже забота сетевой группы консорциума WiMAX.

7.9.2. Особенности на физическом уровне

На физическом уровне в документе IEEE 802.16e [5] не слишком много отличий от IEEE 802.16-2004 [4], но они весьма значимы. Если не вдаваться в технические детали, то суть изменений на физическом уровне — обеспечить большую гибкость для работы в полосах частот различной ширины. Фактически речь идет о максимально эффективном использовании частотного ресурса. Это очень важно для обеспечения мобильности абонентов, поскольку мобильному оператору вряд ли удастся получить частотную полосу шириной 20 МГц. Непосредственные изменения в стандарте коснулись только двух режимов — OFDM и OFDMA. Напомним, что OFDM — это метод модуляции, в то время как OFDMA (OFDM Access) подразумевает использование OFDM не только для модуляции, но и для мультиплексирования каналов (множественного доступа) [2, 6].

Режим OFDM

В режиме OFDM уточнен принцип применения субканалов. Напомним, что суть режима OFDM — использование модуляции OFDM (на основе 256 номинальных поднесущих) во всей полосе частот. Каналы мультиплексируются на основе временного разделения (TDMA), а дуплексирование (разделение приемного и передающего каналов) может быть как временным (TDD), так и частотным (FDD). Но изначально разработчики стандарта предусмотрели возможность использовать так называемые субканалы. Субканал образуют 12 информационных OFDM-поднесущих, всего возможно до 16 субканалов. Причем они могут объединяться. То есть в режиме OFDM предусмотрена опциональная возможность разделять общую полосу частот — некое приближение к OFDMA.

Однако об использовании этого режима в тексте стандарта IEEE 802.16-2004 сказано очень нечетко. Приведены правила канального кодирования при работе с субканалами, но и только. Кроме того, субканальный режим в соответствии с IEEE 802.16-2004 — это опциональная функция для работы только в восходящем канале, от абонентской к базовой станции. Возможность назначения субканала была предусмотрена в управляющем сообщении, входящем в карту восходящего канала UL-MAP. Но ни в одном управляющем сообщении для нисходящего канала выбор субканалов не оговаривался. Почему — загадка: либо предполагалось, что выбор субканалов будет происходить до начала работы, путем начальных установок, либо вообще не предусматривалась такая возможность (например, считали, что субканальный режим будет использоваться только для упрощения начального подключения АС к сети). А может, просто не было ясного понима-

ния и/или единства мнений по этому вопросу, поэтому данную опцию просто обозначили.

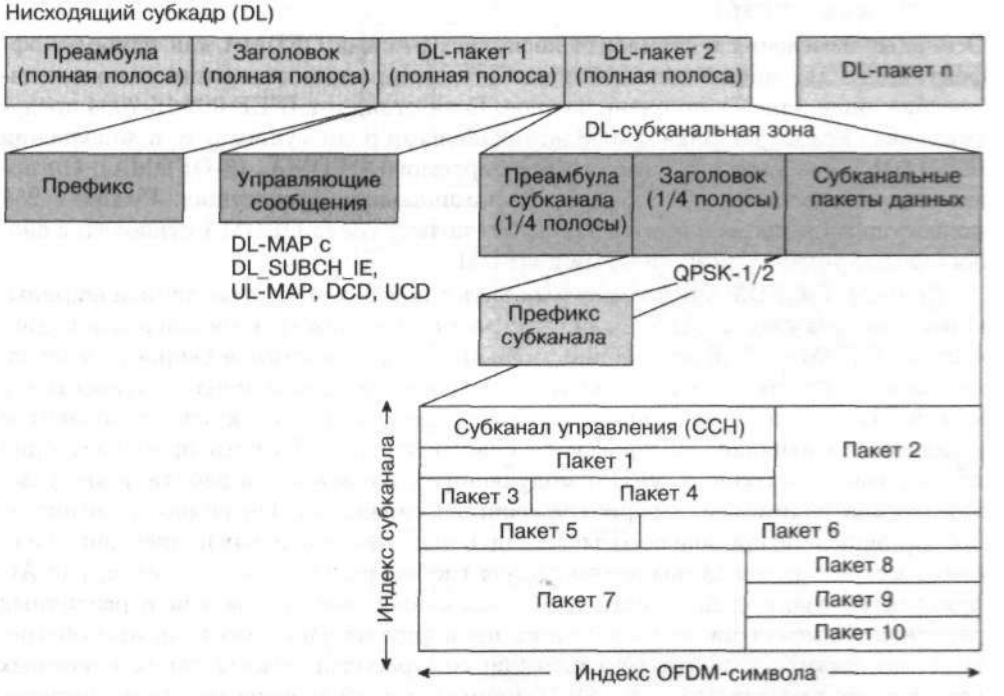


Рис. 7.29. Структура субканальной зоны в нисходящем субкадре в режиме OFDM

В стандарте IEEE 802.15e правила работы в субканальном режиме сформулированы более четко. Причем опциональное использование этой возможности предусмотрено и в нисходящем канале. В документе показано, что эта опция — некое упрощенное приближение к режиму мультиплексирования посредством OFDM, т. е. к режиму OFDMA. В соответствии с IEEE 802.15e, в нисходящем канале можно создавать субканальную зону (рис. 7.29), в которой одновременно передается несколько пакетов, адресованных различным АС. Начало этой зоны обозначают специальные пакеты — преамбула субканальной зоны и заголовок субканальных пакетов (FCH), которые передаются в самом медленном режиме — с модуляцией QPSK и скоростью кодирования 1/2. Заголовок субканальных пакетов должен занимать четыре субканала. Номер группы этих субканалов по умолчанию определяется двумя младшими битами идентификатора БС. Заголовок содержит префикс субканальной зоны, в котором указываются параметры субканала управления (CCH — control subchannel): размер первого пакета в CCH, скорость его передачи, частота повторения CCH (один раз в субканальной зоне либо через каждые 4/8/16 OFDM-символов) и др. CCH — это область, в которой передается управляющая информация для работы в субканальной зоне (субканальные карты нисходящего и восходящего каналов, на основе которых АС определяют назначенные им пакеты). «Субканализированный» трафик следует сразу за заголовком. Само же начало субканальной зоны в восходящем

и нисходящем каналах задается в картах этих каналов (DL-MAP и UL-MAP) управляющими сообщениями (например, DL_SUBCH_IE для нисходящего канала).

Режим OFDMA

Основные изменения в стандарте коснулись режима OFDMA как наиболее эффективного для мобильного доступа. Собственно, принципиально значимое изменение лишь одно — но очень важное. Если стандарт IEEE 802.16-2004 предусматривал режим OFDMA с 2048 номинальными поднесущими, то в дополнении IEEE 802.16e появился режим «масштабируемого OFDMA» (S-OFDMA). Он позволяет использовать 1024, 512 и 128 номинальных поднесущих. Режим с 256 поднесущими выпадает из этого ряда, поскольку тогда OFDMA становится аналогичным субканальному режиму в OFDM.

Принцип S-OFDMA позволяет работать в полосах частот различной ширины. Очевидно, чем уже полоса, тем меньше расстояние между номинальными поднесущими OFDMA. Соответственно уменьшается и допустимая скорость модуляции каждой из них. Но самое главное — снижается возможность противостоять межсимвольной интерференции, особенно для быстро движущихся абонентов (усиливается влияние доплеровского сдвига частот). То есть пропадает одно из ключевых достоинств OFDM-модуляции — возможность работать вне условий прямой видимости, с переотраженными сигналами. Очевидное решение — пропорционально ширине полосы изменять число номинальных поднесущих. Особенно актуальна такая возможность для глобальных мобильных сетей, где АС приходится взаимодействовать с БС различных операторов или в различных регионах. Изменяя число поднесущих пропорционально полосе, можно обеспечить инвариантность относительно ширины рабочей полосы таких ключевых параметров, как длительность OFDM-символа и расстояние между поднесущими. Например, при переходе от 20-МГц полосы к 5-МГц полосе можно с работы на 2048 поднесущих перейти к 512 поднесущим. При этом (при прочих равных) общая пропускная способность системы снизится в четыре раза, но зато все остальные параметры останутся неизменными.

Однако с введением масштабируемой OFDMA стандарт требует, чтобы все мобильные устройства при первичной регистрации и при сканировании окружения могли определять размер рабочей полосы доступных базовых станций и число используемых или OFDM-поднесущих. Также с введением S-OFDMA изменяется и показатель передискретизации (в данном случае — отношение частоты дискретизации сигнала к ширине полосы). Если прежде он был фиксированным и равным 8/7, то теперь он принимает значение 28/25, если ширина рабочей полосы кратна 1,25; 1,5; 2 или 2,75 МГц. Если же полоса пропорциональна 1,75 МГц (и во всех остальных случаях), его значение составляет 8/7.

С появлением четырех различных вариантов номинального числа поднесущих в OFDMA для каждого из них составлены свои таблицы пилотных частот и адаптированы базовые формулы распределения поднесущих по логическим субканалам. Это весьма увеличило объем документа, но никак не изменило общих принципов стандарта IEEE 802.16-2004.

Еще одно значимое отличие — изменился порядок распределения слотов в пакете (рис. 7.30). Если IEEE 802.16-2004 предписывал сначала заполнять доступные слоты в одном субканале, затем — в следующем и т.д., то теперь все наоборот — сначала слотами заполняются все субканалы для одного OFDMA-символа, затем — для следующих OFDMA-символов (с шагом, соответствующим размеру слота).

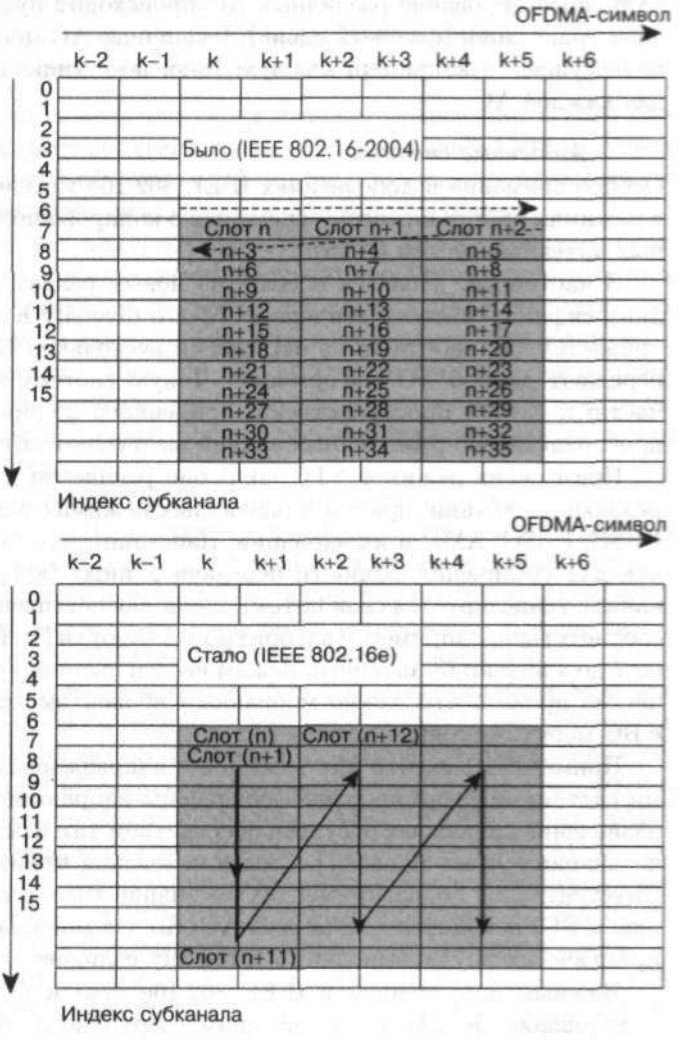


Рис. 7.30. Порядок распределения слотов в OFDMA-пакете (в примере — нисходящий канал, режим PUSC)

В стандарт введены два новых метода распределения несущих по субканалам в нисходящем кадре — так называемые методы TUSC1 и TUSC2 (tile usage of subchannels — фрагментарное использование субканалов). Они практически аналогичны методам распределения несущих для восходящего канала, причем TUSC1 подобен основному методу, а TUSC2 — дополнительному опциональному методу. Оба метода используются только при работе с адаптивными антенными системами (в AAS-зоне).

В стандарт IEEE 802.16e в режиме OFDMA добавлен новый механизм — зондирование восходящего канала (uplink channel sounding). Его суть — по команде БС или по заданному расписанию мобильная АС передает специальное тестовое сообщение (на основе последовательности Голея). Трансляция ведется в особой выделенной зоне в восходящем субкадре. Причем эта зона может быть назначена многим АС. Мобильные станции передают тестовое сообщение либо на всех доступных поднесущих частотах, или на некоторых из них. В первом случае

мультиплексирование различных АС происходит путем смещения момента начала трансляции (фазовый сдвиг). Различные АС могут использовать и разные поднесущие — например, каждую 16-ю, но с уникальным начальным сдвигом для каждой АС.

Антенные системы

Особое внимание в дополнениях IEEE 802.16e уделено системам MIMO — как в режимах пространственно-временного кодирования (STC), так и для адаптивных антенных систем (AAS).

В частности, добавлен абсолютно новый раздел, в котором режим макродиверсифицированного хэндовера (Macro diversity handover — MDHO) рассматривается как работа АС в окружении нескольких БС, которые одновременно передают данной АС информацию. Такую систему можно рассматривать как частный случай пространственно-временного кодирования (STC). В качестве пространственно-разнесенных антенн выступают антенны различных БС.

Применение режимов STC подробно расписано для 2-, 3- и 4-антенных передающих станций, причем в различных режимах распределения субканалов — PUSC, FUSC, AMC и их вариаций. Напомним, что механизм STC используется как для увеличения скорости передачи данных (когда по каждому антенному каналу транслируется свой поток), так и для повышения надежности передачи в соответствии, например, с алгоритмом Аламоути [4]. Возможность использовать от 2 до 4 антенн (3-антенный режим введен впервые) существенно повышает качество приема, что важно мобильным абонентам, для которых условия связи с БС могут постоянно меняться.

Примечательно, что для режимов с направленными адаптивными антеннами системами (формирование диаграммы направленности) исключен механизм назначения канальных ресурсов посредством так называемой направленной сигнализации (Direct Signaling). Смысл изменения лежит на поверхности — метод Direct Signaling подразумевал точное знание того, в каком луче (где по отношению к БС) находится конкретная АС. Но для мобильных абонентов это условие практически трудновыполнимо, а потому излишне.

Важным дополнением в IEEE 802.16e стал и новый механизм канального кодирования. К уже существующим сверточному коду и опциональным блоковым и сверточным турбокодам добавлен код контроля четности с низкой плотностью LDPC (Low Density Parity Check). Этот код базируется на наборе линейных блоковых кодов. Принципиальной особенностью LDPC является гибкость. Он позволяет работать с кодовыми словами переменной длины и с различными скоростями кодирования. Так, в стандарте IEEE 802.16e определены размеры кодовых слов от 72 до 288 байт и скорости кодирования от 1/2 до 5/6. Размер кодируемого блока данных зависит от числа доступных субканалов и метода модуляции (т. е. от доступных частотных ресурсов и условий в канале).

Гибридный режим повторной передачи

Наконец, важное значение в IEEE 802.16e придается методу HARQ — гибридному механизму автоматического запроса на повторную передачу в случае ошибки (Hybrid Automatic Repeat Request). Напомним, что все стратегии трансляции данных с повторной передачей в случае обнаружения ошибки приема (ARQ) можно разделить на три типа:

- остановка и ожидание (Stop and Wait);
- при ошибке передать последние n пакетов (Go Back n, GBN);
- выборочный повтор передачи пакета (Selective Repeat, SR).

Метод Stop and Wait означает, что перед началом трансляции нового блока данных передатчик ожидает подтверждения успешного приема предыдущего блока данных либо сообщения об ошибке. Метод надежный, но требующий существенных канальных ресурсов либо снижающий скорость обмена.

Механизм GBN не требует подтверждения каждого блока данных — передача идет непрерывно. Но если передатчик получает сообщение о пропущенном или принятом с ошибкой пакете, он повторяет передачу данных, начиная с поврежденного пакета. При этом даже успешно переданные пакеты транслируются повторно. Подобный механизм удобен тем, что не требует буферизации и хранения данных в приемнике, но в случае ошибки существенно возрастает загрузка канала. Из-за одного ошибочного пакета нормально принимаемые данные могут транслироваться многократно.

Алгоритм SR позволяет выборочно повторять передачу только поврежденных и пропущенных пакетов. Но в этом случае передатчик должен хранить определенное число последних принятых пакетов. Тем не менее, поскольку данный метод наиболее экономичен в отношении канальных ресурсов, он является основным в беспроводных телекоммуникационных технологиях. Именно метод SR и используется в стандарте IEEE 802.16 в рамках механизма ARQ.

Однако в стандарт вернулся и, казалось бы, забытый механизм Stop and Wait — фактически метод квитированной передачи. Поскольку он требует быстрого подтверждения/сообщения об ошибке, он используется только в режиме OFDMA, который позволяет выделить специальный канал для подтверждения передачи. Данный механизм назвали гибридным ARQ (HARQ). Если HARQ включен, каждый пакет, переданный БС, требует подтверждения от АС по специальному обратному каналу. Такая опция может назначаться выборочным или всем соединениям определенной АС. Причем в случае мобильной АС HARQ позволено назначать только определенным соединениям, а не станции в целом.

Если приемник сообщает об ошибке или подтверждение успешного приема не получено в установленный срок, передатчик приступает к повторной передаче. Здесь можно задействовать один из двух альтернативных механизмов — передачу с увеличивающейся избыточностью (Incremental Redundancy — IR) или с «управляемым комбинированием» (Chase Combining — CC). Суть метода IR — при включенном механизме HARQ для каждого исходного пакета в канальном кодере формируется до четырех так называемых субпакетов, каждый со своим идентификатором (SPID). При использовании стандартного сверточного кодера [2] эти субпакеты отличаются лишь тем, что шаблон перфорирования кодера для каждого последующего SPID циклически сдвигается влево на один разряд. Первым передается пакет с SPID=0. Если произошел сбой, повторно транслируется субпакет с другим SPID — т. е. тот же самый кодированный исходный пакет, но с иными параметрами кодера. Причем стандарт не предполагает изменения скорости кодирования.

Метод CC в случае ошибки предполагает повторную трансляцию одного и того же кодированного пакета. Он может использоваться только с мобильными АС. Почему метод так назван, можно лишь гадать. Вероятно, авторы

документа имели в виду, что возможно корректное декодирование нескольких одинаковых пакетов, но с различными ошибками.

7.9.3. Элементная база систем стандарта IEEE 802.16e

Успех любой современной технологии определяется наличием элементной базы. В полной мере это относится и к технологии IEEE 802.16e. Поскольку это достаточно новый и сложный в реализации стандарт, производителей элементной базы пока не так много, не все обладают соответствующими технологическими возможностями. Из производителей чипсетов с поддержкой «мобильного WiMax» [7] — т.е. режима OFDMA и требований дополнения IEEE 802.16e — можно назвать несколько компаний. Прежде всего это Intel, Runcom, Fujitsu, Sequans, Beceem, GCT Semiconductor, Motorola и др. Тем не менее, по данным аналитической компании In-Stat (www.instat.com), в 2007 году рынок чипсетов для АС стандарта IEEE 802.16 составил 27 млн. долл., а для БС — 130 млн. долл. Причем к 2012 году аналитики данной компании предсказывают его рост до 500 млн. и 1,4 млрд. долл. соответственно.

В частности, Intel представил набор микросхем для абонентского оборудования WiMAX Connection, в состав которого входит ИС телекоммуникационного процессора WiMAX Connection 2400, а также две «системы в корпусе» — трехдиапазонный радиомодуль без блока аналогового сопряжения WiMAX Connection 2300R и полнофункциональный двухдиапазонный радиомодуль WiMAX Connection 2320R (рис. 7.31). Этот чипсет предназначен в первую очередь для модулей Intel «Echo Peak», интегрирующих технологии IEEE 802.16e и IEEE 802.11 (мобильный WiMAX в диапазоне 2,5–2,7 ГГц и Wi-Fi в диапазонах 2,4–2,48 и 5,1–5,8 ГГц).

Процессор WiMAX Connection 2400 реализует физический и MAC-уровни IEEE 802.16, модуль 2300R поддерживает диапазоны 2,4–2,4; 2,5–2,69 и 3,3–3,6 ГГц. Модуль 2320R работает в диапазонах 2,3–2,69 и 3,3–3,6 ГГц. Типовая пропускная способность решения на основе WiMAX Connection 2400 и 2320R — 6 Мбит/с в нисходящем канале и 2 Мбит/с — в восходящем (пиковые скорости — 20 и 5 Мбит/с, соответственно). Поддерживаются частотные полосы шириной 5; 7; 8,75; и 10 МГц. При этом мощность потребления в режиме ожидания — менее 5 мВт. Чипсет предназначен для работы с MIMO в конфигурации 2×2, сам чипсет поддерживает систему с одной передающей и двумя приемными антеннами.

Израильская компания Runcom также представила законченное решение, совместимое с IEEE 802.16e (режим OFDMA), — систему на кристалле RNA200 для абонентских станций (рис. 7.32). ИС включает процессорное ядро семейства ARM11, модули аналоговой обработки, интерфейсные (цифровые и аналоговые) блоки. Чтобы построить законченную систему, дополнительно к RNA200 требуется только ВЧ-модуль и внешняя память (ОЗУ типа SDRAM и РПЗУ программ). Реализован интерфейс SIM-карт. В качестве интерфейса к хост-системе реализован USB и SPI, а также Ethernet-порт.

RNA200 поддерживает работу в полосе до 20 МГц, режим временного дуплексирования, сверточное турбокодирование, ARQ, хэндовер и др. Процессор обеспечивает БПФ на 1024 точки, т.е. работу в режиме OFDMA с 1024 поднесущими. Таким образом, ИС предназначена для поддержки определенного WiMAX-профиля [7], а не режима OFDMA стандарта IEEE 802.16 как такового.

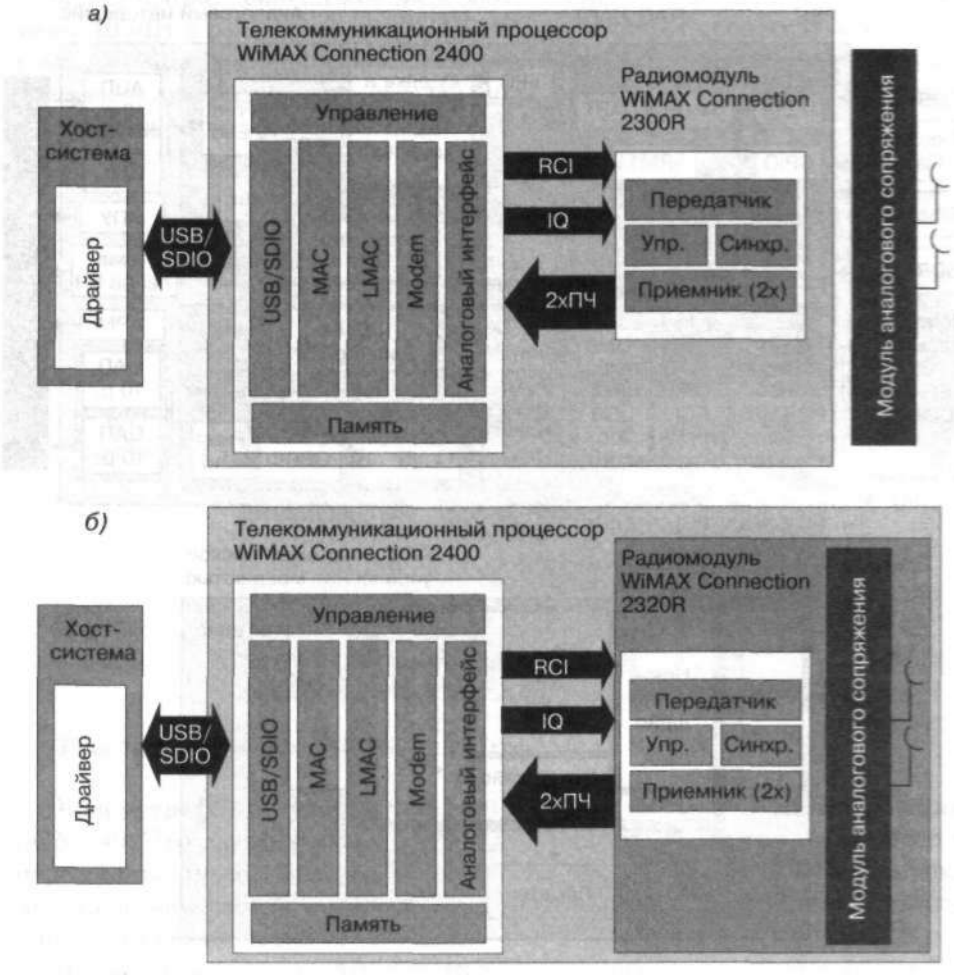


Рис. 7.31. Построение АС на основе чипсета Intel WiMAX Connection: а) с внешним модулем аналогового сопряжения, б) используя радиомодуль WiMAX Connection 2320R

Очень похожую по структуре ИС для мобильного WiMAX представила и компания Fujitsu. Ее система на кристалле MB86K21 (рис. 7.33) для АС поддерживает работу с 512 и 1024 поднесущими в режиме OFDMA, полосы шириной 5 и 10 МГц, режим MIMO 2Ч2 и т. п.

Весьма интересны и решения компании Sequans Communications — известного производителя чипсетов для фиксированного, а теперь — и мобильного WiMAX. Для БС компания предлагает заказную СБИС SQN2130. Ранее компания представила чипсет SQN2110, основанный на трех ПЛИС семейства Stratix II компании Altera и ее IP-блоках. SQN2130 — это телекоммуникационный процессор, поддерживающий физический и MAC-уровни стандарта IEEE 802.16 для режима OFDMA с 512 и 1024 несущими. Чип поддерживает два приемных и два передающих антенных канала, временное и частотное дуплексирование, скорость передачи до 35 Мбит/с. Энергопотребление в режиме MIMO — менее 2 Вт.

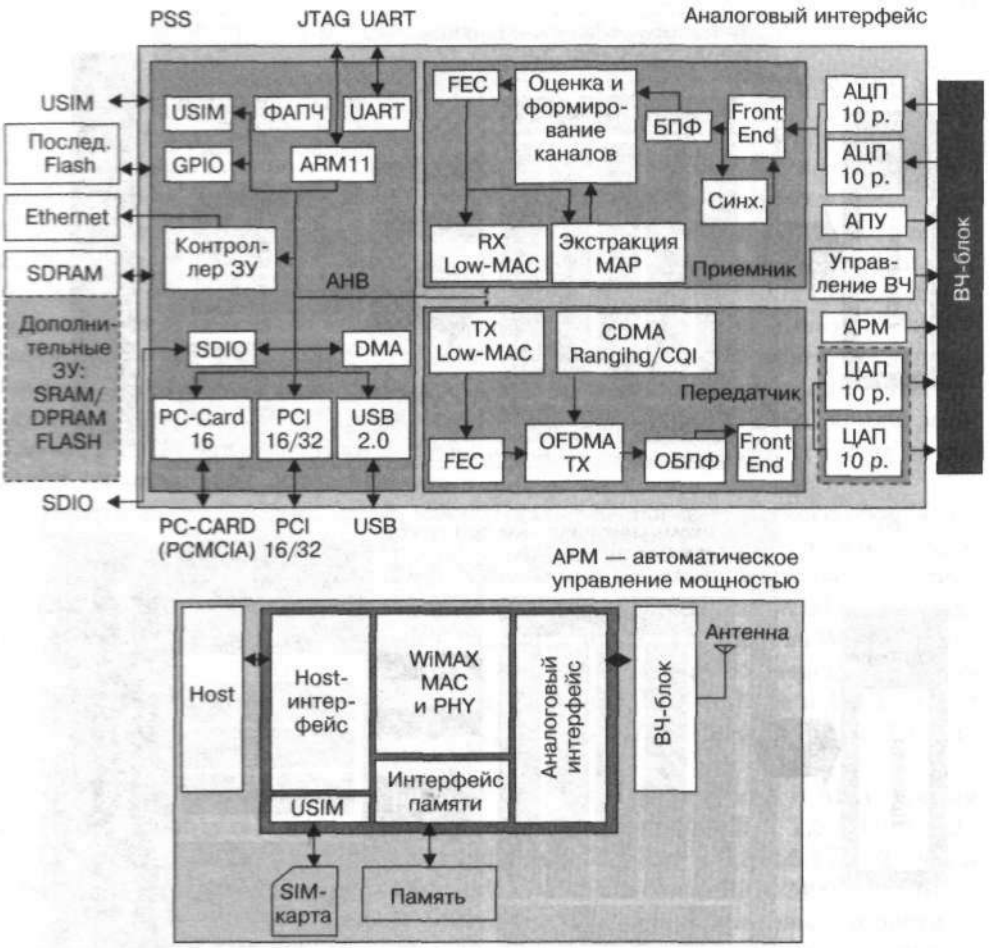


Рис. 7.32. Система на кристалле RNA200 (Runcom)

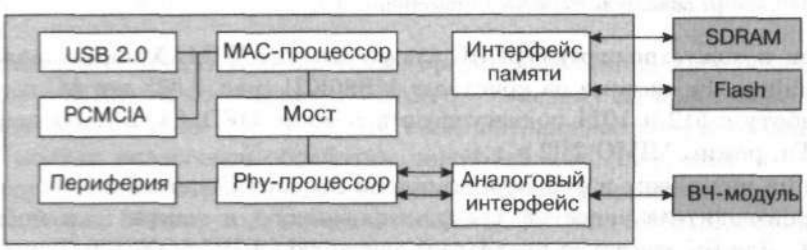


Рис. 7.33. Система на кристалле MB86K21 (Fujitsu)

Для мобильных станций компания Sequans предлагает СБИС SQN1110 и SQN1130. В отличие от процессора для базовой станции, эти СБИС содержат интегрированный модуль аналоговой обработки, поэтому из внешних системных компонентов необходим только ВЧ-трансивер (рис. 7.34). СБИС SQN1130

отличает чрезвычайно низкое энергопотребление — 280 мВт в рабочем режиме, менее 10 мВт в режиме ожидания. СБИС обеспечивает пропускную способность свыше 30 Мбит/с. В ней реализован режим OFDMA с 512 и 1024 поднесущими, временное дуплексирование (в SQN1110 — временное дуплесное и полудуплексное частотное разделение восходящего и нисходящего каналов).

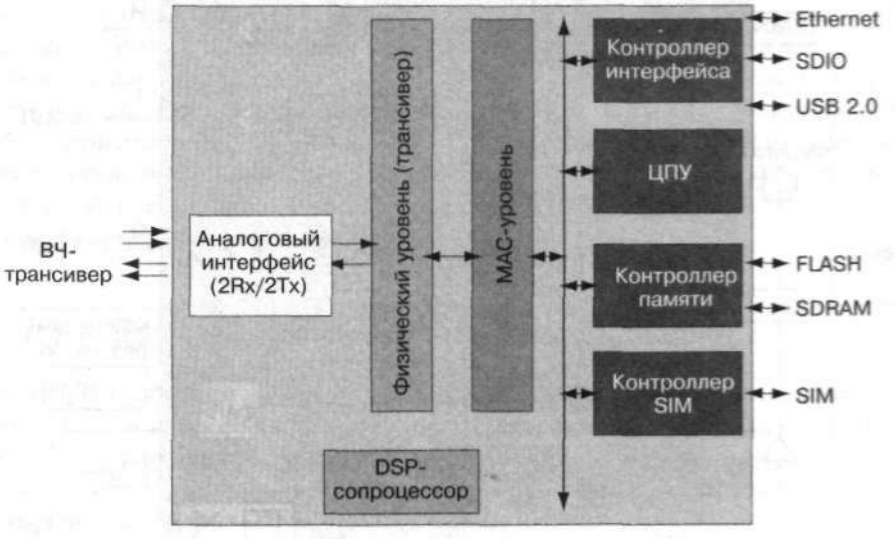


Рис. 7.34. Структура СБИС SQN1130 (Sequans)

В качестве ВЧ-модуля для SQN1130 компания предлагает ИС ВЧ-трансивера SQN1140. Это двухканальный (для поддержки MIMO) трансивер прямого преобразования, предназначенный для работы в диапазоне 2,3–2,7 ГГц. В режиме двухканальной работы его потребление не превышает 290 мВт (195 мВт в одноканальном режиме).

Законченные чипсеты MS120 и BCS200 для мобильных АС предлагает и компания Beceem. Они ориентированы на требования первой и второй волны WiMAX-сертификации [7], соответственно. В состав каждого чипсета входят телекоммуникационный процессор и ВЧ-трансивер. Так, чипсет BCS20 включает процессор BCSB120 и интегрированный ВЧ-трансивер BCSR120. Он позволяет строить системы для работы в 2- и 3-ГГц диапазонах с временным дуплексированием, поддерживает два приемных и один передающий антенный канал. Возможная ширина рабочей полосы — 5; 7; 8,75 и 10 МГц.

Помимо перечисленных производителей, многие компании предлагают решения для универсальных процессорных платформ, позволяющие поддержать требования стандарта IEEE 802.16. Это DSP-процессоры компаний Texas Instruments, Freescale, PicoChip и др. Причем PicoChip хоть и предлагает чрезвычайно оригинальные универсальные процессоры с сотнями (порядка 300) интегрированных DSP-ядер [8], но четко позиционирует их в сегменте WiMAX-продуктов. Неслучайно известный отечественный производитель систем широкополосного доступа — компания InfiNet Wireless (www.infinet.ru) — использует процессоры фирмы PicoChip в своих продуктах.

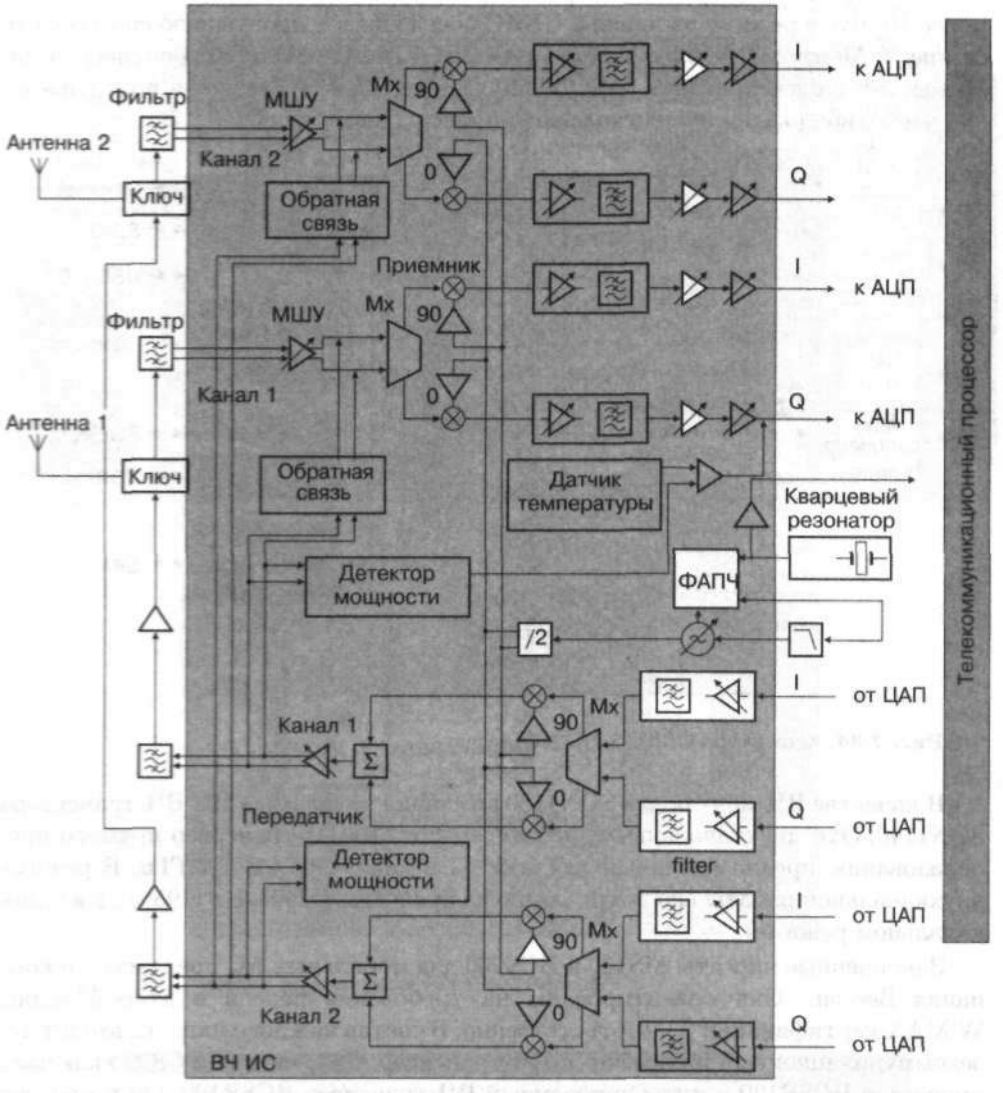


Рис. 7.35. Трансивер прямого преобразования семейства UXA234xx (NXP) (структура и схема включения)

Кроме того, достаточно много предложений аналоговых приборов для WiMAX-систем, прежде всего — специализированных ВЧ-трансиверов. Характерный пример подобных изделий — семейство трансиверов прямого преобразования UXA234xx компании NXP (рис. 7.35). Трансиверы этого семейства перекрывают частотный диапазон от 2,3 до 3,8 ГГц. Благодаря прямому преобразованию не нужны дополнительные фильтры, что упрощает и удешевляет систему в целом. Усиление входного сигнала трансиверов (в зависимости от типа) — от 77 до 87 дБ, уровень шумов — от 2,5 до 3,5 дБ, диапазон усиления передатчика — 74 дБ. В зависимости от типа ИС поддерживается от одного до двух приемных и передающих каналов. Например, трансивер UXA23466 работает в диапазо-

не 2,3–2,7 ГГц, поддерживает по два приемных и передающих канала, а при напряжении питания 2,7–2,9 В ток потребления в режимах приема/передачи составляет от 81/182 мА. Отметим, что аналогичные решения предлагают также компании Maxim, Analog Devices и др.

Таким образом, уже сейчас можно констатировать, что производители элементной базы стремятся поддержать не стандарт IEEE 802.16 как таковой и даже не его отдельные режимы (OFDM, OFDMA и S-OFDMA т.п.), а отдельные профили WiMAX, что далеко не одно и то же. Учитывая, что с частотными диапазонами для WiMAX полной определенности нет, следствием такого подхода может стать то, что в отдельных странах (например, в России) данную элементную базу использовать будет нельзя. Это приведет к необходимости применять универсальные аппаратные средства, неизбежно увеличивающие себестоимость устройств. Вероятно, по мере развития рынка WiMAX подобные лакуны с элементной базой будут заполнены.

7.10. Дальнейшее развитие стандарта IEEE 802.16

Дальнейшим развитием стандарта IEEE 802.16 занимаются несколько рабочих групп комитета IEEE 802.16. Так, группа поддержки трудится над объединением всех существующих дополнений и исправлений в единый документ. Группа безлицензионных диапазонов h занимается проблемами согласованной работы оборудования стандарта IEEE 802.16 в безлицензионных диапазонах (США). Фактически это означает совместимость с оборудованием стандарта 802.11. Но наиболее перспективные работы сосредоточены в группах j и m.

7.10.1. Проект IEEE 802.16j

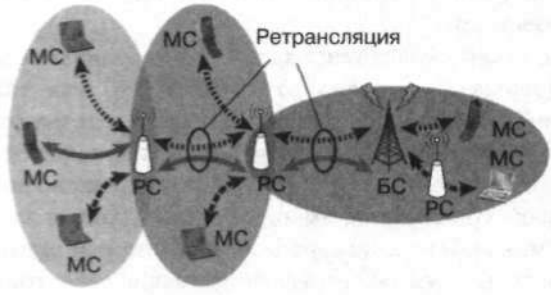
Целевая группа IEEE 802.16j была создана в марте 2006 года для разработки системы мобильной многостадийной ретрансляционной системы в рамках стандарта IEEE 802.16. Фактически эта работа — развитие направления mesh-сетей, уже описанных в IEEE 802.16-2004. Однако в том документе описывались mesh-сети фиксированного доступа. Стандарт IEEE 802.16j должен улучшить производительность сети при многостадийной передаче, причем как ретрансляторы, так и абонентские станции могут быть мобильными. В отличие от mesh-сети, РС может работать в режиме вещания («точка–многоточка»).

Прежде всего, стандарт IEEE 802.16j вводит новое понятие — релейная станция (ретранслятор, РС). Выделяют прозрачный и непрозрачный режим работы РС. В прозрачном режиме РС транслирует только данные и не транслирует преамбулы и управляющие поля, такие, как DL-MAP, UL-MAP и т.п. (рис. 7.36). Эту информацию АС получает непосредственно от БС. При этом АС логически никак не взаимодействует с РС (не знает о ее существовании).

В непрозрачном режиме РС передает не только данные, но и преамбулу, а также все управляющие сообщения. По отношению к АС она выглядит как БС, абонентская станция физически и логически соединена именно с ней.

Кроме того, РС могут обладать возможностями диспетчеризации и защиты передаваемого трафика (распределенное управление) или не обладать таковыми (централизованное управление). По данным ряда исследователей [9], применение

РС в режиме непрозрачности позволяет увеличить общую пропускную способность ячейки более чем на 40% (с 6,2 до 8,8 Мбит/с по данным работы [9]).



Структура ретрансляционной сети

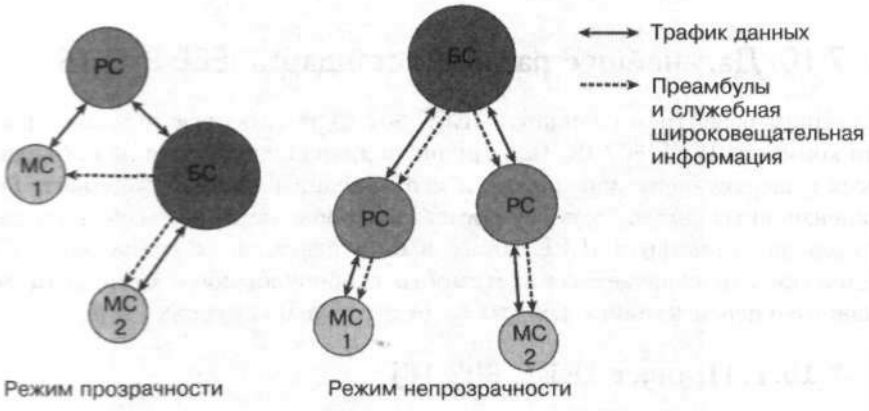


Рис. 7.36. Структура сети ретрансляции

Стандарт IEEE 802.16j обеспечивает полную обратную совместимость с IEEE 802.16e (мы сравниваем с дополнением «e», поскольку IEEE 802.16j ориентирован именно на мобильные сети). Однако он обладает и рядом отличий, как на физическом, так и на MAC-уровне. В частности, на физическом уровне несколько видоизменяется структура кадров. Общая структура кадров режима OFDMA сохранена. Однако и нисходящий, и восходящий субкадры делятся на интервал доступа и интервал ретрансляции (рис. 7.37). В интервале доступа происходит трансляция между МС и ее станциями доступа — базовой или ретрансляционной в режиме непрозрачности. Его структура полностью соответствует структуре кадра OFDMA.

В интервале ретрансляции происходит радиообмен между БС и РС или только между РС. То есть весь передаваемый в этот момент трафик — это ретранслируемый трафик. Интервал ретрансляции начинается с пакета R-FCH (управляющий заголовок ретрансляционного фрейма), за которым следует управляющая информация (аналог карт DL/UL-MAP, в которых расписано назначение слотов для каждого приемного/передающего устройства). Для синхронизации между РС предусмотрена дополнительная ретрансляционная преамбула (Relay Amble). Она передается в конце нисходящего ретрансляционного интервала, причем мо-

жет передаваться не в каждом кадре, но не реже чем раз в 40 мс (8 кадров длительностью 5 мс).

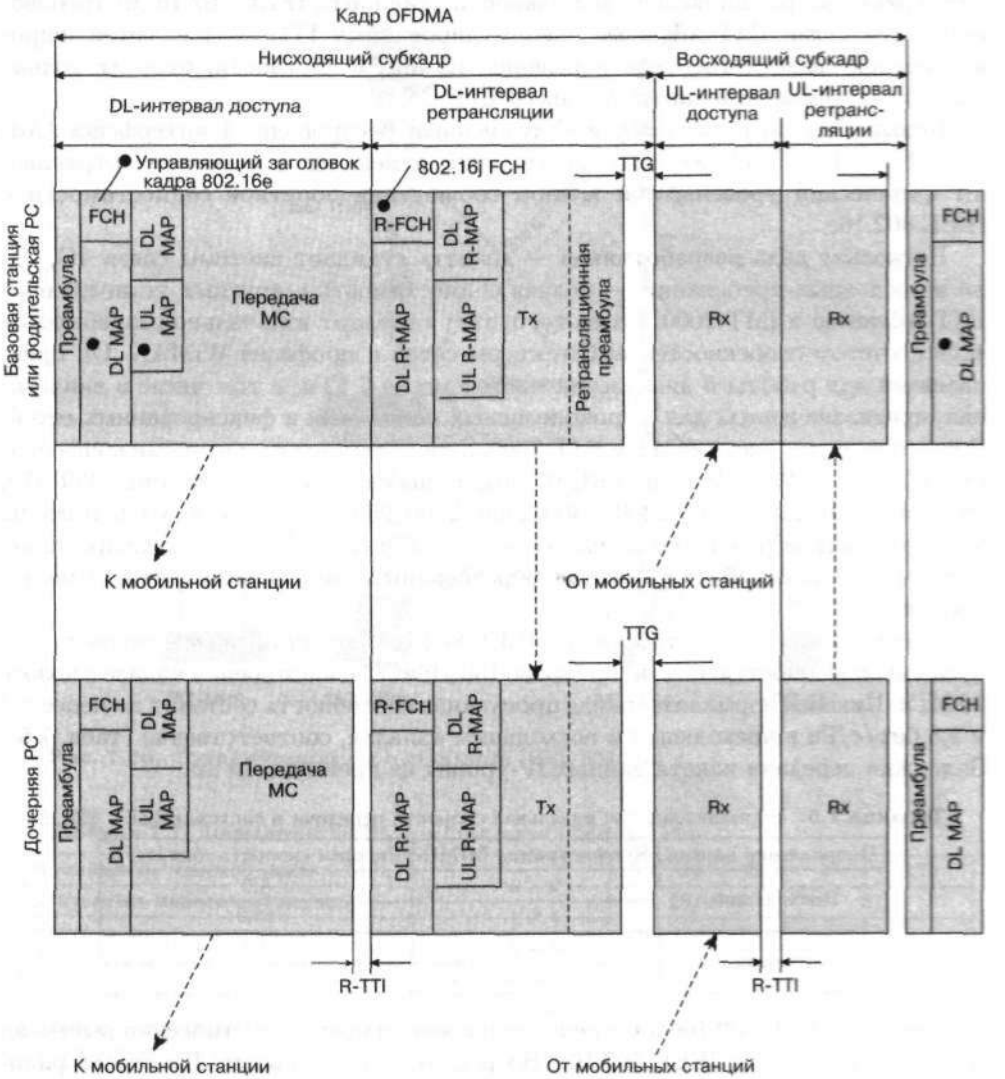


Рис. 7.37. *Интервалы доступа и ретрансляции в кадре стандарта IEEE 802.16j*

На MAC-уровне изменения в IEEE 802.16j коснулись в основном процедур начальной регистрации в сети, авторизации, передачи ключей шифрования и т. п.

Отметим, что РС могут быть фиксированными, номадическими (передвижной ретранслятор) и мобильными. Характерный пример мобильной РС — ретрансляционная станция в скоростном поезде (автобусе). С точки зрения топологии сети РС возможны самые разнообразные варианты. Например, несколько РС могут транслировать данные для одной МС (вариант многоантенной системы MIMO).

7.10.2. Проект IEEE 802.16m

Не менее важными работами занята целевая группа IEEE 802.16m. Она создает спецификации, расширяющие возможности стандарта IEEE 802.16 до требований сетей связи IMT-Advanced, сформулированных ITU — т.е. сетей широкополосной связи четвертого поколения. К ним, в частности, относят только разрабатываемые стандарты Advanced LTE.

Новый стандарт называется «Улучшенный беспроводный интерфейс» (Advanced Air Interface), из чего следует, что изменения в основном затрагивают физический уровень, при полном соблюдении обратной совместимости с IEEE 802.16e.

Поскольку цель разработчиков — создать стандарт системы связи 4G, одно из основных требований — полная совместимость с другими технологиями IMT-Advanced и IMT-2000. Более того, этот стандарт изначально разрабатывается с учетом особенностей архитектуры сетей и профилей WiMAX. Он предназначен для работы в диапазонах частот менее 6 ГГц, в том числе в диапазонах, предназначенных для широкополосных мобильных и фиксированных сетей. В частности, для систем IMT и IMT-2000, в соответствии с решениями конференций WARC-92, WRC-2000 и WRC-07, выделены следующие диапазоны: 450–470, 698–960, 1710–2025, 2110–2200, 2300–2400, 2500–2690 и 3400–3600 МГц. Должны поддерживаться режимы как частотного, так и временного дуплексирования, причем возможна работа с каналом, образованным несколькими частотными полосами.

Оборудование, соответствующее IEEE 802.16m, будет обладать пиковой пропускной способностью на сектор свыше 150 Мбит/с в нисходящем канале в полосе 20 МГц. Пиковая нормализованная пропускная способность составит не менее 8,0 и 2,8 бит/с/Гц в нисходящем и восходящем каналах, соответственно (табл. 7.6). Задержка передачи пакета данных IP-уровня не превысит 10 мс.

Таблица 7.6. Нормализованная идеальная скорость передачи в системах IEEE 802.16m

Направление канала	Конфигурация MIMO	Пиковая скорость, бит/с/Гц
Нисходящий	2 × 2	8,0
	4 × 4	15,0
Восходящий	1 × 2	2,8
	2 × 4	5,6

Системы IEEE 802.16m обеспечат дополнительный энергетический выигрыш 3 дБ по сравнению с IEEE 802.16e. Возрастет зона покрытия. В ячейках радиусом до 5 км система должна работать с заявленными характеристиками. При увеличении радиуса ячейки до 30 км будет происходить постепенная деградация качества. Но система должна сохранять работоспособность на удалении от БС до 100 км (когда ограничения продиктованы принципиальным барьером в виде теплового шума). Скорость МС может составлять до 350 км/ч.

Системы IEEE 802.16m будут обеспечивать определение местоположения МС с точностью от 50 до 150 м.

Стандарт m изначально предусматривает возможности, разрабатываемые в рамках проекта IEEE 802.16j, — т.е. работа в ретрансляционном режиме будет присуща оборудованию IEEE 802.16m (рис. 7.38). При этом ретрансляционные станции IEEE 802.16m не смогут напрямую работать с РС IEEE 802.16j —

но в этом едва ли возникнет необходимость. Более того, можно предположить, что дополнение IEEE 802.16j по сути станет составной частью IEEE 802.16m.

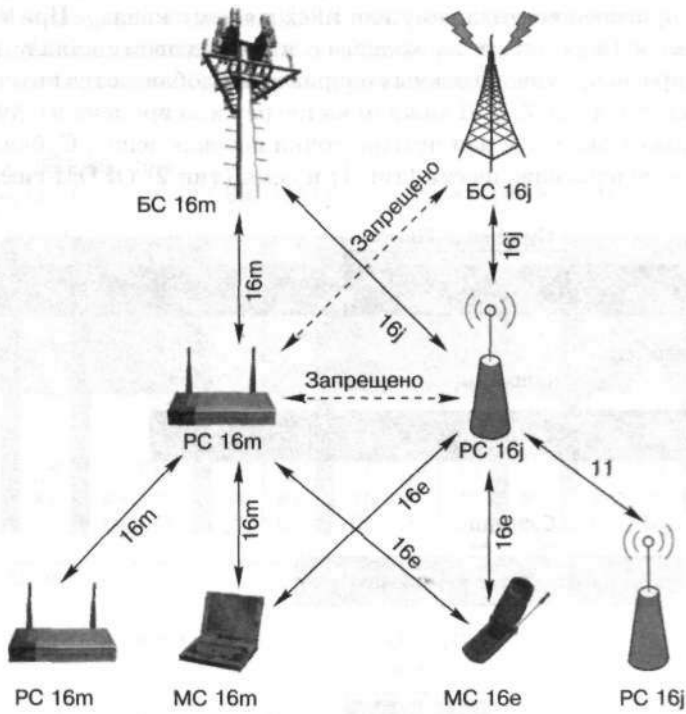


Рис. 7.38. Пути ретрансляции в стандарте IEEE 802.16m

Таблица 7.7. Параметры OFDMA в стандарте IEEE 802.16m

Номинальная ширина канала, МГц		5	10	20	7	8,75
Показатель передискретизации		28/25			8/7	
Частота дискретизации, МГц		5,6	11,2	22,4	8	10
Формальное число поднесущих (для БПФ)		512	1024	2048	1024	1024
Шаг поднесущих, кГц		10,938			7,8125	9,7656
Длительность OFDM-символа без циклического префикса (Tu), мкс		91,429			128	102,4
Циклический префикс длительностью Tg = 1/8 Tu	Длительность символа, мкс	102,86			144	115,2
	Число OFDM-символов в кадре	48			34	43
	Интервал переключения, мкс	62,86			104	46,4
Циклический префикс Tg = 1/16 Tu	Длительность символа, мкс	97,143			Не используется	
	Число OFDM-символов в кадре, не более	51				
	Интервал переключения, мкс	45,71				

Одна из основных особенностей стандарта 802.16m — изменение структуры кадров. Разумеется, стандарт рассматривает исключительно режим OFDMA (табл. 7.7). В IEEE 802.16m введен новый элемент — суперкадр длительно-

стью 20 мс (максимально допустимое время кадра в IEEE 802.16e). Суперкадр (рис. 7.39) делится на четыре кадра длительностью по 5 мс. Если ширина канала составляет 5, 10 или 20 МГц, каждый кадр содержит восемь субкадров. Субкадр может быть присвоен восходящему или нисходящему каналу. При смене направлений передачи (переход от нисходящего к восходящему каналу и наоборот) между субкадрами противоположных направлений добавляется интервал (точка) переключения (см. табл. 7.7). В каждом кадре режима временного дуплексирования (TDD) может быть две или четыре точки переключения. Субкадры бывают двух типов — содержащие шесть (тип 1) и семь (тип 2) OFDM-символов.

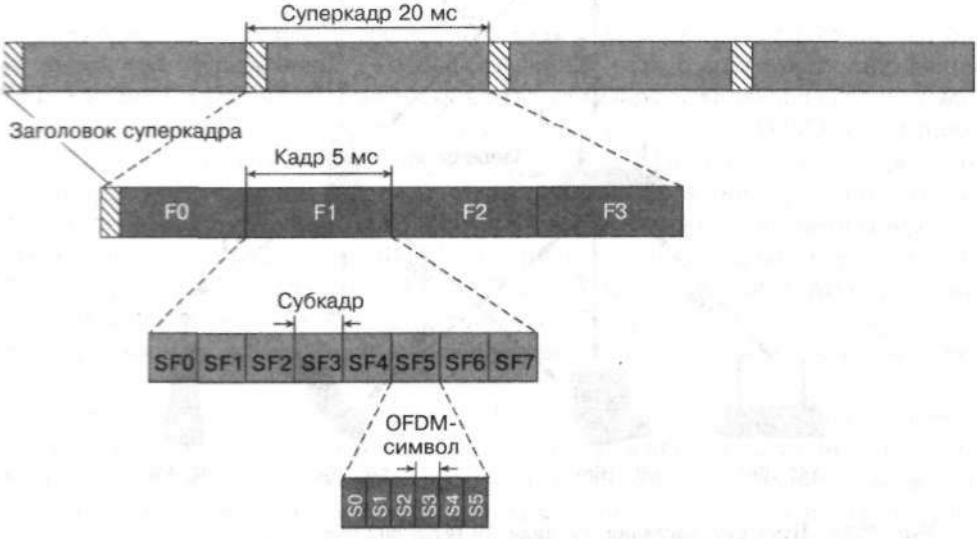


Рис. 7.39. Структура кадров в стандарте IEEE 802.16m

Изменения в кадровой структуре были направлены на обеспечение совместимости как с предыдущими версиями стандарта IEEE 802.16, так и с другими стандартами широкополосной передачи, входящими в пул IMT-2000 и IMT-Advanced (например, LTE). Для поддержки всех этих возможностей вводится понятие временных зон. Зона — это один или несколько смежных субкадров. В каждой такой зоне может передаваться трафик только для устройств IEEE 802.16e или только для устройств IEEE 802.16m (рис. 7.40). Кроме того, каждая зона может дополнительно подразделяться на временные интервалы для поддержки режимов ретрансляции (отдельно зоны IEEE 802.16e/j и зоны IEEE 802.16m, поскольку протоколы ретрансляции в них различны).

Еще одна возможность стандарта IEEE 802.16m — он позволяет работать с широкими каналами, более 20 МГц. Такие каналы представляют собой совокупность нескольких стандартных каналов (рис. 7.41). При этом отпадает необходимость в защитных частотных интервалах между каналами, что увеличивает доступную полосу пропускания. Гибкое регулирование структуры кадра позволяет обеспечивать совместимость с традиционными устройствами IEEE 802.16.

Кроме того, оборудование IEEE 802.16m должно быть совместимым с устройствами IMT-2000 и IMT-Advanced. Совместимость подразумевает, прежде все-

го, отсутствие интерференции сигналов оборудования различных типов. Этого можно добиться, используя схожие структуры кадров в этих стандартах. Необходимо, чтобы нисходящие и восходящие потоки различных систем совпадали по времени. И кадровая структура IEEE 802.16m обеспечивает такую возможность путем добавления пустых символов и выбора соответствующей структуры кадра. В частности, на рис. 7.42 показан режим совместимости с оборудованием стандарта LTE.

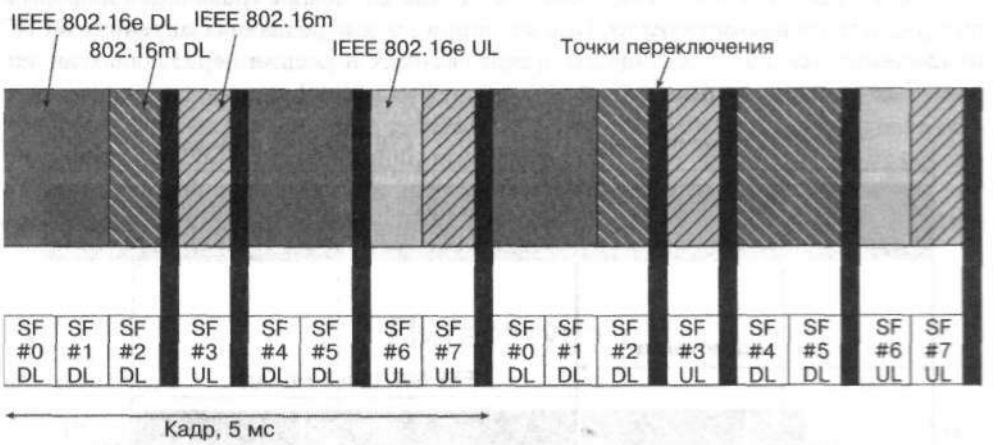


Рис. 7.40. Структура зон в кадрах стандарта IEEE 802.16m

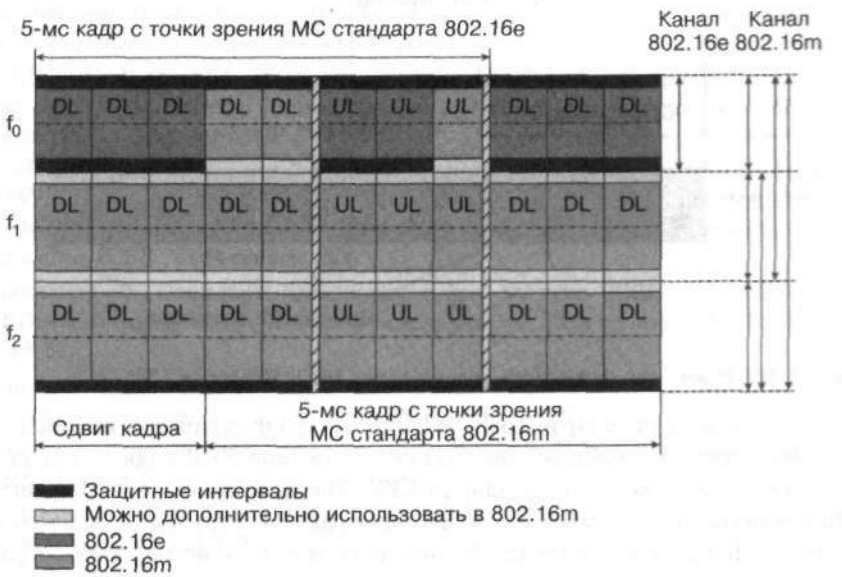


Рис. 7.41. Объединение нескольких стандартных каналов в один канал стандарта IEEE 802.16m

Стандарт IEEE 802.16m содержит ряд других важных изменений. Так, активнее используются системы MIMO. Минимальная конфигурация в нисходящем

канале предусматривает две передающие антенны на БС (в секторе) и две приемные антенны на МС. Всего же возможно до восьми передающих антенн на БС и до восьми приемных — на МС (допустимые конфигурации в DL-канале, передающие × приемные антенны — 2 × 2, 4 × 2, 4 × 4, 8 × 2, 8 × 4, 8 × 8). В восходящем канале на БС должно быть не менее двух приемных антенн, на МС — одна, две или четыре передающие. Помимо числа антенн, расширяются возможности режимов MIMO. Вводится режим Multi-user MIMO, в соответствии с которым единовременно и на одних частотах возможна трансляция информации различным пользователям. При этом при двух передающих антеннах на БС поддерживается до двух абонентов, при четырех и восьми передающих антеннах — до четырех пользователей.

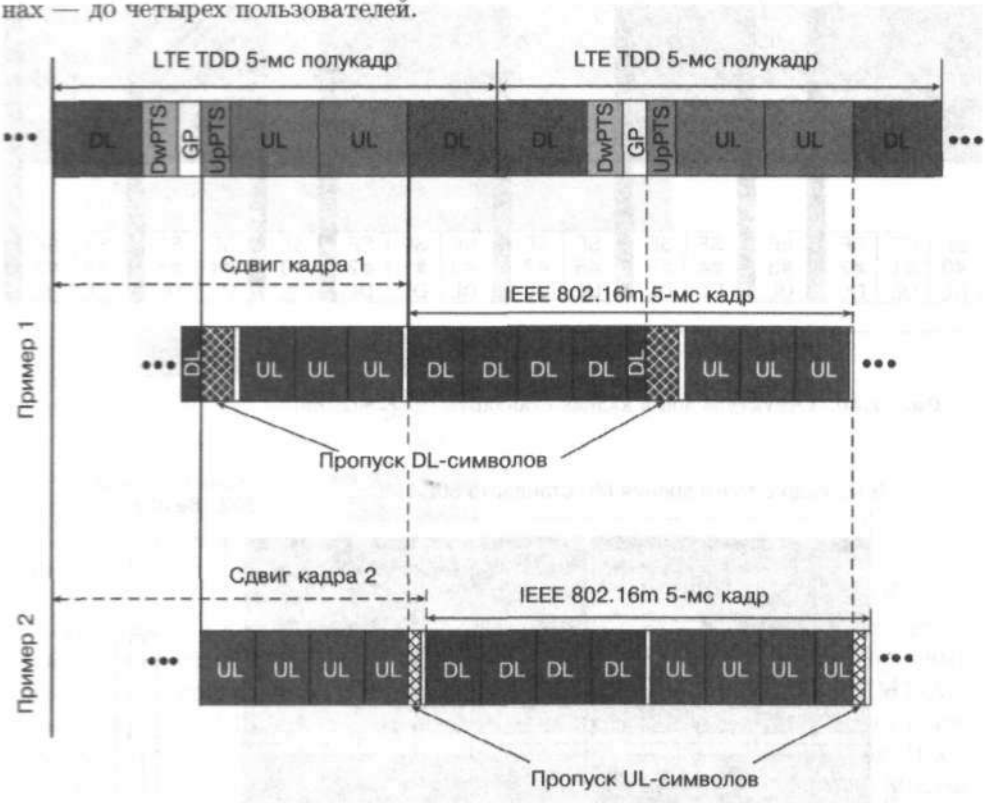


Рис. 7.42. Режим совместимости оборудования IEEE 802.16m и LTE

Таким образом, стандарт IEEE 802.16m — это фактически аналог стандарта LTE-Advanced. А появиться он должен к середине 2010 года — когда только ожидаются первые системы стандарта LTE. Таким образом, с появлением IEEE 802.16m, учитывая его гибкость и имманентные возможности в плане совместимости с другими технологиям и, можно будет реально говорить о создании систем мобильной связи 4G.

В заключение отметим: несмотря на то, что в реальных системах используются лишь отдельные возможности и режимы стандарта IEEE 802.16, новая технология живет и развивается. Подтверждение тому — решение ITU от 19 октября 2007 года о включении режима OFDMA TDD стандарта IEEE 802.16 в пул

глобальных стандартов беспроводной связи IMT-2000 (IMT-2000 OFDMA TDD WMAN). Причем не стоит забывать, что WiMAX — не просто технология. Это новая парадигма построения информационного пространства. Насколько именно эта технология окажется успешной, например в конкурентной борьбе с перспективными стандартами сотовой связи (например, LTE), — большой вопрос. Но то, что в ближайшие несколько лет ее ожидает бурное развитие, несомненно. Кроме того, развитие самого стандарта 802.16 не стоит на месте. В стадии создания и обсуждения находится ряд новых дополнений, таких как проекты j и m.

Не следует забывать, что научная база для практической реализации OFDM-систем впервые была описана российскими учеными [14]. Причем огромный научно-технический потенциал России не утерян и до сих пор. Об этом свидетельствует тот факт, что многие мировые лидеры в области технологии беспроводного оборудования используют коллективы российских разработчиков. Все это позволяет надеяться, что Россия в будущем не только будет полигоном для быстрого и широкомасштабного внедрения WiMax-совместимых систем, но и займет достойное место в мировом разделении рынка по производству таких систем.

Литература

1. Шахнович И. Сети городского масштаба. Решения рабочей группы IEEE 802.16 — в жизнь! — Электроника: НТБ, 2003, № 8, с.50–56.
2. Шахнович И. Стандарт широкополосного доступа IEEE 802.16 для диапазонов ниже 11 ГГц. — Электроника: НТБ, 2005, № 1, с. 8–14.
3. Вишневский В. М., Ляхов А. И., Портной С. Л., Шахнович И. В. Широкополосные беспроводные сети передачи информации. — М.: Техносфера, 2005.
4. IEEE Std 802.16-2004. IEEE Standard for Local and metropolitan area networks. Part 16: Air Interface for Fixed Broadband Wireless Access Systems. — IEEE, 1 October 2004.
5. IEEE Std 802.16e-2005 and IEEE Std 802.16-2004/Cor 1-2005. Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems. Amendment 2:Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands. — IEEE, 28 February 2006.
6. Шахнович И. Стандарт широкополосного доступа IEEE 802.16-2004. Режим OFDMA и адаптивные антенные системы. — Электроника: НТБ, 2005, № 2, с. 46–52.
7. Портной С., Иванов А. Выбор систем ширкополосного беспроводного доступа. Мнение экспертов. — Первая миля, 2007, № 2, с. 8–11.
8. Койнов А. Решения SDR для аппаратуры WiMAX. ПЛИС, DSP или нечто иное? — Электроника: НТБ, 2007, № 2, с. 76–80.
9. M. Okuda et al. Multihop Relay Extension for WiMAX Networks — Overview and Benefits of IEEE IEEE 802.16j Standard. — FUJITSU Sci. Tech. Journal, 44, 3, July 2008, p. 292–302.
10. www.sotovik.ru/library/statistika_world.htm.
11. www.sotovik.ru/ratings/market20010101.htm.
12. Морелос-Сарагоса Р. Искусство помехоустойчивого кодирования. Методы, алгоритмы, применение. — М.: Техносфера, 2005.

13. Alamouti, S. M. A Simple Transmit Diversity Technique for Wireless Communications. — IEEE Journal on Select Areas in Communications. Oct. 1998, Vol. 16, №. 8, p. 1451–1458.
14. Концепция развития связи РФ/ Под ред. В. Б. Булгака, 1996.

ГЛАВА 8

АРХИТЕКТУРА СЕТЕЙ WiMAX

8.1. Основные принципы архитектуры сети WiMAX

Сеть WiMAX представляет собой совокупность беспроводного и базового (опорного) сегментов. Первый описывается в стандарте IEEE 802.16, второй же определяется спецификациями WiMAX-форума. Базовый сегмент — это все, что не относится к радиосети, т. е. связь базовых станций друг с другом, связь с локальными и глобальными сетями (в том числе с Интернетом) и т. п. Базовый сегмент основывается на IP-протоколах (IETF RFC) и стандартах Ethernet (IEEE 802.3-2005). Однако собственно архитектура сети, включая механизмы аутентификации, криптозащиты, роуминга, хэндовера и т. п. (в части, не относящейся к беспроводной сети), описывается в документах WiMAX-форума Network Architecture [1, 2].

Спецификации сети WiMAX основываются на технологии пакетной коммутации, протоколах IP и Ethernet, дополняя их по мере необходимости. Архитектура WiMAX-сети должна обеспечивать независимость архитектуры сети доступа, включая радиосеть, от функций и структуры транспортной IP-сети. Сеть WiMAX должна быть легко масштабируемой и гибко изменяемой, основываться на принципах декомпозиции (т. е. строиться на основе стандартных логических модулей, объединяемых через стандартные интерфейсы). Масштабируемость и гибкость возможна по таким эксплуатационным параметрам, как плотность абонентов, географическая протяженность зоны покрытия (районная, городская или пригородная сети), частотные диапазоны, топология сети (иерархическая, плоская, mesh и т. п.), мобильность абонентов (фиксированные, мобильные, номадические).

8.2. Базовая модель сети

Базовая модель сети WiMAX (БМ) — это логическое представление сетевой архитектуры WiMAX. Термин «логическое» в данном случае означает, что модель рассматривает набор стандартных логических функциональных модулей и стандартных интерфейсов (точек сопряжения этих модулей). При практической реализации одно устройство может включать несколько функциональных элементов или, напротив, функция может быть распределена между различными устройствами.

БМ включает три основных элемента — множество абонентских (мобильных) станций (МС), совокупность сетей доступа (сервисная сеть доступа ASN) и совокупность сетей подключения (CSN). Кроме того, в БМ входят так называемые базовые точки (R1–R8), через которые происходит сопряжение функциональных модулей (рис. 8.1). Сеть (сети) ASN принадлежит провайдеру сети доступа

(NAP) — организации, предоставляющей доступ к радиосети для одного или нескольких сервис-провайдеров WiMAX (NSP). В свою очередь, сервис-провайдер WiMAX — организация, предоставляющая IP-соединения и услуги WiMAX конечным абонентам. В рамках данной модели уже сервис-провайдеры WiMAX заключают соглашения с Интернет-провайдерами, операторами других сетей доступа, соглашения о роуминге и т. п. Сервис-провайдеры по отношению к абоненту могут быть домашними и гостевыми, каждый — со своей сетью CSN.

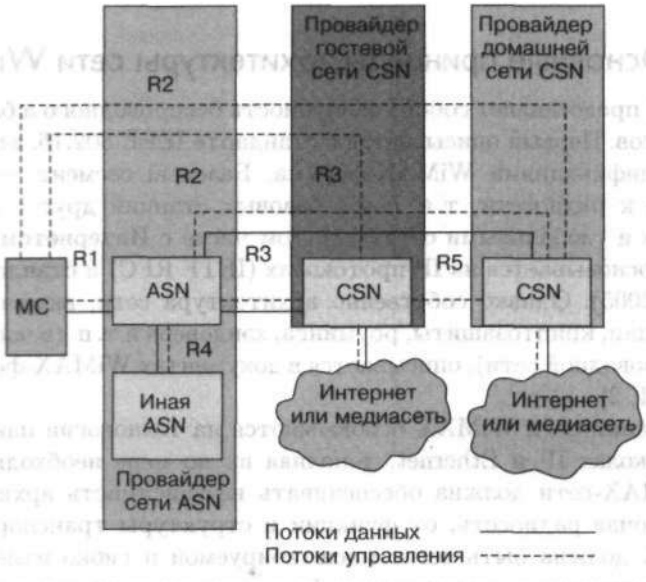


Рис. 8.1. Базовая модель WiMAX-сети в соответствии со стандартом [1]

Сеть доступа ASN представляет собой множество базовых станций (БС) беспроводного доступа по стандарту IEEE 802.16 и шлюзов для связи с транспортной IP-сетью (т. е. с локальной или глобальной сетью передачи информации) (рис. 8.2). Фактически эта сеть связывает радиосеть IEEE 802.16 и IP-сеть. ASN включает как минимум одну БС и как минимум один ASN-шлюз. Но и базовых станций, и шлюзов в одной ASN может быть несколько, причем одна БС может быть логически связана с несколькими шлюзами.

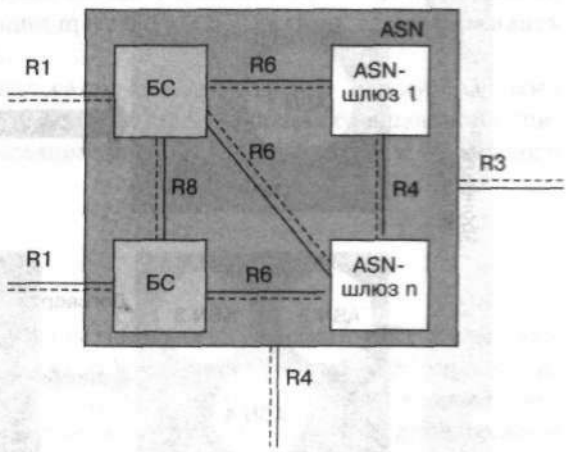
БС в рамках данной модели — это логическое устройство, поддерживающее набор протоколов IEEE 802.16 и функции внешнего сопряжения. Логическая БС — односекторная, с одним частотным номиналом. Очевидно, что реальная базовая станция представляет собой набор нескольких логических БС.

Шлюз ASN — это также логическое устройство, связывающее базовые станции одной ASN с другими сетями доступа и с сетью подключения CSN. Шлюз ASN обеспечивает связность как на уровне каналов передачи данных, так и на уровне управления. Примечательно, что для каждой MC базовая станция логически связана с одним шлюзом. Но реально функции ASN-шлюза для каждой MC могут быть распределены между несколькими шлюзами, принадлежащими одной или нескольким сетям доступа.

Шлюз ASN опционально может быть представлен как совокупность двух групп функциональных элементов — блока решения (DP — Decision Point) и бло-

ка исполнения (EP — Enforcement Point). EP реализует функции, связанные с передачей потока данных, в то время как в DP сосредоточены функции, непосредственно не относящиеся к передаче данных (например, функции контроллера управления радиоресурсами сети). Эти два функциональных модуля соединены через базовую точку R7. Зачем в стандарт введена такая модель, можно только догадываться. Нигде подробнее она не раскрыта, но без упоминания о возможности такой декомпозиции функций ASN-шлюза невозможно объяснить наличие R7. В целом распределение функций между реальными шлюзами и базовыми станциями определяется так называемыми профилями ASN. Сегодня описано три таких профиля (A, B и C), их мы рассмотрим ниже.

Рис. 8.2. Логическая модель сети доступа ASN



Сеть подключения CSN — это собственно сеть оператора WiMAX, именно в ней реализуются функции управления авторизацией, аутентификацией и доступом (AAA), подключение абонентов WiMAX к глобальным IP-сетям, предоставление таких услуг, как IP-телефония, доступ к телефонным сетям общего пользования, доступ в Интернет и частные сети и т. п. Важно отметить, что базовая модель сети WiMAX допускает, что одной сетью доступа ASN могут пользоваться несколько сервис-провайдеров WiMAX (каждый со своей CSN). И напротив — одна CSN может подключаться к сетям доступа разных провайдеров доступа.

В CSN реализованы такие функции, как предоставление мобильным абонентам IP-адресов и других сетевых параметров на период сетевой сессии, сервер политик/контроля доступа и хранения профилей абонентов, передача (туннелирование) данных между сетями доступа и подключения, биллинг абонентов WiMAX и межоператорские расчеты, туннелирование данных между различными CSN при роуминге, обеспечение мобильности при выходе MC за пределы одной ASN. Поддерживаются такие WiMAX-услуги, как соединения «точка-точка», авторизация и/или подключение к мультимедийным IP-сервисам, функции легального перехвата трафика (для России — выполнение требований СОРМ) и т. п.

CSN может включать такие элементы, как маршрутизаторы, серверы (и прокси-серверы) для функций авторизации/аутентификации/доступа, базы данных пользователей, шлюзы и т. п.

В связи с поддержкой мобильности в базовой модели сети WiMAX введены понятия домашних и гостевых сервис-провайдеров — H-CSP и V-CSP, соответственно (рис. 8.3). Домашний NSP — это оператор, заключивший договор об обслуживании с абонентом WiMAX. Именно он реализует функции авторизации, аутентификации и контроля доступа (включая биллинг и взимание абонентской платы). Для поддержки роуминга домашний сервис-провайдер WiMAX заключает роуминговые соглашения с другими NSP.

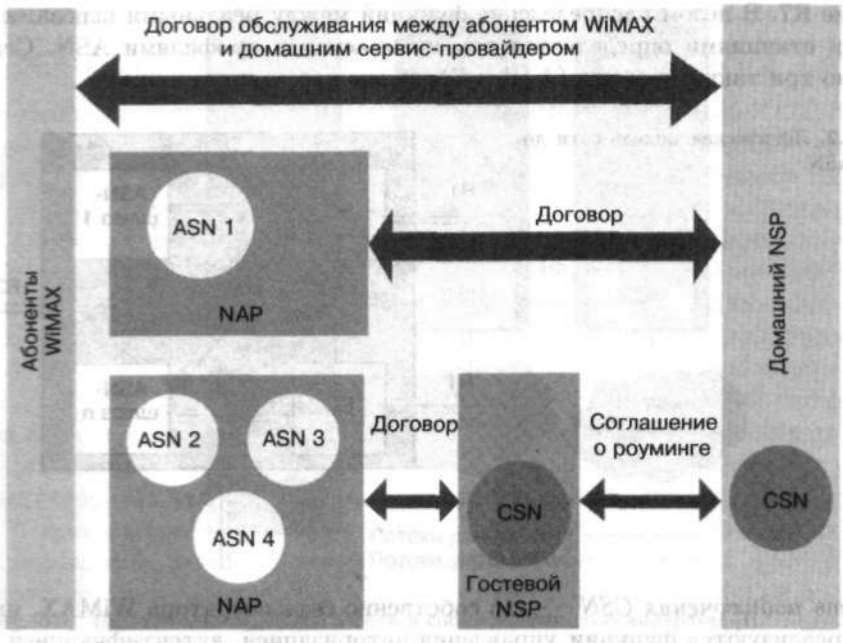


Рис. 8.3. Модель взаимодействия операторов сервисных сетей WiMAX, сетей доступа и абонентов

Гостевой NSP (V-NSP) — это оператор, который предоставляет WiMAX-абоненту услуги роуминга. Прежде всего, V-NSP обеспечивает для такого абонента функции AAA, а также полный или частичный доступ ко всем услугам WiMAX-сети. При этом возможны различные варианты маршрутизации трафика — через домашнюю сеть подключения или непосредственно через гостевую CSN-сеть.

Базовые точки в рамках базовой модели сети WiMAX — это каналы связи между базовыми модулями. Они представляют собой стандартные интерфейсы, причем не обязательно физические, особенно если соединяемые базовой точкой модули конструктивно находятся в одном устройстве.

Базовая точка R1 представляет собой канал связи между мобильной станцией и сетью доступа ASN. Это — беспроводной интерфейс, соответствующий стандарту IEEE 802.16, однако допустимы и дополнительные протоколы управления.

Базовая точка R2 является каналом между MC и CSN. Она включает протоколы и процедуры, связанные с аутентификацией MC, авторизацией и IP-конфигурированием. Это — чисто логический интерфейс, ему нельзя поставить в соответствие никакой конкретный физический интерфейс между MC и CSN.

Базовая точка R3 содержит набор протоколов управления между ASN и CSN для реализации процедур AAA, выполнения различных политик и управления мобильностью. Она также поддерживает функции передачи данных (в том числе — туннелирования) между ASN и CSN.

Базовая точка R4 — это канал связи между ASN-шлюзами различных ASN-сетей или между ASN-шлюзами в пределах одной ASN.

Базовая точка R5 является каналом связи между сетью домашнего и гостевого сервис-провайдера.

Базовая точка R6 служит интерфейсом между БС и ASN-шлюзом.

Базовая точка R7 определена как некий виртуальный канал внутри ASN-шлюза для связи двух групп функций (связанных с каналом передачи информации и не связанных с ним). Конкретизации протоколов R7, видимо, следует ожидать в будущем (или не ожидать вовсе).

Базовая точка R8 — это канал связи непосредственно между базовыми станциями. Он должен поддерживать передачу управляющих сообщений и опционально — непосредственную трансляцию данных (для быстрого и бесшовного хендовера).

8.3. Профили ASN

Профилями ASN называют распределение логических функций ASN-сетей между физическими устройствами. В стандарте описано три типа ASN-профилей. Профиль B подразумевает полную свободу производителя — ему соответствует как концентрация всех функций в одном устройстве, так и их произвольное распределение.

Профили A и C более конкретны. На уровне описания они чрезвычайно похожи — различие в том, что функции контроллера радиоресурсов (RRC) и управления хэндовером в профиле A отнесены к ASN-шлюзу, а в профиле C — к базовой станции. Несмотря на, казалось бы, незначительное формальное различие, на практике оно привело к тому, что профиль A был официально закрыт летом 2007 года на сессии WiMAX-форума в Мадриде, а общепризнанным стандартом стал профиль C (рис. 8.4).

Действительно, профиль A, концентрируя функции управления в ASN-шлюзе, затрудняет совместимость оборудования различных поставщиков. В профиле B интеллект базовых станций возрастает, они играют более существенную роль в управлении трафиком и мобильностью. Профиль C — наиболее открытая и потому перспективная система. В нем, в отличие от профиля A, базовые станции ответственны за все управление радиоресурсами и за обеспечение хендовера. В идеальном случае все элементы такой системы взаимозаменяемы продуктами других поставщиков, сертифицированных WiMAX-форумов.

8.4. Поддержка мобильности

Вся работа по описанию и стандартизации сетей WiMAX, по большому счету, направлена на одно — на обеспечение глобальной мобильности абонентов WiMAX, их свободы перемещаться между различными сетями во всем мире, постоянно

оставаясь «на связи». Для этого необходим механизм глобального распределения общих сетевых ресурсов между различными операторами-провайдерами.

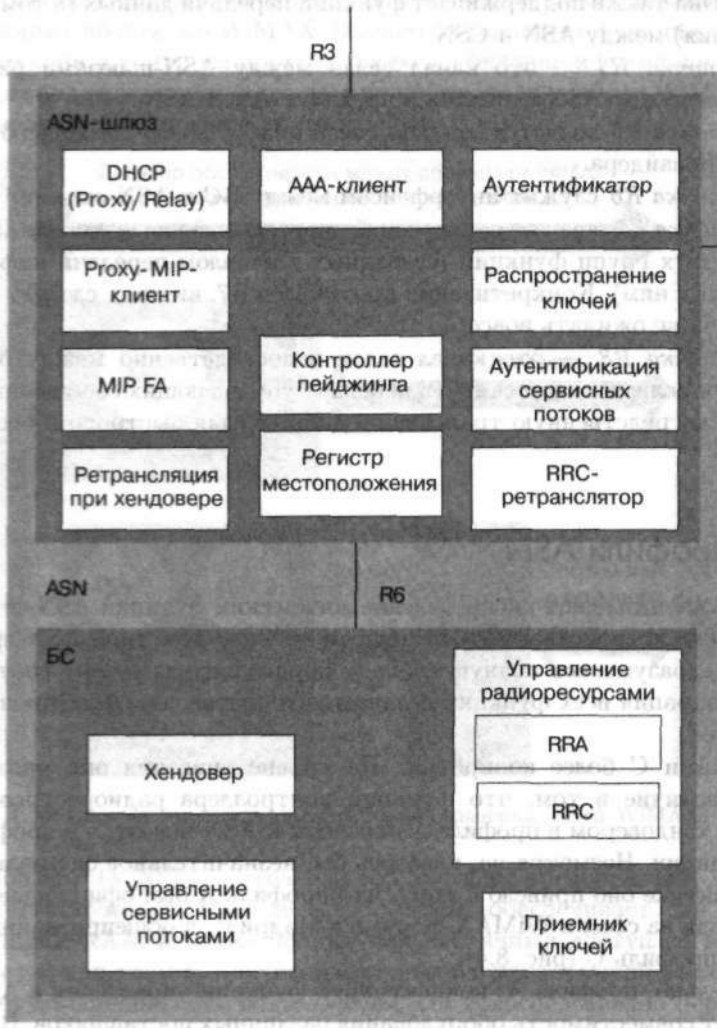
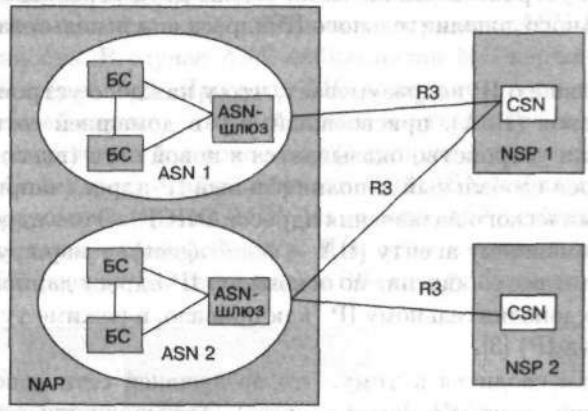


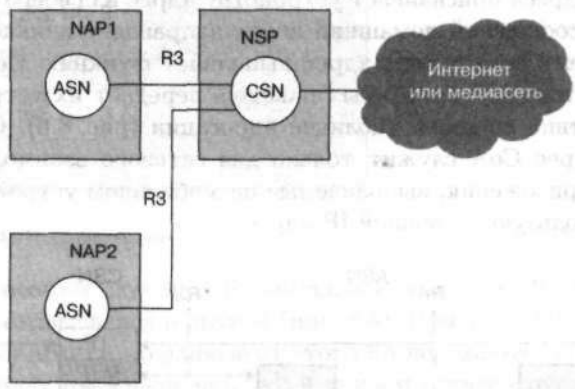
Рис. 8.4. Распределение основных логических функций между базовой станцией и ASN-шлюзом в соответствии с ASN-профилем C

Возможно несколько различных вариантов распределения сетевых ресурсов: одной ASN-сетью пользуются несколько CSN-провайдеров, несколько ASN-сетей (одного или нескольких операторов) взаимодействуют с различными CSN, одному оператору принадлежит ASN и CSN и т. п. (рис. 8.5). Очевидно, что при таком разнообразии вопросы стандартизации процедур при мобильности выходят на первый план.

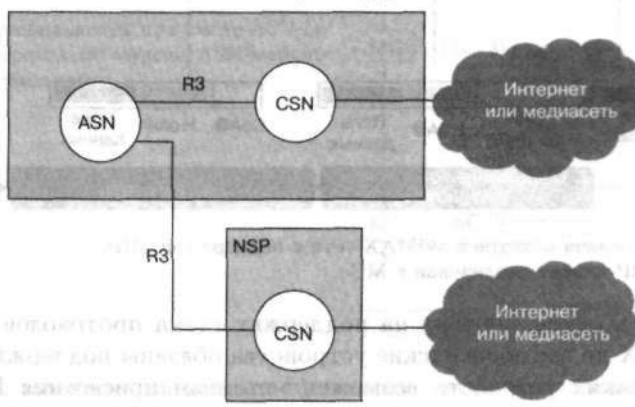
Напомним, что WiMAX-сеть — это TCP/IP-сеть. С точки зрения IP-сети, мобильность — это возможность идентифицировать устройство, подключающееся к различным узлам глобальной сети. Для поддержки мобильности были созданы спецификации мобильных IP-сетей (MIP). В мобильных IP-сетях задача обеспе-



Несколько NSP используют ASN-сети нескольких провайдеров



Один NSP, использующий ASN-сети нескольких провайдеров



ASN и CSN одного оператора с подключением CSN другого провайдера

Рис. 8.5. Некоторые варианты взаимодействия ASN-и CSN-сетей

чения мобильности устройств решается на основе двух основных механизмов – назначения глобального дополнительного IP-адреса или использования внешнего агента.

Протокол мобильного IP подразумевает, что у каждого устройства есть два IP-адреса — основной (HoA), присвоенный ему в домашней сети, и дополнительный (CoA). Если устройство оказывается в новой сети (внешней сети), ему может быть присвоен глобальный дополнительный IP-адрес (например, на основе протокола динамического назначения адресов DHCP). Этот адрес устройство сообщает своему домашнему агенту (HA — home agent) — маршрутизатору, который перехватывает все сообщения по основному IP-адресу данного устройства и направляет их по дополнительному IP (как правило, в режиме туннелирования и инкапсуляции IP-в-IP) [3].

Второй механизм сводится к тому, что во внешней сети используется так называемый внешний агент (FA, foregion agent). Это маршрутизатор, в котором устройство регистрируется при подключении к внешней сети. FA в качестве дополнительного IP-адреса присваивает устройству адрес из своего пула IP-адресов. При передаче сообщений домашний агент направляет инкапсулированные пакеты уже внешнему агенту (его адрес выполняет функцию CoA), который, отбросив оболочку инкапсулированных пакетов, передает их устройству-получателю в соответствии со своей таблицей адресации (рис. 8.6). Отметим, что дополнительный адрес CoA служит только для сетевого взаимодействия. Все пользовательские приложения, выполняемые на мобильном устройстве и в других узлах сети, используют основной IP-адрес.

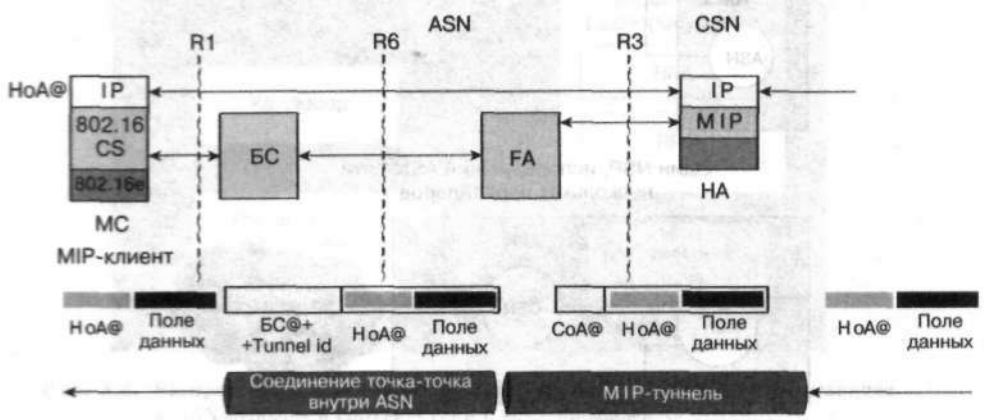


Рис. 8.6. Передача пакетов в WiMAX-сети с поддержкой MIP. MIP-клиент реализован в MC

WiMAX-сеть ориентирована на поддержку стека протоколов MIP. Однако в сетях WiMAX не все абонентские устройства обязаны поддерживать мобильный IP. Для таких устройств возможен механизм присвоения IP-адресов на основе протокола динамического конфигурирования DHCP (IETF RFC4361). Причем DHCP-сервер может находиться как в домашней, так и в гостевой сети. Возможно его размещение и в сети ASN, в этом случае информация об IP-адресе абонентской станции передается в домашнюю сеть при ее подключении и аутентификации.

В WiMAX-сетях выделяют два вида мобильности — микро- и макромобильность. Иначе их называют мобильность в рамках ASN (ASN-мобильность) и CSN-мобильность. В случае *ASN-мобильности* MC перемещается в пределах одной ASN-сети (рис. 8.7). При этом MC обслуживается одним внешним агентом (в простейшем случае — ASN-шлюзом) и с точки зрения CSN-сети никаких перемещений устройства не происходит (маршрут к нему остается неизменным, равно как и CoA-адрес). Таким образом, для ASN-мобильности не требуется поддержка протоколов уровня MIP. На уровне ASN-мобильности реализуется хэндовер в пределах одной ASN-сети. При этом в процесс вовлекаются только интерфейсы R6 (между БС шлюзами) и R8 (между базовыми станциями).

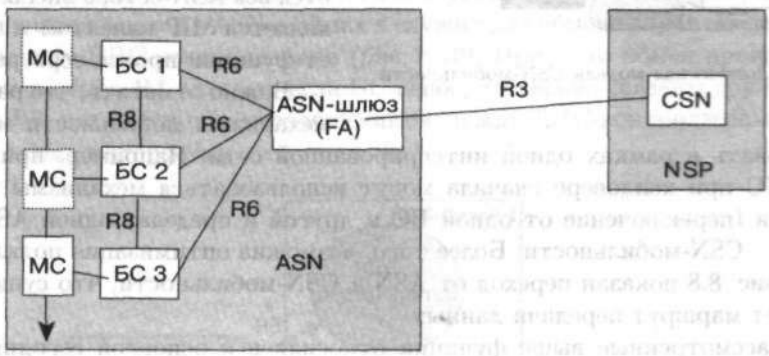
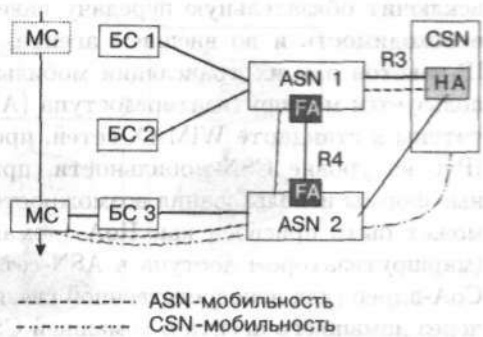


Рис. 8.7. ASN-мобильность при хендовере в пределах одной ASN-сети

Отметим особый случай ASN-мобильности, когда MC выходит за пределы одной ASN и оказывается в другой (рис. 8.8). При этом MC подключается к новому внешнему агенту, но данные от этого FA передаются к прежнему внешнему агенту по каналу R4. Очевидно, что в данном случае с точки зрения сети CSN (т. е. домашнего агента) никаких изменений не произошло.

Рис. 8.8. ASN-мобильность при смене ASN-сетей. Показан переход от модели ASN-мобильности к CSN-мобильности



Макромобильность означает смену внешнего агента, связанного с HA по каналу R3. Это возможно как внутри одной ASN-сети, так и при переходе между различными ASN-сетями (рис. 8.9). Смена внешнего агента однозначно влечет смену CoA-адреса MC. В этом случае изменения затрагивают сетевой уровень, т. е. уровень интерфейса R3. Поэтому данный вид мобильности еще называют R3-мобильность.

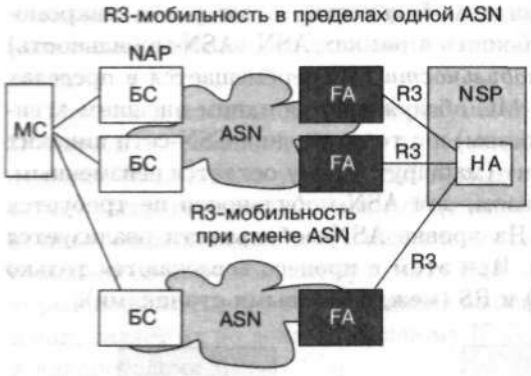


Рис. 8.9. Логическая модель CSN-мобильности

Поскольку MC могут не поддерживать функции мобильного IP, стандарт WiMAX-сетей предусматривает два сценария CSN-мобильности — с поддержкой MIP-клиентов (CMIP) и прокси-мобильный IP (PMIP). В первом случае MIP-клиент реализован в каждой мобильной станции, во втором — в качестве мобильного узла рассматривается вся ASN-сеть, а внешний агент является MIP-клиентом и выполняет функции прокси-сервера MIP.

Важно отметить, что различные механизмы мобильности могут сосуществовать в рамках одной интегрированной сети. Например, при перемещении MC при хендовере сначала могут использоваться механизмы ASN-мобильности (переключение от одной БС к другой в пределах одной ASN-сети), а затем — CSN-мобильности. Более того, возможна оптимизация подключения. Так, на рис. 8.8 показан переход от ASN-к CSN-мобильности, что существенно сокращает маршрут передачи данных.

Все рассмотренные выше функции относились к основной сегодня версии TCP/IP-протоколов — IPv4. Однако на смену им идет новая версия IPv6. Главная причина появления новой версии — нехватка адресного пространства, обусловленного 32-разрядным IP-адресом, а также отсутствие встроенной поддержки QoS. В новой версии предусматриваются 128-разрядные адреса. Кроме того, что важно для мобильных сетей, вводится так называемый альтернативный адрес. Он может быть присвоен группе устройств, распределенных в сети, но доставка пакета производится только до ближайшего узла (от отправителя) с таким адресом. Предусматривается оптимизирующая маршрутизация, которая исключит обязательную передачу пакетов через домашнего агента. Отпадает необходимость и во внешнем агенте, равно как и в инкапсуляции исходных IP-пакетов при их трансляции мобильному узлу. Вместо внешнего агента используется маршрутизатор доступа (AR — access router). Все эти особенности учтены в стандарте WiMAX-сетей, предусматривающем поддержку протокола IPv6 на уровне CSN-мобильности, причем стандарт оговаривает самые разные формы использования возможностей спецификаций IPv6. В частности, MC может быть присвоен как HoA (домашним агентом), так и глобальный CoA (маршрутизатором доступа в ASN-сети). При этом может использоваться или CoA-адрес (для непосредственной связи с заданным узлом), или HoA (для связи через домашнего агента в домашней CSN.)

8.5. Управление радиоресурсами

Функция эффективного управления радиоресурсами — одна из важнейших в любой беспроводной сети. Поскольку стандарт IEEE 802.16 рассматривает только взаимодействие одной БС с окружающими ее абонентскими станциями, вопро-

сы совместной работы нескольких базовых станций относятся к компетенции стандартов WiMAX-сетей. Эти функции сосредоточены в ASN-сегменте, физически — либо в базовой станции (профиль C), либо в ASN-шлюзе (профиль A).

Функции управления радиоресурсами реализуют два логических устройства — контроллер радиоресурсов (RRC — Radio Resource Controller) и агент радиосредств (RRA — Radio Resource Agent). В каждой БС (и только в БС) должен быть свой RRA. Напротив, контроллер RRC может располагаться как в базовой станции, так и в ASN-шлюзах или на отдельных серверах в пределах ASN-сети. Но поскольку фактически стандартным стал ASN-профиль C, будем рассматривать только размещение функций RRC в БС. В этом случае возникает потребность в дополнительном логическом устройстве — RRC-ретрансляторе, который располагается в ASN-шлюзах и служит для обмена управляющей информацией между RRC-контроллерами (рис. 8.10). При этом обмен происходит по интерфейсам R6 и R4. Однако, если БС непосредственно связаны друг с другом каналом R8, возможен обмен сообщениями между RRC-контроллерами данных БС и по этому интерфейсу.

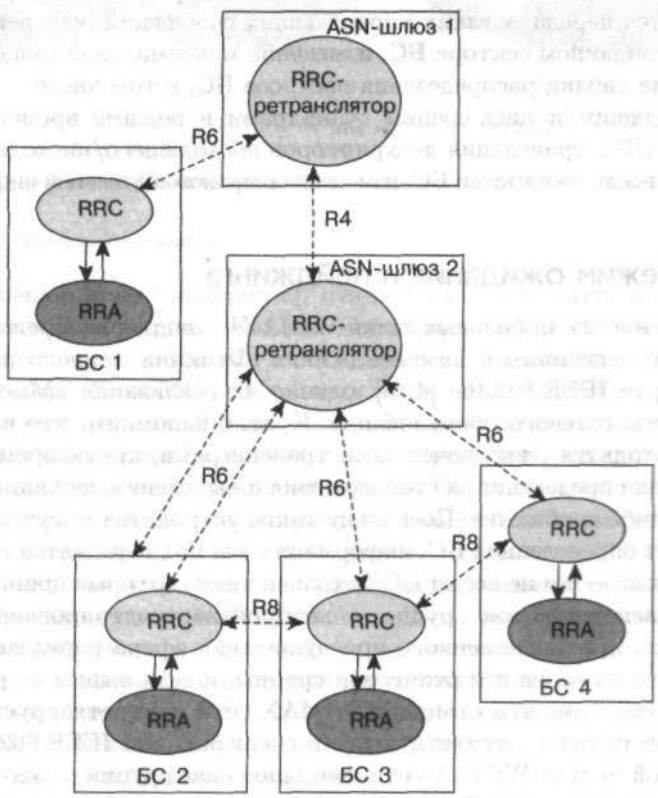


Рис. 8.10. Базовая модель системы управления радиоресурсами (для ASN-профиля C)

Основные функции, которые реализует агент радиосредств, — это управление и сбор информации о радиообстановке вокруг БС. Виды этих измерений и методика их проведения оговорены в стандарте IEEE 802.16. Кроме того, RRA

собирают и информацию об измерениях параметров протоколов верхних уровней, например, интенсивность ошибок передачи пакетов MAC-уровня. В задачу этого устройства входит и трансляция управляющей информации от RRC до мобильных станций по радиоинтерфейсу. Характерный пример такого рода информации — список соседских БС и их параметров. В свою очередь, основная функция контроллера RRC — сбор и хранение информации от связанных с ним RRA и взаимодействие с другими контроллерами RRC.

Таким образом, основная задача управления радиоресурсами — это инициирование процедур измерения характеристик радиосети, сбор этих параметров от всех БС и их сохранение в общедоступной базе данных сети. Эта информация используется для управления хендовером, для обеспечения качества обслуживания QoS и т. п. Основные виды измеряемых/сохраняемых параметров — физические параметры мобильных станций в сети, оценка качества связи с МС, наличные ресурсы базовых станций (число свободных субканалов и символов, усредненных по 200 фреймам, отдельно для восходящего и нисходящего каналов). Кроме того, стандартом оговорена возможность измерения таких параметров, как уровень мощности сигналов БС и уровень интерференции. Предусматривается передача таких управляющих сообщений, как реконфигурация субканалов в заданном секторе БС, изменение максимальной мощности сигнала БС, изменение таблиц распределения ресурсов БС, в том числе — соотношения между восходящим и нисходящим субкадрами в режиме временного дуплексирования (TDD), трансляция дескрипторов восходящего/нисходящего каналов (UCD/DCD) между соседскими БС, изменение широковещательной информации и др.

8.6. Режим ожидания и пейджинга

Важная особенность мобильных сетей WiMAX — поддержка режима ожидания (idle mode) и связанного с ним пейджинга. Функции данного режима описаны в стандарте IEEE 802.16e [4, 5], однако их реализация возможна только с использованием сетевого оборудования. Кратко напомним, что в режиме ожидания МС находятся с выключенными трансиверами, кратковременно включая приемники для определения местоположения и выяснения, не направлено ли в их адрес какое-либо сообщение. Поскольку такие устройства могут выйти из зоны обслуживания определенной БС, информация для них передается в широковещательном режиме — но не всеми БС, а только теми, которые принадлежат к так называемой пейджинговой группе, в которой зарегистрирована данная МС. МС в периоды кратковременного прослушивания эфира определяет, не вышла ли она за пределы своей пейджинговой группы, и если вышла — регистрируется в новой. Отметим, что стандарт WiMAX-сети конкретизирует определение пейджинговой группы, отсутствующее в спецификации IEEE 802.16e. С точки зрения базовой модели WiMAX-сети, пейджинговая группа — это множество из одного или нескольких БС. Причем к одной пейджинговой группе могут принадлежать только БС, находящиеся в пределах ASN-сети (сетей) одного провайдера.

Поддержка режима ожидания возложена на три логических элемента WiMAX-сети (рис. 8.11). Это пейджинговый агент (PA — Paging Agent), контроллер пейджинга (PC — Paging Controller) и регистр местоположения (LR — Location Register). Пейджинговый агент — это логическая функция базовой станции.

Она обеспечивает сопряжение протоколов пейджинга радиосети (MAC-уровень стандарта IEEE 802.16e) и протоколов опорной сети WiMAX. Последние поддерживает контроллер пейджинга. Он может располагаться как непосредственно в БС (ASN-профиль B), так и вне ее (профиль C). В последнем случае для связи PA и PC используется интерфейс R6.

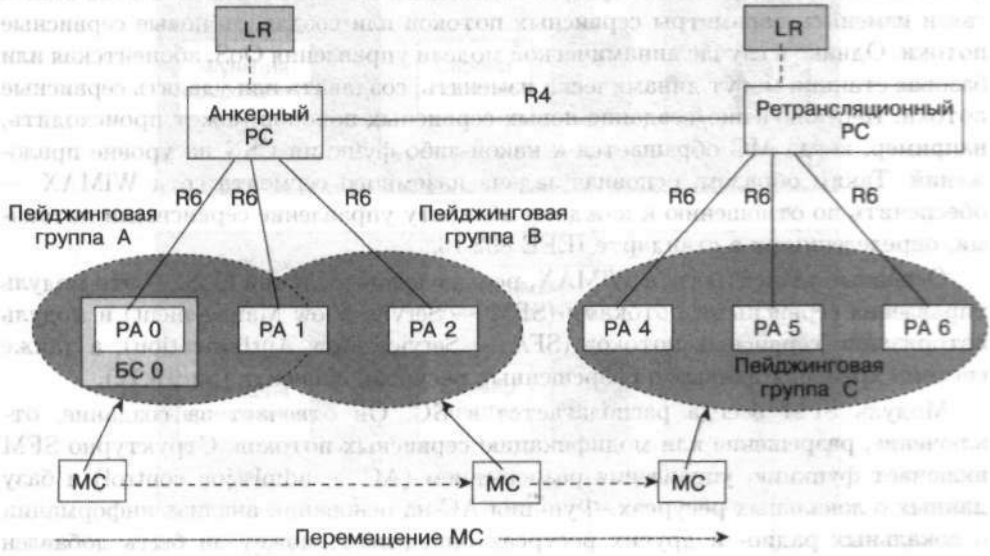


Рис. 8.11. Модель пейджинга

По отношению к MC контроллер пейджинга может быть анкерным (базовым) или ретрансляционным. Каждая мобильная станция в режиме ожидания связана только с одним анкерным PC. Он хранит и обновляет информацию о местоположении MC. Если MC оказывается в зоне действия БС, непосредственно не связанной ее анкерным PC по каналу R6, используется ретрансляционный пейджинговый контроллер (или несколько контроллеров), которые передают информацию о местоположении MC ее анкерному PC по каналам R4 (см. рис. 8.11).

Регистр местоположения — это распределенная база данных, с которой связан каждый анкерный PC. В этой базе данных для каждой MC хранится такая информация, как текущая пейджинговая группа, пейджинговый цикл и смещение, идентификатор последней БС и последнего ретрансляционного PC. Для MC в режиме ожидания в LR также хранятся сетевые настройки (согласно IEEE 802.16e) и информация о сервисных потоках данной MC. Эта информация используется анкерным PC для определения вероятного местоположения MC и передачи ей сообщений, а также для упрощения ее повторного подключения к сети после выхода из режима ожидания.

8.7. Качество обслуживания

Поскольку сети WiMAX изначально рассматривались как сети операторского класса, вопрос обеспечения QoS в них первичен. Стандарт IEEE 802.16 вопросы

QoS связывает с конкретным сервисным потоком. Каждое соединение обслуживается своим сервисным потоком, с заданными параметрами QoS. Абоненту WiMAX доступен заданный набор таких сервисных потоков — QoS-профиль. Информация об этом хранится в системе управления абонентами (например, в базе данных AAA-сервера или в специальном сервере политик). В случае статической модели управления QoS, абонентская станция не может в ходе сеанса связи изменять параметры сервисных потоков или создавать новые сервисные потоки. Однако в случае динамической модели управления QoS, абонентская или базовая станции могут динамически изменять, создавать или удалять сервисные потоки. Переключение/создание новых сервисных потоков может происходить, например, когда МС обращается к какой-либо функции CSN на уровне приложений. Таким образом, основная задача наземного сегмента сети WiMAX — обеспечить по отношению к каждому абоненту управление сервисными потоками, определенными в стандарте IEEE 802.16.

Основные элементы сети WiMAX, реализующие функции QoS, — это модуль управления сервисными потоками (SFM — Service Flow Management) и модуль авторизации сервисных потоков (SFA — Service Flow Authorization), а также система хранения данных о разрешенных ресурсах абонента (рис. 8.12).

Модуль SFM всегда располагается в БС. Он отвечает за создание, отключение, разрешение или модификацию сервисных потоков. Структурно SFM включает функцию управления разрешением (AC — admission control) и базу данных о локальных ресурсах. Функция AC на основании анализа информации о локальных радио- и других ресурсах определяет, может ли быть добавлен новый сервисный поток.

Модуль SFA предназначен для того, чтобы постоянно обеспечивать SFM заданной БС информацией о QoS-параметрах конкретного абонента. То есть он служит мостом между БС и глобальной базой данных о параметрах абонента, хранящейся в AAA-сервере или аналогичном устройстве в домашней CSN-сети абонента. Располагается это устройство в ASN-шлюзе. Поскольку речь идет о мобильных абонентах, вводится понятие анкерного (базового) и сервисного SFA.

Анкерный SFA определяется при подключении МС к сети и не меняется до ее повторной регистрации (остается неизменным в течение всей сессии). В анкерный SFA передается информация о QoS-профиле абонента при его регистрации в сети. Если МС оказывается в зоне другого ASN-шлюза, она взаимодействует уже с новым SFA. Такой SFA, с которым в данный момент связана МС, называется сервисным. Сервисный SFA по каналам R4 выполняет функцию ретранслятора между МС и ее анкерным SFA (точнее — между модулем SFM базовой станции, с которой в данный момент работает МС, и анкерным SFA для данной МС). В функции анкерного и/или сервисного SFM входит реализация так называемой локальной политики QoS для данной ASN-сети, связанной с загрузкой и распределением сетевых ресурсов.

Мы достаточно бегло рассмотрели основные принципы организации WiMAX-сетей. При этом вне пределов нашего внимания остались столь важные вопросы, как реализация процедур AAA для различных конфигураций WiMAX-сетей, механизмы назначения IP-адресов, процедуры соединения, хендовера и т. п. Отметим, что многие вопросы пока вообще не отражены в стандартах WiMAX-сети, но это, видимо, лишь дело времени. Процесс стандартизации WiMAX-сетей

отстает от их практической реализации только потому, что сама по себе технология WiMAX чрезвычайно нова, однако весьма востребована рынком.

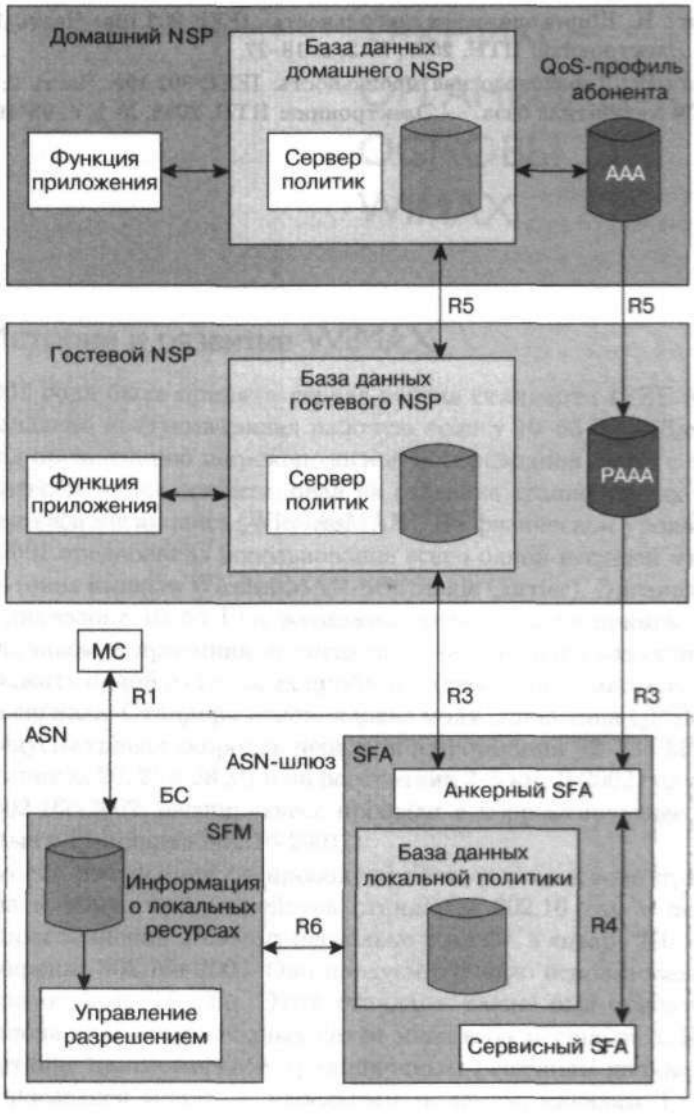


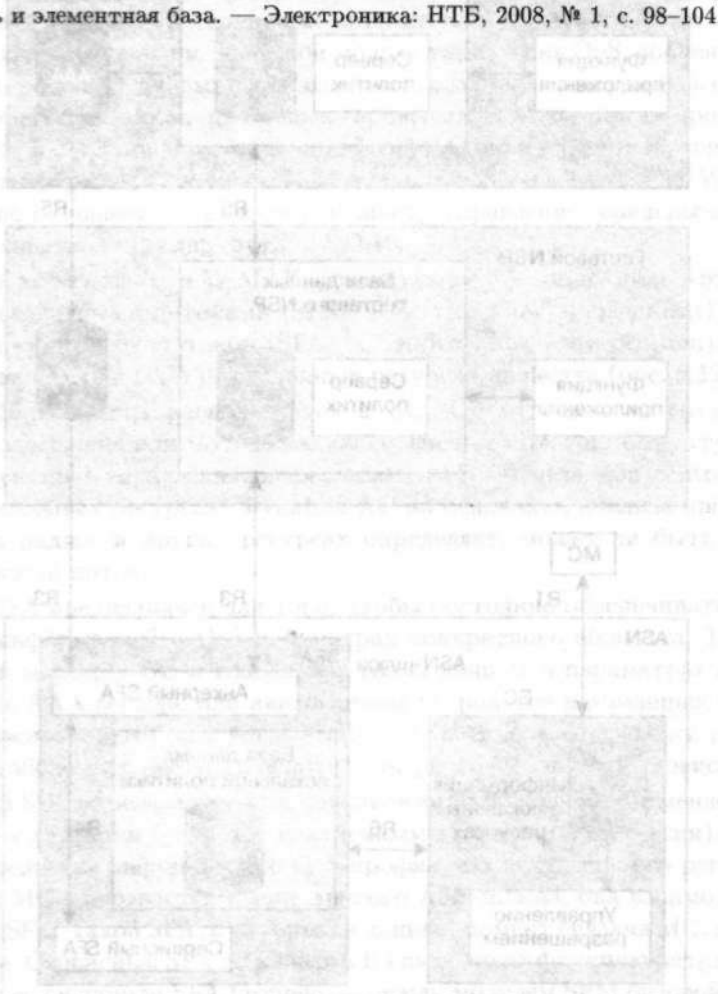
Рис. 8.12. Система обеспечения QoS

Литература

1. WiMAX Forum Network Architecture. (Stage 2: Architecture Tenets, Reference Model and Reference Points). Release 1, Version 1.2. — WiMAX Forum, January 11, 2008.
2. WiMAX Forum Network Architecture. (Stage 3: Detailed Protocols and Procedures).

Release 1, Version 1.2. — WiMAX Forum, January 11, 2008.

3. Дуглас Э. Камер. Сети TCP/IP, Т. 1. Принципы, протоколы и структура, 4-е изд. — М.: Издательский дом «Вильямс», 2003.
4. Шахнович И. Широкополосная мобильность: IEEE 802.16e. Часть 1: MAC-уровень. — Электроника: НТБ, 2007, № 2, с. 18–27.
5. Шахнович И. Широкополосная мобильность: IEEE 802.16e. Часть 2: физический уровень и элементная база. — Электроника: НТБ, 2008, № 1, с. 98–104.



ГЛАВА 9

ТЕХНИКО-ОРГАНИЗАЦИОННЫЕ ОСНОВЫ WiMAX

9.1. История и развитие WiMAX

В декабре 2001 года была принята первая версия стандарта IEEE 802.16-2001, который изначально предусматривал рабочую полосу 10–66 ГГц. Данный стандарт описывал организацию широкополосной беспроводной связи с топологией «точка-многоточка» и был ориентирован на создание стационарных беспроводных сетей масштаба мегаполиса (WirelessMAN). На физическом уровне стандарт IEEE 802.16-2001 предполагал использование всего одной несущей частоты, потому этот протокол назвали WirelessMAN-SC (Single Carrier). Организация связи в частотном диапазоне 10–66 ГГц возможна только в зоне прямой видимости между передатчиком и приемником сигнала из-за быстрого затухания. Но это позволяет избежать одной из главных проблем радиосвязи — многолучевого распространения сигнала. Стандарт рекомендовал модуляцию типа QPSK, 16-QAM, 64-QAM и предусматривал скорость передачи информации 32–134 Мбит/с в радиоканалах шириной 20, 25 и 28 МГц на расстоянии 2–5 км. В 2002 году появилось приложение 802.16c-2002, расширяющее профили и корректирующее погрешности, выявленные в стандарте 802.16-2001.

Необходимость построения беспроводной сети только в зоне прямой видимости привела к тому, что устройства стандарта 802.16 так и не получили широкого распространения. Поэтому несколько позднее, в январе 2003 года, было принято расширение 802.16a-2003. Оно предусматривало использование частотного диапазона от 2 до 11 ГГц. Этот стандарт также был ориентирован на создание стационарных беспроводных сетей масштаба мегаполиса. Планировалось, что он станет альтернативой традиционным решениям широкополосного доступа для «последней мили» — кабельным модемам, каналам T1/E1, xDSL и т. п. Кроме того, предполагалось, что к базовой сети стандарта 802.16a станут подключаться точки доступа стандарта 802.11b/g/a для формирования глобальной сети беспроводного доступа в Интернет.

Основным отличием стандарта 802.16a стала работа в частотном диапазоне, который не требует прямой видимости между приемником и передатчиком. Зона покрытия таких беспроводных сетей значительно шире, чем сетей стандарта 802.16. Использование частотного диапазона 2–11 ГГц потребовало и существенного пересмотра техники кодирования и модуляции сигнала на физическом уровне. Система на базе 802.16a должна была работать с модуляцией QPSK, 16-, 64- и 256-QAM, обеспечивать скорость передачи информации 1–75 Мбит/с

на сектор одной базовой станции в радиоканалах с изменяемой полосой пропускания от 1,5 до 20 МГц на расстоянии 6–9 км (теоретически до 50 км). Типовая базовая станция имела до шести секторов.

Был сохранен режим работы на одной несущей (SCa), предназначенный как для условий прямой видимости, так и вне ее. Но самое главное — предусматривались режимы на основе технологии ортогонального частотного мультиплексирования (OFDM) с 256 поднесущими и режим с технологией многостанционного доступа с ортогональным частотным разделением каналов (OFDMA — Orthogonal Frequency Division Multiple Access) с 2048 поднесущими.

Объединил все нововведения принятый в июле 2004 года стандарт IEEE 802.16-2004. Правда, с полной совместимостью разношерстных режимов мультиплексирования SC, SCa, OFDM и OFDMA, разной ширины радиоканалов, а также FDD, TDD и других требований возникли сложности, поэтому оборудование каждого производителя так и осталось уникальным. Это означало, прежде всего, высокую стоимость абонентских устройств. В результате оборудование IEEE 802.16-2004 фиксированного доступа используется исключительно в нишевых применениях (по отношению к глобальным системам мобильной связи), т. е. там, где другие традиционные методы построения сетей абонентского доступа невозможны или не эффективны.

Параметры физических уровней европейских стандартов HyperMAN (2–11 ГГц, 256 OFDM) и HyperACCESS (11–42 ГГц, SC TDM) во многом соответствуют 802.16-2004.

Стандарт IEEE 802.16e (он же — IEEE 802.16-2005) принят в конце 2005 года. Он рассматривает вопросы поддержки мобильных абонентов, в том числе — систему роуминга между сетями различных беспроводных стандартов. Последнее, в частности, позволяет без разрыва сеанса связи переходить из беспроводных сетей стандарта IEEE 802.11 в сети IEEE 802.16 и обратно. Если стандарт IEEE 802.16-2004 — это протокол операторского класса, то дополнение IEEE 802.16e ориентировано на конечных пользователей, причем мобильных, и в этом смысле оно представляет собой альтернативу стандартам 802.11a/b/g. Таким образом, пользователь, имея ноутбук или КПК со встроенным модемом IEEE 802.16e, сможет постоянно оставаться на связи в любой точке города. Базовая станция мобильного WiMAX поддерживает около 1000 абонентов одновременно. Точное их число зависит от ширины полосы, выделенной для каждого пользователя, и всего спектра, используемого оператором.

Помимо основных стандартов, рабочая группа IEEE 802.16 разработала ряд других документов, рассматривающих хоть и частные, но весьма важные вопросы. Это такие дополнения, как:

- 802.16f-2005 — Информационная база управления (Management Information Base);
- 802.16g-2007 — Процедуры и сервисы уровня управления (Management Plane Procedures and Services);
- 802.16k-2007 — Поправки к 802.16 (Bridging of 802.16).

В стадии разработки находятся:

- 802.16h — Улучшенный механизм сосуществования при безлицензионной работе (Improved Coexistence Mechanisms for License-Exempt Operation);

- 802.16i — Информационная база управления для мобильных сетей (Mobile Management Information Base);
- 802.16j — Спецификация многопролетных ретрансляционных систем (Multihop Relay Specification);
- 802.16m — Улучшенный беспроводной интерфейс (Advanced Air Interface).

Часто стандарт IEEE 802.16 называют WiMAX. Аббревиатура WiMAX расшифровывается как протокол всемирной сети широкополосной радиосвязи (Worldwide Interoperability for Microwave Access). Это — название международной организации WiMAX-форум (www.wimaxforum.org), в которую входит ряд ведущих телекоммуникационных и полупроводниковых компаний (Airspan Networks, Alvarion, Aperto Networks, Fujitsu Microelectronics America, Intel, Proxim Corporation и др.). Однако следует помнить, что на самом деле WiMAX, равно как и европейский HIPERMAN, рассматривает только часть режимов стандарта IEEE 802.16.

WiMAX-форум был организован 11 апреля 2003 года. Его целью является содействие разработке беспроводного оборудования для доступа к широкополосным сетям, скорейшее развертывание сетей во всем мире и сертификация оборудования IEEE 802.16, а также подготовка спецификаций, призванных обеспечить совместимость оборудования разных производителей. Одна из целей WiMAX — дальнейшее разделение труда на рынке производителей беспроводного оборудования. Поставщики элементной базы (Intel, Fujitsu и др.) будут разрабатывать ее для всех производителей оборудования, а те смогут сконцентрировать усилия на оборудовании со стандартной элементной базой. По данным аналитиков, члены WiMAX-форума представляют собой более 75% рынка производителей оборудования широкополосного беспроводного доступа.

В июне 2008 года компании Alcatel-Lucent, Cisco, Clearwire, Intel, Samsung Electronics и Sprint объявили о создании нового стратегического консорциума Open Patent Alliance (OPA), который займется стандартизацией в области технологии WiMAX. Позже к нему присоединились Alvarion и Huawei. Альянс постарается сделать все возможное для снижения стоимости оборудования и услуг, а также для расширения их многообразия. Для осуществления поставленной задачи будет создан так называемый патентный пул — специальное соглашение о взаимном использовании патентов, которыми сможет воспользоваться любой из членов альянса по предсказуемой цене. Данный пул включит все необходимые патенты для производства оборудования и развертывания сетей WiMAX. Членам OPA будет предложена гибкая схема лицензионных отчислений, включая перекрестное лицензирование внутри патентного пула. Помимо этого будут созданы специальные учебные материалы, а также центральный ресурс, на котором можно будет познакомиться с интеллектуальной собственностью компаний.

Основные этапы развития стандарта 802.16 и WiMAX-форума:

- август 2005 года — открытие первой сертификационной лаборатории (AT4 wireless);
- август 2005 года — первый PlugFest («фестиваль подключений», неофициальное тестирование на совместимость, без выдачи сертификата);
- январь 2006 года — сертификация первых продуктов WiMAX;
- март 2007 года — публикация WiMAX сетевой спецификации версии 1.0;

- конец 2007 года — Генеральная ассамблея ITU-R в Женеве приняла IEEE 802.16 в качестве шестого стандарта IMT-2000;
- июнь 2008 года — сертификация мобильного оборудования для диапазона 2,5 ГГц;
- сейчас WiMAX-форум насчитывает более 500 членов, из них 17 — из России и СНГ.

В целом, перед технологиями WiMAX сегодня стоят две главные задачи: в странах с недостаточной инфраструктурой связи — обеспечение доступа в глобальные сети, в развитых странах — обеспечение мобильности высокоскоростного доступа. Именно эти два направления и обусловливают экономическую целесообразность внедрения этой технологии.

9.2. Сертификация WiMAX

Основой идеологии WiMAX-форума являются:

- сертификация оборудования ШБД, т. е. стандартизация требований к оборудованию базовых и абонентских станций различных производителей;
- разработка архитектуры сети, алгоритмов аутентификации и авторизации для обеспечения совместимости;
- требования к физическому уровню: частотному диапазону, режиму дуплексирования и ширине канала.

Основные принципы сертификации:

- возможность совместного использования с оборудованием других производителей;
- обратная совместимость оборудования (поддержка предыдущих версий сертификации в одном и том же профиле);
- учет ситуации с частотами по всему миру при сертификации абонентских устройств.

Данные принципы сертификации выгодны как для производителей оборудования, для телекоммуникационных операторов, так и для пользователей. Первые быстрей отлаживают и разрабатывают оборудование, концентрируются только на самых важных направлениях, быстрей получают финансовую отдачу. Операторы не являются заложниками одного производителя, могут строить более гибкие сети, могут отдать распространение абонентских устройств другим участникам рынка, тем самым также сконцентрироваться на более важных направлениях. Пользователи же получают большую свободу и меньшие цены.

Совокупность требований к физическому уровню получила название «профиль WiMAX» (табл. 9.1).

9.2.1. Процедура сертификации

Сама процедура сертификации осуществляется по следующей цепочке. Сначала производитель оборудования участвует в плагфесте, результаты испытаний на котором обычно конфиденциальны. Дальше происходит выбор лаборатории

(табл. 9.2) и сертификационного профиля, начинается взаимодействие с лабораторией. Проходят основные сертификационные испытания. Издается сертификат, он публикуется и заносится в регистр WiMAX-форума.

Таблица 9.1. Профили WiMAX

	Фиксированный WiMAX	Эволюционный WiMAX	Мобильный WiMAX	
Стандарт радиоинтерфейса	IEEE 802.16-2004	IEEE 802.16e-2005	IEEE 802.16e-2005	
Мультиплексирование	OFDM	OFDM	OFDMA	
Номинальное число поднесущих	256	256	512, 1024	
Дуплексный режим	TDD, FDD, HFDD	TDD, FDD, HFDD	TDD	
Модуляция	BPSK, QPSK, 16-QAM, 64-QAM	BPSK, QPSK, 16-QAM, 64-QAM (опционально)	QPSK, 16-QAM, 64-QAM (в восходящем канале — опционально)	
Классы мощности, дБм	0–14	0–14	16 QAM	QPSK
	14–17	14–17	18–21	20–23
	17–20	17–20	21–25	23–27
	20–23	20–23	25–30	27–30
	Свыше 23	Свыше 23	Свыше 30	Свыше 30

Вся дополнительная информация доступна для членов WiMAX-форума по адресу http://members.wimaxforum.org/members/certification.

Таблица 9.2. Сертификационные лаборатории WiMAX-форума

№	Название лаборатории	Адрес
1	AT4 Wireless (Испания, США)	www.at4wireless.com Parque Tecnologico de Andalucia (Spain) Calle Severo Ochoa 2 29590 Campanillas, Málaga, Spain и 520B Huntmar Park Drive, Herndon, VA 20170 (USA)
2	Bureau Veritas ADT (Тайвань)	www.adt.com.tw No. 19, Hwa Ya 2nd Rd, Wen Hwa Tsuen, Kwei Shan Hsiang, Taoyuan Hsien 333, Taiwan
3	CCS/TTC (Китай)	Compliance Certification Services (CCS) www.ttc.org.tw 11 Wugong 6th Rd., Wugu Industrial Park Taipei County 248, Taiwan и Telecom Technology Center (TTC) www.ccsemc.com.tw 4F, No.300, Yangguang St. Neihu District, Taipei City 114, Taiwan
4	China Academy of Telecommunication Research (Китай)	www.catr.cn 52 Hua Yuan Bei Lu, Haidian District, Beijing 100083, China
5	Telecommunications Technology Association (Южная Корея)	www.tta.or.kr 267-2 Seohyun-dong, Bundang-gu, Seongnam-City, Gyeonggi-do, 463-824 Korea

Собственно сертификация оборудования в WiMAX-форуме предусматривает три этапа. На первом этапе производитель, желающий участвовать в сертификации, регистрирует свое оборудование в соответствующем профиле. Для начала

сертификации необходимо три производителя в выбранном профиле для фиксированного WiMAX и четыре производителя — для мобильного WiMAX. На втором этапе оборудование каждого производителя проходит аттестационные испытания, которые состоят из двух основных частей: проверка на соответствие стандарту MAC-уровня и PHY-уровня. Завершающий этап — тестирование на функциональную совместимость оборудования разных изготовителей: к базовой станции одного производителя подключаются абонентские устройства других производителей.

Последний этап тестирования разделен WiMAX-форумом на две волны, различные для фиксированного и мобильного WiMAX. В фиксированном WiMAX в первой волне сертификации проверяются ключевые возможности подключения к сети, предоставление сервиса, распределение полосы пропускания. Во второй волне тестируются процедуры аутентификации, совместимости протоколов радиосвязи, возможность выделять гарантированную полосу пропускания каждому пользователю (QoS). Также проверяются алгоритмы шифрования AES и корректность работы алгоритма автоматического запроса повторной передачи (ARQ).

В случае мобильного WiMAX в первой волне проверяется соответствие требованиям стандарта параметров режима OFDMA, протоколов QoS, AES, H-ARQ (гибридный ARQ), протокола управления ключами шифрования для обеспечения авторизации абонентских устройств на базовой станции (PKMv2), алгоритма контроля мощности абонентских устройств, режимов ожидания (sleep и idle mode) и сжатия заголовков.

Во второй волне испытаний проверяется хэндовер, QoS более высокого уровня, поддержка протокола IPv6, максимальный размер передаваемого фрейма (MBS). Кроме этого, возможна проверка механизмов MIMO, таких, как многоканальный прием/передача (как на базовой, так и на абонентской станциях) и формирование луча диаграммы направленности антенн.

При переходе от фиксированного к мобильному WiMAX предъявляются дополнительные требования к оборудованию и технологии в целом. Так, в мобильном WiMAX уже используется мультиплексирование OFDMA, позволяющее передавать в одном кадре информацию для нескольких абонентских станций. Для увеличения зоны покрытия вне прямой видимости требуется увеличить мощность передатчика. Такой подход применим на базовой станции, но увеличение мощности абонентского устройства невозможно по санитарным нормам. Поэтому необходимо внедрение технологии MIMO-антенн на базовой станции, существенно увеличивающее стоимость системы в целом, но обеспечивающее дополнительный энергетический выигрыш в 3–6 дБ по уровню принимаемого сигнала от абонентской станции. Дополнительно увеличить зону покрытия возможно за счет формирования диаграммы направленности антенны в требуемом направлении. Эта технология позволяет получить более мощный сигнал за счет высокого коэффициента анизотропного усиления антенной системы. С появлением мобильности возникает потребность в алгоритмах безопасной аутентификации и авторизации, как абонентских устройств, так и пользователей и быстрой перерегистрации при переходе от одной базовой станции к другой без потери связи.

Для прохождения сертификационных испытаний базовая станция должна работать как минимум с тремя абонентскими устройствами различных про-

изводителей. А абонентское оборудование — как минимум с двумя базовыми станциями различных изготовителей (разумеется, при условии, что профиль оборудования одинаков).

По прогнозам WiMAX-форума, к 2011 году сертифицированных продуктов станет более 1000 (рис. 9.1). Впервые на рынке сетей ШБД операторского класса клиенту предоставляется свобода выбора оборудования различных производителей. В этих целях WiMAX-форум окончательно утвердил в 2008 году сертификационную программу интенсивного тестирования абонентских устройств и базовых станций. Работа ведется в сотрудничестве с институтами IEEE и ETSI, с тестовыми лабораториями и более чем с 500 членами WiMAX-форума. На первых этапах сертификация концентрируется на физическом и MAC-уровнях сети, позже она достигнет и более высоких уровней.

Рис. 9.1. Прогноз сертификации WiMAX-форума

При сертификации должны выполняться следующие тесты (табл. 9.5):

- тестирование на физическом уровне Radio Conformance Testing (RCT);
- тестирование на MAC уровне Protocol Conformance Testing (PCT), включая подуровни безопасности, управления и сервиса;
- тесты на совместную работу (IOT), в том числе — и в мобильном режиме (MIOT);
- тесты на сетевое взаимодействие Network Conformance Testing (NCT);
- с июля 2009 года вводятся измерения уровней электромагнитного излучения абонентских устройств Radiated Performance Testing (RPT);

– со второй половины 2009 года начались тесты элементов инфраструктуры WiMAX-сетей Infrastructure Interoperability Testing (IIOT). Внимание обращается на ASN- и CSN-элементы сети и их совместную работу.

Сегодня действуют пять сертификационных лабораторий WiMAX-форума (табл. 9.2).

9.2.2. Сертификационные профили. Динамика сертификации

Сразу оговоримся, что список сертификационных профилей WiMAX-форума постоянно претерпевает изменения. С одной стороны, в него добавляются новые профили, отражающие появление новых спецификаций, новых версий стандартов, а также увеличение функциональности устройств. С другой стороны, отдельные WiMAX-профили перестают получать поддержку со стороны производителей аппаратуры — соответствующие им устройства перестают разрабатывать (хотя уже разработанные могут выпускаться). Поэтому возникает понятие «активный WiMAX-профиль» — т. е. такой профиль, который реально используется разработчиками новой аппаратуры.

На начало 2007 года WiMAX-продукты могли быть сертифицированными в соответствии с пятью фиксированными WiMAX-профилями и тринадцатью мобильными WiMAX-профилями (табл. 9.3). Каждый сертификационный профиль соответствует специфическому распределению частот.

Таблица 9.3. Сертифицированные профили WiMAX

Системы WiMAX	Сертифицированные профили			
	Название	Частотный диапазон, ГГц	Дуплекс	Ширина канала, МГц
Фиксированный WiMAX (IEEE 802.16-2004, OFDM)	Air1	3,4–3,6	TDD	3,5
		3,4–3,6	TDD	7
	Air1	3,4–3,6	FDD	3,5
		3,4–3,6	FDD	7
		5,725–5,850	TDD	10
Эволюционный WiMAX (IEEE 802.16e-2005, OFDM)	ETG8	4,935–4,990	TDD	5
Мобильный WiMAX (IEEE 802.16e-2005, OFDMA)	1B	2,3–2,4	TDD	5 и 10
	1A	2,3–2,4	TDD	8,75
	3A	2,496–2,690	TDD	5 и 10
	5AL	3,4–3,6	TDD	5
	5BL	3,4–3,6	TDD	7

В 2007 году состоялись две волны сертификации оборудования фиксированного WiMAX, были сертифицированы 28 типов оборудования следующих производителей:

– **Alvarion** — АС: BMAX PRO-S CPE, BreezeMAX Si; БС: BreezeMAX Macro Modular Base Station, BreezeMAX Micro Base Station;
– **Airspan** — АС: EasyST, MicroMAX-SoC BSR; БС: MacroMAX;

- **Axxcelera** — АС: ExcelMAX FD CPE;БС: ExcelMAX BS;
- **Aperto** — АС: PacketMAX, ProST; БС: PacketMAX 5000;
- **Proxim** — Tsunami MP16 3500;
- **Redline** — АС: RedMAX Subscriber Station; БС: RedMAX Base Station;
- **Siemens** — АС: Gigaset SE461 WiMAX; БС: WayMAX@vantage;
- **Sequans** — АС: SQN 1010-RD (FDD); БС: SQN2010-RD;
- **SR Telecom** — БС: Symmetry Base Station;
- **Selex** — АС: YSEMAX БС: YSEMAX;
- **Telsima** — БС: StarMAX 2140-3.5G, StarMAX 4120-3.5G;
- **Wavesat** — АС: Wavesat miniMAX 3.5GHz (FDD), Wavesat miniMAX 3.5GHz (TDD).

Все вышеперечисленное оборудование было сертифицировано в двух фиксированных профилях. Профиль Air1 характеризуется частотным диапазоном 3,5 ГГц, шириной канала 3,5 МГц и режимом дуплексирования с временным разделением (TDD). Профиль Air2 отличается от него режимом дуплексирования с частотным разделением (FDD).

В конце 2007 года на Тайване прошел четвертый WiMAX Forum Plugfest, где были проведены многочисленные тесты и было принято руководство по сценариям для сопряжения поставщиков оборудования (в том числе по биллингу). Там были проведены испытания по шести профилям мобильного WiMAX (табл. 9.4).

Таблица 9.4. Сертификационные профили мобильного WiMAX на четвертом WiMAX Forum Plugfest (Тайвань) и число сертифицированных устройств

Профили	Частотный диапазон, ГГц	Ширина полосы, Мгц	Вид дуплекса	Сертифицировано	
				абонентских устройств	базовых станций
1A	2,3–2,4	8,75	TDD	5	3
3A	2,496–2,69	5		20	10
3A	2,496–2,69	10		24	19
5A	3,4–3,8	5		2	1
5AL	3,4–3,6	5		7	6
5BL	3,4–3,8	7		3	2
Всего	6 профилей			61	41

В 2008 году и в начале 2009 года было сертифицировано оборудование мобильного WiMAX диапазонов 2,3, 2,5 и 3,5 ГГц в профилях 1A, 1B, 3A, 5AL, 5BL и ETG8 (см. табл. 9.3). Уже более 60 компаний разрабатывают абонентское оборудование и чипсеты и около 40 — базовое и инфраструктурное оборудование. Информация о новом сертифицированном оборудовании регулярно обновляется на официальном сайте WiMAX-форума www.wimaxforum.org и www.wimaxforum.ru.

В июне 2008 года в Амстердаме WiMAX-форум сертифицировал оборудование для базовых станций IEEE 802.16e, производимых компаниями Samsung, Alvarion и Motorola. Кроме того, также сертифицировано оборудование NEC и Alcatel-Lucent. Проблемы были с освоением оборудования, соответствующего ASN-профилю C. Еще в мае 2008 года компании не представляли на сертификацию такие устройства, однако уже к осени его освоили все ведущие производители (или включили в состав своего оборудования сертифицированные базовые станции других поставщиков, как сделал Nortel, использовав продукцию Alvarion).

WiMAX-форум сегодня занимается продвижением исключительно систем IEEE 802.16e, работающих в указанных выше диапазонах частот. Отметим, что системы ШБД в диапазоне 5–6 ГГц, распространенные в РФ, не относятся к WiMAX. И рассматривать диапазон 5 ГГц WiMAX-форум не намерен, несмотря на запросы производителей, работающих на российском рынке. Этот рынок слишком мал и потому не очень интересен для международного консорциума: по данным WiMAX-форума, на примерно 160 развернутых в России сетей фиксированного WiMAX и pre-WiMAX приходится едва ли 50 тыс. абонентов.

К середине 2009 года продукты сертифицировались в соответствии с релизом 1.0, который основан на стандарте IEEE 802.16e-2005. На этот момент были активны два сертификационных профиля для фиксированного WiMAX и шесть — для мобильного WiMAX (табл. 9.6). Это означает, что по ним реально проводится сертификация оборудования. Однако это никоим образом не отменяет предшествующие сертификационные профили — просто в настоящий момент нет разработок соответствующего им оборудования.

В 2009 году ожидается появление релиза 1.5 (рис. 9.2) как расширение релиза 1.0 в области сетевых спецификаций. Он будет основан на стандарте IEEE 802.16e Revision 2 и соответствующей модификации системного профиля мобильного WiMAX. В релиз 1.5 войдут новые FDD-профили в соответствии с требованиями тех стран, где регулятор обязывает использовать именно частотное дуплексирование каналов.

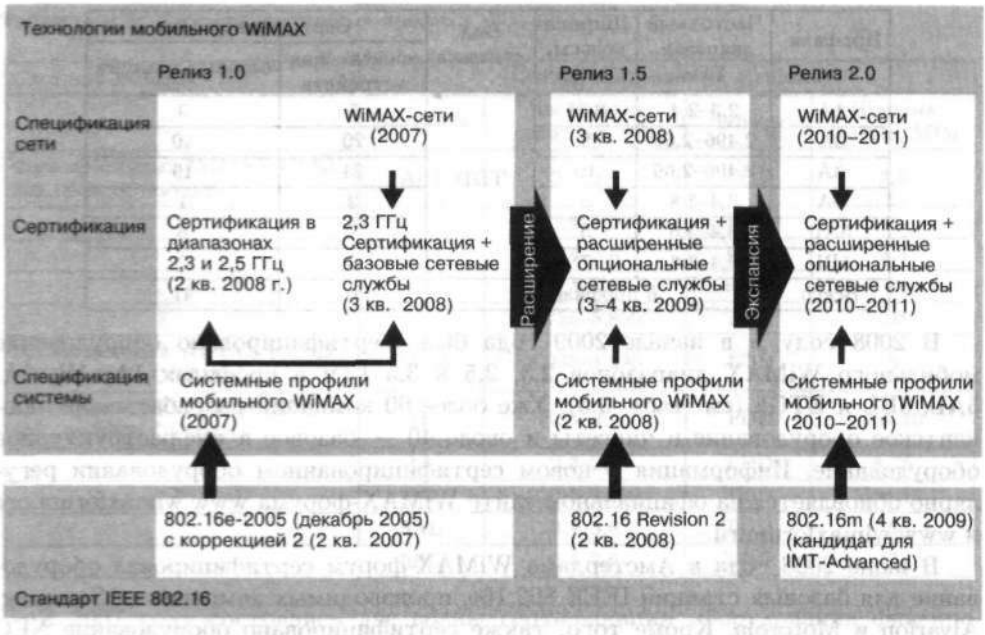


Рис. 9.2. Развитие программы WiMAX-сертификации

Появление релиза 2 ожидается в 2010–2011 годах. Он должен опираться на новый стандарт IEEE 802.16m, который будет одним из основных кандидатов на включение в перечень технологий IMT-Advanced.

К концу марта 2009 года WiMAX-форум зарегистрировал 94 модели сертифицированного WiMAX-оборудования 36 различных производителей (табл. 9.7) (www.wimaxforum.org/productshowcase). Хотя в основном эта таблица содержит оборудование для фиксированного доступа, доля мобильного WiMAX постоянно растет. Широко представлено как базовое, так и абонентское оборудование. Далеко не все оборудование в табл. 9.7 коммерчески доступно, многие сертифицируют только системный дизайн или «предкоммерческие» образцы.

Таблица 9.5. Сертификационные тесты на различных уровнях представления WiMAX-сети

Уровни модели OSI	WiMAX-спецификации	Сертификационные модули
Приложений	Сетевые спецификации WiMAX-форума	NCT, IIOT
Представления		
Сеансовый		
Транспортный		
Сетевой		
Уровень соединения — подуровень логического соединения (LLC)		
Уровень соединения — MAC-подуровень	Системные спецификации WiMAX-форума (основанные на IEEE 802.16)	PCT, IOT, MIOT
Физический		RCT, IOT, MIOT, RPT

Из наиболее значимых производителей оборудования для мобильного WiMAX отметим компании Alvarion (Breeze0MAX 4Motion), Alcatel-Lucent (серия 97xx), Cisco System (BWX 8305 и BWX 2305), Huawei (DBTS 3900 и WASN9970), Motorola (wi4 WiMAX), Samsung (mobile WiMAX Udicell), ZTE и др. Оборудование большинства из них сертифицировано WiMAX-форумом.

Таблица 9.6. Активные сертификационные профили WiMAX 2009 года

Название профиля	Частотный диапазон, ГГц	Ширина канала, МГц	Дуплексирование
Фиксированные			
ET01	3,4–3,6	3,5	TDD
ET02	3,4–3,6	3,5	FDD
Мобильные			
MP01 (M2300T-01)	2,3–2,4	8,75	TDD
MP02 (M2300T-02)	2,3–2,4	5 и 10	TDD
MP05 (M2500T-01)	2,496–2,69	5 и 10	TDD
MP09 (M3500T-02)	3,4–3,6	5	TDD
MP10 (M3500T-03)	3,4–3,6	7	TDD
MP12 (M3500T-05)	3,4–3,6	10	TDD

Таблица 9.7. Оборудование, сертифицированное WiMAX-форумом к апрелю 2009 года

№	Компания	Дата сертификации (год.месяц.число)	Профиль сертификации	Название продукта	Название модели	Тип (БС/АС)	Частотный диапазон, ГГц	Ширина полосы, МГц	Тип дуплексирования
1	2	3	4	5	6	7	8	9	10
76	SEQUANS Communications	2006.01.17	Air 1	SQN2010-RD	SQN2010-RD	БС	3,4–3,6	3,5	TDD
59	Redline Communications	2006.01.17	Air 1	RedMAX Base Station	AN-100U	БС	3,4–3,6	3,5	TDD
57	Redline Communications	2006.01.30	Air 1	RedMAX Subscriber Station	SU-OIA 29-00018-00 (Integrated Antenna) SU-ORF 29-00019-00 (External Antenna via Type N connector)	АС	3,4–3,6	3,5	TDD
71	SEQUANS Communications	2006.01.31	Air 1	SQN1010-RD	SQN1010-RD	АС	3,4–3,6	3,5	TDD
47	Nokia Siemens Networks	2006.02.17	Air 2	WayMAX@vantage	WayMAX@vantage IDU/3.5 GHz FDD ODU	БС	3,4–3,6	3,5	FDD
1	Airspan Networks	2006.03.17	Air 2	MarcoMax	MMAX-3.5	БС	3,4–3,6	3,5	FDD
20	Axxcelera	2006.03.17	Air 2	ExcelMAX BS	EMAXFDBTS-B3DC	БС	3,4–3,6	3,5	FDD
74	SEQUANS Communications	2006.03.17	Air 2	SQN1010-RD	SQN1010-RD	АС	3,4–3,6	3,5	FDD
78	Siemens	2006.03.17	Air 2	Gigaset SE461 WiMAX	WayMAX@vantage IDU/3.5 GHz FDD ODU	АС	3,4–3,6	3,5	FDD
5	Airspan Networks	2006.03.17	Air 2	EasyST	EST350F	АС	3,4–3,6	3,5	FDD
94	Wavesat Inc.	2006.03.17	Air 2	Wavesat miniMAX 3.5GHz (FDD)	Subscriber Platform ASIC Base	АС	3,4–3,6	3,5	FDD

Таблица 9.7 (продолжение)

1	2	3	4	5	6	7	8	9	10
2.0	Airspan Networks	2006.04.05	Air 2	ProST	PST350F	БС	3,4–3,6	3,5	FDD
21	Axxcelera Broadband Wireless	2006.04.20	Air 2	ExcelMAX FD CPE	EMAXFD-CPE3200	АС	3,4–3,6	3,5	FDD
70	Selex Communications	2006.05.23	Air 2	YSEMAX	WRY035-B	БС	3,4–3,6	3,5	FDD
69	Selex Communications	2006.05.31	Air 1	YSEMAX (subscriber station)	WRY035-C	АС	3,4–3,6	3,5	TDD
81	SR Telecom	2006.06.05	Air 2	SSU5000 Symmetry Subscriber Station	SSU5000	АС	3,4–3,6	3,5	FDD
14	Alvarion	2006.06.14	MP05	BreezeMAX®	BreezeMAX 802.16e 2.5 GHz	БС	2,496–2,69	5; 10	TDD
12	Alvarion	2006.06.15	Air 2	BMAX PRO-S CPE	BMAX PRO-S Outdoor CPE family	АС	3,4–3,6	3,5	FDD
13	Alvarion	2006.06.23	Air 2	BreezeMAX Si	BreezeMAX Si CPE family	АС	3,4–3,6	3,5	FDD
16	Aperto Networks	2006.07.06	Air 1	Packet MAX	PM100 и PM300	АС	3,4–3,6	3,5	TDD
10	Alvarion	2006.07.28	Air 2	Macro Modular Base Station	BreezeMAX	БС	3,4–3,6	3,5	FDD
11	Alvarion	2006.07.28	Air 2	Micro Base Station	BreezeMAX Micro Base Station	БС	3,4–3,6	3,5	FDD
84	Telsima	2006.08.31	Air 1	StarMAX	StarMAX 2140-3.5G	БС	3,4–3,6	3,5	TDD
3	Airspan Networks	2006.09.21	Air 2	MicroMAX-SoC BSR	BSR350LF-12V-1	АС	3,4–3,6	3,5	FDD
85	Telsima	2006.10.21	Air 1	StarMAX	StarMAX 4120-3.5G	БС	3,4–3,6	3,5	TDD
19	Aperto Networks	2006.12.17	Air 1	PacketMAX 5000	PM 5000	БС	3,4–3,6	3,5	TDD

Таблица 9.7 (продолжение)

1	2	3	4	5	6	7	8	9	10
88	Wavesat Inc.	2006.12.17	Air 1	miniWIMAX 3.5 GHz	Subscriber Platform ASIC Base	АС	3,4–3,6	3,5	TDD
80	SR Telecom	2007.06.08	Air 2	Symmetry Base Station	CBS5000	БС	3,4–3,6	3,5	FDD
25	E.T. Industries	2007.07.05	Air 2		Apollo	БС	3,4–3,6	3,5	FDD
26	E.T. Industries	2007.07.05	Air 2	Apollo Subscriber Station	Apollo-SU & Apollo Max -SU	БС	3,4–3,6	3,5	FDD
18	Aperto Networks	2007.08.10	Air 1	PacketMAX 3000	PM3000	БС	3,4–3,6	3,5	TDD
82	SR Telecom	2007.11.14	Air 1	SymmetryMX	CBS5000	БС	3,4–3,6	3,5	TDD
83	SR Telecom	2007.11.14	Air 1	SSU5000	CBS5000	АС	3,4–3,6	3,5	TDD
56	Redline Communications	2008.02.22	Air 1	RedMAX	AN-100U	АС	3,4–3,6	3,5; 7	TDD
58	Redline Communications	2008.02.22	Air 1	RedMAX Outdoor Subscriber Unit (SU-O)	U-OIA 29-00018-00 (Integrated Antenna) SU-ORF 29-00019-00 (External Antenna via Type N connector)	АС	3,4–3,6	3,5; 7	TDD
75	SEQUANS Communications	2008.03.03	MP01	SEQUANS Communications	SQN1110-RD	АС	2,3–2,4	8,75	TDD
51	POSDATA	2008.04.03	MP01	FLYVO	P-RAS 1002	БС	2,3–2,4	8,75	TDD
52	POSDATA	2008.04.03	MP01	FLYVO	U-100	АС	2,3–2,4	8,75	TDD
61	Runcom Technologies Ltd.	2008.04.03	MP01	Runcom Mobile Base Station	RNU2000N	БС	2,3–2,4	8,75	TDD
62	Samsung	2008.04.03	MP01	SAMSUNG Mobile WiMAX	SWT-P230	АС	2,3–2,4	8,75	TDD
77	SEQUANS Communications	2008.04.03	MP01	SEQUANS Communications	SQN2130-RD	БС	2,3–2,4	8,75	TDD

Таблица 9.7 (продолжение)

1	2	3	4	5	6	7	8	9	10
63	Samsung	2008.04.03	MP01	SAMSUNG Mobile WiMAX	SWT-P230	AC			
68	Samsung	2008.04.3	MP01	Samsung Mobile WiMAX Base Station	SPI-2110	БС	2,3–2,4	8,75	TDD
41	Motorola	2008.06.14	MP05	Motorola Inc.	WAP 25400	БС	2,496–2,69	5; 10	TDD
65	Samsung	2008.06.14	MP05	SAMSUNG ELECTRONICS Co., Ltd.	WiMAX Wave2 Base Station SPI-2211	БС	2,496–2,69	5; 10	TDD
66	Samsung	2008.06.14	MP05	SAMSUNG Mobile WiMAX	SWC-E100	AC	2,496–2,69	5; 10	TDD
72	SEQUANS Communications	2008.06.14	MP05	SEQUANS Communications	SQN2130-RD	БС	2,496–2,69	5; 10	TDD
93	ZyXEL Communications Inc.	2008.06.14	MP05	ZyXEL	MAX-206M2	AC	2,496–2,69	5; 10	TDD
4	Airspan Networks	2008.06.17	MP05	MiMAX	MiMAX-USB-Q1-1	AC	2,496–2,69	5; 10	TDD
37	Intel Corporation	2008.07.14	MP05	Intel WiMAX/WiFi Link 5350	Intel WiMAX/WiFi Link 5350	AC	2,496–2,69	5; 10	TDD
46	NEC	2008.07.30	MP05	PasoWings BS	NWA-027932-001 (ODU) & NWA-024297-001 (IDU)	БС	2,496–2,69	5; 10	TDD
44	NEC	2008.08.04	MP05	WiMAX PC Card	TRP-2GW-2A	AC	2,496–2,69	5; 10	TDD

Таблица 9.7 (продолжение)

1	2	3	4	5	6	7	8	9	10
87	Telsima Corparetion	2008.08.13	MP05	StarMax 6400	StarMax 6022/8200-25	БС	2,496–2,69	5; 10	TDD
91	ZTE Corporation	2008.08.22	MP05	ZTE	TU25 USB Modem	АС	2,496–2,69	5; 10	TDD
28	GCT Semiconductor, Inc	2008.08.25	MP05	Mobile WiMAX Single Chip SoC	GDM7205K	АС	2,496–2,69	5; 10	TDD
67	Samsung	2008.09.03	MP05	SAMSUNG Mobile WiMAX	SWC-U200; SWC-U201; SWC-U202; SWC-U203; SWC-U204; SWC-U205; SWC-U206; SWC-U207; SWC-U208; SWC-U209	АС	2,496–2,69	5; 10	TDD
8	Alcatel-Lucent	2008.09.13	MP05	Alcatel-Lucent	9710 Compact Base Station for WiMAX	БС	2,496–2,69	5; 10	TDD
31	Huawei Technologies	2008.09.15	MP05	Huawei EchoLife	BM625	АС	2,496–2,69	5; 10	TDD
32	Huawei Technologies	2008.09.15	MP05	Huawei EchoLife	BM325	АС	2,496–2,69	5; 10	TDD
55	Redline Communications	2008.09.17	MP05	RedMAX 4C RPM	RPM2500M	АС	2,496-2,69	5; 10	TDD
33	Huawei Technologies	2008.09.25	MP05	Huawei WiMAX Base Station	DBS3900 WiMAX	БС	2,496–2,69	5; 10	TDD
53	POSDATA Co., Ltd.	2008.09.25	MP05	POSDATA Mobile WiMAX Wave 2 Base Station	FLYVO RAS6000	БС	2,496–2,69	10	TDD

Таблица 9.7 (продолжение)

1	2	3	4	5	6	7	8	9	10
23	Cisco Systems	2008.10.08	MP05	Cisco	Cisco BWX 8305 Basestation + TTA 2496-2620 MHz	БС	2,496–2,69	5; 10	TDD
24	Cisco Systems	2008.10.15	MP05	Cisco	Cisco BWX 2305 Basestation + TTA 2496-2620 MHz	БС	2,496–2,69	5; 10	TDD
9	Alcatel-Lucent	2008.10.20	MP05	CARC	CARC 500 Series		2,496–2,69	5; 10	TDD
39	MODACOM Co., Ltd.	2008.10.21	MP05	WiMAX USB Modem	MW-U2500 ; MW-U2510 ; MW-U2520 ; MW-U2530 ; MW-U2540 ; MW-U2550	АС	2,496–2,69	5; 10	TDD
40	Motorola	2008.10.30	MP05	WTM1000 Reference Design	WTM1000 Reference Design	АС	2,496–2,69	5; 10	TDD
7	Alcatel-Lucent	2008.11.13	MP05	Alcatel-Lucent	9715 Light Base Station for WiMAX (LWBSA25)	БС	2,496–2,69	5; 10	TDD
54	Proxim Wireless Corporation	2008.11.24	Air 1	Tsunami MP16 3500	3500-B00	БС	3,4–3,6	3,5	TDD
22	Beceem	2008.11.24	MP05	Beceem	USB200	АС	2,496–2,69	5; 10	TDD
73	SEQUANS Communica-tions	2008.11.24	MP05	SEQUANS Communications	SQN1130-RD	АС	2,496–2,69	5; 10	TDD
86	Telsima Corparetion	2008.11.24	MP05	StarMAX 3160	StarMAX 3160-25	АС	2,496–2,69	5; 10	TDD

Таблица 9.7 (продолжение)

1	2	3	4	5	6	7	8	9	10
92	ZTE Corporation	2008.11.24	MP05	ZTE Wimax Base Station System (ZXMBW-A250)	ZXMBW-A250	БС	2,496–2,69	5; 10	TDD
60	Runcom Technologies Ltd.	2008.11.24	MP01	Tornado CPE	01-0000000038	АС			
79	SOMA Networks, Inc.	2008.12.05	MP05	FlexMax Mobile WiMAX System	FlexMAX ABS-1000 Base Station	БС	2,496–2,69	5; 10	TDD
48	Nortel Networks Corp.	2008.12.09	MP05	Nortel WiMAX BTS	WiMAX BTS 5020	БС	2,496–2,69	5; 10	TDD
34	Intel Corporation	2008.12.1	MP05	Intel WiMAX/WiFi Link 5350	Echo Peak Prime	АС	2,496–2,69	5; 10	TDD
36	Intel Corporation	2008.12.11	MP05	Echo Peak Volume	Intel WiMAX — WiFi Link 5150 MC	АС	2,496–2,69	5; 10	TDD
49	OKI Networks Co., Ltd.	2008.12.17	MP05	UD01OK	BR3001	АС	2,496–2,69	5; 10	TDD
17	Aperto Networks	2008.12.2	3.5T2	PacketMAX	PM120 / PM320	АС	3,4–3,6	3,5	TDD
35	Intel Corporation	2008.12.4	MP05	Echo Peak Volume	Intel WiMAX — WiFi Link 5150 HMC	АС	2,496–2,69	5; 10	TDD
29	Gigaset Communications GmbH	2008.12.5	MP05	Gigaset	Gigaset SE680/SE681/SX682 WiMAX	АС	2,496–2,69	5; 10	TDD

Таблица 9.7 (окончание)

1	2	3	4	5	6	7	8	9	10
27	Fujitsu Microelectronics Limited	2008.12.9	MP05	Fujitsu Mobile WiMAX SoC	MB86K21-UD1	АС	2,496–2,69	5; 10	TDD
90	ZTE Corporation	2009.01.15	MP09, MP 10 & MP12	ZXMBW	B9100(BBU), R9100(RRU)	БС	3,4–3,6	5	TDD
64	SAMSUNG ELECTRONICS Co., Ltd.	2009.01.15	MP12	SAMSUNG Mobile WiMAX RAS SPI-2213	SPI-2213	БС	3,47–3,49 и 3,568–3,578	10	TDD
38	MODACOM Co., Ltd.	2009.01.21	MP05	Express Card WiMAX Modem	MW-C2500E ; MW-C2510E ; MW-C2520E ; MW-C2530E ; MW-C2540E ; MW-C2550E	АС	2,496–2,69	5; 10	TDD
43	NEC	2009.01.21	MP05	Aterm WM3200U	PA-WM3200U; UD01NA	АС	2,496–2,69	5; 10	TDD
45	NEC	2009.01.21	MP05	Aterm WM3200C	PA-WM3200C; UD02NA	АС	2,496–2,69	5; 10	TDD
6	Alcatel-Lucent	2009.02.17	MP09, MP10, MP12	Alcatel-Lucent 9710 Compact Base Station for WiMAX (9710 CWBS)	WBSA34, WBSD34	БС	3,4–3,6	5; 7; 10	TDD
15	Alvarion LTD.	2009.02.18	MP12	BreezeMAX	BreezeMAX 802.16e 3.5 GHz	БС	3,4–3,6	10	TDD
50	PointRed Telecom Pvt Ltd	2009.03.07	M2500T-01	Wimax BS	PointMAX10000	БС	2,496–2,69	5; 10	TDD
30	Huawei Technologies	2009.03.09	MP09, MP10, MP12	DBS3900 WiMAX	Huawei WiMAX Base Station	БС	3,4–3,6	5; 7; 10	TDD
42	NEC	2009.03.12	M3500T-03 & M3500T-05	PasoWings	NWA-027932-001 (ODU) / NWA-024297-001 (IDU)	БС	3,4–3,6	7; 10	TDD
89	WiNetworks	2009.03.12	M2500T-01	pBST	WiN7225-2	БС	2,496–2,69	5; 10	TDD

9.3. Оборудование WiMAX на примере платформы BreezeMAX 4Motion

Подробнее рассмотрим построение оборудования мобильного WiMAX на примере базовой и абонентской станций системы BreezeMAX 4Motion израильской компании Alvarion. Система 4Motion — это полнофункциональное решение мобильного WiMAX операторского класса, с открытой архитектурой, позволяющее сопрягать оборудование различных производителей в одной сети. В качестве сетевого профиля, описывающего взаимодействие элементов сети, выбран ASN-профиль C как полностью открытый.

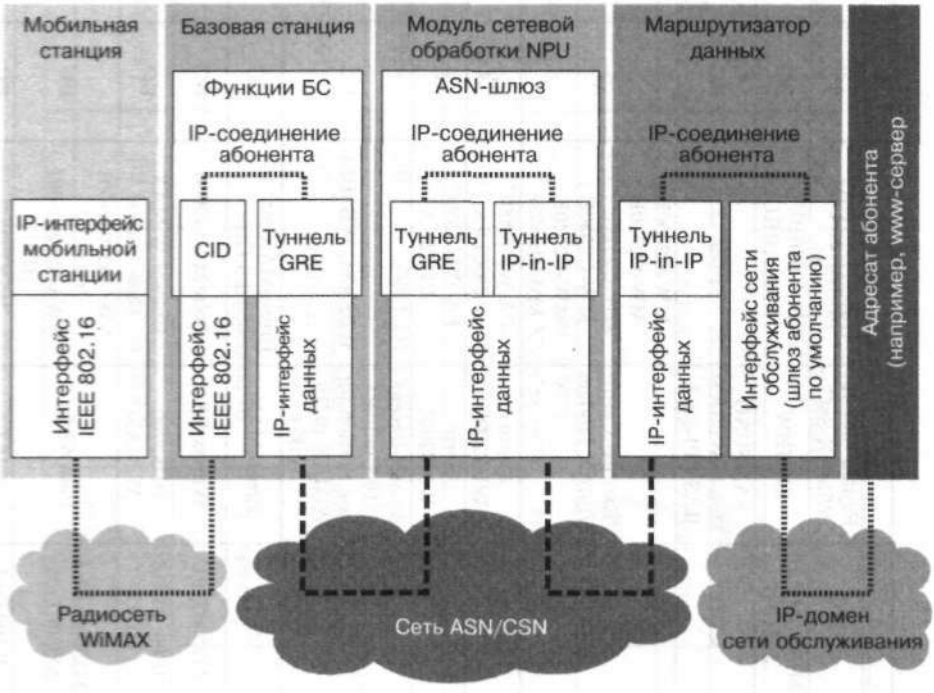


Рис. 9.3. Схема передачи данных в сети WiMAX на основе оборудования 4Motion компании Alvarion

Напомним, архитектура сети WiMAX включает три основных элемента — абонентские (мобильные) станции (MC), совокупность сетей доступа (сервисная сеть доступа, ASN) и совокупность сетей подключения (CSN). Сеть доступа ASN включает базовые станции и шлюзы (ASN-шлюзы). Сеть подключения CSN — это наземная IP-сеть оператора WiMAX, именно в этой сети размещены AAA-серверы.

Платформа BreezeMAX 4Motion включает четыре основные составляющие — абонентские станции, базовые станции, шлюзы сети доступа (ASN-шлюзы) и серверы системы управления авторизацией, аутентификацией и доступом (AAA-серверы). Последние представляют собой достаточно стандартные сетевые серверы (производители, которые не имеют своих AAA-серверов, обычно используют оборудование компаний Bridgewater и Cisco), вся их функциональность реализуется программно, поэтому мы не будем останавливаться на них

сколь-нибудь подробно. Остальные три элемента обеспечивают прохождение данных пользователя между оконечными устройствами (мобильными станциями, узлами IP-сетей и т. п.) (рис. 9.3).

9.3.1. ASN-шлюзы

Система BreezeMAX 4Motion может быть реализована с двумя типами ASN-шлюзов: распределенным и централизованным. В случае распределенной модели функции ASN-шлюзов реализуют устройства в составе БС (модуль устройства сетевой обработки NPU) (рис. 9.4, *а*). Такое решение предназначено в первую очередь для сетей малого масштаба: не более 3 тыс. абонентов и до 200 Мбит/с на один ASN-шлюз. Число абонентов в сети можно повысить за счет увеличения числа распределенных ASN GW в WiMAX-сети. Это позволяет постепенно увеличивать сеть (и расходы на ее развертывание). Централизованный ASN-шлюз предназначен для сетей большого масштаба с сотнями базовых станций и десятками тысяч абонентов внутри сети (рис. 9.4, *б*). Данное решение реализуется, в частности, на маршрутизаторах операторского класса компании Cisco серии 7600 (рис. 9.5).

Применение централизованного ASN-шлюза позволяет довольно просто масштабировать сеть. Оборудование Cisco строится по модульной blade-архитектуре SAMI (Service and Application Module for IP). Один SAMI-модуль позволяет подключить до 100 тыс. абонентов с суммарной пропускной способностью до 5 Гбит/с. Всего blade-модулей SAMI в корзине может быть до шести, а это уже 600 тыс. абонентов и до 30 Гбит/с агрегированного трафика от абонентов. Ну и, конечно, централизованный ASN-шлюз поддерживает набор приложений безопасности, VPN и QoS.

9.3.2. Базовая станция

Базовая станция обеспечивает все необходимые функции для организации соединений по радиоканалу с абонентскими устройствами станции и по каналу GB Ethernet — для подключения к магистральному каналу сети провайдера. Она полностью соответствует всем требованиям стандарта IEEE 802.16 и сертификационным профилям WiMAX. Станция поддерживает режимы масштабируемой OFDMA, т. е. может работать с каналами шириной 20, 10 и 5 МГц (2048, 1024 и 512 формальных поднесущих, соответственно). Поддерживается режим использования неполного набора поднесущих (PUSC-режим) в нисходящем канале.

Базовая станция BreezeMAX обладает модульной архитектурой, что позволяет легко масштабировать систему и воплощать требуемую конфигурацию (рис. 9.6). Оборудование БС построено на основе шасси Compact PCI высотой 8U (рис. 9.7), предназначенного для установки инсталляции в 19- или 22-дюймовые стойки. В шасси монтируется девять двойных (6U) и шесть одиночных (3U) слотов. Возможна горячая замена любого модуля, что позволяет заменять отдельные модули, не прерывая работу всей БС. Каждое функциональное устройство может быть резервировано по схеме $N+1$, т. е. только одно устройство данного типа может быть резервным.

Шесть одиночных слотов предназначены для установки двух модулей интерфейса питания (PIU, Power Interface Units) (основной и резервный) и до четырех модулей источников питания (PSU, Power Supply Unit). Источник питания — это

3U-модуль, выдающий 48 В постоянного тока. Корзина БС может содержать до четырех PSU-модулей. В зависимости от числа устройств доступа необходимо два (до 4 AU) или три модуля источника питания. Еще один модуль PSU используется как резервный.

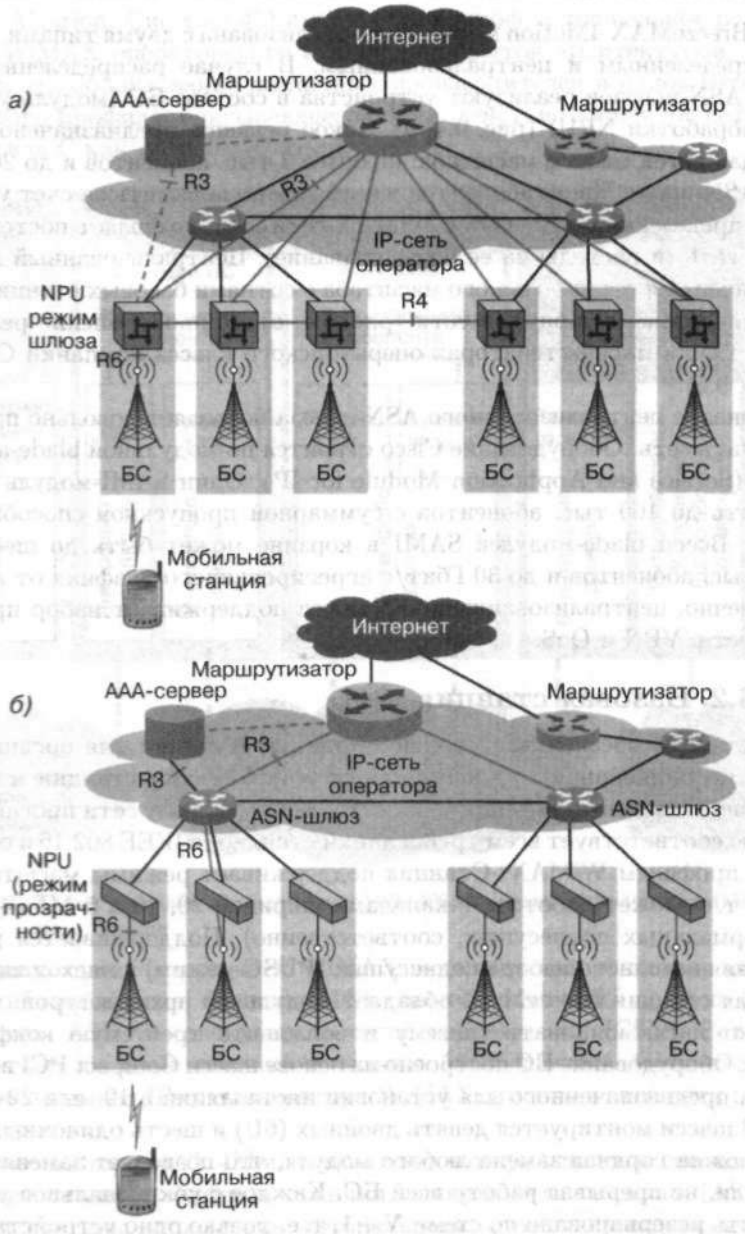
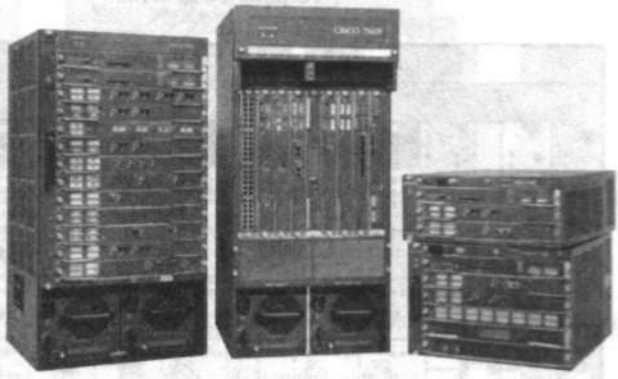


Рис. 9.4. WiMAX-сеть с распределенными *(а)* и централизованными *(б)* ASN-шлюзами

Рис. 9.5. Маршрутизаторы серии Cisco 7600



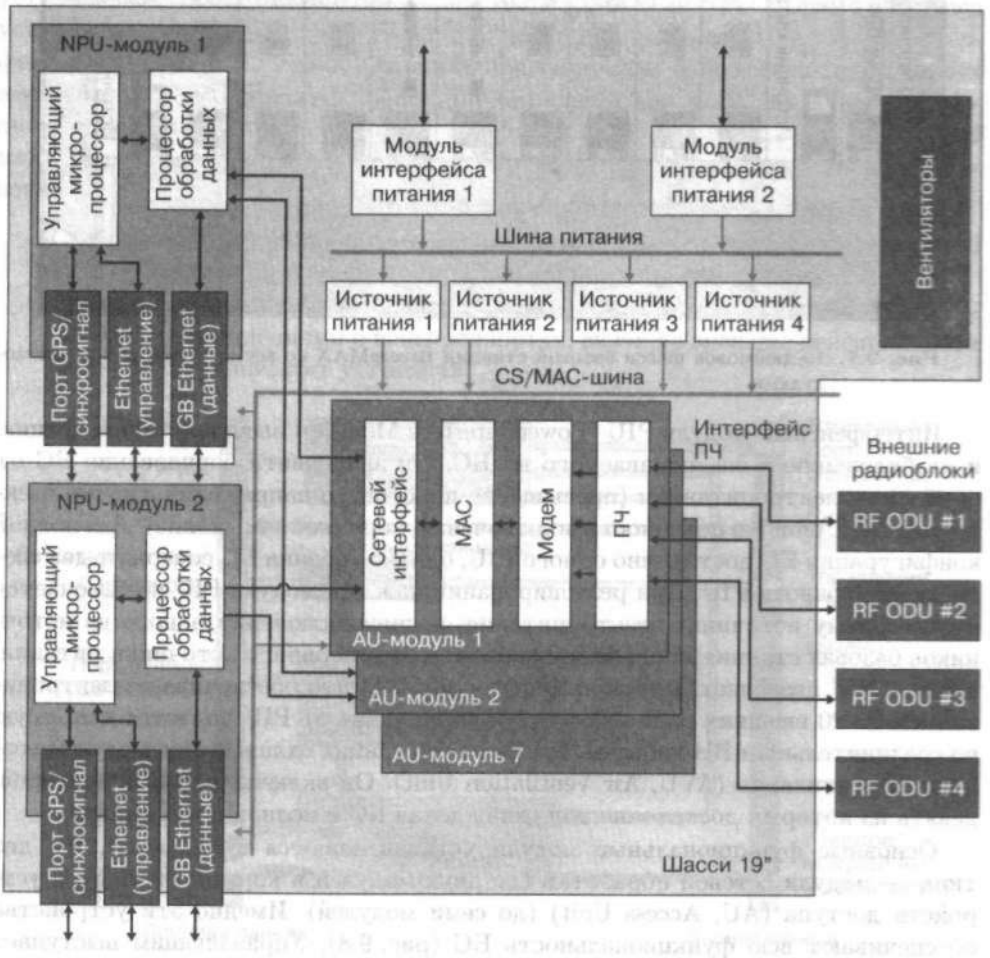


Рис. 9.6. Общая архитектура базовой станции BreezeMAX

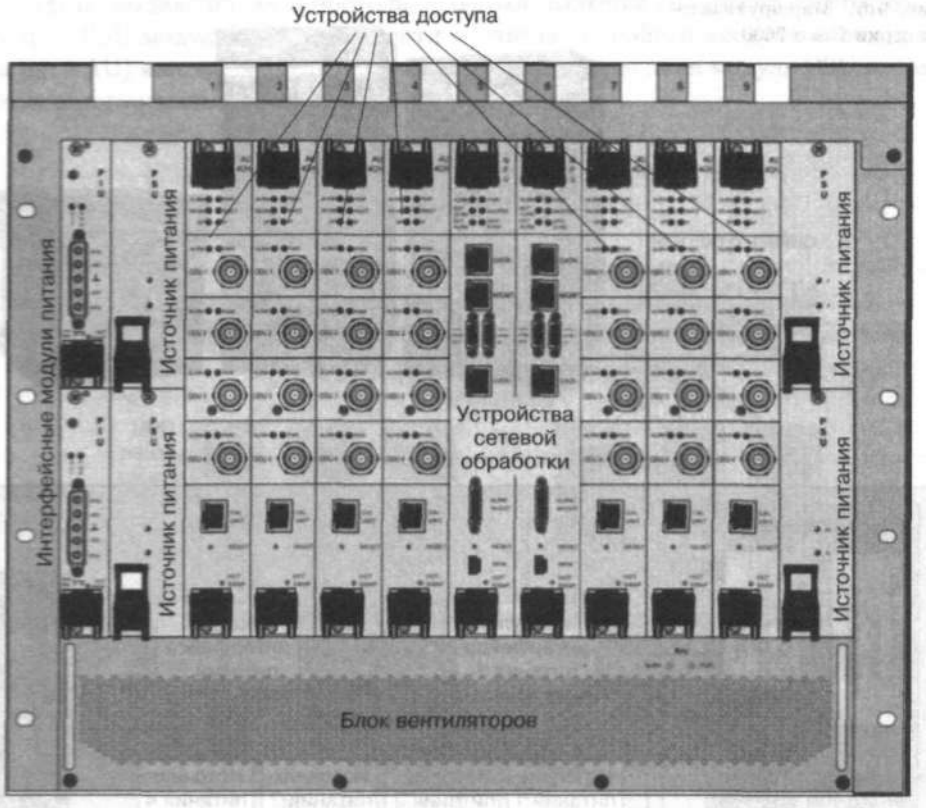


Рис. 9.7. 19-дюймовое шасси базовой станции BreezeMAX со всеми установленными модулями

Интерфейсный модуль PIU (Power Interface Module) выполняет фильтрацию и стабилизацию тока, подаваемого на БС. Он защищает оборудование БС от проблем с электропитанием (превышение порогового напряжения, скачки электропитания, ошибки полярности подключения, короткое замыкание). Для полной конфигурации БС достаточно одного PIU, однако корзина БС содержит два слота для резервного PIU. При резервировании каждый модуль PIU подключается к отдельному источнику электропитания, и при отключении одного из источников базовая станция продолжает работать от резервного источника питания. Модуль PIU позволяет подать на БС ток до 58 А, что обеспечивает электропитанием до 20 внешних радиоблоков. К ним питание от PIU подается напрямую по соединительным ВЧ-кабелям. Вся базовая станция охлаждается модулем воздушной вентиляции (AVU, Air Ventilation Unit). Он включает 10 вентиляторов, девять из которых достаточно для охлаждения БС в полной комплектации.

Основные функциональные модули устанавливаются в 6U-слоты. Их два типа — модули сетевой обработки (до двух модулей в корзине) и модули устройств доступа (AU, Access Unit) (до семи модулей). Именно эти устройства обеспечивают всю функциональность БС (рис. 9.8). Управляющим выступает устройство сетевой обработки, которое связывает наземную сеть (Gigabit/Fast Ethernet) и радиосеть (через устройства доступа).

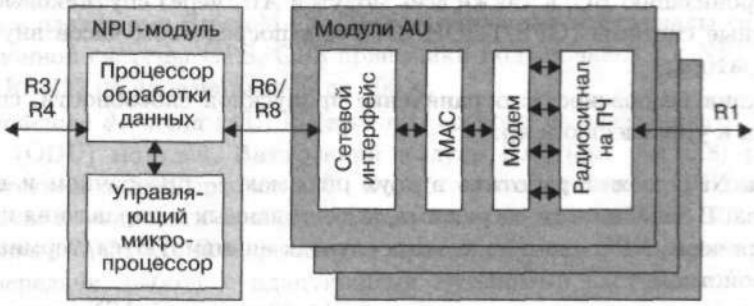


Рис. 9.8. Взаимодействие устройств сетевой обработки и устройств доступа с интерфейсами базовой модели сети WiMAX

Модуль сетевой обработки (NPU, Network Processing Unit) (рис. 9.9) управляет всеми компонентами оборудования базовой станции и абонентскими устройствами, обслуживаемыми данной БС. NPU объединяет трафик от модулей устройств доступа и передает его в IP-магистраль через выделенный интерфейс Gigabit/Fast Ethernet. Отдельно формируются Ethernet-потоки данных и управления. NPU реализует общее управление модулями БС, включая управление и диагностику AU, мониторинг состояния источников питания, управление модулем вентиляторов и поддержку избыточности для резервирования. Он обеспечивает:

- локальное и удаленное управление через интерфейс командной строки CLI (telnet, SSH) и протокол SNMP, включая загрузку программ;
- управление производительностью и выявление неисправностей, управление предупреждениями о неисправностях, включая внешние неисправности и активацию внешних устройств;

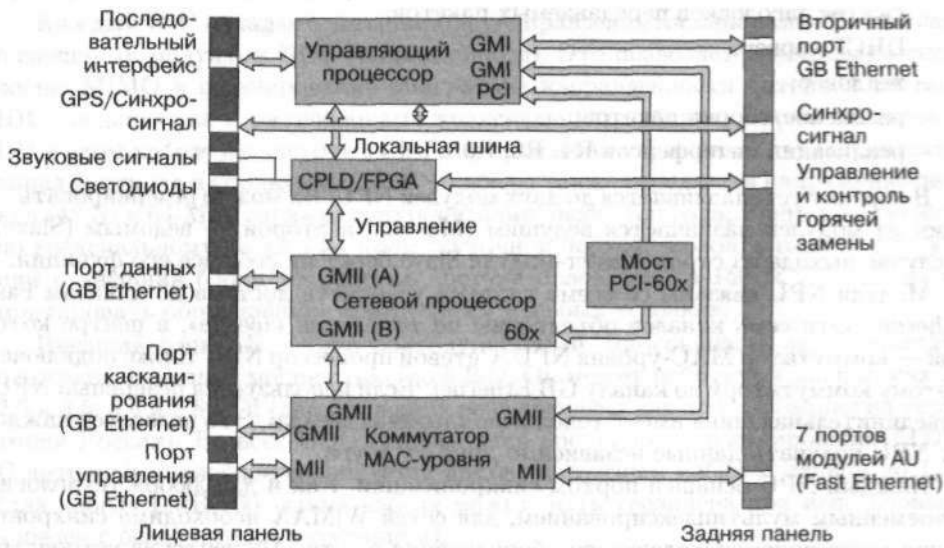


Рис. 9.9. Модуль сетевой обработки

- синхронизацию БС, а также всех модулей AU, через спутниковые навигационные системы (GPS/ГЛОНАСС) или посредством часов внутреннего генератора;
- функции безопасности: ограничение пропускной способности, список доступа к управлению и др.

Модуль NPU может работать в двух режимах — прозрачном и в режиме ASN-шлюза. В зависимости от режима, потоки данных и управления прозрачно передаются через NPU или в отдельных случаях инициируются/терминируются этим устройством.

В *прозрачном режиме* используется централизованный ASN-шлюз, на который передается весь трафик через Gigabit/Fast Ethernet (абонентский и управляющий) (через интерфейс R6). Помимо этого, в прозрачном режиме NPU обеспечивает каскадное подключение БС, коммутацию трафика на MAC-уровне, VLAN-инкапсуляцию для внутреннего и внешнего трафика, маркировку пакетов согласно QoS.

В *режиме ASN-шлюза* NPU, помимо перечисленных задач, выполняются функции шлюза сети доступа, а именно:

- аутентификация (EAP, Extensible Authentication Protocol — расширяемый протокол аутентификации);
- клиент AAA RADIUS-сервера;
- AAA-клиент биллинга;
- хранение профилей политики подключения абонентских станций;
- авторизация сервисных потоков QoS;
- GRE инкапсуляция/декапсуляция (Generic Routing Encapsulation — общая инкапсуляция маршрутов);
- IP-in-IP инкапсуляция/декапсуляция;
- сжатие заголовков передаваемых пакетов;
- DHCP-сервер;
- хендовер;
- разбиение/сборка пакетов;
- реализация интерфейсов R4, R6, R3.

В корзину устанавливается до двух модулей NPU. Их можно резервировать — один из модулей назначается ведущим (Master), а второй — ведомым (Slave). В случае выхода из строя Master-модуля Slave берет на себя все его функции.

Модули NPU связаны со всеми слотами устройств доступа по каналам Fast Ethernet. Эти семь каналов объединены по топологии «звезда», в центре которой — коммутатор MAC-уровня NPU. Сетевой процессор NPU также подключен к этому коммутатору по каналу GB Ethernet. Если используется резервный NPU, объединительная шина имеет топологию «двойная звезда». Это позволяет каждому NPU получать данные независимо друг от друга.

Каждый NPU оснащен портом синхронизации. Как и для любой технологии с временным мультиплексированием, для сетей WiMAX необходима синхронизация всего приемопередающего оборудования в сети. Абонентские устройства синхронизируются с БС посредством специальных синхронизирующих последовательностей в преамбулах кадров. Базовые станции в пределах сети могут

синхронизироваться только от внешнего источника синхронизации. В качестве такового в платформе BreezeMAX 4Motion используются сигналы спутниковой навигационной системы GPS. GPS-приемники подключаются к NPU через интерфейс RS-422 по кабелю длиной до 100 м.

Устройство доступа (AU, Access Unit) состоит из внутреннего (IDU) и наружного (ODU) модулей. Внутренний модуль IDU (см. рис. 9.8) реализован в соответствии с ASN-профилем C. Он обеспечивает все необходимые функции работы в радиосети в соответствии со стандартом IEEE 802.16e и спецификациями WiMAX. IDU реализует пространственно-временное разнесение приема-передачи, работу с адаптивными антенными системами, гибкую настройку ширины канала (до 20 МГц), управление соединением (подключение к сети, основное согласование совместимости, аутентификацию и регистрацию, управление), планирование (вычисление возможной выделяемой полосы пропускания для всех типов доставляемых данных), формирование фреймов, управление хендовером, контроль и управление мощностью передачи на абонентских и базовой станциях БС.

Это устройство выполняет такие функции безопасности и контроля доступа, как аутентификация и шифрование трафика, пересылка аутентификационных запросов, получение ключей безопасности. Устройство доступа поддерживает интерфейсы R1, R6 и R8. Одно AU может работать с 512 абонентскими станциями.

В одной корзине БС может быть до шести устройств доступа — соответственно, до шести секторов БС. Седьмой слот используется для резервного IDU.

Внешним входным сигналом для IDU является поток MAC-уровня от NPU по каналу Fast Ethernet. Устройство доступа преобразует его в поток физического уровня в соответствии со стандартом IEEE 802.16 и формирует выходной сигнал на промежуточной частоте (ПЧ) 240 МГц. При приеме данных из радиосети в IDU по тому же ВЧ-каналу поступает сигнал с ПЧ 140 МГц, который IDU преобразует в поток MAC-уровня.

Каждый IDU обладает четырьмя двунаправленными каналами ПЧ для связи с внешними модулями ODU (радиоблоками). Это позволяет реализовать технологию MIMO и формирование диаграммы направленности антенной системы. IDU соединяются с радиоблоками коаксиальным кабелем. Кроме сигналов на ПЧ в этом кабеле на частоте 14 МГц организован служебный двунаправленный канал контроля и управления. По ВЧ-кабелю также передается напряжение питания (48 В) и 64-МГц сигналы синхронизации часов. Благодаря передаче сигналов по коаксиальному кабелю на промежуточной частоте можно разнести внутренний и внешний блоки устройства доступа более чем на 150 м. Это позволяет монтировать оборудование практически в любых условиях.

Внешний блок AU — это полнодуплексный, многоканальный радиомодуль с высокой выходной мощностью (до 5 Вт). Он может работать в диапазонах 2,3; 2,5 и 3,5; 3,6 и 3,8 ГГц (два последних не разрешены для использования на территории России). Во всех ODU используется временное дуплексирование (TDD). С антенной радиоблок соединяется высокочастотным кабелем длиной до 1,5 м (для кабеля LMR-400, длина кабеля может быть увеличена при использовании кабелей с более низким затуханием).

ODU выпускаются в виде моноблоков в трех исполнениях — один приемник / один передатчик (1 × 1), два приемника / один передатчик (2 × 1) и четыре

Рис. 9.10. Наружный радиомодуль AU: *а)* 1 × 1, *б)* 4 × 2

приемника / два передатчика (4 × 2) (рис. 9.10). Возможны различные варианты подключения ODU к IDU. Так, к одному IDU может подключаться до четырех ODU 1 × 1 (реализуется схема 4 передающих × 4 приемных канала) или один ODU 2 × 4 (2 передающих × 4 приемных канала) (рис. 9.11).

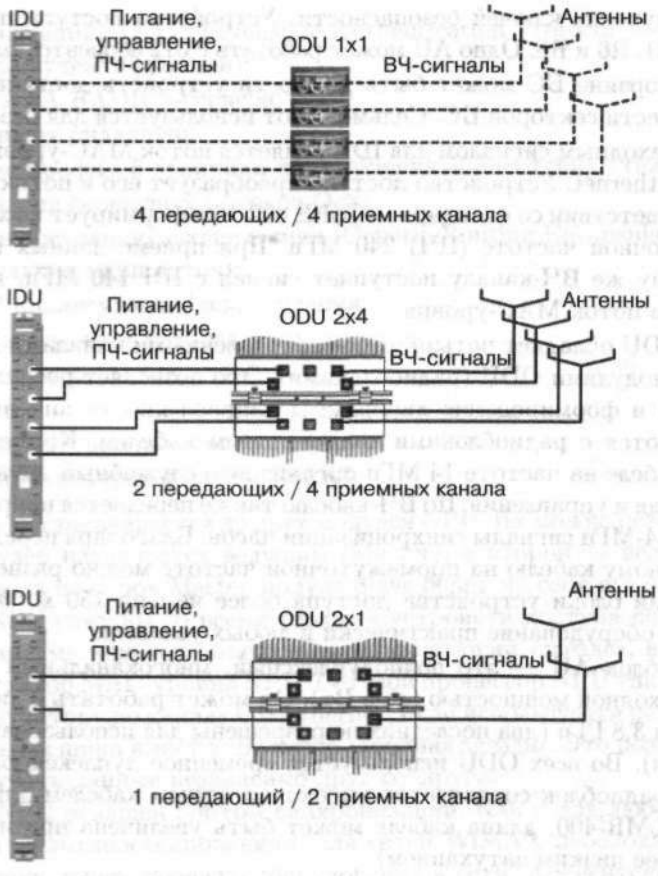


Рис. 9.11. Варианты подключе- ния внешних радиоблоков устройств доступа

9.3.3. Антенные системы

Для реализации технологии MIMO (пространственно-временное кодирование по алгоритму Аламоути или передача независимого потока по каждому из антенных каналов) или адаптивного диаграммообразования в платформе 4Motion предполагается использовать несколько конфигураций антенн. Так, для формирования независимых потоков в каждом антенном канале предлагается три варианта: разнесенные антенны с различной поляризацией (взаимная поляризация 90°, наклон к горизонту — ±45°). Антенны должны быть разнесены на расстояние не менее 10 длин волн (λ). Как правило, для этого используются две двухэлементные антенны с взаимной поляризацией элементов 90°, но подключаются только по одному элементу в антенне. Использование кроссполяризационных антенн с поляризацией ±45° относительно линии горизонта объясняется тем, что при переотражении сигналов изменяется их поляризация.

Второй вариант подразумевает применение X-образной антенны с двумя элементами со взаимно-ортогональной поляризацией. Такая поляризация обеспечивает разнесение каналов не менее чем на 20 дБ. Оба этих варианта позволяют организовать передачу по двум независимым каналам.

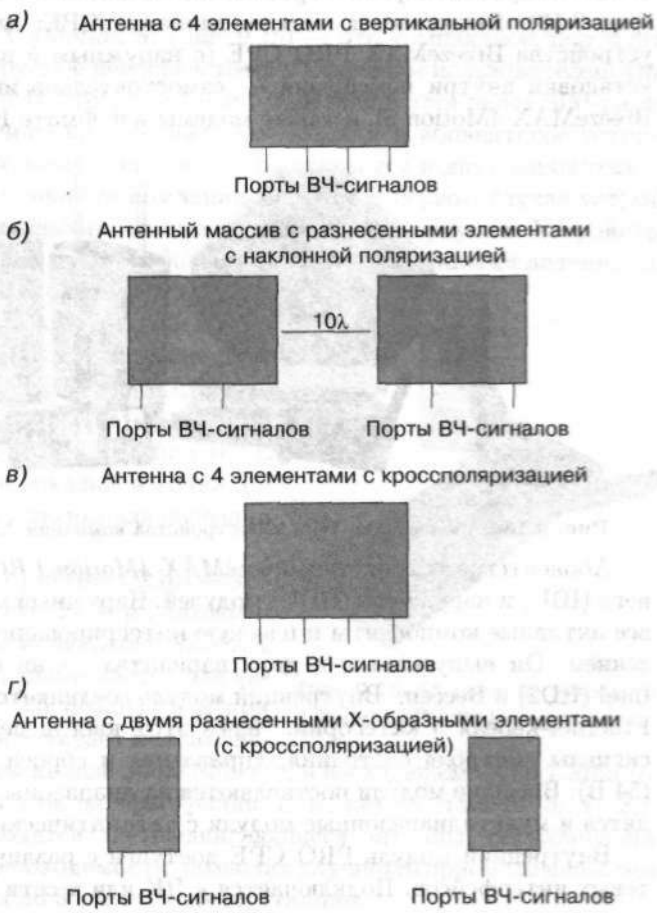


Рис. 9.12. Варианты антенных массивов для реализации адаптивных антенных систем

Для реализации передачи по четырем каналам рекомендована четырехэлементная антенная система — две X-образные антенны (как в предыдущем варианте), разнесенные друг от друга не менее чем на 10λ.

Во всех этих вариантах подразумевается, что каждый антенный элемент формирует луч шириной 65° в азимутальной плоскости и 7° — в вертикальной (по уровню 3 дБ), уровень боковых лучей до −30 дБ в азимутальной плоскости и −17 дБ — в вертикальной.

Для задач адаптивного формирования диаграммы направленности используют антенные массивы из четырех близко расположенных элементов с вертикальной поляризацией (рис. 9.12, *а*). Однако для смешанных режимов (формирование ДН и MIMO) этот вариант не оптимален. Если требуются совмещенные режимы, предпочтительны варианты на рис. 9.12, *б–г*. Причем варианты *б* и *в* позволяют, помимо формирования ДН, организовать два MIMO-канала, вариант *г* — четыре канала.

9.3.4. Абонентское оборудование

С платформой BreezeMAX 4Motion предлагается несколько вариантов оборудования конечного пользователя (CPE), которые позволяют операторам эффективно обслуживать разнообразных пользователей в деловых и жилых секторах (рис. 9.13). Выпускается четыре варианта CPE: для наружного монтажа — устройства BreezeMAX PRO CPE (с наружным и внутренним модулями), для установки внутри помещений — самостоятельно инсталлируемые устройства BreezeMAX 4Motion Si, а также модемы в формате PC Card и USB Dangle.

Рис. 9.13. Абонентские WiMAX-устройства компании Alvarion

Абонентское устройство *BreezeMAX 4Motion PRO CPE* состоит из внутреннего (IDU) и наружного (ODU) модулей. Наружный модуль (рис. 9.14) содержит все активные компоненты и плоскую интегрированную антенну с высоким усилением. Он выпускается в двух вариантах — на основе чипсетов компаний Intel (RD2) и Beceem. Внутренний модуль соединяется с наружным посредством Ethernet-кабеля 5 категории. Через этот кабель передаются Ethernet-данные, сигналы контроля состояния, управления и сброса от IDU, а также питание (54 В). Внешние модули поставляются на диапазоны 2,3; 2,5 и 3,5 ГГц. Производятся и мультидиапазонные модули с автоматическим определением частоты.

Внутренний модуль PRO CPE доступен с различными конфигурациями сетевых интерфейсов. Подключается к ПК или к сети передачи данных абонента

через стандартный IEEE 802.3 Ethernet 10/100-BaseT (RJ 45) интерфейс. Также IDU может включать в себя два (опционально) голосовых порта (RJ-11) для передачи VoIP. Предусмотрен и Wi-Fi модуль (IEEE 802.11b/g) для организации локальной точки доступа.

Рис. 9.14. Наружный модуль PRO CPE

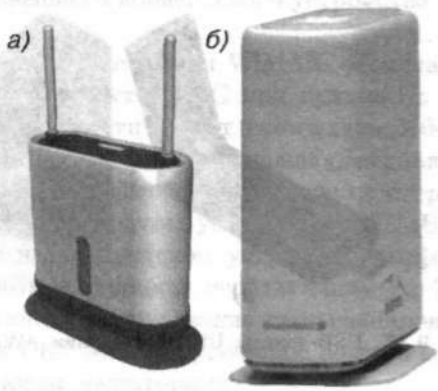


Рис. 9.15. Абонентское устройство Si: *а)* на чипсете Beceem, *б)* на чипсете Intel

Устройство *BreezeMAX 4Motion Si* (рис. 9.15) — это компактное, портативное устройство, инсталлируемое непосредственно конечным пользователем. Оно напрямую подключается к ПК (plug and play) и активируется через SIM-карту или с помощью специального приложения. Так же, как и абонентское устройство PRO, портативная абонентская станция выпускается в двух вариантах — на чипсете Intel RD2 и на чипсете компании Beceem. В первом случае устройство оснащено шестью антеннами, расположенными под корпусом. Устройство на чипсете Beceem оснащено двумя небольшими всенаправленными антеннами.

Выпускается несколько вариантов устройств BreezeMAX 4Motion Si для каждого из диапазонов 2,3; 2,5 и 3,5 ГГц. Все они включают обязательный интерфейс IEEE 802.3 Ethernet 10/100-BaseT (от 1 до 4 портов RJ-45). Опционально устройства оснащаются модулем IEEE 802.11b/g для организации локальной точки доступа, а также голосовым шлюзом для передачи VoIP.

Рис. 9.16. Абонентское устройство в формате PC card

BreezeMAX 4Motio PC Card (рис. 9.16) — это сетевой адаптер на чипсете Beceem, позволяющий подключать к сети мобильного WiMAX переносной компьютер. Он выпускается для каждого из диапазонов 2,3; 2,5 и 3,5 ГГц и при ширине канала 10 МГц обеспечивает максимальную скорость в нисходящем канале до 20 Мбит/с, в восходящем — до 7 Мбит/с. Ширина канала задается при конфигурации и может составить 5; 7; 8,75 и 10 МГц. На карте находятся две выдвигающиеся антенны, регулирование положения которых, при необходимости, позволит улучшить прием сигнала. Возможна работа на удалении до 5 км от базовой станции.

К середине 2009 года было известно о нескольких USB-устройствах (донглах), прошедших тест на совместимость с базовыми станциями BreezeMAX 4Motion. Среди них — устройство US210 компании AWB и WU211 компании Quanta computers (рис. 9.17).

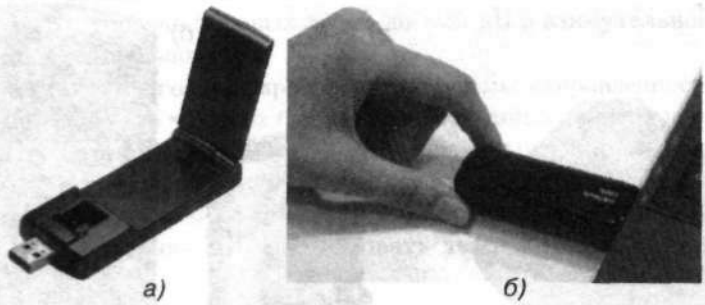


Рис. 9.17. USB-донглы US210 компании AWB *(а)* и WU211 компании Quanta computers *(б)*

Устройство US210 — это WiMAX USB-адаптер для ПК. Адаптер полностью соответствует стандарту IEEE 802.16e и поддерживает мобильное беспроводное соединение на скорости до 130 км/ч. Устройство инсталлируется и настраивается конечным пользователем, пиковая скорость в нисходящем канале — до 33 Мбит/с, в восходящем — до 7 Мбит/с. Работает в частотных диапазонах 2,3; 2,5 и 3,5 ГГц. Мощность передатчика — 23 дБм, усиление антенны — 2 дБ от изотропной мощности. Благодаря одной передающей и двум приемным антеннам US210 поддерживает MIMO-технологию. Энергопотребление — 2,4 Вт при мощности в антенне 23 дБм.

WiMAX-адаптер WU211 от Quanta computers аналогичен рассмотренному выше устройству. Он работает в диапазоне 2,496–2,69 ГГц. Максимальная выходная мощность на антенне: 23±1 дБм, усиление антенны — 2 дБ от изотропной мощности.

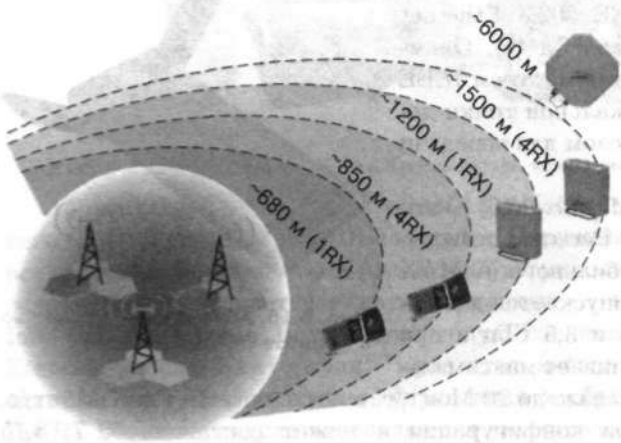


Рис. 9.18. Типовые зоны обслуживания для различных абонентских устройств

Зона покрытия БС зависит (рис. 9.18) не только от мощности передачи БС и абонентского устройства, но и от типа абонентского устройства, в также

условий работы (рельеф и тип застройки). Существенно влияют на дальность работы и условия видимости. При прямой видимости теоретическое ограничение дальности составляет 54 км, практически удавалась получать устойчивую связь с пропускной способностью порядка 3 Мбит/с на расстоянии 30 км от базовой станции с использованием внешнего абонентского устройства с направленной антенной.

Опыт эксплуатации оборудования 4Motion в сети WiMAX «Комстар» показал, что в городе зона действия одного сектора БС при приеме на антенну WiMAX USB-адаптера вне прямой видимости может составлять 300–1500 м, а при прямой видимости — до 6 км. Такие значения дальности объясняются несимметричностью восходящего и нисходящего каналов: если со стороны БС используются антенны с высоким коэффициентом усиления (15–17 дБи) и передатчики мощностью до 5 Вт, то на абонентском устройстве коэффициент усиления всенаправленных антенн 3 дБи, а мощность передатчика — до 200 мВт.

Следует отметить, что все описанное оборудование сертифицировано и в России. Сейчас на территории РФ и СНГ это оборудование, помимо компании «Комстар», тестируется еще несколькими операторами.

9.4. Проблемы радиочастотного ресурса

Дефицит радиочастотного ресурса — наиболее деликатная тема в сетях беспроводной связи для всех промышленно-развитых стран мира.

Анализ WiMAX-форума, проведенный еще в 2007 году, позволил оценить оптимальную полосу радиочастот на одного оператора в 30–40 МГц. Оптимальной величиной частотного ресурса является полоса 30 МГц (TDD) и 2x30 МГц (FDD). При этом удается покрыть территорию базовыми станциями, состоящими из трех секторов по 10 МГц каждый. Следует отметить, что в последующих релизах WiMAX, когда будет доступна полоса по 20 МГц на сектор, естественно 30 МГц могут превратиться в 60 МГц. Поэтому отечественные реалии дефицита и распределения радиочастот существенно снижают бизнес-потенциал операторов, которым, как правило, приходится довольствоваться полосой в 15–20 МГц. Дефицит частотного ресурса из-за высокого уровня внутрисистемных помех снижает радиус действия базовых станций. В результате растет их число и, соответственно, затраты на создание транспортной сети. А капитальные затраты на транспортные сети составляют 65–75% в сети WiMAX. При ширине выделенной полосы 5 МГц, чтобы достичь того же качества услуг, что и при полосе 30 МГЦ, потребуется в 6–7 раз больше инвестиций.

Отметим, что все радиочастотные диапазоны (2–3 ГГц), выделенные по остаточному принципу под наиболее перспективные сети (мобильный WiMAX, LTE и т. п.), не очень приспособлены для передачи в условиях сложной городской застройки многомегабитных информационных потоков. А учитывая, что предельные технические характеристики этих технологий достигаются при полосах рабочих частот 20 МГц и более, дефицит частотного ресурса становится лишь острее. Использование же систем типа LTE или IEEE 802.16m в узких полосах хоть и возможно, но коммерчески едва ли нецелесообразно, поскольку при этом они практически не будут иметь преимущества перед уже действующими беспроводными системами.

Поэтому дальнейшая конверсия радиочастот — важная часть пути к 4G. Хотим мы того или нет, понижение рабочих частот — это естественный путь кардинального улучшения качества обслуживания для всех перспективных сетей широкополосной беспроводной связи. И наиболее логичный путь — переход в радиочастотные диапазоны 450–900 МГц, где затухание радиоволн существенно ниже. Такой вариант построения используется очень мало, поскольку все частоты в этом диапазоне давно заняты. Но именно с этим путем связаны успехи сетей CDMA-450 (450 МГц) в России или сетей UMTS в Австралии (850 МГц). Под аналогичные цели в США освободили диапазон 700–790 МГц от телевидения и уже распродали его на аукционе.

В Европе стратегическим радиочастотным ресурсом считается диапазон 770–862 МГц, который в ближайшие годы также планируется освободить от аналогового ТВ-вещания. Министерства связи Евросоюза и Европейская комиссия в декабре 2007 года одобрили использование GSM-диапазона частот 900 МГц для сетей передачи данных третьего поколения, включающих, наряду со стандартами EVDO и HSPA, также WiMAX. Однако при действующих сверхзагруженных сетях GSM такой переход без снижения качества обслуживания абонентов представляет трудноразрешимую проблему. Поэтому во многих странах, в частности, в Индии, также есть предложения по выделению под перспективные сети 4G диапазона частот 700–790 МГц, как в США.

С другой стороны, при уменьшении значений рабочих частот объем доступного частотного ресурса ощутимо снижается (20-МГц полос в диапазоне 0,9 ГГц можно разместить меньше, чем в диапазоне 3,5 ГГц). При очевидном дефиците операторского ресурса вообще переход в низкочастотные диапазоны нельзя рассматривать как единственный вариант развития перспективных сетей ШБД. Очевидно, необходима комбинация низкочастотных и высокочастотных сегментов, например, с разделением по виду застройки местности, по степени мобильности абонентов или по виду трафика. В перспективе, наверное, такой подход сводится к концепции когнитивного радио. Этот термин ввел в 1999 году Дж. Митола [1]. Он означает парадигму беспроводных телекоммуникаций, при которой каждая сеть или узел модифицирует свои приемопередающие параметры для достижения максимальной эффективности и предотвращения взаимных помех между пользователями. Для этого системы когнитивного радио ведут постоянный мониторинг внешней и внутренней (по отношению к сети) радиообстановки, включая такие параметры, как загруженность радиоспектра, поведение пользователей (тип трафика) и состояние сети. Однако подобные системы — это пока еще будущее, хоть и не слишком отдаленное. Поэтому рассмотрим основные принципы и процедуры выделения радиочастотного ресурса в России.

9.4.1. Принципы выделения частотного ресурса в России

Радиоизлучающая аппаратура, такая, как радиорелейные системы, системы спутниковой связи, беспроводные устройства сетей передачи информации и др., требует определенного порядка выделения частот для их применения. В случае свободного использования радиочастот различные беспроводные системы могут создавать друг другу мешающие их работе помехи. Чтобы этого не случилось,

в каждой стране мира имеется национальный регулятор, занимающийся распределением радиочастот. Регулирование использования радиочастотного спектра является исключительным правом государства.

В Российской Федерации сегодня регулятор представлен такими ведомствами, как Министерство связи и массовых коммуникаций РФ (Минкомсвязи), Государственная комиссия по радиочастотам (ГКРЧ) при Минкомсвязи, Агентством «Россвязь» Минкомсвязи, ФГУП «Главный радиочастотный центр», Роскомсвязьнадзор, Минобороны, МВД, ФСО и ФСБ. В этом процессе также участвуют Росприроднадзор, Рострой, Роспотребнадзор, Ростехрегулирование и таможенные органы. Регулятор устанавливает основные принципы и общие условия назначения (присвоения) радиочастот радиоэлектронным средствам (РЭС) различного применения, порядок их согласования в целях обеспечения электромагнитной совместимости (ЭМС). Важно, что частоты в РФ используются либо совместно гражданским обществом с правительственными организациями (41,85%), либо для правительственных целей (58,15%).

Полосы частот для РЭС гражданского применения выделяет ГКРЧ, а присвоение номиналов радиочастот в выделенных полосах — государственная радиочастотная служба в лице ФГУП «Главный радиочастотный центр». Операторы связи вправе использовать присвоенную часть частотного ресурса для осуществления разрешенного им вида деятельности при выполнении условий, определенных в разрешительных документах, выданных ГКРЧ и органами государственной радиочастотной службы.

Радиочастотные органы при назначении радиочастот РЭС различного применения руководствуются:

- требованиями законов Российской Федерации, указов и распоряжений Президента РФ, постановлений и распоряжений Правительства РФ, таблицей распределения полос частот между радиослужбами Российской Федерации, решениями ГКРЧ, определяющими условия использования полос радиочастот;
- международными обязательствами, принятыми Российской Федерацией в рамках ITU и двусторонних или многосторонних соглашений с администрациями иностранных государств, а также обязательствами, принятыми в рамках международных организаций, членом которых является Российская Федерация;
- условиями, определенными решениями ГКРЧ о совместном использовании полос радиочастот РЭС гражданского применения и РЭС, обеспечивающих потребности правительственной связи, правопорядка, безопасности и обороны России;
- условиями обеспечения электромагнитной совместимости РЭС различного применения и условиями совмещения различных радиослужб в Российской Федерации.

Радиочастотные органы при назначении радиочастот РЭС различного применения соблюдают следующий порядок.

В полосах радиочастот категории «ПР» РЭС правительственного применения пользуются правом преимущественного использования радиочастот по отношению к РЭС гражданского применения. Назначают радиочастоты в этих полосах

для РЭС правительственного применения радиочастотные органы Минобороны России. Возможность и условия использования радиочастот в этих полосах РЭС гражданского применения определяются отдельными решениями ГКРЧ. На основании этих решений органы государственной радиочастотной службы назначают радиочастоты конкретным РЭС гражданского применения.

В полосах радиочастот категории «ГР» РЭС гражданского применения имеют преимущество по отношению к РЭС правительственного применения. Радиочастоты в этих полосах для РЭС гражданского применения назначают органы государственной радиочастотной службы без согласования с Минобороны России. Возможность и условия использования радиочастот в этих полосах РЭС правительственного применения определяются отдельными решениями ГКРЧ. На основании этих решений назначение радиочастот конкретным РЭС правительственного применения осуществляет Минобороны России.

Радиочастоты для РЭС гражданского применения, входящие в состав спутниковых или космических систем, назначает ФГУП «Главный радиочастотный центр» на условиях, определяемых соответствующими решениями ГКРЧ.

Радиочастоты РЭС различного применения назначаются на основании расчетов на электромагнитную совместимость с другими РЭС, работающими в совмещенных полосах частот и расположенными в районах размещения данных РЭС, исходя из реальной электромагнитной обстановки. Необходимость разработки условий совместного использования и/или норм частотно-территориального разноса определяется соответствующими решениями ГКРЧ.

Разработка норм частотно-территориального разноса (ЧТР) и методик расчетов ЭМС обеспечивается заявителями. К разработке норм ЧТР и методик расчета ЭМС привлекаются научно-исследовательские организации Минкомсвязи и Минобороны России. Нормы ЧТР и методики расчета ЭМС РЭС, утвержденные ГКРЧ, могут применяться радиочастотными органами для РЭС любого заявителя (пользователя). Нормы ЧТР по мере необходимости уточняются и дополняются с учетом изменений технических характеристик РЭС.

9.4.2. Выделение частотного ресурса для систем ШБД

В связи с быстрым развитием новых беспроводных технологий в области связи и вещания вопрос о выделении соответствующего спектра частот становится одним из основных. Дефицит частот наблюдается при внедрении таких технологий, как широкополосная передача данных, цифровое телевидение, цифровое вещание, сотовая связь следующих поколений.

Однако получение разрешительных документов является длительной процедурой, требующей значительных материальных затрат. Проиллюстрируем сказанное примером получения разрешительных документов для одной из первых в РФ сетей ШБД — сети Radionet, разработанной и внедренной под руководством одного из авторов книги.

Сеть Radionet [2, 3] была предназначена для подключения к Интернету многочисленных организаций науки и образования Москвы. В начале 1996 года была направлена заявка для получения разрешения ГКРЧ на использование полосы частот. Спустя пять месяцев было получено Решение ГКРЧ № 1658-ОР от 25 ноября 1996 года «Об использовании радиочастот для закупаемой по импорту аппаратуры». Согласно этому документу разрешалось использование радиочастот в

пределах полосы 2400–2483,5 МГц для закупаемой аппаратуры, предназначенной для создания на территории Москвы сети передачи данных. Получению разрешений на использование конкретных номиналов частот в Главсвязьнадзоре (в 2001 году вопросы присвоения номиналов радиочастот в выделенных полосах были переданы ФГУП «Главный радиочастотный центр») предшествовало проведение расчетов на ЭМС с РЭС военного и гражданского назначения. Наконец, 27 января 1998 года для базовых станций сети Radionet было получено решение Главгоссвязьнадзора о назначении радиочастот для установки РЭС (№ 07-5-18/00781). Далее более полугода оформлялось временное разрешение на строительство объекта связи. После дополнительных натурных испытаний полученное ранее временное разрешение было переоформлено на разрешение на использование радиочастот для эксплуатации. Следует отметить, что в течение всего периода оформления разрешительной документации сеть Radionet успешно функционировала в «опытном» режиме, предоставляя услуги доступа в Интернет многочисленным научным и образовательным учреждениям Москвы.

Длительная процедура получения разрешений в значительной мере сдерживала разворачивание на территории РФ беспроводных сетей, хотя потребность в них, учитывая обширность территории РФ и слабость наземной кабельной инфраструктуры, весьма велика. Упрощение доступа к радиочастотам для беспроводных сетей передачи информации можно ожидать после того, как будут разработаны соответствующие модели совместного использования перспективных радиотехнологий и уже существующих, с введением пространственно-временного разделения частот между потребителями с разными технологиями использования радиоспектра. Несомненно, увеличение доступного спектра станет возможным при успешной конверсии радиочастотного спектра. Необходимы также другие меры административного характера, которые предпринимаются ГКРЧ.

Учитывая это, в 1998 году ГКРЧ приняла очень важное для развертывания сетей ШПД решение (протокол № 7/6 от 29.06.98г.) об условиях использования радиочастот в диапазоне 2400–2483,5 МГц. Признавая необходимость упрощения процедуры согласования радиочастот для РЭС, ГКРЧ постановил разрешить использование юридическим и физическим лицами на вторичной основе отдельных радиочастот в пределах полосы радиочастот 2400–2483,5 МГц для закупаемой за границей аппаратуры беспроводной передачи данных без оформления частных решений ГКРЧ для каждого конкретного заявителя при условии выполнения ряда требований. Конкретные номиналы рабочих частот для этой аппаратуры должны были назначаться, как и прежде, в установленном порядке Главсвязьнадзором России.

В 2003 году ФГУП «Главный радиочастотный центр» принял меры для дальнейшего упрощения разрешительной системы, в частности для некоторых видов заявок введена «одношаговая» процедура оформления. Основным результатом внедрения «одношаговой» процедуры явилось сокращение сроков рассмотрения на назначение радиочастот, особенно когда разработку частотно-территориальных планов и расчеты ЭМС с РЭС гражданского применения производит сам «Главный радиочастотный центр», без соисполнителей. Разрешительная процедура стала занимать 4–5 месяцев против 6–18 месяцев без применения «одношаговой» процедуры. Если заявка требует проведения расчетов ЭМС, которые осуществляют внешние организации, сроки увеличиваются.

Таблица 9.8. Полосы радиочастот, одобренные решениями ГКРЧ для использования в широкополосных беспроводных сетях связи

Полосы радиочастот, МГц	Ширина полосы, МГц	Решение ГКРЧ
1787,5–1802,5	15	От 23.10.2006 № 06-17-03-001
2400–2483,5	83,5	Внутриофисные СБД — от 6.12.2004 № 04-03-04-003; уличные СБД — от 28.11.2005 № 05-10-01-001
2500–2530, 2560–2570, 2620–2630, 2660–2670, 2680–2690	70	От 4.09.2006 № 06-16-03-001
3400–3450, 3500–3550	100	От 23.12.2002 № 23/3; От 1.12.2003 № 30/4; От 28.11.2005 № 05-10-01-001
5150–5350, 5650–6425	975	От 23.12.2002 № 23/5; От 28.11.2005 № 05-10-01-001
Итого	1160	

Полосы радиочастот, одобренные ГКРЧ для использования в широкополосных беспроводных сетях, представлены в табл. 9.8. Приведенные в ней решения ГКРЧ и соответствующие тактико-технические характеристики РЭС, ориентированы в основном на внедрение систем фиксированного беспроводного доступа. В целях содействия процессу внедрения новых радиотехнологий (включая мобильный WiMAX) на заседании ГКРЧ от 19 марта 2009 года (протокол № 09-02) был принят ряд важных решений:

1. Принять к сведению сообщение по вопросу «Об использовании полос радиочастот в диапазонах частот 2,5 ГГц и 3,5 ГГц РЭС беспроводного доступа».
2. Признать возможным создание опытно-коммерческих сетей мобильного беспроводного доступа на базе действующих сетей беспроводного доступа в диапазоне частот 2,5 ГГц на территории Москва и Санкт-Петербурга, а также на территории, ограниченной расстоянием 30 км от границ указанных городов, с целью оценки изменения электромагнитной обстановки в результате увеличения мощности излучения базовых станций до 40 Вт.
3. Признать возможным создание опытных зон беспроводного доступа на базе действующих сетей беспроводного доступа в диапазоне частот 3,5 ГГц на территории городов Красногорск, Щербинка и Подольск Московской области, а также на территории, ограниченной расстоянием 5 км от границ указанных городов, с целью проведения экспериментально-исследовательских работ по определению допустимо возможной мощности излучения базовых станций и оценки изменения электромагнитной обстановки в результате изменения мощности излучения базовых станций.
4. Операторам связи до 10 мая 2009 г. представить в ФГУП «Научно-исследовательский институт радио» результаты эксплуатации опытно-коммерческих сетей и экспериментально-исследовательских работ, проводимых в соответствии с пунктами 2 и 3 настоящего решения ГКРЧ, для их обобщения и учета при подготовке работ, проводимых в соответствии с пунктом 5 настоящего решения ГКРЧ.

5. Рекомендовать ФГУП «Научно-исследовательский институт радио» совместно с научно-исследовательскими учреждениями Минобороны России и заинтересованными операторами связи провести работы по определению:
   - минимально необходимого частотного ресурса для создания сетей беспроводного доступа в диапазонах частот 2,5 ГГц и 3,5 ГГц;
   - возможности применения РЭС мобильного беспроводного доступа в указанных полосах радиочастот;
   - возможности изменения технических параметров РЭС беспроводного доступа (включая возможность увеличения мощности базовых станций), утвержденных решением ГКРЧ от 28 ноября 2005 года № 05-10-01-001 и от 4 сентября 2006 года № 06-16-03-001;
   - ФГУП «Научно-исследовательский институт радио» по результатам выполненных работ до 1 июня 2009 года направить в ГКРЧ предложения для принятия соответствующего решения Комиссии.

6. Поручить Федеральной службе по надзору в сфере связи, информационных технологий и массовых коммуникаций провести проверку выполнения пользователями радиочастотного спектра условий выделения полос радиочастот в указанных полосах радиочастот для РЭС беспроводного доступа, указанных в соответствующих решениях ГКРЧ.

   Результаты выполненных работ представить в ГКРЧ до 10 мая 2009 года.

7. Федеральной службе по надзору в сфере связи, информационных технологий и массовых коммуникаций продолжить выполнение работ по подготовке заключений экспертизы о возможности использования РЭС беспроводного доступа и об их электромагнитной совместимости с действующими и планируемыми для использования РЭС с целью расширения действующих сетей беспроводного доступа только в границах населенных пунктов Российской Федерации, на территории которых пользователь радиочастотного спектра выполнил все условия выделения полос радиочастот, указанные в соответствующих решениях ГКРЧ.

   Технические характеристики РЭС должны соответствовать тактико-техническим характеристикам, указанным в приложении № 2 к решению ГКРЧ от 28 ноября 2005 года № 05-10-01-001 и в приложении № 1 к решению ГКРЧ от 4 сентября 2006 года № 06-16-03-001.

8. Федеральной службе по надзору в сфере связи, информационных технологий и массовых коммуникаций продолжить осуществлять присвоение (назначение) радиочастот для РЭС беспроводного доступа в диапазонах частот 2,5 ГГц и 3,5 ГГц только на территории тех населенных пунктов Российской Федерации, в которых пользователь радиочастотного спектра выполнил все условия выделения полос радиочастот, указанные в соответствующих решениях ГКРЧ.

9. Пункт 6 решения ГКРЧ от 20 января 2009 года № 09-01-07 изложить в следующей редакции:

   «Приостановить рассмотрение радиочастотных заявок граждан Российской Федерации и российских юридических лиц на выделение полос радиочастот в диапазонах частот 2,5 ГГц и 3,5 ГГц с целью создания новых сетей

беспроводного доступа и применения дополнительных типов РЭС в действующих сетях беспроводного доступа, а также приостановить действие пункта 3 решения ГКРЧ от 4 сентября 2006 года № 06-16-03-001 и пункта 3 решения ГКРЧ от 28 ноября 2005 года № 05-10-01-001 в части выделения полос радиочастот в диапазонах частот 2,5 ГГц и 3,5 ГГц радиоэлектронным средствам беспроводного доступа для указанных целей до рассмотрения результатов работ, указанных в пункте 4 настоящего решения ГКРЧ».

Все отмеченное выше характеризует, с одной стороны, весьма сложную и запутанную ситуацию с радиочастотным спектром в РФ. А именно:

- процедура получения частот по-прежнему привязана к адресам базовых станций и типам оборудования. Что требует большого объема проектных и изыскательских работ на начальном этапе построения сети;
- практически отсутствуют формально свободные полосы, хотя многие операторы не используют или слабо используют выданные частоты;
- операторы слабо мотивированы освобождать неиспользуемые частоты, не приносящие доход;
- в одних и тех же диапазонах работают РЭС различных стандартов и типов, что затрудняет анализ ЭМС.

С другой стороны, последние мероприятия регулятора РФ в области связи свидетельствуют о том, что все приведенные проблемы понятны и есть надежда на постепенное исправление ситуации. По крайней мере, формально в России присутствует достаточно много операторов WiMAX и ШБД и в диапазоне 2,5 и 3,5 ГГц. Есть надежда, что после приостановки выдачи новых разрешений и наведения порядка со старыми создастся обстановка для получения частотных присвоений со стимуляцией быстрого построения сетей.

Литература

1. Mitola, J., III; Maguire, G. Q., Jr. Cognitive radio: making software radios more personal. — IEEE Personal Communications, Volume 6, Issue 4, Aug. 1999, p. 13–18.
2. Вишневский В. М. Теоретические основы проектирования компьютерных сетей. — М.: Техносфера, 2003. — 512 с.
3. Вишневский В. М., Семенова О. В. Системы поллинга: теория и применение в широкополосных беспроводных сетях. — М.: Техносфера, 2007. — 320 с.

ГЛАВА 10

РЕАЛИЗОВАННЫЕ ПРОЕКТЫ WiMAX

10.1. Развертывание сетей WiMAX в мире

Как полагают аналитики корпорации IDC, в 2012 году число пользователей беспроводного Интернета превысит 30% от общего населения Земли. При этом отмечается драматический рост Интернет-трафика — к примеру, только сервис YouTube генерирует трафик больше, чем весь Интернет восемь лет назад. На это вынуждены реагировать владельцы телекоммуникационной инфраструктуры увеличением пропускной способности сетей и развитием новых технологий. Человек в 21 веке хочет иметь доступ к сети всегда, везде и с хорошей скоростью.

Сети WiMAX развиваются лишь несколько лет, но весьма стремительно. На февраль 2009 года в 139 странах мира было развернуто более 468 WiMAX-сетей (включая сети стандарта 802.16 в диапазоне 5 ГГц) (рис. 10.1, табл. 10.1 и 10.2). Более 430 млн. человек живет в зоне покрытия WiMAX. Причем в 2008 году было развернуто 200 WiMAX-сетей. В 2009 году ожидается запуск как минимум 100 новых коммерческих сетей WiMAX. Прошли WiMAX-аукционы в развивающихся странах — Индии, Малайзии, Бразилии и др.

Рис. 10.1. Динамика роста сетей WiMAX

В конце третьего квартала 2008 года, по данным компании Maravedis, насчитывалось 2,68 млн. абонентов сетей WiMAX (включая сети ШБД, например, на основе оборудования *Сапору, компания Motorola*) — рост по сравнению

с третьим кварталом 2007 года составил 91%. По данным компании Informa Telecoms & Media, это число на конец 2008 года составило 3,6 млн. абонентов (с ростом к 2013 году до 103 млн.). Эксперты компании J&P Consulting ожидают, что к 2012 году объем мирового рынка WiMAX достигнет 60 млн. абонентов. Из них порядка 65% будет приходиться на мобильные сети. Столь большой разброс данных говорит лишь о молодости и высокой динамике данного рынка. Несмотря на высокую динамику роста, абсолютные показатели развития абонентской базы остаются весьма скромными — в среднем 15 тыс. абонентов на сеть.

Таблица 10.1. Распределение сетей ШБД по регионам

Регион	Число сетей
Африка	95
Азиатско-Тихоокеанский регион	72
Центральная и Южная Америка	96
Восточная Европа	75
Ближний Восток	18
Северная Америка (США и Канада)	48
Западная Европа	64
Всего	468

В первом квартале 2008 года был отмечен значительный рост рынка WiMAX-оборудования. Согласно отчету исследовательской компании Infonetics Research, объем данного рынка в тот период увеличился на 59% (по сравнению с первым кварталом 2007 года) и достиг 363 млн. долл. Главным движущим фактором такого роста стал крайне высокий уровень продаж оборудования для мобильного WiMAX. В первом квартале 2008 года этот сегмент вырос на 141%, впервые обойдя по доходности фиксированный WiMAX. Столь стремительному развитию способствовало развертывание большого количества новых WiMAX-сетей, а также расширение уже существующих. К технологии WiMAX стали проявлять интерес национальные операторы: Sprint-Clearwire в США, SK Telecom и KT в Южной Корее, Wateen в Пакистане, BSNL и Tata Communications в Индии, а также Vodafone и Orange в Европе.

В первом квартале 2008 года лидером на рынке WiMAX по уровню доходов стала компания Motorola. Американский производитель первым на этом рынке заработал свыше 50 млн. долл. за один квартал. В сегменте оборудования для фиксированного WiMAX первенство удерживает компания Alvarion.

Таблица 10.2. Распределение сетей ШБД по диапазонам частот (считались только сети, для которых этот параметр известен)

Диапазон, ГГц	Число сетей
2,3	26
2,5	58
3,3	9
3,5	230
5–6	18
Итого	341

WiMAX-технологии продолжают опережать LTE на три года. По отзывам независимых экспертов, сервисы WiMAX по скорости и качеству обслуживания в развернутых мобильных сетях, даже на весьма требовательных рынках Южной Кореи и США, более чем в три раза превышают аналогичные возможности сетей сотовой связи 3G.

Хотя пока число абонентов сетей фиксированного ШБД продолжает доминировать (по данным Market Intelligence & Consulting Insti-

tute (Тайвань) абоненты фиксированных сетей в третьем квартале 2008 года составляли 76,9%), мир начинает стремительно переходить на мобильные сети. Уже известны несколько действующих мобильных WiMAX-сетей. Одна из крупнейших из них — сеть Korea Telecom (KT), насчитывающая свыше 400 тыс. абонентов.

Впечатляют планы компании Clearwire. Этот оператор после объединения своих ресурсов с оператором сотовой связи Sprint-Nextel стал ведущим в США оператором WiMAX, намеревающимся построить мобильную сеть ШБД национального масштаба. Среди стратегических инвесторов Clearwire, помимо Sprint, такие компании, как Google, Intel Capital и др., которые вложили в этот проект 3,2 млрд. долл. (без учета доли Sprint). Уже развернуты сети в Балтиморе и Портленде, в 2009 году должна начаться эксплуатация сетей еще в 10 городах США — в Атланте, Шарлотте, Чикаго, Далласе, Форт-Уэрте, Гонолулу, Лас-Вегасе, Филадельфии, и Сиэтле. В 2010 году планируется запустить сети в Бостоне, Хьюстоне, Нью-Йорке, Сан-Франциско и Вашингтоне. К концу 2008 года у Clearwire насчитывалось 475 тыс. абонентов, но не все они были пользователями сетей WiMAX. Однако в 2010 году компания обещает создать сеть, в зоне действия которой будет проживать свыше 120 млн. человек.

Значительную роль в реализации проекта сети Sprint Clearwire сыграла компания Intel, для которой развитие технологии мобильного WiMAX является, без сомнения, одним из самых важных проектов.

Японская компания KDDI получила лицензию на сеть WiMAX (2,5–2,7 ГГц) в конце 2007 года и планировала инвестировать в нее 1,3 млрд. долл. в течение пяти лет. Группа KDDI, включающая в себя японского производителя электроники Kyocera планирует запуск услуг WiMAX на 2009 год и предполагает к 2013 году охватить 90% территории Японии, включая 5,6 млн. пользователей.

Индийская компания Tata Communications разворачивает крупную WiMax-сеть, инвестировав в ее строительство 500 млн. долл. К марту 2009 года было подключено порядка 200 тыс. домашних клиентов. Потенциальный рынок широкополосного доступа в Индии, насчитывающей 1,1 млрд. жителей, огромен. По данным к концу января 2008 года, в стране было 3,24 млн. абонентов широкополосных сетей подписчиков. Tata Communications рассчитывает занять большую часть этого рынка.

Мобильные сети развернуты во многих странах Восточной и Западной Европы, а также Азии и Америки (см. цв. вклейку). Однако процесс этот находится в самой начальной стадии развития. Рассмотрим некоторые из них (подробнее смотри www.wimaxforum.org).

Сеть WiMAX (WiBro) компании KT

Южнокорейская компания KT, также известная под названием Korea Telecom, была основана (выделилась из Министерства связи Кореи) в 1981 году, как оператор проводных сетей. В последующие последние несколько лет KT, используя свою проводную сеть, оказалась лидирующим оператором DSL-услуг. После окончательной приватизации в 2002 году, KT полностью сосредоточилась на высокоскоростном Интернет-доступе. В январе 2005 года KT была одной из трех компаний, получивших общенациональную лицензию на полосу шириной 27 МГц в диапазоне 2,3–2,4 ГГц для предоставления широкополосных беспроводных услуг. KT приступила к строительству сети мобильного ШБД

на основе технологии WiBro (разработка компании Samsung на базе стандарта IEEE 802.16e).

В декабре 2005 года KT впервые продемонстрировала работу мобильного широкополосного подключения — с мобильной станцией, движущейся со скоростью 120 км/ч, было установлено соединение со средней пропускной способностью в 3 и 1 Мбит/с в нисходящем и восходящем каналах, соответственно.

Пробный коммерческий запуск сервиса WiBro состоялся в июне 2006 года, за которым последовал полноценный ввод в коммерческую эксплуатацию в апреле 2007 года, сделавший KT первой компанией в Корее запустившей сервис WiBro в коммерческое использование. Фактически, это была первая коммерческая сеть мобильного ШБД в мире. Серьезным недостатком WiBro было неполное соответствие рекомендациям WiMAX-форума (в частности, не поддерживались некоторые технологии кодирования, заложенные в IEEE 802.16e), а также не поддерживались голосовые сервисы (только передача данных). В 2007 году компания Samsung доработала WiBro до требований «полноценного» WiMAX, в результате чего последние сети Korea Telecom делаются уже на этом оборудовании.

К концу 2007 года KT установила приблизительно 600 базовых станций, охвативших территорию с 12 млн. жителей. Абонентская база сети WiBro насчитывала тогда 106 тыс. пользователей, представляющих 43% мобильного широкополосного рынка Кореи. К июню 2008 года, спустя 14 месяцев после коммерческого запуска, KT имела 200 тыс. клиентов сети WiBro, а к концу 2008 года — порядка 400 тыс. С привлекательной ценой и высокой скоростью передачи данных, намного превосходящей альтернативные мобильные сервисы, такие как HSDPA, WiBro стала наиболее динамично развивающимся беспроводным сервисом в Корее (рис. 10.2).

Рис. 10.2. Динамика роста абонентской базы сети WiBro

В 2008 году KT предоставляла доступ в Интернет со средней скоростью в нисходящем канале 3 Мбит/с и в восходящем — 1 Мбит/с за 10,84 долл. при ограничении 1 Гбайт/мес. и за 21,45 долл. при ограничении 30 Гбайт/мес. KT также предлагает ряд смешанных тарифных планов: Wi-Fi + WiBro, фиксированный высокоскоростной доступ (ADSL, VDSL или FTTH) + WiBro, фиксированный

высокоскоростной доступ + Wi-Fi + WiBro. Также KT имеет соглашения с сетевыми операторами услуг сотовой связи 3G (cdma200 EV-DO и WCDMA HSPA), прежде всего — со своей дочерней компанией KT Freetel (KTF), благодаря чему предлагает двухдиапазонные устройства 3G/WiBro: WiBro + HSDPA для пользователей USB-модемов и WiBro + EV-DO/CDMA для пользователей сматрфонов. Эти тарифные планы предоставляют 20%-ную скидку на сервис KT WiBro и 20%-ную скидку на базовую услугу KTF.

Почти 80% абонентов для доступа в сеть WiBro используют ноутбуки, поэтому USB-адаптеры — наиболее часто применяемые устройства. Некоторые модели этих адаптеров могут, кроме того, работать как MP3-плееры, приемники мобильного ТВ (T-DMV, Terrestrial Digital Multimedia Broadcasting) и т. д. Аналогичные функции поддерживают некоторые смартфоны и коммуникаторы. Это было очень важно для развития сети мобильного WiMAX. Мобильное телевещание, бесплатное в Южной Корее, предоставляет 7 телевизионных каналов, 12 радиоканалов и 8 информационных каналов, вещающих в СВЧ-диапазоне.

KT строила сеть WiMAX на основе своей, уже существующей опорной оптоволоконной сети. Кроме того, ее дочерняя компания KTF позволяет использовать ее инфраструктуру сотовой телефонной сети. Это не только обеспечивает очевидные экономические преимущества, но и сокращает время выхода на рынок, поскольку не нужно, например, создавать новую инфраструктуру для базовых станций.

Первая стадия развертывания сети WiBro компанией KT была завершена в марте 2007 года, обеспечив покрытие Сеула с 12 млн. жителей. Вторая стадия завершилась к сентябрю 2008 года с помощью оборудования WiMAX второго поколения, которое обеспечило покрытие 20 дополнительных провинциальных городов вокруг Сеула. К концу 2010 года сервис WiBro должен быть доступен более чем 40 млн. людей.

Южная Корея — это прекрасный пример высокого спроса на услуги мобильного широкополосного сервиса. Интересно, что 80% абонентов KT WiBro перестают дома пользоваться DSL-услугами (также предоставляемыми KT). К концу 2012 года KT намеревается покрыть сетью WiBro всю территорию страны и обслуживать 4 млн. абонентов.

WiMAX-сеть компании Danske Telecom

Штаб-квартира компании располагается в Копенгагене. Частотный диапазон — 3,5 ГГц, сеть состоит из 72 базовых станций, расположенных в 7 крупных городах Дании. В октябре 2005 года широкополосные сервисы были запущены в трех крупнейших городах Дании: Копенгагене, Орхусе и Оденсе. К середине февраля 2008 года Danske Telecom увеличила покрытие до 7 городов, в зоне ее покрытия оказалось более 550 тыс. семей Дании и около 40% всего населения страны. Таким образом, компания сфокусировалась на густонаселенных городских территориях, несмотря на жесткую конкуренцию со стороны крупных DSL-операторов. Общее число абонентов Danske Telecom сегодня составляет 13,5 тыс.

Спектр, выделенный для развертывания WiMAX-сети Danske Telecom, включает две общенациональные лицензии в полосе 3,5 ГГц каждая, с двумя спаренными каналами по 28 МГц — всего 112 МГц. Дополнительно выделен частотный ресурс в полосах 3,6; 10,5 и 26 ГГЦ для непосредственной связи базовых станций по схеме «точка-точка».

Предлагаемые сервисы включают множество тарифных планов с ежемесячной оплатой, соответствующей скорости передачи. Опции VoIP предоставляют третьи компании по специальным соглашениям. Типичные потребительские тарифы колеблются от 99 датских крон (~ 15 долл.) в месяц за 1 Мбит/с в нисходящем канале до 199 датских крон (30 долл.) в месяц за 3 Мбит/с. Danske Telecom пришла к выводу, что агрессивная ценовая политика — ключевой фактор в победе над крупными DSL-операторами на потребительском рынке.

Тарифы для бизнес-клиентов, очевидно, выше. Имея свою собственную общенациональную оптоволоконную сеть и дополнительные лицензии на полосы спектра 3,6; 10,5 и 26 ГГц, Danske Telecom предоставляет выделенные каналы для большого бизнеса, требующего высокой пропускной способности — до 100 Мбит/с в обоих направлениях.

Восходящие каналы в сети WiMAX варьируются в соответствии с загрузкой трафика и соглашением о QoS. Базовые станции, обслуживающие большое число клиентов, могут быть связаны одновременно и оптоволоконными, и радиорелейными линиями на случай неисправности одного из каналов.

Один из ключевых элементов сети Danske Telecom — программный инструмент управления широкополосным трафиком WiMOSS, позволяющий контролировать сетевой трафик, включая тип модуляции, интерференцию и действия пользователей. По наблюдениям Danske Telecom, 5% активных пользователей во время часов наибольшей нагрузки могут создать неблагоприятные условия для остальных 95%. Чтобы гарантировать хорошее обслуживание остальным пользователям, WiMOSS позволяет автоматически ограничивать доступ наиболее активных абонентов во время пиковой загрузки периода — но тогда и только тогда, когда сектор активных пользователей перегружен. Такой подход помогает снизить как число жалоб, так и сбоев. Кроме того, WiMOSS позволяет контролировать эффективность использования спектра и производительности оборудования каждого сектора БС, что необходимо для точечного инвестирования в увеличение пропускной способности сети.

По данным Организации экономического сотрудничества и развития (OECD), Дания с населением приблизительно 5,5 млн. человек, уже сейчас одна из наиболее обеспеченных широкополосным покрытием стран мира. Danske Telecom за несколько лет продемонстрировала, что с правильной технологией и деловым подходом можно добиться успеха в среде с высокой конкуренцией.

Danske Telecom развернула свою сеть с широкополосным беспроводном доступом на базирующемся на платформе Motorola Expedience. Это позволило Danske Telecom быстро увеличить свое присутствие на рынке фиксированного ШБД. С появлением WiMAX сертифицированного оборудования Danske Telecom смотрит на расширение и реорганизацию сети, которая будет поддерживать мобильные сервисы.

WiMAX-сеть компании DBD

DBD (Deutsche Breitband Dienste) — телекоммуникационная компания, работающая на всей территории Германии, предоставляющая высокоскоростной доступ в Интернет. Штаб-квартира находится в Гейдельберге. Основанная в 2003 году, DBD выбрала беспроводную технологию WiMAX для предоставления широкополосных услуг. Развертывание сети WiMAX началось в 2005 году, параллельно с созданием опорной сети, головного центра данных и т. п.

DBD были предоставлены лицензии на использование частот в полосе 3,5 ГГц по всей Германии (83 млн. жителей). Ширина доступных полос составляет 42 МГц по всей стране, 70 МГц — в некоторых больших городах общим населением 21 млн. человек.

Сейчас у DBD 180 базовых станций WiMAX, имеющих от трех до шести секторов, охватывающих приблизительно 400 тыс. семей в городских центрах и пригородных районах. Компания DBD — не только первый и самый большой WiMAX-оператор в Германии, но также и крупнейший владелец спектра в полосе 3,5 ГГц.

Через свой сервис DSLonair компания DBD предлагает высокоскоростной Интернет-доступ в городской и сельской местности, где ограничены услуги DSL. В столичной зоне DBD предлагает широкополосный WiMAX доступ в Интернет и IP-телефонию (VoIP) (сервис «MAXXonair»). Например, только доступ в Интернет стоит 9,99 евро в месяц, а единый тарифный план с доступом в Интернет и голосовыми сервисами — 29,99 евро в месяц.

Список продаваемых устройств для конечных пользователей включает в себя домашний модем Airspan EasyST WiMAX со встроенной антенной мощностью 7,5 дБи, а для клиентов с плохим уровнем сигнала, находящихся далеко от базовой станции, — с внешней антенной мощностью 9 дБи.

Сеть компании DBD состоит из региональных точек присутствия, связанных общенациональной опорной сетью (рис. 10.3). Каждая точка присутствия, включающая ASN-шлюз, связана арендуемой линией с пропускной способностью свыше 100 Мбит/ с кольцевой сетью узлов агрегации. Каждый из узлов агрегации (высокопроизводительный маршрутизатор/коммутатор) связан сетью с несколькими базовыми станциями. Эта сеть произвольной топологии строится на основе радиорелейных линий.

На июнь 2008 года компания обслуживала свыше 25 тыс. постоянных клиентов сети фиксированного доступа. Осенью 2008 года DBD приступила к запуску мобильной WiMAX-сети. Учитывая высокую конкуренцию (сети широкополосного доступа охватывают 19 млн. абонентов, через 3–5 лет их число возрастет до 28 млн., проникновение сотовой связи достигло 100%), компания DBD ставит целью охватить мобильной WiMAX-сетью порядка 5% абонентов в Германии. Но выручка при этом составит около 1,5 млрд. евро в год.

WiMAX-сеть компании Digital Bridge Communications

Компания Digital Bridge Communications (DBC) является поставщиком телекоммуникационных услуг, базирующихся в Эшберне (шт. Вирджиния, США). Основанная в 2005 году, корпорация DBC предоставляет услуги широкополосного доступа в Интернет в удаленных регионах страны на основе технологии WiMAX. Ее основное внимание концентрируется на населенных пунктах с численностью от 150 тыс. человек. Компания имеет представительства в 13 городах пяти штатов, где предоставляет услуги в основном частным потребителям, а также малому и среднему бизнесу. Но часть клиентов компании — крупные предприятия и правительственные организации.

Основа деятельности DBC — предоставление стационарных и мобильных Интернет-услуг на основе WiMAX. DBC является владельцем (в некоторых случаях арендатором) частот в диапазоне 2,5 ГГц в некоторых областях штатов Айдахо, Монтана, Вайоминг, Индиана и Южная Дакота. В Вирджинии DBC арендует частоту 2,3 ГГц. Сеть развернута на оборудовании компании Alvarion.

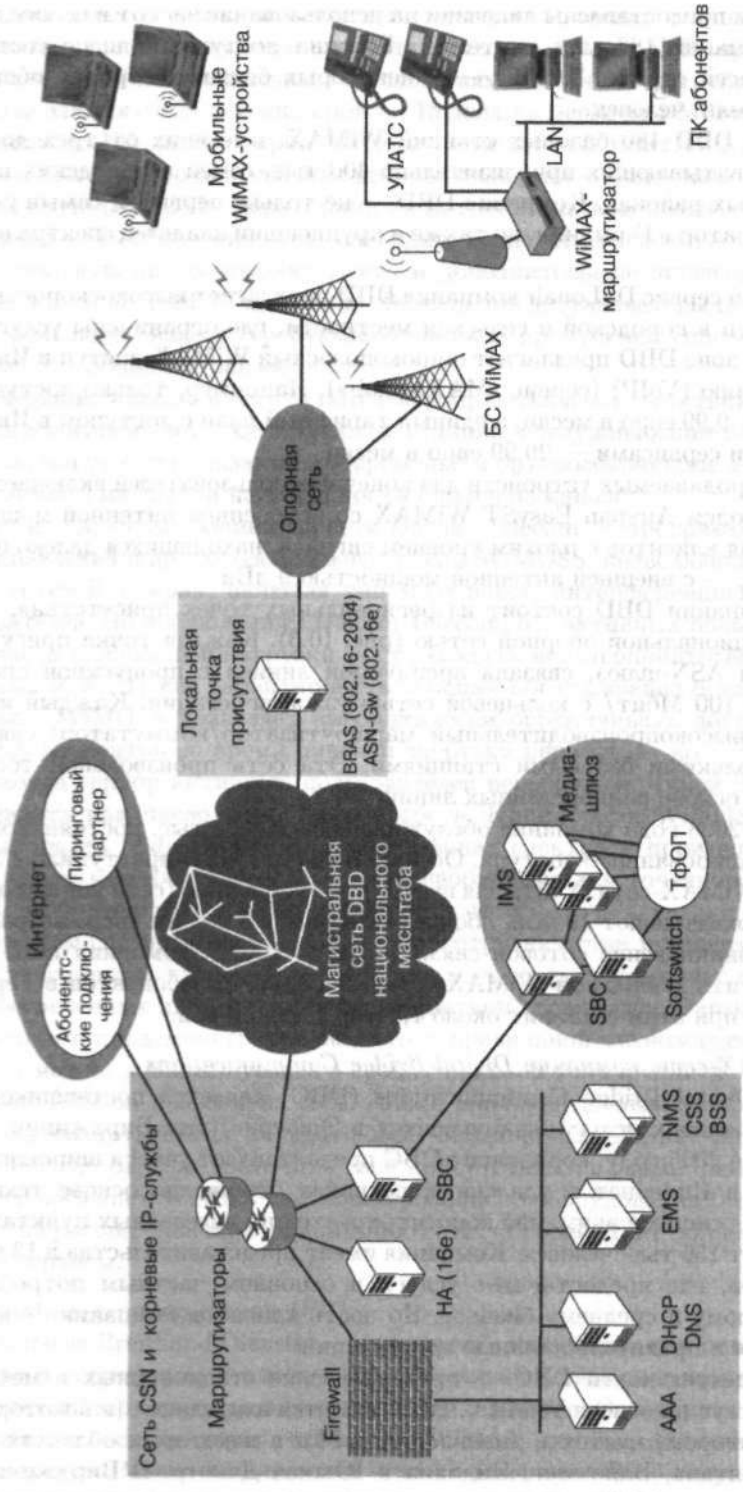


Рис. 10.3. Сеть компании DBD

Сеть DBC основана на принципе plug-and-surf (подключился-и-работай), с использованием самонастраивающегося WiMAX-оборудования Alvarion. Для клиентов, расположенных на краю зоны покрытия, DBC использует специально установленные антенны с повышенным коэффициентом усиления.

Во второй половине 2008 года DBC запускает сеть гостевого доступа, состоящую из «горячих точек» и точек временного доступа. Первоначально такие сети будут запущены в таких туристических центрах, как Сан-Волли (шт. Айдахо) и в Джексоне (шт. Вайоминг). Кроме того, DBC намерена предоставлять и голосовую связь (VoIP) для клиентов своих WiMAX-сетей.

DBC удалось построить эффективную партнерскую сеть, которая предоставляет экономически эффективные вспомогательные службы и сетевые инфраструктуры. Эта стратегия позволяет DBC сосредоточиться на создании и обслуживании WiMAX-сетей.

Типичная небольшая сеть WiMAX будет состоять из 1–6 базовых станций, часто с несколькими секторами для обеспечения полного покрытия и возможности удовлетворения потребительского спроса. При полном подключении 15 городов, WiMAX-сеть DBC сможет охватить свыше 2,5 млн. пользователей.

WiMAX-сеть компании Iberbanda

Компания Iberbanda была основана в 2000 году в Испании, ее штаб-квартира находится в Мадриде. Iberbanda предлагает широкий спектр услуг высокоскоростного Интернета, телефонии, передачи данных и др. Компания имеет общенациональную лицензию в диапазоне 3,5 ГГц, две полосы по 20 МГц (всего 40 МГц). Около 900 БС WiMAX охватывают почти 30% территории Испании. Это делает Iberbanda обладателем одной из крупнейших действующих сетей WiMAX в Европе. Базовое оборудование поставляет компания Alvarion.

Целевой сегмент рынка Iberbanda направлен на клиентов в пригородных и сельских районах, где DSL и кабельные сети немногочисленны или совсем отсутствуют. За 2007 год компания увеличила число пользователей на 82% и закончила год с 33 тыс. абонентов. Iberbanda рассчитывала достичь 50 тыс. абонентов к концу 2008 года.

WiMAX-сеть компании Liberty Technologies

Корпорация Liberty Technologies была основана в 2002 году в Панаме, стране с населением, превышающим 3 млн. человек. Liberty предоставляет услуги широкополосного WiMAX под собственным брендом WIPET. Компания имеет доступ к спектру шириной 150 МГц в диапазоне 3,4–3,6 ГГц.

Целью Liberty на рынке услуг WiMAX является доступ в Интернет. В настоящее время WiMAX-сети состоят из 31 сектора базовых станций Cisco 802.16e на 16 участках, охватывающих примерно 1,2 млн. жителей Панамы.

WiMAX-сеть компании Max Telecom

Компания Max Telecom является оператором мобильного WiMAX в Болгарии. Она предоставляет широкий спектр телекоммуникационных услуг, в том числе: мобильный и стационарный доступ в Интернет, VPN, голосовую связь, видео и IPTV.

Max Telecom стала первым оператором в Болгарии, предлагающим коммерческие решения для мобильного широкополосного доступа с использованием технологии WiMAX. Компания имеет лицензии на спектр частот шириной 42 МГц в диапазоне 3,5 ГГц. К середине 2008 года компания установила около 180 базовых

станций WiMAX в густонаселенных городах, что помогло привлечь более 30% населения. В первые четыре месяца работы Max Telecom подключила 4000 клиентов, а к концу 2008 года абонентская база достигла 13 тыс.

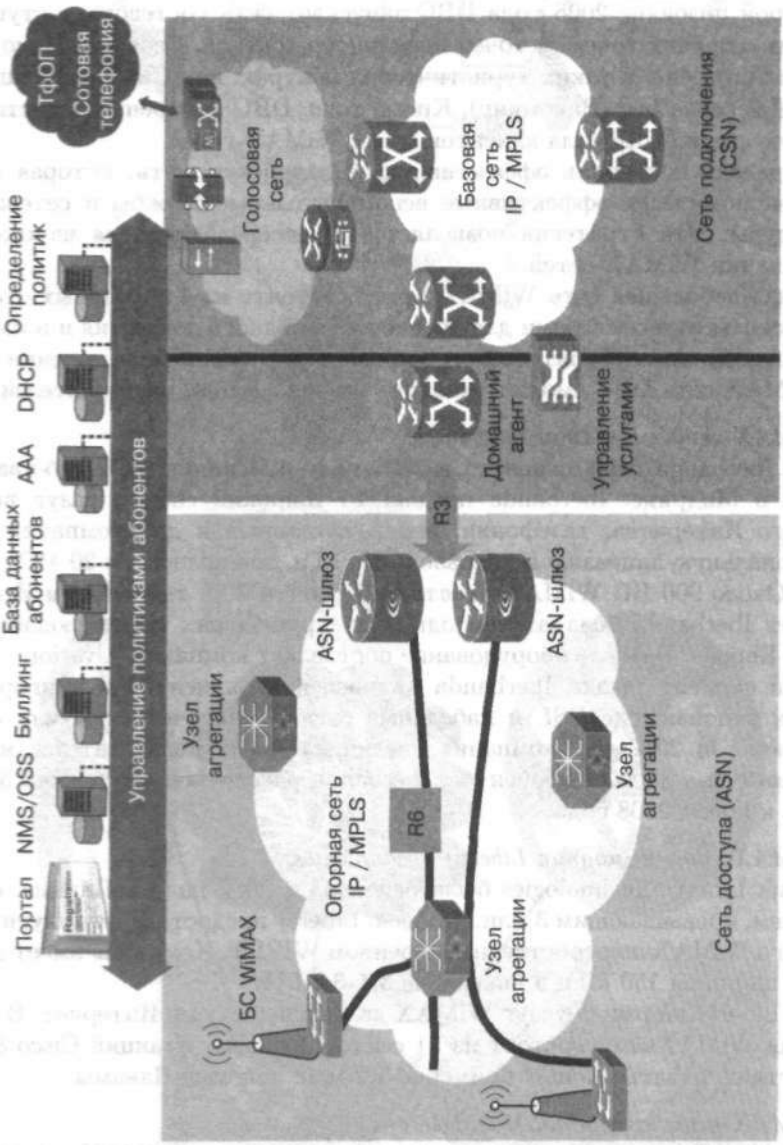


Рис. 10.4. Архитектура сети WiMAX компании Max Telecom

Max Telecom предлагает широкий спектр услуг для частных и бизнес-клиентов. Ежемесячная плата для частных лиц в среднем составляет 19 евро за доступ в Интернет и 6 евро — за услуги голосовой связи.

Базовые станции в сети Max Telecom связаны каналами «точка-точка» с узлами агрегации (посредством волоконно-оптических или радиорелейных линий) (рис. 10.4). И уже эти устройства сопрягаются с ASN-шлюзами посредством

оптоволоконной сети, охватывающей всю страну. Max Telecom планирует развернуть около 1500 базовых станций для повсеместного покрытия сети WiMAX.

WiMax-сеть компании Packet One Networks

Packet One Networks (P1) — это телекоммуникационная компания, предоставляющая беспроводные услуги по всей территории Малайзии. Зарегистрирована в феврале 2002 года как филиал корпорации Green Packet. Ей принадлежит лицензия на полосу шириной 30 МГц в диапазоне 2,3 ГГц. P1 начала строить WiMAX-сеть в марте 2007 года. В январе 2008 года компания объявила о контракте с Alcatel-Lucent на поставку оборудования WiMAX, на сумму порядка 71 млн. долл. К августу 2008 года P1 завершила пилотный проект и стала первым WiMAX-оператором в Малайзии, запустившим сеть в коммерческую эксплуатацию. Коммерческий запуск был произведен в отдельных районах и вокруг Куала-Лумпура, и к концу 2008 года сеть протянулась до южного региона Джохора. P1 утверждает, что за первую половину 2009 года она охватит сетью 30% территории страны, 40% — к 2010 году и 60% — к 2012. Ожидается, что суммарные инвестиции в WiMAX в течение следующих пяти лет составят порядка 323 млн. долл. В ближайшие 10 лет P1 собирается покрыть сетью 100% территории Малайзии.

Packet One предлагает одинаковые условия подключения как частных, так и бизнес-клиентов. Пакет услуг со скоростью 1,2 Мбит/с стоит 51 долл. в месяц, а 2,4 Мбит/с — 109 долл. Однако с января 2009 года цены на эти пакеты упали до 32 и 74 долл., соответственно, при заключении договора на год.

WiMax-сеть компании TransTelecom

TransTelecom Bulgaria — еще один WiMAX-оператор в Болгарии, фирма принадлежит болгарской нефтяной компании Petrol Holding AD. TransTelecom Bulgaria работает под торговой маркой 1one в полосе шириной 42 МГц в диапазоне 3,5 ГГц. Компания приступила к испытаниям сети в 2007 году, используя WiMAX-оборудование Alvarion (фиксированный доступ, 18 четырехсекторных БС). После успешных испытаний TransTelecom в качестве поставщика оборудования выбрала китайскую компанию Huawei. В середине 2008 года состоялся коммерческий запуск сети в составе 120 базовых станций. TransTelecom продолжала развертывать мобильный WiMAX на оборудовании Huawei в городах, в то время как оборудование для фиксированного доступа от Alvarion устанавливается в сельских районах.

В настоящее время 1one предлагает стандартный пакет услуг WiMAX 802.16e и планирует предложить номадический широкополосный беспроводной доступ во втором квартале 2009 года. Основной пакет услуг стоит 12,5 евро в месяц за канал 2 Мбит/с. Канал в 4 Мбит/с обойдется в 14,50 евро/мес. при заключении годового контракта.

В связи с предстоящим внедрением номадического широкополосного доступа TransTelecom продолжает участвовать в испытаниях с несколькими поставщиками оборудования. Она планирует отказаться от использования PCMCIA-карт в пользу USB-модемов, таких, как Seowon.

WiMax-сеть компании WiMAX Telecom

Основанная в 2004 году, компания WiMAX Telecom является одним из ведущих международных операторов беспроводного доступа в Европе. Ее штаб-квартира

расположена в Швейцарии, а филиалы действуют в Австрии, Германии (под названием Inquam Broadband), Хорватии и Словакии. Во всех этих пяти странах компания обладает спектральной полосой шириной как минимум 42 МГц в диапазоне 3,5 ГГц. Развертывание сетей WiMAX началось в Австрии и Словакии, а в июне 2008 года и в Хорватии. В разработке находятся сети Германии и Швейцарии. Кроме того, WiMAX Telecom активно проектирует сети для других стран Центральной и Восточной Европы.

Сегодня сети WiMAX Telecom включают в общей сложности 140 базовых станций WiMAX и охватывают территорию, на которой проживает около 500 тыс. человек в Австрии и около 1 млн. — в Словакии. Имея более 13 тыс. абонентов, WiMAX Telecom активно развивает рынки ШБД этих стран.

В австрийской сети WiMAX Telecom базовые станции WiMAX соединяются преимущественно через Ethernet/IP-каналы (радиорелейные). Региональные и центральные точки присутствия соединяются посредством арендуемых линий Ethernet, оптических волокон, а иногда и через высокоскоростные радиорелейные каналы. Ядро сети построено по технологии MPLS на основе маршрутизаторов и коммутаторов Cisco. В сети обеспечивается поддержка QoS, что позволяет реализовывать сервисы телефонии.

WiMAX Telecom имеет четыре центра обработки данных: два в Вене, в Братиславе, и в Загребе. Все они соединены линиями высокой пропускной способности или оптоволоконными линиями. Все основные поставщики услуг, а также службы управления операциями (OSS) и бизнесом (BSS) размещаются в Вене.

Потенциал сети WiMAX в Австрии — 30 тыс. абонентов, в Словакии — 40 тыс. абонентов. В каждой из этих стран компания WiMAX Telecom обладает полосой 56 МГц, что позволяет легко нарастить емкость сетей до нескольких сотен тысяч абонентов путем простого добавления новых каналов передачи информации и секторов к существующим базовым станциям.

Проникновение мобильной широкополосной связи составляет уже 8% в Австрии и около 1% в Словакии и Хорватии. WiMAX Telecom планирует внедрять услуги мобильного доступа с 2009 года, начиная с Хорватии. Ближайшие планы WiMAX Telecom по расширению сетей услуг включают установку 350 базовых станций 802.16e-2005 для обеспечения стационарных и мобильных услуг в Хорватии в 2009 году, внедрение мобильных сетей WiMAX во всех странах, где WiMAX Telecom обладает лицензией на частотный ресурс, поиск возможностей для получения частотных лицензий в других странах Центральной и Восточной Европы.

WiMAX-сеть компании Wateen Telecom

Пакистанская компания Wateen Telecom — это результат последних инвестиций консорциума Abu Dhabi Group, на счету которого удачный запуск компании Warid Telecom, крупнейшего GSM-оператора в Пакистане. Шатб-квартира Wateen Telecom располагается в Лахоре. Wateen Telecom с успехом развернула одну из крупнейших в мире национальных сетей WiMAX (возможно, первую в мире сеть IEEE 802.16e). Сеть строится на основе оборудования компании Motorola (WiMAX-платформа wi4 и оборудование для сервисных сетей с мультисервисной архитектурой IMS). Она включает 842 четырехсекторные БС в 22 городах, покрывает 20% территории Пакистана, на которой проживает 164 млн. человек. После начала коммерческой эксплуатации в декабре 2007 года сеть обслуживает

52 тыс. абонентов. Сеть расширяется, в 2009 году число БС планируют довести до 1300, а в более отдаленной перспективе покрыть сетью еще 70 городов. Месячный доступ в Интернет (канал порядка 1 Мбит/с) стоит от 6,3 до 1,1 долл.

Компания Wateen Telecom на всей территории страны обладает спектром в диапазоне 3,5–3,6 ГГц. В пяти регионах страны ширина полосы составляет 42 МГц, еще в девяти — 21 МГц. Сеть WiMAX использует каналы на принадлежащей компании Wateen Telecom оптоволоконной сети национального масштаба (рис. 10.5). По соглашению с компанией Warid, Wateen Telecom может использовать принадлежащую ей инфраструктуру GSM-сети, что существенно упрощает и ускоряет развертывание WiMAX-сети.

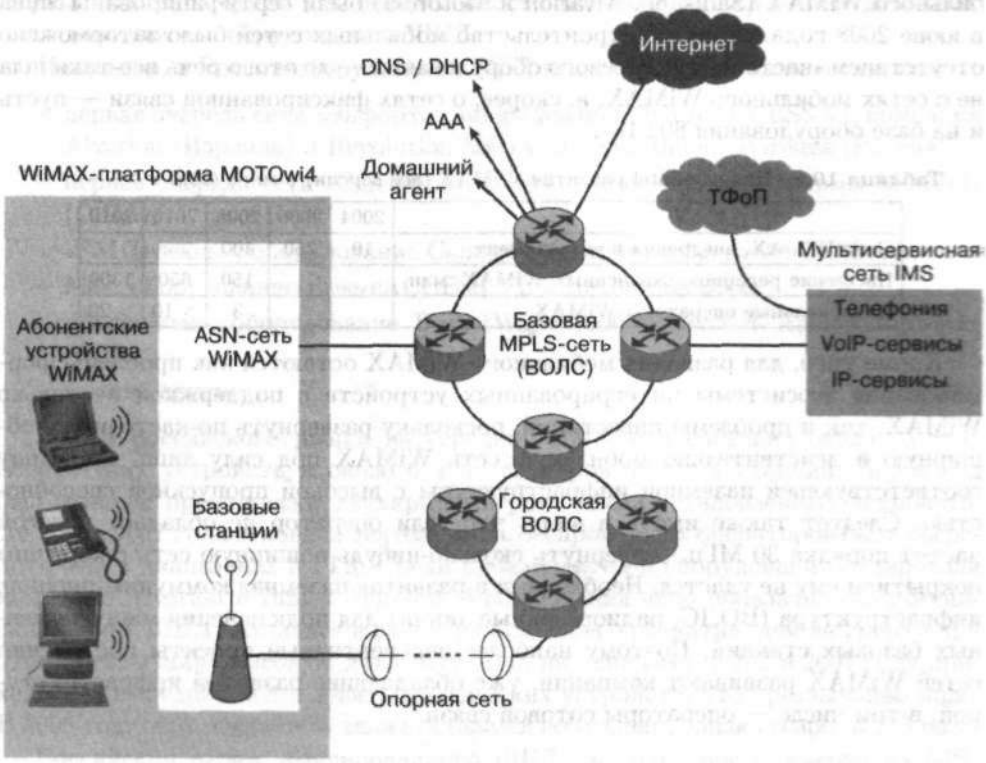


Рис. 10.5. Архитектура сети WiMAX компании Wateen Telecom

Wateen Telecom, являясь собственником глобальной беспроводной и оптоволоконной сети, создает в Пакистане NGN-сеть на основе технологии MPLS, с поддержкой мультисервисной архитектуры IMS. Действительно, этой компании в стране принадлежат свыше 5 тыс. км волоконно-оптических кабелей (24 оптических волокна G.652/G.655), охватывающих 71 город. Сеть имеет топологию в виде пяти полностью резервированных колец, реализована технология DWDM, потенциально масштабируемая до 160 несущих. В городах реализована оптическая сеть доступа на основе кабелей с 96 парами оптических волокон (G.652/G.655). Во всех крупных городах — кольца Ethernet с поддержкой FTTC.

Потенциал Пакистана огромен. В июне 2008 года в стране насчитывалось порядка 5 млн. пользователей Интернета, и 95% из них подключались посредством

телефонных модемов (dial-up). Поэтому в Пакистане так активно развиваются компании, предоставляющие широкополосный доступ. И есть все шансы, что к 2010–2011 годам 2/3 всех широкополосных подключений будут беспроводными. Поэтому компания Wateen Telecom, обладая столь существенными возможностями, в том числе — в области интеграции сервисов, обоснованно рассчитывает через пять лет обладать абонентской базой широкополосного доступа в 1,5 млн. пользователей, и существенная доля в ней будет принадлежать абонентам мобильной WiMAX-сети.

Не подлежит сомнению, что сети WiMAX в ближайшие годы продолжат интенсивно развиваться (табл. 10.3). Кроме того, первые базовые станции мобильного WiMAX (Samsung, Alvarion и Motorola) были сертифицированы лишь в июне 2008 года. До этого строительство мобильных сетей было заторможено отсутствием «настоящего» сетевого оборудования — до этого речь все-таки шла не о сетях мобильного WiMAX, а, скорее, о сетях фиксированной связи — пусть и на базе оборудования 802.16e.

Таблица 10.3. Прогноз Intel развития WiMAX (все версии) в мире (2007 г.)

	2004	2006	2008	2010	2012
Сети WiMAX, внедрения и эксперимент	10	250	400		
Население регионов, охваченных WiMAX, млн.			150	650	1300
Национальные операторы WiMAX			3	> 10	> 20

Кроме того, для развития мобильного WiMAX остаются как проблемы формирования экосистемы интегрированных устройств с поддержкой не только WiMAX, так и проблемы инвестиций, поскольку развернуть по-настоящему обширную и действительно мобильную сеть WiMAX под силу лишь владельцу соответствующей наземной инфраструктуры с высокой пропускной способностью. Следует также иметь в виду, что если оператор не обладает полосой частот порядка 30 МГц, развернуть сколько-нибудь обширную сеть с хорошим покрытием ему не удастся. Необходима и развитая наземная коммуникационная инфраструктура (ВОЛС, радиорелейные линии) для подключения многочисленных базовых станций. Поэтому наиболее перспективные проекты построения сетей WiMAX развивают компании, уже обладающие развитой инфраструктурой, в том числе — операторы сотовой связи.

10.2. Развитие сетей WiMAX в России и СНГ

Авторы данной монографии давно принимают участие в различных измерениях рынка России и СНГ беспроводного широкополосного доступа и WiMAX. Из наиболее известных проектов такого рода отметим исследования «Беспроводная BWA география» (июнь, 2007 год) совместно с журналом «Информкурьер-Связь» и «Карта WiMAX» (ноябрь, 2008 год) совместно с журналом «Стандарт» (см. цв. вклейку). В первом проете было обследовано 156 операторов беспроводного доступа, а во втором — 50 самых крупных.

Сразу оговоримся, что в рамках исследования рассматривались все сети широкополосного доступа — как использующие сертифицированное оборудование WiMAX, так и не сертифицированное. В целом, все оборудование ШБД на российском рынке можно разделить на четыре большие группы:

- прошедшее сертификацию WiMAX-форумом;
- не входящее в список сертифицированного WiMAX-оборудования, но соответствующее какому-либо сертификационному профилю WiMAX-форума;
- соответствующее стандарту IEEE 802.16, но не соответствующее ни одному сертификационному профилю WiMAX-форума (например, работающее в диапазоне 5 ГГц, в России характерный пример — продукция компании BreezeMAX 5,2 ГГц от Alvarion);
- оборудование ШБД с нестандартными протоколами работы (например, оборудование компании Infinet Wireless).

Первые два типа будем для определенности называть WiMAX-оборудованием, последние два — оборудованием ШБД.

Например, в России оборудование ШБД используется в таких сетях, как:

- первая очередь сети «Энфорта», оборудование BreezeACCESS VL компании Alvarion (Израиль) и Revolution 5000 компании Infinet Wireless (Россия);
- первая очередь сети «Синтерра», оборудование NextNET фирмы Motorola (США),
- «РМ Телеком», оборудование Canopy (Motorola) и BreezeACCESS VL;
- Flex, оборудование BreezeACCESS VL и Revolution 5000;
- «Инфосети», оборудование PacketWave 1000 компании Aperto Networks (США);
- многие другие.

Объем российского рынка на середину 2007 года составил около 45 тыс. абонентских устройств и около 6 тыс. секторов базовых станций. Через год был отмечен практически двухкратный рост числа установленных устройств. До конца 2007 года основная деятельность беспроводных операторов была сосредоточена в диапазонах 5 ГГц. Среди производителей оборудования лидировали компании Alvarion и Infinet Wireless. Исследования явно показали, что российский рынок ШБД находился в самом начале своего развития, демонстрируя при этом чрезвычайно высокую динамику. Так, на один сектор БС в 2007 году приходилось в среднем всего девять абонентских устройств, что чрезвычайно мало. В 2008 году этот показатель также оставался небольшим, однако возрос в 2–3 раза.

Российский рынок фиксированного ШБД ежегодно рос примерно на 50%. Фиксированные сети ШБД уже развернуты в плотно населенных регионах — Екатеринбургской, Челябинской, Тюменской, Нижегородской областях, в Алтайском крае. По объемам этого рынка лидируют Приволжский округ, за ним следует Москва и Санкт-Петербург, а по темпам роста Южный (615%), Центральный (463%) и Северо-Западный (409%) федеральные округа. Основная клиентская база фиксированного ШБД — корпоративные клиенты, а также домохозяйства с доходами выше среднего, где есть проблемы с подключением провода или кабеля.

Реально в России сотни операторов занимаются фиксированным широкополосным беспроводным доступом. Среди них лидировала компания «Энфорта», занимавшая от 1/5 до 1/4 российского рынка ШБД. По мере развития рынка отмирают мелкие операторы, объединяются средние. И тех, и других поглощают крупные компании.

Рынок мобильного WiMAX в конце 2008 года хотя и находился в зачаточном состоянии, но к середине 2009 года составил уже около 4500 секторов, что сравнимо со всеми сетями фиксированного доступа за все время развития. По данным компании J'son & Partners, рынок услуг фиксированного WiMAX в домашнем сегменте (около 40 тыс. абонентов) составляет порядка 25 млн. долл. Рынок абонентских устройств может составить несколько сот тысяч устройств в год, что существенно больше всего остального рынка фиксированного доступа. Пока реально сети мобильного WiMAX развивают всего несколько операторов (компании «Скартел», «Комстар», «Синтерра» в диапазоне 2,5 ГГц и некоторые в диапазоне 3,5 ГГц (например, компания Lythgoe). Ожидается проведение конкурсов на частотные присвоения мобильного WiMAX в диапазонах 2,3 и 2,5 ГГц в 2009 году.

Как показали исследования, среди крупнейших операторов ШБД в РФ выделяются несколько компаний, претендующие на статус федеральных операторов — «Энфорта» (ООО «Престиж-Интернет»), Unitline («Медиасети»), «Престиж» и др., но они работают с «фиксированным» WiMAX (3,5 ГГц). Рассмотрим некоторые наиболее значимые проекты.

10.2.1. Сети фиксированного доступа

Сеть «Энфорта»

«Энфорта» (головная компания — ООО «Престиж-Интернет», голландский холдинг Enforta) — национальный оператор связи, является безусловным лидером российского рынка фиксированного широкополосного беспроводного доступа в Интернет (рыночная доля — 24%). По версии Cnews, «Энфорта» входит в 30 крупнейших российских телекоммуникационных компаний по величине выручки. Это первый оператор беспроводной связи, попавший в данный рейтинг.

На март 2009 года сеть компаний развернута в 68 городах России (рис. 10.6): Ангарске, Армавире, Артеме, Астрахани, Барнауле, Бийске, Биробиджане, Братске, Ванино, Владивостоке, Волгограде, Воронеже, Дзержинске, Димитровграде, Екатеринбурге, Златоусте, Ижевске, Иркутске, Каменск-Уральском, Кемерово, Комсомольске-на-Амуре, Краснодаре, Красноярске, Липецке, Магнитогорске, Миассе, Москве и Московской области, Нижнем Новгороде, Нижнем Тагиле, Находке, Новокузнецке, Новокуйбышевске, Новомосковске, Новороссийске, Новосибирске, Новотроицке, Новочеркасске, Омске, Оренбурге, Орске, Пензе, Перми, Прокопьевске, Ростове-на-Дону, Рубцовске, Рыбинске, Рязани, Самаре, Санкт-Петербурге, Саратове, Сочи, Стерлитамаке, Сызрани, Таганроге, Тольятти, Томске, Туле, Тюмени, Улан-Удэ, Ульяновске, Уссурийске, Уфе, Челябинске, Хабаровске, Шахтах, Южно-Сахалинске, Ярославле.

Таким образом, на март 2009 года сеть беспроводного доступа компании покрывает практически все областные центры и города-«миллионники», где получен необходимый частотный ресурс в диапазоне 3,5 и 5,2 ГГц. В основном — в силу особенностей российской политики в области распределения частотного ресурса — компания работает в диапазоне 5,15–5,35 ГГц. В сети используется оборудование трех типов — WiMAX (SMAX от Airspan и BreezeMAX от Alvarion в диапазоне 3,5 ГГц), оборудование стандарта IEEEE 802.16 (BreezeMAX в диапазоне 5,2 ГГц) и пропроетарное оборудование ШБД (BreezeACCESS от

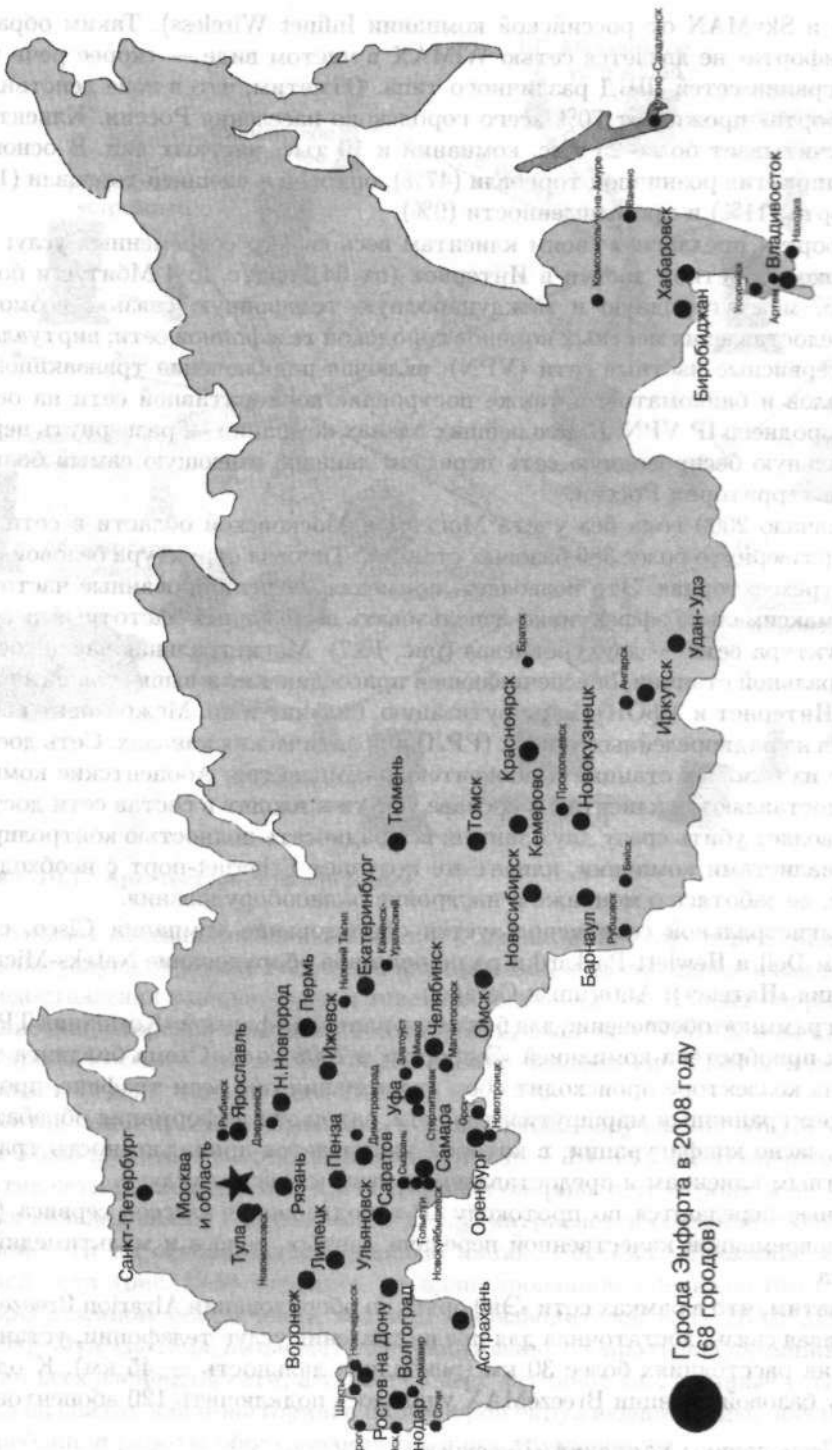


Рис. 10.6. Сеть фиксированного ШПД «Энфорта»

Alvarion и SkyMAN от российской компании Infinet Wireless). Таким образом, сеть «Энфорта» не является сетью WiMAX в чистом виде — скорее речь идет об интеграции сетей ШБД различного типа. Отметим, что в зоне действия сети «Энфорта» проживает 70% всего городского населения России. Клиентская база насчитывает более 25 тыс. компаний и 10 тыс. частных лиц. В основном это предприятия розничной торговли (47%), оптовой и внешней торговли (17%), транспорта (11%) и промышленности (9%).

«Энфорта» предлагает своим клиентам весь спектр современных услуг связи: высокоскоростной доступ в Интернет (от 64 кбит/с до 4 Мбит/с и более); местную, междугородную и международную телефонную связь с возможностью предоставления местных номеров городской телефонной сети; виртуальные мультисервисные частные сети (VPN), включая подключение транзакционных терминалов и банкоматов; а также построение корпоративной сети на основе междугороднего IP VPN. В дальнейших планах компании — развернуть первую национальную беспроводную сеть передачи данных, имеющую самый большой охват на территории России.

На начало 2009 года без учета Москвы и Московской области в сети «Энфорта» развернуто более 386 базовых станций. Типовая структура базовой станции — трехсекторная. Это позволяет, применяя масштабированные частотные планы, максимально эффективно использовать выделенный частотный ресурс.

Структура сети — двухуровневая (рис. 10.7). Магистральная часть состоит из центральной станции, обеспечивающей присоединение к вышестоящим операторам (Интернет и ТфОП), маршрутизацию, биллинг и пр. Межбазовые каналы строятся на радиорелейных линиях (РРЛ) или оптических каналах. Сеть доступа состоит из базовых станций и абонентских комплектов. Абонентские комплекты предоставляются клиентам в составе услуги и входят в состав сети доступа, что позволяет убить сразу двух зайцев: вся радиосеть полностью контролируется специалистами компании, клиент же получает Ethernet-порт с необходимой услугой, не заботясь о монтаже и настройке радиооборудования.

В магистральной сети используется оборудование компаний Cisco, серверы фирм Dell и Hewlett-Packard и радиорелейное оборудование Nateks-Microlink (компания «Натекс»), Anterum и Ceragon.

Программное обеспечение для биллинга написано фирмой «Компания ТРОН», которая приобретена компанией «Энфорта» в 2006 году. Схема биллинга типовая — на коллекторе происходит сбор информации обо всем трафике, проходящем через граничный маршрутизатор сети, затем эта информация обрабатывается согласно конфигурации, в которой указывается принадлежность трафика конкретным клиентам и предоставляемые этим клиентам услуги.

Данные передаются по протоколу IP с поддержкой классов сервиса (QoS) для одновременной качественной передачи данных, голоса и мультимедийного трафика.

Отметим, что в рамках сети «Энфорта» на оборудовании Alvarion BreezeMAX устойчивая связь, достаточная для предоставления услуг телефонии, устанавливалась на расстояниях более 30 км (рекордная дальность — 45 км). К одному сектору базовой станции BreezeMAX удавалось подключить 120 абонентов.

Сеть группы компаний «Престиж»

Группа компаний проекта «Престиж» (головная компания — ООО «Престиж») была основана в 2006 году швейцарским инвестиционным фондом ABT Group.

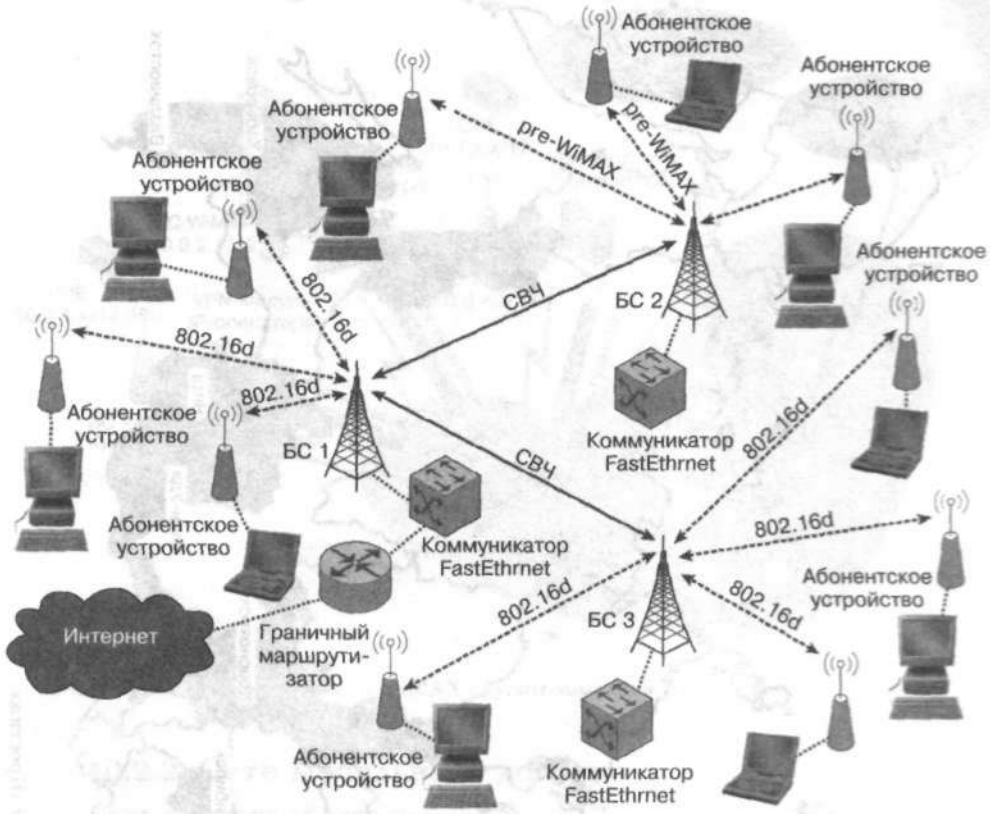


Рис. 10.7. Архитектура сети «Энфорта»

Она обладает лицензиями Минкомсвязи РФ на оказание услуг передачи данных, исключая услуги передачи голосовой информации, а также лицензиями на услуги предоставления каналов связи и телематические услуги. Проект направлен на создание операторской компании и сети фиксированного WiMAX в 33 крупнейших города России с населением более 21 млн. человек, где в целом в регионах проживает 94 млн. человек (рис. 10.8).

В сетях связи проекта «Престиж» используется оборудование WayMAX компании Nokia Siemens Network. Диапазон частот — 3,4–3,6 ГГц. На первом этапе развития сети региональная и магистральная опорная сеть строится на арендованных каналах связи. В дальнейшем предусматривается создание собственной опорной сети на основе оптоволоконных линий. Система управления и мониторинга сети (рис. 10.9) строится на адаптированном специалистами компании программном обеспечении ведущих производителей (HP, IBM, Microsoft, Siemens). Эта система позволяет централизовано собирать информацию о состоянии всех элементов сети, а также со всех датчиков БС, которые устанавливаются на сайтах для мониторинга параметров окружающей среды, обеспечения бесперебойной работы оборудования БС и его безопасности.

Рис. 10.8. Сеть фиксированного WiMAX группы компаний «Престиж»

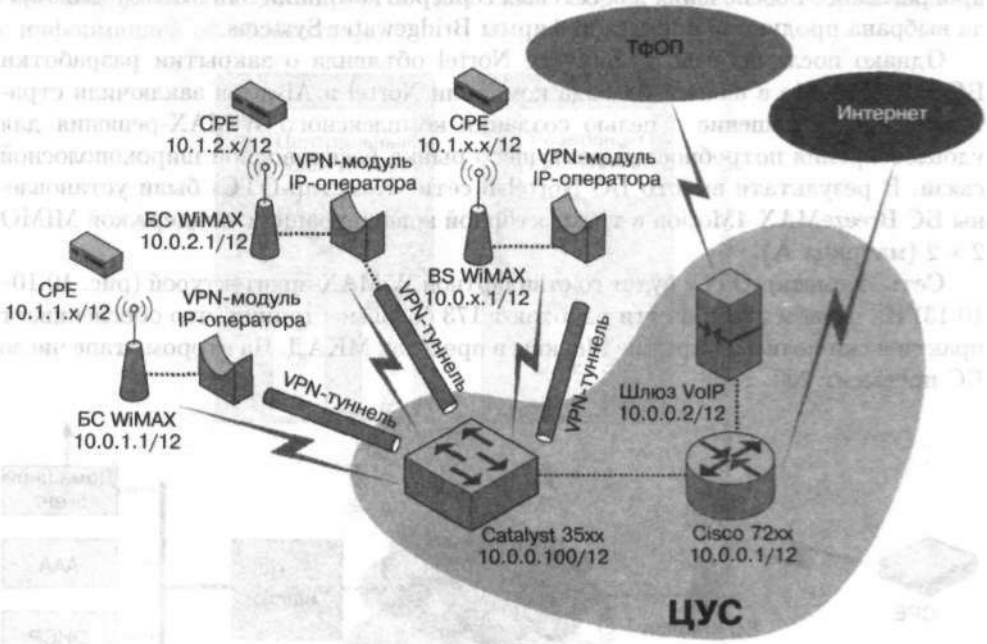


Рис. 10.9. Архитектура сети WiMAX группы компаний «Престиж»

10.2.2. Сети мобильного доступа

Сеть компании «Комстар ОТС»

Компания «Комстар — Объединенные ТелеСистемы» («Комстар-ОТС») сделала ставку на технологию WiMAX как на конвергентную услугу, позволяющую пользователю иметь доступ к сети широкополосного доступа через WiMAX. Будучи крупным оператором проводного широкополосного доступа, компания решила, что вне дома и в офисе ее абоненты также должны иметь возможность беспроводного подключения к сети. В декабре 2007 года «Комстар-ОТС» и корпорация Intel подписали соглашение о стратегическом сотрудничестве по развитию технологии мобильного WiMAX в России. В соответствии с соглашением, на первом этапе усилия будут сосредоточены на московском регионе. «Комстар-ОТС» планирует построить и запустить в коммерческую эксплуатацию сеть мобильного WiMAX в радиочастотном диапазоне 2,5–2,7 ГГц, охватывающую всю территорию Москвы. Со своей стороны, корпорация Intel будет способствовать расширению поставок клиентских устройств с интегрированной поддержкой WiMAX. Основная услуга на первом этапе — доступ в Интернет, в дальнейшем набор услуг будет расширен (VoIP, мультимедиа в реальном времени, организация VPN-каналов для корпоративных пользователей и др. услуги, характерные для мультисервисных сетей).

В начале 2008 года был проведен конкурс среди производителей оборудования. В нем приняли участие такие известные компании, как Alcatel-Lucent, Alvarion, Motorola, Samsung, Nortel Networks, Cisco/Navini. Но победила компания Nortel Networks (базовая станция BTS5020, ASN-шлюз ASG 5100). В качестве

программного обеспечения для сетевых серверов компании Sun Microsystems была выбрана продукция канадской фирмы Bridgewater Systems.

Однако после победы в конкурсе Nortel объявила о закрытии разработки БС BTS5020. Но в июне 2008 года компании Nortel и Alvarion заключили стратегическое соглашение с целью создания комплексного WiMAX-решения для удовлетворения потребностей растущего рынка беспроводной широкополосной связи. В результате вместо БС Nortel в сети «Комстар-ОТС» были установлены БС BreezeMAX 4Motion в трехсекторной конфигурации с поддержкой MIMO 2×2 (матрица **A**).

Сеть «Комстар-ОТС» будет со стандартной WiMAX-архитектурой (рис. 10.10–10.13) На первом этапе в сети работают 173 базовые станции, что обеспечивает практически полное покрытие Москвы в пределах МКАД. На втором этапе число БС превысит 200.

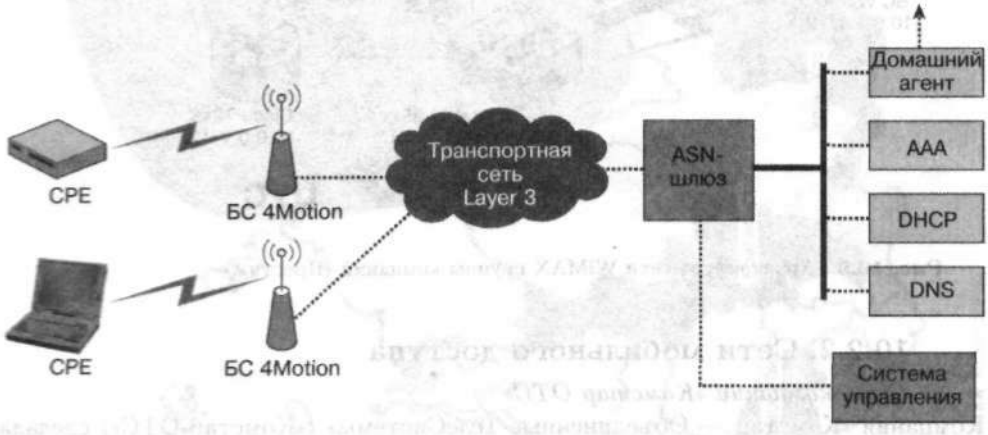


Рис. 10.10. Архитектура мобильной WiMAX-сети компании «Комстар-ОТС»

Компания «Комстар-ОТС» имеет разрешение на использование трех частотных номиналов с шириной полосы 10 МГц в диапазоне 2,5–2,7 ГГц (центральные частоты 2540, 2550 и 2560 ГГц). Это позволило ей реализовать схему переиспользования частотного спектра (reuse 3). В этом случае все сектора, ориентированные на север, имеют частоту F1, на юго-восток — F2 и на юго-запад — F3. Это позволяет при небольшом частотном ресурсе покрыть территорию всей Москвы.

Один сектор БС поддерживает до 512 абонентских устройств. Считая, что в каждой БС три сектора, получаем для 173 БС теоретический максимум для первого этапа развертывания сети 265 728 абонентов. Но число БС — не единственное ограничение абонентской емкости сети. Не менее важный фактор — это производительность ASN-шлюзов и серверов авторизации AAA. Поэтому на первом этапе предполагается ограничить число абонентов на уровне 150 тыс. В дальнейшем потребуется увеличивать емкость ASN-шлюзов и добавлять лицензии на AAA-серверы.

В качестве типового решения (в силу юридической близости с компанией МГТС) «Комстар-ОТС» предполагает устанавливать БС преимущественно на АТС сети МГТС (см. рис. 10.12). Вся сеть спроектирована с довольно высокой степенью резервирования. Ядро сети, включающее ASN-шлюзы и сетевые

серверы, полностью дублируется (как само оборудование, так и базы данных с информацией об абонентах).

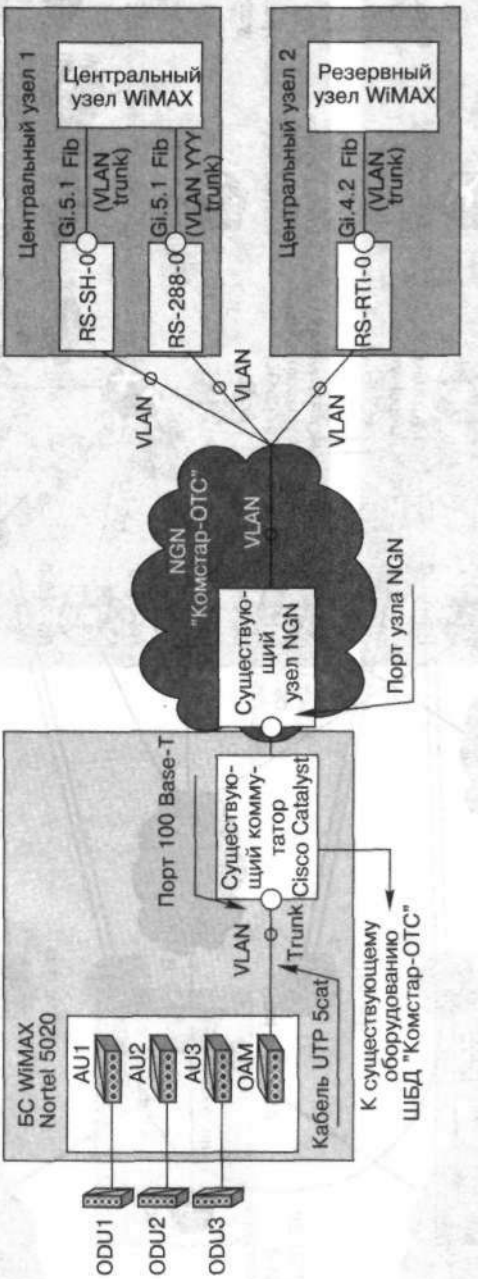


Рис. 10.11. Типовая схема включения БС, размещенных на существующих площадках ШБД с использованием транспортных каналов NGN в сети «Комстар-ОТС»

Для биллинга используется стандартная для сети «Комстар-ОТС» система расчета, но при этом используются возможности AAA-сервера по подготовке такой информации, как число вхождений в сеть, предоставленный сервис, скачанная информация и др. Это позволяет корректно тарифицировать каждого абонента.

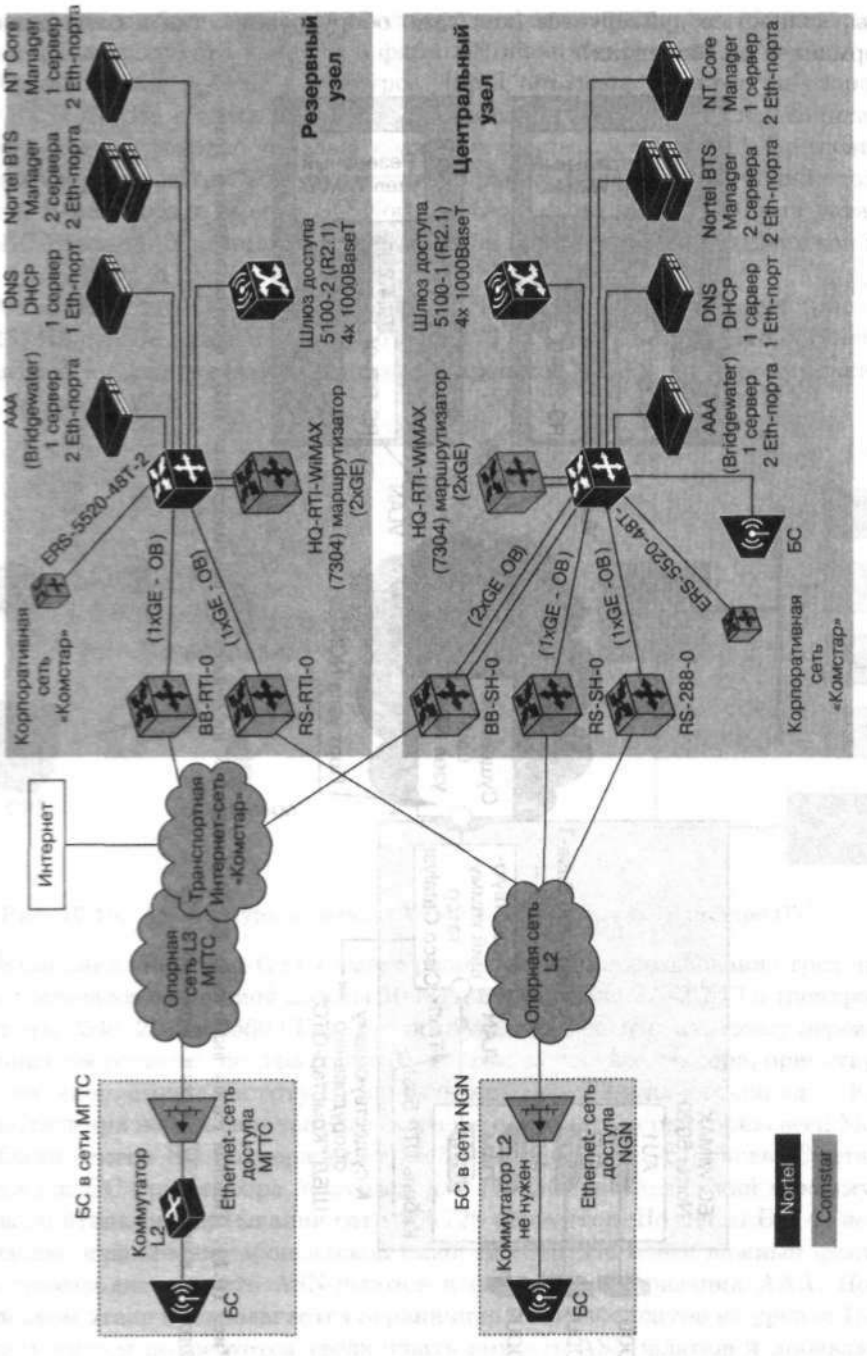


Рис. 10.12. Физическая структура WiMAX-сети «Комстар-ОТС»

В процессе тестовой эксплуатации было выявлено, что в одном абонентском канале пиковая пропускная способность может достигать 15 Мбит/с для нисходящего трафика и до 7 Мбит/с — для восходящего. При этом в нисходящем канале используется модуляция 64-QAM при скорости кодирования 5/6, в восходящем — 16-QAM со скоростью кодирования 3/4. Таким образом, спектральная

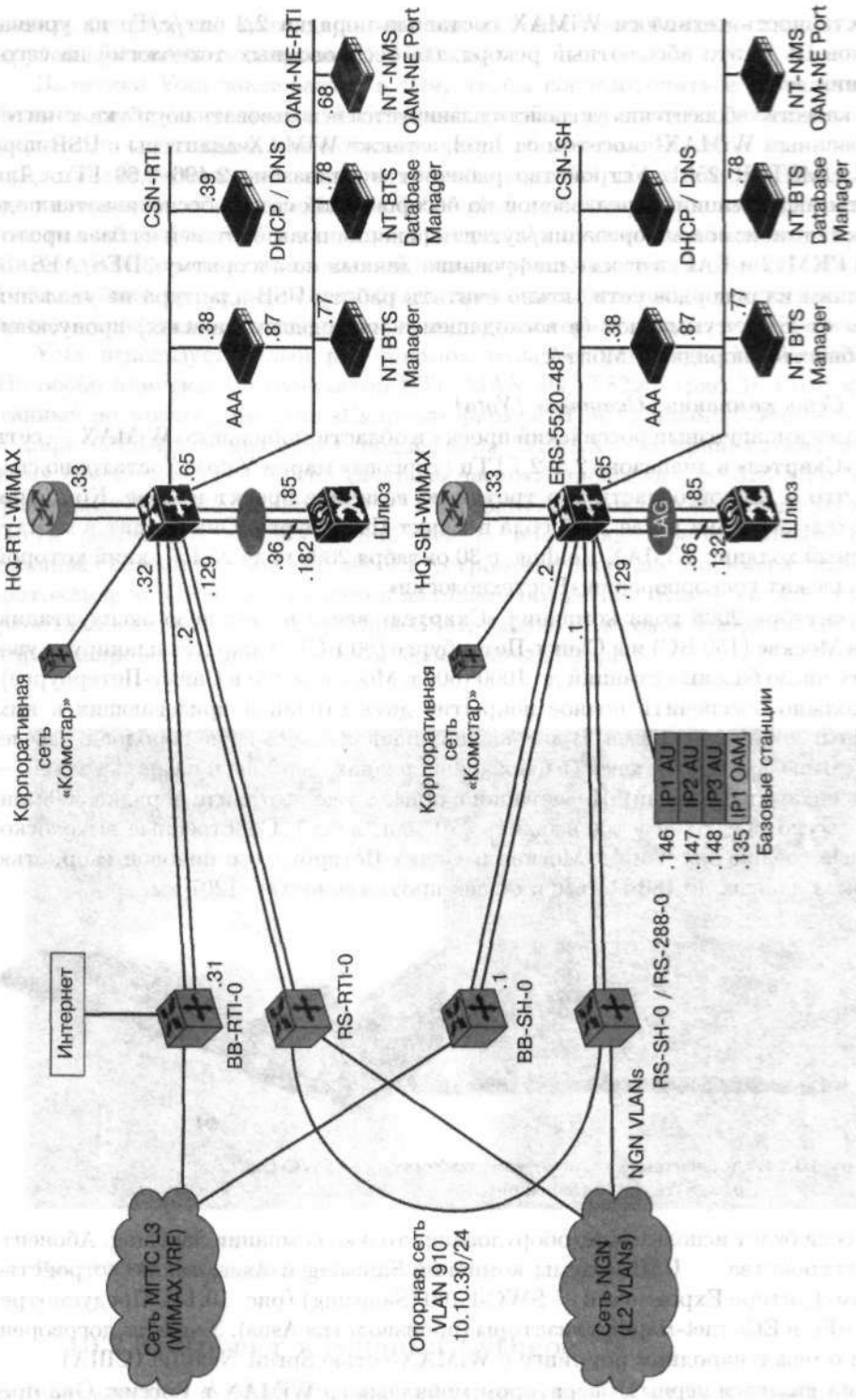


Рис. 10.13. Логическая схема WiMAX-сети «Комстар-ОТС»

эффективность технологи WiMAX составила порядка 2,2 бит/с/Гц на уровне потоков IP — это абсолютный рекорд для беспроводных технологий на сегодняшний день!

В качестве абонентских устройств планируется использовать ноутбуки с интегрированным WiMAX-чипсетом от Intel, а также WiMAX-адаптеры с USB-портом COMSTAR 2501. Устройство работает в диапазоне 2,496–2,69 ГГц. Для защиты информации, передаваемой по беспроводной сети, обеспечивается поддержка механизмов авторизации/аутентификации пользователей на базе протоколов PKMv2 и EAP, а также шифрования данных по алгоритму 3DES/AES.

Одним из рекордов сети можно считать работу USB-адаптера на удалении 5,5 км от БС с суммарной (в восходящем и нисходящем каналах) пропускной способностью порядка 2 Мбит/с.

Сеть компании «Скартел» (Yota)

Наиболее амбициозный российский проект в области мобильного WiMAX — сеть ООО «Скартел» в диапазоне 2,5–2,7 ГГц (торговая марка Yota). Достаточно сказать, что в данной области это третий по величине проект в мире. Компания «Скартел» основана в мае 2007 года в Санкт-Петербурге. Она входит в международный холдинг WiMAX Holding, с 30 октября 2008 года 25,1% акций которой принадлежат госкорпорации «Ростехнологии».

2 сентября 2008 года компания «Скартел» ввела в тестовую эксплуатацию сети в Москве (150 БС) и в Санкт-Петербурге (80 БС). «Скартел» планирует увеличить число базовых станций до 1000 (600 в Москве и 400 в Санкт-Петербурге), что должно обеспечить полное покрытие двух столиц и прилегающих к ним основных автомагистралей. В дальнейших планах — освоение городов с населением свыше 500 тыс. человек (в ближайших планах — в Уфе и в Сочи, а затем — в Краснодаре и в Казани). Инвестиции в проект уже составили порядка 300 млн. долл. (всего предполагается вложить 470 млн. долл.). Собственные высокоскоростные соединения Yota в Москве и Санкт-Петербурге с пиковой скоростью передачи данных до 180 Гбит/с и общей протяженностью 1207 км.

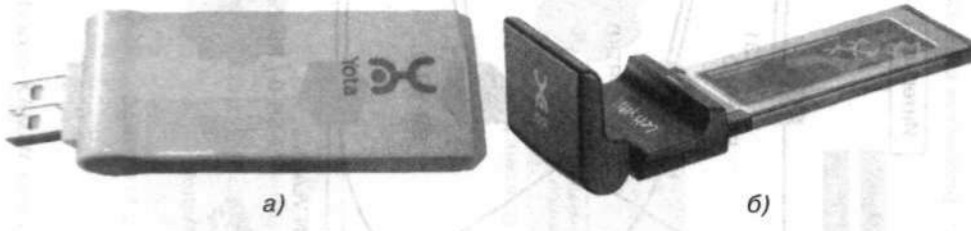


Рис. 10.14. Абонентские устройства сети Yota: *а)* — SWC-U200, *б)*— SWC-E100 (Samsung)

В сети будет использовано оборудование только компании Samsung. Абонентские устройства — USB-модемы компаний Samsung и Asus, также устройство в форм-факторе Express-card — SWC-E100 (Samsung) (рис. 10.14). Предусмотрены Wi-Fi- и Ethernet-маршрутизаторы (производства Asus). Уже есть договоренности о международном роуминге с WiMAX-сетью Sprint NextTel (США).

Yota является первым оператором мобилльного WiMAX в России. Она предоставляет доступ в Интернет на скорости до 10 Мбит/с для каждого пользо-

вателя, без разрыва сессии при переходе из зоны одной базовой станции в зону другой при скорости движения до 120 км/ч.

Политика Yota заключается в том, чтобы сосредоточиться на услугах мобильной связи в качестве своей основной деятельности, считая сети лишь транспортными средствами для доставки услуг абонентам. По этой причине Yota представляется более чем интернет-провайдером. В частности, владельцы сети позиционируют ее как «мобильный Google», для которого компанией разработан ряд приложений. Компания уже предлагает музыкальный каталог и услуги Интернет-ТВ для своих пользователей (оснащенных коммуникатором HTC MAX 4G). В ближайшее время абонентам будет предложен ряд новых услуг, таких, как видеозвонки, видеоконференции, расширенный сервис «Видео по запросу» и др.

Yota использует целый ряд пользовательских устройств Mobile WiMAX. Но особо отметим коммуникатор HTC MAX 4G (T8290) (рис. 10.15), разработанный по заказу компании «Скартел» фирмой HTC (Тайвань). Это — первый в мире GSM-коммуникатор с поддержкой WiMAX. Он принадлежит известному семейству HTC Touch. Тактовая частота процессора — 528 МГц, объем флеш-памяти — 512 Мбайт, ОЗУ — 256 Мбайт. Диагональ сенсорного экрана — 9,5 см, разрешение — 480×800 пикселей. На коммуникаторе установлена операционная система Windows Mobile 6.1, устройство оборудовано 5-мегапиксельной фотокамерой, GPS-навигатором и датчиком положения. Коммуникатор сможет работать в сети любого оператора GSM, однако голосовой трафик в сети Yota тарифицироваться не будет («свой собственный Skype»).

Рис. 10.15. Коммуникатор HTC MAX 4G

10.2.3. Проект компании Lythgoe

К числу крупных WiMAX-проектов на территории СНГ следует отнести компанию Lythgoe. Более 70% компании Lythgoe принадлежит американскому фонду

ICON PE. Он также владеет WiMAX-ресурсами во многих других странах, например, в Италии, Бангладеш, Мозамбике, Сомали. Компания Lythgoe владеет украинской компанией UHT, которой эксклюзивно принадлежит спектр 3,4–3,6 ГГц по всей Украине, а также около 30 МГц аналогичного спектра в России. Lythgoe будет работать под торговой маркой FreshTel.

Стратегическим приоритетом компании является создание оператора мобильного широкополосного доступа № 1 на территории Украины и оператора № 2 — в России (оператором № 1 они считают компанию «Скартел»). На Украине компания планирует быстрый захват рынка мобильного широкополосного доступа в Киеве и экспансию в города-«миллионники». В России Lythgoe намерена разворачивать WiMAX-сети в городах с высокими доходами населения и низким проникновением Интернета (Центральный и Южный федеральные округа), а также проводить спецпроекты с госструктурами и федеральными компаниями. Компания уже владеет развернутыми сетями стандарта IEEE 802.16-2004 в шести крупнейших городах Украины (на обрудовании Alvarion BrezeMAX). В России ею реализован ряд пилотных проектов.

В целом, российские сети WiMAX находятся в самом начале своего развития, и говорить о динамике рынка пока не слишком уместно. Отметим лишь, что появление сетей мобильного WiMAX снизит наблюдавшуюся до сих пор скорость роста сетей фиксированного ШБД на уровне 50% в год. По подсчетам аналитиков компании J'son & Partners, к 2010 году объем доходов фиксированного WiMAX в стране составит более 100 млн. долл. (до 160 тыс. абонентов). В 2008 году он был на уровне 25 млн. долл. (~ 40 тыс. абонентов), в 2009 — 60 млн. долл. (до 100 тыс. абонентов). О рынке мобильного WiMAX говорить совсем рано. Но можно указать, что если на 2008 год в России насчитывалось порядка тысячи абонентов таких сетей, то в 2009 году их число существенно превысит 100 тыс.

Безусловно, это очень скромные показатели по сравнению с числом абонентов сетей мобильной телефонной связи (190,8 млн. активных SIM-карт на 1 марта 2009 года, уровень проникновения — 131,4%, данные компании AC&M Consulting). Но, вне всякого сомнения, свою нишу на российском рынке технологии WiMAX займут прочно и в ближайшие годы будут активно развиваться.

ГЛАВА 11

СРАВНЕНИЯ И ВЫВОДЫ

11.1. Состояние и прогнозы рынка ШБД

Число пользователей беспроводного широкополосного доступа в мире к 2015 году достигнет 2,1 млрд. человек, утверждают эксперты аналитической компании Analysys Mason (рис. 11.1) [1]. К этому моменту совокупный годовой доход от услуг ШБД в мире составит 784 млрд. долл. (рис. 11.2) Лидирующие позиции по числу пользователей беспроводного широкополосного доступа в ближайшее время будет занимать технология HSPA. В конце мая 2009 года ею пользовались свыше 123 млн. абонентов (по данным UMTS Forum), а еще в конце 2008 года — лишь 61 млн. [1], что тогда составляло 88% всех абонентов сетей мобильного ШБД. Для сравнения, всего абонентов WCDMA в мае 2009 года было свыше 374 млн. К 2015 году число абонентов сетей HSPA возрастет до 1,1 млрд. человек, что составит 54% всех абонентов мобильного ШБД. Данные компании Ericsson аналогичны — 70% всех абонентов мобильного ШБД будут использовать HSPA (рис. 11.3) [2].

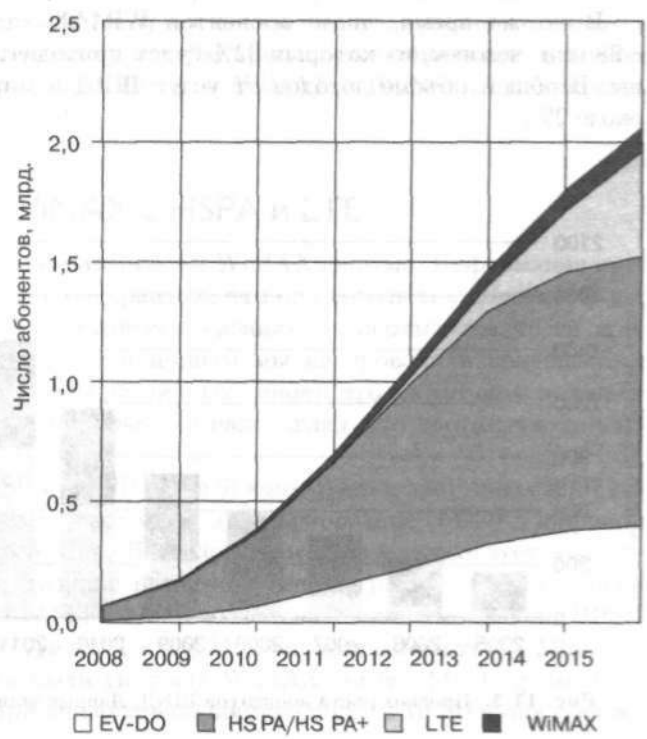


Рис. 11.1. Прогноз роста абонентской базы сетей ШБД. Данные компании Analysys Mason

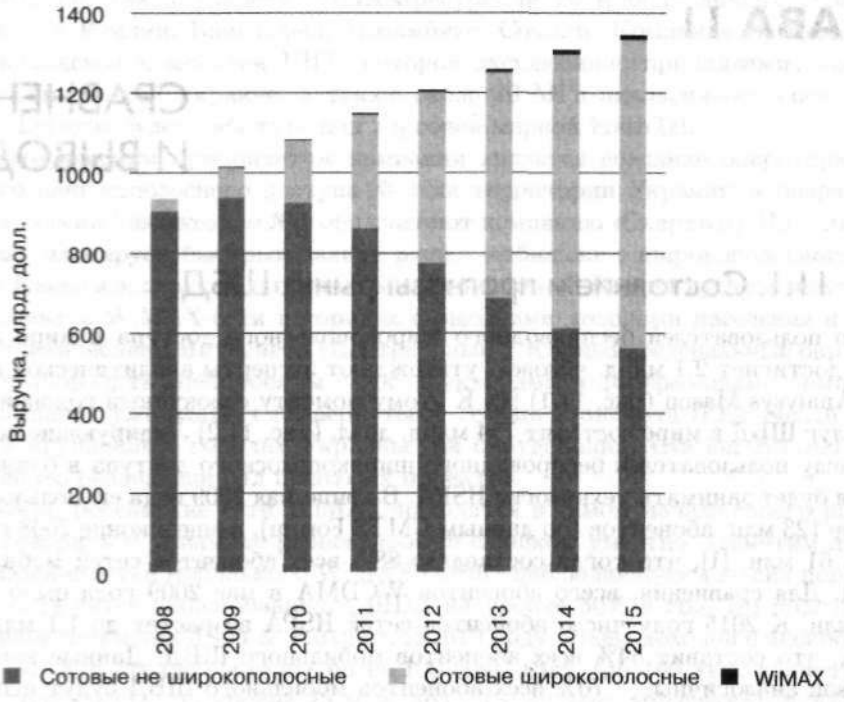


Рис. 11.2. Выручка от услуг беспроводной связи. Данные компании Analysys Mason

В то же время, число абонентов WiMAX-сетей к 2015 году достигнет ~98 млн. человек, из которых 92% будет приходиться на развивающиеся страны. В общем объеме доходов от услуг ШПД в мире WiMAX будет занимать около 2%.

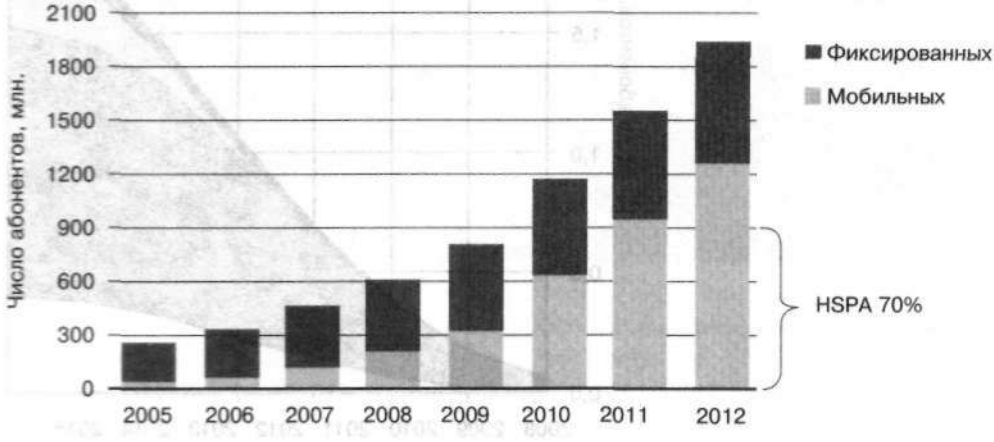


Рис. 11.3. Прогноз роста абонентов ШБД. Данные компании Ericsson

По данным Analysys Mason, к 2015 году общий доход операторов сетей связи стандарта LTE составит 194 млрд. долл. (~15% доходов мирового рынка услуг сотовой связи), а число их абонентов по всему миру превысит 440 млн. Ожидается, что развертывание коммерческих сетей LTE начнется не позже 2010 года, а широкое предоставление услуг на их основе — с 2011 года. Отсутствие развитой проводной инфраструктуры связи приведет к тому, что доля развивающихся регионов в общемировой абонентской базе беспроводного ШПД вырастет к 2015 году до 57% против 17% в 2008 году.

В Европе к 2014 году число абонентов мобильных сетей ШБД составит 148 млн. — порядка 58% всех абонентов ШБД против 8% в 2008 году. Прибыль от услуг мобильного ШБД составит в 2014 году 23 млрд. евро против 6 млрд. евро в 2008 году.

Кроме того, рано списывать со счетов технологии cdma2000 EV-DO. Мировая абонентская база операторов сетей EV-DO rev. A выросла во втором квартале 2008 года в восемь раз по сравнению с тем же периодом 2007 года (ABI Research). В марте 2009 года в мире было 474 млн. абонентов cdma2000 (по данным 3GPP2). На конец 2008 года их насчитывалось почти 456 млн. Из них свыше 112 млн. — абоненты сетей EV-DO [3]. Наибольший рост наблюдается на рынках США и Южной Кореи. По оценкам ABI Research, к 2013 году число абонентов сетей EV-DO rev.A превзойдет 54 млн. человек, а сетей rev.B — 25 млн.

Все это, казалось бы, говорит о том, что технологии WiMAX отводится нишевая роль. Но так ли однозначен этот вывод? Обратим внимание, что каждый кулик хвалит свое болото. И прогнозы с красивыми картинками часто показывают то, что хотят их авторы. Особенно это относится к сравнению таких якобы объективных показателей, как спектральная эффективность, скорость передачи данных и т. п. Удивительно, но и эти показатели разительно отличаются от того, к какому технологическому лагерю принадлежит аналитик. Сравним наиболее перспективные технологии ШБД — WiMAX и HSPA с LTE, опираясь на материалы WiMAX-форума [4].

11.2. Сравнение WiMAX с HSPA и LTE

Данное сравнение составлено по материалам WiMAX-форума. Информация претендует на объективность, но нет гарантии, что она абсолютно полная. Кроме того, сравнение основано на доступных сведениях, экспериментально не проверенных авторами (по очевидным причинам мы не проводили полноценных натурных испытаний оборудования различных типов), что влечет определенный риск неверной трактовки. Тем не менее, на наш взгляд, это достаточно объективные данные.

Системы с технологией HSPA (3GPP релиз 6) коммерчески доступны с 2007 года. Технология предусматривает частотное дуплексирование (FDD) с шириной каждого дуплексного канала 5 МГц. В нисходящем канале используется модуляция QPSK либо 16-QAM, двойное пространственное разнесение на приеме (1×2 SIMO), пиковая скорость 14 Мбит/с. В восходящем канале модуляция BPSK либо QPSK, антенная конфигурация 1 × 2 SIMO, пиковая скорость 5,8 Мбит/с.

В то же время на рынке были системы WiMAX (релиз 1.0) с временным дуплексированием (TDD). При аналогичной ширине полосы 10 МГц они обеспе-

чивали скорость в нисходящем канале в 2–3 раза более высокую, чем у HSPA (поскольку в WiMAX при TDD общая пропускная способность динамически распределяется между нисходящим и восходящим каналами, точное значение привести невозможно).

Следующим шагом в эволюции систем HSPA являются технологии HSPA+ (HSPA релиз 7 и отдельные поправки релиза 8). Системы HSPA+ стали доступны в конце 2008 года. В нисходящем канале их отличает модуляция 64-QAM с SIMO (1 × 2) или 16-QAM с MIMO (2 × 2). В восходящем канале добавлена модуляция 16-QAM и улучшены возможности для VoIP. Поправки в соответствии с релизом 8 (внедрение ожидается не ранее 2009 года) позволяют использовать в нисходящем канале режим MIMO (2 × 2) с модуляцией 64-QAM, рассматривается возможность использования MIMO больших порядков в нисходящем канале и MIMO (2 × 2) — в восходящем канале.

Сравнивая мобильный WiMAX и HSPA+ (табл. 11.1), можно сделать следующие выводы:

- Мобильный WiMAX (релиз 1.5) имеет сравнимые с HSPA+ (релиз 8) пиковые скорости в нисходящем канале при одинаковых модуляции, скорости кодирования и ширине канала. При этом у мобильного WiMAX в восходящем канале пиковая скорость выше в 2–3 раза.
- Системы HSPA+ ограничены шириной канала 2 × 5 МГц в традиционных спектральных условиях сетей 3G. Мобильный WiMAX поддерживает ширину канала до 20 МГц, как частотное, так и временное дуплексирование. Его частотные профили планируются в диапазонах 700, 1700, 2300, 2500, и 3500 МГц. Мобильный WiMAX обеспечивает «гладкую» IP-сеть (из конца в конец).

Следующим шагом в эволюции систем 3GPP, причем стратегическим шагом, являются системы Long Term Evolution (LTE). Их отличает технология OFDMA в нисходящем канале и SC-FDMA — в восходящем. Модуляция — до 64-QAM, ширина канала — до 20 МГц, дуплексирование TDD и FDD. Применены адаптивные антенные системы, гибкая сеть доступа. Сетевая архитектура — полностью IP-сеть. В системе LTE используются технологии и методы, уже применяемые

Таблица 11.1. Сравнение систем HSPA (релизы 7 и 8) и WiMAX (релиз 1.5)

Параметры	HSPA			WiMAX	
Версия	Релиз 7		Релиз 8	Релиз 1.5	
Диапазон, ГГц	2,0			2,5	
Дуплексирование	FDD			FDD	TDD
Ширина канала, МГц	2 × 5			2 × 5	10
Антенны БС	1 × 2	2 × 2		2 × 2	
Антенны АС	1 × 2			1 × 2	
Модуляция и скорость кодирования					
В нисходящем канале	64-QAM, 5/6	16-QAM, 3/4	64-QAM, 5/6	64-QAM, 5/6	
В восходящем канале	16-QAM, 3/4			64-QAM, 5/6	
Пиковая скорость, Мбит/с					
В нисходящем канале	17,5	21	35	36	48
В восходящем канале	8,3	8,3	8,3	17	24

Таблица 11.2. Сравнение параметров реальных систем LTE (по отчетам производителей) и мобильного WiMAX (релиз 1.5) в одинаковых частотных условиях при FDD с полосами 2 × 20 МГц

Параметры	LTE				WiMAX Релиз 1.5	
	Motorola		T-Mobile	Qualcomm		
Нисходящий канал						
Антенна БС	2 × 2	4 × 4	2 × 4	4 × 2	2 × 2	4 × 4
Модуляция и скорость кодирования	64-QAM, 5/6		64-QAM, 5/6	64-QAM, нет данных	64-QAM, 5/6	
Скорость, Мбит/с	117	226	144	277	144,6	289
Восходящий канал	Нет данных					
Антенна АС			1 × 2	1 × 2	1 × 2	
Модуляция и скорость кодирования			64-QAM	16-QAM	64-QAM, 5/6	
Скорость, Мбит/с			50,4	75	69,1	

в мобильном WiMAX, поэтому следует ожидать схожей эффективности систем LTE (табл. 11.2 и 11.3).

Следует отметить, что системы LTE — это революционное улучшение 3G. LTE представляет переход от систем CDMA к системам OFDMA, а также переход к полностью IP-системе с коммутацией пакетов. Поэтому внедрение этой технологии на существующих сетях сотовой связи означает, как минимум, необходимость новых радиочастотных ресурсов для получения преимущества от широкого канала. Кроме того, для обеспечения обратной совместимости необходимы двухрежимные абонентские устройства. Поэтому плавный переход от систем 3G к LTE весьма проблематичен.

Дальнейшее развитие мобильного WiMAX будут описывать спецификации релиза 2.0. Он будет основан на стандарте IEEE 802.16m, который отражает требования IMT-Advanced. В соответствии с ними, по сравнению с параметрами WiMAX релиз 1.0 вдвое увеличатся спектральная эффективность в нисходящем (до 2,6 бит/с/Гц) и восходящем (1,3 бит/с/Гц) каналах (рис. 11.4). Этот параметр

Таблица 11.3. Сравнение ключевых параметров LTE и WiMAX

Параметры	LTE	WiMAX Релиз 1.5
Дуплексирование	FDD и TDD	FDD и TDD
Частотный диапазон для анализа	2000 МГц	2500 МГц
Ширина канала	до 20 МГц	до 20 МГц
От базы	OFDMA	OFDMA
К базе	SC-FDMA	OFDMA
Спектральная эффективность, бит/Гц/с		
Нисходящий канал, MIMO (2 × 2)	1,57	1,59
Восходящий канал, SIMO (1 × 2)	0,64	0,99
Максимальная скорость мобильной станции, км/ч	350	120
Длительность кадра, мс	1	5
Антенные системы		
Нисходящий канал	2 × 2, 2 × 4, 4 × 2, 4 × 4	2 × 2, 2 × 4, 4 × 2, 4 × 4
Восходящий канал	1 × 2, 1 × 4, 2 × 2, 2 × 4	1 × 2, 1 × 4, 2 × 2, 2 × 4

возрастет вдвое и на границе соты базы — до 0,09 и 0,05 бит/с/Гц для нисходящего и восходящего каналов, соответственно.

Станут возможными более 60 одновременных голосовых сессий на мегагерц для речевого кодека AMR (12,2 кбит/с). Появится режим расширения каналов за счет интеграции отдельных частотных полос — как смежных, так и нет (всего до 100 МГц). Допустимая скорость перемещения мобильных терминалов возрастет до 500 км/ч. Сократится время установления соединения, общая задержка радиосети и время переключения при хендовере. При этом гарантируется полная обратная совместимость с системами WiMAX релиза 1.0 и 1.5. технологий WiMAX и LTE.

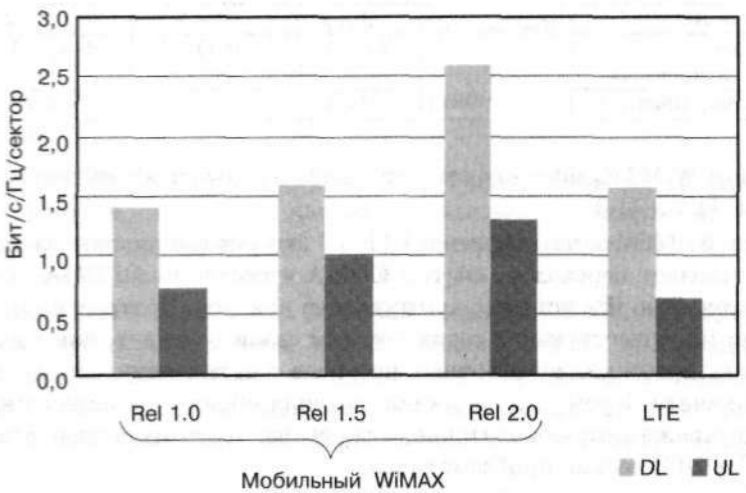


Рис. 11.4. Сравнение средней спектральной эффективности

Отметим, что преимущество в спектральной эффективности означает выигрыш в стоимости развертывания сети (в том числе в удельной стоимости по отношению к пропускной способности сети). Кроме того, возрастает канальная емкость, что позволяет операторам вводить дополнительные сервисы.

Мобильный WiMAX представляет гладкую IP-сеть, сеть LTE более сложна (рис. 11.5). Если сеть WiMAX основывается полностью на IP-протоколах IETF, то сеть LTE более сложна, включает больше протоколов, в том числе — проприетарные протоколы 3G. Немаловажно, что интеллектуальная собственность в области технологий WiMAX, соответствующие патенты распределены среди многих компаний, создан открытый патентный альянс, что позволяет снижать цены абонентских устройств.

Но самое главное преимущество мобильного WiMAX — время выхода на рынок (рис. 11.6). К концу 2008 года только сертифицированных продуктов WiMAX было почти 100, к 2011 году их число возрастет на порядок. Ряд сетей мобильного WiMAX уже введены в коммерческую эксплуатацию. Сети же LTE только планируется начать разворачивать в 2009 году. Притом, что объем инвестиций для апгрейда уже существующих 3G-сетей в сети LTE сравним с затратами на развертывание WiMAX-сетей, фактор времени, а именно выигрыш в 2–3 года, становится решающим при выборе технологий 4G.

Таким образом, можно в целом говорить, что с технической точки зрения и WiMAX, и LTE представляют собой примерно одинаковый класс систем. И весь вопрос — какая технология окажется коммерчески более успешной.

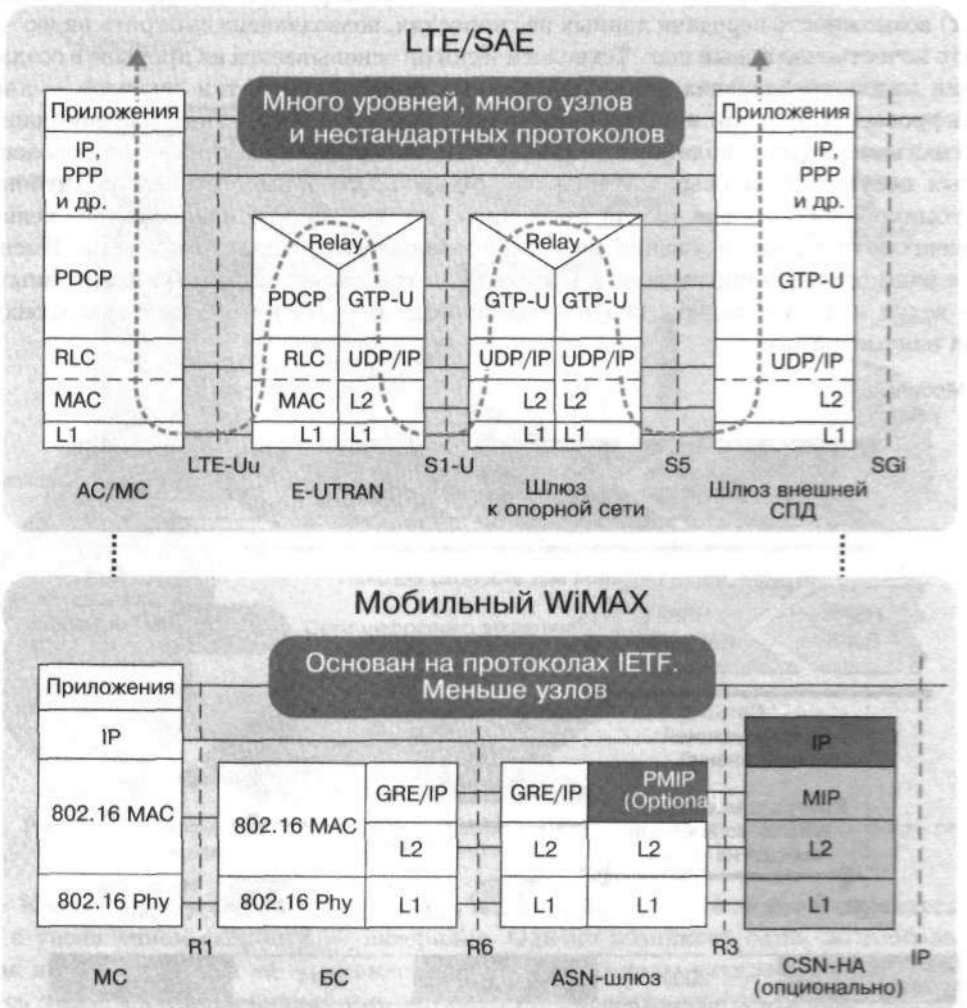


Рис. 11.5. Сравнение системных архитектур сетей WiMAX и LTE

11.3. Что такое 4G?

Вопрос о том, что такое технологии беспроводной связи четвертого поколения (4G) совсем не однозначен. Кто-то даже пытается называть 4G сети WiMAX и HSPA. Но это совершенно неверно. Недаром специалисты себе подобных высказываний не позволяют, используя термины 3,5G, 3,9G (по отношению к LTE) и т. п. Поколение от поколения должно отличаться качественно, причем на всех

уровнях — как технологическом, так и потребительском. В свое время переход от технологий сотовой связи первого поколения ко второму означал переход к цифровым технологиям на техническом уровне и к сервисам передачи данных (пусть и очень простым) — на пользовательском. Переход к 3G означал (означает) возможность передачи данных на скоростях, позволяющих смотреть видео — это качественно новый шаг. Технологически он основывается на прорыве в создании малопотребляющих микроэлектронных средств обработки сигналов — как цифровых (DSP), так и аналоговых (например, высокочастотные малошумящие усилители, полупроводниковые приборы на основе GaAs и других широкозонных полупроводниковых материалов). Микроэлектронные технологии глубоко субмикронного уровня (65–45 нм и ниже) — это ни что иное, как снижение энергопотребления и увеличение функциональности в заданном объеме. Именно ради создания портативных устройств, в том числе телекоммуникационных, и ведут мировые лидеры полупроводниковых технологий пресловутую «гонку за нанометрами».

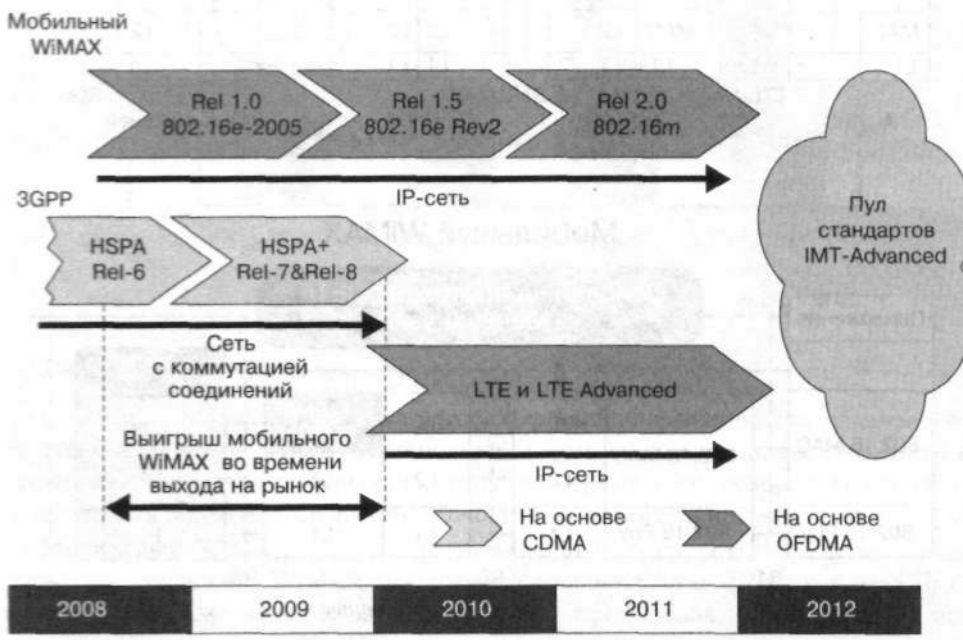


Рис. 11.6. Развитие стандартов мобильного WiMAX и 3GPP

Четвертое поколение устройств, по нашему мнению, — это полноценный мультимедийный «офис в кармане». Именно на это и направлены требования IMT-Advanced по обеспечению такими системами скорости в нисходящем канале до 100 Мбит/с для мобильных и 1 Гбит/с — для номадических и фиксированных абонентов (рис. 11.7). Изначально они были сформулированы в рекомендации ITU-R M.1645 [7], сейчас пребывают в стадии постоянного уточнения. Это — возможность устанавливать голосовые соединения, одновременно возможность для различных информационных сервисов — работа в Интернете, обмена большими массивами данных, просмотр ТВ-трансляций (IPTV), видео по запросу и т. п. То есть все то, что пользователь имеет сегодня у себя дома (в офисе).

И все это — за очень небольшие деньги. Как сотовая телефония позволила быть на связи всегда и везде (почти), так и системы 4G должны обеспечить всех и каждого надежным высокоскоростным доступом к различным сетям передачи данных.

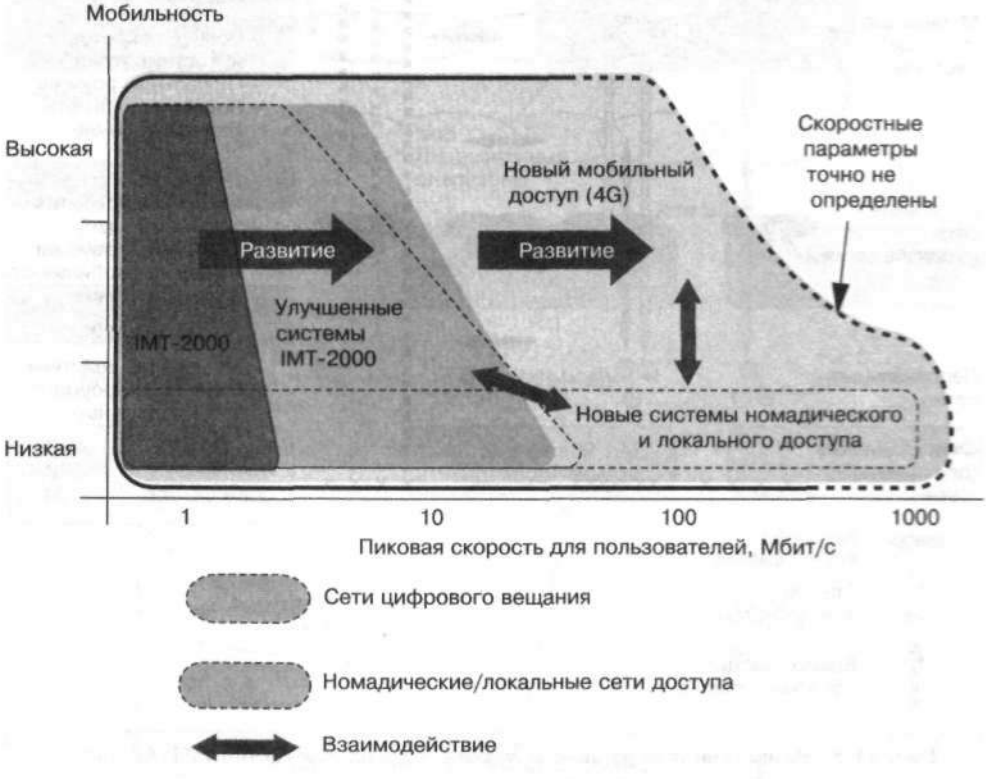


Рис. 11.7. Параметры систем IMT-Advanced. Низкая мобильность означает скорость пешехода, высокая — от 60 до 250 км/ч

Казалось бы, и мобильный WiMAX, и LTE/UMB с такой задачей справятся. А с увеличением скорости — наверняка. Однако возникает одна, но глобальная проблема. И имя ей — совместимость. Много воды утекло с тех пор, как весь телекоммуникационный мир устами ITU провозглашал концепцию единой общемировой беспроводной сети. Теперь уже ясно — протоколов и технологий глобальных систем связи всегда будет несколько. Например, в пуле IMT-2000 — шесть различных стандартов с соответствующими им частотами. Группа стандартов IMT-Advanced также будет представлена различными технологиями, и среди них наверняка будут и WiMAX релиз 2.0, и LTE Advanced, и UMB. Все это — широкополосные технологии, но ни одна из них заведомо не получит 100%-ного распространения. Поэтому им нужно будет не просто совместно сосуществовать, а комфортно сосуществовать. И не мешая, а дополняя друг друга. Только такие технологии — ориентированные на совместную работу и интеграцию на системном уровне — и можно относить к поколению 4G. Причем речь идет о взаимодействии технологий всех уровней — от широковещательных (например, DVB-T2) до сетей фиксированной связи (рис. 11.8).

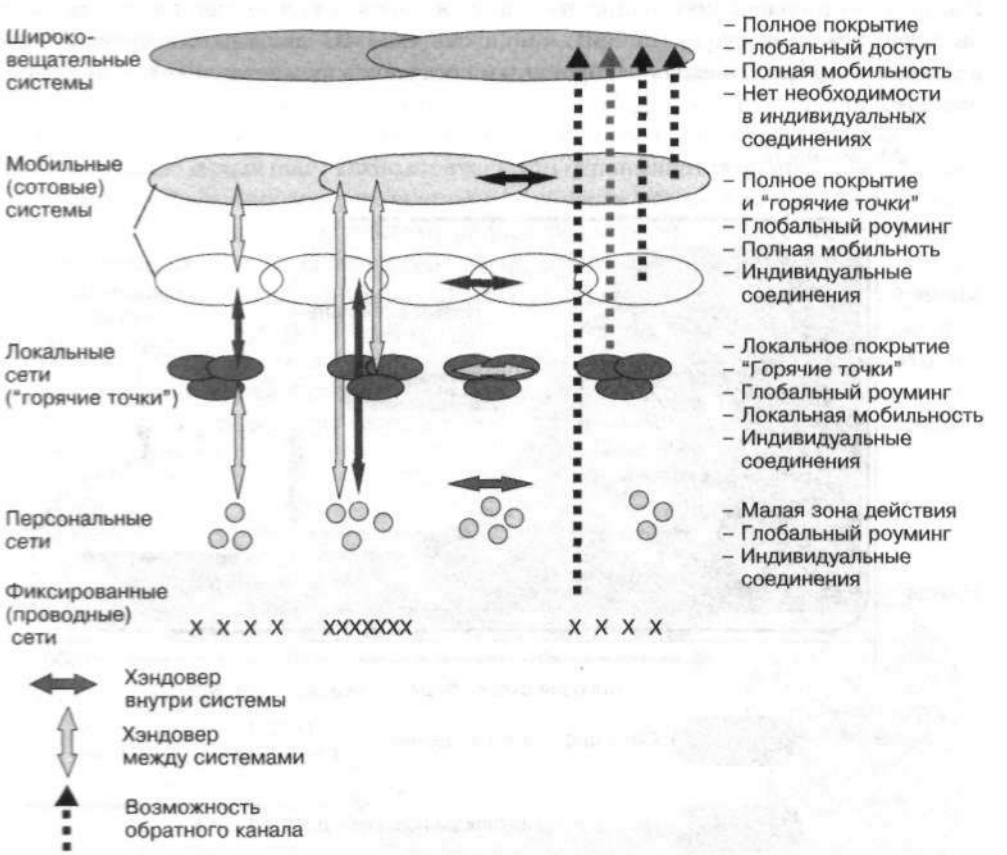


Рис. 11.8. Взаимодействие различных технологий в рамках систем IMT-Advanced

Но за счет чего реально обеспечивается совместимость различных технологий? Для этого, прежде всего, необходимы согласованные протоколы работы в радиосети. Например, должна совпадать длительность кадров и зоны нисходящих и восходящих каналов при временном дуплексе. Обязательна масштабируемость по частотным полосам, причем с одинаковой для разных технологий кратностью (или поддерживаться работа в одинаковых по ширине полосах). Нужны средства гибкой адаптации и перестройки системы, в том числе — на уровне антенных систем. Для этого все технологии IMT-Advanced должны поддерживать работу с адаптивными антенными системами, включая функции формирования диаграмм направленности антенных систем. А в перспективе — и поддерживать динамическое цифровое диаграммообразование (ЦДО). На уровне опорных сетей интеграция должна быть еще более полной, вплоть до прозрачного обмена потоками между сетями с различными радиоинтерфейсами. Отметим, что все эти требования поддерживают перспективные стандарты, разрабатываемые как в рамках LTE-Advanced, так и WiMAX.

Таким образом, системы 4G можно определить как технологии, которые войдут в пул стандартов IMT-Advanced (рис. 11.9). На пользовательском уровне их будет отличать:

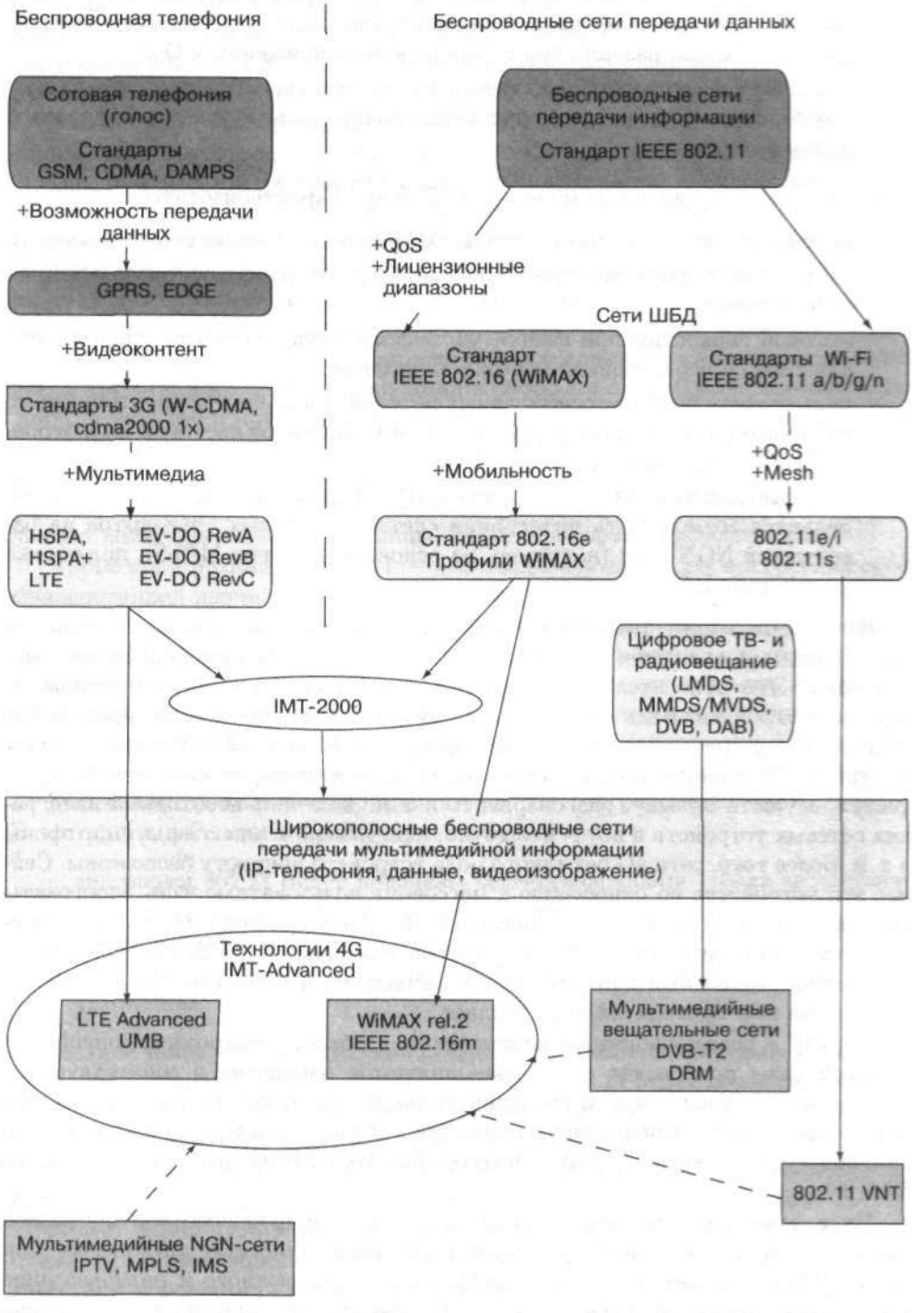


Рис. 11.9. Формирование технологий 4G

- высокая (от 100/1000 Мбит/с для мобильных/номадических абонентов) скорость. Это означает работу одновременно с несколькими мультимедийными потоками, различными по природе и требованиям к QoS;
- взаимная совместимость и активное взаимодействие. Пользователь не должен ощущать ни помех от других сетей, ни проблем с межсетевой передачей данных.

На уровне технологическом системы 4G будут характеризоваться:

- полным переходом к модуляции OFDM (работа в условиях переотражений);
- согласованностью совместной работы на уровне радиопротоколов физического уровня;
- высокой гибкостью при выборе частотных полос, частотных диапазонов, адаптивной перестройкой методов модуляции;
- применением наиболее совершенных методов канального корректирующего кодирования (каскадные коды, коды LDPC, развитой системой многоуровневого интерливинга и т. п.);
- опорные/базовые сети — полностью IP (с переходом к протоколам IPv6), появится возможность интеграции систем различных стандартов на базе единой NGN-сети (например, на основе технологии MPLS), поддержка платформы IMS.

Все перечисленные технологические особенности опираются на поистине революционные достижения последних лет в области микроэлектронной элементной базы. Это относится не только к функциональности самих приемопередающих устройств. Ведь сами по себе данные пользователю, как правило, не нужны — ему требуются средства их обработки. Конечный потребитель хочет смотреть ТВ и видео, обмениваться сообщениями, прикрепляя к ним большие файлы, слушать музыку, разговаривать и т. п. Для чего необходима интеграция сетевых устройств в ноут- и нетбуки, наладонные компьютеры, смартфоны и т. п. Более того, сети 4G без такого рода устройств попросту бесполезны. Сейчас эти устройства по отношению к массовому пользователю чуть дороговаты (в отношении нетбуков это уже неверно). Но через три года их цена неминуемо упадет до стоимости обычного сотового телефона. И тогда потребность на скоростной мобильный контент станет глобальной и всеобщей. Отметим этот факт — он нам пригодится чуть позднее.

Наряду с революционными изменениями в области микроэлектронной элементной базы происходят и не менее значимые изменения в сопутствующем программном обеспечении и создании принципиально нового интерфейса «человек–компьютер». Специалисты в области мобильного контента называют это заменой существующего сегодня интерфейса Web 2.0 на интерфейс будущего Web 3.0.

Ну и, конечно, огромное влияние на развитие широкополосной мобильной связи 4G оказывают (и будут оказывать) такие смежные технологии, как IEEE 802.11, а также технологии цифрового телевизионного и радиовещания. Стандарты 802.11 продолжают активно совершенствоваться, их уже давно нельзя позиционировать только как технологии для беспроводных локальных сетей. Известны многочисленные примеры, когда на основе методов 802.11

строились сети ШБД городского масштаба, причем с объединением нескольких регионов. А с появлением стандарта mesh-сетей 802.11s, высокоскоростных стандартов 802.11n и 802.11 VNT (в перспективе) это направление будет только развиваться.

Разумеется, стандартам 802.11 нет места в пуле IMT-Advanced. Для этого у них отсутствует ряд важных свойств — прежде всего, нет поддержки мобильности и высокой плотности абонентов. Не говоря уж об отсутствии единых частотных полос в лицензируемых диапазонах. Но ведь сети 802.11 специально создавались для работы в безлицензионных диапазонах частот. И их основное свойство, которое неизменно сохраняется во всех новых системах — простота инсталляции и низкая стоимость. Сегодня эта технология — доминирующая и фактически безальтернативная для беспроводных локальных сетей. Уже крайне сложно найти смартфон без поддержки 802.11 и практически невозможно найти такой ноутбук. Абонентское оборудование стоит порядка 10 долл. и менее, точки доступа (для работы внутри помещений) — порядка 100 долл. Огромная армия пользователей «бесплатно» оснащена адаптерами 802.11, этот интерфейс воспринимается как должное в мобильных устройствах (например, как USB-порт). Поэтому стандарты 802.11 незаменимы для формирования сети доступа в различных локальных зонах (гостиницы, кафе, аэропорты, вокзалы и т.п.).

Кроме того, технология 802.11 — действенное решение для построения сетей фиксированной широкополосной связи в локальных зонах (город с населением порядка 100 тыс. жителей, «горячие зоны», а также регионы, где невозможны проекты с большими объемами инвестиций). Разумеется, в таких проектах, если идет речь о предоставлении операторского качества услуг, необходима работа в лицензируемом диапазоне и с существенно большими мощностями передатчиков, чем в случае домашних/офисных сетей. Вопрос о перспективности таких решений спорен, однако подобный опыт есть. И тут много будет зависеть от дальнейших решений национального регулятора радиочастотного спектра. В частности, в России возможны два варианта:

- диапазоны 802.11 будут признаны безлицензионными (тогда автоматически возникнут ограничения на уровень эквивалентной изотропной мощности в антенне);
- будут выделяться частоты для "802.11-образных" сетей в нестандартных диапазонах.

Каждое такое решение способно существенно повлиять на судьбу технологии 802.11, но не кардинально. В любом случае сети 802.11 будут продолжать развиваться и сосуществовать с сетями 4G, оказывая на них большое влияние. Ведь все основные перспективные решения, которые заложены в стандарты 3,5–4G, изначально воплощались в оборудовании 802.11. Это и MIMO, и mesh-сети, и агрегация/фрагментация пакетов, и OFDM — перечислять можно долго.

Может быть, чуть меньшее, но тем не менее существенное влияние на 4G окажут и развивающиеся технологии цифрового вещания. Прежде всего, они в известной мере окажутся конкурентами ряду услуг 4G. Ведь в последних предусмотрена передача видео- и аудиоконтента, включая ТВ- и радиотрансляции. Кроме того, изменяются и сами стандарты цифровых широковещательных сетей. Они становятся все более мультимедийными, с возможностью обратной

связи и т. п. Как будет строиться взаимодействие таких сетей с сетями 4G? В ответ на этот вопрос можно только фантазировать. Возможно, в чуть более отдаленном будущем их ожидает слияние в единую технологию, возможно — возникнут различные интегральные решения. Может сохраниться и существующий паритет. Ведь вещательные технологии специально рассчитаны на возможность приема слабого сигнала, что дает им несомненное преимущество. Но одно несомненно — независимо существовать, не замечая друг друга, они не смогут.

И, разумеется, нельзя не упомянуть о развитии технологий сетей пакетной передачи. Собственно, развитие беспроводных сетей лишь отражает основные тенденции сетей проводных. А в этой области вот уже лет 20 неуклонными темпами происходят поистине революционные изменения. Причем настолько стабильно, что все успели к этому привыкнуть и перестали воспринимать как революцию. Технология пакетной коммутации уверенно вышла на уровень транспортных сетей, пройдя путь до мультисервисной транспортной платформы (MSTP) — технологии SDH следующего поколения. Пакеты Gigbit Ethernet (и последующих Ethernet-технологий) уже передают непосредственно по магистральным каналам волоконно-оптических линий связи (ВОЛС). Для предоставления мультимедийных услуг создана платформа IMS (IP Multimedia Subsystem) и протокол SIP (Session Initiation Protocol) [8]. Сети NGN с поддержкой MPLS (как наиболее эффективного сегодня механизма обеспечения QoS в мультимедийных сетях) уже получили широчайшее развитие [9]. И все эти технологии, вкупе с не менее бурно прогрессирующими технологиями ВОЛС, обеспечивают для сетей ШБД прочнейший фундамент — как идеологический, так и формируя собственно наземную сетевую инфраструктуру. Недаром поддержка платформы IMS прямо прописана в спецификациях как LTE, так и WiMAX. А опорная сеть на основе ВОЛС с DWDM и поддержкой IP-MPLS — это не исключение, а скорее стандарт для современных беспроводных сетей широкополосного доступа.

11.4. Перспективы WiMAX

Можно уверенно утверждать, что технология WiMAX обладает всеми указанными выше признаками 4G. Это мы показали и в описании проекта стандарта IEEE 802.16m, и в сравнительном анализе с LTE. Однако войти в пул IMT-Advanced мало, нужно еще и добиться массового применения. Но сегодня очень многие аналитики сходятся на том, что WiMAX уготована участь нишевой технологии, что ее доля в общем пироге сетей ШБД едва ли превысит 5% (вспомним приведенные в начале главы прогнозы — к 2015 году всего 2100 млн. абонентов сетей ШБД, из них 98 млн. — абонентов WiMAX-сетей). Но тут всегда интересно смотреть, кто и зачем такие прогнозы делает. И на чем при этом основывается. Итак, кому это выгодно? Ответ очевиден — держателям основных патентов на технологии сотовой связи 3G. Таких компаний, как известно, две — Ericsson и Qualcomm. Последняя, например, при объеме выручки в 2008 году свыше 11 млрд. долл. почти 4 млрд. получила в виде роялти и лицензионных отчислений. Один из крупнейших лицензиатов Qualcomm — компания Nokia. Отметим, что Qualcomm с Ericsson долго вели патентные войны, пока к 2006 году не договорились о взаимном признании патентов. Очевидно, что двум этим монстрам невыгодно развитие мощного конкурента. К тому же, компания Eric-

sson сегодня практически не выпускает электронные компоненты, но является крупным производителем телекоммуникационного оборудования, а также владельцем сетей сотовой связи. И годовой оборот у нее — порядка 25 млрд. долл. Телефоны Ericsson производит в рамках совместного с Sony предприятия — компании Sony-Ericsson. В 2008 году — неудачном, убыточном — объем продаж у этого предприятия составил почти 11,25 млрд. евро. Очевидно, что эти гиганты очень не хотят терять свои позиции на рынке. В этот же лагерь можно отнести крупнейшего в мире производителя сотовых телефонов — компанию Nokia, чрезвычайно тесными узами связанную с технологиями 3G.

И с подачи этих корпораций все аналитики, предрекающие WiMAX относительно незавидную участь, исходят из двух постулатов:

- сотовые компании (операторы) предпочтут плавный переход к LTE и 4G, нежели выберут совершенно новую технологию;
- бурное (если не сказать — взрывное) распространение технологий HSPA свидетельствует о выборе потребителей в пользу технологий 3GPP/3GPP2.

Оценивая общий объем рынка беспроводных широкополосных систем и экстраполируя существующую динамику продаж телефонов с HSPA, и получают столь безрадостную для WiMAX картину. Но насколько достоверен такой подход?

Мы не будем рассуждать о пиковых скоростях и спектральной эффективности различных технологий — данные от разных аналитиков зачастую прямо противоположны. Сойдемся на том, что они схожи. Но идеологически HSPA — это далеко не мобильный WiMAX. Хотя бы потому, что использует опорную сеть с коммутируемыми каналами (наследие телефонной природы) или *одночастотную технологию передачи* (CDMA). *Нам важно не сравнивать* конкретные показатели каких-либо конкретных сетей, а оценить уровень технологии (т.е. перспективность, способность с меньшими инвестициями повысить ее характеристики) Потенциально же технология HSPA уступает сетям на основе IEEE 802.16e. Этот факт признают все, иначе не стали бы разрабатывать технологии LTE и LTE Advanced (правда, по данным из презентации компании Qualcomm [5] следует, что пропускная способность систем HSPA+ окажется даже выше, чем у систем LTE — 42 против 37 Мбит/с в 5-МГц полосе нисходящего канала, MIMO 2 × 2. Но это понятно — Qualcomm в CDMA-системах заинтересована кровно).

Поэтому успех систем HSPA говорит только об одном — пользователям нужен высокоскоростной доступ. И неважно какой. Реальной же конкуренции с WiMAX у этой технологии еще не было. Поэтому некорректно экстраполировать темпы ее развития на будущее. Это все равно, что оценить скорость машины ночью, на пустой трассе, и на этом основании делать прогнозы о скорости проезда днем, когда появятся другие машины со всеми вытекающими дорожными коллизиями. Утверждение же, что пользователи, привыкнув к HSPA, захотят другие технологии 3GPP, вообще лишено смысла — пользователь меняет свой телефон (коммуникатор, смартфон и т. п.) не реже чем раз в три года, и ему глубоко все равно, на каких принципах строится технология передачи данных. А аналитики сегодня считают объем проданных HSPA-устройств (порядка 100 млн. на конец 2008 года), соотносят их с числом абонентов WiMAX (4 млн.) и полагают, что такая пропорция (3–5%) сохранится вечно. Очень спорное утверждение.

Также сомнителен тезис о том, что операторам UMTS-сетей проще переходить на технологию LTE. Если не сказать — неверен. Обратной совместимости между UMTS и LTE нет — ни на уровне опорной сети, ни на уровне радиоинтерфейса. А инвестиции в создание новых сетей и WiMAX, и LTE сопоставимы. То есть операторам с этой точки зрения все равно. Но им далеко не все равно с точки зрения рыночной конкуренции. WiMAX есть уже сегодня, а LTE нужно ждать. А ждать нельзя. Поэтому ряд компаний — владельцев сотовых сетей успешно разворачивают сегодня сети WiMAX. Причем это — наиболее продвинутые по своему масштабу и идеологии сети. В США к таким компаниям относится проект Sprint-Cleaware, в Пакистане — сеть Wateen, в Южной Корее — Korea Telecom, в Индии — сеть Tata Communications и т. д. А вот в рамках существующих систем WiMAX переход на стандарты следующих уровней гораздо проще, чем с 3G на LTE. Даже с сетей фиксированного доступа. То есть операторы считают выбор WiMAX как перспективной технологии оправданным. Хотя бы потому, что сегодня нет иной альтернативы. Но только ли поэтому?

Что, по большому счету, является при прочих равных для операторов массовых сетей решающим фактором при выборе технологии? Наличие частотного ресурса и массового производства абонентских устройств. Насколько эти факторы влияют на выбор WiMAX или иной технологии? Вопрос с радиоспектром по отношению к WiMAX достаточно нейтрален — проблемы аналогичны и у операторов WiMAX, и у сотовых операторов, поскольку в любом случае нужны новые частотные полосы. А вот с абонентскими устройствами ситуация иная. Вспомним, вся сотовая связь — это результат достижений микроэлектроники. И на рынке абонентских устройств повелевают, прежде всего, производители элементной базы. Им, в общем-то, все равно, что производить — WiMAX-, LTE- или HSPA-чипсеты. Но хорошо известно, что основная прибыль формируется не в производственных цехах. Прибыль заводов в Юго-Восточной Азии составляет (очень примерно) процентов 10 от себестоимости. А основная прибыль оседает у разработчиков технологии, у держателей патентов на используемые технические решения (например, у Qualcomm она превышает 35% от общей выручки в 11,13 млрд. долл.). Вот почему важно не просто производить продукты, но и создавать технологии. Отсюда оправданность многомиллиардных инвестиций всех ведущих производителей в научные изыскания. Уж больно привлекателен приз.

В истории мировой электроники за последние 30 лет было две ярко выраженные инновационные волны, на гребне которых эта отрасль развилась и приняла современный вид — персональные компьютеры и мобильные телефоны. Безусловным вдохновителем первой была компания Intel — не единственным, но одним из основных. Именно благодаря микропроцессорам для ПК Intel стала тем, кем стала — бессменным технологическим лидером и законодателем мод с оборотом 34 млрд. долл. (вдвое больше, чем ближайший конкурент в области полупроводников, имя которому — Samsung Electronics). Но вторую волну этот гигант пропустил — нет его решений в сотовых телефонах, не используется в них его элементная база, коммуникаторы не в счет. И тут поднимается третья волна — широкополосная беспроводная связь и связанные с ними персональные устройства с потребностью в огромной вычислительной мощности. Необъятный рынок. И уж эту волну компания Intel пропустить не могла.

А теперь посмотрим, кто по сути вдохновляет WiMAX-движение в последние годы, кто выступает одним из самых активных членов WiMAX-форума?

Компания Intel. Вспомним, IEEE 802.16 начинался в 1990-е годы как стандарт операторского класса, о массовости речи не шло. Но появился WiMAX-форум во главе с Intel, и все кардинально изменилось — фиксированный WiMAX задвинут, развивается мобильный стандарт, ориентированный на конечного потребителя. Вот она, массовость, в котором заинтересован Intel и его соратники. А кто ж не хочет быть соратником Intel — компании, в микроэлектронике ни в какие технологические альянсы не вступающей?

Возьмем на себя смелость утверждать, что Intel — это один из основных действующих факторов, делающих технологию WiMAX не просто перспективной, но едва ли не более перспективной, чем LTE. Обладая абсолютно недоступными кому бы то ни было технологическими возможностями (не столько даже самими возможностями, сколько способностью постоянно опережать всех в области технологий), этот монстр способен самостоятельно сформировать рынок абонентских WiMAX-устройств и монопольно определять цены на них. Сейчас Intel выпускает модемы WiMAX в виде отдельных модулей. Но при некотором желании ни один центральный процессор для ПК не будет выпускаться без поддержки WiMAX — заметим, практически даром. Но этого уже и не требуется — у Intel немало соратников-производителей. Кроме того, сами мобильные абонентские устройства 4G должны обладать высокой вычислительной мощностью, оснащаться мощным центральным процессором с чрезвычайно низким энергопотреблением. А в этой области у Intel конкурентов не наблюдается. Хотя бы просто потому, что низкое удельное энергопотребление — это новые технологии уровня 45 и менее нанометров [6]. А в перспективе — переход на новые материалы. В этой сфере Intel пока в одиночку успешно конкурирует с объединением всех ведущих полупроводниковых компаний. Конечно, можно, как и сегодня, использовать в коммуникаторах и смартфонах центральный процессор от Intel, а телекоммуникационные модули — от других производителей. Но только 4G означает массовый переход на такого рода устройства с соответствующим снижением цен. И старые подходы могут не сработать.

Обратим внимание, производство устройств WiMAX уже началось, а патентных скандалов, характерных для развития почти каждой новой технологии, удалось избежать. Даже создан патентный альянс Open Patent Alliance, куда вошли практически все лидеры в своих направлениях — производство элементной базы, абонентских устройств, базовых станций и сетевой инфраструктуры.

Небезынтересно, что едва ли не первый (вопрос первенства всегда сложен) создатель оборудования для сотовых сетей (в далеких 1970-х) — компания Motorola — и в области WiMAX остается лидером по объему производства. 50 млн. долл. за квартал — вроде бы и не много, но ведь и сети только начали строиться. Первая в мире мобильная сеть WiMAX (тогда еще WiBro) была создана в Южной Корее на оборудовании гиганта Samsung Electronics (объем продаж в 2008 году, по утверждению компании — 105 млрд. долл.). А WiBro — это технология Samsung. Если эта корпорация пошла на модификацию технологии и приведение ее в соответствие профилям WiMAX, то уж наверняка интерес Samsung в ее развитии велик. Так что с соратниками у Intel все благополучно.

И вот этот фактор — мощная и заинтересованная поддержка производителей элементной базы и оборудования — для операторов и является определяющей. Конечно, и производители оборудования стандартов 3GPP/3GPP2 не дремлют,

но три года форы (с учетом кризиса — точно не меньше) — это очень много. Это — смена технологического поколения в микроэлектронике. Через три года появятся возможности, о которых сегодня нельзя даже мечтать. Например, (может быть) реальностью станут абонентские устройства с динамическим формированием диаграммы направленности приемной/передающей антенн, появится доступная элементная база для каналов в миллиметровых диапазонах (порядка 60 ГГц) для микро-mesh-сетей, дешевыми и массовыми станут фемто-соты и т. п. И во всех этих инновациях кровно заинтересованы все производители, входящие в WiMAX-форум. Поэтому стимулы для развития технологии WiMAX очень велики, а перспективы — весьма привлекательны. Ведь навязали же в свое время производители — именно они — реально не нужные пользователям и операторам технологии 3G. А в WiMAX есть и реальная потребность, и рыночные выгоды.

Разумеется, WiMAX вряд ли станет доминирующей системой ШБД. Скорее, можно будет говорить об альянсе WiMAX/LTE/UMB. Не забудем и про системы 802.11, а также про цифровое ТВ- и радиовещание. Поэтому особую роль приобретут многомодовые, интегрированные устройства. Так, уже сегодня Samsung выпускает 3G/WiMAX-совместимые устройства, по аналогичному пути уже идут многие другие производители. И это относится не только к абонентскому оборудованию, но и к сетевой инфраструктуре. А в чуть более отдаленной перспективе, наверное, появятся системы когнитивного радио, когда устройства будут поддерживать все существующие технологии ШБД и сами смогут динамически выбирать тип протокола, частоту, ширину полосы и другие параметры в зависимости от ситуации. Но при этом число базовых технологий вряд ли сократится. Не нужно ограничивать мир — пусть он будет разнообразным. И пусть пользователь берет лучшее от каждой его грани!

Литература

1. Brydon A., Heath. M. Wireless broadband forecasts for 2008–2015: HSPA, HSPA+, EV-DO, LTE and WiMAX. — Analysys Mason, июль 2008.
2. Mobile Broadband Evolution: the roadmap from HSPA to LTE. A White Paper from the UMTS Forum. — UMTS Forum, февраль 2009.
3. www.cdg.org.
4. Gray. D. Comparing Mobile WiMAX with HSPA+, LTE, and Meeting the Goals of IMT-Advanced. — WiMAX Forum, февраль 2009.
5. HSPA+ for Enhanced Mobile Broadband. — Qualcomm, февраль 2009.
6. Шахнович И. Технологии уровня 45 нм: 45, 32, далее везде. — Электроника: НТБ, 2008, № 2, с. 102–109.
7. RECOMMENDATION ITU-R M.1645. Framework and overall objectives of the future development of IMT-2000 and systems beyond IMT-2000. — ITU-R, 2003.
8. Абрагин Д. Телекоммуникационные сети нового поколения: решения НТЦ «Натекс». — Первая миля, 2009, № 2, с.28–31.
9. Гольдштейн А. Б., Гольдштейн Б. С. Технология и протоколы MPLS. — СПб.: БХВ, 2005.

Глоссарий

3-DES	Triple data encryption standard	Тройной стандарт шифрования данных
3GPP	Third Generation Partnership Project	Проект партнерства третьего поколения
3GPP2	Third Generation Partnership Project 2	Второй проект партнерства третьего поколения
AA	Anchor Authenticator also called Network Authenticator Server (NAS)	Опорная точка аутентификации, также называют сетевым аутентификационным сервером
AAA	Authentication, Authorization, and Accounting	Аутентификация, авторизация и учет
AAA Proxy	An intermediary for transparently routing and/or processing AAA messages sent between a AAA client and a AAA server	Посредник для прозрачной маршрутизации и/или обработки AAA-сообщений, пересылаемых между AAA-клиентом и AAA-сервером
AAA Server	Computer system performing AAA services (authentication, authorization, accounting)	Компьютерная система, осуществляющая AAA обслуживание (авторизацию, аутентификацию и учет)
AAA-V	AAA proxy server located within the visited network	AAA прокси-сервер, расположенный в гостевой сети
AAS	Adaptive Antenna System also Advanced Antenna System	Адаптивная антенная система или улучшенная антенная система
AASN	Anchor ASN	Опорная ASN. Относится к ASN, которые осуществляют функции опорного маршрута данных для определенной станции
AC	Admission Control	Управление доступом
ACK	Acknowledge	Подтверждение, квитанция
ADPF	Anchor Data Path Function	Функция опорного маршрута данных
AES	Advanced encryption standard	Улучшенный стандарт шифрования
AF	Application Function	Функция приложения
AG	Absolute Grant	Абсолютный допуск
AGC	Automatic gain control	Автоматическая регулировка уровня (АРУ)
AK	Authorization Key	Ключ авторизации
AK SN	Derivation from PMK and PMK2 SN	Производная от PMK и PMK2 SN
AKA	Authentication and Key Agreement	Соглашение об аутентификации и ключе
AM	Authorization Module	Модуль авторизации
AMC	Adaptive modulation and coding	Адаптивные модуляция и кодирование
A-MIMO	Adaptive Multiple Input Multiple Output (Antenna)	Адаптивная MIMO
AMS	Adaptive MIMO Switching	Адаптивное MIMO-коммутация

APC	Anchor paging Controller	Опорный контроллер пейджинга
APCF	Anchor paging controller function	Функция опорного контроллера пейджинга
API	Application Program Interface	Интерфейс программ приложения
AR	Access Router	Маршрутизатор доступа
ARP	Address resolution protocol	Протокол переопределения (преобразования) адресов
ARQ	Automatic Retransmission Request	Автоматический запрос на повторную передачу
AS	Authentication Server	Сервер аутентификации
ASA	Authentication and service authorization	Аутентификация и авторизация услуги
ASN	Access Service Network	Сервисная сеть доступа
ASP	Application Service Provider	Сервис провайдер приложений
ATDD	Adaptive time division duplexing	Адаптивное временное дуплексирование
ATM	Asynchronous transfer mode	Режим асинхронной передачи
BCC	Block convolutional code	Блоковый сверточный код
BCE	Binding Cache Entry	Связанная запись в кэш
BE	Best Effort	Канал с не гарантированной скоростью, но наилучший из возможных
BER	Bit error ratio	Коэффициент (частота, вероятность) ошибки в символе
BNI	Base station network interface	Сетевой интерфейс базовой станции
BR	Bandwidth request	Запрос полосы
BRAN	Broadband Radio Access Network	Сеть широкополосного радиодоступа
BRAS	Broadband Remote Access Server	Сервер широкополосного удаленного доступа
BS	Base Station	Базовая станция
BSID	Base Station Identifier	Идентификатор базовой станции
BSN	Block Sequence Number	Номер блоковой последовательности
BTC	Block turbo code	Блоковый турбокод
BU	Binding Update	Компоновка обновления
BW	Bandwidth	Полоса частот
BWA	Broadband wireless access	Широкополосный беспроводный доступ
BWAA	Bandwidth allocation/ access	Распределение полосы частот и доступа
C/(I+N)	Carrier-to/(interference plus noise) ratio	Отношение несущая / (помеха плюс шум)
C/I	Carrier-to-interference ratio	Отношение несущая / помеха
C/N	Carrier-to-noise ratio	Отношение несущая / шум
CA	Certification authority	Центр сертификации
CAC	Connection Admission Control	Контроль за установлением соединений
CBC	Cipher block chaining	Последовательность блоков шифра
CBC-MAC	Cipher block chaining message authentication code	Последовательность блоков шифра для сообщения кода аутентификации

CC	Chase Combining (also Convolutional Code)	Комбинированный метод Чейза (или сверточный код)
CC	Confirmation code	Код подтверждения
CCH	Control subchannel	Управление подканалами
CCI	Co-channel interference	Межканальная интерференция
CCM	Counter with Cipher-block chaining Message authentication code	Счетчик последовательности блоков шифра кода аутентификации сообщения
CCoA	Collocated Care of Address	Связанное предоставление адреса
CCS	Common channel signaling	Общий канал сигнализации
CCV	Clock comparison value	Значение сравнения отсчетов
CDF	Cumulative Distribution Function	Кумулятивная функция распределения
CDMA	Code division multiple access	Множественный доступ с кодовым разделением
CDMA2000	3rd Generation Code Division Multiple Access Radio Technology	Радиотехнология множественного доступа с кодовым разделением третьего поколения
CEPT	European Conference of Postal and Telecommunications Administrations	Европейская конференция почтовых и телекоммуникационных администраций
CG	Continuous grant	Непрерывный допуск
ChID	Channel Identifier	Идентификатор канала
CID	Connection IDentifier	Индикатор соединения
CINR	Carrier to noise and interference ratio	Отношение несущая / шум и помеха
CIR	Channel impulse response	Импульсный отклик канала
CLP	Cell Loss Priority	Приоритет потери элемента данных
CMAC	Block Cipher-based Message Authentication Code	Аутентификационный код сообщения, основанный на блочном шифре
CMIP	Client Mobile IP	Мобильный IP-клиент
COA	Change of Authority	Изменение полномочий
CoA	Care of Address	Адрес для передачи (временный IP-адрес, используемый мобильным узлом при подключении к внешней сети)
COS	Class of Service	Класс обслуживания
CP	Cyclic Prefix	Циклический префикс
CPS	Common part sublayer	Общий подуровень MAC-уровня
CQI	Channel Quality Indicator (или information)	Индикатор (или информация) качества канала
CQICH	Channel quality information channel	Информационный канал оценки качества канала
CRC	Cyclic redundancy check	Циклическая избыточная проверка
CS	Convergence Sublayer	Подуровень конвергенции (преобразования) MAC-уровня

CSCF	Centralized scheduling configuration	Конфигурация с централизованным планированием
CSCH	Centralized schedule	Централизованное планирование
CSIT	Channel state information at the transmitter	Информация о состоянии канала в передатчике
CSN	Connectivity Service Network	Сеть подключения
CSTD	Cyclic Shift Transmit Diversity	Разнесение передачи циклическим сдвигом
CTC	Convolutional Turbo Code	Сверточные турбокоды
CTR	Counter mode encryption	Счетчик режима шифрования
CUI	Chargeable User Identity	Изменяемые особенности (профиль) пользователя
DAD	Duplicate Address Detection	Определение дуплицированного адреса
DAMA	Demand assigned multiple access	Множественный доступ с предоставлением каналов по требованию
DARS	Digital audio radio satellite	Спутник цифрового радиовещания
dBi	Decibels of gain relative to the zero dB gain of a free-space	Усиление (дБ) по отношению к излучению изотропной антенны с нулевым усилением
dBm	Decibels relative to one milliwatt	дБ относительно мВт
DCD	Downlink Channel Descriptor	Дескриптор нисходящего канала
DES	Data encryption standard	Стандарт шифрования данных
DFS	Dynamic frequency selection	Динамический выбор частоты
DHCP	Dynamic Host Configuration Protocol	Протокол динамического конфигурирования узла (хост-машины)
diffserv	Differentiated services	Дифференциальное обслуживание
DIUC	Downlink Interval Usage Code	Код профиля пакетов в нисходящем канале
DL	Down Link	Нисходящий канал
DLFP	Downlink frame prefix	Префикс кадра нисходящего канала
DNS	Domain Name Service	Служба доменных имен
DOCSIS	Data Over Cable Service Interface Specification	Протокол передачи данных по коаксиальным (телевизионным) кабелям
DoS	Denial of Service	Отказ в обслуживании
DP	Decision Point Data Path	Точка принятия решения по маршруту данных
DSA	Dynamic service addition	Динамическое добавление услуг
DSC	Dynamic service change	Динамическое изменение услуг
DSCH	Distributed schedule	Распределенный каталог
DSCP	Differentiated services code-point	Код указателя службы при дифференциальном обслуживании
DSD	Dynamic service deletion	Динамическое исключение услуги

DSL	Digital Subscriber Line	Цифровая абонентская линия
DSLAM	Digital Subscriber Link Access Multiplexer	Мультиплексор доступа посредством DSL
DSx	Dynamic service addition, change, or deletion	Динамическое добавление, изменение или удаление услуги
DVB	Digital Video Broadcast	Цифровое видеовещание
E2E	End-to-End	Непрерывная, из конца в конец (про IP-сеть)
E911	US Emergency Services	Служба неотложной помощи США
EAP	Extensible Authentication Protocol	Расширяемый протокол аутентификации
EAP-AKA	EAP Authentication and Key Agreement to be used with USIM	EAP-аутентификация и соглашения о ключе для использования с USIM
EAP-PSK	Extensible Authentication Protocol — Pre Shared Key	EAP с предварительным ключом
EAP-SIM	EAP Subscriber Identity Module to be used with SIM	SIM-карта с EAP
EC	Encryption control	Управление шифрованием
ECB	Electronic code book	Электронная кодовая книга
EDE	Encrypt-decrypt-encrypt	Шифрация–дешифрация–шифрация
EESM	Exponential Effective SIR Mapping	Экспоненциально эффективное отображение отношения сигнал/шум
EESS	Earth exploratory satellite systems	Спутниковая система исследования Земли
EIRP	Effective isotropic radiated power	Эффективная изотропная излучаемая мощность
EKS	Encryption key sequence	Последовательность ключа шифрования
EMSK	Extended Master Session Key	Расширенный ключ мастер-сессии
ErtPS	Extended Real-Time Polling Service	Расширенная поллинг-служба реального времени
ETSI	European Telecommunications Standards Institute	Европейский телекоммуникационный институт стандартов
EUI	Extended unique identifier	Расширенный уникальный идентификатор
EUI-64	Extended Unique Identifier (64-bit)	Расширенный уникальный идентификатор (64 бита)
EVM	Error vector magnitude	Величина вектора ошибки
FA	Foreign Agent	Внешний агент
FBSS	Fast Base Station Switching	Быстрое переключение базовой станции
FC	Fragmentation control	Управление фрагментацией
FCAPS	Fault Configuration Accounting Performance and Security	Модель ISO, отражающая ключевые функции администрирования и управления сетями: отказами, конфигурацией, учетом, производительностью, безопасностью

FCH	Frame Control Header	Заголовок управления кадром
FDD	Frequency division duplex	Частотный дуплекс
FEC	Forward error correction	Кодирование с исправлением ошибок
FFT	Fast Fourier transform	Быстрое преобразование Фурье
FHDC	Frequency hopping diversity coding	Разнесенное кодирование с прыжками по частоте
FIPS	Federal Information Processing Standard	Федеральный стандарт обработки информации
FQDN	Fully Qualified Domain Name	Полное доменное имя узла
FRD	Fast Router Discovery	Быстрое обнаружение маршрутизатора
FSH	Fragmentation subheader	Подзаголовок фрагментации
FSN	Fragment sequence number	Номер фрагмента последовательности
FSS	Fixed satellite service	Фиксированная спутниковая служба
FUSC	Fully Used Sub-Carrier	Полное использование поднесущих
FWA	Fixed wireless access	Фиксированный беспроводный доступ
GF	Galois field	Поле Галуа
GKEK	Group key encryption key	Ключевой групповой ключ шифрации
GM	Grant management	Управление ресурсами (канальными)
GMH	Generic MAC header	Общий MAC заголовок
GPC	Grant per connection	Ресурс на соединение
GPRS	General Packet Radio Services	Обобщенные пакетные радиослужбы
GPS	Global positioning satellite	Глобальное спутниковое позиционирование
GPSS	Grant per subscriber station	Ресурсы на абонентскую станцию
GRE	Generic Routing Encapsulation	Общая инкапсуляция маршрутов (протокол Cisco)
GS	Guard symbol	Защитный символ
GSA	Group Security Association	Ассоциация групповой безопасности
GSM	Global System for Mobile communication	Глобальная система мобильной связи
GTEK	Group traffic encryption key	Групповой ключ шифрации трафика
GW	Gateway	Шлюз
HA	Home Agent	Домашний агент
HARQ	Hybrid Automatic Repeat reQuest	Гибридная система с автозапросом повторной передачи
HCS	Header check sequence	Проверочные символы заголовка
HEC	Header error check	Проверка ошибок заголовка
H-FDD	Half-duplex FDD	Полудуплекс
HHO	Hard Hand-Off	Жесткая передача управления (хендовер)
HiperMAN	High Performance Metropolitan Area Network	Высокопроизводительная городская (или региональная) сеть
HLA	Hot-Line Application или Hot-Lining Application	Линия прямой связи

HLD	Hot-Line Device Hot-lining Device	Устройство прямой связи
HMAC	Keyed-Hashing for Message Authentication Code	Хэширование ключа для кода аутентификации сообщения
HO	Hand-Off or Hand Over	Передача управления или обслуживания (хендовер)
HO ID	Handoff Identifier	Идентификатор хендовера
HoA	MS Home Address	Домашний адрес мобильной станции
HSDPA	High Speed Downlink Packet Access	Высокоскоростной пакетный доступ в нисходящем канале
HT	Header type	Тип заголовка
HTTP	HyperText Transfer Protocol	Протокол передачи гипертекста
HUMAN	High-Speed Unlicensed Metropolitan Area Network	Высокоскоростная безлицензионная городская (или региональная) сеть
I	inphase	Синфазный
IANA	Internet Assigned Numbers Authority	Полномочный орган по цифровым адресам в Интернете
IBS	Integrated Base Stations. Refers to a BS that can instantiate all the ASN functions for a given MS. Such an Integrated BS can also be labeled a Profile B ASN	Интегрированная базовая станция отсылает к базовой станции, которая может подтвердить все ASN-функции для данной мобильной станции. Такая интегрированная базовая станция может быть отмечена как ASN профиля B
ICMPv6	Internet Control Message Protocol for (IPv6) Specification [RFC 2463]	Интернет-протокол управления сообщениями для (IPv6) спецификации RFC 2463
IE	Information Element	Информационный элемент
IEEE	Institute of Electrical and Electronics Engineers	Институт инженеров по электротехнике и электронике
IEEE 802.3	IEEE standard specification for Ethernet	Спецификация стандарта IEEE для Ethernet
IETF	Internet Engineering Task Force	Рабочая группа разработки Интернета
IFFT	Inverse fast Fourier transform	Обратное быстрое преобразование Фурье
IGMP	Internet Group Management Protocol	Групповой Интернет-протокол управления
IID	Interface Identifier	Идентификатор интерфейса
IIP3	Input intercept point of third order	Входная точка пересечения третьего порядка
IK	Integrity Key	Надежный ключ
IKEv2	Internet Key Exchange protocol version 2	Интернет-протокол обмена ключами, версия 2
IMS	IP Multimedia Subsystem	Мультимедийная подсистема IP
IMSI	International Mobile Subscriber Identity	Международная идентификация мобильного абонента

IP	Internet Protocol	Интернет-протокол
IPsec	IP Security	IP-безопасность
IPv4	Internet Protocol Version 4	Интернет-протокол, версия 4
IPv6	Internet Protocol Version 6	Интернет-протокол, версия 6
IR	Incremental Redundancy	Пошагово возрастающая избыточность
ISF	Initial Service flow	Начальный сервисный поток
ISI	Inter-Symbol Interference	Межсимвольная интерференция
ITU	International Telecommunications Union	Международный союз электросвязи
IUC	Interval Usage Code	Код профиля пакетов
IWF	Internetworking Function	Межсетевая функция
IWG	Inter-working Gateway	Межсетевой шлюз
I-WLAN	Interworking with Wireless LANs	Межсетевое взаимодействие с беспроводной локальной сетью
IWU	Internetworking Unit	Межсетевой блок
KEK	Key encryption key	Ключ шифрования ключа
LAN	Local area network	Локальная сеть
LBS	Location Based Services	Сервисы, предоставляемые с учетом местоположения пользователя
LDPC	Low-Density-Parity-Check	Низкая плотность проверок на четность (коды)
LE	License-Exempt	Безлицензионные диапазоны
LFSR	Linear feedback shift registers	Регистр сдвига с линейной обратной связью
LLC	Logical link control	Управление логическим соединением (подуровень уровня звена данных модели OSI)
LMDS	Local multipoint distribution service	Локальная многоточечная вещательная служба
LOS	Line of sight	Прямая видимость
LPF	Local Policy Function	Функция локальной политики
LR	Location Register MSID, BSID	Регистр локализации (MSID, BSID)
LSB	Least Significant Bit	Младший бит
MAC	Medium access control layer	Уровень управления медиадоступом
MAI	Multiple Access Interference	Интерференция множественного доступа
MAN	Metropolitan area network	Городская или региональная сеть
MAP	Media Access Protocol	Протокол медиадоступа
MBMS	Multimedia Broadcast/ Multicast Service	Мультимедийная широковещательная/много-адресная услуга
MBS	Multicast and Broadcast Service	Многоадресная и широковещательная услуга
MCC	Mobile Country Code	Мобильный код страны

MCS	Modulation coding scheme	Сигнально-кодовая конструкция
MDHO	Macro Diversity handoff	Многомерная передача обслуживания
MIB	Management information base	База данных управляющей информации
MIC	Message integrity check	Проверка целостности сообщения
MIMO	Multiple input multiple output	Множественный вход — множественный выход
MIP	Mobile IP	Мобильный IP (относится и IPv4, и к IPv6)
MIP6	Mobile IP version 6	Мобильный Интернет-протокол, версия 6
MM	Mobility Management	Управление мобильностью
MMDS	Multichannel multipoint distribution service	Многоканальная многоточечная вещательная служба
MMS	Multimedia Message Service	Служба мультимедийных сообщений
MNC	Mobile Network operator Code	Код оператора мобильной сети
MPEG	Moving Pictures Experts Group	Экспертная группа подвижных изображений
MPLS	Multi Protocol Label Switching	Мультипротокольная коммутация на основе меток (признаков)
MS	Mobile Station или Mini-Slot	Мобильная станция или мини-слот
msb	Most significant bit	Старший бит
MSH	Mesh	Mesh-сеть
MSID	Mobile Station Identifier	Идентификатор мобильной станции
MSK	Master Session Key	Ключ мастер-сессии
MSO	Multi-Services Operator	Мультисервисный оператор
NA	Neighbor Advertisements	Объявление о состоянии соседей
NACK	Not Acknowledge	Сообщение неподтверждения
NAI	Network Access Identifier	Идентификатор сетевого доступа
NAP	Network Access Provider	Провайдер сети доступа
NAPT	Network Address Port Translation	Преобразование сетевого адреса и порта
NAS	Network Access Server	Сервер сетевого доступа
NAT	Network Address Translation	Преобразование сетевого адреса
NCFG	Network configuration	Конфигурация сети
NENT	Network entry	Вход в сеть
NLOS	Non line of sight	Отсутствие прямой видимости
NMS	Network Management System	Система управления сетью
NNI	Network-to-network interface (or network node interface)	Интерфейс сеть-сеть (или узел-узел)
NRM	Network Reference Model	Эталонная модель сети
nrtPS	Non-real-time Polling Service	Служба поллинга не реального времени
NS	Neighbor Solicitation	Запрос от соседа
NSP	Network Service Provider	Сетевой сервис-провайдер

NUD	Neighbor Unreachability Detection	Обнаружение недоступности соседа
OAM	Operations and Maintenance	Эксплуатация и техническое обслуживание
OFDM	orthogonal frequency division multiplexing	Ортогональное частотное мультиплексирование
OFDMA	orthogonal frequency division multiple access	Множественный доступ с ортогональным частотным разделением
OID	object identifier	Идентификатор объекта
OTA	Over-The-Air	По эфиру
OUI	Organization Unique Identifier	Уникальный идентификатор организации
PA	Paging Agent	Пейджинговый агент
PAK	primary authorization key	Основной ключ авторизации
PAPR	peak to average power ratio	Отношение пиковой к средней мощности
PBR	piggyback request	Запрос прибавки к уже выделенной полосе
PC	Paging Controller	Контроллер пейджинга
P-CSCF	Proxy-Call Session Control Function	Функция управления сессией прокси вызова
PDFID	packet data flow ID	Идентификатор пакета потока данных
PDG	Packet Data Gateway	Шлюз пакетов данных
PDH	plesiochronous digital hierarchy	Плезиохронная цифровая иерархия
PDU	protocol (или packet) data unit	Единица данных протокола (пакет, сформированный для передачи на более низкий уровень модели OSI или принимаемый с более низкого уровня. Например, на MAC-уровне PDU — это пакет MAC-уровня, передаваемый на физический уровень)
PEAP	Protected EAP	Защищенный наращиваемый протокол аутентификации
PER	Packet Error Rate	Пакетная вероятность ошибки
PF	Proportional Fair (Scheduler)	Достаточной пропорциональности (планировщик)
PF	Policy Function	Политика
PG	Paging Group	Группа пейджинга
PG ID	Paging Group Identifier	Идентификатор группы пейджинга
PHS	Payload Header Suppression	Удаление заголовков (их повторяющихся частей) из поля данных
PHSF	Payload Header Suppression Field	Зона удаления заголовков из поля данных

PHSI	Payload Header Suppression Index	Индекс удаления заголовков из поля данных (8-разрядная маска, показывающая, какие байты удаляются из заголовков, какие — нет)
PHSM	Payload Header Suppression Mask	Маска удаления заголовков из поля данных
PHSS	Payload Header Suppression Size	Размер поля заголовков, из которого производится удаление байтов (то же, что и размер поля PHSF)
PHSV	Payload Header Suppression Valid	Флаг, говорящий передающему устройству о необходимости верифицировать все байты заголовка, которые подлежат удалению
PHY	physical layer	Физический уровень
PKM	Privacy Key Management	Управление частным ключом
PM	poll-me bit	Бит запроса на опрос (поллинг)
PMD	physical medium dependent	Зависимый от среды передачи
PMIP	Proxy-Mobile IP	Протокол прокси-мобильного IP
PMK	Pairwise Master Key	Мастер-ключ для определения ключей более низкого уровня
PMN	Proxy Mobile Node	Мобильный узел прокси
PMP	point-to-multipoint	Точка-многоточка
PoA	Point of Attachment	Точка вложения
PPAC	prepaid accounting capability	Возможность организации предоплаты
PPC	Prepaid Client	Клиент предоплаты
ppm	parts per million	Одна миллионная часть
PPP	Point-to-Point Protocol	Протокол точка-точка
PPS	Prepaid Server	Сервер предоплаты
PRBS	pseudo random binary sequence	Псевдослучайная двоичная последовательность
Proxy-ARP	Proxy Address Resolution Protocol	Протокол прокси-преобразования адреса
PS	physical slot	Физический слот
PSH	Packing Subheader	Упаковочный заголовок
PSK	PreShared Key	Предварительно распределенный ключ
PSTN	Public Switched Telephone Network	Телефонная сеть общего пользования (ТФоП)
PTI	Payload Type Indicator	Индикатор типа полезной нагрузки
PtP	Peer to Peer	«Каждый с каждым» — соединение точка-точка или узел-узел, т. е. одноранговая сеть
PUSC	Partially Used Sub-Carrier	Использование части поднесущих

PUSC-ASCA	PUSC adjacent subcarrier allocation	PUSC со смежными поднесущими
PVC	Permanent Virtual Circuit	Перманентный виртуальный канал
Q	quadrature	Квадратура (квадратурная координата)
QAM	quadrature amplitude modulation	Кавадратурная амплитудная модуляция (КАМ)
QoS	Quality of Service	Качество обслуживания
QPSK	Quadrature phase-shift keying	Квадратурная фазовая модуляция (ФМ-4)
RA	Router Advertisement	Сообщение обнаружения маршрутизатора
RADIUS	Remote Access Dial In User Service	Сетевой протокол, обеспечивающий централизованные процедуры AAA для всех устройств сети
REQ	Request	Запрос
RLAN	Radio Local Access Network	Локальная сеть радиодоступа
RNG	Ranging	Ранжирование (первичная настройка на физические параметры трансивера БС)
RO	Route optimization	Оптимизация пути
RP	Reference Point	Опорная (базовая) точка
RR	Round Robin (Scheduler)	Карусельный (циклический) планировщик
RR	Resource-Reservation	Резервирование ресурса
RRA	Radio Resource Agent	Агент радиоресурсов
RRC	Radio Resource Controller	Контроллер радиоресурсов
RRI	Reverse Rate Indicator	Индикатор уменьшения скорости
RRM	Radio Resource Management	Управление радиоресурсами
RS	Reed–Solomon	Рида-Соломона (алгоритм)
RS	Router Solicitation	Принудительное назначение маршрутизатора
RSP	Response	Ответ
RSSI	Receive signal strength indicator	Индикатор уровня мощности принимаемого сигнала
RSVP	Resource Reservation Protocol	Протокол резервирования ресурсов
RTG	Receive/Transmit Transition Gap	Интервал между субкадрами приема и передачи
rtPS	Real-time Polling Service	Услуга поллинга в реальном времени
RUIM	Removable User Identity Module	Сменяемый идентификационный модуль пользователя, аналог SIM-карты в телефонах стандарта CDMA (IS-95a)
Rx	Reception	Прием

RxDS	Receiver Delay Spread Clearing Interval	Интервал «молчания» в конце последовательности пакетов в кадре, в течение которого передатчик выключается (как правило, снижая мощность по линейном закону), а приемник «собирает» пришедшие с различной задержкой замыкающие элементы последовательности пакетов
SA	Security Association	Набор параметров безопасности для БС и связанных с ней АС, поддерживающих защищенное соединение
SAE	System architecture evolution	Эволюция архитектуры системы, архитектура сети LTE
SAID	Security association identifier	Идентификатор набора параметров безопасности SA
SAP	Service access point	Сервисная точки доступа
SAR	Synthetic aperture radar	Радар с синтезированной апертурой
SC	Single carrier	Одиночная несущая
SCI	Spare capacity indicator	Индикатор дополнительной емкости
S-CSCF	Serving-Call Session Control Function	Сервер управления вызовами и сеансами, центральный узел сети IMS
SCTE	Society of Cable Telecommunications Engineers	Общество инженеров кабельной связи
SDFID	Service data flow ID	ID сервисного потока данных
SDH	Synchronous Digital Hierarchy	Синхронная цифровая иерархия
SDMA	Space (or Spatial) Division (or Diversity) Multiple Access	Множественный доступ с пространственным разделением (или разнесением)
SDU	Service data unit	Сервисный блок данных (пакет, передаваемый на верхний уровень модели OSI или принимаемый с верхнего уровня)
SF	Spreading Factor	Коэффициент расширения
SF	Service flow	Сервисный поток
SFA	Service Flow Authorization	Авторизация сервисного потока
SFID	Service Flow Identifier	Идентификатор сервисного потока
SFM	Service Flow Management	Управление сервисным потоком
SFN	Single Frequency Network	Одночастотная сеть
SGSN	Serving GPRS Support Node	Узел обслуживания абонентов GPRS
SHA	Secure hash algorithm	Безопасный шумовой алгоритм
SHO	Soft Hand Off	Мягкое переключение между секторами или базами (мягкий хендовер)

SI	Slip indicator	Индикатор скольжения (флаг, показывающий, что входной поток превысил допустимую глубину очереди приемника)
SI	Subscriber Identity	Идентичность пользователя
SII	System Information Identity or Service Identity Information	Идентичность информации системы или сервиса
SIM	Subscriber Identity Module	Модуль идентификации пользователя, интеллектуальные карты, используемые GSM-операторами
SIMO	Single Input Multiple Output	Одиночный вход — множественный выход
SINR	Signal to Interference + + Noise Ratio	Отношение сигнал / (помеха+шум)
SIR	signal-to-interference ratio	Отношение сигнал / помеха
SLA	Service Level Agreement	Соглашение об уровне сервиса
SM	Spatial Multiplexing	Пространственное мультиплексирование
SMS	Short Message Service	Служба коротких сообщений
SMTP	Simple Mail Transport Protocol	Простой почтовый транспортный протокол
SN	sequence number	Номер последовательности
SNIR	Signal to Noise + Interference Ratio	Отношение сигнал / (шум+помеха)
SNMP	Simple Network Management Protocol	Простой протокол управления сетью
SNR	signal-to-noise ratio	Отношение сигнал / шум
S-OFDMA	Scalable Orthogonal Frequency Division Multiple Access	Масштабируемый множественный доступ с ортогональным частотным разделением
SS	Subscriber Station	Абонентская станция
SS7	Signaling System 7	Система сигнализации № 7 (ОКС-7)
SSID	subscriber station identification (MAC address)	Идентификатор абонентской станции (MAC-адрес)
SSL	Secure Sockets Layer	Уровень безопасности соединений
SSTG	subscriber station transition gap	Интервал (в кадре) между зонами отдельных абонентских станций
STC	space time coding	Пространственно-временное кодирование
STTD	space time transmit diversity	Пространственно-временное разнесение на передаче
SUBC	Subscriber Credentials	Полномочия пользователя
SVC	Switched Virtual Circuit	Виртуальный коммутируемый канал
TBS	Target BS	Целевая БС

TC	Transmission convergence sublayer	Конвергентный подуровень передачи
TCM	Trellis coded modulation	Решетчатая сигнально-кодовая конструкция
TCP	Transmission Control Protocol	Протокол управления передачей
TDD	Time division duplex	Временной дуплекс
TDM	Time division multiplex	Временное мультиплексирование
TDMA	Time division multiple access	Множественный доступ с временным разделением
TE	Terminal Equipment	Терминальное оборудование
TEK	traffic encryption key	Ключ шифрования трафика
TFTP	Trivial File Transfer Protocol	Тривиальный протокол передачи файлов (тривиальный FTP)
TLS	Transport Layer Security, a variant of SSL	Безопасность транспортного уровня, как вариант SSL
TLV	Type-length-value	Тип – длина – значение
TTG	Transmit/Receive Transition Gap	Интервал между субкадрами передачи и приема
TTI	Transmission Time Interval	Временной интервал передачи
TTLS	Tunneled TLS	Туннельный TLS
TU	Typical Urban (as in channel model)	Типичная городская застройка (в канальной модели)
TUSC	Tile usage of subchannels	Мозаичное использование подканалов
Tx	Transmission	Передача
UCD	Uplink Channel Descriptor	Дескриптор восходящего канала
UDP	User Datagram Protocol	Протокол дейтаграмм пользователя
UE	User Equipment	Оборудование пользователя
UEP	Unequal error protection	Неравная кодовая защита
UGS	Unsolicited Grant Service	Доступ к каналу без запроса
UICC	Universal Integrated Circuit Card	Универсальная карта на интегральной схеме, мультисервисная платформа, позволяющая на одной и той же карте запускать различные для смарт-карт приложения. Стандарт USIM-карт
UID	User-identity	Идентификация пользователя
UIUC	Uplink Interval Usage Code	Код профиля пакетов в восходящем канале
UL	uUplink	Восходящий канал (к базовой станции)
UMTS	Universal Mobile Telecommunications System	Универсальная мобильная телекоммуникационная система
UNI	User-to-network interface	Интерфейс пользователь-сеть
U-NII	Unlicensed National Information Infrastructure	Национальная информационная безлицензионная инфраструктура

USIM	Universal Subscriber Identity Module	Универсальный идентификационный модуль абонента. Интеллектуальные карты, используемые UMTS-операторами
UTC	Coordinated Universal Time	Всеобщее скоординированное время
UW	Unique Word	Уникальное слово
VC	virtual channel	Виртуальный канал
VLAN	Virtual LAN	Виртуальная локальная сеть
VoIP	Voice over Internet Protocol	Голос через IP
VPN	Virtual Private Network	Виртуальная частная сеть
VSA	Vendor Specific Attributes	Специфические атрибуты производителя
VSF	Variable Spreading Factor	Переменный коэффициент расширения
VSM	Vertical Spatial Multiplexing	Вертикальное пространственное мультиплексирование
WAG	WLAN Access Gateway	Шлюз доступа беспроводной локальной сети
WAP	Wireless Application Protocol	Беспроводный протокол приложений
WATSP	WiMAX ASN Transport Signaling Protocols	Транспортные протоколы сигнализации ASN-сети WiMAX
WCDMA	Wideband Code-Division Multiple Access	Широкополосный множественный доступ с кодовым разделением
WEP	Wired Equivalent Privacy	Конфиденциальность, эквивалентная проводной (т. е. в проводных сетях)
WiBro	Wireless Broadband (Service)	Беспроводной широкополосный сервис (стандарт компании Samsung)
Wi-Fi	Wireless Fidelity	Достоверность данных. Международная организация по сертификации на совместимсть устройств стандартов 802.11 a / b / g
WiMAX	Worldwide Interoperability for Microwave Access	Всемирная совместимость при высокочастотном доступе
WLAN	Wireless local area network	Беспроводная локальная сеть
WPA	Wi-Fi Protected Access	Wi-Fi защищенный доступ
WWAN	Wireless Wide Area Network	Беспроводная распределенная сеть
X.509	ITU standard for digital public-key certificate issued by a CA	Стандарт ITU для цифрового сертификата системы с открытыми ключами, изданный CA