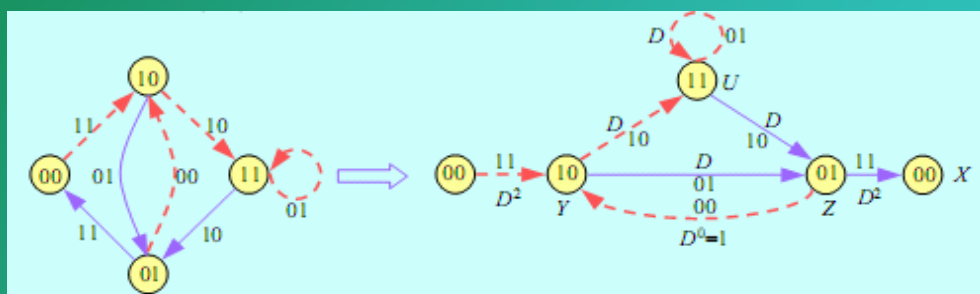


Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное
учреждение высшего профессионального образования

Новгородский государственный университет
имени Ярослава Мудрого

Рассветалов Л.А.

РАДИОТЕХНИЧЕСКИЕ СИСТЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ



Великий Новгород
2013 г.

Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования

Новгородский государственный университет
имени Ярослава Мудрого

Рассветалов Л.А.

РАДИОТЕХНИЧЕСКИЕ СИСТЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

УЧЕБНОЕ ПОСОБИЕ

Великий Новгород
2013 г.

ББК

УДК 621.396; 681.3.07

Рассветалов Л.А. Радиотехнические системы передачи информации: Учебное пособие /

ФГБОУ «Новгородский государственный университет им. Ярослава Мудрого», Великий Новгород, 2013 г. - 195 с.

Рецензенты: д-р техн. наук, проф. Голик Ф.В.

В учебном пособии рассмотрены вопросы: представление информации в цифровом виде, каналы передачи цифровой информации, цифровые модуляционные форматы и модуляторы, кодирование в цифровых системах связи, синхронизация в радиосистемах связи, организация множественного доступа, широкополосные сигналы в системах связи, широкополосные методы борьбы с многолучевостью, сотовые системы мобильной связи, технологии беспроводной передачи интернет-данных.

Учебное пособие отвечает новым образовательным стандартам и предназначено для подготовки магистров по направлению 210400.68 «Системы и устройства передачи, приема и обработки сигналов».

Учебное пособие одобрено советом института Электронных и Информационных Систем Новгородского государственного университета имени Ярослава Мудрого.

© Федеральное государственное бюджетное образовательное
учреждение высшего профессионального образования
Новгородский государственный университет
имени Ярослава Мудрого, 2013

Оглавление

Предисловие.....	7
ВВЕДЕНИЕ.....	8
Классификация систем передачи информации	9
1. Структурная схема ЦСС и преобразования сигналов.....	10
1.1. Сообщения, знаки и символы.....	13
Основная терминология цифровой связи.....	13
1.1.1. Пример сообщений, знаков и символов	14
1.1.2. Цифровые и аналоговые критерии производительности	16
1.1.3. Основные характеристики ЦСС.....	16
1.1.4. Предельные возможности ЦСС. Выбор вида сигналов.....	17
1.2. Выбор основания системы счисления	19
1.3. Основные коды	20
1.4. Модели каналов	23
Обобщенная схема ЦСПИ	23
1.4.1. Модели непрерывных каналов	26
1.4.2. Модели дискретных каналов.....	30
2. Модуляция.....	35
2.1. Общие положения	35
2.2. Энергетическая и частотная эффективность	36
2.3. Квадратурная амплитудная манипуляция.....	38
2.4. Квадратурная фазовая манипуляция со сдвигом.....	43
2.5. МЧМ – минимальная частотная манипуляция	44
2.6. Модуляторы КАМ высоких порядков	47
2.7. OFDM.....	49
2.8. Треллис-модуляция	50
3. Кодирование в ЦСС	53
3.1. Корректирующие коды	53
3.2. Линейные коды	54
3.3. Совершенные и квазисовершенные коды	55
3.4. Метод перемежения	56
3.5. Системы с обратной связью	56
3.6. Мягкое и жесткое декодирование.....	57
3.7. Сверточные коды.....	58
3.8. Алгоритм декодирования Витерби	65
3.9. Декодирование с мягкими решениями.....	75
3.10. Сверточные коды с выкалыванием.....	77
Практическое использование сверточных кодов	77
4. СИНХРОНИЗАЦИЯ В РАДИОСИСТЕМАХ ПЕРЕДАЧИ ДИСКРЕТНОЙ ИНФОРМАЦИИ	79
4.1. Принципы построения и основные характеристики систем синхронизации....	79
4.2. Влияние точности оценки синхропараметров на качество работы систем	80
4.3. Фазовая синхронизация	82
4.3.1. Фазовая автоподстройка частоты	82
4.3.2. Устройства фазовой синхронизации	84
4.4. Тактовая синхронизация.....	86
4.5. Цикловая синхронизация.....	89
4.6. Кадровая синхронизация	91
4.7. Синхронизация в системах с широкополосными сигналами.....	94
5. Широкополосные сигналы в системах связи.....	98
5.1. Узкополосная помеха.....	98

5.2.	Электромагнитная совместимость.....	100
5.3.	Энергетическая скрытность.....	101
5.4.	Структурная скрытность сигнала	105
5.5.	Дискретные широкополосные сигналы.....	106
5.6.	Псевдослучайные кодовые последовательности.....	110
5.6.1.	Основные характеристики.....	110
5.6.2.	Классы псевдослучайных последовательностей	112
5.6.2.1.	М-последовательности.....	112
5.6.2.2.	Последовательности Голда.....	113
5.6.2.3.	Последовательности Касами	114
5.6.2.4.	Нелинейные последовательности	114
5.6.3.	Практическое применение ШПС в системах связи и управления.....	117
5.6.3.1.	Сотовые системы подвижной радиосвязи.....	118
5.6.3.2.	Локальные радиосети	119
5.6.3.3.	Спутниковые радионавигационные системы (СРНС)	119
5.6.4.	Широкополосные методы борьбы с многолучевостью	120
5.6.4.1.	Оптимальный приемник при многолучевом распространении сигнала..	122
5.6.4.2.	Оптимальный приемник многолучевого сигнала. Случай известных запаздываний	124
5.6.4.3.	Оптимальный приемник многолучевого сигнала. Случай неизвестных запаздываний	125
5.6.4.4.	Приемник ОФМ-сигналов	127
5.6.4.5.	Приемники, зондирующие канал.....	129
5.6.4.6.	Межсимвольная интерференция.....	132
5.6.4.7.	Приемник с последетекторным интегрированием сигналов.....	134
5.6.4.8.	Цифровой вариант приемника «Рейк»	135
5.6.4.9.	Установка пакетной радиосвязи (УПР).....	137
6.	ОРГАНИЗАЦИЯ МНОЖЕСТВЕННОГО ДОСТУПА	140
6.1.	Множественный доступ с частотным разделением каналов.....	140
6.2.	Множественный доступ с временным разделением каналов.....	141
6.3.	Множественный доступ с кодовым разделением каналов	143
6.3.1.	Варианты реализации кодового разделения	143
7.	Системы мобильной связи.....	144
7.1.	Структура сетей мобильной связи	144
7.2.	Классификация и эволюция систем мобильной радиосвязи.....	145
7.2.1.	Основные типы систем мобильной связи	145
7.2.2.	Транкинговые системы связи.....	145
7.2.3.	Системы персонального радиовызова.....	147
7.3.	Сотовые системы мобильной связи	149
7.3.1.	Понятие о модели взаимодействия открытых систем	154
7.3.2.	Эволюция систем и стандартов сотовой связи	156
7.3.3.	Описание ССМС на примере GSM.....	158
7.3.4.	Структура TDMA кадров и формирование сигналов в стандарте GSM	162
7.4.	Радиоинтерфейс системы cdmaOne (IS-95).....	165
7.4.1.	Общая характеристика системы CDMAOne	165
7.4.2.	. Архитектура линии "вниз"	168
7.4.3.	Пилотный канал.....	169
7.4.4.	Канал синхронизации.....	171
7.4.5.	Регулировка мощности в стандарте IS-95.....	173
7.5.	Деятельность МСЭ по развитию сетей UMTS	175
7.6.	характеристика 3G – 4G поколений ССМС	179
7.6.1.	Четвертое поколение МС.....	183

7.6.2. Сравнение Wi-Fi и WiMAX	186
Список литературы.....	188

Предисловие

Обилие учебников и монографий, посвященных передаче цифровой информации радиотехническими методами, не стимулирует, казалось бы, написание учебного пособия по столь популярной теме, тем более что оно носит компилятивный характер. Главным побудительным мотивом составителя пособия явилось появление нового образовательного стандарта специальности «Радиотехника», в соответствии с которым в Новгородском университете курсу «Радиотехнические системы передачи информации» предшествуют дисциплины «Статистическая теория РТС», «Статистическая теория связи», «Теория информации и кодирование» и «Радиотехнические системы». Существующие на сегодняшний день книги учебной направленности по цифровым системам передачи информации содержат, как правило, разделы цитированных дисциплин наряду с информацией более общего характера. В результате такие методически прекрасно написанные книги, как [1, 3], имеют весьма большой объем, который зачастую просто отпугивает студентов. Поэтому составитель взял на себя смелость выбора из различных источников отдельных фрагментов и написания связывающую их часть текста, предполагая получить компактный объем настоящего пособия. Нет нужды подчеркивать, что выбор источников отражает личные пристрастия составителя.

Пособие состоит из семи глав. Первая глава содержит общие сведения о структуре цифровой системы связи (ЦСС), используемых сигналах и кодах, моделях непрерывных и дискретных каналов. В этой главе использованы материалы популярных учебников [1, 3, 4].

Во второй главе рассматриваются модуляционные форматы цифровой связи и их аппаратная реализация. Описание базируется на материалах [1,2,5].

В третьей главе рассмотрены основные методы кодирования, используемые в ЦСС. Используются материалы [2,5,6].

Четвертая глава содержит сведения о системах синхронизации в ЦСС: фазовой, тактовой, цикловой и кадровой. Отдельно рассмотрены вопросы синхронизации в системах с широкополосными сигналами. Здесь использованы материалы [1,2,3,5].

В пятой главе рассмотрены вопросы применения широкополосных сигналов (ШПС) в ЦСС: новые свойства, приобретаемые системами связи при их использовании, классы дискретных широкополосных сигналов. При описании общих вопросов использованы фрагменты книг [2,5]. Основное внимание уделено псевдослучайным последовательностям. Рассмотрены (по материалам обзора [7]) широкополосные методы борьбы с многолучевостью, структура оптимального приемника, которая иллюстрируется описанием установки пакетной радиосвязи [8] с остроумным способом создания опорного напряжения для детектирования широкополосных сигналов с абсолютной фазовой манипуляцией в многолучевом канале подвижной радиосвязи.

Шестая глава посвящена организации множественного доступа с частотным, временным и кодовым разделением каналов.

Седьмая глава «Системы мобильной связи» написана по материалам [2,5,10,11], а также многочисленных статей на соответствующих сайтах Интернета.

ВВЕДЕНИЕ

Телекоммуникации являются одной из наиболее быстро развивающихся областей современной науки и техники. Жизнь современного общества уже невозможно представить без тех достижений, которые были сделаны в этой отрасли за немногим более ста лет развития. Отличительная особенность нашего времени - непрерывно возрастающая потребность в передаче потоков информации на большие расстояния. Это обусловлено многими причинами, и в первую очередь тем, что связь стала одним из самых мощных рычагов управления экономикой страны. Одновременно, претерпевая значительные изменения, становясь многосторонней и всеобъемлющей, электросвязь каждой страны становится все более интегрированной в мировое телекоммуникационное пространство. Радиотехнические системы передачи информации (РТСПИ) широко применяются также для передачи команд управления, телеметрической информации, визуальной информации и т.д.

Краткий исторический очерк. Теоретическими основами радиосвязи можно считать «Трактат об электричестве и магнетизме», опубликованный Д.К. Максвеллом в 1873 г., из которого следовало, что любой проводник с током излучает в пространство электромагнитные волны, распространяющиеся со скоростью света.

В 1887 г. Г.Герц экспериментально доказал существование электромагнитных волн, а в 1895 г. А.С. Попов и Г. Маркони использовали их для беспроводной передачи информации.

Изобретатель электрического телеграфа и азбуки Морзе (1837г.). Самюэль Морзе (S.Morse) предложил двоичный код переменной длины, в котором буквы английского алфавита представлены последовательностью точек и тире (кодовые слова). В этом коде часто встречающиеся буквы представлены короткими кодовыми словами, тогда как появляющиеся редко – длинными. Таким образом, код Морзе предвосхитил методы экономного неравномерного кодирования.

В 1875 г. Эмиль Бодо (E.Baudot) изобрел буквопечатающий телеграф и код, в котором каждая буква кодировалась двоичным кодом фиксированной длины 5. В коде Бодо элементы двоичного кода имеют равную длину и именуются посылкой и паузой.

Следующим этапом в развитии радиосвязи является переход от радиотелеграфии к радиотелефонии, развитие которой началось с появлением электронной вакуумной лампы – основы генераторов и усилителей высокочастотных электромагнитных колебаний.

Развитию радиосвязи способствовали работы выдающихся ученых Найквиста (1924) и Хартли (1928), положивших начало современной теории цифровой связи, Шеннона (1947 – 1948 гг.), установившего математические основы передачи информации по каналам связи и фундаментальные ограничения для систем цифровой связи, Котельникова (1953), одного из создателей теории потенциальной помехоустойчивости.

В настоящем курсе излагаются идеи и технологии, являющиеся фундаментальными для систем цифровой связи. Последние становятся (и стали) все более привлекательными из-за того, что цифровая передача предлагает возможности обработки информации, недоступные при использовании аналоговой передачи.

Отличительной особенностью систем цифровой связи – СИЦ (DCS – digital communication system) является передача конечного набора элементарных сигналов за конечный промежуток времени. Задачей приемника при этом является не точное воспроизведение переданного сигнала, а определение, какой именно сигнал из конечного набора был передан.

В аналоговых системах сигнал состоит из бесконечного множества элементарных сигналов, что вызывает значительные трудности достоверного приема при наличии помех.

Основным преимуществом цифрового подхода является легкость восстановления пораженных помехой цифровых сигналов по сравнению с аналоговыми. Так, при

двоичном представлении информации помеха должна быть достаточно мощной, чтобы трансформировать сигнал из одного состояния в другое.

Цифровые каналы менее подвержены искажению и интерференции, т.к. наличие всего двух состояний облегчает восстановление сигнала и предотвращает накопление шумов в процессе передачи. Применение процедур выявления и коррекции ошибок делают возможным высокую точность передачи сообщения. С аналоговыми технологиями подобные процедуры недоступны.

Существуют и другие важные преимущества цифровой связи. Цифровые каналы надежнее и имеют более низкую стоимость по сравнению с аналоговыми. Цифровое программное обеспечение допускает более гибкую реализацию, чем аналоговое (микропроцессоры, цифровые коммутаторы и большие интегральные схемы – БИС). Уплотнение каналов с временным разделением – ВРК (time-division multiplexing – TDM) проще уплотнения с частотным разделением – ЧРК (frequency-division multiplexing – FDM) в аналоговых системах. При передаче различные типы цифровых сигналов (данные, телеграф, телефон, телевидение) могут рассматриваться как идентичные – бит есть бит.

Классификация систем передачи информации

Современные радиотехнические системы передачи информации (РТСПИ), отличающиеся большим разнообразием, могут классифицироваться по многим признакам.

- По числу каналов – одноканальные и многоканальные.
- По режиму использования канала – *системы односторонней связи (симплексные), системы двусторонней связи (дуплексные) и полудуплексные системы*. В первых связь возможна только в одном направлении, во вторых - одновременно в двух направлениях (как в телефоне), в третьих – передача и прием ведется поочередно.
- По виду передаваемых сообщений – *системы передачи дискретных и непрерывных сообщений*.
- По назначению – *телефонные, телеграфные, фототелеграфные (факсимильные), телевизионные, телеметрические, системы телеуправления, системы передачи данных*.
- По механизму распространения радиоволн – *ионосферные, тропосферные, метеорные и космические*.

1. Структурная схема ЦСС и преобразования сигналов

Функциональная схема, приведенная на рис. 1.1, иллюстрирует распространение сигнала и этапы его обработки в типичной системе цифровой связи (ЦСС).

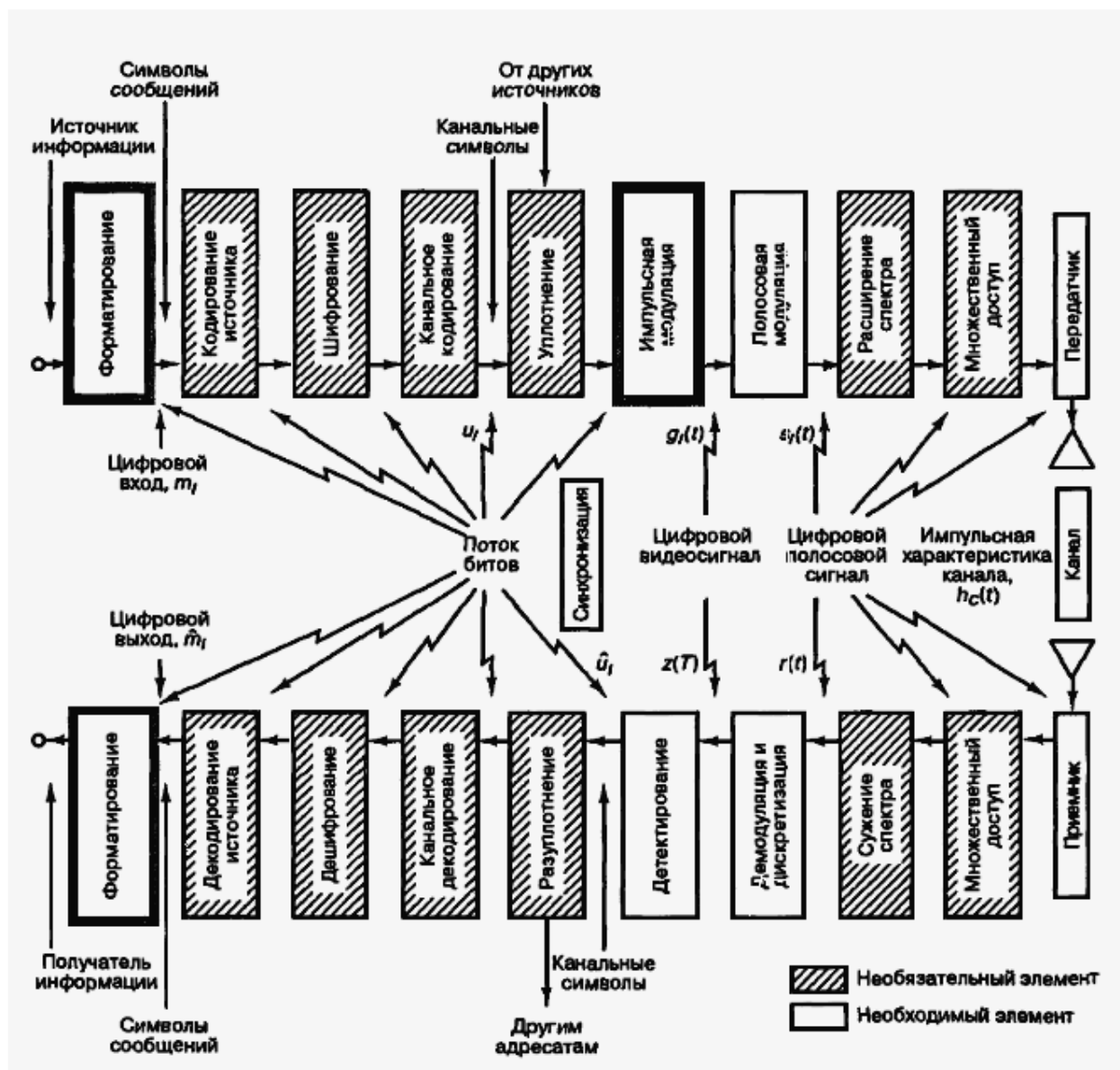


Рис. 1.1. Общая структурная схема цифровой связи.

Верхние блоки – форматирование, кодирование источника, шифрование, канальное кодирование, уплотнение, импульсная модуляция, полосовая модуляция, расширение спектра и множественный доступ — отражают преобразования сигнала на пути от источника к передатчику. Нижние блоки— преобразования сигнала на пути от приемника к получателю информации, и, по сути, они противоположны верхним блокам, блоки модуляции и демодуляции/детектирования вместе называются модемом. Термин "модем" часто объединяет несколько этапов обработки сигналов, показанных на рис. 1.1; в этом случае модем можно представлять как «мозг» системы. Передатчик и приемник можно рассматривать как "мускулы" системы. Для беспроводных приложений передатчик состоит из схемы повышения частоты в область радиочастот, усилителя мощности и

антенны, а приемник – из антенны и малошумящего усилителя. Обратное понижение частоты производится на выходе приемника и/или демодулятора.

На рис. 1.1 исходная информация преобразуется в двоичные цифры (*биты*); после этого биты группируются в *цифровые сообщения* или *символы сообщений*. Каждый такой символ (m_i , где $i = 1, \dots, M$) можно рассматривать как элемент конечного алфавита, содержащего M элементов. Следовательно, для $M = 2$ символ сообщения m_i является бинарным (т.е. состоит из одного бита).

Рассмотрим ключевые блоки обработки сигналов, изображенные на рис. 1.1. Необходимыми для систем цифровой связи являются только этапы форматирования, модуляции, демодуляции/детектирования и синхронизации.

Форматирование преобразовывает исходную информацию в биты, обеспечивая, таким образом, совместимость информации и функций обработки сигналов с системой ЦСС. С этой точки рисунка и вплоть до блока импульсной модуляции информация остается в форме *потока битов*.

Модуляция — это процесс, посредством которого символы сообщений или каналные символы (если используется каналное кодирование) преобразуются в *сигналы*, совместимые с требованиями, налагаемыми каналом передачи данных. Канальное *не помехоустойчивое* кодирование используется в многоканальных системах, передающих сообщения от нескольких источников одновременно. В этом случае цель канального кодирования – придание сигналам различных каналов свойств, которые позволяют на приемной стороне выделить нужный канал из *группового* сигнала. Обычно канальное кодирование осуществляется с помощью *канальных поднесущих*, отличающихся частотой или формой колебаний. Подвергнувшиеся канальному кодированию сигналы разных источников поступают в блок *уплотнения*, где формируется групповой сигнал. *Помехоустойчивое* кодирование может осуществляться как при канальном кодировании, так и над групповым сигналом. Специфические вопросы многоканальной передачи будут рассмотрены отдельно; дальнейшее описание предполагает наличие одного источника (канала).

Импульсная модуляция — это еще один необходимый этап, поскольку каждый символ, который требуется передать, вначале нужно преобразовать из двоичного представления в *видеосигнал* (*модулированный сигнал*). Блок импульсно-кодовой модуляции обычно включает фильтрацию с целью достижения минимальной полосы передачи. При использовании импульсной модуляции для обработки двоичных символов результирующий двоичный сигнал называется ИКМ-сигналом (ИКМ – импульсно-кодовая модуляция, в англоязычной литературе – PCM - pulse-code modulation). Существует несколько типов ИКМ-кодированных сигналов; в приложениях телефонной связи эти сигналы часто называются *кодами канала*. При применении импульсной модуляции к не двоичным символам результирующий сигнал именуется *многопозиционным* (М-арным, М-ичным) импульсно-модулированным. Существует несколько типов подобных сигналов, основное внимание будем уделять *амплитудно-импульсной модуляции* АИМ (англ: pulse-amplitude modulation — PAM).

После импульсной модуляции каждый символ сообщения или каналный символ принимает форму видеосигнала $g_i(t)$, где $i = 1, \dots, M$. В любой электронной реализации поток битов, предшествующий импульсной модуляции, представляется уровнями напряжений. Может возникнуть вопрос, почему существует отдельный блок для импульсной модуляции, когда фактически уровни напряжения для двоичных нулей и единиц уже можно рассматривать как идеальные прямоугольные импульсы, длительность каждого из которых равна времени передачи одного бита? Существует два важных отличия между подобными уровнями напряжения и видеосигналами, используемыми для модуляции. Во-первых, блок импульсной модуляции позволяет использовать бинарные и М-арные сигналы. Во-вторых, фильтрация, производимая в блоке импульсной модуляции,

формирует отличные от прямоугольных импульсы, длительность которых больше времени передачи одного бита. Фильтрация позволяет использовать импульсы большей длительности; таким образом, импульсы расширяются на соседние временные интервалы передачи битов. Этот процесс иногда называется формированием импульсов; он используется для поддержания полосы передачи в пределах некоторой желаемой области спектра.

Для систем передачи радиочастотного диапазона следующим важным этапом является *полосовая модуляция*; она необходима всегда, когда среда передачи не поддерживает распространение видеосигналов, а требует радиочастотного сигнала $s_i(t)$, где $i = 1, \dots, M$. Далее сигнал s_i проходит через канал, причем связь между входным и выходным сигналами канала полностью определяется *импульсной характеристикой канала* $h_c(t)$. Кроме того, в различных точках вдоль маршрута передачи дополнительные случайные шумы искажают сигнал, так что сигнал на входе приемника $y(t)$ отличается от переданного сигнала $s_i(t)$:

$$y(t) = s_i(t) * h_c(t) + n(t), \quad i = 1, \dots, M \quad (1.1)$$

где знак «*» представляет собой операцию свертки, а $n(t)$ — случайный процесс.

При обработке полученного сигнала в принимающем устройстве входной каскад приемника и/или демодулятор обеспечивают понижение частоты каждого полосового сигнала $y(t)$. В качестве подготовки к детектированию демодулятор восстанавливает $y(t)$ в виде оптимальной огибающей видеосигнала $z(t)$. Обычно с приемником и демодулятором связано несколько фильтров — фильтрация производится для удаления нежелательных высокочастотных составляющих (в процессе преобразования полосового сигнала в видеосигнал) и формирования импульса. Выравнивание можно описать как разновидность фильтрации, используемой в демодуляторе (или после демодулятора) для удаления всех эффектов ухудшения качества сигнала, причиной которых мог быть канал. Выравнивание (*англ.*: equalization) необходимо в том случае, если импульсная характеристика канала $h_c(t)$ настолько плоха, что принимаемый сигнал сильно искажен.

Эквалайзер (устройство выравнивания) реализуется для компенсации (т.е. для удаления или ослабления) всех искажений сигнала, вызванных неидеальной характеристикой $h_c(t)$. И последнее, на этапе дискретизации сформированный импульс $z(t)$ преобразовывается в выборку $z(T)$ для восстановления (приблизительно) символа канала и/или символа сообщения m_i (если не используется канальное кодирование). Некоторые авторы используют термины "демодуляция" и "детектирование" как синонимы. Здесь под *демодуляцией* подразумевается восстановление сигнала (полосового импульса), а под *детектированием* — принятие решения относительно цифрового значения этого сигнала. Далее, однако, будем считать, что обе операции — демодуляция и детектирование выполняются в одном устройстве — демодуляторе.

Остальные этапы обработки сигнала в модеме являются необязательными и направлены на обеспечение специфических системных нужд.

Кодирование источника — это преобразование аналогового сигнала в цифровой (для аналоговых источников) и удаление избыточной (ненужной) информации, т.е. *статистическое кодирование*. Отметим, что типичная система ЦСС может использовать либо *кодирование источника* (для оцифровывания и сжатия исходной информации), либо более простое *форматирование* (только для оцифровывания). Система не может одновременно применять и кодирование источника, и форматирование, поскольку первое уже включает необходимый этап оцифровывания информации.

Шифрование, которое используется для обеспечения секретности связи, предотвращает понимание сообщения несанкционированным пользователем и введение в систему ложных сообщений.

Канальное кодирование при данной скорости передачи данных может снизить вероятность ошибки P_E (помехоустойчивое кодирование) или уменьшить отношение сигнал/шум,

необходимое для получения желаемой вероятности P_E за счет увеличения полосы передачи или усложнения декодера.

Процедуры *уплотнения* и *множественного доступа* объединяют сигналы, которые могут иметь различные характеристики или могут поступать от разных источников, с тем, чтобы они могли совместно использовать часть ресурсов связи (например, спектр, время).

Расширение полосы частот может давать сигнал, относительно неуязвимый для узкополосной помехи (как естественной, так и умышленной), и может использоваться для повышения конфиденциальности сеанса связи. Также оно является ценной технологией, используемой для множественного доступа и работы в многолучевых каналах.

Блоки обработки сигналов, показанные на рис. 1.1, представляют типичную функциональную схему системы цифровой связи; впрочем, эти блоки иногда реализуются в несколько ином порядке. Например, уплотнение может происходить до канального кодирования *или* модуляции *либо* — при двухэтапном процессе модуляции (поднесущая и несущая) — оно может выполняться между двумя этапами модуляции. Подобным образом блок расширения частоты может находиться в различных местах верхнего ряда рис. 1.1; точное его местонахождение зависит от конкретной используемой технологии. Синхронизация и ее ключевой элемент, синхронизирующий сигнал, задействованы во всех этапах обработки сигнала в системе ЦСС. Для простоты блок синхронизации на рис. 1.1 показан безотносительно к чему-либо, хотя фактически он участвует в регулировании операций практически в каждом блоке, приведенном на рисунке.

1.1. Сообщения, знаки и символы

Основная терминология цифровой связи

Ниже приведены некоторые основные термины, часто используемые в области цифровой связи.

Источник информации. Устройство, передающее информацию посредством системы ЦСС. Источник информации может быть *аналоговым* или *дискретным*. Выход аналогового источника может иметь любое значение из непрерывного диапазона амплитуд, тогда как выход источника дискретной информации — значения из конечного множества амплитуд. Источники аналоговой информации преобразуются в источники цифровой информации посредством *дискретизации* или *квантования*. Методы дискретизации и квантования называются форматированием и кодированием источника.

Текстовое сообщение. Последовательность символов. При цифровой передаче данных сообщение представляет собой последовательность цифр или символов, принадлежащих конечному набору символов или алфавиту.

Знак. Элемент алфавита или набора символов. Знаки могут представляться последовательностью двоичных цифр. Существует несколько стандартизованных кодов, используемых для знакового кодирования, в том числе код ASCII (American Standard Code for Information Interchange — Американский стандартный код для обмена информацией), код Холлерита (Hollerith code), код Бодо (Baudot code), код Муррея (Murray code) и код (азбука) Морзе (Morse code).

Двоичная цифра — *бит* (binary digit — bit). Фундаментальная единица информации для всех цифровых систем. Термин "бит" также используется как единица объема информации. *Поток битов* (bit stream) — последовательность двоичных цифр (нулей и единиц). Поток битов часто называют *видеосигналом*, или *низкочастотным сигналом*.

Символ (цифровое сообщение). Символ — это группа из k бит, рассматриваемых как единое целое. Далее мы будем называть этот блок *символом сообщения* m_i ($i = 1, \dots, M$) из конечного набора символов или алфавита (рис. 1.2) Размер алфавита M равен 2^k , где k — число битов в символе. При *низкочастотной* передаче каждый из сигналов m_i будет представлен одним из набора видеопульсов $g_1(t), g_2(t), \dots, g_M(t)$. Иногда при передаче последовательности таких импульсов для выражения скорости передачи импульсов

(скорости передачи символов) используется единица *бод* (baud), характеризующая *техническую* скорость передачи. Для типичной *полосовой* передачи каждый импульс $g_i(t)$ будет представляться одним из набора полосовых импульсных сигналов $s_1(t), s_2(t), \dots, s_M(t)$. Таким образом, для беспроводных систем символ m_i посылается путем передачи цифрового сигнала $s_i(t)$ в течение T секунд (T — длительность символа). Следующий символ посылается в течение следующего временного интервала, T . То, что набор символов, передаваемых системой ЦСС, является конечным, и есть главное отличие этих систем от систем аналоговой связи. Приемник ЦСС должен всего лишь определить, какой из возможных M сигналов был передан; тогда как аналоговый приемник должен точно определять значение, принадлежащее непрерывному диапазону сигналов.

Цифровой сигнал. Описываемый уровнем напряжения или силы тока, сигнал (импульс — для низкочастотной передачи или синусоида — для полосовой передачи), представляющий цифровой символ. Характеристики сигнала (для импульсов — амплитуда, длительность и положение или для синусоиды — амплитуда, частота и фаза) позволяют его идентифицировать как один из символов конечного алфавита.

Скорость передачи данных. Эта величина в битах в секунду (бит/с) дается формулой $R = k/T = (1/T) \log_2 M$ (бит/с), где k бит определяют символ из $M = 2^k$ -символьного алфавита, а T — это длительность k -битового символа.

Текстовые сообщения состоят из последовательности буквенно-цифровых знаков. При цифровой передаче знаки вначале кодируются в последовательность битов, которая называется *поток*ом битов, или *видеосигналом*. После этого формируются группы из k бит, именуемые *символами*, причем число всех символов конечно ($M = 2^k$), а их совокупность называется *алфавитом*. Система, использующая символьный набор размера M , называется *M-арной* (*M-ичной*). Выбор величины k или M является важным первоначальным этапом проектирования любой цифровой системы связи. При $k = 1$ система является *бинарной*, размер набора символов равен $M = 2$, а модулятор использует один из двух различных сигналов для представления двоичного значения "один", а другой — для представления двоичного значения "ноль". В этом частном случае символ и бит — это одно и то же. При $k = 2$ система именуется *четверичной*, или *4-уровневой* ($M = 4$). В каждый момент формирования символа модулятор использует один из четырех возможных сигналов для представления символа. Разделение последовательности битов сообщения определяется размером алфавита M . Ниже приведен пример, который поможет лучше понять связь между следующими терминами: "сообщение", "знак", "символ", "бит" и "цифровой сигнал".

1.1.1. Пример сообщений, знаков и символов

На рис. 1.2 приведен пример разбиения потока битов, определяемого спецификацией системы для различных значений k и M . Текстовое сообщение на рисунке — это слово "ДУМАЙ!". Использование 8-битовой кодировки ANSI дает поток битов, состоящий из 48 бит. Кодировка ANSI Windows 1251 приведена в Таблице 1.

Символы 0 – 127 совпадают с кодировкой ASCII, приведенной в Приложении 1.

На рис. 1.2, *a* размер набора символов, M , был выбран равным 8 (каждый символ представляет восьмеричное число). Таким образом, биты группируются по три ($k = \log_2 8$); полученные в результате 16 чисел представляют 16 готовых к передаче восьмеричных символов.

Таблица 1. Кодировка Windows 1251

128 Ъ	144 ђ	160	176 .	192 А	208 Р	224 а	240 р
129 Ѓ	145 '.	161 ѱ	177 ±	193 Б	209 С	225 б	241 с
130 .	146 '.	162 ѱ	178 і	194 В	210 Т	226 в	242 т
131 ґ	147 ".	163 Ј	179 і	195 Г	211 У	227 г	243 у
132 ".	148 ".	164 ѿ	180 ґ	196 Д	212 Ф	228 д	244 ф
133 ...	149 .	165 Ґ	181 μ	197 Е	213 Х	229 е	245 х
134 †	150 —	166 і	182 ¶	198 Ж	214 Ц	230 ж	246 ц
135 ‡	151 —	167 §	183 .	199 З	215 Ч	231 з	247 ч
136 '.	152 '.	168 Ё	184 ё	200 И	216 Ш	232 и	248 ш
137 ‰	153 ™	169 ©	185 №	201 Й	217 Щ	233 й	249 щ
138 Љ	154 љ	170 €	186 €	202 К	218 Ъ	234 к	250 ъ
139 '.	155 '.	171 «	187 »	203 Л	219 Ы	235 л	251 ы
140 Њ	156 њ	172 ~	188 ј	204 М	220 Ь	236 м	252 ь
141 ґ	157 ґ	173 -	189 S	205 Н	221 Э	237 н	253 э
142 Ң	158 ҥ	174 ©	190 s	206 О	222 Ю	238 о	254 ю
143 Ҫ	159 ҫ	175 İ	191 і	207 П	223 Я	239 п	255 я

Передатчик должен иметь набор из восьми сигналов $s_i(t)$, где $i = 1, \dots, 8$, сопоставляемых со всеми возможными символами, причем передача каждого сигнала возможна в течение времени символа. В последней строке рис. 1.2, а указаны 16 сигналов, передаваемых восьмеричной системой модуляции для представления текстового сообщения "ДУМАЙ!".

Сообщение (текст): «ДУМАЙ!»

Знаковое кодирование

Д У М А Й !

Кодировка Windows 1251

1100010011010011111001100110000001100100110000100

8-ричные цифры

6 1 1 5 1 7 1 4 6 0 1 4 4 6 0 4

8-ричные сигналы

$s_1(t)$ $s_2(t)$ $s_3(t)$ $s_4(t)$ $s_5(t)$ $s_6(t)$ $s_7(t)$ $s_8(t)$ $s_9(t)$ $s_{10}(t)$ $s_{11}(t)$ $s_{12}(t)$ $s_{13}(t)$ $s_{14}(t)$ $s_{15}(t)$ $s_{16}(t)$

а)

Знаковое кодирование

Д У М А Й !

Кодировка Windows 1251

1100010011010011111001100110000001100100110000100

32-ричные числа

24 19 9 28 25 16 6 9 16 16

32-ричные сигналы

$s_1(t)$ $s_2(t)$ $s_3(t)$ $s_4(t)$ $s_5(t)$ $s_6(t)$ $s_7(t)$ $s_8(t)$ $s_9(t)$ $s_{10}(t)$

б)

Рис. 1.2. Сообщения, знаки и символы: а) 8-ричный пример; б) 32-ричный пример.

На рис. 1.2, б размер набора символов, M , был выбран равным 32 (каждый символ представляет 32-ричную цифру). Следовательно, биты берутся по пять, а результирующая группа из десяти чисел представляет десять готовых к передаче 32-ричных символов. Отметим, что границы символов и знаков не обязательно должны совпадать. Первый символ представляет 5/8 первого знака, "Д", второй символ — оставшуюся 3/8 знака "Д" и 2/8 следующего знака, "У", и т.д. Более эффективное разбиение знаков совсем не обязательно, поскольку система рассматривает знаки как строку символов, которую необходимо передать; только конечный пользователь (или телетайп пользователя) приписывает текстовое значение полученной последовательности битов. В 32-ричном примере передатчик должен содержать набор из 32 символов $s_i(t)$, где $i=1, \dots, 32$, сопоставляемых со всеми возможными символами. На рис. 1.2, б указаны десять сигналов,

передаваемых 32-ричной системой модуляции для представления текстового сообщения “ДУМАЙ!”.

1.1.2. Цифровые и аналоговые критерии производительности

Принципиальное отличие систем аналоговой и цифровой связи связано со способом оценки их производительности. Сигналы аналоговых систем составляют континуум, так что приемник должен работать с бесконечным числом возможных сигналов. Критерием производительности аналоговых систем связи является критерий достоверности, такой как отношение сигнал/шум, процент искажения или ожидаемая среднеквадратическая ошибка между переданным и принятым сигналами.

В отличие от аналоговых, цифровые системы связи передают сигналы, представляющие цифры. Эти цифры формируют конечный набор или алфавит, и этот набор известен приемнику априорно. Критерием качества цифровых систем связи является вероятность неверного детектирования цифры или вероятность ошибки (P_E).

1.1.3. Основные характеристики ЦСС

Любая система связи характеризуется скоростью передачи R и достоверностью; последний параметр для ЦСС определяется вероятностью ошибок P_E , т.е. вероятностью перепутывания символов. Однако этих характеристик недостаточно для оценки работы всей системы связи в целом. Всегда желательно, чтобы линия связи обеспечивала передачу информации с требуемым качеством и скоростью наиболее экономно, т.е. с наименьшими затратами энергетического и частотного ресурсов.

В теории связи наиболее широко применяются критерии качества работы ЦСС, в соответствии с которыми она оценивается величиной затрат на передачу единицы количества информации при заданном качестве ее приема. Такие критерии называются удельными, а под затратами понимают расход энергии, полосы частот, вес и габариты системы связи и ее стоимость. В дальнейшем будем учитывать только затраты энергии и полосы частот на передачу одной двоичной единицы передачи информации – бита.

Удельные затраты энергии будем характеризовать величиной

$$\beta_E = E_b / N_0, \quad (1.2)$$

где E_b – энергия входного сигнала приемника, соответствующая передаче одного бита информации с заданной достоверностью, N_0 – спектральная плотность мощности помехи на входе приемника.

Удельные затраты полосы частот будем оценивать величиной

$$\beta_W = W/R, \quad (1.3)$$

где W – эквивалентная (эффективная) полоса частот спектра сигнала, которой соответствует полоса пропускания приемного устройства, R – информационная скорость передачи (бит/с).

Т.о., качество работы системы цифровой связи можно характеризовать следующими показателями:

1. вероятностью ошибки при передаче одного элемента сообщения;
2. информационной и технической скоростью передачи (техническая скорость передачи R_t обратно пропорциональна длительности T_t одного элемента дискретного сообщения $R_t = 1/T_t$;
3. энергетическими затратами на передачу одного бита информации (удельные затраты энергии);
4. затратами полосы частот на передачу одного бита информации (удельные затраты полосы)

1.1.4. Предельные возможности ЦСС. Выбор вида сигналов

Для гауссовского канала с белым шумом и ограниченной средней мощностью сигнала К.Шеннон [4] получил широко известную формулу, определяющую максимум средней скорости передачи информации по такому каналу:

$$\max R = C = W \log_2(1 + P_c / P_{ш}) = W \log_2(1 + P_c / N_0 W), \quad (1.4)$$

где C – пропускная способность канала, P_c – средняя мощность сигнала, $P_{ш}$ – средняя мощность шума в канале, N_0 – спектральная плотность белого шума.

Эта формула характеризует предельные возможности канала, в котором реализованы оптимальные кодирование и декодирование, обеспечивающие согласование производительности источника с пропускной способностью канала. В этом случае ошибки отсутствуют. Вместе с тем заметим, что (1.4) справедлива только при физически нереализуемых условиях: бесконечном времени передачи информации и неизменных условиях работы канала¹.

Представим введенные выше затраты энергии и полосы на передачу одного бита информации в идеальном канале величинами

$$\beta_E = E_b / N_0 = P_c T_0 / N_0, \quad (1.5)$$

$$\beta_W = W / R_{\max} = W / C, \quad (1.6)$$

где T_0 – время передачи одного бита информации.

Т.к. по определению $\max RT_0 = CT_0 = 1$ бит, то (1.4) с учетом (1.5) и (1.6) преобразуем к

$$\frac{W}{C} \log_2(1 + \frac{E_b}{N_0} \frac{C}{W}) = 1 \quad (1.7)$$

или

$$\log_2(1 + \frac{\beta_E}{\beta_W})^{\beta_W} = 1 \quad (1.8)$$

Решая (1.8) относительно β_E , получим:

$$\beta_E = \beta_W (2^{1/\beta_W} - 1). \quad (1.9)$$

Выражение (1.9) определяет функциональную связь между удельными затратами энергии и полосы в идеальном гауссовском канале связи. Исследуя его на экстремум, получим, что при увеличении удельных затрат полосы удельные затраты энергии в идеальном канале уменьшаются, стремясь в пределе к величине

$$\beta_E \rightarrow \ln 2 \approx 0,7 \text{ при } \beta_W \rightarrow \infty. \quad (1.10)$$

Зависимость (1.10) изображена на рис.1.3. Эта зависимость называется границей Шеннона для идеального гауссовского канала связи.

¹ Открытые недавно турбо-коды позволяют вплотную приблизиться к пределу Шеннона.

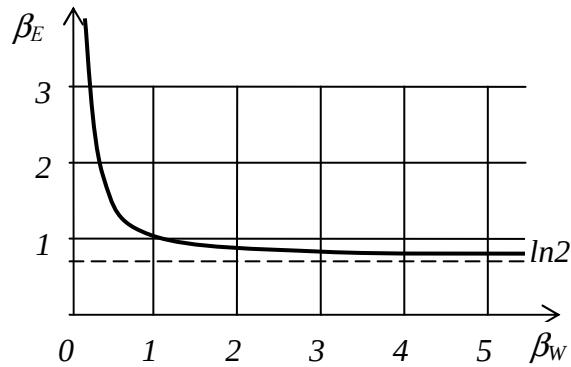


Рис. 1.3.

Она показывает, что существует бесконечное множество оптимальных систем, каждая из которых характеризуется своей парой β_E и β_W . Из графика на рис. 1.3 следует, что желание улучшить один из показателей неизбежно ухудшает другой. Оптимальные системы с малыми затратами энергии требуют значительных затрат полосы и наоборот. Понятно, что потенциально реализуемым системам соответствуют те точки на плоскости $\beta_E - \beta_W$, которые лежат выше границы Шеннона.

Проблема выбора сигнала

Из общей теории информации известно, что максимальная средняя производительность H_t источника информации определяется величиной

$$\max H_t = \lim_{T \rightarrow \infty} (\log_2 M / T), \quad (1.11)$$

где M – число различных сообщений длительностью T , вырабатываемых источником. Введем обозначения:

k – количество символов (элементов) длительностью T_0 в сообщении (слове),

n – алфавит источника сообщений,

$M = n^k$ – количество сообщений источника. Перепишем (1.11) с учетом сделанных обозначений:

$$\max R = C = 1/T_0 = \max H_t = \lim_{k \rightarrow \infty} (\log_2 n^k / kT) = \log_2 n / T, \quad (1.12)$$

которое показывает, что максимальная производительность источника в данном случае определяется только основанием алфавита n и длительностью информационного символа T .

Из (1.12) непосредственно следует

$$T_0 = T / \log_2 n. \quad (1.13)$$

Так, для кодировки ASCII, содержащей 128 символов, один информационный символ (кодированное слово) содержит 7 бит, т.е. (1.13) устанавливает связь между временем T_0 на передачу одного бита и длительностью T информационного символа. Запишем выражение для удельных затрат полосы в идеальном канале с учетом (1.13):

$$\beta_W = W / \max R = W T_0 = WT / \log_2 n = B_C / \log_2 n, \quad (1.14)$$

где $B_C = WT$ – база сигнала.

Рассмотрим два характерных случая.

1) Малые затраты полосы ($\beta_W \leq 1$). В соответствии с (1.14) используемые сигналы характеризуются соотношением $B_C \leq \log_2 n$, откуда следует

$$n \geq 2^{B_C} \quad (1.15)$$

и применение сложных сигналов с большой базой потребует огромных оснований алфавита. Использование простых сигналов с базой $B_C = 1$ приводит к следующему требованию для алфавита: $\beta_W = 1 / \log_2 n$, или

$$n = 2^{1/\beta_W} \geq 2 \quad (1.16)$$

Т.о., для получения малой полосы следует выбирать простые сигналы и использовать алфавиты с основанием больше двух. Такие M -ичные сигналы часто называют *многоосновными*.

Рис. 1.3 показывает, что малые затраты полосы требуют больших затрат энергии, что иллюстрирует Таблица 2.

Таблица 2. Затраты в идеальной системе с простыми сигналами

n	2	4	8	16	32	64
β_W	1	0,5	0,33	0,25	0,2	0,167
$1/\beta_W$	1	2	3	4	5	6
β_E	1	1,5	2,31	3,75	6,2	10,5

2) Малые затраты энергии ($\beta_E \leq 1$). Это условие соответствует $\beta_W \rightarrow \infty$. Используемые при этом сигналы характеризуются соотношением

$$B_C \gg \log_2 n. \quad (1.17)$$

Т.к. минимум n равен двум, для малых затрат энергии необходимо применять сложные сигналы с $B_C \gg 1$.

Малые затраты полосы и энергии.

Введем коэффициент обобщенных затрат $\gamma_{EW} = \beta_E \beta_W$, который в равной степени зависит как от затрат энергии, так и полосы. Подставляя в него выражение для β_E , будем иметь

$$\gamma_{EW} = \beta_W^2 (2^{1/\beta_W} - 1) \quad (1.18)$$

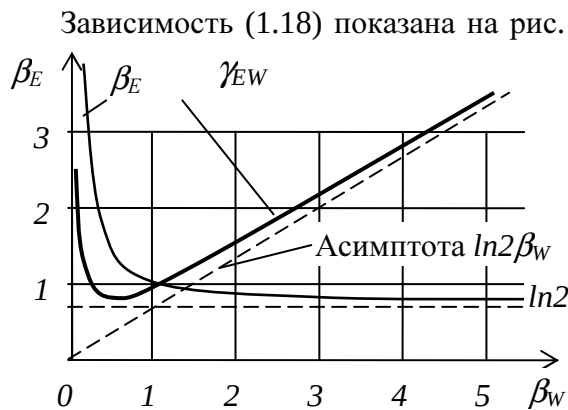


Рис. 1.4.

Зависимость (1.18) показана на рис. 1.4 и имеет минимум при $\beta_{Wopt} = 0,435$, равный $\min \gamma \approx 0,735$.

Область $\beta_W \approx 0,35..0,5$ соответствует классу простых сигналов с основанием от 4 до 8 ($n = 4..8$, см. (1.16) и величину $1/\beta_W$ в Таблице 1). В этой области, как видно из рисунка, коэффициент γ_{EW} минимален и практически не меняется. Значит, простые сигналы с основанием кода от четырех до восьми являются оптимальными с точки зрения минимизации совместных затрат полосы частот и энергии.

Если рассматривать область, в которой обобщенные затраты $\gamma_{EW} \leq 1$, то ей соответствуют простые сигналы с основанием $n = 2..16$.

1.2. Выбор основания системы счисления

В своей повседневной деятельности мы пользуемся десятичной системой. В информационных технологиях используются системы счисления, основания которых кратно двум: двоичная, восьмеричная, шестнадцатеричная, и т. д. Возникает вопрос, какая система счисления наиболее экономно (компактно) представляет любые числа?

Если n – основание системы; m – число разрядов для записи числа, то количество элементов, используемых системой для записи числа, определяется, как $v = mn$. Максимальное число, которое может быть записано в такой системе, есть $\max N(n) = n^m - 1$.

Откуда

$$n^m = \max N(n) + 1.$$

Логарифм последнего выражения:

$$m = \ln[\max N(n) + 1] / \ln(n)$$

Умножим его на n и получим:

$$v = \frac{n}{\ln(n)} \ln[\max N(n) + 1].$$

Исследуем на экстремум:

$$\frac{dv}{dn} = \ln[\max N + 1] \frac{\ln(n) - 1}{(\ln(n))^2} = 0,$$

откуда

$$n = n_{opt} = e \approx 2,72.$$

Итак, мы получили, что наиболее компактно числовые величины представляются в системе счисления с основанием $n = e$. В Таблице 3 представлены характеристики некоторых употребительных систем счисления в сравнении с оптимальной при $n = e$.

Таблица 3. Сравнение систем счисления.

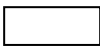
n	2	2,72	3	4	8	10	12
$n/\ln n$	2,88	2,72	2,735	2,88	3,85	4,3	4,83
v/v_{opt}	1,06	1	1,006	1,06	1,42	1,58	1,77

Видно, что системы счисления с основанием от двух до четырех практически не проигрывают оптимальной. С технологической точки зрения двоичная система несравненно удобнее остальных, что и обусловило ее широкое распространение.

1.3. Основные коды

Код – это система соответствий между элементами дискретных сообщений и кодовыми комбинациями. Будем рассматривать двоичные коды, используемые при форматировании символов источника. Встречаются различные обозначения символов двоичного кода, представленные таблицей 4.

Таблица 4. Обозначения символов двоичного кода

x_1	1	+1	+	Z	M	
x_2	0	-1	-	A	S	

Обозначения «1» и «0» как «Z» и «A» рекомендованы Международным Консультативным Комитетом по телефонии и телеграфии (МККТТ), M и S – “mark” и “space”.

Наиболее употребительные коды перечислены на рис. 1.5.

Различают коды с возвратом к нулю – ВН и без возврата к нулю – БВН. В англоязычной литературе и ряде переводных книг ВН соответствует RZ (return-to-zero), а БВН – NRZ (no return-to-zero). На диаграммах 1, 2 показаны униполярные коды ВН, на диаграммах 3-5 полярные коды. Коды ВН могут принимать три значения (диаграмма 3), коды БВН – только два. Коды с расщепленной фазой передают «1» импульсом в первой половине тактового интервала (длительности бита), а «0» - во второй.

Кодировка БВН, диаграмма 4 – биполярные импульсы с активной паузой – широко используется в цифровых логических схемах и известна также под наименованием NRZ-L (L – level – уровень). Двоичная единица при этом представляется одним уровнем напряжения, а двоичный ноль – другим.

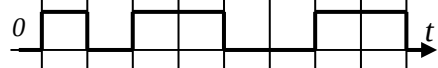
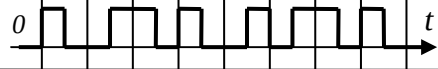
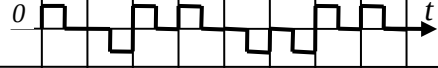
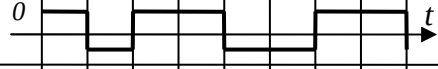
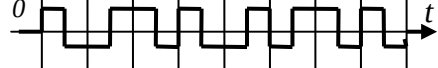
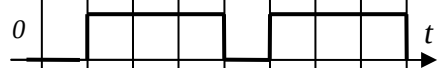
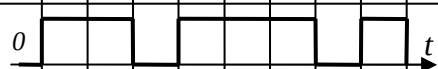
№	Вид последовательности символов								Метод формирования кода
	1	0	1	1	0	0	1	1	
1									ВН Униполярные символы с возвратом к нулю
2									ВН Униполярные символы с расщепленной фазой
3									ВН Полярные символы с расщепленной фазой
4									БВН Полярные символы с активной паузой
5									БВН Полярные символы с расщепленной фазой
6									ВН Относительный код S
7									ВН Относительный код M

Рис. 1.5. Виды наиболее употребительных кодов

Группа кодировок с расщепленной фазой применяется в системах связи, тактовая синхронизация в которых осуществляется принимаемым сигналом – в их спектре всегда присутствует тактовая частота. Кодировка 5 известна также под названием *манчестерская* (Manchester Encoding); используется в системах магнитной записи и оптической связи, а также в некоторых спутниковых каналах телеметрической передачи данных.

Относительные коды *M* и *S* применяются в телеграфии и системах связи с относительной фазовой манипуляцией (ОФМ). Код *S* изменяет свою полярность с приходом «1», а код *M* – с приходом «0».

Код называется *равномерным*, если все кодовые комбинации содержат одинаковое количество элементов, например, код Бодо, ASCII. Код Морзе – *неравномерный*.

Код называется *простым* или *первичным*, если для передачи числа *N* используются *m* разрядов, причем $n^{m-1} < N \leq n^m$, где *n* – основание кода. Например, для передачи максимального числа 130, имеющегося в сообщении, *натуральным двоичным кодом* требуется 8 разрядов, тогда как такой код может кодировать 256 чисел. Значит, 256-130 = 126 чисел не будут использованы.

Невзвешенные коды.

Натуральному двоичному коду присущ недостаток, приводящий к значительным ошибкам при действии помех. Т.к. такой код является взвешенным, т.е. разряды кодового слова имеют разный вес, ошибка в старшем разряде приводит к гораздо большему искажению сообщения, нежели ошибка в младших разрядах.

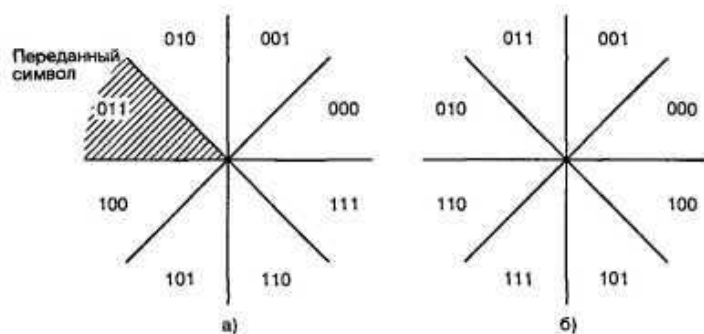


Рис.1.6. Пространства решений для: а) натурального двоичного кода; б) кода Грея.

Сказанное иллюстрируется рис. 1.6, где представлены пространства решений при использовании натурального двоичного кода (а) и невзвешенного кода Грея (б). При передаче М-ичными фазоманипулированными сигналами (МФМ) каждый сигнальный вектор не является равноудаленным от всех остальных.

На рис. 1.6, а показано восьмеричное пространство решений, где области решений обозначены 8-ричными символами в двоичной записи. При передаче символа (0 1 1) и появлении в нем ошибки наибольшую вероятность превратиться в тот же символ имеют ближайшие соседние символы, (0 1 0) и (1 0 0). Вероятность превращения символа (0 1 1) вследствие ошибки в символ (1 1 1) относительно мала. Если биты распределяются по символам согласно двоичной последовательности, показанной на рис. 1.6, а, то некоторые символьные ошибки всегда будут давать две (или более) битовые ошибки, даже при значительном отношении сигнал/шум.

Для неортогональных схем, таких как МФМ, часто используется код преобразования бинарных символов в М-арные, такой, что двоичные последовательности, соответствующие соседним символам (сдвигам фаз), отличаются только одной битовой позицией; таким образом, при появлении ошибки в М-арном символе высока вероятность того, что ошибочным является только один из k прибывших битов. Кодом, обеспечивающим подобное свойство, является код Грея; на рис. 1.6, б для восьмеричной схемы ФМ показано распределение битов по символам с использованием кода Грея. Можно видеть, что соседние символы отличаются одним двоичным разрядом. Следовательно, вероятность появления многобитовой ошибки при данной символьной ошибке значительно меньше по сравнению с не кодированным распределением битов, показанным на рис. 1.6, а.

Код Грея образуется из суммы по модулю 2 кодовой комбинации натурального кода с ней же, сдвинутой на один разряд вправо, при этом младший разряд сдвинутой комбинации отбрасывается. Пример 4-разрядных кодов – натурального и кода Грея показан в таблице 5.

Таблица 5

Десятичное число	Натуральный двоичный код	Код Грея	
0	0000	0000	
1	0001	0001	
2	0010	0011	
3	0011	0010	
4	0100	0110	Ось симметрии
5	0101	0111	
6	0110	0101	
7	0111	0100	
8	1000	1100	Главная ось симметрии
9	1001	1101	
10	1010	1111	
11	1011	1110	
12	1100	1010	Ось симметрии
13	1101	1011	
14	1110	1001	
15	1111	1000	

Свойства кода Грея:

1. Соседние комбинации отличаются в одной позиции.
2. Смена элементов в каждом разряде происходит в 2 раза реже.
3. В сумме соседних комбинаций по mod 2 число единиц = $m-3$, где m – значность кода.
4. Имеются оси симметрии (отражения) – код является рефлексным.

В заключение отметим, что натуральный двоичный код удобен для различных преобразований, ввода в компьютер, при декодировании и т. п., а код Грея – для передачи по линиям связи.

1.4. Модели каналов

В начале главы, на рис. 1.1, показана структурная схема ЦСПИ, составной частью которого является канал передачи. Чтобы установить терминологию и обозначения, используемые в дальнейшем, рассмотрим более общую модель системы цифровой связи.

Обобщенная схема ЦСПИ

Специфика радиотехнических систем передачи информации (РТСПИ) связана с особенностями распространения радиоволн, которая учитывается при выборе модели канала связи. В остальном процессы, протекающие в РТСПИ, не отличаются от процессов в других системах передачи – системах проводной связи, гидроакустических и оптических. Обратимся к схеме РТСПИ с одним источником и одним получателем (рис. 1.7).

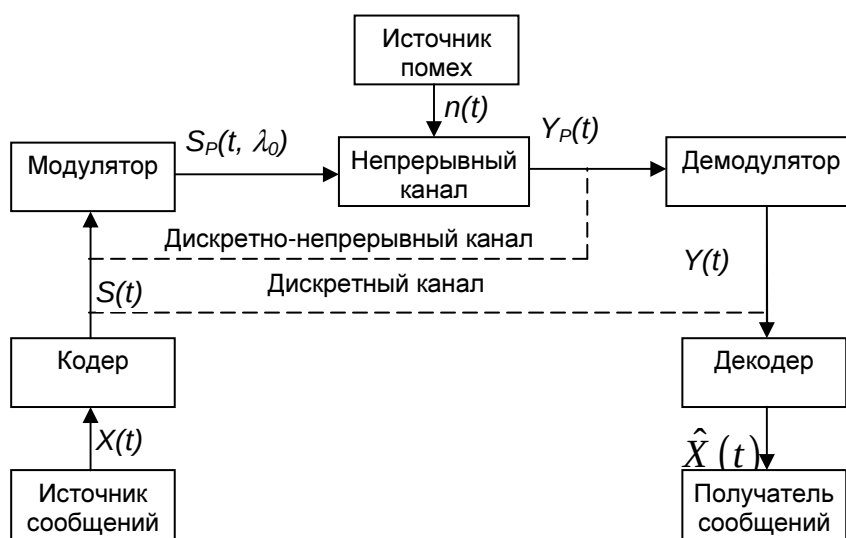


Рис. 1.7. Обобщенная схема цифровой системы передачи.

На этом рисунке использованы обозначения:

$X(t)$ – выход источника сообщений, вход кодера;

$S(t)$ – выход кодера, вход модулятора;

$S_p(t, \lambda_0)$ – выход модулятора (радиосигнал). λ_0 – параметры сигнала;

$Y_p(t)$ – выход канала (радиосигнал), вход демодулятора;

$Y(t)$ – выход демодулятора, вход декодера;

$\hat{X}(t)$ – выход декодера, к получателю сообщений.

Назначение и свойства отдельных блоков схемы были рассмотрены ранее (см. раздел 1.1).

Сигнал $S_p(t, \lambda_0)$ на входе канала имеет параметры $\lambda_0 = \{\omega_0, \varphi, A\}$, которые, в общем случае, зависят от времени. Сигнал $Y_p(t)$ на выходе канала содержит искаженный входной сигнал $S_p(t, \lambda)$ и помеху $n(t)$:

$$Y_P(t) = h_c * S_P(t, \lambda_0) + n(t) = S_P(t, \lambda) + n(t) \quad (1.19)$$

Вектор параметров λ , кроме λ_0 , содержит дополнительные составляющие: время запаздывания, доплеровский сдвиг частоты, уменьшение амплитуды и т.п.

Канал будем называть *непрерывным*, если на его входе и выходе действуют непрерывные сигналы; в *дискретном* канале, соответственно, дискретные сигналы. Можно также обозначить *дискретно-непрерывные* и *непрерывно-дискретные* каналы.

Возможна также классификация каналов по назначению радиотехнических систем передачи информации (РСПИ) (телеграфные, телефонные, телевизионные, телеметрические и др.), по виду физической среды распространения (проводные, кабельные, волноводные и др.) и по диапазону используемых ими частот. К радиодиапазону относят частоты в пределах $30...3 \cdot 10^{12}$ Гц, что соответствует длинам волн от 10^8 м до 0,1 мм. Кроме радиодиапазона, в настоящее время широко используется оптический диапазон волн. В силу дискретного характера электромагнитного излучения в оптическом диапазоне волн такие каналы принято называть квантовыми. Данные о радиодиапазонах приведены в табл. 1.1. Далее рассматриваются особенности распространения и использования радиоволн различных диапазонов.

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

В диапазонах ОНЧ и НЧ на небольших расстояниях поле в месте приема создается в результате дифракционного огибания волнами выпуклой поверхности Земли. На больших расстояниях радиоволны распространяются в своеобразном сферическом волноводе, внутренняя стенка которого образуется поверхностью Земли, а внешняя — ионосферой. Такой механизм распространения радиоволн позволяет принимать сигналы в любой точке Земли, причем параметры принятых сигналов отличаются достаточно высокой стабильностью. Особенностью этих диапазонов является также способность волн проникать в толщу Земли и воды на глубину в десятки метров. Принципиальным недостатком таких каналов являются ограниченная полоса частот (единицы герц) и очень большие линейные размеры антенных устройств, соизмеримых с длиной волны, составляющей километры. Сверхдлинные волны находят применение в навигации и передаче ограниченного объема информации на подводные объекты.

В распространении волн диапазона ВЧ участвует ионосфера. Однако если волны, длиннее 1 км, отражаются от нижнего ее слоя практически зеркально, то декаметровые волны достаточно глубоко проникают в ионосферу, что приводит к эффекту многолучевости, когда в точку приема приходит одновременно несколько сигналов с разными временами запаздывания. Многолучевость может носить дисперсионный или дискретный характер. Дисперсия (рассеяние) сигнала определяется отражением радиоволн от некоторого объема ионосферы, а дискретная многолучевость — отражением от разных слоев ионосферы. Поскольку глубина проникновения в ионосферу зависит от длины волны, то для передачи информации между двумя пунктами можно указать оптимальную рабочую частоту, на которой связь будет наиболее надежной (максимум мощности принимаемого сигнала, минимум эффекта многолучевости). Значения оптимальной рабочей частоты рассчитывают для определенных трасс и времени связи. Для этого составляют долговременные и кратковременные прогнозы по данным мировой сети станций ионосферного зондирования. Декаметровые волны широко применяются для глобальной связи и радиовещания. С их помощью можно передавать информацию сравнительно большого объема в пределах всего земного шара при ограниченной

мощности передатчика и небольших по размеру антеннах. Полоса частот передаваемых сигналов в декаметровом канале не превышает 10 кГц.

До появления спутниковых систем связи этот диапазон был единственно пригодным для организации связи между двумя любыми пунктами на Земле без промежуточной ретрансляции. Однако эффект глобального распространения коротких волн имеет и свою отрицательную сторону: в точке приема могут появиться сильные помехи от дальних радиостанций.

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

Волны диапазона частот выше 30 МГц слабо дифрагируют и поэтому распространяются в пределах прямой видимости. Расстояние прямой видимости (радиогоризонт) в километрах по поверхности Земли примерно

равно $5\sqrt{A+B}$, где A и B — высоты передающей и приемной антенн в метрах. Если предположить, что $A = 30$ м, а $B = 3$ м, то радиогоризонт равен 30 км и для всех радиоволн, имеющих путь распространения менее 30 км, потери сигнала не будут связаны с кривизной поверхности Земли. Для всех остальных лучей возникает дополнительное затухание, обусловленное экранированием сигнала земной поверхностью. Однако резкое уменьшение уровня сигнала на расстояниях, превышающих прямую видимость, обуславливает снижение уровней помеховых сигналов и возможность использования одинаковых частот для организации связи в зонах, удаленных на расстояния, превышающие радиогоризонт. Некоторого увеличения дальности можно достичь, применив поднятые антенны, а для организации связи на расстояния, превышающие прямую видимость, — ретрансляцию сигналов. Системы с ретрансляцией сигналов называются радиорелейными линиями. Одним из основных достоинств высокочастотных диапазонов является большой частотный ресурс, что позволяет создавать радиосистемы передачи информации с высокой скоростью передачи и радиосети с большим числом одновременно работающих радиостанций. В последние 30 лет диапазон ОВЧ и СВЧ нашел широкое применение для создания сетей мобильной связи.

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

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

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

Из всего многообразия орбит ИСЗ особый интерес представляет экваториальная круговая орбита, удаленная от поверхности Земли на расстояние около 36 тыс. км (стационарная экваториальная орбита). Когда направление движения ИСЗ по такой орбите совпадает с направлением вращения Земли, спутник будет казаться наземному

наблюдателю неподвижным (стационарный спутник). При использовании трех стационарных спутников, расположенных в экваториальной плоскости через 120° по дуге, оказывается принципиально возможным организовать глобальную систему связи. Максимальный, от горизонта до горизонта, обзор земной поверхности одним ИСЗ, или, иначе говоря, максимальное расстояние вдоль поверхности Земли между двумя станциями, будет практически составлять 15... 17 тыс. км. Существенные преимущества стационарной орбиты заключаются в снижении требований к системам слежения за спутником, сведении к минимуму доплеровских сдвигов частоты сигналов, что упрощает приемное устройство при большом обзоре поверхности Земли. Недостатком стационарной орбиты является плохой охват приполярных зон. Поэтому в России для систем связи широко применяются сильно вытянутые эллиптические орбиты с большой полуосью до пяти земных радиусов с эксцентриситетом 0,8—0,9 и углом наклона примерно 65° . Три спутника, выведенные через равномерные интервалы времени на аналогичные эллиптические орбиты, восходящие узлы которых смещены относительно друг друга на 120° , могут обеспечить круглосуточную непрерывную связь между земными станциями, расположенными в Северном полушарии Земли, на глобальные расстояния. Другая возможность в создании глобальных РСПИ заключается в использовании «созвездий» спутников, орбиты которых выбраны так, что над любой точкой на поверхности Земли находится по крайней мере один спутник.

Выбор рабочих частот для линии радиосвязи через ИСЗ определяется следующими факторами: условиями распространения и поглощения радиоволн, уровнем внешних помех, принимаемых антенной, техническими средствами (коэффициент шума приемного устройства, ширина лепестка диаграммы направленности антенны, точность ориентации и т. п.), взаимными помехами между системами связи через ИСЗ и другими службами, работающими в смежных или совмещенных диапазонах частот. Ограничение диапазона частот снизу определяется экранирующим действием ионосферы, а сверху — поглощением в тропосфере. Эти два фактора предопределили диапазон рабочих частот 40 МГц.. 40 ГГц. В настоящее время наибольшее использование находит диапазон частот 1... 12 ГГц.

1.4.1. Модели непрерывных каналов

Непрерывные каналы можно классифицировать по виду помех и характеру преобразования $S_P(t, \lambda_0)$ в полезный принятый сигнал $S_P(t, \lambda)$. Если ограничиться предположением, что в канале действует аддитивный нормальный белый шум $n(t)$, то непрерывные каналы подразделяются по виду преобразования $S_P(t, \lambda_0)$ в $S_P(t, \lambda)$, т.е. по виду искажений сигнала.

В большинстве радиотехнических систем излучаемые сигналы являются узкополосными:

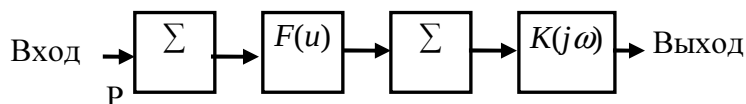
$$S_P(t, \lambda_0) = A(t)\cos[\omega_0 t + \varphi(t)], \quad (1.20)$$

где $A(t)$ и $\varphi(t)$ – функции, отражающие законы амплитудной и угловой модуляции; ω_0 – несущая частота сигнала, значительно превышающая ширину его спектра. Совокупностью параметров λ_0 является множество $\lambda_0 = \{A(t), \varphi(t), \omega_0(t)\}$.

Искажения излученного сигнала принято рассматривать отдельно для однолучевых и многолучевых каналов. В однолучевых каналах отсутствуют замирания, вызванные интерференцией нескольких сигналов, распространяющихся по различным путям. Многолучевые каналы будут рассмотрены отдельно.

Физическая модель непрерывного канала связи включает в свой состав технические средства, расположенные между выходом модулятора и входом демодулятора (рис. 1.8).

Проходя по непрерывному каналу связи, сигнал претерпевает ряд изменений. Эти изменения сводятся к ослаблению, искажению сигнала и наложению на него помех. В отдельных случаях искажению подвергается смесь сигнала и помех, например во входных цепях приемника или при ретрансляции в радиорелейных линиях. Для анализа системы важно знать характер искажений и уметь их моделировать. Реальные искажения имеют достаточно сложный характер. Однако для решения большинства задач непрерывный канал можно смоделировать в виде последовательно включенных линейных инерционных и нелинейных безынерционных четырехполюсников, обуславливающих соответственно линейные и нелинейные искажения сигналов (см. рис. 1.8). Помехи могут накладываться на сигнал в любой точке цепи. Несмотря на кажущуюся простоту такой модели канала, нахождение отклика на ее выходе в тех



ис. 1.8. Модель непрерывного канала связи

случаях, когда помеха действует на входе нелинейного звена, является сложной математической задачей. Поэтому часто при решении подобных задач обращаются к различным методам моделирования.

Линейные искажения проявляются в изменении спектра (корреляционной функции) сигналов и помех. В зависимости от того, каковы эти искажения: регулярны или случайны, различают соответственно каналы с детерминированными или случайными линейными искажениями. Детерминированные линейные искажения в реальных каналах связаны с наличием частотно-избирательных цепей (фильтров во входных каскадах приемника и в выходных каскадах передатчика, коаксиальных и волноводных трактов, антенн и т. д.).

Случайные линейные искажения определяются средой распространения и связаны с прохождением сигнала от передающей антенны к приемной антенне разными путями (лучами). Этот эффект называется рассеянием сигнала. Различают два вида рассеяния сигнала: дискретное, когда запаздывание между сигналами в соседних лучах принимает конкретное значение (многолучевый канал), и дисперсное, когда запаздывание между соседними лучами бесконечно мало, а число лучей бесконечно велико.

Характер рассеяния сигнала определяется диапазоном используемых частот и типом системы. Если раньше типичными каналами с рассеянием сигналов являлись тропосферный и ионосферный, в которых связь за пределами прямой видимости достигалась рассеянием радиоволн на неоднородностях тропосферы и ионосферы и для которых были найдены способы борьбы с многолучевостью, то в последние десятилетия с развитием мобильных систем связи, действующих в условиях городской застройки, борьба с многолучевостью приобрела еще большую актуальность. Искажения сигналов, особенно в мобильных системах, имеют достаточно сложный характер. Однако даже упрощенные модели каналов позволяют разобраться в характере искажений сигналов и находить способы повышения качества передачи информации по каналам с рассеянием.

Пусть в точку приема приходят сигналы, переотраженные от совокупности бесконечно малых по размеру отражателей, размещенных в некотором пространстве, которые случайным образом перемещаются, сохраняя постоянным, в среднем, объем занимаемого пространства. Тогда на входе приемника будем иметь сумму сигналов с разной амплитудой и временем прихода, которые, в свою очередь, случайно изменяются с некоторой скоростью. Максимальную разность времен прихода сигналов называют **временем рассеяния сигнала, расширением задержки** или **памятью канала**. Употребляется также термин **длина профиля многолучевости**. Естественно, что эта величина также носит случайный характер, но можно указать ее среднее значение. Рассмотрим характер искажения гармонического сигнала при прохождении по такому

каналу. Сигнал на входе приемника является суммой синусоид со случайными амплитудами и фазами.

Если число переотраженных сигналов велико, то по центральной предельной теореме теории вероятностей суммарные ортогональные сигналы будут иметь нормальные законы распределения амплитуд, а результирующий сигнал будет иметь случайную огибающую и фазу, изменяющиеся соответственно по рэлеевскому и равномерному законам.

Скорость изменения огибающей (замираний сигнала) определяется доплеровским сдвигом по частоте при движении отражателей. Обычно в мобильных системах связи переотраженные сигналы действуют на фоне достаточно мощного прямого сигнала. Тогда результирующий сигнал будет иметь райсовский закон распределения огибающей.

Если число лучей ограничено, например только два, то результирующий сигнал будет представлять биения колебаний последних. Вероятностные характеристики огибающей и фазы результирующего сигнала будут определяться соответствующими характеристиками амплитуд и фаз суммируемых сигналов.

Для сигналов с фиксированной шириной спектра F_c канал с рассеянием можно представить как фильтр со случайно изменяющимися во времени параметрами..

В общем случае сигнал на выходе линейного канала с изменяющимися параметрами можно найти, используя интеграл Дюамеля

$$u(t) = \int_0^t h(t, \tau) s(t - \tau) d\tau,$$

где $h(t, \tau)$ — импульсная характеристика канала. Таким образом, для оценки линейных искажений необходимо знать функцию $h(t, \tau)$ или связанную с ней преобразованием Фурье комплексную частотную характеристику $K(j\omega, t)$.

Решение задач анализа и синтеза устройств обработки сигналов существенно упрощается, если перейти к дискретной модели канала и сигналов. Дискретное представление математических моделей каналов основывается на конечном времени τ_n рассеяния сигнала (памяти канала), определяемом протяженностью импульсной характеристики $h(t, \tau)$, и конечной ширине спектра передаваемого сигнала F_c . Формальным способом введения дискретной модели может быть разложение функции в ряды Котельникова, Фурье и т. п. Если полоса частот сигнала, передаваемого по каналу, ограничена интервалом $F_c = F_B - F_H$, где F_B и F_H — верхняя и нижняя частоты спектра сигнала, то достаточно рассматривать функцию $K(j2\pi f, t)$ переменной f только в интервале F_c . При этом импульсную характеристику $h(t, \tau)$ можно представить в виде ряда Котельникова для сигнала с полосовым спектром:

$$\begin{aligned} h(t, \tau) &= \sum_{i=-\infty}^{\infty} h_i(t, \tau) = \\ &= \sum_{i=-\infty}^{\infty} H(i, t) \frac{\sin[\pi F_c (t - i / F_c)]}{\pi F_c (t - i / F_c)} \cos \left[\pi (F_B + F_H) \left(\tau - \frac{i}{F_c} \right) + \varphi(i, t) \right], \end{aligned}$$

(9.2-1.20)

где $H(i, t)$ — значения огибающей импульсной характеристики при $\tau = i/F_c$; $\varphi(i, t)$ — значения фазы.

Физическая модель канала (рис. 1.9), построенная в соответствии с формулой (9.2), содержит линию задержки с L отводами, усилители, комплексный коэффициент которых $h(i, f) = H(i, t) \exp(j\varphi(i, t))$ может изменяться, и сумматор. Временная задержка между соседними отводами равна $1/F_c$.

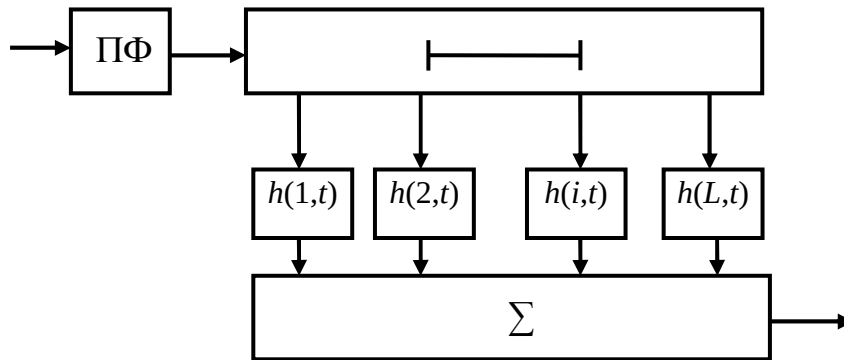


Рис. 1.9. Модель канала с рассеянием для сигналов с ограниченной шириной спектра

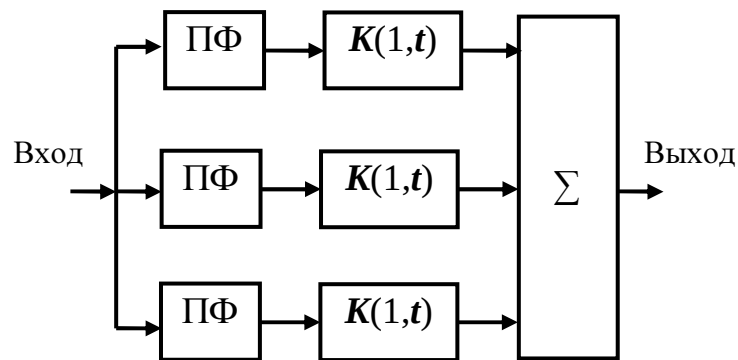


Рис. 1.10. Модель канала с ограниченным временем рассеяния

В частотной области модель канала можно построить, предположив конечность времени τ_n рассеяния сигнала. Тогда функция $K(j\omega, t)$ переменной $f = \omega/(2\pi)$ может быть задана комплексными значениями: $K(i, t) = K(j2\pi\Delta f)$, где $\Delta f = 1/\tau_n$. Дискретная модель канала содержит набор полосовых фильтров с примыкающими частотными характеристиками, полоса пропускания каждого из которых равна $1/\tau_n$, и усилителей с управляемыми комплексными коэффициентами передачи $K(i, t)$ (рис. 1.10). Величину $1/\tau_n = F_K$ иногда называют **полосой когерентности**. Гармонические сигналы с разносом по частоте, превышающим $1/\tau_n$, будут иметь некоррелированные случайные огибающую и фазу. Этот параметр определяет и характер замираний сигнала. Если ширина спектра передаваемого по каналу сигнала меньше F_K , то все спектральные составляющие сигнала изменяются одновременно и такие замирания называются **общими**. В том случае, когда $F \gg F_K$, отдельные участки спектра сигнала изменяются независимо и замирания называются **селективными**.

Информацию о динамике изменения коэффициента передачи дает корреляционная функция или спектральная плотность мощности этого коэффициента. Время корреляции τ_k или ширина спектра флуктуации $F_{фл}$ характеризуют скорость изменения параметров канала. Так, для КВ канала ширина $F_{фл}$ составляет 0,1...1 Гц. В мобильных системах с частотами в несколько сотен МГц спектр флуктуаций оказывается значительно шире.

Принятый полезный сигнал по отношению к излученному характеризуется дополнительными параметрами: случайным ослаблением $\alpha(t)$, средним временем

запаздывания τ_3 , доплеровским смещением частоты Ω , случайной начальной фазой θ и имеет вид

$$S_p(t, \lambda) = \alpha(t)A(t - \tau_3)\cos[(\omega_0 + \Omega)t + \varphi(t - \tau_3) - \theta]. \quad (1.21)$$

Т.о., совокупность параметров принятого полезного сигнала $\lambda = \{A(t), \varphi(t), \omega_0(t), \alpha(t), \tau_3, \Omega, \theta\}$.

В зависимости от того, какие параметры принятого сигнала на приемной стороне известны, можно выделить несколько моделей непрерывных каналов:

1. **Гауссовский канал** – канал, в котором помеха имеет вид аддитивного нормального белого шума, а искажения полезного сигнала несут существенны, так как могут быть скомпенсированы. Компенсация искажений возможна, если на приемной стороне дополнительные параметры полностью известны или могут быть измерены достаточно точно (т.е. известна или доступна для измерения импульсная характеристика канала h_c). Поэтому можно считать, что $S_p(t, \lambda_0) = S_p(t, \lambda)$, а выходной сигнал гауссовского канала есть

$$Y_p(t) = S_p(t, \lambda_0) + n(t) \quad (1.22)$$

2. **Гауссовский канал с неизвестной фазой** сигнала определяется параметрами $\Omega, \tau_3, \alpha(t) = \alpha$, которые постоянны и известны. Начальная фаза неизвестна и обычно считается равномерно распределенной в интервале $(0, 2\pi)$. Такой сигнал хорошо описывает процессы в линии связи на расстоянии прямой видимости.

3. **Канал с амплитудными замираниями** является дальнейшим усложнением канала с неизвестной фазой в предположении, что $\alpha(t)$ – случайная функция времени. Его выходной полезный сигнал имеет вид

$$S_p(t, \lambda) = \alpha(t)A(t - \tau_3)\cos[(\omega_0 + \Omega)t + \varphi(t) - \theta] \quad (1.22)$$

Случайная функция $\alpha(t)$ перемножается с сигналом и поэтому называется *мультипликативной помехой*. Мультипликативную помеху можно рассматривать как функцию, модулирующую амплитуду полезного сигнала. Модуляция приводит к расширению спектра принятого сигнала относительно спектра излученного сигнала. Вследствие этого модель канала (1.22) называют также *каналом с рассеянием энергии по частоте*.

Для задания канала с замираниями требуется описать мультипликативную помеху $\alpha(t)$. Считается достаточным, если указывается одномерная плотность вероятности $W(\alpha)$ и время корреляции τ_k . Если $W(\alpha)$ описывается релеевским законом распределения, то канал называется релеевским или каналом с релеевскими замираниями.

По времени корреляции мультипликативные помехи разделяются на *медленные* и *быстрые*. О медленных замираниях говорят в случае, если время корреляции процесса $\alpha(t)$ значительно превышает интервал наблюдения сигнала. Соответственно, быстрые замирания характеризуются временем корреляции меньшим, чем интервал наблюдения сигнала.

Причинами медленных замираний являются изменения свойств среды распространения радиоволн в зависимости от метеоусловий, времени суток, года, от солнечной активности и т.п. Причиной быстрых замираний может быть, например, наличие в канале нескольких путей распространения радиоволн.

При описании непрерывных каналов важны также ограничения, накладываемые на среднюю или пиковую мощность излучаемых сигналов, на полосу используемых частот.

Т. о., непрерывный канал считается заданным, если указаны мощность сигналов, полоса частот, дано описание моделей помех и искажений сигналов.

1.4.2. Модели дискретных каналов

Дискретными называются каналы, входные и выходные сигналы которых принимают конечное число мгновенных значений. На рис. 1.7 дискретный канал (ДК) включен между

кодером и декодером. Очевидно, что дискретный канал всегда содержит непрерывный, так что искажения сигналов и помехи, действующие в дискретном канале, определяются непрерывным каналом.

Переход от дискретных сигналов к непрерывным на передающей стороне осуществляется модулятором; на приемной стороне дискретные сигналы появляются на выходе демодулятора. Поэтому дискретный канал всегда содержит модем.

Таким образом, свойства дискретного канала определяются непрерывным каналом и структурой модема.

ДК задается множеством входных $\{s_i\}, i = [1, L_S]$ и выходных $\{y_j\}, j = [1, L_y]$ символов длительностью T и условными вероятностями $P(y_j/s_i)$ преобразования входных символов в выходные. Здесь L_S и L_y – объемы алфавитов входных и выходных символов. В общем случае $L_y \geq L_S$, однако, обычно $L_y = L_S = m$.

Для дискретных каналов широко используется представление принятой последовательности символов $Y = \{y_1, y_2, y_3, \dots, y_m\}$ в виде суммы переданной последовательности $S = \{s_1, s_2, s_3, \dots, s_m\}$ и комбинации помехи (вектора ошибки) $E = \{e_1, e_2, e_3, \dots, e_m\}$

$$Y = S \oplus E, \quad (1.23)$$

где знак $\oplus$ означает поразрядное сложение по модулю L_S .

В бинарном случае ($L_S = 2$) нулевой символ вектора ошибки $e_i = 0$ означает, что i -й символ принят правильно ($y_i = s_i$), а $e_i = 1$ – ошибку в приеме ($y_i \neq s_i$).

Классификацию дискретных каналов удобно вести по вектору ошибки E . Наиболее распространены следующие модели.

Канал без памяти – канал, в котором символы e_i являются независимыми случайными величинами. Прием каждого символа в таком канале не зависит от предыстории.

Канал с памятью – канал, в котором прием символа зависит от результата приема предыдущего символа.

Стационарных канал – канал, в котором вероятность ошибочного приема символов не изменяется с течением времени.

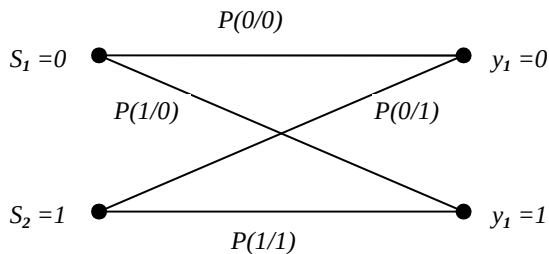


Рис. 1.8.

Двоичные (бинарные) каналы – каналы, сигналы в которых представляются двоичным кодом с символами 0 и 1. Они задаются с помощью графа, представленного на рис. 1.8.

Вероятности $P(0/0)$ и $P(1/1)$ определяют правильный прием символов 0 и 1 соответственно, а $P(1/0)$ и $P(0/1)$ – вероятности ошибок при приеме символов 0 и 1.

Симметричным двоичным называется канал, в котором вероятности ошибок при приеме 0 и 1 одинаковы, $P(1/0) = P(0/1) = p$. Следовательно, $P(0/0) = P(1/1) = 1 - p = q$, где q – вероятность правильного приема.

Например, вероятность ошибочного приема символа в трехбитовом сообщении $P(000/001) = qq(1-q) = q^2p$.

Учитывая (1.23), можно записать $P(\{y_i\}/\{s_i\}) = P(\{e_i\}/\{s_i\}) = P(\{e_i\})$, $e_i \in \{0,1\}$, т.е. вероятность трансформации символов определяется вектором ошибок.

Найдем $P_n(t)$ – вероятность t ошибок при приеме n символов. Обозначим через $P_n^*(t)$ вероятность одного сочетания t ошибок в n символах: $P_n^*(t) = p^t(1-p)^{n-t}$. Тогда искомая вероятность определится с учетом всех возможных сочетаний ошибок:

$$P_n(t) = C_n^t P_n^*(t) = n! p^t (1-p)^{n-t} / t!(n-t)! \quad (1.24)$$

Большинство реальных каналов имеют «память», которая проявляется в том, что вероятность ошибки в символе зависит от того, какие символы передавались до него и как они были приняты. Первый факт обусловлен межсимвольными искажениями –

результатом рассеяния сигнала в канале, второй – изменением отношения сигнал/шум в канале или характера помех.

Рассмотрим подробнее канал с памятью. Условная вероятность приема сигнала y_i в канале зависит от передачи предыдущих символов: $P(y_i | s_i, s_{i-1}, s_{i-2}, \dots)$. Если обозначить через $(s_{i-1}, s_{i-2}, s_{i-3}, \dots, s_{i-N}) = C_{i-1}$ – состояние канала в $(i-1)$ -й момент времени, то памятью канала N будем называть максимальное количество предыдущих символов, влияющих на прием текущего символа. Увеличение N не приводит к изменению условной вероятности:

$$P(y_i / s_i, s_{i-1}, \dots, s_{i-N}) = P(y_i / s_i, s_{i-1}, \dots, s_{i-N-j}), \quad j \geq 1 \quad (1.25)$$

Канал задается совокупностью переходных вероятностей $P(C_i / C_{i-1})$ и вероятностью ошибок $P_e^{(i)}$ в i -ом состоянии: $P(y_i, C_i / s_i, C_{i-1})$, $C_i \in L$ – множество состояний канала. Разумно предположить, что передаваемые сигналы s_i и состояние канала C_i – независимы. Тогда условную вероятность приема текущего символа можно записать, как

$$P(y_i, C_i / s_i, C_{i-1}) = P(y_i / s_i, C_{i-1}) P(C_i / C_{i-1}) \quad (1.26)$$

Если текущее состояние канала зависит только от его предыдущего состояния ($N=1$), то канал называют *марковским*, а последовательность состояний описывается простой цепью Маркова.

Рассмотрим *модель Гильберта* – простейшую модель канала с памятью.

1.4.2.1. Модель Гильберта

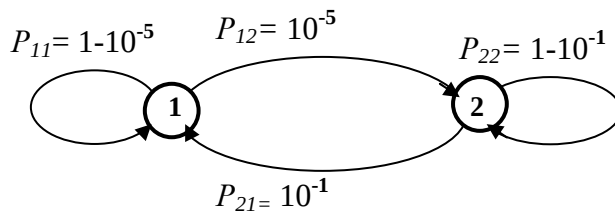
В этой модели канал характеризуется двумя состояниями.

Состояние 1 – хорошее, ошибки отсутствуют.

Состояние 2 – плохое, (например, вероятность ошибок $P_e^{(2)}=0,4$).

Канал можно описать матрицей переходных состояний

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}, \text{ или графом (числовые значения приведены для примера):}$$



Когда $P_{11} \neq P_{22}$, возникают пакеты ошибок.

Вероятности, составляющие полную группу событий:

$$\begin{aligned} P_1 + P_2 &= 1 \\ P_{11} + P_{12} &= 1 \\ P_{22} + P_{21} &= 1 \end{aligned} \quad (1.27)$$

Вероятности пребывания канала в состояниях 1 и 2:

$$\begin{aligned} P_1 &= P_1 P_{11} + P_2 P_{21}; \\ P_2 &= P_2 P_{22} + P_1 P_{12}. \end{aligned} \quad (1.28)$$

Выразим в (1.28) P_2 через P_1 :

$$\begin{aligned} (1-P_1)P_{21} + (1-P_{12})P_1 &= P_1, \text{ откуда} \\ P_1 &= P_{21} / (P_{12} + P_{21}). \end{aligned} \quad (1.29)$$

Аналогично, для состояния 2 получим

$$P_2 = P_{12} / (P_{12} + P_{21}). \quad (1.30)$$

Так как ошибки возникают только в плохом состоянии канала, то среднюю вероятность ошибок найдем, как

$$P_e = P_2 P_e^{(2)} = P_e^{(2)} P_{12} / (P_{12} + P_{21}) \quad (1.31)$$

Найдем длину пакета ошибок как среднее число символов, переданных по каналу в его плохом состоянии:

$$\bar{l}_2 = \sum_{i=1}^{\infty} i P_2(i), \quad (1.32)$$

где $P_2(i) = P_{22}^{i-1} P_{21}$ - вероятность того, что возникшее плохое состояние канала будет длиться в течение i символов. Тогда (1.32) переписывается в виде:

$$\bar{l}_2 = \sum_{i=1}^{\infty} i P_{22}^{i-1} P_{21} = 1 \cdot P_{21} + 2 P_{22} P_{21} + 3 P_{22}^2 P_{21} + \dots, \quad (1.33)$$

который представляет собой арифметико-геометрическую прогрессию.

В справочнике [12] приведена сумма такой прогрессии:

$$\sum_{k=0}^{\infty} (a + kr) q^k = \frac{a}{1-q} + \frac{rq}{(1-q)^2}. \text{ Тогда искомое выражение в окончательном виде:}$$

$$\bar{l}_2 = \frac{1}{P_{22}} \frac{P_{21} P_{22}}{(1 - P_{22})^2} = \frac{1}{P_{21}} \quad (1.34)$$

Аналогично найдем среднюю длину интервала между ошибками:

$$\bar{l}_1 = \sum_{i=1}^{\infty} i P_{11}^{(i-1)} P_{12} = \frac{1}{P_{12}} \quad (1.35)$$

Для нашего примера $\bar{l}_{21} = 10$, $\bar{l}_{12} = 10^5$.

1.4.2.2. Характерные искажения в канале

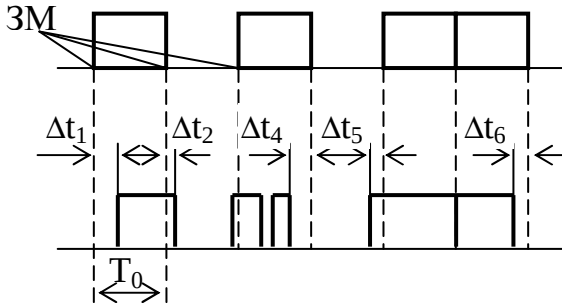


Рис. 1.9. Искажения сигналов

Отметим два характерных вида искажений: *краевые искажения* и *дробления*. На верхней диаграмме рисунка показаны излученные импульсы, а на нижней – сигнал, восстановленный демодулятором. Начало и конец тактового интервала называются значащими моментами (ЗМ). Если фронты восстановленного сигнала не совпадают с ЗМ, то имеют место краевые искажения, которые характеризуются смещениями Δt_i . Если в течение тактового интервала наблюдаются переходы от 0 к 1 и обратно, имеют место дробления сигнала.

1.4.2.3. Регистрация сигналов

Регистрация принятых сигналов осуществляется демодулятором, выход которого является цифровым. В простейшем случае (например, в проводной связи) решение о принятом символе – ноль или единица – производится демодулятором на основе анализа восстановленного сигнала, подверженного краевым искажениям и дроблениям.

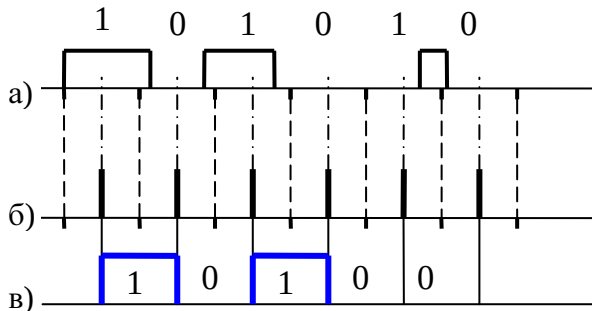


Рис. 1.10. Метод стробирования

Существует два основных вида регистрации сигналов – метод стробирования и метод интегрирования.

1) Метод стробирования

Верхняя диаграмма рис.1.10 *a* – сигналы на выходе демодулятора. Присутствуют краевые искажения. Для регистрации символов производится взятие отсчетов в середине тактовых интервалов, как показано на второй диаграмме (б). Диаграмма (в) показывает результат регистрации. Очевидно, что метод стробирования будет давать ошибки при наличии дроблений, т.е. всякий раз, когда в центре тактового интервала сигнал

искажен помехой. Такая ситуация, впрочем, возможна как в результате дроблений, так и в результате краевых искажений.

2) Метод интегрирования

В этом методе используется накопление (интегрирование) сигналов в течение тактового интервала, т.е. предполагается использование интегратора со сбросом. Сброс интегратора производится в точках ЗМ (рис. 1.11).

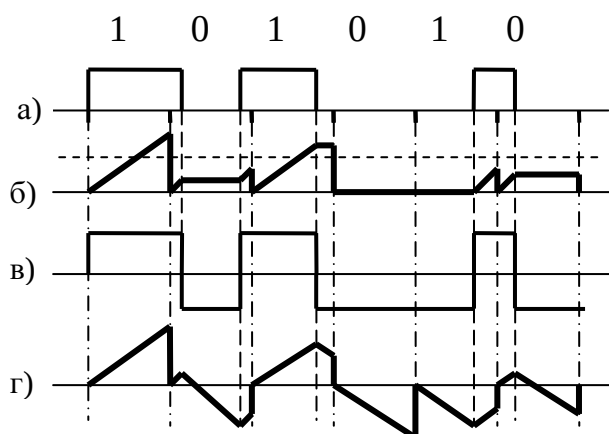


Рис. 1.11 Метод интегрирования

При униполярном сигнале (диаграмма *a*) выход интегратора показан на диаграмме *б*). Моменты взятия отсчетов выходного напряжения интегратора совпадают с ЗМ. После взятия отсчета интегратор сбрасывается в нуль и процесс накопления возобновляется. Для решения о приеме символа требуется пороговая схема; величина порогового уровня составляет половину максимально возможного напряжения интегратора. Это максимальное напряжение получается при входном сигнале без искажений. Чтобы избавиться от необходимости выставления порога,

обычно используют полярные сигналы (диаграммы *в*), *г*)). При этом пороговый уровень равен нулю.

2. Модуляция

2.1. Общие положения

На рис. 2.1 показаны примеры модуляции несущего колебания прямоугольными импульсами. Предполагается, что данные представляют собой последовательность "нулей" и "единиц", передаваемых со скоростью $1/T$. Временные диаграммы левого столбца рисунка демонстрируют сам прямоугольный модулирующий сигнал и результаты амплитудной (АМ), частотной (ЧМ) и фазовой (ФМ) модуляций. В моменты времени $(T/2, 3T/2, 5T/2...)$ происходят скачкообразные изменения амплитуды, частоты или фазы несущего колебания соответственно. Правый столбец показывает результаты расчета спектра модулирующего сигнала и модулированных колебаний.

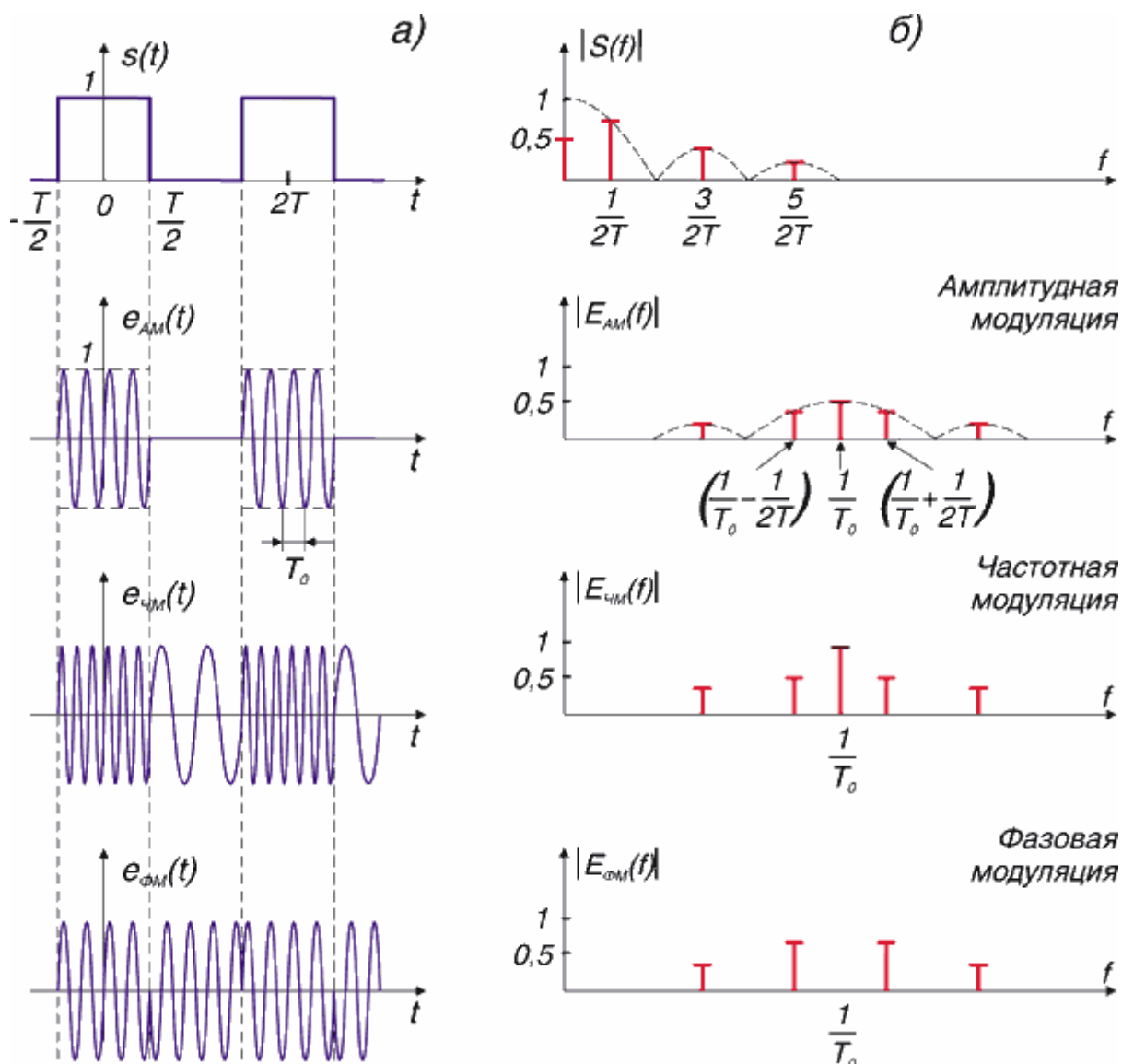


Рис. 2.1 Модуляция прямоугольными импульсами

Часто спектр видеосигнала называют однополосным, или находящимся в первичной полосе, а спектр радиосигнала – двухполосным.

Как видно, в спектре модулированных колебаний вокруг несущей располагаются целые боковые полосы, связанные со спектром модулирующего сигнала. При амплитудной модуляции верхняя боковая полоса является сдвинутой по оси частот

копией спектра модулирующего сигнала, нижняя - его зеркальной копией. Для частотной модуляции связь между спектром модулирующего сигнала и боковыми составляющими спектра модулированного колебания является более сложной. Однако можно отметить, что теоретически спектр модулированного колебания бесконечно широкий. Если принять меры по ограничению спектра модулирующего сигнала (как при передаче низкочастотного сигнала) и не использовать слишком глубокие частотную и фазовую модуляции (в техники связи говорят - использовать индекс модуляции, не превышающий единицу), то можно добиться того, что ширина спектра модулированного колебания будет приблизительно равна удвоенной частоте следования импульсов передаваемых данных, то есть $2/T$. Наличие двух боковых полос приводит к тому, что удельная скорость передачи данных при использовании модуляции оказывается в два раза меньше, чем при передаче низкочастотного сигнала.

Следует упомянуть, что в русской терминологии радиосвязи существует традиция использовать для модуляции цифровым сигналом термин «манипуляция».

Существуют три фундаментальных типа цифровой модуляции и один гибридный:

ASK – Amplitude shift keying (АМ, АМн – амплитудная манипуляция)

FSK – Frequency shift keying (ЧМ, ЧМн – частотная манипуляция).

PSK – Phase shift keying (ФМн – фазовая манипуляция).

ASK/PSK – комбинированный способ модуляции.

Каждая из схем имеет свои сильные и слабые стороны.

АМ (ASK) хороша с точки зрения эффективности использования полосы частот, но подвержена искажениям при наличии шума и недостаточно эффективна с точки зрения потребляемой мощности.

ЧМ (FSK) – с точностью до наоборот, энергетически эффективна, но не эффективно использует полосу частот.

ФМ (PSK) – энергетически эффективна, но частотная эффективность зависит от формы посылок.

АМ/ФМ (ASK/PSK) – комбинация двух схем. Она позволяет еще лучше использовать полосу частот.

2.2. Энергетическая и частотная эффективность

Хорошую частотную эффективность можно получить передачей амплитудно-манипулированных колебаний в одной боковой полосе (АМ/ОБП). Но этот тип модуляции, как мы видели, обладает низкой энергетической эффективностью.

В простейшем случае бинарной фазовой манипуляции – БФМ, которая реализует наиболее помехоустойчивый способ двоичной телеграфии, в отношении расходования частотного ресурса она, однако, оказывается чрезвычайно неэффективной. Дело в том, что спектр мощности сигнала БФМ совпадает по форме с энергетическим спектром посылки $S_0(t)$, и когда последняя прямоугольна, убывает с частотой f весьма медленно – пропорционально $1/f_0^2$. Если, как это часто делается в технике связи при регламентации использования радиоспектра, измерять полосу сигнала шириной окна Δf_{99} , в котором удерживается не менее 99% излучаемой мощности, для БФМ получится цифра, многократно превосходящая традиционный ориентир $1/T$: $\Delta f_{99} \approx 18.5/T$. По этой причине БФМ с прямоугольными посылками не применяется в цифровой связи.

Для повышения спектральной эффективности можно наметить несколько путей. Простейший из них состоит в увеличении длительности прямоугольной посылки T с сохранением прежней скорости передачи R_t в числе бит на единицу времени. При БФМ один бит передается за время $T_b = T$, так что $R_t = 1/T_b$. Для удержания этой скорости при

"удлинении" посылки следует увеличить число возможных значений фазы ϕ_i . Так, при удвоении длительности посылки ($T = 2T_b$) в течение временного отрезка T придется передавать 2 бита информации, т. е. 4 различных сообщения, что может быть достигнуто за счет использования четырех разрешенных значений фазы вместо двух, например $0, \pi, \pi/2, -\pi/2$. Такой способ манипуляции и называют квадратурной ФМ – КФМ (QPSK).

Можно видеть, что КФМ в 2 раза экономнее БФМ в отношении использования частотного ресурса, поскольку имеет спектр той же формы, но суженный вдвое за счет двукратного растяжения посылки. Особо подчеркнем, что указанный выигрыш достигнут без ухудшения помехоустойчивости приема. В самом деле, пусть энергия посылки при БФМ равна E_b . Тогда евклидово расстояние между противоположными посылками (геометрически представляющими собой противоположные векторы длины $\sqrt{E_b}$), определяющее вероятность их перепутывания, составит $2\sqrt{E_b}$ (рис. 2.2, а).

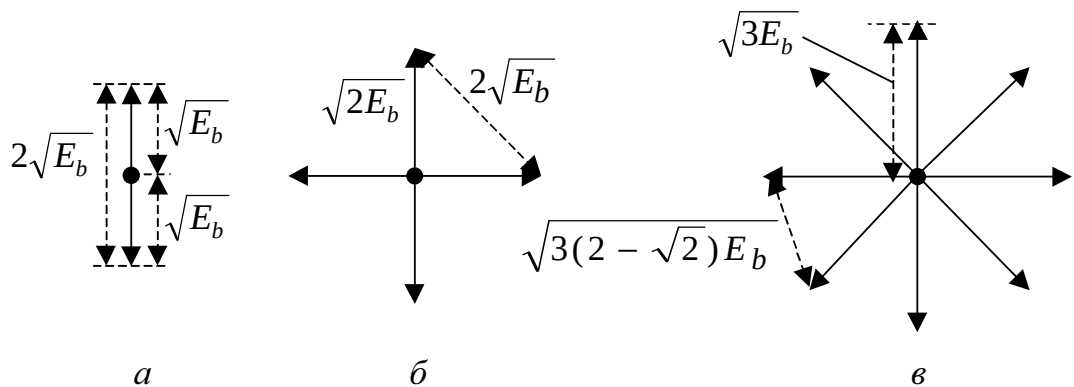


Рис. 2.2. Геометрическая интерпретация фазовой манипуляции

При КФМ четырем сообщениям отвечает четверка биортогональных векторов длины $\sqrt{2E_b}$ (рис. 2.2, б), так как при неизменной мощности энергия посылки E_q удвоится по сравнению со случаем БФМ за счет удвоения длительности: $E_q = 2E_b$. При этом расстояние между соседними векторами, определяющее наибольшую из вероятностей перепутывания сообщений, останется прежним ($\sqrt{2E_q} = 2\sqrt{E_b}$), что и означает отсутствие сколько-нибудь заметного ухудшения помехоустойчивости приема при переходе от БФМ к КФМ.

Не составляет труда убедиться, что при дальнейшем увеличении длительности посылки требование поддержания постоянства скорости приведет к сближению соседних векторов. Так, утроение длительности без снижения скорости означает передачу одной посылкой восьми сообщений, так что трехкратный рост энергии посылки по сравнению с БФМ будет нивелироваться уменьшением угла между соседними сигналами до 45° (рис. 2.2, в), т.е. уменьшением минимального евклидова расстояния до $\sqrt{3E_b(2-\sqrt{2})}$. Таким образом, в данном случае трехкратный выигрыш в полосе приобретается ценой энергетических потерь порядка 3,5 дБ (именно таким должно быть увеличение энергии, компенсирующее сближение векторов и снижающее вероятность ошибки до прежнего уровня). Дальнейшее повышение спектральной эффективности подобным способом окажется тем более невыгодным в плане энергозатрат: M -кратный выигрыш в полосе при 2М-ичной ФМ будет сопровождаться энергетическим проигрышем в

$$\gamma = \frac{2}{M(1 - \cos \frac{\pi}{2^{M-1}})} \text{ раз.}$$

2.3. Квадратурная амплитудная манипуляция

В принципе существует возможность определенного снижения упомянутых энергетических потерь за счет оптимизации созвездия сигнальных векторов на плоскости, максимизирующей минимальное расстояние между сигнальными векторами. При этом последние имеют неодинаковую длину, т.е. фазовая манипуляция дополняется параллельной амплитудной. Подобные форматы, известные под названиями амплитудно-фазовая и квадратурная амплитудная манипуляции (АФМ и КАМ – ASK/PSK и QAM), широко распространены во многих телекоммуникационных сетях (кабельная, радиорелейная связь и т. д.).

В КАМ для одновременной передачи двух отдельных k -битовых информационных блоков используется две несущих, находящихся в квадратуре ($\cos\omega_c t$ и $\sin\omega_c t$). Более высокие уровни КАМ строятся по тому же принципу, что и ФМ. Теоретическая эффективность использования полосы пропускания:

Формат	Эффективность (бит/с/Гц)
БФМ (BPSK)	1
КФМ (QPSK)	2
8-ФМ (8-PSK)	3
16-КАМ (16-QAM)	4
32-КАМ (32-QAM)	5
64-КАМ (64-QAM)	6
256-КАМ (256-QAM)	8

Чем сложнее схема модуляции, тем меньше минимальное расстояние между сигналами и тем более пагубное воздействие на нее оказывают искажения при передаче.

Теоретически возможны ФМ и КАМ схемы еще более высокого уровня, но на практике при их использовании возникает слишком большое количество ошибок.

Сигналы КАМ можно выразить так:

$$s_m(t) = \operatorname{Re} \left[(A_{mc} + jA_{ms}) g(t) e^{j\omega_c t} \right] = A_{mc} g(t) \cos \omega_c t - A_{ms} g(t) \sin \omega_c t,$$

$$m = 1, 2, \dots, M, \quad 0 \leq t \leq T,$$

где A_{mc} и A_{ms} – информационные амплитуды сигнала для квадратурных несущих, а $g(t)$ – форма импульса.

Сигнал КАМ можно также выразить и так:

$$s_m(t) = \operatorname{Re} \left[V_m e^{j\theta_m} g(t) e^{j\omega_c t} \right] = V_m g(t) \cos(\omega_c t + \theta_m), \frac{n!}{r!(n-r)!}$$

где $V_m = \sqrt{A_{mc}^2 + A_{ms}^2}$ и $\theta_m = \arctg(A_{ms} / A_{mc})$. Из этой формы представления видно, что сигнал КАМ можно рассматривать как комбинацию амплитудной и фазовой модуляции.

Действительно, мы можем образовать определенную комбинацию M_1 -уровневой АМ и M_2 -позиционной ФМ, чтобы сконструировать комбинированное АМ-ФМ сигнальное созвездие, содержащее $M = M_1 M_2$ точек пространства сигналов. Если $M_1 = 2^n$ и $M_2 = 2^m$, то сигнальное созвездие комбинированной АМ-ФМ сводится к мгновенной передаче $m+n=\log M_1 M_2$ двоичных символов, возникающих со скоростью $R/(m+n)$. Примеры сигнальных пространственных диаграмм для комбинированной АМ-ФМ показаны на рис. 2.3 для $M=8$ и $M=16$.

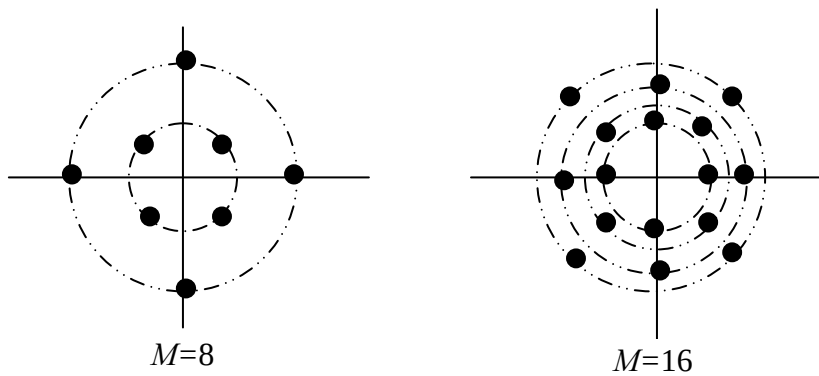


Рис. 2.3. Примеры пространственных диаграмм для комбинированной АМ-ФМ.

Для частного случая, когда амплитуда сигнала принимает ряд дискретных значений $\{(2m-1-M)d, m = 1, 2, \dots, M\}$, пространственная диаграмма сигналов является прямоугольной, как показано на рис. 2.4.

Т.о., цифровая модуляция несущей по амплитуде и фазе позволяет сконструировать сигналы, которые соответствуют двумерным векторам и пространственным диаграммам сигналов. Если мы хотим сконструировать сигнал, соответствующий вектору большей размерности, можно использовать или временную, или частотную, или обе области для того, чтобы увеличить размерность пространства.

Предположим, что мы имеем N -мерные сигнальные векторы. Для любого N можем разделить интервал времени длины $T_1 = T_1/N$. В каждом интервале длины T можем использовать двоичную АМ (БАМ – одномерный сигнал), чтобы передать элемент N -мерного сигнального вектора. Таким образом, N временных отрезков используется для передачи N -мерного сигнального вектора.

Если N четно, отрезок длиной T можно использовать для мгновенной передачи двух компонент N -мерного вектора путем независимой модуляции амплитуды квадратурных несущих соответствующими компонентами. Таким путем N -мерный сигнальный вектор передается за $NT/2$ секунд ($N/2$ временных отрезков).

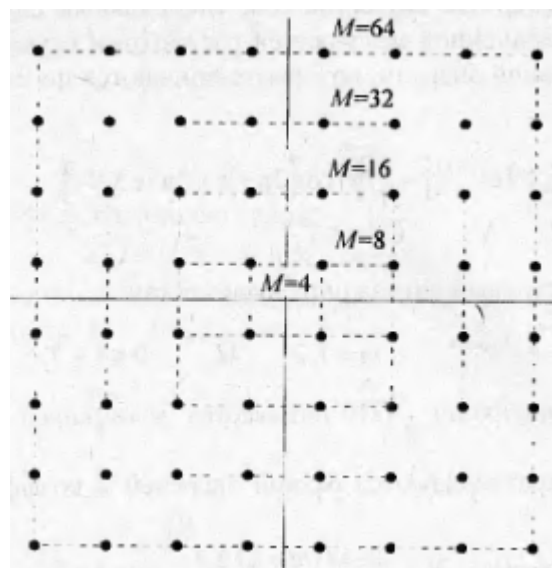


Рис. 2.4. Пространственные диаграммы для прямоугольной КАМ.

Альтернативно полоса частот $N\Delta f$ может быть подразделена на N частотных отрезков, каждый шириной Δf . N -мерный сигнальный вектор можно передать через канал путем одновременной (параллельной) модуляции амплитуды N несущих, одна на каждый из N частотных отрезков. Надо позаботиться о достаточном частотном разnose Δf между

смежными несущими с тем, чтобы не возникала взаимная интерференция между сигналами на N несущих. Если используются квадратурные несущие на каждом частотном отрезке, то N -мерный вектор (N – чётно) можно передать на $(1/2)N$ частотных отрезков, что сокращает используемую каналом полосу частот вдвое.

В более общем виде можно использовать совместно временную и частотную области для передачи N -мерного сигнального вектора. Так, рис. 2.5 иллюстрирует разделение частотно-временной области на 12 ячеек. Таким образом, можно передать $N=12$ -мерный сигнальный вектор при АМ или $N=24$ -мерный сигнальный вектор с использованием двух квадратурных несущих (КАМ) на каждом отрезке.

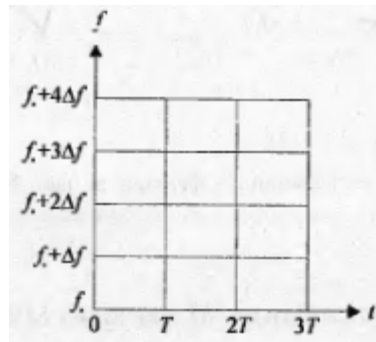


Рис. 2.5. Разделение осей времени и частоты на индивидуальные отрезки

Рассмотрим более подробно квадратурную модуляцию. В качестве общей модели сигнала примем последовательность одинаковых по форме импульсов (посылок), повторяющихся с постоянным интервалом T :

$$s(t) = \sum_{i=-\infty}^{\infty} S_0(t - iT) \cos(2\pi f_0 t + \varphi_i), \quad (2.1)$$

где $S_0(t)$ – огибающая посылки; φ_i – начальная фаза i -й посылки.

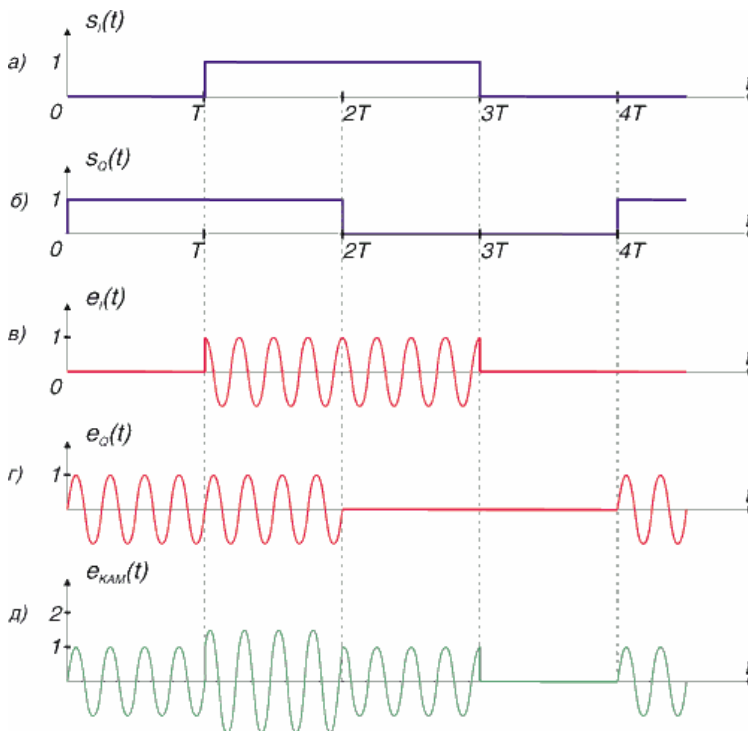


Рис. 2.6. Квадратурная модуляция прямоугольными импульсами

Временные диаграммы, иллюстрирующие квадратурную модуляцию, приведены на рис. 2.6. Сигнал исходных передаваемых данных разделяется на два потока - s_I и s_Q (рис. 2.6, а и 2.6, б). Сигнал s_Q модулирует синусоидальное несущее колебание, а сигнал s_I - косинусоидальное (рис. 2.6, в и 2.6, г). Затем два модулированных колебания складываются, образуя единое квадратурно-модулированное колебание (рис. 2.6, д). Как видно, модуляция является сложной: в моменты T и $2T$ мы видим как скачки амплитуды, так и скачки фаз.

Полезно рассмотреть работу квадратурного модулятора на примере схемы формирования сигналов четырехфазной ФМ из потока двухбитовых символов (рис.2.7).

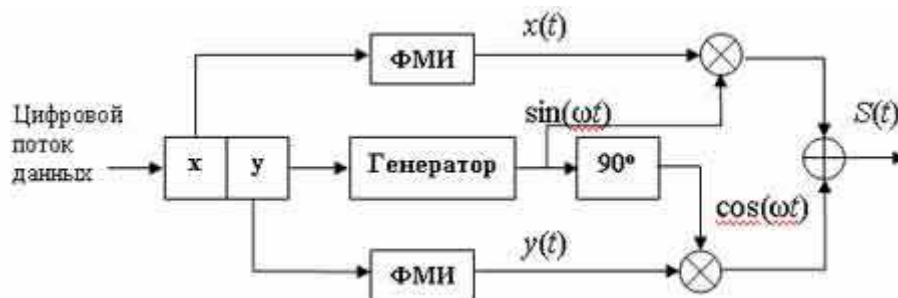


Рис. 2.7. Схема квадратурного модулятора

Исходная последовательность двоичных символов длительностью T при помощи регистра сдвига разделяется на нечетные y и четные x импульсы, которые поступают на входы формирователей манипулирующих импульсов (ФМИ) соответственно квадратурного ($\cos\omega t$) и синфазного ($\sin\omega t$) каналов. На выходах ФМИ образуются последовательности биполярных импульсов $x(t)$ и $y(t)$ с амплитудой $\pm A_m$ и длительностью $2T$, которые поступают на входы каналов перемножителей, где они независимо друг от друга модулируют по амплитуде два одинаковых несущих колебания, сдвинутых по фазе на 90° , т.е. находящихся в квадратуре. В результате, на их выходах формируются двухфазные $(0, \pi)$ ФМ колебания. После суммирования они образуют сигнал ФМ-4 или квадратурный ФМ-сигнал (Quadrature Phase Shift Keying – QPSK). Поскольку в каждом канале осуществляется амплитудная манипуляция, этот вид модуляции называют еще квадратурной амплитудной манипуляцией (QASK – Quadrature Amplitude Shift Keying) или просто квадратурной амплитудной модуляцией (КАМ).

Разделение двух модулирующих сигналов s_I и s_Q на приемной стороне возможно с помощью синхронных детекторов, так как квадратурные несущие ортогональны, то есть их среднее (по времени) произведение равно нулю. В синхронном детекторе квадратурно-модулированное колебание умножается на косинусоидальный сигнал и результат перемножения усредняется во времени. В результате подавляется квадратурная компонента (e_Q) и выделяется огибающая синфазной (s_I). Аналогичным образом выделяется и огибающая квадратурной компоненты модулированного колебания s_Q .

Квадратурно-модулированное колебание, изображенное на рис. 2.6, представлено в векторной форме на рис. 2.8. Как видно, пространство комплексной плоскости используется не слишком эффективно - занят только один квадрант. На рис. 2.9 показано квадратурно-модулированное колебание, вектор которого при модуляции двух квадратурных компонентов занимает также 4 точки, но уже в 4 квадрантах, что повышает помехоустойчивость системы модуляции.

Временные диаграммы сигналов, соответствующих этим четырем положениям вектора, также приведены на рисунке. Поскольку амплитуды всех сигналов одинаковы, то такую модуляцию можно назвать и четырехпозиционной фазовой модуляцией, или

манипуляцией (QPSK). Развивая этот подход, можно прийти и к другим системам многопозиционной квадратурной модуляции.

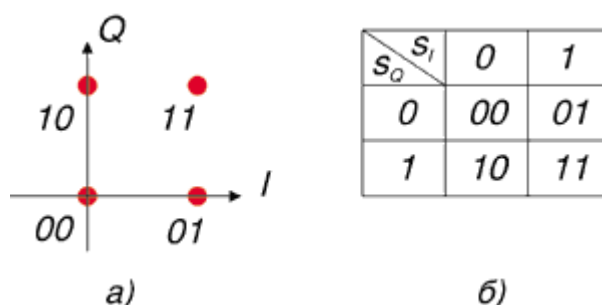


Рис. 2.8 Сигнальная диаграмма КАМ, соответствующая рис. 2.6

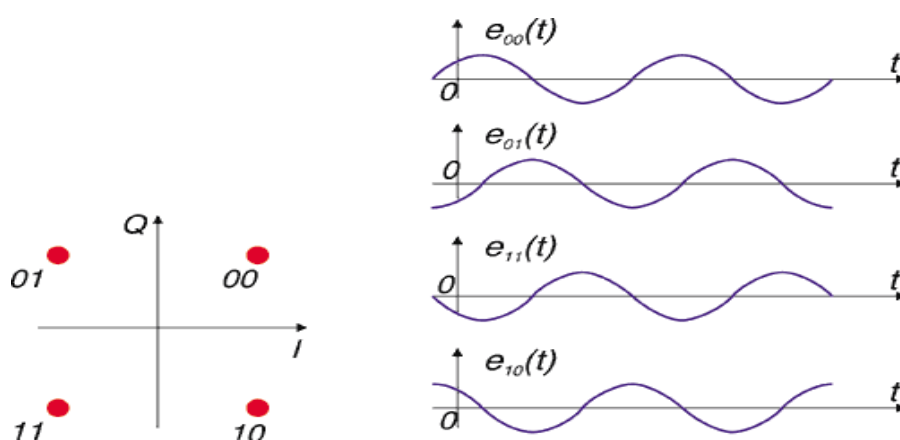


Рис. 2.9 КАМ с более высокой помехоустойчивостью

При одновременной смене символов в обоих каналах модулятора (с 10 на 01, или с 00 на 11) в сигнале ФМ-4 происходит скачок фазы на 180° (π). Такие скачки фазы вызывают паразитную амплитудную модуляцию огибающей сигнала. В результате этого при прохождении сигнала через узкополосный фильтр возникают провалы огибающей несущего колебания до нуля. Такие изменения сигнала нежелательны, поскольку приводят к увеличению энергии боковых полос и помех в канале связи.

Другой негативной стороной скачков амплитуд и фаз являются снижение КПД передатчика, что особенно критично для систем цифровой мобильной связи.

Поясним подробнее смысл последнего фактора.

Энергопотребление передатчика в значительной мере зависит от режима оконечного усилителя мощности или, иначе говоря, от пик-фактора сигнала, равного отношению его пиковой и средней мощностей. Наиболее благоприятным считается режим класса С, когда активный элемент усилителя постоянно находится вблизи точки насыщения, выполняя, по существу, функции ключа, и требования к его линейному динамическому диапазону минимальны. Для приближения к подобному режиму передаваемый сигнал должен быть свободен от глубокой амплитудной модуляции, т.е. иметь пик-фактор, близкий к единице. Данное условие, в свою очередь, означает, что скачки мгновенной фазы, сопровождающие переход от данной посылки к последующей, должны быть по возможности минимизированы.

Обратимся за подтверждением к рис. 2.10. Когда значения передаваемых символов 0 и 1 при БФМ (рис. 2.10, а) равновероятны, в среднем в половине переходов между последовательными посылками наблюдается скачок фазы на 180° , т.е. смена полярности на противоположную. Изменение мгновенной комплексной огибающей от A до $-A$ или наоборот должно быть без искажений передано усилителем мощности, которому тем самым навязывается чрезмерное требование в части линейного динамического диапазона ($2A$). Поскольку при линейной передаче перепадов большая часть мощности рассеивается на активном элементе усилителя, становится понятным желание снизить частоту и размах скачков фазы в манипулированном сигнале. Переход от БФМ к стандартной КФМ (рис. 2.10, б) с сохранением прежней мощности не ослабляет требований к динамическому диапазону усилителя, поскольку максимальные скачки фазы в 180° сохраняются, хотя и случаются значительно (примерно вдвое) реже.

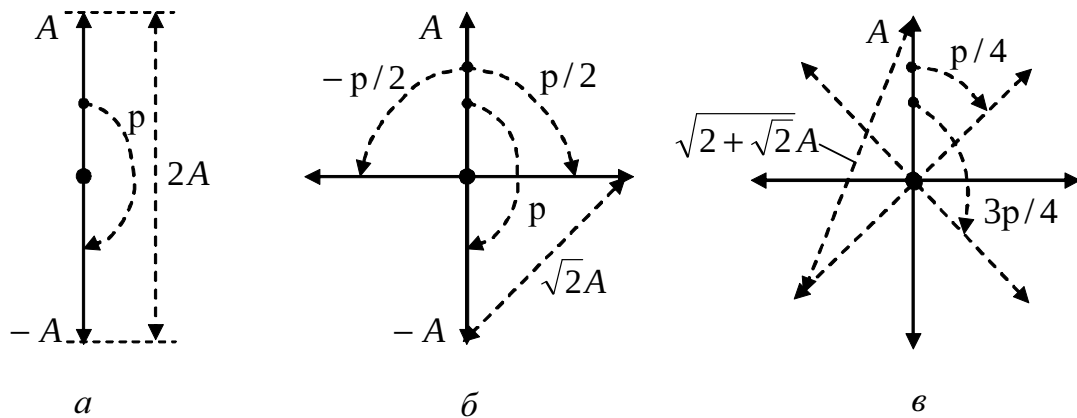


Рис. 2.10. Перепады мгновенной амплитуды при различных видах ФМ

2.4. Квадратурная фазовая манипуляция со сдвигом

Смягчение требований к динамическому диапазону передатчика достигается в формате КФМ *со сдвигом* (OQPSK – *offset QPSK*), для понимания смысла которого разумно сначала ввести в (2.1) начальную фазу в $\pi/4$ (разумеется, это никак не повлияет на свойства КФМ и лишь повернет целиком диаграмму рис. 2.10, б на 45°), а затем переписать результат в форме:

$$s(t) = \sum_{i=-\infty}^{\infty} S_0(t-iT) \cos(2\pi f_0 t + \varphi_i + \frac{\pi}{4}) =$$

$$= \frac{1}{\sqrt{2}} \left[\sum_{i=-\infty}^{\infty} a_i S_0(t-iT) \cos(2\pi f_0 t) + \sum_{i=-\infty}^{\infty} b_i S_0(t-iT) \sin(2\pi f_0 t) \right],$$

где $a_i = \cos \varphi_i - \sin \varphi_i$ и $b_i = -\cos \varphi_i - \sin \varphi_i$ – бинарные символы, принимающие значения +1 и -1. Как видно, КФМ сигнал можно интерпретировать как сумму двух квадратурных БФМ сигналов, что дает возможность демодулировать каждую из последовательностей символов $\{a_i\}, \{b_i\}$ отдельно от другой, поскольку синусная и косинусная квадратурные компоненты ортогональны и не создают друг другу перекрестных помех при когерентном приеме. Сдвинем теперь синусную квадратурную составляющую во времени на половину длительности посылки, придя к результату

$$s(t) = \frac{1}{\sqrt{2}} \left[\sum_{i=-\infty}^{\infty} a_i S_0(t - iT) \cos(2\pi f_0 t) + \sum_{i=-\infty}^{\infty} b_i S_0(t - iT - \frac{T}{2}) \sin(2\pi f_0 t) \right], \quad (2.2)$$

сохраняющему ортогональность квадратурных компонент и, следовательно, все характеристики обычной КФМ в части качества передачи данных. Теперь, однако, изменения бинарных символов, манипулирующих квадратурные несущие, происходят неодновременно: в момент смены символа a_i символ другой квадратуры b_i остается неизменным и наоборот. Благодаря этому при каждой смене символа a_i или b_i сигнальный вектор на рис. 2.10, б может перейти только в соседний, но никак не в противоположный. Тем самым требуемый линейный динамический диапазон $\sqrt{2}A$ оказывается в $\sqrt{2}$ раз меньшим по сравнению со случаем БФМ или стандартной КФМ. Именно по этой причине КФМС была избрана для построения обратного (от МС к БС) канала в CDMA стандарте второго поколения IS-95.

Другой вариант достижения той же цели реализован в цифровых стандартах D-AMPS (США, известен также под названиями IS-136 и US-TDMA) и PDC (Япония, прежнее наименование JDC), где вместо сдвига посылок введен поворот на угол $\pi/4$ алфавита значений фаз при переходе от четных посылок к нечетным. Благодаря такому смещению при $i = 2k, k = \dots, -1, 0, 1, \dots$ φ_i в (2.1) принимает значения из множества $0, \pi, \pm\pi/2$, а при $i = 2k + 1$ – из множества $\pm\pi/4, \pm3\pi/4$ (пунктир на рис. 2.10, в). Разумеется, на приемной стороне этот сдвиг легко учитывается, так что демодуляция выполняется практически так же, как при обычной КФМ. Подобный вид КФМ получил название $\pi/4$ -КФМ ($\pi/4$ -QPSK). Его преимущество в сравнении с КФМС состоит именно в отсутствии усложнений демодулятора, свойственных КФМС, хотя в отношении смягчения требований к линейному динамическому диапазону $\pi/4$ -КФМ не столь эффективна: переходам на угол $\pm3\pi/4$ соответствует перепад комплексной огибающей, равный $\sqrt{2} + \sqrt{2}A$ и примерно в 1,31 раза превышающий аналогичный показатель КФМС.

2.5. МЧМ – минимальная частотная манипуляция

Рассмотренные трансформации КФМ никоим образом не влияют на спектральные характеристики модулированного сигнала, которые в рамках случайной модели передаваемого потока данных определяются исключительно формой посылки. Поскольку причиной медленного спада спектра мощности сигналов (2.1) и (2.2) с частотой является разрывность прямоугольной посылки, радикального сужения спектра можно добиться только за счет сглаживания ее формы. Если, в частности, принять в (2.2) за посылку импульс с огибающей в виде положительной полувольты косинуса (см. рис 2.11) с амплитудой $\sqrt{2}A$

$$S_0(t) = \begin{cases} \sqrt{2}A \cos(\frac{\pi t}{T}), & |t| \leq \frac{T}{2}, \\ 0, & |t| > \frac{T}{2}, \end{cases} \quad (2.3)$$

КФМ преобразуется в свой подвид, известный как *минимальная частотная манипуляция* – МЧМ (MSK – *minimal shift keying*).

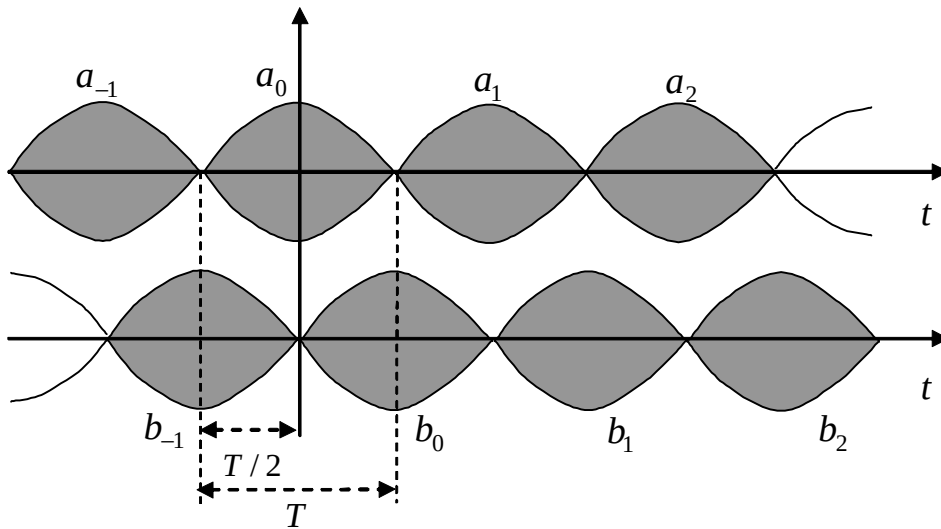


Рис 2.11. Квадратурные составляющие МЧМ сигнала

Для того, чтобы убедиться в преимуществах МЧМ и попутно выяснить корни ее наименования, найдем комплексную огибающую $\dot{S}(t)$ сигнала (2.2) с формой посылки, задаваемой (2.3). При этом в силу полной однородности поведения сигнала во времени достаточно рассмотреть лишь два смежных отрезка $[-T/2, 0]$ и $[0, T/2]$. Используя элементарную тригонометрию, формулу Эйлера и бинарность символов a_i, b_i , в силу которой $a_i / b_i = a_i b_i = \pm 1$, получаем

$$\dot{S}(t) = \begin{cases} a_0 A \cos(\frac{\pi t}{\Delta}) + j b_{-1} A \cos[\frac{\pi}{T}(t + T - \frac{T}{2})] = a_0 A \exp(-j \frac{\pi a_0 b_{-1} t}{T}), & -\frac{T}{2} \leq t < 0, \\ a_0 A \cos(\frac{\pi t}{\Delta}) + j b_0 A \cos[\frac{\pi}{T}(t - \frac{T}{2})] = a_0 A \exp(j \frac{\pi a_0 b_0 t}{T}), & 0 \leq t < \frac{T}{2}, \end{cases} \quad (2.4)$$

Как показывает (2.4), действительная огибающая $S(t) = |\dot{S}(t)| = A$ постоянна и, следовательно, МЧМ не сопровождается амплитудной модуляцией, обеспечивая равенство пик-фактора сигнала единице и, как результат, оптимальность режима усилителя мощности передатчика.

Другой вывод из (2.4) состоит в том, что рассматриваемый вид модуляции сводится, по существу, к бинарной частотной манипуляции прямоугольных посылок длительности $T/2$, поскольку линейное изменение фазы с угловым коэффициентом π/T означает сдвиг частоты на $\pm 1/2T$. Принципиальным является то, что переключение частоты между посылками каждые $T/2$ с происходит без скачков фазы, т.е. МЧМ является модуляцией с непрерывной фазой. В этом легко убедиться, подставив $t = 0$ в обе строки (2.4) и придя к одному и тому же результату:

$$\dot{S}(t = 0^-) = a_0 = \dot{S}(t = 0^+),$$

где верхние индексы "−" и "+" символизируют приближение к точке $t = 0$ соответственно слева и справа.

Таким образом, при любом текущем передаваемом символе очередная посылка начинается с той фазы, которая "набежала" в течение предыдущей. Сказанное иллюстрируется "деревом" траекторий фазы на рис 2.12. В течение каждого отрезка

$[kT/2, (k+1)T/2]$ фаза линейно растет или убывает в соответствии с текущим приращением частоты $\pm 1/2T$ (определяемым, в свою очередь, комбинацией символов a_i, b_i сигнала (2.2) на данном отрезке). Значения фазового угла в момент $(k+1)T/2$, отвечающие двум возможным частотам, отличаются ровно на π . При этом любая из возможных траекторий фазы оказывается непрерывной функцией.

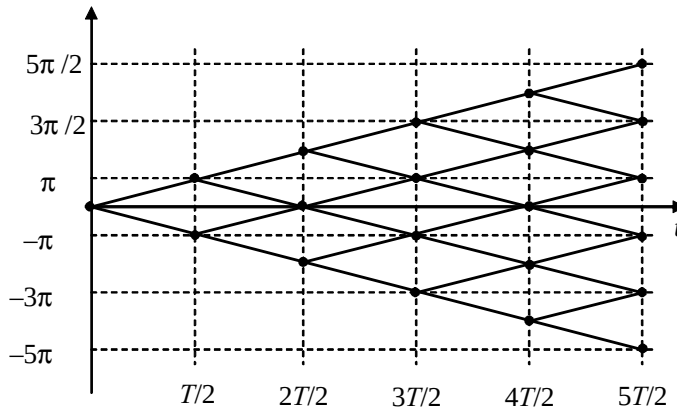


Рис. 2.12. Траектории мгновенной фазы МЧМ сигнала

Отсутствие разрывов фазы обеспечивает спектру МЧМ значительно большую степень компактности по сравнению со стандартной КФМС: спектр мощности МЧМ сигнала убывает примерно пропорционально $1/f^4$, так что занимаемая им полоса сокращается более чем на порядок:

$$\Delta f_{99} \approx \frac{1.2}{T_b}.$$

Резервы дальнейшего сжатия спектра модулированного сигнала кроются в устранении разрывов не только самой фазы, но и ее производных (частоты, скорости изменения частоты и т.д.). Иными словами, линейно-ломаные траектории на рис. 2.12 могут быть заменены более гладкими. В стандарте GSM применен именно такой вариант модуляции – *гауссовская* МЧМ (GMSK – *Gaussian* MSK), при которой закон изменения фазы в течение посылки повторяет ход гауссовской интегральной функции распределения, чем обеспечивается плавность изменения фазы и частоты, а значит, высокая степень компактности спектра. Технически гауссовская (как и обычная) МСК может быть реализована разными средствами, однако для объяснения смысла ее параметров согласно спецификации GSM разумно считать, что исходный поток битовых прямоугольных посылок длительности T_b пропускается через низкочастотный фильтр с гауссовской амплитудно-частотной характеристикой и полосой W (на уровне -3 дБ), после чего сглаженный сигнал модулирует частоту задающего генератора. В стандарте жестко зафиксировано значение $WT_b = 0,3$, отвечающее полосе

$$\Delta f_{99} \approx \frac{0.92}{T_b}.$$

С упрощениями, игнорирующими второстепенные детали, дерево траекторий фазы гауссовской МЧМ показано на рис. 2.13.

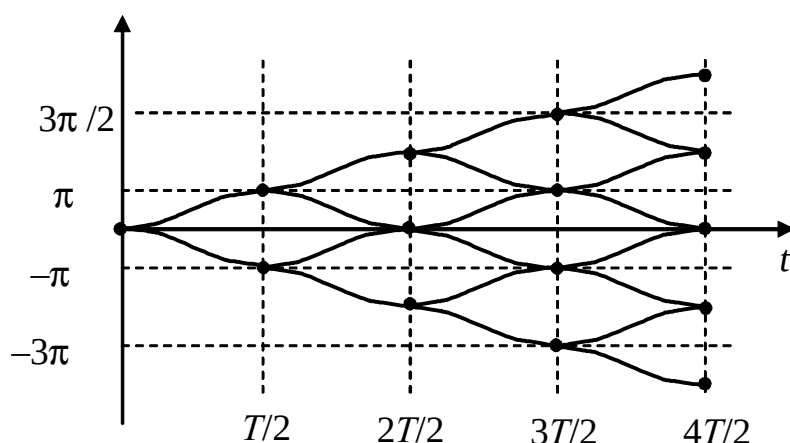


Рис. 2.13. Траектории фазы при гауссовской МЧМ

Отметим, что в эволюции видов модуляции от БФМ до МЧМ обмен между спектральной эффективностью и достоверностью передачи фактически отсутствовал: повышение спектральной эффективности не сопровождалось снижением минимального евклидова расстояния в сигнальном созвездии, что было продемонстрировано на примере сравнения КФМ с БФМ. Случай гауссовской МЧМ несколько выпадает из этого ряда в связи с тем, что неограниченность отклика гауссовского фильтра во времени не позволяет, строго говоря, трактовать данный вид модуляции как сглаженную версию КФМС. В литературе можно встретить упоминание об энергетическом проигрыше гауссовской МЧМ относительно БФМ в 0,46 дБ, однако не исключено, что ее теоретический потенциал выше с учетом возможности демодуляции сообщения "в целом", т.е. на интервале, охватывающем несколько посылок (например, с помощью алгоритма Витерби).

Приведем итоговую таблицу, связывающую ширину полосы Δf_{99} модулированного сигнала со скоростью передачи данных R_t (бит/с) для рассмотренных форматов.

Таблица 2.1. Соотношения между полосой и скоростью передачи для различных видов модуляции

Вид модуляции	БФМ	КФМ, КФМС, $\pi/4$ -КФМ,	МЧМ	гауссовская МЧМ
$\Delta f_{99} / R_t$	18,5	9,2	1,2	0,92

Как показывает приведенный обзор, применяемые методы модуляции отличаются заметным разнообразием. Их перечень станет еще более обширным, если обратиться и к другим телекоммуникационным системам (модемной, радиорелейной, спутниковой связи, персонального вызова, бесшнурового телефона и т. п.). Все это лишний раз свидетельствует о неоднозначности и многовариантности задач идеологического проектирования и существовании параллельных путей достижения оптимальных системных показателей.

2.6. Модуляторы КАМ высоких порядков

При квадратурной амплитудной модуляции формируются два логически независимых канала, т. е. единичному уровню в одном канале может соответствовать единичный или нулевой уровень в другом канале. Благодаря этому два выходных сигнала не влияют друг на друга при прохождении по одной и той же физической среде.

В общем случае, для системы, поддерживающей m амплитудных уровней для каждого потока двоичных символов, можно образовать m^2 различных комбинаций нуля и единицы. Аналогично, диаграмма, отображающая все возможные значения сигнала (комбинации нуля и единицы) системы многоуровневой модуляции, называется диаграммой конstellации (constellation – созвездие) или диаграммой совокупности состояний.

На рис. 2.14 показана структурная схема модулятора и диаграмма состояний (сигнальное созвездие) системы КАМ-16, в которой $x(t)$ и $y(t)$ принимают значения $\pm 1, \pm 3$ (4-х уровневая КАМ). Как видно из рис. 2.14, созвездие содержит 16 сигнальных точек, каждая из которых соответствует четырем передаваемым информационным битам.

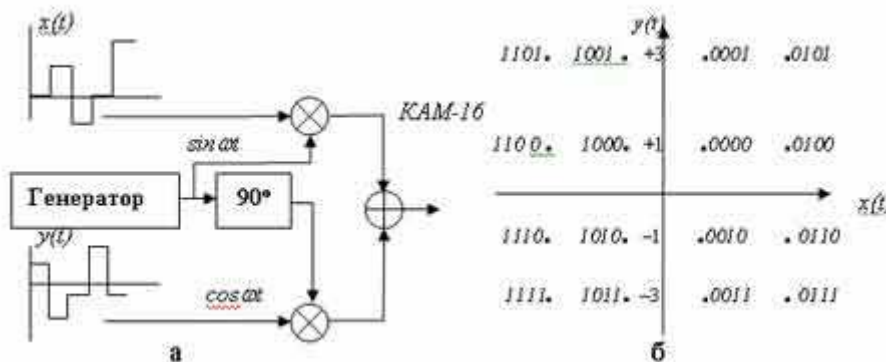


Рис. 2.14. Схема модулятора и сигнальная диаграмма КАМ-16

Существует несколько способов практической реализации 4-уровневой КАМ. Наиболее распространенным из них является так называемый способ модуляции наложением (SPM – Superposed Modulation).

В схеме, реализующей данный способ, используются два одинаковых 4-фазных модулятора. Структурная схема модулятора SPM и диаграммы, поясняющие его работу, приведены на рис. 2.15.

Если задается шесть амплитудных уровней ($\pm 1, \pm 3, \pm 5$) для каждого потока, то можно получить систему КАМ с совокупностью состояний, равной $6^2=36$. Однако из них в протоколах ITU-T используется только 32 равномерно распределенных в сигнальном пространстве точек.

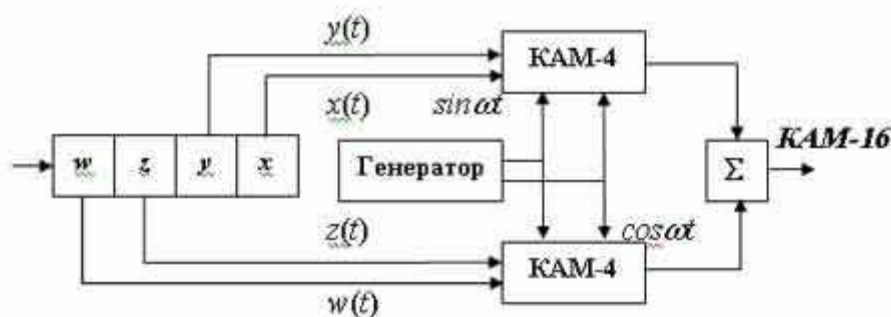


Рис. 2.15. Схема модулятора SPM

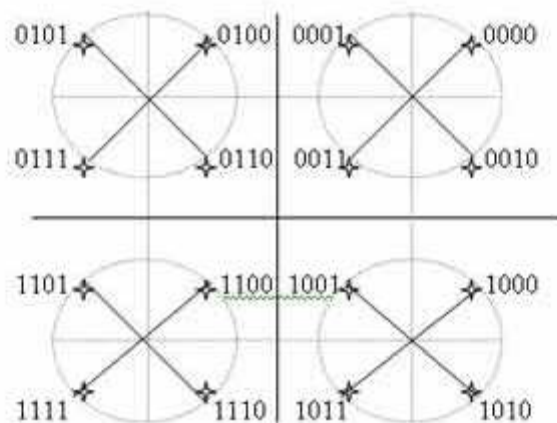


Рис. 2.16. Сигнальная диаграмма модулятора КАМ-16

В общем случае, при формировании сигналов многопозиционной QAM модуляция ортогональных сигналов осуществляется в цифровом виде. Для этих целей используется два цифровых полосовых фильтра с одинаковой амплитудой входных колебаний, но различающихся фазовым сдвигом в 90° . Уровни усиления амплитуды для каждого потока устанавливают независимо. Для системы, поддерживающей m амплитудных уровней для каждого потока можно образовать m^2 различных комбинаций нуля и единицы.

Из теории связи известно, что при равном числе точек в сигнальном созвездии спектр сигналов КАМ идентичен спектру сигналов ФМ. Однако помехоустойчивость систем ФМ и КАМ различна. При одинаковом числе точек сигналы системы КАМ имеют лучшую помехозащищенность, чем сигналы системы ФМ. Основная причина этого состоит в том, что расстояние между сигнальными точками в системе ФМ меньше расстояния между сигнальными точками в системе КАМ.

Расстояние d между соседними точками сигнального созвездия в системе КАМ с L уровнями модуляции определяется выражением: $d = \sqrt{2} / (L - 1)$

Аналогично при ФМ: $d = 2 \sin(\pi / M)$, где M – число фаз. Так, например, при $M=16$ ($L=4$) $d_{\text{кам}}=0,471$ и $d_{\text{фм}}=0,390$, а при $M=32$ ($L=6$) $d_{\text{кам}}=0,283$; $d_{\text{фм}}=0,196$.

2.7. OFDM

Новым способом модуляции, предложенным сравнительно недавно, является OFDM (Orthogonal Frequency Division Multiplexing) - частотное уплотнение с ортогональными несущими. Подобно квадратурной модуляции, этот способ использует

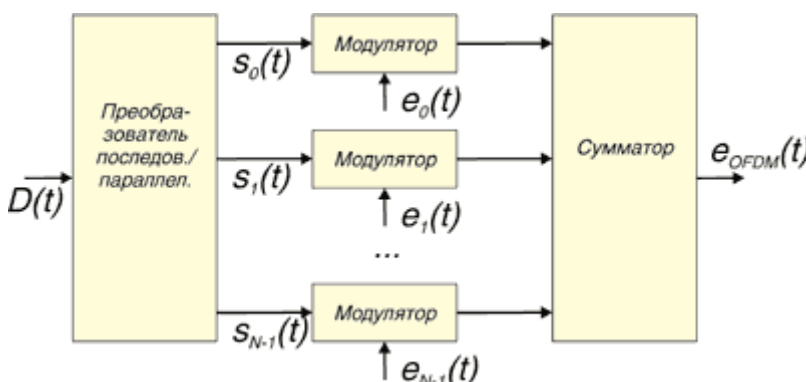


Рис. 2.17. Модуляция частотным уплотнением

ортогональные несущие, но в отличие от квадратурной модуляции частоты этих несущих не являются одинаковыми, они расположены в некотором диапазоне, отведенном для передачи данных путем модуляции. Частоты несущих соответствуют уравнению $e_n(t) = \cos(2\pi(f_0 + n/T_s)t)$, где f_0 - начало интервала, в котором

производится частотное уплотнение, n - номер несущей, находящийся в диапазоне от 0 до $(N-1)$, то есть всего несущих - N , T_s - длительность интервала передачи одного символа.

Как нетрудно убедиться, эти колебания действительно являются ортогональными, то есть их среднее произведение равно нулю, что, как и в случае квадратурной модуляции, означает возможность их разделения на приеме даже при частичном перекрытии их боковых полос.

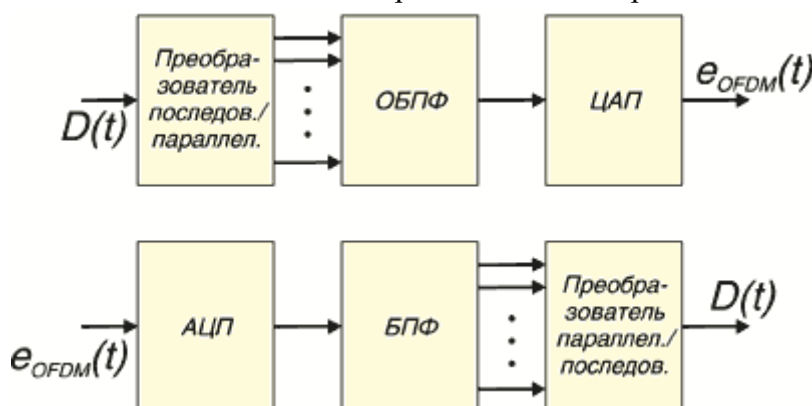


Рис. 2.18. Пример реализации частотного уплотнения

Схема, иллюстрирующая принцип модуляции с частотным уплотнением, приведена на рис. 2.17. Сигнал исходных данных разделяется на N потоков, трансформируясь в параллельную форму. Каждый из параллельных сигналов поступает на свой модулятор, в котором одна из

ортогональных несущих подвергается модуляции. После сложения модулированных

колебаний формируется сигнал OFDM.

Преимущества, заложенные в способе модуляции OFDM, проявляются при большом числе несущих. Например, если несущих 8000, то скорость данных, модулирующих каждую несущую, уменьшается в 8000 раз. Следовательно, в 8000 раз увеличивается длительность передачи одного символа (правда, импульс становится сложным). При этом появляется возможность вставить еще и защитный интервал между передачей соседних символов, что позволяет бороться с межсимвольными искажениями. Еще к преимуществам этого способа относится возможность борьбы с многолучевым приемом. Реализовать схему рис. 2.17 при таком числе несущих ранее было практически невозможно. Но разработки алгоритмов и интегральных схем быстрого преобразования Фурье позволили решить эту проблему (рис. 2.18). Ведь перемножение некоторых коэффициентов на гармонические колебания разных частот, удовлетворяющих вышеприведенным условиям, и суммирование полученных произведений представляет собой не что иное, как вычисления обратного преобразования Фурье (на схеме рис. 22 соответствующий блок обозначен как ОБПФ - обратное быстрое преобразование Фурье), коэффициентами для вычисления которого являются распараллеленные потоки данных. Поскольку все вычисления производятся в цифровой форме, то на выходе появляется цифроаналоговый преобразователь. Демодуляция может быть построена на базе прямого преобразования Фурье (БПФ). Естественно, что на входе должен стоять аналого-цифровой преобразователь.

2.8. Треллис-модуляция

Применение многопозиционной КАМ в чистом виде сопряжено с проблемой недостаточной помехоустойчивости. Поэтому во всех высокоскоростных протоколах КАМ используется совместно с решетчатым кодированием - специальным видом сверточного кодирования. В результате появился новый способ модуляции, называемый треллис-модуляцией (TCM - Trellis Coded Modulation, trellis - решетка).

Выбранная определенным образом комбинация конкретной КАМ и помехоустойчивого кода в отечественной технической литературе носит название

сигнально-кодовой конструкции (СКК), которая позволяет повысить помехозащищенность передачи информации одновременно со снижением требований к отношению сигнал-шум на 3-6 дБ. Это объясняется увеличением вдвое числа сигнальных точек за счет добавления к информационным битам одного избыточного, образованного путем сверточного кодирования. Расширенный таким образом блок битов подвергается все той же КАМ. В процессе демодуляции принятый сигнал декодируется по алгоритму Витерби. Именно этот алгоритм за счет введенной избыточности и знания предыстории процесса приема позволяет по критерию максимального правдоподобия выбрать из сигнального пространства наиболее достоверную точку.

Все применяемые сегодня СКК используют сверточное кодирование со скоростью $(n-1)/n$, т.е. при передаче одного сигнального элемента используется только один избыточный двоичный символ. Типичная схема сверточного кодирования со скоростью, равной $2/3$, приведена на рис. 2.19.

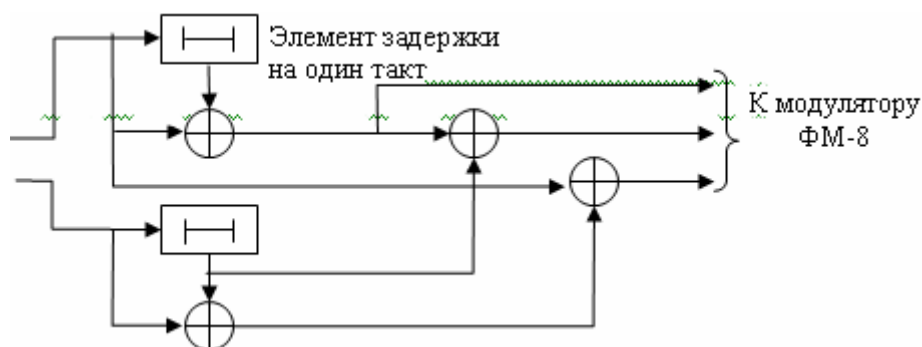


Рис. 2.19. Схема сверточного кодера (со скоростью кода $2/3$)

В схеме, представленной на рис. 2.19, сверточный кодер с относительной скоростью кода, равной $2/3$, используется совместно модулятором ФМ-8. Каждым двум информационным битам на входе кодер сопоставляет трехсимвольные двоичные блоки на выходе, которые поступают на модулятор ФМ-8.

Применение сигналов ФМ связано с разрешением проблемы неоднозначности фазы восстановленной на приеме несущей. Данная проблема решается за счет относительного (дифференциального) кодирования, что в системах передачи без помехоустойчивого кодирования приводит к размножению ошибок. В системах с помехоустойчивым кодированием относительное кодирование также используется. В этом случае имеет значение последовательность включения относительного и помехоустойчивого кодеров.

Различают внешнее и внутреннее относительное кодирование. При внутреннем кодировании относительный кодер расположен на выходе помехоустойчивого кодера, а на приемной стороне относительный декодер включен на входе помехоустойчивого декодера (рис. 2.20, а). Внешнее относительное кодирование в ряде случаев является более выгодным, так как источник размножения ошибок — относительный декодер — включен на выходе помехоустойчивого декодера (рис. 2.20, б). Однако при этом возникают трудности декодирования, вызванные неоднозначностью фазы опорного колебания при демодуляции. При ФМ-2 неоднозначность фазы опорного колебания (0 или π) приводит к явлению «обратной работы», заключающейся в том, что передаваемые единичные биты принимаются нулевыми, а нулевые — наоборот единичными.

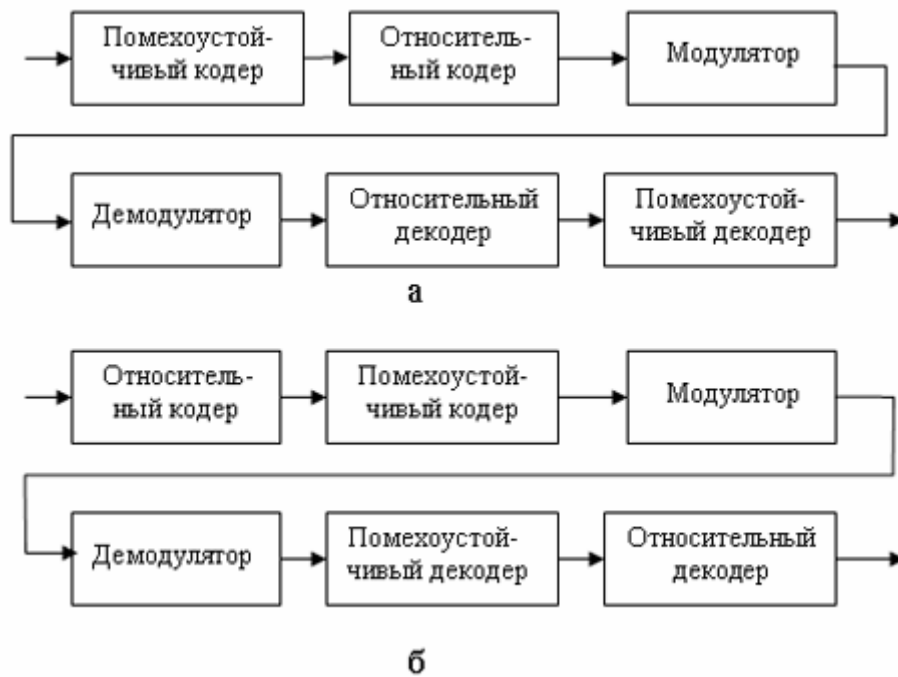


Рис. 2.20. Схема включений относительного кодера:
а) внутреннего; б) внешнего

При большем числе позиций фазы возможна не только инверсия, но и перестановка двоичных символов. Решение этой проблемы заключается в использовании помехоустойчивых кодов, «прозрачных», т.е. нечувствительных к неопределенности фазы опорного колебания.

3. Кодирование в ЦСС

Вопросы построения корректирующих кодов и их декодирования подробно рассматривались в курсе «Теория информации и кодирование». Поэтому в этой главе приводятся краткие сведения о помехоустойчивых кодах и области их применения в цифровых системах связи. Относительно подробно будут рассмотрены сверточные коды и методы их декодирования, в частности, алгоритм Витерби.

3.1. Корректирующие коды

Корректирующие коды строятся так, что для передачи сообщения используются не все кодовые комбинации mn (m – основание кода, n – длина кодового слова), а лишь некоторая часть их (так называемые разрешенные кодовые комбинации). Тем самым создается возможность обнаружения и исправления ошибки при неправильном воспроизведении некоторого числа символов. Корректирующие свойства кодов достигаются введением в кодовые комбинации дополнительных (избыточных) символов.

Декодирование состоит в восстановлении сообщения по принимаемым кодовым символам. Устройства, осуществляющие кодирование и декодирование, называют соответственно кодером и декодером. Как правило, кодер и декодер выполняются физически в одном устройстве, называемым кодеком.

Рассмотрим основные принципы построения корректирующих кодов или помехоустойчивого кодирования.

Напомним, что расстоянием Хэмминга между двумя кодовыми n -последовательностями, b_i и b_j , которое будем далее обозначать $d(i; j)$, является число разрядов, в которых символы этих последовательностей не совпадают.

Говорят, что в канале произошла ошибка кратности q , если в кодовой комбинации q символов приняты ошибочно. Легко видеть, что кратность ошибки есть не что иное, как расстояние Хэмминга между переданной и принятой кодовыми комбинациями, или, иначе, вес вектора ошибки.

Рассматривая все разрешенные кодовые комбинации и определяя кодовые расстояния между каждой парой, можно найти наименьшее из них $d = \min d(i; j)$, где минимум берется по всем парам разрешенных комбинаций. Это минимальное кодовое расстояние является важным параметром кода. Очевидно, что для простого кода $d=1$.

Обнаруживающая способность кода характеризуется следующей теоремой. Если код имеет $d>1$ и используется декодирование по методу обнаружения ошибок, то все ошибки кратностью $q<d$ обнаруживаются. Что же касается ошибок кратностью $q \geq d$, то одни из них обнаруживаются, а другие нет.

Исправляющая способность кода при этом правиле декодирования определяется следующей теоремой. Если код имеет $d>2$ и используется декодирование с исправлением ошибок по наименьшему расстоянию, то все ошибки кратностью $q<d/2$ исправляются. Что же касается ошибок большей кратности, то одни из них исправляются, а другие нет.

Задача кодирования состоит в выборе кода, обладающего максимально достижимым d . Впрочем, такая формулировка задачи неполна. Увеличивая длину кода n и сохраняя число кодовых комбинаций M , можно получить сколь угодно большое значение d . Но такое "решение" задачи не представляет интереса, так как с увеличением n уменьшается возможная скорость передачи информации от источника.

Если длина кода n задана, то можно получить любое значение d , не превышающее n , уменьшая число комбинаций M . Поэтому задачу поиска наилучшего кода (в смысле максимального d) следует формулировать так: при заданных M и n найти код длины n , содержащий M комбинаций и имеющий наибольшее возможное d . В общем виде эта задача в теории кодирования не решена, хотя для многих значений n и M ее решения получены.

На первый взгляд помехоустойчивое кодирование реализуется весьма просто. В память кодирующего устройства (кодера) записываются разрешенные кодовые комбинации выбранного кода и правило, по которому с каждым из M сообщений источника сопоставляется одна из таких комбинаций. Данное правило известно и декодеру.

Получив от источника определенное сообщение, кодер отыскивает соответствующую ему комбинацию и посылает ее в канал. В свою очередь, декодер, приняв комбинацию, искаженную помехами, сравнивает ее со всеми M комбинациями списка и отыскивает ту из них, которая ближе остальных к принятой.

Однако даже при умеренных значениях n такой способ весьма сложный. Покажем это на примере. Пусть выбрана длина кодовой комбинации $n=100$, а скорость кода примем равной 0.5 (число информационных и проверочных символов равно). Тогда число разрешенных комбинаций кода будет $2^{50} \gg 10^{15}$. Соответственно размер таблицы будет $100 \cdot 10^{15} = 10^{17}$ бит $\gg 10^{16}$ байт = 10000 Тбайт.

Таким образом, применение достаточно эффективных (а значит, и достаточно длинных) кодов при табличном методе кодирования и декодирования технически невозможно.

Поэтому основное направление теории помехоустойчивого кодирования заключается в поисках таких классов кодов, для которых кодирование и декодирование осуществляются не перебором таблицы, а с помощью некоторых регулярных правил, определенных алгебраической структурой кодовых комбинаций.

3.2. Линейные коды

Одним из таких классов являются линейные блочные коды. Линейными называются такие двоичные коды, в которых множество всех разрешенных блоков является линейным пространством относительно операции поразрядного сложения по модулю 2.

Если записать k линейно-независимых блоков в виде k строк, то получится матрица размером $n \times k$, которую называют порождающей или производящей матрицей кода G .

Множество линейных комбинаций образует линейное пространство, содержащее 2^k блоков, т.е. линейный код, содержащий 2^k блоков длиной n , обозначают (n, k) . При заданных n и k существует много различных (n, k) -кодов с различными кодовыми расстояниями d , определяемых различными порождающими матрицами. Все они имеют избыточность

$e_k = 1 - k/n$ или относительную скорость $R_k = k/n$.

Чаще всего применяют систематические линейные коды, которые строят следующим образом. Сначала строится простой код длиной k , т.е. множество всех k -последовательностей двоичных символов, называемых информационными. Затем к каждой из этих последовательностей приписывается $r = n - k$ проверочных символов, которые получаются в результате некоторых линейных операций над информационными символами.

Простейший систематический код $(n, n-1)$ строится добавлением к комбинации из $n-1$ информационных символов одного проверочного, равного сумме всех информационных символов по модулю 2. Такой код $(n, n-1)$ имеет $d=2$ и позволяет обнаружить одиночные ошибки и называется кодом с одной проверкой на четность.

Преимуществом линейных, в частности систематических, кодов является то, что в кодере и декодере не нужно хранить большие таблицы всех кодовых комбинаций, а при декодировании не нужно производить большое количество сравнений.

Однако, для получения высокой верности связи следует применять коды достаточно большой длины. Применение систематического кода в общем случае, хотя и позволяет упростить декодирование по сравнению с табличным способом, все же при значениях n порядка нескольких десятков не решает задачу практической реализации.

3.3. Совершенные и квазисовершенные коды

Совершенными (плотно упакованными) называют коды, в которых выполняются соотношения $\sum_{i=0}^q C_n^i (b-1)^i = b^{r-1}$, где q - максимальная кратность исправляемых ошибок; b - основание кода; r - число проверочных символов.

Они отличаются тем, что позволяют исправлять все ошибки кратностью q или меньше и ни одной ошибки кратности больше q .

Число известных совершенных кодов ограничено кодами Хэмминга значности $n = \frac{b^r - 1}{b - 1}$ и бинарным циклическим кодом Голея.

Квазисовершенными принято называть коды, исправляющие все ошибки кратности q и $A \leq (b-1)^i C_n^i$ ошибок кратности $q + 1$ при условии, что $\sum_{i=0}^q (b-1)^i C_n^i + A = b^r + 1$.

Класс квазисовершенных кодов значительно шире, чем класс плотно упакованных кодов.

Совершенные и квазисовершенные коды обеспечивают максимум вероятности правильного приема комбинации при равновероятных ошибках в канале связи.

Был предложен ряд кодов и способов декодирования, при которых сложность декодера растет не экспоненциально, а лишь как некоторая степень n . В классе линейных систематических двоичных кодов это - циклические коды. Циклические коды просты в реализации и при невысокой избыточности обладают хорошими свойствами обнаружения ошибок. Циклические коды получили очень широкое распространение, как в технике связи, так и в компьютерных средствах хранения информации. В зарубежных источниках циклические коды обычно называют избыточной циклической проверкой (CRC, Cyclic Redundancy Check).

При использовании в циклических кодах декодирования с исправлением ошибок остаток от деления может играть роль синдрома. Нулевой синдром указывает на то, что принятая комбинация является разрешенной. Всякому ненулевому синдрому соответствует определенная конфигурация ошибок, которая и исправляется.

Однако обычно в системах связи исправление ошибок при использовании циклических кодов не производится, а при обнаружении ошибок выдается запрос на повтор испорченной ошибками комбинации. Такие системы называются системами с обратной связью.

Среди циклических кодов особое значение имеет класс кодов, предложенных Боузом и Рой-Чоудхури и независимо от них Хоквингемом. Коды Боуза - Чоудхури - Хоквингема (обозначаемые сокращением БЧХ) отличаются сравнительно просто реализуемой процедурой декодирования.

Относительно простой является процедура мажоритарного декодирования, применяемая для некоторого класса двоичных линейных, в том числе циклических кодов. Основана она на том, что в этих кодах каждый информационный символ можно несколькими способами выразить через другие символы кодовой комбинации.

Мощные коды (т.е. коды с длинными блоками и большим кодовым расстоянием d) при сравнительно простой процедуре декодирования можно строить, объединяя несколько коротких кодов. Так строится, например, итеративный код из двух линейных систематических кодов (n_1, k_1) и (n_2, k_2) . Минимальное кодовое расстояние для двумерного итеративного кода $d = d_1 d_2$, где d_1 и d_2 - соответственно минимальные кодовые расстояния для кодов 1-й и 2-й ступеней.

На итеративный код похож каскадный код, но между ними имеется существенное различие. Первая ступень кодирования в каскадном коде является линейным систематическим двоичным кодом (внутренний код), каждая комбинация которого рассматривается как один символ недвоичного кода второй ступени (внешнего). При приеме сначала декодируются (с обнаружением или исправлением ошибок) все строки (блоки) внутреннего кода, а затем декодируется блок внешнего m -ичного кода, причем исправляются ошибки и стирания, оставшиеся после декодирования внутреннего кода. В качестве внешнего кода используют обычно m -ичный код Рида-Соломона, который является подклассом кодов БЧХ и обеспечивает наибольшее возможное d при заданных n_2 и k_2 , если $n_2 < m$. Каскадные коды во многих случаях наиболее перспективны среди известных блочных помехоустойчивых кодов.

3.4. Метод перемежения

Для каналов с группированием ошибок часто применяют метод перемежения символов, или декорреляции ошибок. Он заключается в том, что символы, входящие в одну кодовую комбинацию, передаются не непосредственно друг за другом, а перемежаются символами других кодовых комбинаций.

Если интервал между символами, входящими в одну комбинацию, сделать больше максимально возможной длины группы ошибок, то в пределах комбинации группирования ошибок не будет. Группа ошибок распределится в виде одиночных ошибок на группу комбинаций. Одиночные ошибки будут легко обнаружены (исправлены) декодером.

3.5. Системы с обратной связью

Нередко встречаются случаи, когда информация может передаваться не только от одного абонента к другому, но и в обратном направлении. В таких условиях появляется возможность использовать обратный поток информации для существенного повышения верности сообщений, переданных в прямом направлении. При этом не исключено, что по обоим каналам (прямому и обратному) в основном непосредственно передаются сообщения в двух направлениях ("дуплексная связь") и только часть пропускной способности каждого из каналов используют для передачи дополнительных данных, предназначенных для повышения верности.

Возможны различные способы использования системы с обратной связью в дискретном канале. Обычно их подразделяют на два типа: системы с информационной обратной связью и системы с управляющей обратной связью.

Системами с информационной обратной связью (ИОС) называются такие, в которых с приемного устройства на передающее поступает информация о том, в каком виде принято сообщение. На основании этой информации передающее устройство может вносить те или иные изменения в процесс передачи сообщения:

- повторить ошибочно принятые отрезки сообщения;
- изменить применяемый код (передав предварительно соответствующий условный сигнал и убедившись в том, что он принят);
- прекратить передачу при плохом состоянии канала до его улучшения.

В системах с управляющей обратной связью (УОС) приемное устройство на основании анализа принятого сигнала само принимает решение о необходимости повторения, изменения способа передачи, временного перерыва связи и передает об этом указание передающему устройству. Возможны и смешанные методы использования обратной связи, когда в некоторых случаях решение принимается на приемном устройстве, а в других случаях на передающем устройстве на основании полученной по обратному каналу информации.

Наиболее распространены системы с УОС при использовании одновременно с обнаружением ошибок. Такие системы часто называют системами с переспросом, или с автоматическим запросом ошибок (ARQ, Automatic Repeat reQuest).

3.6. Мягкое и жесткое декодирование

Рассмотрим ситуацию, когда двоичный сигнал передается за отрезок времени $(0, T)$, причем двоичная единица представляется сигналом $s_1(t)$, а двоичный нуль — сигналом $s_2(t)$. Принятый сигнал имеет вид $y(t) = s_i(t) + n(t)$, где $n(t)$ представляет собой вклад гауссового шума с нулевым средним. В главе 1 мы описывали детектирование $y(t)$ в два основных этапа. На первом этапе принятый сигнал переводится в число $z(T) = a_i + n_0$, где a_i — это компонент сигнала $z(T)$, а n_0 — компонент шума. Компонент шума n_0 — это *случайная переменная*, значения которой имеют *гауссово* распределение с нулевым средним. Следовательно, $z(T)$ также будет *случайной гауссовой величиной* со средним a_1 или a_2 , в зависимости от того, какая величина была отправлена — двоичная единица или двоичный нуль. На втором этапе процесса детектирования принимается решение о том, какой сигнал был передан. Это решение принимается на основе сравнения $z(T)$ с порогом. Условные вероятности $z(T)$, $p(z|s_1)$ и $p(z|s_2)$, показанные на рис. 3.1, обозначены как правдоподобие s_1 и s_2 . Демодулятор преобразует упорядоченный по времени набор случайных переменных $\{z(T)\}$ в кодовую последовательность Z и подает ее на декодер. Выход демодулятора можно настроить по-разному. Можно реализовать его в виде *жесткой схемы принятия решений* относительно того, представляет ли $z(T)$ единицу или нуль. В этом случае выход демодулятора квантуется на два уровня, нулевой и единичный, и соединяется с декодером. Поскольку декодер работает в режиме жесткой схемы принятия решений, принятых демодулятором, такое декодирование называется *жестким*.

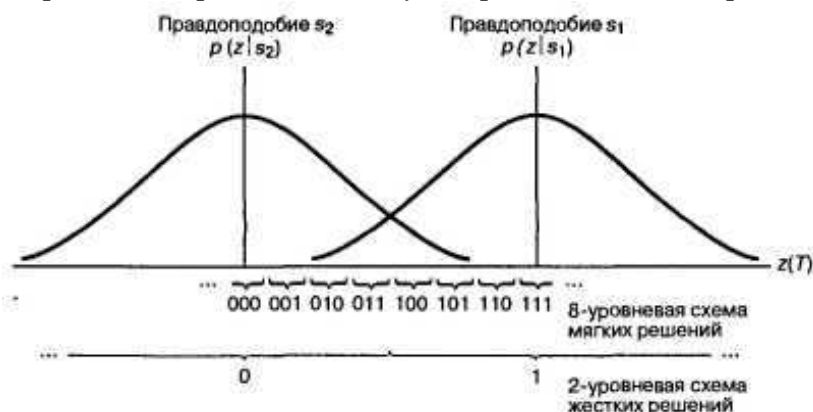


Рис. 3.1. Жесткая и мягкая схемы декодирования.

Аналогично демодулятор можно настроить так, чтобы он подавал на декодер значение $z(T)$, квантованное более чем на два уровня. Такая схема обеспечивает декодер большим количеством информации, чем жесткая схема решений. Если выход демодулятора имеет более двух уровней квантования, то декодирование называется *мягким*. На рис. 3.1 на оси абсцисс изображено восемь (3-битовых) уровней квантования. Если в демодуляторе реализована жесткая схема принятия двоичных решений, он отправляет на декодер только один двоичный символ. Если в демодуляторе реализована мягкая двоичная схема принятия решений, квантованная на восемь уровней, он отправляет на декодер 3-битовое слово, описывающее интервал, соответствующий $z(T)$. По сути, поступление такого 3-битового слова, вместо одного двоичного символа, эквивалентно передаче декодеру *меры достоверности* вместе с решением относительно кодового символа. Согласно рис. 3.1, если с демодулятора поступила на декодер последовательность 111, это равносильно утверждению, что с очень высокой степенью достоверности кодовым символом была 1, в то время как переданная последовательность 100 равносильна утверждению, что с очень низкой степенью достоверности кодовым символом была 1. Совершенно ясно, что, в конечном счете, каждое решение, принятое

декодером и касающиеся сообщения, должно быть жестким; в противном случае на распечатках компьютера можно было бы увидеть нечто, подобное следующему: "думаю, это 1", "думаю, это 0" и т.д. То, что после демодулятора *не принимается жесткое решение* и на декодер поступает больше данных (мягкое принятие решений), можно понимать как промежуточный этап, необходимый для того, чтобы на декодер поступило больше информации, с помощью которой он затем сможет восстановить последовательность сообщения (с более высокой достоверностью передачи сообщения по сравнению с декодированием в рамках жесткой схемы принятия решений).

3.7. Сверточные коды

Наряду с блоковыми кодами существует большой и эффективный класс древовидных или решетчатых кодов, среди которых особый интерес представляют *сверточные коды*. Отличительной чертой последних (в сравнении с блоковыми) служит способ отображения потока информационных символов (бит) в кодовые. При блоковом кодировании последовательность не пересекающихся k -символьных информационных блоков заменяется последовательностью не пересекающихся n -символьных кодовых слов, причем каждое кодовое слово «защищает» только свой k -символьный информационный блок и занимает в реальном времени интервал, отводимый для передачи именно этих k символов. Иллюстрацией формирования (7,4) блокового кода служит рис. 3.2.

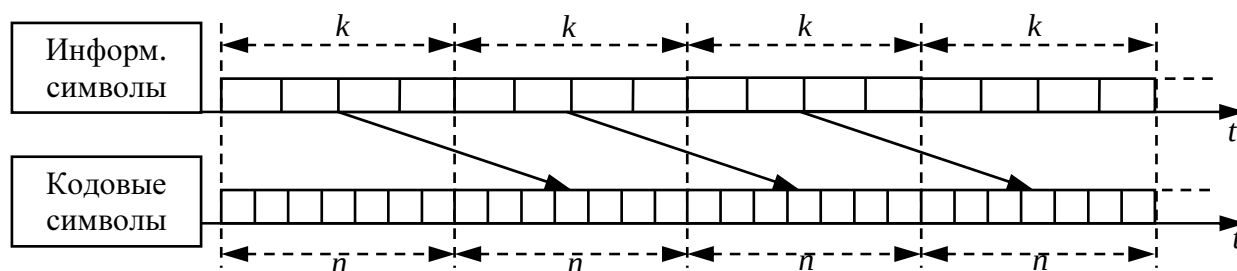


Рис. 3.2. Схема блокового кодирования

. Различие между сверточными и блоковыми кодами до некоторой степени размыто: любой сверточный код может трактоваться как блоковый код соответственно большой длины. Более разумно рассмотреть особенности и причины чрезвычайной популярности сверточных кодов в их рекуррентной форме, позволяющей реализовать значительно более

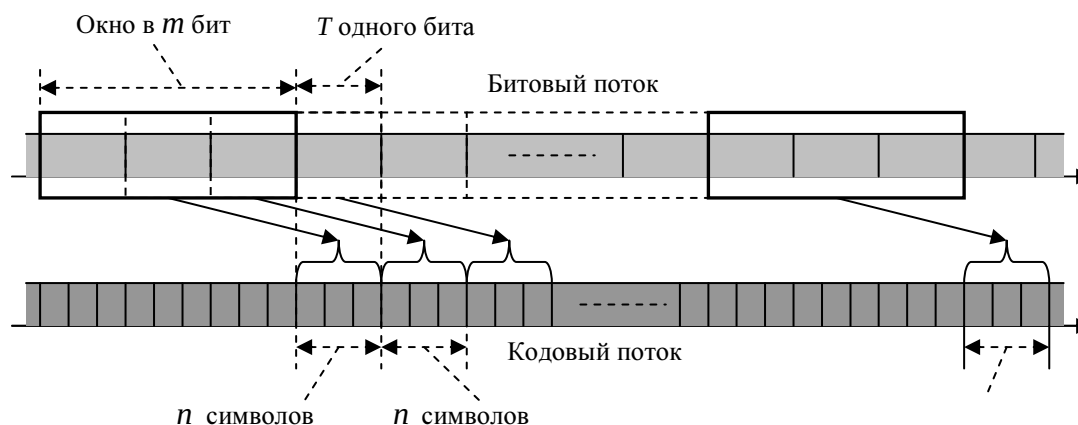


Рис. 3.3. Иллюстрация сверточного кодирования.

практичные по сравнению с другими процедуры декодирования с исправлением ошибок (алгоритм Витерби).

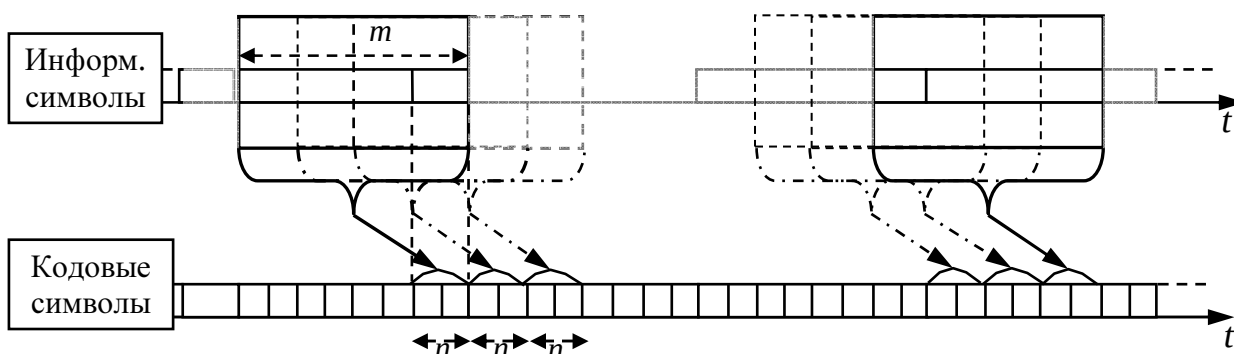


Рис. 3.3а. Формирование сверточного кода с $m = 4$ и $n = 2$

Сверточный кодер

Идея сверточного кодирования в общих словах может быть описана следующим образом. Возьмем блок (вектор) из m последовательных бит источника и линейно преобразуем его в $n > 1$ выходных двоичных кодовых символов, занимающих временной интервал, равный длительности одного бита источника. Линейность по отношению к векторам означает просто суммирование по модулю два выбранных компонентов. После этого обновим блок бит источника, включив в него один новый и исключив один наиболее старый. Снова имеем блок из m бит источника, задержанный относительно исходного блока на один бит (и содержащий $m - 1$ предшествующих бит и один новый), который кодируется n новыми кодовыми символами. Данная процедура непрерывно повторяется раз за разом, в каждый момент времени вовлекая один новый бит и отбрасывая старейший. Рис. 3.3 иллюстрирует указанную процедуру для параметров: $m = 3$, $n = 3$; кодер как будто просматривает битовый поток источника через скользящее окно шириной k и кодирует все биты, которые в текущий момент видит, в n кодовых символов (см. также рис. 3.3а).

После каждого шага процедуры окно сдвигается на длительность одного бита источника и выполняется следующий шаг. Число бит источника, определяющих кодовые символы на одном шаге, называется *длиной кодового ограничения*. Описанный выше принцип может быть реализован структурой, представленной на рис. 3.4, где регистр сдвига, состоящий из $m - 1$ ячейки памяти, хранит $m - 1$ предшествующих бит источника. Вместе с поступающим битом они подаются в линейную логическую цепь, содержащую n сумматоров по модулю два. Выход каждой ячейки памяти и вход кодера может быть, а может и не быть (что объясняет обозначение связей пунктирной линией) соединены с каждым сумматором, причем схема соединений определяет конкретную зависимость выходных кодовых символов от m бит источника, т.е. правило кодирования.

При поступлении на вход текущего бита источника b_i на выходах сумматоров параллельно появляются n кодовых символов $u_i^1, u_i^2, \dots, u_i^n$. После каждого такта битовый образец в регистре смещается вправо на одну ячейку, подготавливая цепь к формированию следующих n кодовых символов. Выходной ключ в течение одного бита поочередно подключается к выходам n сумматоров и преобразует параллельную форму представления кодовых символов в последовательную, создавая выходной кодовый поток $u_0^1, u_0^2, \dots, u_0^n, u_1^1, u_1^2, \dots, u_1^n, \dots$. Видно, что в устойчивом состоянии кодер на рис.3.4 на каждый новый бит источника отвечает формированием n кодовых символов (см. также рис.3.3), так что скорость кода R_c , измеренная в битах на один кодовый символ, равна $1/n$.

Если на вход будет подана информационная последовательность из k бит, то на выходе кодера появится кодовая последовательность не из kn , а $(k + m - 1)n$ символов, поскольку регистр сдвига обнулится только после $(m - 1)$ дополнительного такта. Вследствие этого реальная *скорость* сверточного кода составляет величину $R = k / (k + m - 1)n$, однако учитывая, что $k \gg 1$, а $m \ll k$, то $R = 1/n$.

Нижеприведенный пример поможет лучше понять принцип сверточного кодирования.

Пример 3.1.1 Рис. 3.5 иллюстрирует реализацию сверточного кодера с длиной

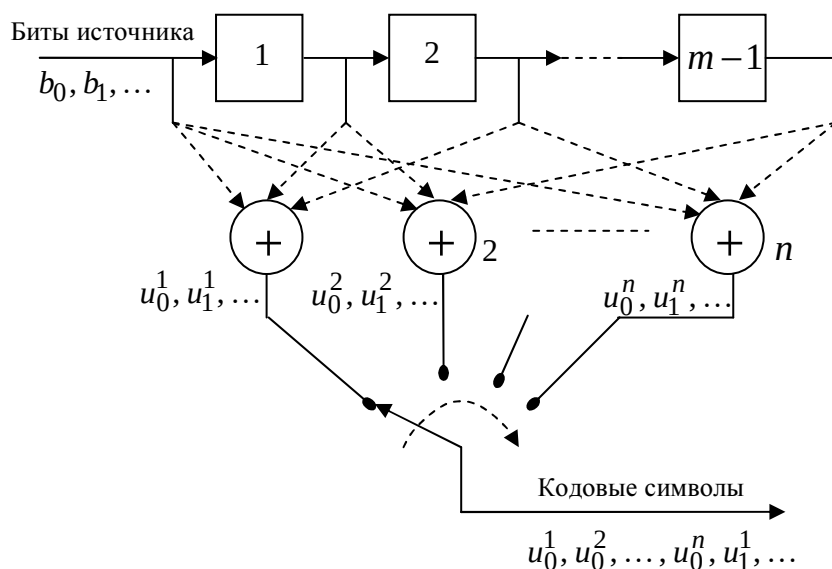


Рис. 3.4. Сверточный кодер.

кодировочного ограничения $m=3$ и скоростью $R_c=1/2$. Входные биты $b_0, b_1, \dots$ порождают два потока кодовых символов $u_0^1, u_1^1, \dots$ и $u_0^2, u_1^2, \dots$, которые затем мультиплексируются таким образом, что u_i^1 и u_i^2 занимают соответственно четные и нечетные позиции в общем кодовом потоке. Например, битовый поток $\{b_i\}=10100100\dots$ порождает последовательности $\{u_i^1\}=1101111\dots$ и $\{u_i^2\}=10001101\dots$, которые затем мультиплексируются в кодовый поток $\{(u_i^1, u_i^2)\}=1110001011111011\dots$.

Естественно возникает следующий вопрос: возможна ли при сверточном кодировании скорость R_c , отличная от $1/n$, т.е. равная k/n , где $1 < k < n$? Существуют два классических способа решения этой задачи. Первый состоит в обобщении вышеприведенного принципа: на каждом шаге $m-1$, а не m , бит линейно преобразуется в n кодовых символов, после чего k наиболее старых бит (вместо одного) заменяются k новыми и кодер переходит к следующему шагу. Второй способ, называемый *выкалыванием* (*puncturing*), использует коды со скоростью $1/n$ в качестве подсобного материала и удаляет некоторые кодовые символы в соответствии с заранее назначенным образцом. При соответствующей организации процедуры выкалывания происходит уменьшение числа кодовых символов на бит данных, что обеспечивает достижение скорости k/n . Исходя из реализационных возможностей, процедура выкалывания часто считается предпочтительной, подчеркивая еще раз доминирующий интерес в отношении кодов со скоростями вида $1/n$. Следуя этой тенденции, сконцентрируем наше внимание

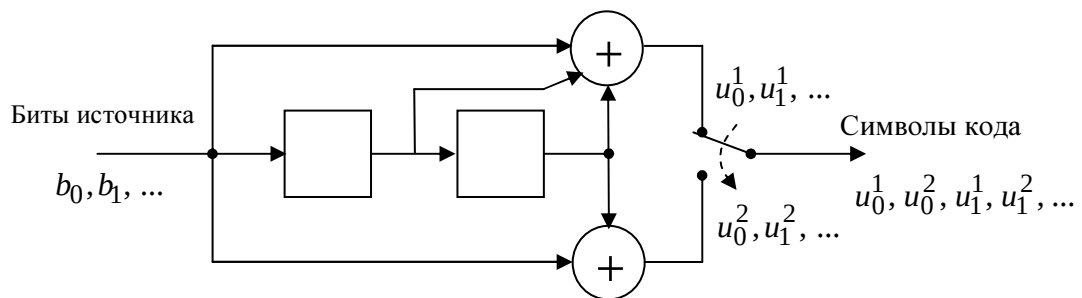


Рис. 3.5. Сверточный кодер со скоростью $1/2$.

только на сверточных кодах со скоростью $1/n$.

Очевидно, что схема, содержащая регистр сдвига и отдельные сумматоры по модулю два вместе со всеми их соединениями, представляет собой ничто иное, как *фильтр с конечным импульсным откликом* (КИО) (*finite impulse response (FIR) filter*), сигнал на выходе которого представляет собой свертку входного потока битов и импульсной характеристикой фильтра, что объясняет название рассматриваемых кодов. Данный факт лежит также в основе одного из способов формального описания сверточного кодера. Свертка, связывающая кодовый символ u_i^l (т.е. появляющийся на выходе l -го сумматора при поступлении b_i бита источника) с входным битовым потоком, определяется как

$$u_i^l = \sum_{t=0}^{m-1} b_{i-t} g_t^l, \quad i=0,1,\dots; \quad l=1,2,\dots,n, \quad (3.1)$$

где $g_t^l=1$, если l -й сумматор соединен с t -й ячейкой памяти ($t=0$ соответствует входу кодера) и $g_t^l=0$ в противном случае; $b_i=0$ при $i < 0$. В частотной области подходящим

инструментом для описания дискретных систем является z -преобразование. В z -области свертке соответствует произведение z -преобразований, так что соотношение (3.1) может быть представлено эквивалентной формой

$$u^l(z) = \sum_{i=0}^{\infty} u_i^l z^i = b(z)g_l(z), \quad l=1,2,\dots,n, \quad (3.2)$$

где $b(z) = \sum_{i=0}^{\infty} b_i z^i$ – z -преобразование входного потока бит, а

$$g_l(z) = g_0^l + g_1^l z + \dots + g_{V_c-1}^l z^{m-1}, \quad l=1,2,\dots,n \quad (3.3)$$

– передаточная функция l -го КИО-фильтра (т.е. формирующая l -й кодовый символ), называемая также l -м *порождающим полиномом* сверточного кода. Множество из n порождающих полиномов полностью определяет сверточный код, поскольку их ненулевые коэффициенты конкретизируют соединения сумматоров с регистром сдвига.

Пример 3.1.2. Кодер, изображенный на рис. 3.5, характеризуется порождающими полиномами $g_1(z) = 1 + z + z^2$ и $g_2(z) = 1 + z^2$.

Полагая $g_1(z) = 1$, получаем *систематический* сверточный код, в котором биты источника данных занимают вполне определенные позиции. Однако проблема заключается в том, что в рамках фиксированной структуры, изображенной на рис. 3.4, систематические коды, как правило, не оказываются лучшими с точки зрения исправляющей способности. Последнее обосновывает модификацию регистра сдвига сверточного кодера в структуру с обратной связью, если свойство систематичности играет важную роль.

Решетчатая диаграмма, свободное расстояние и асимптотический выигрыш от кодирования

Регистр сдвига сверточного кодера обладает 2^{m-1} возможными состояниями и имеет только два состояния, в которые он может перейти из текущего состояния после очередного такта. Именно значение входного бита источника b_i определяет выбор одного из двух возможных переходов. Если на i -м тактовом интервале состояние регистра определяется, как $(b_{i-1}, b_{i-2}, \dots, b_{i-m+1})$, то следующее состояние будет либо $(0, b_{i-1}, \dots, b_{i-m+2})$ при поступлении на вход бита источника $b_i = 0$, либо $(1, b_{i-1}, \dots, b_{i-m+2})$, если $b_i = 1$. Аналогично, регистр приходит в состояние $(b_i, b_{i-1}, \dots, b_{i-m+1})$, если его предыдущее состояние было либо $(b_{i-1}, b_{i-2}, \dots, b_{i-m+2}, 0)$, либо $(b_{i-1}, b_{i-2}, \dots, b_{i-m+2}, 1)$. Для схематического изображения всех деталей поведения регистра адекватным инструментом служит *решетка (trellis)*. Она включает две колонки из 2^{m-1} узлов, причем левая колонка отвечает текущему состоянию регистра, а правая – следующему. Ребра (стрелки) выходят из каждого узла левой колонки и входят в некоторые два узла следующей, причем сплошные и пунктирные ребра показывают пути, соответствующие приходу бита нуля или единицы соответственно. Аналогично, в каждый узел правой колонки входят два ребра, отмеченные или сплошной (нулевой входной бит), или пунктирной (входной бит равен единице) линиями. Каждое ребро помечено n -разрядным блоком, который представляет собой группу из n кодовых символов, выдаваемых кодером в момент, когда входной бит источника переводит его из одного состояния в другое.

Нижеприведенный пример поясняет способ построения решетки для случая сверточного кодера на рис. 3.5.

Пример 3.1.3. Двухразрядный регистр сдвига может находиться в одном из четырех возможных состояниях: (00), (10), (11) и (01), причем запись состояния ведется с левой

ячейки памяти. На рис. 3.6 приведена решетка, построенная согласно выше приведенному описанию. Например, ребро из состояния (10) в (01) отмечено сплошной линией, а в состояние (11) – пунктирной. Ребро, идущее из состояния (11) в (01), помечено кодовыми символами 01, поскольку при единицах в обеих ячейках и нулевым битом на входе верхний сумматор на рис. 3.5 вырабатывает нулевой выходной сигнал, а нижний выдает единицу и т.п. В течение интервала каждого такта состояние кодера изменяется согласно движению вдоль какого-то ребра решетки, одновременно выдавая на выход кодовые символы, которыми помечено данное ребро. Отслеживая этот процесс по диаграмме, подобно представленной на рис. 3.6, необходимо производить скачки из правой колонки в тот же самый узел левой колонки на каждом последующем шаге. Для избежания этого повторим решетку столько раз, сколько необходимо, используя правую колонку текущего

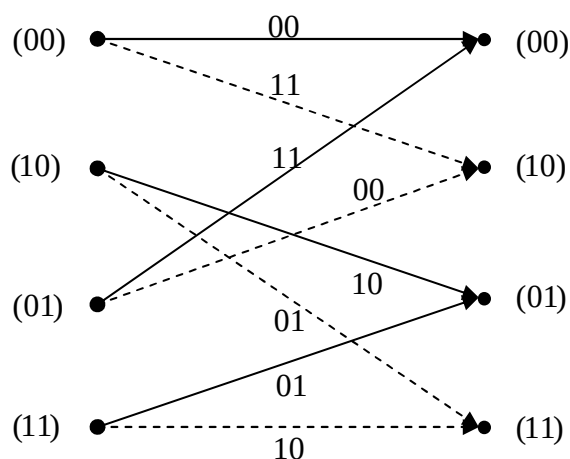


Рис. 3.6. Решетка кодера на рис. 3.4.

шага в качестве левой для последующего шага.

Тогда кодирование может быть отождествлено с движением вдоль полученной *решетчатой диаграммы*, регулируемым текущим входным битом, который направляет кодер по сплошному или пунктирному ребру в зависимости от его конкретного значения: нуля или единицы. Для иллюстрации сказанного обратимся к рис. 3.6, на котором представлена решетчатая диаграмма кода из примера 3.1.1. Каждая последовательность–кандидат входных битов данных выбирает определенный путь по решетчатой диаграмме, который можно отследить, например, для последовательности $\{b_i\}=10100100\dots$. Ее первый бит равен единице, что направляет кодер из узла (00) в узел (10), выдавая выходные кодовые символы 11. Второй бит, равный нулю, переводит кодер из узла (10) в (01), генерируя кодовые символы 10. Третий бит изменяет состояние кодера с (01) на (10), выдавая на выход кодовые символы 00 и т.д. Жирной чертой выделен результирующий путь, отвечающий кодовому слову 1110001011111011.

Ранее уже было отмечено, что сверточный код можно трактовать как блочный код соответственно большой длины. Кодовые слова этого блочного кода есть просто различные пути по решетчатой диаграмме, минимальное хэммингово расстояние среди всех их пар дающее кодовое расстояние. В свою очередь, линейность кода упрощает задачу нахождения кодового расстояния: минимум расстояния между путями определяется наименьшим весом Хэмминга среди всех ненулевых слов. Предположим теперь, что закодированный битовый поток заканчивается после некоторого большого (не меньше чем m) числа бит и дополняется $m-1$ хвостовыми нулями для установки регистра в нулевое состояние. Реализованное на практике (одним из примеров служит cdmaOne) это дополнение не приведет к значительным затратам, если длина кодированного потока бит во много раз больше длины кодового ограничения. С другой стороны, введенное дополнение приведет к слиянию всех путей в один, который соответствует нулевому состоянию регистра, как показывает рис. 3.6 для кода из примера 3.1.1. Если теперь взять дополненный битовый поток, начинающийся с некоторого числа n_0 нулей, за которыми следует бит, равный «1», и переместить эти n_0 начальных нулей в конец потока, то это будет означать простой сдвиг соответствующего пути на n_0 шагов влево на диаграмме без изменения его веса.

После такого сдвига путь отклонится от нулевого пути на самом первом шаге и сольется с ним не позднее, чем за n_0 шагов до последнего хвостового нуля. Вдоль этого пути могут произойти неоднократные возвращения к нулевому пути и последующие

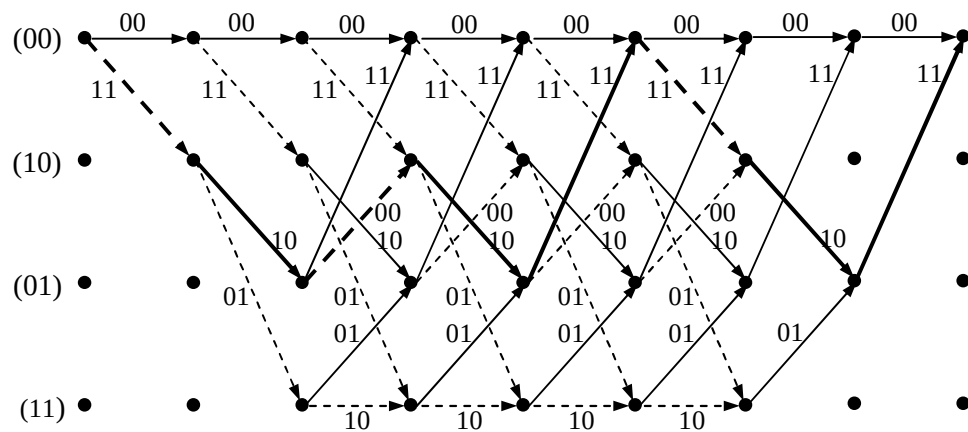


Рис. 3.7 Решетчатая диаграмма кодера на рис. 3.5.

отклонения от него (см. рис. 3.8), каждое из которых только увеличивает вес пути, и поскольку целью является отыскание пути минимального веса, то любые отклонения от нулевого пути, кроме первого, должно игнорироваться. Суммируя все вышесказанное, заключаем, что для отыскания расстояния сверточного кода следует исследовать только пути, отклоняющиеся от нулевого пути в начале решетчатой диаграммы и не отходящие от него после первого объединения. В теории сверточных кодов эта величина традиционно называется *свободным расстоянием*.

Обозначив ее как d_f , можно увидеть, например, что среди всех путей с единственным отклонением от нулевого пути, изображенных на рис. 3.7, кодовое слово 11101100..., кодирующее битовый поток 100..., имеет минимальный вес, так что $d_f = 5$.

Очевидно, что свободное расстояние d_f гарантирует исправление любых $\left\lfloor \frac{d_f - 1}{2} \right\rfloor$ ошибок в символах, однако, как правило, исправляются также и многие образцы с большим

числом ошибок. Существует только несколько примеров эффективных алгебраических правил сверточного кодирования. Большинство же известных сверточных кодов с хорошей корректирующей способностью были найдены в результате компьютерного поиска.

Вследствие специфичности алгоритмов кодирования аналитическое нахождение всех возможных весов слов (весового спектра) произвольного сверточного кода оказывается не столь трудно, как в случае многих линейных блоковых кодов. В частности, непосредственно из решетки (или, что эквивалентно, диаграммы состояний) строится система линейных уравнений, решение которой приводит к точному выражению для весового спектра.

Выигрыш от кодирования, показывающий во сколько раз может быть уменьшена энергия сигнала на бит или мощность сигнала в результате кодирования при фиксированной вероятности ошибки, является универсальной мерой для оценки эффективности того или иного кода. Понятно, что асимптотический выигрыш от кодирования в случае АБГШ канала определяется выигрышем в евклидовом расстоянии. При передаче с использованием бинарной ФМ любое различие в символах двух сигналов увеличивает квадрат евклидова расстояния на величину, равную $4E_s$, где E_s – энергия одного символа. Имеется пара слов сверточного кода, имеющая d_f различных символов, и отсутствует пара с меньшим различием (расстоянием Хэмминга). Следовательно,

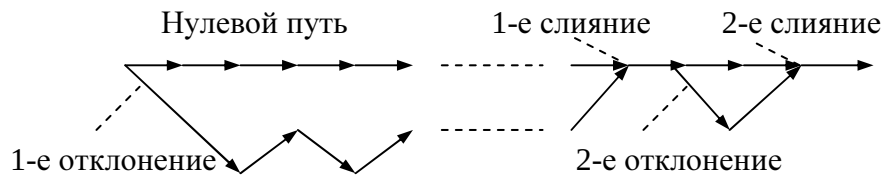


Рис. 3.8. Иллюстрация отклонения и слияния путей.

минимум квадрата евклидова расстояния между словами сверточного кода, передаваемыми с помощью бинарной ФМ, составляет $d_{\min, cc}^2 = 4d_f E_s$. В то же время аналогичная характеристика при некодированной передаче составляет $d_{\min, u}^2 = 4E_b$, определяя асимптотический выигрыш от кодирования сверточным кодом, как

$$G_a = \frac{d_{\min, cc}^2}{d_{\min, u}^2} = \frac{d_f E_s}{E_b} = d_f R_c,$$

где R_c – скорость кода, измеренная в бит/сим. Для кода из примера 3.1.1 $R_c = 1/2$, $d_f = 5$, так что $G_a = 2.5$ или около 4 дБ. Напомним еще раз, что G_a выводится для АБГШ (не двоичного симметричного канала – ДСК) канала, другими словами, для случая применения мягкого декодирования. Жесткое декодирование ухудшает эту цифру на 2..3 дБ в зависимости от параметров кода и отношения сигнал-шум на символ.

3.8. Алгоритм декодирования Витерби

Как уже упоминалось, среди причин широкой популярности сверточных кодов особую роль играет существование эффективного алгоритма декодирования. Рассмотрение его начнем со следующего положения.

Утверждение 1. Максимально правдоподобное жесткое декодирование с исправление ошибок бинарного кода эквивалентно правилу минимума расстояния Хэмминга:

$$d_H(\hat{\mathbf{u}}, \mathbf{y}) = \min_{\mathbf{u} \in U} d_H(\mathbf{u}, \mathbf{y}) \Rightarrow \hat{\mathbf{u}} \text{ принимается за принятое слово.} \quad (3.4)$$

Как можно видеть, данное правило очень похоже на правило минимума расстояния с одним изменением: в случае ДСК расстояние Хэмминга заменяет Евклидово расстояния, адекватное для АБГШ канала. Для доказательства (3.4) достаточно заметить, что для независимых символьных ошибок, если ошибки в символах произойдут на всех $d_H(\mathbf{u}, \mathbf{v})$ позициях, где $\mathbf{u}$ и $\mathbf{v}$ отличаются, и не будут повреждены на $n - d_H(\mathbf{u}, \mathbf{v})$ позициях, где символы отличаются:

$$P(\mathbf{y}|\mathbf{u}) = p^{d_H(\mathbf{u}, \mathbf{y})} (1-p)^{n-d_H(\mathbf{u}, \mathbf{y})} = \left(\frac{p}{1-p} \right)^{d_H(\mathbf{u}, \mathbf{y})} (1-p)^n$$

есть просто переходная вероятность ДСК, т.е. вероятность трансформации ДСК посланного кодового вектора $\mathbf{u}$ длины n в двоичное наблюдение $\mathbf{y}$. Поскольку перекрестная вероятность ДСК $p < 0.5$, то переходная вероятность $P(\mathbf{y}|\mathbf{u})$ есть убывающая функция расстояния Хэмминга между наблюдением $\mathbf{y}$ на выходе ДСК и кодовым вектором $\mathbf{u}$. Следовательно, максимально правдоподобной оценкой $\hat{\mathbf{u}}$ кодового слова является слово, ближайшее по Хэммингу к $\mathbf{y}$.

Прямая реализация правила (3.4) для произвольного кода заключается в сравнении M расстояний Хэмминга между наблюдением $\mathbf{y}$ и всеми кодовыми словами. Поскольку, как правило, M достаточно велико, то такое решение может оказаться не осуществимым. С другой стороны, для сверточных кодов вследствие их благоприятной структуры МП декодирование не представляет особой трудности, по крайней мере, для умеренных значений длины кодового ограничения.

Процедура декодирования Витерби реализует максимально правдоподобную стратегию в рекуррентной форме, последовательно осуществляя поиск пути по решетчатой диаграмме, ближайший к двоичному наблюдению $\mathbf{y}$. Каждый новый шаг декодирования начинается с приема следующей группы из n символов наблюдения. На i -м шаге декодер вначале вычисляет расстояние между n пришедшими символами наблюдения и каждым ребром решетчатой диаграммы, а затем увеличивает расстояния всех путей, рассчитанных за $i-1$ предшествующих шагов. Для произвольного кода расстояния можно вычислять аналогично сразу же по приходу новых символов наблюдения, однако рекуррентная природа сверточных кодов позволяет осуществить эту обычную вычислительную работу более экономно по причине отбрасывания многих путей на каждом шаге.

Рассмотрим все пути, проходящие через фиксированный узел A на i -м шаге, как это показано на рис. 3.8. Продолжение любого пути после i -го шага не зависит от маршрута прихода в A , так что все различные пути, достигающие A , могут в дальнейшем считаться слившимися в один. Однако это означает, что из всех путей, проходящих через A и продолжающихся после i -го шага, один, обладающий минимальным расстоянием от $\mathbf{y}$ вплоть до i -го шага, навсегда останется ближайшим к $\mathbf{y}$, поскольку общее продолжение внесет равный вклад во все расстояния.

Тогда зачем продолжать вычисление расстояний для остальных, если совершенно очевидно, что у них нет шансов оказаться в итоге ближайшими к наблюдению? Вместо этого отбросим их, оставляя только один путь с минимальным расстоянием, достигший узла A . Последний называется *выжившим* (survivor) путем и на данный момент будем считать, что имеется только один выживший путь для каждого из узлов решетчатой диаграммы на i -м шаге (см. комментарий ниже). Текущее, т.е. вычисленное по всем символам наблюдения вплоть до i -го шага, расстояние узла A выжившего пути от наблюдения y называется *метрикой узла A* .

Теперь вспомним, что имеются только два ребра, входящие в любой узел, что и показывает рис. 3.9 для некоторого узла A . Два ребра, входящие в него, исходят из узлов B , C предшествующего этапа и, следовательно, продолжают выжившие пути B и C . Можно вычислить расстояния двух путей, достигающих A , в результате вычисления только *метрик ребер*, т.е. расстояний ребер от принятой группы из n наблюдаемых символов, и суммирования их с метриками узлов B и C . Путь с меньшим расстоянием считается выжившим для узла A и запоминается вместе с его расстоянием (метрикой

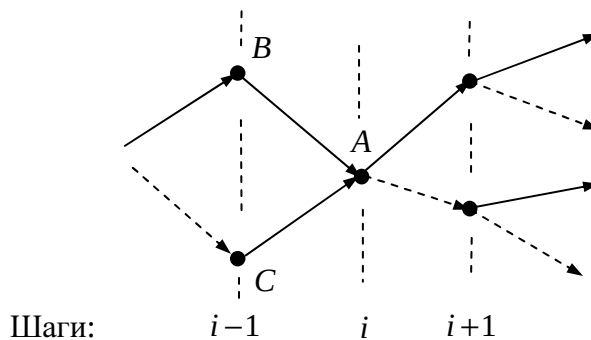


Рис. 3.9. Пути, проходящие через узел A на i -м шаге.

узла), тогда как другой отбрасывается. По выполнению этих операций для всех узлов решетчатой диаграммы декодер переходит к следующему этапу.

Резюмируя сказанное, декодер Витерби на каждом шаге вычисляет метрики ребер, добавляет их ко всем метрикам узлов, накопленных ранее, а затем отбрасывает большее расстояние из двух, ведущих к каждому узлу. Поскольку всего имеется 2^{m-1} узлов (т.е. состояний регистра), сложность декодера определяется только длиной кодового ограничения m и остается фиксированной вне зависимости от теоретически неограниченного числа кодовых слов (путей).

Возвращаясь к сделанному предположению о единственности выживающего пути для каждого узла, отметим, что поскольку расстояние Хэмминга целочисленно, то всегда существует вероятность того, что два пути, ведущие к одному и тому же узлу, обладают одинаковыми текущими расстояниями относительно y . Возможны различные стратегии преодоления неопределенности этого типа. Одна из них использует случайный выбор: подбрасывание монеты и провозглашение выжившим пути, в случае выпадания ассоциированной с ним стороной монеты. Подобное действие, конечно, нарушает оптимальность декодирования, хотя сопровождающие его энергетические потери, как правило, незначительны. Альтернативой является признание обоих конкурирующих путей выжившими и запоминание их до тех пор, пока дальнейшие шаги не разрешат неопределенность. Последняя стратегия сохраняет оптимальность декодирования ценой увеличения требуемой памяти.

Рекуррентный характер процесса сверточного кодирования открывает путь к также рекуррентному и эффективному с вычислительной точки зрения алгоритму декодирования, получившему наименование алгоритма Витерби (по имени автора, предложившего в 1967 г. данный способ декодирования). Сразу же следует подчеркнуть, что алгоритм Витерби является оптимальным, т.е. максимально правдоподобным, и заключается в нахождении кодового слова, наиболее близкого к принятому наблюдению. Первоначально рассмотрим его применение для случая жестких решений, т.е. для двоичного симметричного канала, когда наблюдение Y декодируется в такое кодовое слово, которое находится на наименьшем расстоянии Хэмминга относительно принятого Y . Основная идея декодирования по Витерби заключается в поэтапном сравнении всех путей решетчатой диаграммы (которые в точности представляют собой кодовые слова) с наблюдением Y и отбрасывании тех из них, которые находятся на большем расстоянии от Y , чем некоторые другие. Если два пути, входящих в один и тот же узел, характеризуются различными расстояниями от наблюдения до данного узла, то у пути, обладающим большим расстоянием, отсутствует возможность стать впоследствии более близким к наблюдению, оставаясь на большем расстоянии при любом общем продолжении обоих путей. Следовательно, из двух входящих в узел путей более удаленный от наблюдения может быть исключен из поиска ближайшего пути. Более подробно процедура декодирования может быть описана следующим образом.

Назовем i -м шагом (этапом) декодирования временной интервал, в течение которого принимается i -я n -символьная кодовая группа (кадр) наблюдения Y . Как раз перед этим моментом рассматриваемые пути (т.е. кодовые слова) могут проходить через один из 2^{m-1} узлов (состояний) решетчатой диаграммы. На i -м шаге:

1. Определяются хэмминговы расстояния между принятой n -символьной кодовой группой и каждой из ветвей решетчатой диаграммы. Поскольку из каждого из 2^{m-1} узлов (состояний) выходят две ветви, всего вычисляется 2^m расстояний.

2. Рассматриваются две ветви, идущие из разных предшествующих состояний к каждому из 2^{m-1} узлов:

- 2.1. Отвечающие указанным ветвям расстояния Хэмминга прибавляются к накопленным до i -го шага расстояниям Хэмминга двух соответствующих путей для получения новых значений расстояний. Указанное накапливаемое расстояние пути называется *метрикой*.

- 2.2. Сравняются метрики двух соревнующихся путей, идущих в одно и то же состояние. Путь, находящийся на большем расстоянии от наблюдения, чем другой, отбрасывается и больше не учитывается в процедуре декодирования. Оставшийся путь называется *выжившим путем*.

3. Для всех 2^{m-1} выживших путей запоминаются значения их метрик и декодер готов к переходу на $i+1$ -й шаг процедуры.

На основании рассмотренных операций становится ясным, что ресурсосбережение алгоритма заключается в отбрасывании на каждом шаге ровно половины из 2^m возможных путей, ведущих в 2^{m-1} узлов решетки. В результате число выживших путей остается постоянным и равным 2^{m-1} различных состояний вне зависимости от величины соревнующихся кодовых слов, число которых удваивается на каждом шаге алгоритма декодирования. Для лучшего понимания алгоритма Витерби рассмотрим в подробностях его функционирование на конкретном примере.

Пример 3.1.4. Обратимся к сверточному коду, рассмотренному в примере 3.1.2. Пусть на приемной стороне наблюдается последовательность $Y=(100100000000000000000000)$.

Первые два этапа процедуры декодирования являются тривиальными. На первом

шаге возможны два перехода, описываемые ветвями, которые выходят из состояния 00 и

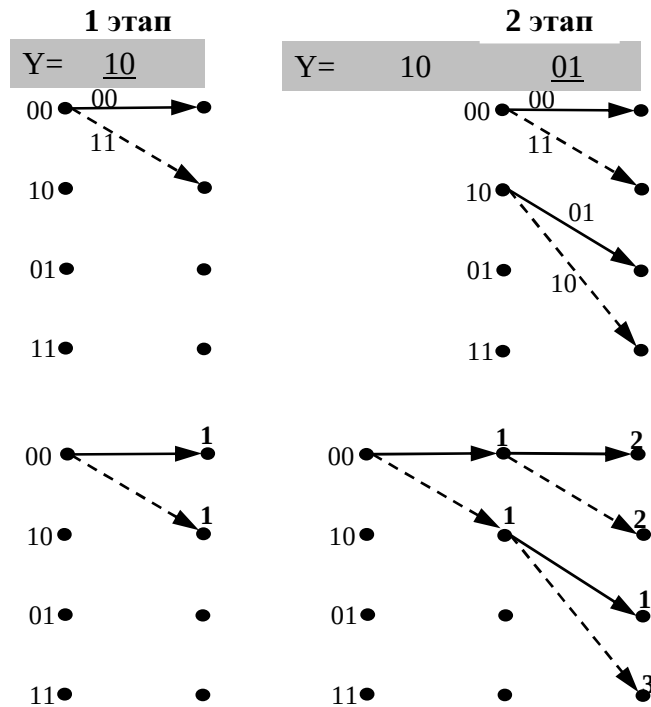


Рис. 3.10.

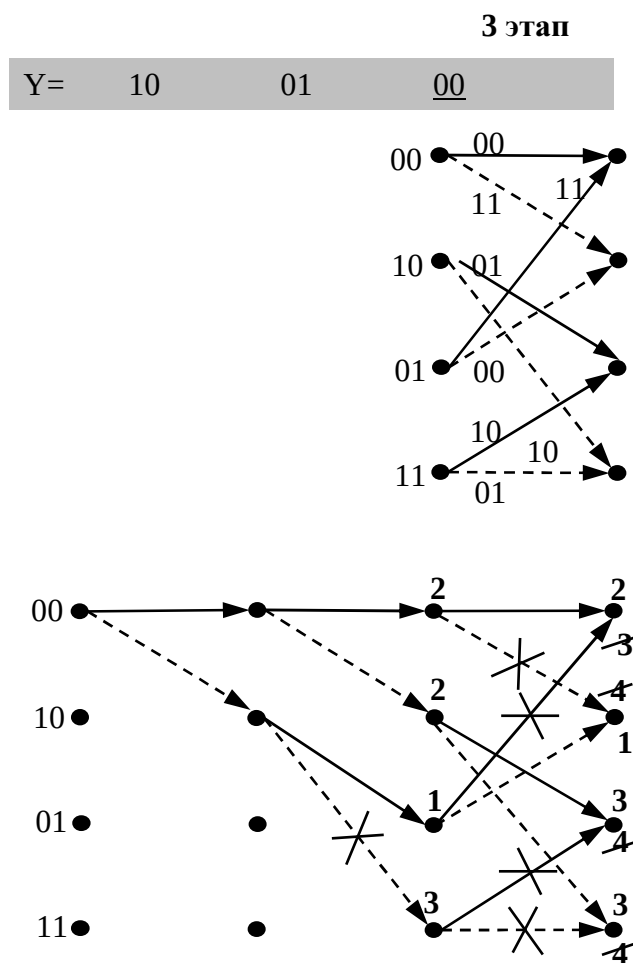


Рис. 3.11.

заканчиваются соответственно в состоянии 00 и 10 с метрикой, равной 1. (см. рис. 3.10). На втором шаге метрика пути, приходящего в состояние 00, становится равным 2, поскольку соответствующий единственный путь начинается в состоянии 00 с метрикой, равной 1, и ее значение увеличивается на единицу, так как расстояние между ветвью (00 → 00) и наблюдаемой 2-х символьной группой Y (в текущий момент 01) также равно 1. Аналогичным образом определяются метрики путей, ведущие в остальные состояния. Так, например, метрика пути, заканчивающегося в состоянии 11, равна 3, поскольку переход из состояния 10 с метрикой 1 в состояние 11 отличается от наблюдения Y на двух позициях.

Критическим для понимания алгоритма является третий этап данного примера (см. рис. 3.11, в верхней части которого изображена решетка кодера). Рассмотрим состояние 00. Достижение его возможно двумя путями: из состояния 00 или 01. Значение накопленной метрики первого пути составит 2, поскольку данный путь начинается из состояния 00 с метрикой 2, а расстояние Хэмминга между ветвью (00 → 00) и 2-х символьной группой наблюдения Y (00) равно нулю. С другой стороны, путь, приходящий из состояния 01, характеризуется аккумулярованной метрикой, равной 3, так как метрика предшествующего состояния 01 равнялась 1, а расстояние Хэмминга между переходом (01 → 00) и наблюдением (00) – 2. В случае дальнейшего рассмотрения обоих путей второй всегда будет находиться на большем расстоянии от наблюдения Y, чем первый, поскольку их общее продолжение в равной мере будет увеличивать их метрики. Однако декодирование по минимуму расстояния предполагает нахождение кодового слова, ближайшего к наблюдению, а значит, совершенно

оправдано отбрасывание пути, который в большей степени отличается от Y , чем альтернативный. Следовательно, путь, идущий через состояние 01 на предшествующем этапе, отбрасывается (перечеркнуто крестом на рис. 3.11), и сохраняется только путь из состояния 00, который является *выжившим путем* для состояния 00 на данном шаге.

Подобным же образом анализируется состояние 10 и определяется, что для него выжившим является путь, выходящий из состояния 01, тогда как путь из состояния 00 отбрасывается, как находящийся на большем расстоянии от наблюдения Y , и т.д.

Этапы декодирования 4, 5 и 6, изображенные на рис. 3.12–3.14, абсолютно аналогичны третьему и не нуждаются в дополнительных комментариях.

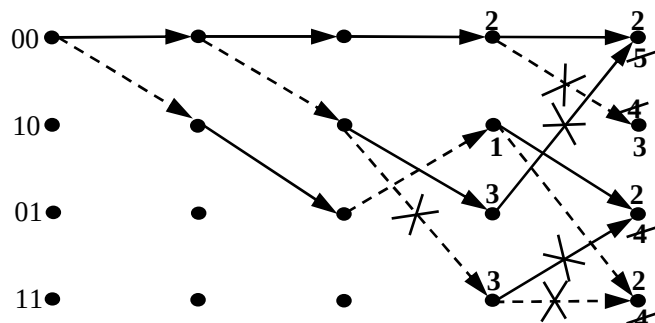
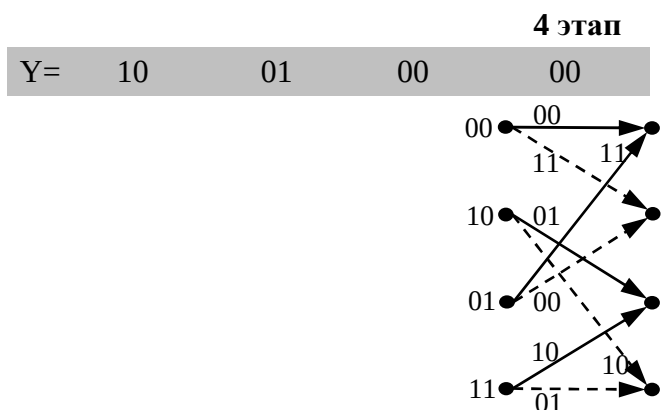


Рис. 3.12.

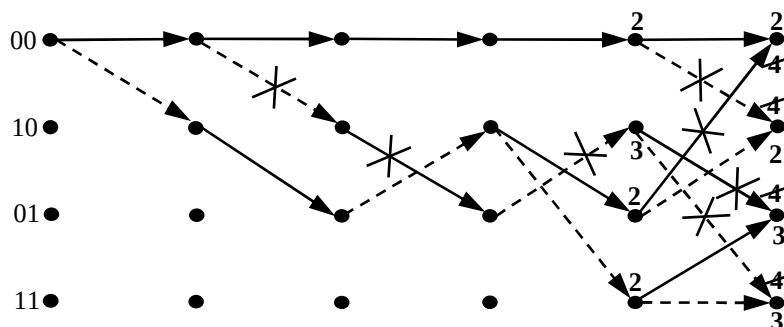
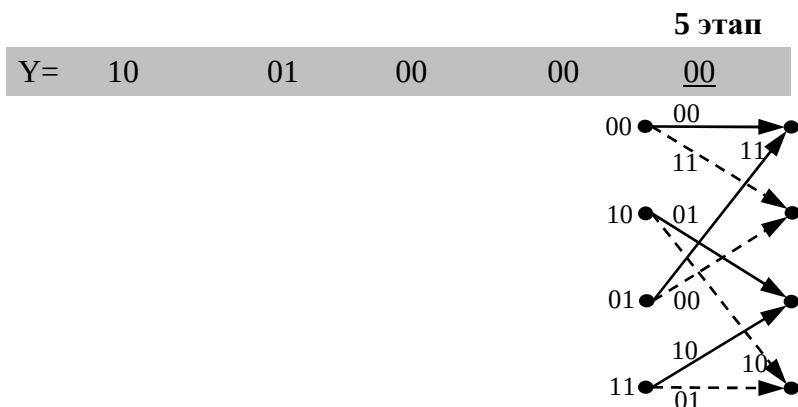


Рис. 3.13.

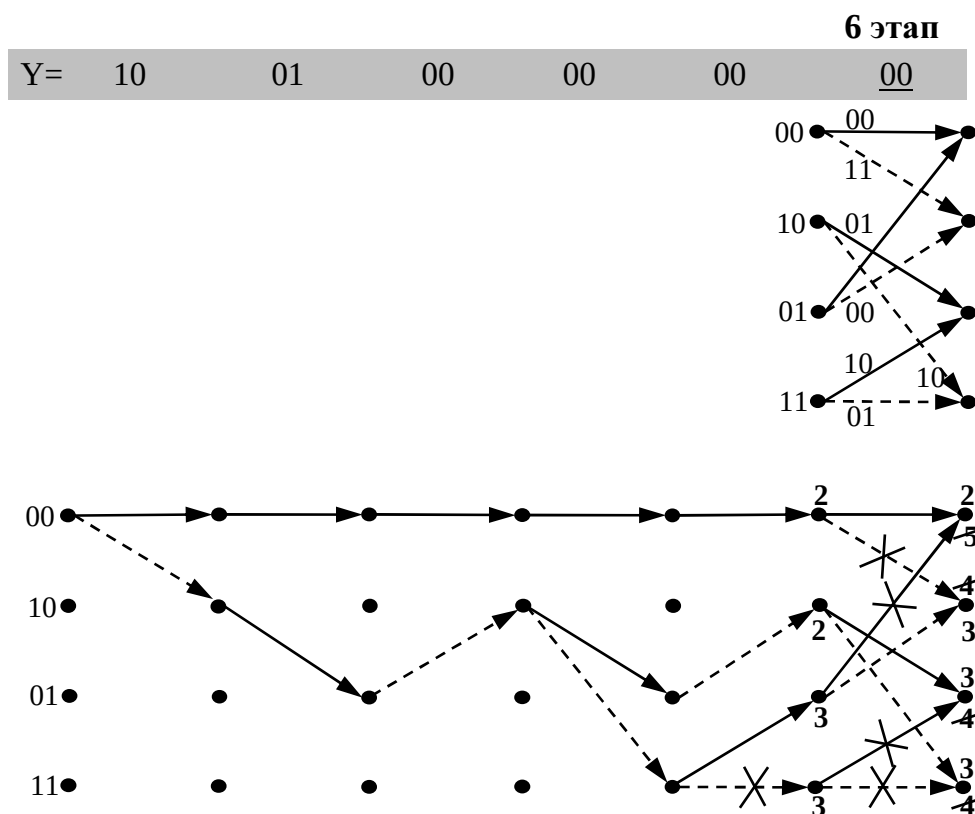


Рис. 3.14.

Особенностью 7-го этапа, изображенного на рис. 3.15, является двузначность, обусловленная равенством метрик конкурирующих путей, сходящихся в одно и то же состояние: пути, приходящие в состояние 01, имеют одну и ту же метрику, равную 4, точно так же, как и пути, сходящиеся в состоянии 11.

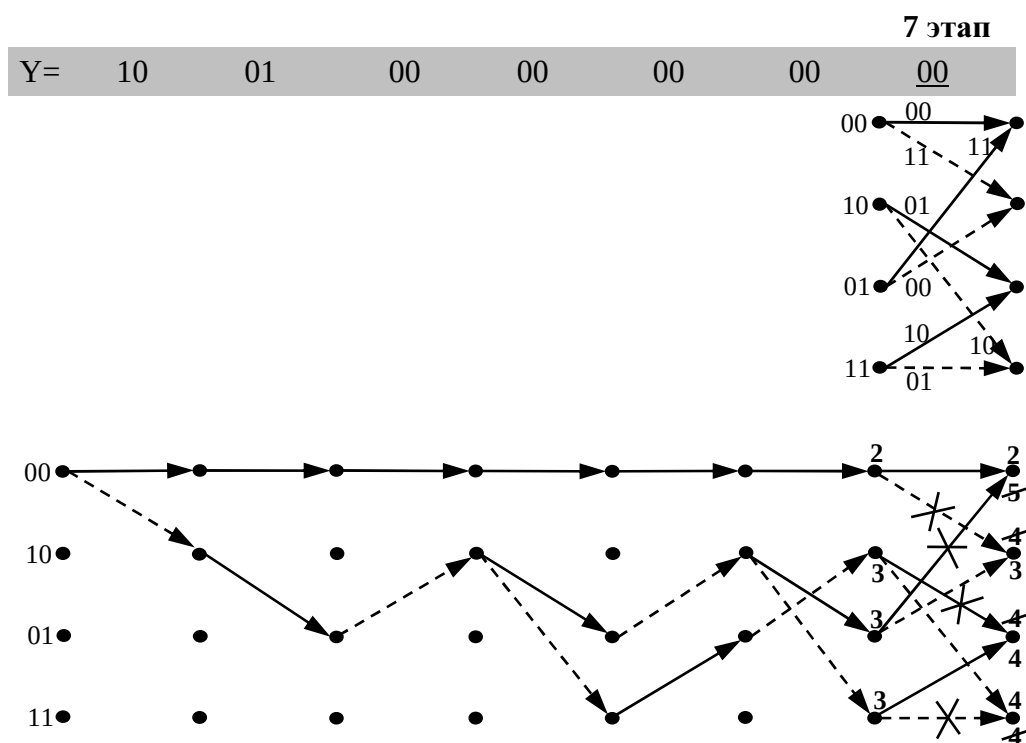


Рис. 3.15.

8 этап
Y= 10 01 00 00 00 00 00 00

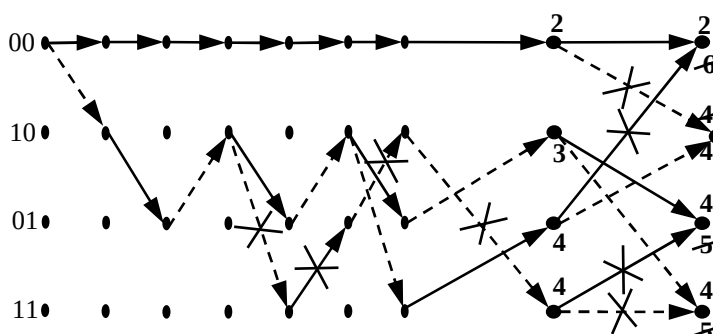
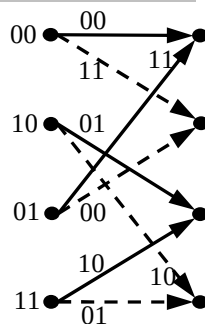


Рис. 3.16.

становится очевидным, что случайный выбор пути в состоянии 11 на предыдущем этапе не имел последствий, поскольку отвечающие ему пути исключаются из дальнейшего рассмотрения.

Ситуации, подобные ранее рассмотренной, встречаются и на последующих 9–м и 10–м этапах, возможные переходы на которых изображены на рис. 3.17–3.18. В частности,

9 этап
Y= 10 01 00 00 00 00 00 00 00 00

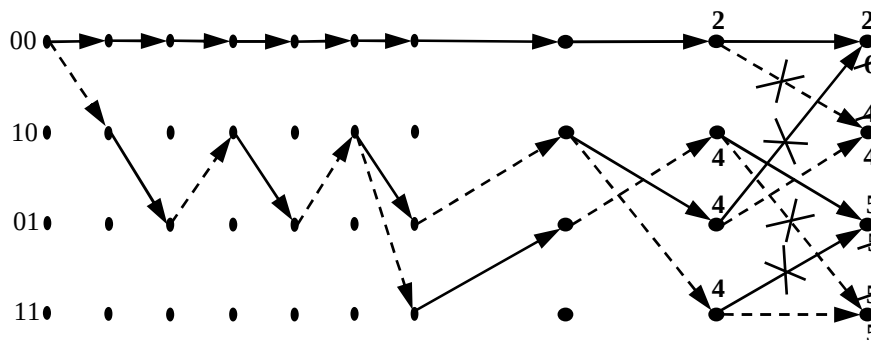
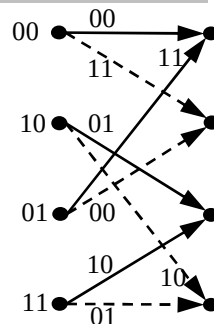


Рис. 3.17.

Для разрешения подобной неопределенности могут быть использованы различные стратегии. Простейшей является стратегия, основанная на «выбрасывании монеты», т.е. случайного выбора одного из двух путей с вероятностью 0,5. Применяя указанную стратегию, предположим, что выжившими остаются пути, изображенные на рис. 3.15.

На 8–м этапе, представленном на рис. 3.16, в состоянии 10 вновь возникает аналогичная двузначная ситуация. Однако, более интересным представляется тот факт, что оба пути, прошедшие на предыдущем этапе через состояние 11, отбраковываются, т.е. они не остаются в числе выживших. С учетом этого

на 9–м этапе неопределенность в пути, возникающая в состояниях 10, 01 и 11, вновь разрешается случайным выбором. Однако уже на 10–м шаге этот произвольный выбор для состояний 01 и 11 теряет свое значение, поскольку траектории, проходящие через них, не выживают.

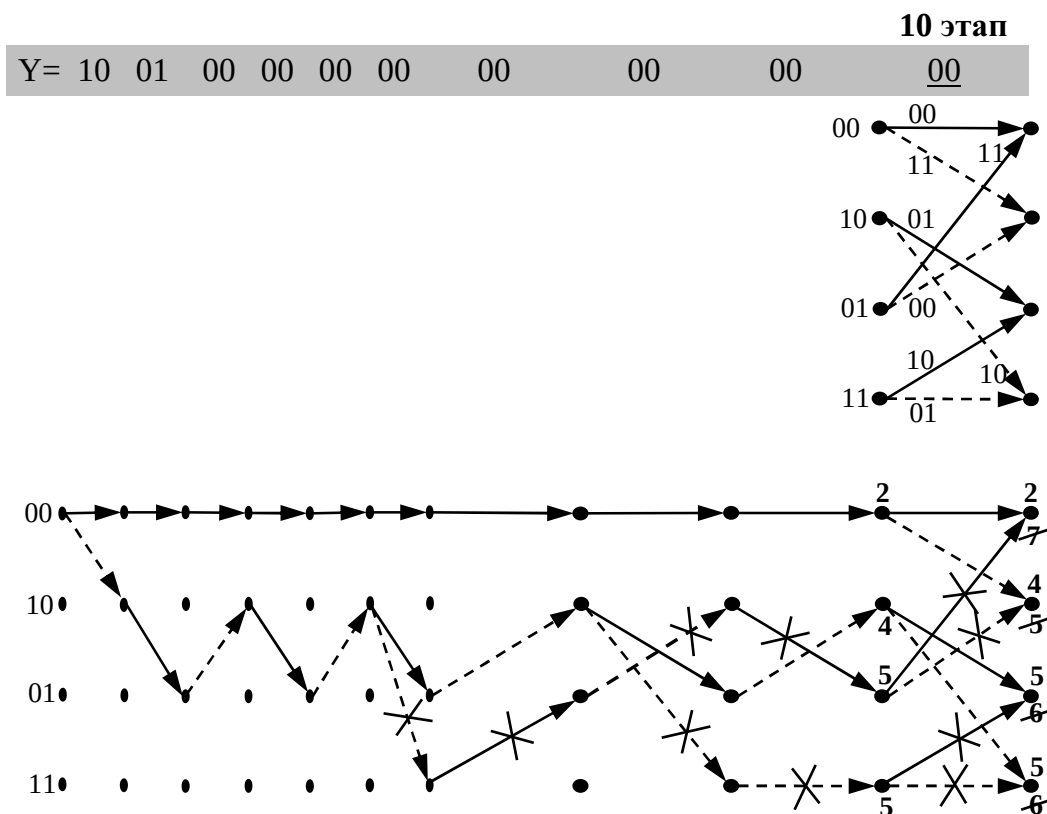


Рис. 3.18.

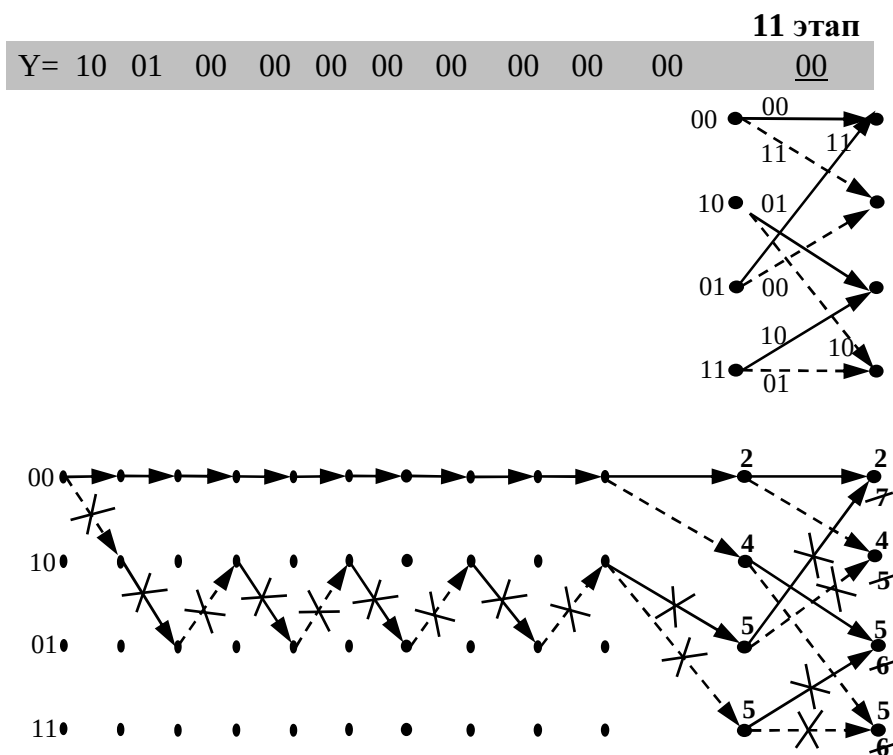


Рис. 3.19.

На одиннадцатом этапе получается модель переходов, имеющая чрезвычайно определенный вид, которую рассмотрим более подробно (Рис. 3.19). Во-первых, маршруты с предшествующими неопределенными решениями отбраковываются, и значит, возможное влияние не очевидных решений на более ранних этапах полностью исключается. Во-вторых, все выжившие пути за один или два этапа от текущего слились в состоянии

00. Следовательно, любой будущий выживший путь будет иметь один и тот же начальный сегмент от 1-го до, по крайней мере, 9-го шага. Наличие общего начального сегмента

означает, что осуществлено декодирование первых 9 бит переданной кодовой последовательности: 000000000.

После 12-го этапа, изображенного на рис. 3.20, вновь остаются выжившими только траектории, проходящие через состояние 00 одним или двумя шагами ранее, что позволяет декодировать еще один бит (10-й по счету), так что декодированной считается последовательность вида 0000000000.

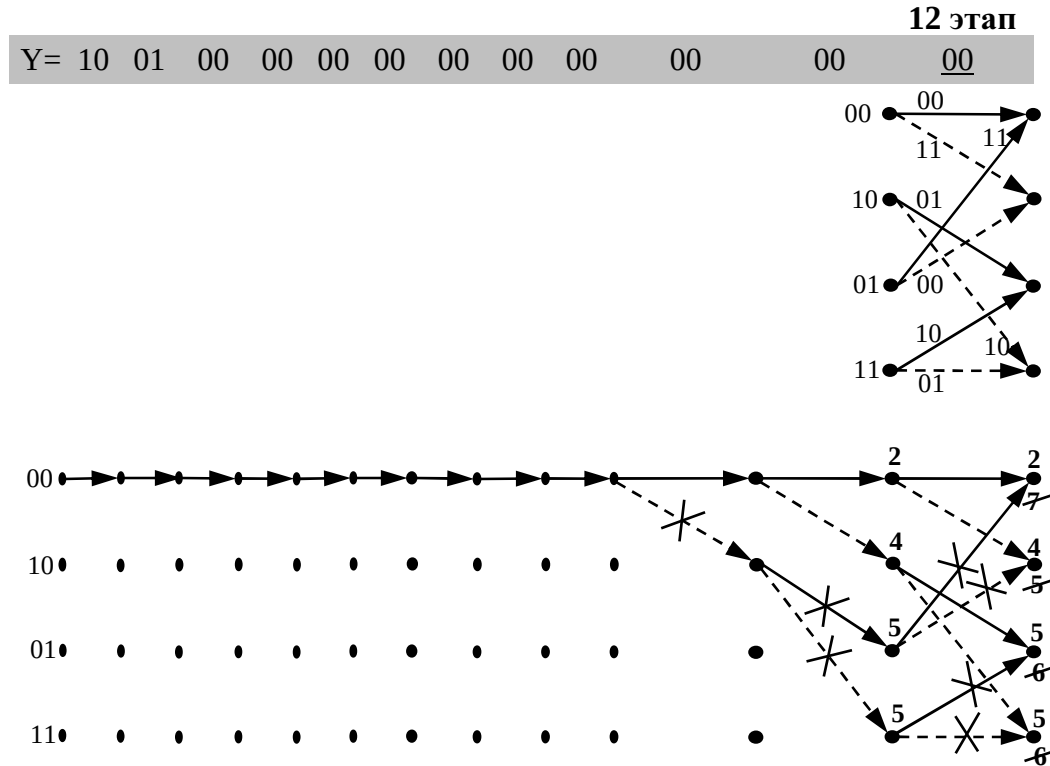


Рис. 3.20.

Обсудим теперь, что происходит после прекращения потока передаваемых битовых символов. Положим, что длина потока равна 12. После прекращения информационного потока кодер продолжает свою работу, поскольку ему необходимы еще $m-1$ тактов для своего обнуления. Можно считать, что на время обнуления на вход кодера дополнительно поступают $m-1$ нулевых символов, причем декодер знает об их конкретных значениях. Однако, поскольку кодовые символы продолжают зависеть от предшествующих, то продолжается и их передача. Для рассмотренного примера $m-1=2$ и кодовые символы, отвечающие входным битам «обнуления», изображены на решетчатой диаграмме бесцветными. Поскольку декодеру известно, что, начиная с 13-го шага, идет процесс обнуления кодера, то возможными являются только переходы в состояния 00 или 01, что нашло отражения в диаграмме, представленной на рис. 3.21. Тем самым исключаются из рассмотрения 4 из 8 ветвей в решетчатой диаграмме и только для двух узлов вычисляются метрики. В результате остаются выжившими только два пути и декодируется еще один (11-й) бит: 00000000000. На 14-м (окончательном) шаге декодирования (см. рис. 3.21) осуществляется выбор между двумя путями, сливающимися в состоянии 00, устанавливая окончательный результат декодирования: 000000000000.

Теперь становится очевидным, что единственным выжившим (а значит, и декодированным) путем является нулевой. Если же нулевая последовательность действительно передавалась по каналу, то произошло исправление двукратной ошибки в полном соответствии с корректирующей способностью кода (свободное расстояние рассматриваемого кода $d_f = 5$).

С точки зрения практической реализации выдавать решение о декодированной последовательности символов только тогда, когда все выжившие пути сольются в одном единственном состоянии, не может считаться удобным. Вследствие этого широко распространена практика принятия решения после прохождения определенного числа t_d этапов по решетчатой диаграмме. Неоднократно экспериментально и с помощью компьютерного моделирования было подтверждено, что задержка в выдаче результата декодирования, равная $t_d = 5m$, т.е. пяти длинам кодового ограничения, является достаточной для сохранения практической оптимальности алгоритма Витерби. Только что завершённый пример может служить частичным подтверждением данного утверждения.

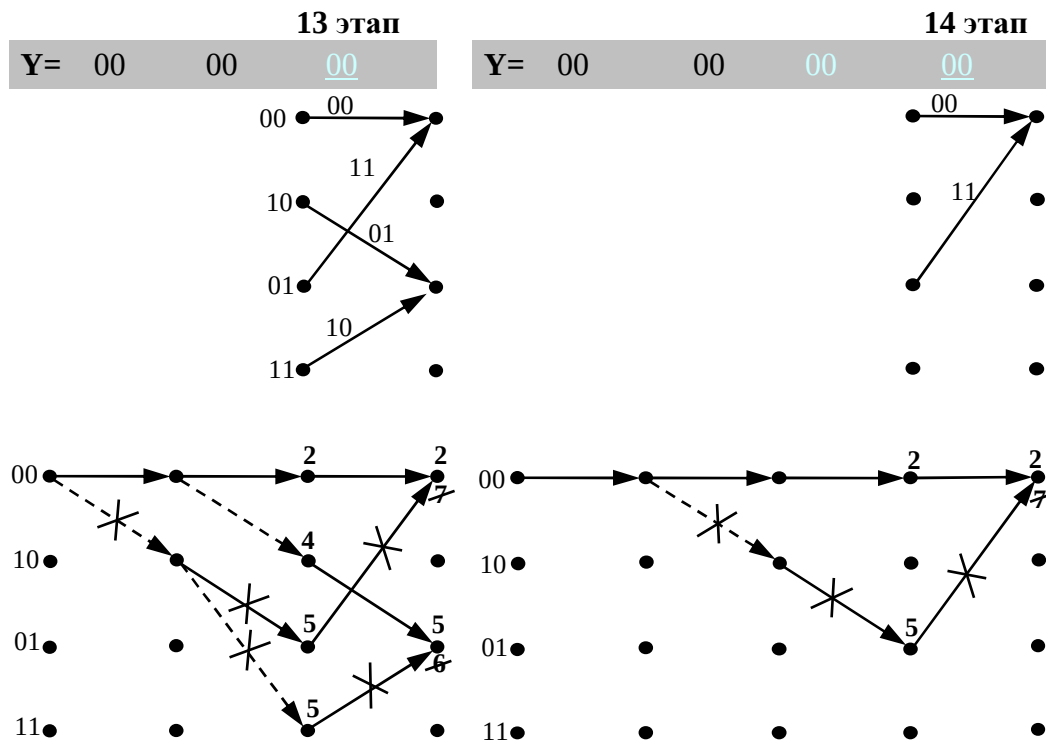


Рис. 3.21.

3.9. Декодирование с мягкими решениями

При передаче потока данных по гауссовскому каналу наилучшая возможная (максимально правдоподобная) стратегия состоит в декодировании принятого непрерывного наблюдения $y(t)$ в кодовое слово, ближайшее к нему в смысле евклидова расстояния. Для блочных кодов подобный метод декодирования в большей степени является бесперспективным вследствие значительной сложности практической реализации декодера. Например, рассмотрим код БЧХ, исправляющий любые трехкратные и менее ошибки. Применение стратегии мягких решений при его декодировании потребует вычисления $2^{16} > 6,4 \cdot 10^4$ евклидовых расстояний между наблюдением $y(t)$ и каждым из конкурирующих кодовых слов, что, как правило, считается неприемлемой платой в обмен на 2...3 дБ выигрыша по сравнению с использованием жестких решений.

Очень важным свойством сверточных кодов, делающим их еще более привлекательными, является сравнительная простота реализации мягкого декодирования. В самом общем случае блочного кода с M кодовыми словами осуществление мягкого декодирования означает непосредственное вычисление M евклидовых расстояний или корреляций, что представляется достаточно сложным алгебраически по сравнению с

синдромным декодированием. При значительных величинах M , типичных для множества приложений, данный факт часто превращает указанную задачу в полностью не осуществимую. В то же время платой за упрощения приемника, использующего жесткие решения, служат энергетические потери, оцениваемые для АБГШ канала в 2 дБ: цифрой, рассматриваемой сегодня, как достаточно значительной. Обратимся вновь к алгоритму Витерби и просто заменим расстояние Хэмминга евклидовым. Очевидно, что данная операция превратит декодирование в его вариант с мягкими решениями, оптимальный для АБГШ канала. Тогда метрики ребер и узлов в точности отвечают соответствующим евклидовым расстояниям (или корреляциям). Подобная модификация ни в малейшей степени не затронет реализационные преимущества алгоритма Витерби. Действительно, метрики узлов вычисляются, как и ранее рекурсивно, путем пошагового увеличения за счет метрик ветвей, а путь, входящий в узел с худшей метрикой, чем альтернативный, снова может быть отброшен на каждом шаге, как не выживший.

Вместе с тем, среди 2^m ветвей число отличающихся по расстоянию не превосходит 2^n , что и определяет максимальное число вычисляемых евклидовых расстояний на каждом шаге. Например, рассмотрим сверточный код с длиной кодового ограничения $m=5$, скоростью $R=1/2$ и свободным расстоянием $d=7$. По своим скоростным и корректирующим способностям данный сверточный код в значительной мере идентичен упомянутому ранее БЧХ-коду. Решетчатая диаграмма сверточного кода в каждом сечении содержит $2^4=16$ узлов и требует на каждом этапе вычисления не более, чем 32 евклидовых метрик. Значение каждой из этих метрик складывается из величины расстояния ветвей текущего этапа (всего их $2^2=4$) и соответствующих метрик путей, сохранившихся к предшествующему этапу. После этого один из двух путей, приходящих в каждый узел, отбрасывается, и процедура переходит к следующему этапу. Как следует из описания алгоритма, не существует никаких технических препятствий к использованию мягкого декодирования. В случае замены евклидова расстояния вычислением корреляции между отсчетами наблюдения и тестовым путем процедура мягкого декодирования становится еще более осуществимой. Вместе с тем следует помнить, что при использовании мягкого декодирования маркировка путей решетчатой диаграммы должна быть соответствующим образом трансформирована. Например, при передаче двоичных кодовых символов с помощью бинарных фазоманипулированных сигналов отметки на путях изменятся с 0 на 1 и с 1 на -1 и т.п.

Конечно, цифровая реализация декодеров является наиболее предпочтительной, подразумевая квантование входного наблюдения. Общепринятой является тенденция классификации декодеров двоичных кодов, использующих квантование более чем на два уровня, как декодеры с мягким декодированием. Углубленный анализ показывает, что в большинстве случаев 3-битовое (8-ми уровневое) квантование достаточно для достижения почти потенциальных (характерных для непрерывной обработки) параметров.

Возможность применения стратегии мягких решений часто служит одной из основных причин предпочтения сверточных кодов перед блоковыми.

3.10. Сверточные коды с выкалыванием

В том случае, когда при проектировании системы возникает необходимость в

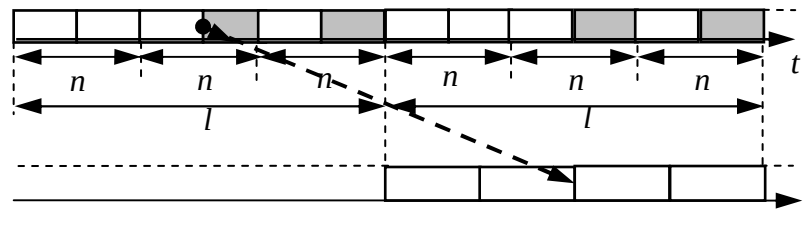


Рис. 10.22.

использовании сверточного кода со скоростью t/n ($t > 1$), он может быть построен с использованием алгоритма, описанного в 3.1. Однако более перспективным с практической точки зрения представляется метод формирования *сверточных кодов с выкалыванием*. Под операцией *выкалывания* понимается удаление из кода некоторых символов по правилу, согласованному между передающей и приемной сторонами. Возьмем сверточный код со скоростью $1/n$ и поделим его на блоки, состоящие из l групп по n символов. Затем удалим в каждом блоке по одному кодовому символу в g n -символьных группах. Описанная процедура поясняется примером с параметрами $n = 2, l = 3, g = 2$, представленным на рис. 3.22, где удаляемые символы окрашены в серый цвет.

Код, формируемый подобным образом, содержит l информационных бит в каждом блоке, длина которого после удаления g символов будет составлять $(l - g)n + g(n - 1) = ln - g$ символов, так что в результате получаем код со скоростью

$$R = \frac{l}{ln - g}.$$

Изменением параметров l и g при фиксированном n можно регулировать величину скорости в значительном диапазоне, добиваясь ее желаемого значения. Так, в рассмотренном примере $R = 3/4$.

Очевидно, что после подобной трансформации кода его корректирующие способности могут значительно измениться, поэтому для оптимизации кодов с выкалыванием был произведен значительный машинный поиск, в результате которого получен целый ряд кодов с хорошими характеристиками.

Выкалывание символов не приводит к какому-нибудь усложнению процедуры декодирования. В самом деле, приемная сторона заранее знает о том, какие кодовые символы удаляются, и поэтому у нее не возникнет никаких трудностей при применении алгоритма декодирования Витерби.

Практическое использование сверточных кодов

Примеры использования сверточных кодов в современных информационных технологиях чрезвычайно обширны. Кратко остановимся только на тех, которые относятся к системам мобильной радиосвязи. Так, в системе второго поколения стандарта GSM канальное кодирование осуществляется на основе сверточного кода с длиной кодового ограничения, равной 5, скоростью $R = 1/2$ и свободным расстоянием $d_f = 7$. Сотовая система связи стандарта IS-95 на основе CDMA технологии, также относящаяся ко второму поколению, и 3G стандарт UMTS использует два различных сверточных кода в прямом и обратном каналах связи. В канале «вниз» применяется код с длиной кодового ограничения $m = 9$, скоростью $R = 1/2$ и свободным расстоянием $d_f = 12$, тогда как в

канале «вверх» используется более мощный код с параметрами $m=9$, $R=1/3$ и $d_f=18$. Широко представлены сверточные коды, коды с выкалыванием, а также турбо-коды (представляющие собой дальнейшее развитие сверточных) в существующих проектах систем третьего поколения (UMTS, cdma2000 и др.). Помимо их собственного значения сверточные коды служат основой для построения турбо-кодов, приближающих надежность передачи данных к пределу Шеннона.

4. СИНХРОНИЗАЦИЯ В РАДИОСИСТЕМАХ ПЕРЕДАЧИ ДИСКРЕТНОЙ ИНФОРМАЦИИ

4.1. Принципы построения и основные характеристики систем синхронизации

Система синхронизации в РСПИ в общем случае должна определять следующие синхропараметры сигнала:

- фазу высокочастотного несущего колебания (фазовая синхронизация ФС);
- временные границы принимаемых посылок (тактовая синхронизация ТС);
- моменты, соответствующие началу кодовых слов (цикловая синхронизация ЦС);
- моменты, соответствующие началу и концу групповых сигналов в многоканальной СПИ (кадровая синхронизация);
- начало и конец передаваемого сообщения.

В подавляющем большинстве случаев сигналы тактовой, кадровой и цикловой синхронизации связаны по фазе между собой (синхронны). Частота повторения кодовых слов $f_{ц}$ получается делением тактовой частоты f_T на число разрядов в кодовом слове ($f_{ц} = f_T/n$), частота повторения кадров - делением частоты повторения кодовых слов на число кодовых слов в кадре $f_k = f_{ц}/k_{сп}$

Все устройства синхронизации (УС), входящие в систему синхронизации (СС), можно разделить на два принципиально различных типа. Первый тип УС служит для синхронизации отсчетов времени (фазовая и тактовая синхронизация). С их помощью формируются временные шкалы. Второй тип устройств служит для устранения неоднозначности отсчетов времени при определении начала слова, кадра и сообщения. Устройства синхронизации отсчетов времени должны функционировать непрерывно, отслеживая изменение фазы входного колебания, а функции устройств устранения неоднозначности отсчетов времени сводятся к периодическому, а иногда и к однократному фазированию.

Принципиальное различие указанных двух типов УС, естественно, сказывается на методах поиска и оценки синхропараметров, а также на характеристиках качества их работы. Качество работы УС в общем случае должно определяться степенью соответствия фазы входного колебания и колебания местного генератора. До начала работы СС неопределенность в оценке фазы φ принимаемых синхроколебаний может быть задана плотностью распределения $W_{вх}(\varphi)$. Если нет дополнительной информации о параметре φ , то логично предположить, что закон распределения $W_{вх}(\varphi)$ равномерный на интервале $[-\pi, \pi]$. В результате работы СС неопределенность уменьшается, причем происходит это поэтапно. Сначала при когерентном приеме осуществляется ФС, затем ТС и только потом устраняется неоднозначность отсчетов в устройствах ЦС и КС.

Погрешности синхронизации отсчетов времени при случайных внешних воздействиях могут быть заданы законом распределения $W(\varphi)$, а скорость их изменения - корреляционной функцией $R_{\varphi}(\tau)$ или спектром флуктуации $G_{\varphi}(f)$. Чаще всего полагают закон распределения $W(\Delta\varphi)$ гауссовским с дисперсией σ^2 .

Удобной количественной характеристикой оценки погрешности синхронизации является вероятность попадания фазы колебания местного генератора в некоторую область допустимых значений $\Delta\varphi_{сх}$. Эту область можно определить как *область синхронизма*. Если $\Delta\varphi_{сх}$ задана, то можно ввести еще ряд показателей качества работы УС. Такими показателями следует считать время достижения синхронизма $T_{сх}$ (длительность переходных процессов до достижения области синхронизма), вероятность срыва синхронизма и время поддержания синхронизма при пропадании сигнала на входе.

При наличии на входе сигнала решение о синхронизме может быть правильным

или нет. Поэтому качество синхронизации следует характеризовать вероятностью ложного синхронизма $p_{лс}$ при заданном времени анализа T_a .

В ряде РТСПИ важно знать начало передаваемого сообщения. Для этого в начале сеанса связи передается специальный сигнал (преамбула), по которому оценивается факт передачи сообщения и его временное положение. В этом случае возможны ошибки следующих видов: ложное обнаружение сигнала с вероятностью $p_{лв}$, пропуск сигнала с вероятностью $p_{пс}$ и ложная синхронизация, когда выносится решение, что сигнал есть, но его временное положение оценивается неправильно с вероятностью $p_{лс}$.

Обработывая входной сигнал, СС получает информацию о синхропараметрах. В простейшем случае, когда в спектре принимаемого сигнала содержится составляющая требуемой частоты и фазы (например, при ФС), обработка сводится к фильтрации синхроколебания. Если в спектре сигнала отсутствуют составляющие, несущие информацию о синхропараметрах, то сигнал предварительно подвергается в приемнике нелинейным преобразованиям.

В общем случае УС должно содержать входной преобразователь ВП, в выходном сигнале которого содержится колебание требуемой частоты и фазы. Этим колебанием синхронизируется местный управляемый генератор УГ. Для уменьшения дисперсии флуктуации фазы, возникающих из-за действия помех, синхросигнал предварительно фильтруется полосовым фильтром ПФ.

Различают УС *разомкнутые* (рис. 4.1, а), в которых синхроколебание фильтруется полосовым фильтром (аналоговым или цифровым), и *замкнутые* (рис. 4.1, б), построенные на базе систем фазовой автоподстройки частоты (ФАПЧ). В замкнутых системах колебание синхронизируемого УГ в фазовом детекторе ФД сравнивается по фазе с входным колебанием. Затем сигналом рассогласования УГ управляется по частоте так, чтобы свести это рассогласование к минимуму. Сглаживание сигнала рассогласования осуществляется низкочастотным фильтром ФНЧ. Принципиальным достоинством системы ФАПЧ является то, что она представляет собой следящий фильтр, в котором шумовая полоса $F_{ш}$ может быть достаточно узкой независимо от диапазона изменения частоты входного колебания.

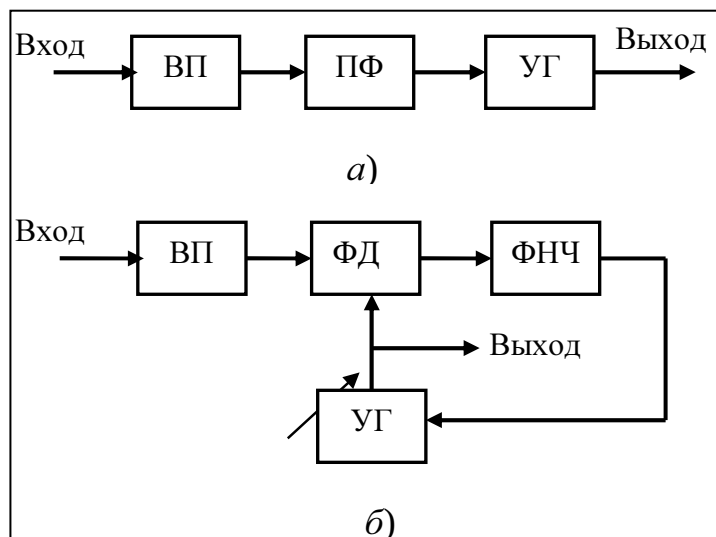


Рис. 4.1. Структурные схемы разомкнутой (а) и замкнутой (б) систем

4.2. Влияние точности оценки синхропараметров на качество работы систем

Работа отдельных ступеней СС по-разному сказывается на качестве работы РСПИ в целом. Погрешности в работе устройства цикловой синхронизации (УЦС) и устройства кадровой синхронизации (УКС) приводят к неправильному приему всего сообщения или

части его, а устройства фазовой синхронизации (УФС) и устройства тактовой синхронизации (УТС) - к снижению его достоверности. Оценим влияние погрешностей УФС и УТС на достоверность принимаемого сообщения. Устройство фазовой синхронизации входит в состав когерентного демодулятора и обеспечивает совпадение по фазе напряжения местного генератора и несущей частоты. Погрешности $\Delta(\varphi)$ УФС приводят к уменьшению отношения сигнал/шум q на выходе устройства обработки, а при фазовой модуляции к изменению коэффициента корреляции между сигналами в ансамбле

$$q^2(\Delta\varphi) = (2E / N_0) \cos^2(\Delta\varphi).$$

Поэтому, оценивая достоверность принимаемого сообщения с учетом ошибок ФС, можно ввести условную вероятность ошибки $p_{\text{ош}}(\Delta\varphi)$. Если скорость флуктуации ошибок ФС мала (интервал корреляции случайного процесса $\Delta\varphi(t)$ много больше длительности посылок, $T_k \gg T_c$), средняя вероятность ошибок

$$\bar{p}_{\text{ош}}(\Delta\varphi) = \int_{-\pi}^{\pi} p_{\text{ош}}(\Delta\varphi) W(\Delta\varphi) d(\Delta\varphi). \quad (4.1)$$

Сложность зависимости $p_{\text{ош}}(\Delta\varphi)$ не позволяет получить точной формулы для $\bar{p}_{\text{ош}}(\Delta\varphi)$ (рис. 4.2, а). Поэтому пользуются либо приближенными соотношениями, найденными при тех или иных упрощениях, либо численными методами с привлечением ЭВМ.

Влияние ошибок ФС на помехоустойчивость можно оценить по приближенной формуле, справедливой для систем с фазовой модуляцией любой кратности m

$$\bar{p}_{\text{ош}}(\Delta\varphi) \approx 1 - F \left[\frac{\pi}{m} \frac{h}{\sqrt{1 + 2\sigma_{\varphi}^2 h^2}} \right] \quad (4.2)$$

Значение σ_{φ} на рис. 4.2,а и в формуле (4.2) задано в радианах.

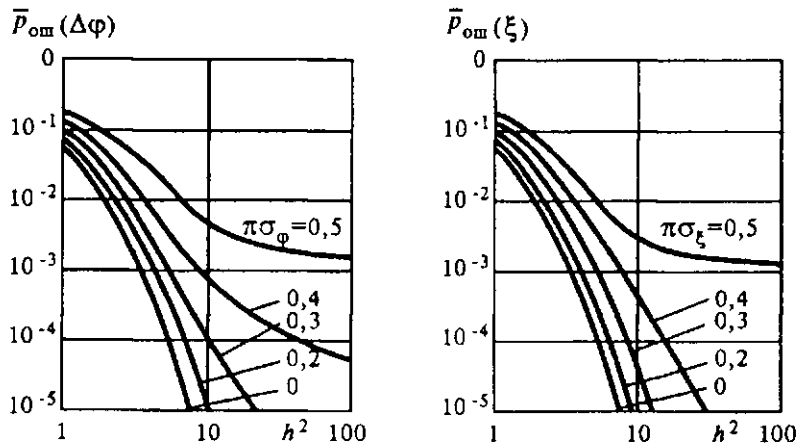


Рис. 4.2. Зависимости вероятности ошибки когерентного приема противоположных двоичных сигналов при наличии погрешностей фазовой (а) и тактовой (б) синхронизации

Погрешности тактовой синхронизации τ приводят к тому, что момент отсчета напряжения на выходе коррелятора или согласованного фильтра не соответствует моменту окончания посылки. Тогда при смене символов на входе решающей схемы будут накапливаться сигналы от посылок разного знака и отношение сигнал/шум q^2 уменьшится

$$q^2(\xi) = \frac{2E}{N_0} \left(1 - 2 \frac{|\tau|}{T_c} \right) = q^2(1 - 2|\xi|).$$

Введя условную вероятность ошибки $p_{ош}(\xi)$ и зная закон изменения $W(\xi)$, можно вычислить среднюю вероятность ошибки $\bar{p}_{ош}(\xi)$ аналогично (4.1). Как и в предыдущем случае, количественные результаты можно получить численными методами. Ошибки тактовой синхронизации, в отличие от фазовой, будут влиять только на прием чередующихся символов. Для случайной последовательности двоичных символов вероятность смены знака равна 0,5, а средняя вероятность ошибки (рис. 4.2, б)

$$\bar{p}_{ош}(\xi) = 0,5[p_{ош}(\xi) + p_{ош}],$$

где $P_{ош}$ - вероятность ошибки при идеальной тактовой синхронизации.

Интересно сравнить степень влияния на помехоустойчивость ошибок фазовой и тактовой синхронизации при заданном отношении сигнал/шум в канале. Для этого допустим, что тактовая частота формируется из опорной, равной частоте несущей, путем деления на k_d . Тогда флуктуации фазы тактовой частоты будут меньше флуктуации фазы опорной частоты в k_d раз. Из этого следует, что в режиме слежения, когда неоднозначность отсчета фазы тактовой частоты устранена, в основном будут влиять на помехоустойчивость ошибки устройства фазовой синхронизации. Это утверждение справедливо и тогда, когда опорная и тактовая частоты формируются от разных генераторов.

4.3. Фазовая синхронизация

4.3.1. Фазовая автоподстройка частоты

Система ФАПЧ (см. рис. 4.1, б) является основным звеном устройств синхронизации отсчетов времени. Она в том или ином виде входит в УФС и УТС демодулятора и служит для фильтрации сигнала синхронизации. Работу ФАПЧ можно характеризовать режимами слежения и захвата. В режиме слежения частоты входного и опорного сигналов одинаковы. С фазового детектора снимается напряжение, пропорциональное $\cos(\varphi)$, где φ - разность фаз между входным и опорным сигналами. Этим напряжением управляется по частоте (фазе) УГ так, чтобы свести рассогласование к $\pi/2$. Включив фазовращатель на $\pi/2$, можно компенсировать постоянный фазовый сдвиг между входным и опорным сигналами. В установившемся режиме значение φ определяется параметрами ФАПЧ, начальной разностью частот между входным и опорным сигналами и уровнем шумов на входе. Если систему ФАПЧ рассматривать как четырехполосник, в котором выходным сигналом являются колебания УГ, то по своим свойствам она подобна следящему полосовому фильтру. Характеризовать работу такого следящего фильтра можно следующими параметрами:

- шумовой полосой $F_{ш}$ (полосой пропускания);
- полосой удержания F_y (максимальным изменением частоты входного колебания, при котором ФАПЧ, находясь в режиме слежения, сохраняет работоспособность);
- полосой захвата F_z (максимальной расстройкой между частотой входного колебания и частотой управляемого генератора, при которой после включения ФАПЧ начинает следить за фазой входного колебания, входит в синхронизм);
- временем ввода в синхронизм $T_{сх}$, которое характеризует продолжительность переходного процесса от момента включения сигнала на входе до момента установления равенства частот входного сигнала и сигнала управляемого генератора.

Анализ работы системы ФАПЧ, особенно в режиме захвата и при воздействии помех, представляет сложную математическую задачу. Сложность ее решения определяется тем, что поведение ФАПЧ описывается нелинейным дифференциальным уравнением. Обозначим частоты входного и опорного сигналов соответственно через

$\omega_0 + \Delta\omega_{\text{вх}}(t)$ и $\omega_0 + \Delta\omega_{\text{уг}}(t)$, где ω_0 - центральная частота, а $\Delta\omega_{\text{вх}}(t)$ и $\Delta\omega_{\text{уг}}(t)$ - изменение ее во времени для входного и опорного сигналов. Тогда разность частот входного сигнала и опорного $\Delta\omega(t) = \Delta\omega_{\text{вх}}(t) - \Delta\omega_{\text{уг}}(t)$. Рассматривая $\Delta\omega_{\text{вх}}(t)$ как входное воздействие на систему, а $\Delta\omega(t) = d\varphi(t)/dt$ как реакцию системы, можно записать дифференциальное уравнение, описывающее ее поведение

$$\partial\varphi/\partial t + u\Omega F(\varphi)K(p) = \Delta\omega_{\text{вх}}(t),$$

где φ - текущая разность фаз между сигналами управляемого генератора и входным; u - напряжение входного сигнала; $F(\varphi)$ - нормированная характеристика фазового детектора; $K(p)$ - коэффициент передачи фильтра в операторной форме; $p = \partial/\partial t$ - оператор дифференцирования; Ω - коэффициент передачи управляемого генератора, характеризующий приращение частоты управляемого генератора при подаче на его вход постоянного напряжения 1 В.

Характеристика фазового детектора $F(\varphi)$ нелинейна, периодична и определяется формой опорного и входного сигналов. Если аппроксимировать ее полигональной кривой (рис. 4.3), то $F(\varphi) = -(2/\pi)\varphi$, $-\pi/2 < \varphi < \pi/2$. При малых внешних воздействиях, когда частоты входного и опорного сигналов совпадают, а фазовое рассогласование не превышает $\pi/2$, систему ФАПЧ можно заменить эквивалентной линеаризованной моделью (рис. 4.4). Управляемый генератор, для которого частотный сдвиг выходного сигнала пропорционален входному сигналу, а фазовый сдвиг – интегралу от него, выполняет роль интегрирующего звена с коэффициентом передачи в операторной форме $u\Omega/p$.

Передаточная функция замкнутой линеаризованной системы ФАПЧ $H(p)$ связана с передаточной функцией фильтра системы $K(p)$ в операторной форме соотношением

$$H(p) = \frac{\varphi_{\text{уг}}(p)}{\varphi_{\text{вх}}(p)} = \frac{2u\Omega K(p)}{(2/p)u\Omega K(p) + p}$$

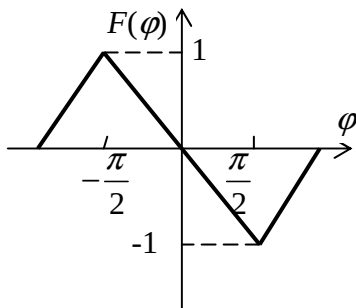


Рис. 4.3. Характеристика фазового детектора

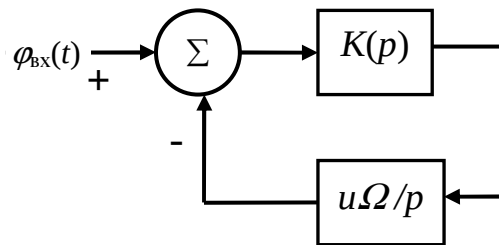


Рис. 4.4. Линеаризованная модель системы ФАПЧ

Анализ этого выражения позволяет сделать вывод, что передаточная функция $H(p)$ определяется как внутренними параметрами ФАПЧ (Ω , $K(p)$), так и внешними (u). Зависимость $H(p)$ от u нежелательна, и для ее устранения или ослабления применяют автоматическую регулировку усиления (АРУ). В дальнейшем будем полагать $u = 1$. Выбор оптимальной формы передаточной характеристики $K(p)$ и коэффициента Ω является основной задачей при проектировании ФАПЧ. Для определения $K(p)$ достаточно задать $H(p)$ и затем через нее выразить $K(p)$

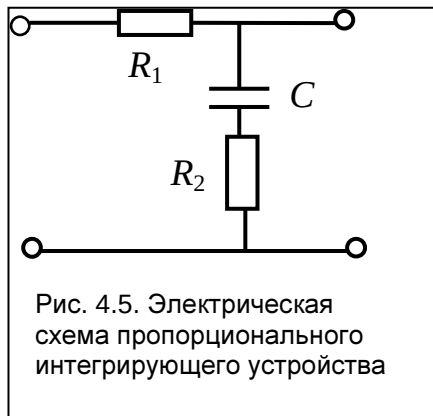
$$K(p) = \frac{pH(p)}{(2/\pi)\Omega[1 - H(p)]}.$$

Так как ФАПЧ в УС выполняет роль полосового фильтра с частотной характеристикой $H(j\omega)$, то важнейшим показателем качества ее работы является шумовая

полоса $F_{ш} = \frac{1}{2\pi} \int_0^{\infty} |H(j\omega)|^2 d\omega$. Выбрав соответствующим образом $K(p)$ и Ω , можно

получить требуемое значение шумовой полосы. Однако при этом необходимо помнить, что фильтр, реализующий характеристику $H(p)$, должен быть устойчивым. Стремление сузить шумовую полосу приводит к уменьшению полосы захвата. Доказано, что полоса захвата в системе ФАПЧ с любым ФНЧ не может быть больше шумовой полосы [9]. Уменьшение шумовой полосы влечет за собой также увеличение времени ввода в синхронизм $T_{сх}$. Для ФАПЧ без ФНЧ шумовая полоса, полоса захвата и полоса удержания равны между собой

$$F_{ш} = F_3 = F_y = \frac{1}{2\pi} \int_0^{\infty} \frac{(2\Omega/\pi)^2}{(2\Omega/\pi)^2 + \omega^2} d\omega = \frac{\Omega}{2\pi}.$$



Используя пропорционально интегрирующий фильтр (рис. 4.5), у которого

$$F_{ш} = \frac{\Omega}{2\pi} \cdot \frac{\pi + 2\Omega T_1^2 / T_2}{\pi + 2\Omega T_1},$$

где $T_1 = R_2 C$, $T_2 = (R_1 + R_2) C$, можно получить достаточно узкую шумовую полосу, если соответствующим образом выбрать параметры R_1 , R_2 , C . Время ввода в синхронизм в этом случае $T_{сх} \approx 3,5 \Delta f^2 / F_{ш}^3$, где Δf – расстройка между частотами опорного и управляемого генераторов.

Противоречие между улучшением фильтрующих свойств ФАПЧ и уменьшением полосы захвата может быть разрешено, если применить дополнительные поисковые процедуры или изменить параметры ФНЧ (адаптация) на этапе ввода в синхронизм. В первом случае на вход УГ следует подать пилообразное напряжение, которое изменяет его частоту. В определенный момент разность частот входного и опорного сигналов оказывается такой, что происходит захват и система входит в синхронизм. Амплитуда пилообразного напряжения должна обеспечивать перестройку генератора в требуемом диапазоне частот, а его период должен быть достаточен для завершения переходных процессов в системе ФАПЧ. При этом методе можно время поиска $T_{п}$ ориентировочно оценить по формуле $T_{п} \approx \pi \Delta f / F_{ш}^2$. Поисковая процедура ввода в синхронизм особенно эффективна, когда $\Delta f \gg F_{ш}$.

Добиться уменьшения $T_{сх}$ можно также изменением параметров ФАПЧ Ω и $K(p)$. На этапе ввода в синхронизм можно расширить полосу пропускания ФАПЧ или увеличить коэффициент передачи Ω . После того как поиск будет завершен, эти параметры необходимо сделать такими, чтобы обеспечивалось нужное качество слежения. Выбор оптимального режима требует моделирования системы ФАПЧ.

4.3.2. Устройства фазовой синхронизации

Системы передачи дискретной информации, в приемном тракте которых для демодуляции сигнала используется когерентное опорное колебание на несущей частоте, получили название *когерентных*. В качестве демодуляторов в них применяется согласованный фильтр или коррелятор, а основной вид модуляции – фазовый. Опорное напряжение несущей частоты формируется в устройстве фазовой синхронизации (УФС), реализуемом, как правило, на базе ФАПЧ.

Спектр сигнала $A_0 \sin[\omega_0 t + \{x_i\}\psi + \varphi(t)]$ при равновероятной передаче символов x_i и девиации фазы $\psi = 2\pi/m$ не содержит дискретной составляющей на частоте ω_0 . Один из способов ее восстановления основан на том, что в спектре передаваемого радиосигнала

оставляется составляющая достаточной мощности на несущей частоте. Для этого девиацию делают равной $(2\pi - \varphi)/m$. Другой способ предусматривает использование сигнала с полностью подавленной несущей. Чтобы при этом восстановить несущую, сигнал подвергают нелинейным преобразованиям. Рассмотрим способы ФС и оценим их целесообразность в тех или иных конкретных ситуациях.

Спектр сигнала с неполностью подавленной несущей состоит из дискретной и непрерывной составляющих, т. е. смешанный. Для двоичной ФМ сигнала дискретная составляющая на частоте ω_0 определяется как $A_0^2 \cos^2 \varphi / 2$. Непрерывная часть спектра связана со случайным чередованием нулей и единиц в сообщении и определяется как

$$G(\omega) = 2A_0^2 T_c \sin^2 \left(\frac{\psi}{2} \right) \frac{\sin^2 [(\omega - \omega_0) T_c / 2]}{[(\omega - \omega_0) T_c / 2]}.$$

Для двоичной ФМ с произвольной девиацией фазы φ при идеальном канале синхронизации вероятность ошибки

$$p_{ou} = 1 - F \left[\sqrt{h^2 (1 - r_{12})} \right], \quad (4.3)$$

где $r_{12} = \frac{1}{E_c} \int_0^{T_c} s_1(t) s_2(t) dt = \cos(\psi)$ - коэффициент взаимной корреляции сигналов s_1 и s_2 .

Из (4.3) следует, что вероятность ошибки оказывается минимальной при $\varphi = \pi$. Однако в этом случае в спектре сигнала будет отсутствовать дискретная составляющая на частоте ω_0 . Соотношение между мощностью дискретной составляющей и мощностью непрерывной составляющей зависит от девиации фазы φ . Изменяя значение φ , можно добиться оптимального режима работы системы передачи в целом. Под оптимальным режимом следует понимать такой, для которого при прочих равных условиях достигается минимальное значение вероятности ошибки. В канале синхронизации дисперсия флуктуации фазы

$$\sigma_\varphi^2 = \frac{F_w [N_0 + G(f_0)]}{P_c \cos^2(\psi/2)} = \frac{N_0 F_w}{P_c \cos^2(\psi/2)} + F_w T_c \operatorname{tg}^2 \left(\frac{\psi}{2} \right)$$

Влиянием ошибок синхронизации на $\bar{p}_{ou}(\Delta\varphi)$, как ранее указывалось, можно пренебречь при $\sigma_\varphi \leq (0,2 \dots 0,3)$. Поэтому, если допустить, что на передачу информации отводится большая часть мощности сигнала, то рассматриваемый способ формирования синхросигнала можно применять при $Q = (1 / F_w T_c) \geq 10$. Это означает, что полоса пропускания фильтра F_w в канале синхронизации должна быть существенно меньше ширины спектра информационного сигнала. В противном случае помеха, создаваемая ФМ сигналом, делает прием неэффективным. Оптимальное соотношение между мощностью синхросигнала P_{cc} и полной мощностью сигнала P_c с учетом ошибок, вносимых СС, можно оценить по формуле [9]

$$\left(\frac{P_{cc}}{P_c} \right)_{опт} \approx \frac{1}{\left(h^2 - \frac{1}{4} \right) \left\{ \left[1 + \frac{h^2 (Q+2)}{\left(h^2 - \frac{1}{4} \right)^2} \right]^2 - 1 \right\}}$$

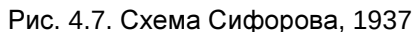
Заметим, что когда Q велико по сравнению с h^2 ,

$$(P_{cc} / P_c)_{опт} \approx \frac{1}{\sqrt{h^2 Q}}.$$

- с квадратичной нелинейной обработкой сигнала (схемы Пистолькорса, Сифорова);
- ФАПЧ с квадратурными каналами (схема Костаса);
- с обратной связью по решению.

- Для схемы с квадратичной обработкой (рис. 4.6, 4.7) получаемый на выходе нелинейного элемента сигнал будет содержать гармонику частоты $2f_0$, которую можно отфильтровать полосовым фильтром (схема Пистолькорса) или ФАПЧ (схема Сифорова). Разделив эту частоту на 2, получим частоту f_0 . Так как во всяком нелинейном элементе при малом отношении сигнал/помеха помеха подавляет сигнал, то перед схемой возведения в квадрат желательно включить фильтр.

Рис. 4.6. Схема Пистолькорса, 1933



4.4. Тактовая синхронизация

86

посылок. Тактовые импульсы (ТИ), временное положение которых совпадает с моментами окончания посылок, управляют работой интеграторов при корреляционной обработке сигнала или используются для снятия отсчета напряжения с выхода согласованных фильтров. При неоптимальном приеме ТИ используются при регенерации посылок. Помимо этого, ТС необходима тогда, когда квазисинхронные потоки символов разных источников объединяются в один поток.

Поскольку при случайном характере передаваемой информации, спектр радиосигнала сплошной и расположен в области несущей частоты, то он не содержит составляющей тактовой частоты. Поэтому для обеспечения ТС сигнал должен быть соответствующим образом обработан. Необходимо отметить, что информацию о тактовой частоте в СПИ с простыми сигналами можно выделить только из сигнала, в котором модулирующие посылки меняют свое значение. Сигнал, модулированный посылкой одного знака, информации о тактовой частоте не несет. Чтобы предотвратить появление длинных последовательностей одного знака, часто используют специальные устройства рандомизации потока. Например, в кодере СПИ с ОФМ выходные символы y_i , связаны с входными x_i соотношением $y_i = y_{i-1} \oplus x_i$, при этом последовательности одного знака преобразуются в меандр.

При создании устройств тактовой синхронизации (УТС) необходимо найти алгоритм, обеспечивающий наилучшую (в смысле выбранного критерия) оценку временного положения сигнала. Из теории оценок известно, что эта задача сводится к определению максимума функции правдоподобия $\Lambda(\tau)$. Максимум функции $\Lambda(\tau)$ можно найти устройством с параллельным анализом на интервале неопределенности $(0, T_c)$ или с последовательным (следающие УТС). Первый тип устройств позволяет определять τ за минимальное время, однако из-за сложности реализации применяется редко.

В следящих УТС в произвольной точке вычисляется значение производной функции правдоподобия $\left. \frac{\partial \Lambda(t)}{\partial \tau} \right|_{\tau=\tau_i}$ (иногда дополнительно значение $\Lambda(\tau_i)$), а затем по

этому значению в решающем устройстве оценивается наиболее вероятное положение максимума $\Lambda(\tau_i) = \max$. Следующее вычисление производится в точке, которая позволяет оценивать положение максимума с наибольшей достоверностью. Ею могла бы являться координата максимума функции $\Lambda(\tau)$. Однако система слежения в этой точке оказывается нечувствительной к изменению временного положения входного сигнала, так как

$\left. \frac{\partial \Lambda(t)}{\partial \tau} \right|_{\tau=\tau_i} = 0$. Поэтому целесообразно следить за точкой, где производная $\partial \Lambda(\tau) / \partial \tau$ и

значение $\Lambda(\tau)$ достаточно большие. Если передаваемый сигнал известен, то определение $\Lambda(\tau)$ заключается в нахождении модуля функции взаимной корреляции принимаемого сигнала и опорного. В СПИ это принципиально невозможно, так как передаваемая информация носит случайный характер. При этом оптимальный алгоритм вычисления $\Lambda(\tau)$ оказывается слишком сложным и его целесообразно применять лишь для получения оценок потенциально достижимой точности измерения фазы.

На практике используют квазиоптимальные алгоритмы, реализуемые на базе демодулятора посылок без синхронизации. Сигналы с выходов согласованных фильтров детектируются и их разность затем подается на решающую схему (РС). Момент смены знака содержит информацию о фазе тактовой частоты. В качестве примера рассмотрим работу демодулятора двоичных ЧМ сигналов (рис. 4.9 а). В отсутствие шумов сформированные импульсы ТС имеют постоянный временной сдвиг $T_c / 2$ относительно тактовых импульсов посылок (рис. 4.9 б). При действии шумов их временное положение изменяется. Дисперсия флуктуации определяется отношением h^2 , видом модуляции и способом обработки. При когерентной обработке амплитуда сигнала в точке 1 при

отсутствии шума равна E , а флуктуации имеют гауссовский закон распределения $w(u)$ с дисперсией EN_0 . Поэтому, как это следует из рис. 4.9 в, при больших значениях $h^2 = E/N_0$ флуктуации фазы сформированных импульсов ТС также будут подчиняться гауссовскому закону $w(x)$ с дисперсией $\sigma_t^2 \approx T_c^2 / h^2$.

Для повышения точности окончательную оценку фазы тактовой частоты производят по ряду измерений в следящем фильтре. Для этого во временном (фазовом) дискриминаторе ВД сравнивают последовательности сформированных и опорных импульсов. Напряжение с выхода дискриминатора определяется разностью фаз. Опорный генератор управляется по фазе так, чтобы свести это рассогласование к минимуму.

По своей структуре, алгоритму работы и характеристикам следящее УТС подобно ФАПЧ. Принципиальное отличие здесь заключается в форме фазовой характеристики и во входном сигнале, представляющем собой случайную последовательность импульсов, статистические характеристики которой определяются передаваемой информацией.

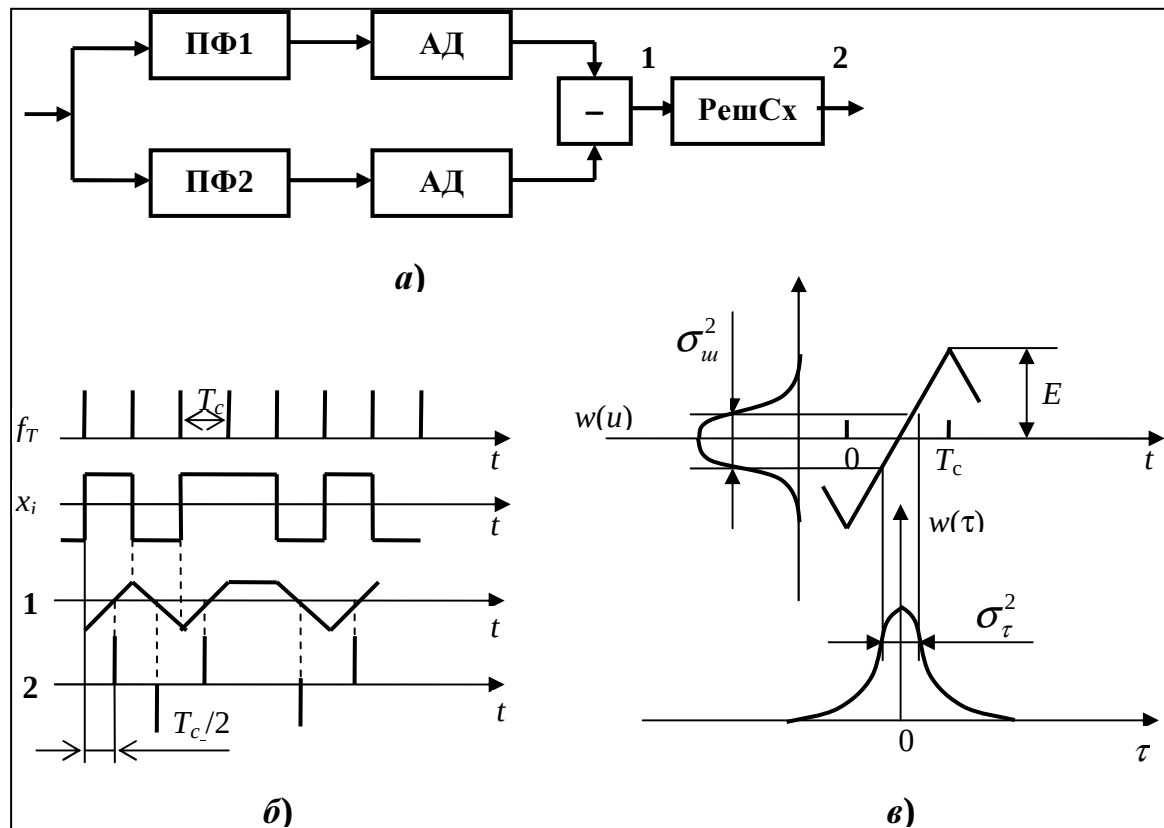


Рис. 4.9. а) – Структурная схема демодулятора БЧМ; б) – формирование тактовых импульсов; в) – определение дисперсии фазы импульсов

Существуют различные способы практической реализации следящих УТС, из которых наиболее распространены устройства с дискретным управлением (рис. 4.10). Последние позволяют получать высокую точность слежения и могут быть реализованы на современной элементной базе с цифровыми методами обработки сигналов. Принцип работы УТС с дискретным управлением основан на смещении фазы сигнала, формируемого управляемым делителем частоты УДЧ при добавлении или исключении одного импульса на его входе. Точность управления фазой (шаг подстройки) определяется коэффициентом деления $K_{\text{дел}}$. В качестве фильтра в цепи обратной связи используется интегратор, выполненный на основе реверсивного счетчика. В реверсивном счетчике подсчитывается разность числа импульсов, поступивших с выходов ВД 1 и 2. Если она превышает емкость счетчика $N_{\text{сч}}$, то на соответствующем выходе формируется команда, которая в УДЧ приводит к смещению фазы опорного сигнала в требуемую

сторону. Емкость реверсивного счетчика $N_{сч}$ определяет число импульсов, по которому выносятся оценка о знаке рассогласования, и соответственно помехоустойчивость УТС.

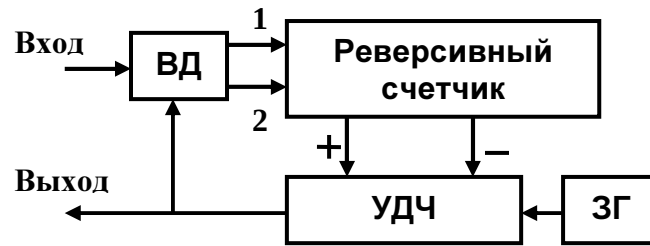


Рис. 4.10. Структурная схема устройства тактовой синхронизации с дискретным управлением

При проектировании дискретного УТС важно правильно выбрать значения $N_{сч}$ и $K_{дел}$. Зная $K_{дел}$ и длительность посылки T_c , можно определить шаг подстройки $\Delta t = T_c / K_{дел}$. Время, через которое произойдет коррекция фазы опорного сигнала, $T_k = 2N_{сч}T_c$, где коэффициент 2 учитывает случайный характер появления импульсов на входе ВД. Устойчивую работу УТС можно обеспечить только тогда, когда шаг подстройки $T_{кор}$ достаточен для компенсации временной расстройки ΔT , возникающей за время между двумя подстройками. Причиной появления этой расстройки является нестабильность тактовой частоты -

$$K_f = \frac{\Delta f_T}{f_T} = \frac{\Delta T}{T_c}$$

. Коэффициент K_f характеризует скорость «скольжения» частот. Следовательно, предельное значение допустимой относительной нестабильности частот $K_f < \Delta t / T_k$. Это неравенство является условием синхронизма и позволяет связать основные параметры УТС между собой

$$K_f \leq \frac{2}{K_{дел} N_{сч}} = \frac{2\Delta t}{T_c N_{сч}}.$$

Важный параметр УТС - время ввода в синхронизм $T_{сх}$, поскольку оно определяет длительность вхождения в связь. При расчете $T_{сх}$ будем считать, что $K_f = 0$, а временной сдвиг наибольший и равен $T_0/2$. Тогда получим

$$T_{сх} = \frac{T_c}{2\Delta t} T_k = K_{дел} N_{сч} T_c.$$

4.5. Цикловая синхронизация

Устройства цикловой синхронизации (УЦС) предназначены для определения начала кодовых слов. Поскольку при передаче сообщения безыбыточным кодом последовательность символов случайна и информации о начале и конце кодовых слов не несет, то предпринимают специальные меры для внесения этой информации. Цикловую синхронизацию обеспечивают либо с помощью специальных синхросигналов, либо с помощью внутренней избыточности кодовых слов. В обоих случаях цикловая синхронизация реализуется за счет снижения скорости передачи информации.

В качестве циклового синхросигнала можно использовать периодически повторяющиеся от слова к слову сосредоточенные или распределенные синхрогруппы. На приемной стороне синхросигнал, генерируемый местным генератором, сравнивается со входной последовательностью символов при различных взаимных временных положениях. Совпадение символов принимаемого и опорного сигналов фиксируется как режим синхронизма. Этот режим можно обнаружить, просматривая все временные позиции одновременно или последовательно. Выбор того или иного метода передачи и обработки цикловой синхроинформации при реализации УЦС определяется

необходимостью быстрого обеспечения синхронизма, высокой помехоустойчивости, минимального объема синхроинойформации в цикле, а также простоты реализации УЦС.

Наиболее просто реализуется ЦС при передаче односимвольного синхросигнала в начале каждого кодового слова или группы кодовых слов (рис. 4.11). Импульсы с частотой $f_{Ц}$ (частотой следования слов), формируемые с помощью счетчика Сч, подаются на схемы совпадения СС, на другие входы которых поступают кодовые символы с выхода регенератора посылок. В зависимости от знака этих символов на реверсивный счетчик РСч поступают импульсы по одному из двух входов. Счетчик импульсов интервала анализа СчИИА определяет отрезок времени длиной в несколько слов, через который число, записанное в реверсивный счетчик, сравнивается с порогом в решающей схеме РС. Если порог не превышен, то в решающей схеме формируется строб, которым с помощью схемы запрета СЗ вычеркивается один из подаваемых на счетчик тактовых импульсов (частоты f_T) и точка анализа синхронизации в кодовом слове смещается на один символ.

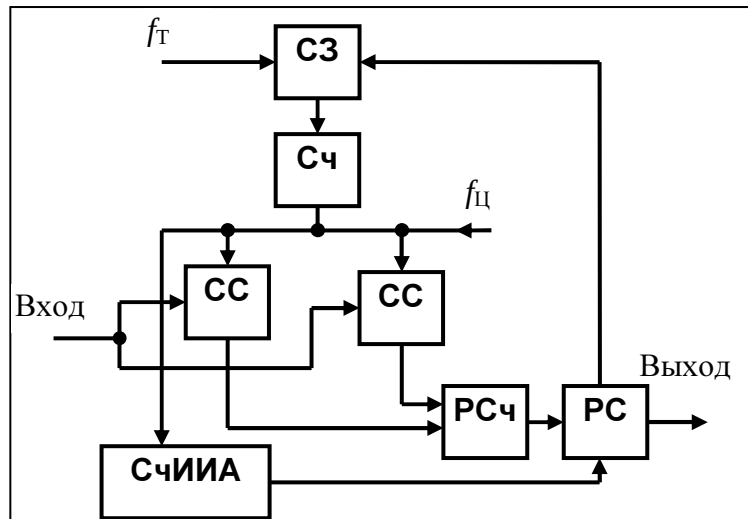


Рис. 4.11. Структурная схема устройства цикловой синхронизации, реализующего шаговый поиск односимвольного синхросигнала

Таким образом, в зависимости от принимаемого решения устройство цикловой синхронизации находится либо в режиме поиска синхронизации, либо в установившемся режиме контроля за появлением символов синхросигнала на синхропозиции цикла.

Важной характеристикой УЦС является среднее время установления синхронизма $\bar{T}_л$ (время поиска). Процесс поиска позиции, на которой передается синхросигнал, продолжается до тех пор, пока единица не повторится на проверяемой позиции требуемое число раз / на интервале анализа в g циклов. Величины l и g , по которым принимается решение о наличии циклового синхронизма, зависят от вероятности появления символа 1 на информационных позициях цикла, вероятности искажения синхросимволов $p_{ош}$ и заданных характеристик принятия решения. Используя известные формулы для обнаружения двоичных сигналов, можно определять вероятности ложного синхронизма $p_{лс}$ и пропуска синхронизма $p_{пс}$:

$$p_{лс} = n \sum_{i=1}^g C_g^i p^i (1-p)^{g-1}, \quad (4.4, 4.5)$$

$$p_{пс} = 1 - \sum_{i=1}^g C_g^i (1-p_{ош})^i p_{ош}^{g-1},$$

где n – число символов в кодовом слове.

Для равноправного появления единичных и нулевых символов на информационных позициях цикла $p_{лс} = \frac{n}{2^n} \sum_{i=1}^g C_g^i$. Если значения $p_{лс}$ и $p_{пс}$ заданы, то, решая систему уравнений, можно определить минимальное значение $g_{\min}$ и соответствующий ему оптимальный порог $l_{\text{опт}}$, а затем вычислить среднее время поиска

$$\bar{T}_{\Pi} = \frac{(n-1)^2}{2} T_c g. \quad (4.6)$$

Систему уравнений (4.4) и (4.5) можно решить только численными методами. Определение характеристик УЦС существенно упрощается, если заданы $\bar{T}_{\Pi}$ и $p_{пс}$. Тогда по формулам (4.6) и (4.4) можно определить требуемый объем выборки g и порог l , а затем из формулы (4.5) найти вероятность пропуска синхронизма $p_{пс}$.

Основой для построения УЦС, использующих синхронизирующие свойства кодов, является то, что при отсутствии цикловой синхронизации вероятность появления обнаруживаемой избыточным (помехоустойчивым) кодом ошибки значительно больше, чем при синфазной работе. Если на приемной стороне СПИ декодер, проверяя правильность поступающих на него кодовых слов, установит, что число искаженных кодовых слов превысит пороговое значение, то управляющее устройство УЦС переключится в режим поиска циклового синхронизма. В этом режиме управляющее устройство дискретно (на один такт за каждый цикл) изменяет момент начала записи в декодер принятого кодового слова. Как только слово будет записано правильно (от начала до конца), обнаружение ошибок прекращается и на выходе декодера сформируется импульс, блокирующий работу управляющего устройства. Если число правильно принятых слов превысит соответствующее пороговое значение, то УЦС выйдет из режима поиска.

В дальнейшем наличие цикловой синхронизации будет проверяться по правильности приема информации.

Подобный способ фазирования можно реализовать в СПИ, где для обнаружения ошибок используется код, пригодный для синхронизации. Пригодным кодом для синхронизации считается такой, у которого вероятность появления разрешенной комбинации в последовательности из символов, входящих в два соседних кодовых слова (пересечение двух слов), очень мала. Например, при передаче сообщений вида

$$\begin{array}{c} \text{1-е слово} \quad b_1^{(i)}, \dots, b_n^{(i)}, b_1^{(i+1)}, b_2^{(i+1)}, \dots, b_j^{(i+1)} \\ \underbrace{\hspace{10em}}_{\text{пересечение двух слов}} \\ b_1^{(i)}, b_2^{(i)}, \dots, b_j^{(i)}, \dots, b_n^{(i)}, b_1^{(i+1)}, b_2^{(i+1)}, \dots, b_j^{(i+1)}, \dots, b_n^{(i+1)} \end{array}$$

вероятность появления разрешенной комбинации в последовательности символов $b_1^{(i)}, \dots, b_n^{(i)}, b_1^{(i+1)}, b_2^{(i+1)}, \dots, b_j^{(i+1)}$ должна быть очень малой.

Код, у которого эта вероятность равна нулю, называется кодом «без запятой». Ограничения при использовании рассмотренного метода цикловой синхронизации связаны с большой избыточностью самосинхронизирующихся кодов.

4.6. Кадровая синхронизация

В многоканальных РСПИ с временным уплотнением каналов необходимо передавать информацию о начале и конце кадра. Это обеспечивается устройством кадровой синхронизации (УКС). При наличии цикловой синхронизации для синхронизации кадров используется одно определенное кодовое слово, передаваемое в начале кадра. Кадровое синхрослово (КСС) по своей структуре должно существенно отличаться от всех возможных кодовых комбинаций, образуемых при передаче

дискретной информации, и обеспечивать наилучшие условия его поиска и обнаружения в информационной последовательности символов даже при наличии искажений принимаемых посылок. Если цикловой синхронизации нет, то неопределенность возрастает и тогда для обеспечения этого условия под КСС должна выделяться довольно значительная часть кадра (иногда до десяти и более процентов от общего числа символов в кадре). Это объясняется тем, что к помехоустойчивости кадровой синхронизации многоканальных СПИ предъявляются высокие требования, так как нарушение синхронизма по кадрам влечет за собой потерю связи во всех каналах системы.

Для выделения КСС в приемнике СПИ используется дискретный согласованный фильтр (ДСФ), настроенный на КСС. В него поочередно записываются принимаемые кодовые слова и в момент превышения выходным напряжением порога выделяется импульс кадровой синхронизации. Используя повторяемость импульсов синхронизации кадров, можно, накапливая их, увеличить помехоустойчивость УКС.

Во избежание ложного выделения КСС из информационной последовательности посылок необходимо предъявлять определенные требования к структуре и регулярности его повторения. Желательно, чтобы КСС или вообще в информационном сигнале не встречалось, или его появление в нем было маловероятным. Если информационный сигнал кодируется безызбыточным кодом и источник информации может выдавать все кодовые слова с равной вероятностью, то единственным отличием КСС от информационного кодового слова может быть регулярность его появления на одних и тех же позициях кадра. Когда повторение одних и тех же кодовых слов в информационном сигнале имеет высокую вероятность, синхросигнал образуют путем передачи двух чередующихся КСС. Например, в одном кадре посылается заданное КСС, а во втором инверсное ему, в третьем кадре снова повторяется первоначальное КСС и т. д. Инверсное КСС образуется путем замены всех единиц нулями, а нулей единицами.

Перспективным методом кадровой синхронизации является метод, при котором специальные кодовые группы используются для синхронизации как кадров, так и слов. В этом случае для синхронизации СПИ тратится меньшее число посылок в кадре. Однако требования к длине и, главным образом, к структуре КСС существенно ужесточаются. Действительно, в рассматриваемой ситуации УКС должно выделить КСС на основе анализа всей поступающей на его вход последовательности символов, в которой теперь уже не известны границы отдельных кодовых слов.

Для уменьшения времени установления кадровой (и одновременно цикловой) синхронизации необходимо, чтобы вероятность ложного появления КСС в принимаемой последовательности посылок была минимальна. Вероятность ложного фазирования определяется длиной (числом разрядов) и структурой КСС, а также числом информационных посылок в передаваемом сообщении.

Определим основные требования к длине и структуре КСС, если цикловой синхронизации нет. Предположим, что КСС состоит из n разрядов $a_1, a_2, \dots, a_n$, информационная же последовательность имеет вид $b_1, b_2, \dots, b_n$. Тогда анализируемая кодовая комбинация будет иметь вид

$$a_1, a_2, \dots, a_n, b_1^{(i)}, b_2^{(i)}, \dots, b_n^{(i)}, b_1^{(i+1)}, b_2^{(i+1)}, \dots,$$

где $b_j^{(i)}$ и $b_j^{(i+1)}$ - информационные разряды соответствующих кодовых комбинаций.

В процессе скользящего посимвольного поиска ложное выделение КСС может быть как из информационной последовательности $b_1^{(i)}, b_2^{(i)}, \dots, b_n^{(i)}$, так и из последовательности, полученной на пересечении КСС и информационного слова $a_j, a_{j+1}, b_1^{(i)}, b_2^{(i)}, \dots, b_n^{(i)}$. Вероятность ложного выделения КСС из последовательности информационных посылок определяется длиной КСС и числом информационных посылок, а на пересечении КСС и информационного слова определяется также структурой КСС. Специальным выбором структуры КСС можно добиться того, чтобы эта вероятность была достаточно малой.

Кодовые последовательности, удовлетворяющие этому требованию, должны иметь корреляционные функции с низким уровнем боковых выбросов. В качестве КСС в настоящее время широко используются М-последовательности, формируемые генераторами, включающими в себя сдвигающие регистры, охваченные логической обратной связью.

Как было сказано, вероятность выделения КСС из информационной последовательности определяется числом посылок в КСС n и числом слов в кадре k_{cl} . Определим требуемую длину КСС n в зависимости от k_{cl} при использовании в качестве обнаружителя дешифратора. Поскольку вероятность появления 0 и 1 на любой позиции одинакова и равна 0,5, то при длине КСС в n посылок

$$p_{nc}(1) = 0,5^n.$$

Вероятность ложной синхронизации и вероятность пропуска синхросигнала при просмотре кадра соответственно равны:

$$p_{лс} = 1 - [1 - p_{пс}(1)]^{k_{cl}} \approx k_{cl} / 2^n,$$

$$p_{пс} = 1 - (1 - p_{ош})^n \approx np_{ош}.$$

Ложное фазирование приводит к ошибочному приему всей информации, поэтому с точки зрения уменьшения вероятности ошибки за счет фазирования значение n должно быть как можно больше. Однако при длинном КСС увеличивается избыточность кадра а, следовательно, ухудшается использование пропускной способности СПИ. Поэтому необходимое значение n следует выбирать с учетом двух важных характеристик: требуемой достоверности и эффективности использования пропускной способности системы.

При достаточной длине КСС, когда требуемая вероятность $p_{лс}$ обеспечивается с запасом, можно снизить порог l в решающей схеме, т. е. отказаться от дешифратора. В этом случае характеристики обнаружения можно рассчитать по формулам, аналогичным (4.4) и (4.5):

$$P_{лс} = \frac{k_{cl}}{2^2} \sum_{i=1}^n C_n^i,$$

$$p_{пс} = 1 - \sum_{i=1}^n C_n^i (1 - p_{ош})^i p_{ош}^{n-i}.$$

При больших значениях n пользоваться этими формулами трудно. Поэтому, переходя при $n \rightarrow \infty$ от биномиального закона к гауссовскому, можно записать

$$p_{лс} = 1 - F(1/\sqrt{n}),$$

$$p_{пс} = 1 - f\left\{[n(1 - 2p_{ош}) - 1] / 2\sqrt{np_{ош}(1 - p_{ош})}\right\}.$$

В ряде СПИ цикловую и кадровую синхронизацию обеспечивают передачей в начале сообщения специального синхросигнала, называемого *командой фазового пуска*. Такой режим оправдан, когда длительность передаваемого сообщения ограничена и потеря даже части сообщения недопустима. Поскольку в точке приема сведения о начале сообщения могут быть весьма ограниченными, то для его надежного определения команда фазового пуска должна быть достаточно продолжительной. Это требование противоречит простоте реализации устройства обработки сигнала фазового пуска, память которого по крайней мере не может быть меньше числа символов сигнала. Выход здесь находят в использовании специальных сигналов и устройств их обработки. Если каждый символ команды фазового пуска связан с группой предыдущих рекуррентным правилом, то, приняв правильно группу символов, можно воссоздать весь сигнал и соответственно определить его начало и конец. Генерируя таким образом сигнал, на приемной стороне с помощью корреляционного приемника можно с высокой достоверностью определить его совпадение по фазе с принимаемым сигналом и тем самым перепроверить первоначально

принятое по группе символов решение. В качестве рекуррентных последовательностей применяют М-последовательности, а сам метод назван методом *последовательной оценки*.

4.7. Синхронизация в системах с широкополосными сигналами

Оптимальный алгоритм обработки широкополосных сигналов (ШПС) при синхронизации не имеет каких-то принципиальных отличий от алгоритма обработки простых сигналов. Для оценки времени запаздывания следует использовать метод максимального правдоподобия. Если помеха представляет собой гауссовский случайный процесс с равномерной спектральной плотностью, то максимум функции правдоподобия совпадает с максимумом модуля взаимной корреляционной функции принимаемого и опорного сигналов $R(\tau)$. Если для передачи символов сообщения используют сигналы разной формы, то число каналов вычисления взаимной корреляционной функции должно быть равно основанию кода.

При наличии помех положение максимума ВКФ становится случайным и его определение сопровождается ошибками. Ошибки могут быть двух видов: нормальными, когда их значение не превышает ширины пика взаимной корреляционной функции, и аномальными в случае превышения. Аномальные ошибки, при которых прием информации корреляционным приемником ШПС вообще невозможен, вносят принципиальное отличие характеристик и алгоритма работы системы тактовой синхронизации ШПС от системы тактовой синхронизации простых сигналов. Погрешности оценки временного положения принимаемого сигнала определяются его корреляционной функцией и отношением сигнал/шум на выходе СФ. Идеальными сигналами для синхронизации следует считать такие, у которых боковые выбросы корреляционной функции невелики, например периодические М-последовательности.

Оценим приближенно, как влияют нормальные и аномальные ошибки на точность синхронизации сложных сигналов с идеальными корреляционными функциями. Временное положение выходного сигнала СФ с наибольшей достоверностью можно оценить по положению точки, где крутизна $R(\tau)$ максимальна (рис. 4.12). Для сигналов с треугольной формой корреляционной функции крутизна постоянна и равна EB/T_c , где B - база сложного сигнала. Поэтому координата этой точки выбрана равной $E/2$. Флуктуации шума с законом распределения $w(g_{ш})$ и дисперсией $\sigma_{ш}^2 = EN_0/2$ приведут к погрешностям измерения временного положения с дисперсией σ_τ^2

$$\sigma_\tau^2 = T_c^2 / 2h^2 B^2 \quad (4.7)$$

Этот результат практически совпадает с оценками, получаемыми при точном анализе.

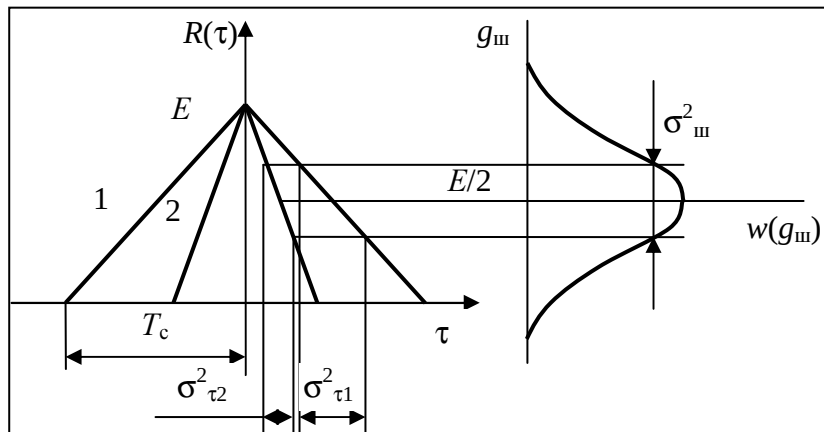


Рис.4.12. Принцип оценки точности определения временного положения простого (1) и сложного (2) сигналов

Теперь оценим аномальные ошибки. На интервале T_c выходное напряжение СФ можно представить во времени B независимыми отсчетами, один из которых принадлежит пику корреляционной функции, а остальные $(B - 1)$ - шуму. Тогда для сигналов с идеальной формой КФ при вынесении решения о наличии пика в каждой позиции вероятность ложного синхронизма $p_{лс}(1)$ численно равна вероятности ошибки при различении двух ортогональных сигналов. Так как аномальные ошибки могут возникнуть независимо в любой из $(B - 1)$ точек, то их результирующая вероятность

$$p_{ан} = 1 - [1 - p_{лс}(1)]^{B-1} \approx (B - 1)p_{лс}(1),$$

а дисперсия

$$\sigma_{ан}^2 = T_c^2 / 12.$$

Дисперсия ошибок синхронизации с учетом нормальных и аномальных ошибок

$$\sigma_{\Sigma}^2 = p_{ан} \sigma_{ан}^2 + (1 - p_{ан}) \sigma_{\tau}^2 \quad (4.8)$$

Подставив в (4.8) выражение для вероятности аномальных ошибок

$$p_{ан} \approx \frac{B-1}{2} \exp\left(-\frac{h^2}{2}\right)$$

и приняв во внимание, что в реальных ситуациях $p_{ан} \ll 1$, окончательно получим

$$\sigma_{\Sigma}^2 \approx T_c^2 \left[\frac{B-1}{24} \exp\left(-\frac{h^2}{2}\right) + \frac{1}{2h^2 B^2} \right].$$

С увеличением базы сигнала B нормальные ошибки уменьшаются, а аномальные растут. Нетрудно найти значение $B_{опт}$, при котором достигается минимум дисперсии ошибки

$$B_{опт} = \sqrt[3]{\frac{24}{h^2} \exp\left(\frac{h^2}{2}\right)}$$

Например, для $h^2 = 2$ база $B_{опт} = 3$, а для $h^2 = 24$ база $B_{опт} = 50$.

Когда сложные сигналы обрабатываются в корреляторах, синхронизацию выполняют в два этапа. Сначала, изменяя временное положение опорного сигнала, последовательно просматривают область неопределенности и грубо находят положение центрального пика взаимной корреляционной функции (этап поиска), а затем следящим устройством уточняют положение его максимума (этап слежения). Центральный пик необходимо искать с шагом Δt , не превышающим $0,5/F_c$. При этом исключается возможность его пропуска в отсутствие помех. Общее время синхронизации складывается из времени поиска $T_{п}$ и слежения $T_{сл}$. Дисперсию ошибок слежения можно приближенно оценить по формуле (4.7), подставив $h^2 = P_{сл}/N_0$.

При неподвижных относительно друг друга приемнике и передатчике и высокой стабильности генераторов опорных частот требуемая точность σ^2 , достигается соответствующим выбором времени слежения $T_{сл}$.

Этап поиска характеризуется продолжительностью $T_{п}$ и вероятностью его правильного завершения $p_{прав} = 1 - p_{лс}$. Существующие процедуры поиска можно условно разбить на два класса: с фиксированным временем анализа на каждой позиции и с переменным. Минимальное время поиска можно получить для переменного времени анализа в каждой позиции (последовательная процедура оценки). Реализуется этот метод наиболее просто для схемы принятия решения с двумя порогами $u_{п1}$ и $u_{п2}$ ($u_{п1} < u_{п2}$). Если при анализе в течение времени T_{a1} превышен порог $u_{п1}$, но не превышен $u_{п2}$, то наблюдение повторяется. Поиск считается завершенным, если превышен порог $u_{п2}$.

Пропуск синхронизма с вероятностью $p_{лс}$ приводит к повторному просмотру области временной неопределенности, при этом продолжительность поиска оказывается случайной величиной. Среднее время поиска с учетом пропуска

$$\bar{T}_\Pi = MT_a \left(1 + 2 \sum_{i=1}^{\infty} p_{\text{лс}}^i \right) = MT_a \left(1 + \frac{2p_{\text{лс}}}{1 - p_{\text{лс}}} \right),$$

где M – среднее число шагов при поиске без пропусков.

При выборе шага $\Delta t = 0,5/F_c$ число $M \approx B$. Зная связь между вероятностью ложной синхронизации в одной точке анализа $p_{\text{лс}}(1)$, вероятностью пропуска $p_{\text{лс}}$ и временем анализа T_a

$$\frac{P_c T_a}{N_0} = \sqrt{\ln \left(\frac{1}{p_{\text{лс}}(1)} \right)} + \sqrt{\frac{1}{p_{\text{лс}}} - 1}, 4,$$

можно найти зависимость продолжительности поиска от времени анализа T_a при фиксированных значениях B и $p_{\text{лс}}(1)$. Для определения оптимального значения времени анализа целесообразно использовать численные методы или моделирование процедуры поиска. В качестве ориентировочной оценки воспользуемся результатами расчета времени поиска для процедуры с однократным просмотром области неопределенности и принятием решения в точке с максимальным значением напряжения на выходе коррелятора. В этом случае вероятность ложной синхронизации на одном шаге

$$p_{\text{лс}}(1) \approx \frac{1}{2} \exp \left(-\frac{P_c T_\Pi}{2N_0 B} \right),$$

откуда несложно найти время поиска

$$T_\Pi = \frac{4N_0 B}{P_c} \ln \left(\frac{1}{2p_{\text{лс}}(1)} \right). \quad (4.9)$$

Формула (4.9) наглядно характеризует влияние отношения сигнал/шум P_c/N_0 , вероятности ложной синхронизации $p_{\text{лс}}(1)$ и степени сложности сигнала B на продолжительность поиска T_Π .

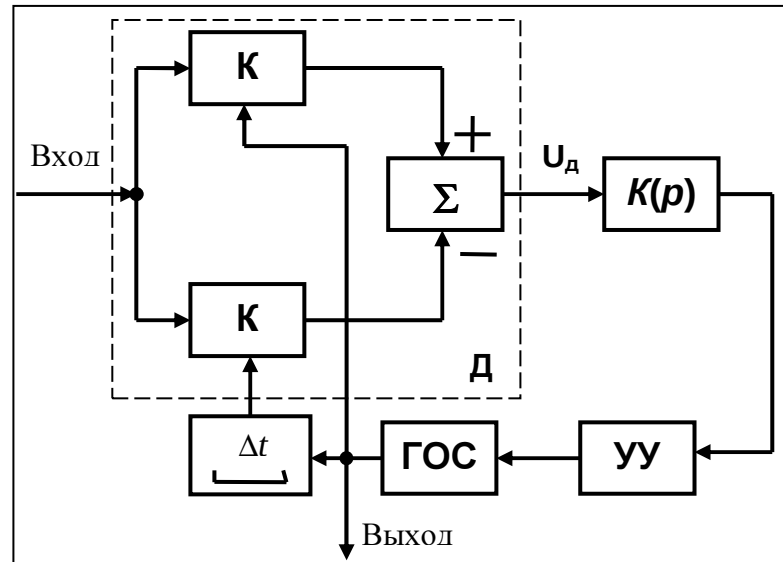


Рис.4.13. Устройство слежения за временным положением сложного сигнала

Устройство слежения (рис. 4.13) по принципу работы подобно устройству тактовой синхронизации. Оно включает в свой состав дискриминатор Δ , определяющий рассогласование по времени между входным и опорным сигналами, устройство управления временным положением опорного сигнала $УУ$ и низкочастотный фильтр $K(p)$ цепи обратной связи. Для получения дискриминационной характеристики $u_d = K\Delta\tau$ (при малых значениях $\Delta\tau$), где $\Delta\tau$ - временное рассогласование между входным и опорным сигналами, определяется разность напряжений с выходов двух корреляторов K , на

которые поданы опорные сигналы со сдвигом во времени. Временной сдвиг Δt_c чаще всего выбирают так, чтобы обеспечить максимальную крутизну дискриминационной характеристики, т. е. минимальную ошибку слежения.

5. Широкополосные сигналы в системах связи

Детерминированный сигнал, для которого произведение длительности на полосу частот $B = WT \gg 1$ и полосой которого можно управлять независимо от длительности, называется сигналом с *расширенным спектром* или *широкополосным* (ШПС). Система, использующая сигналы с расширенным спектром, является широкополосной.

Величина B , называемая *базой сигнала*, определяет важнейшие характеристики радиотехнических систем.

Задача этой главы – показать, почему и как широкополосная технология может помочь в борьбе с помехами. По этой причине мы ограничимся лишь простейшими оценками, основанными на отношении мощностей сигнала и суммы АБГШ и помехи. Такие оценки вполне адекватны для БАМ и БФМ в гауссовских каналах; что касается более сложных ситуаций, их можно принять как оценки первого приближения.

При дальнейшем рассмотрении будем иметь в виду, что в силу независимости длительности и полосы ШПС подчеркивается роль угловой модуляции (частотной или фазовой) в любых технологиях расширения спектра.

5.1. Узкополосная помеха

Данный тип помех обусловлен работой соседствующих РТС в перекрывающемся диапазоне частот, то есть помехами искусственного происхождения. На рис. 5.1 в виде прямоугольников представлены амплитудный спектр $|\tilde{S}(f)|$ полезного сигнала и спектральная плотность мощности помехи $J(f)$ на фоне равномерной спектральной плотности мощности $N_0/2$ АБГШ. Назовем помеху *узкополосной* только по той причине, что занимаемая ей полоса W_j меньше полосы W , занимаемой сигналом, и имеются области, где спектр сигнала не подвержен искажению помехой. Узкополосная помеха, кроме того, может быть определена как частично-полосная, гармоническая и т.д., однако в

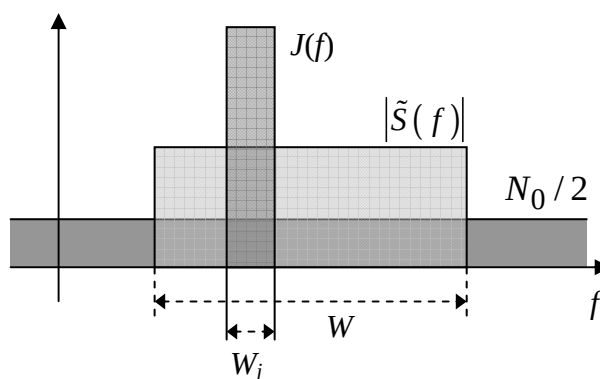


Рис. 5.1. Спектр сигнала, помехи и фонового АБГШ.

нашей ситуации надобность в подобной классификации отсутствует.

Предположим, что приемник системы всегда использует только фильтр, согласованный с АБГШ, невзирая на наличие или отсутствие помехи на входе.

При определении отношения q_I^2 мощностей сигнала и мешающего воздействия (SIR - signal-to-interference ratio) на выходе согласованного фильтра отметим тот факт, что при прямоугольном спектре сигнала ($|\tilde{S}(f)|$ равен константе $\tilde{s}$ в пределах полосы сигнала W и нулю вне ее) амплитудная передаточная функция фильтра равномерна в полосе сигнала

W и равна нулю вне ее. Без потери общности положим, что ее ненулевое значение равняется единице. Следовательно помеха, трактуемая как случайный процесс, проходит на выход фильтра без изменения своей мощности J , тогда как отфильтрованное значение мощности шума составит $N_0 W$. С другой стороны, фильтр согласован с сигналом и, значит, когерентно суммирует все гармонические составляющие сигнала, обеспечивая максимальное значение амплитуды на выходе $A_{out} = \int_{-\infty}^{\infty} |\tilde{S}(f)| df = 2W\tilde{S}$, где учтена равномерность спектра в полосе сигнала W , а удвоение обусловлено вкладом «отрицательных» частот. При аналогичных обозначениях энергия сигнала вычисляется с помощью теоремы Парсеваля как $E = \int_{-\infty}^{\infty} |\tilde{S}(f)|^2 df = 2W\tilde{S}^2$. В итоге имеем

$$q_I^2 = \frac{A_{out}^2}{J + N_0 W} = \frac{4W^2 \tilde{S}^2}{J + N_0 W} = \frac{2E}{N_0 + J/W} \quad (5.1)$$

Из последнего равенства можно увидеть, что, несмотря на конкретное значение полосы помехи W_j , отношение SIR на выходе согласованного фильтра ведет себя так, как если бы мощность помехи была равномерно распределена в полосе W сигнала (не помехи), образуя дополнительный «АБГШ» со спектральной плотностью J/W .

Рассмотрим широко используемый вариант режекции превосходящей АБГШ узкополосной помехи (рис. 5.2).

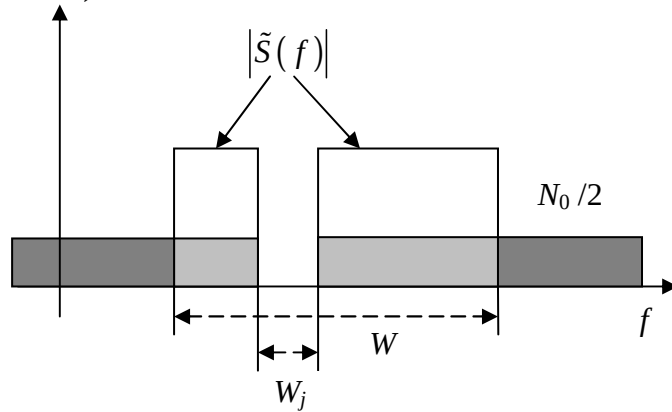


Рис. 5.2. Модель спектра на выходе режекторного фильтра

Рис. 5.2 иллюстрирует вид спектра на выходе режекторного фильтра. Наблюдается полное отсутствие помех, однако, вырезаются и частотные компоненты сигнала в пределах полосы помехи, точно также как и шум. Модель спектра может трактоваться, как образец, в котором исходный сигнал занимал бы только часть полосы W , свободную от помех, и обладал энергией, равной $E(1 - W_j/W)$. Тогда согласованный фильтр, очищая этот остаточный сигнал от АБГШ, обеспечивал бы выходное отношение мощностей сигнала и шума в виде (индекс «J» отвечает помехе)

$$q_J^2 = \frac{2E(1 - W_j/W)}{N_0} = q^2(1 - W_j/W), \quad (5.2)$$

Соотношения (5.1)-(5.2) свидетельствуют о достоинствах широкополосных сигналов с точки зрения противостояния воздействию помех. Действительно, чем шире полоса сигнала W в сравнении с полосой помехи W_j , тем меньше дополнительная спектральная плотность в первом случае и энергетические потери – во втором (при постоянстве мощности помехи J), и следовательно, больше q_I^2 и q_J^2 . Однако, если пиковая мощность сигнала P ограничена и не может быть увеличена, то расширение полосы сигнала не реализуется тривиальным укорочением длительности сигнала, поскольку в этом случае произойдет уменьшение энергии сигнала, а также отношений

SNR и SIR. В результате приходим к следующему выводу: достижение высокой помехоустойчивости по отношению к узкополосной помехе, не прибегая к увеличению энергии сигнала или пиковой мощности, возможно только расширяя спектр сигнала независимо от его длительности, т.е. *при использовании широкополосной технологии.*

5.2. Электромагнитная совместимость.

Проблема электромагнитной совместимости (ЭМС) является одной из наиболее приоритетных в современной беспроводной технике. ЭМС подразумевает бесконфликтное сосуществование различных систем в эфире, несмотря на то, что каждая из них принимает не только свой собственный сигнал, но и сигналы остальных систем. Очевидно, что невозможно полностью исключить взаимное влияние, когда несколько систем одновременно функционируют в относительно малой области. Любая активная, т.е. излучающая электромагнитную волну, система неизбежно окажет воздействие на все соседние, и в задачу разработчика систем входит сделать это потенциально вредное воздействие терпимым.

В задаче ЭМС принимают участие две стороны. Первая из них, которую можно назвать «излучающей», старается минимизировать свое вмешательство, обусловленное излучаемой мощностью, на другую соседнюю, называемую «принимающей» системой. Мотивацией подобного поведения являются не только этические принципы, но и строгие международные и национальные правила, за исполнением которых тщательно следят службы, уполномоченные налагать соответствующие санкции. Любая из систем второй, «принимающей» стороны также предпринимает собственные меры, нацеленные на нейтрализацию чужих сигналов, попадающих в ее приемник.

К числу традиционных способов обеспечения ЭМС относятся строгое частотное планирование, контроль за которым осуществляют национальные и международные институты, использование антенн с узкой направленностью, тщательная разработка высокочастотных узлов и др. В данном параграфе кратко рассмотрим вопрос, почему технология распределенного спектра также может быть отнесена к приведенному перечню мероприятий.

Со стороны излучающей системы оправданной является следующая логика поведения. Сделать излучаемый сигнал, насколько это возможно, почти незаметным для специальных приемников радиоперехвата (см. 5.2), путем усложнения закона модуляции, т.е. расширением спектра, в результате чего подобный сигнал почти не будет оказывать вредного влияния на обычную стороннюю систему, работающую в той же полосе частот. Задача состоит только в выборе такого выигрыша от обработки, который гарантировал бы достаточно низкий уровень спектральной плотности мощности используемого сигнала относительно спектральной интенсивности естественного шума на входе приемника сторонней системы. Как правило полагают, что «достаточно низким» считается уровень в -7 дБ, т.е. $N_s / N_0 \leq 0.2$. Подстановка $N_s = P/W = E/WT$ в последнее соотношение приводит к критерию ЭМС $E/WTN_0 \leq 0.2$ или $q^2 / WT \leq 0.4$, где снова целевой параметр определен в терминах отношения сигнал-шум скрытного приемника q^2 и выигрыша от обработки WT . Если, например, отношение сигнал-шум скрытной системы на входе стороннего приемника было бы 20 дБ, то величину $WT \geq 400$ можно было бы считать удовлетворительной в смысле ЭМС. На практике подобные оценки должны координироваться с дальностью, определяющей радиус окружности вокруг излучающей системы, вне которой сигнал от излучателя считается практически безвредным для других систем [16].

Со стороны принимающей системы любой сторонний сигнал на входе приемника может трактоваться как узкополосная или широкополосная помеха, и все доводы о пользе широкополосности в борьбе с помехами (см. 5.1) снова применимы. Следовательно,

широкополосная философия хорошо ложится в русло обеспечения ЭМС.

5.3. Энергетическая скрытность

Предположим, что некоторая система использует сигнал с нетривиальным законом модуляции, параметры которого неизвестны противоборствующей стороне, лишая последнюю возможности применения согласованного фильтра или коррелятора для обнаружения сигнала. В этих условиях естественно полагать, что у противостоящей системы нет иного выбора, как рассматривать перехватываемый сигнал в виде случайного и основывать его обнаружение только на факте появления или отсутствия некоторого избытка энергии в подозрительном участке частотного диапазона. Таким образом, *энергетический детектор*, называемый также *радиометром*, который является оптимальным с точки зрения обнаружения ограниченного по полосе шумового сигнала на фоне АБГШ, принимается в качестве рабочего инструмента на стороне перехвата. На рис. 5.3 изображена структурная схема подобного энергетического детектора. Полосовой фильтр с полосой W_i , пропускающей весь спектр сигнала или только его часть, фильтрует наблюдение с целью устранения внеполосного шума. Затем квадратичный амплитудный детектор осуществляет оценку мгновенной мощности, которая в дальнейшем интегрируется для выработки оценки энергии $\hat{E}$ в пределах интервала наблюдения T_{ob} . Полученная оценка энергии сравнивается затем с порогом E_t , а выполнение неравенства $\hat{E} \geq E_t$ влечет принятие решения о наличии в наблюдении сигнала наряду с естественным шумом, тогда как противоположное свидетельствует об отсутствии сигнала. В реальной практике перехватчику могут быть неизвестны заранее сведения о частотном диапазоне и

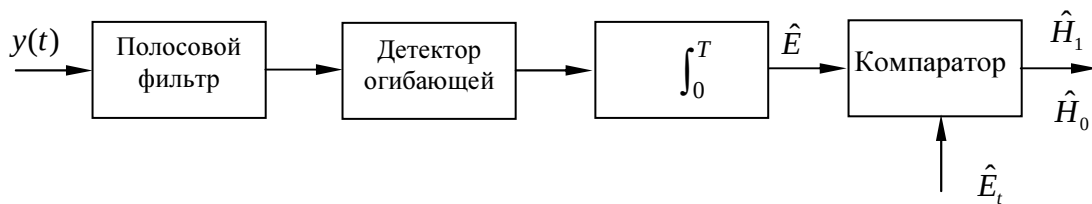


Рис. 5.3. Энергетический детектор.

интервале времени, занимаемом сигналом.

Учитывая эти обстоятельства, его стратегия будет заключаться в комбинировании указанных параметров, осуществляя процедуру обнаружения либо путем сканирования частотно-временной области, либо используя набор параллельных каналов, каждый из которых ответственен за анализ ограниченного участка частотно-временной области. В любом случае качество работы приемника системы-перехватчика будет полностью определяться характеристикой энергетического детектора, настроенного на истинную для перехватываемого сигнала частотно-временную зону. Это позволяет идеализировать ситуацию с априорными сведениями перехватчика и предполагать, что он осведомлен о возможном месте появления энергии на плоскости время–частота. Поскольку наблюдение вне длительности сигнала не содержит информации о его присутствии, можно положить $T_{ob} = T$, как это было сделано в схеме на рис. 5.3.

На рис. 5.4 представлена прямоугольная аппроксимация спектра сигнала вместе с равномерной спектральной плотностью мощности естественного АБГШ (а) и амплитудно-частотная передаточная функция полосового фильтра радиометра (б). С точки зрения перехватчика указанием на присутствие сигнала служит избыточная (сигнальная) спектральная плотность мощности $N_s/2 = P/2W$, дополнительная к спектральной

плотности мощности теплового шума $N_0/2$. В отсутствии сигнала случайный процесс на выходе полосового фильтра радиометра характеризуется мощностью $\sigma_n^2 = N_0 W_i$, а при его наличии – $\sigma_n^2 + \sigma_s^2 = (N_s + N_0)W_i = (P/W + N_0)W_i$.

Найдем среднее и дисперсию на выходе детектора огибающей. Прежде всего, выходное напряжение u_d детектора равняется мгновенной входной мощности. Следовательно, математическое ожидание $\bar{u}_d$ напряжения u_d в отсутствии сигнала есть просто средняя мощность отфильтрованного шума $\bar{u}_{d0} = \sigma_n^2$, тогда как при наличии сигнала его значение возрастает до $\bar{u}_{d1} = \sigma_n^2 + \sigma_s^2$. Очевидно, что именно приращение $\bar{u}_d$, обусловленное сигналом и равное

$$\Delta \bar{u}_d = \bar{u}_{d1} - \bar{u}_{d0} = \sigma_s^2 = N_s W_i = P W_i / W, \quad (5.3)$$

позволяет перехватчику надеяться зарегистрировать факт появления сигнала.

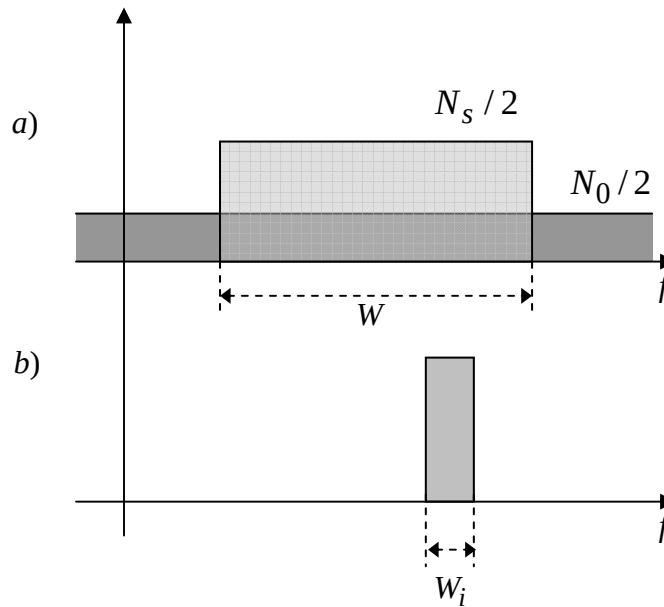


Рис. 5.4. Спектр наблюдения на входе радиометра и ширина полосы полосового фильтра.

С другой стороны, этот полезный (с точки зрения перехватчика) эффект маскируется случайными флуктуациями u_d , величина которых оценивается дисперсией $D\{u_d\}$. Последняя характеристика может быть найдена на основании того факта, что мгновенная мощность полосного процесса совпадает с мгновенным значением его огибающей Y , возведенной в квадрат и умноженной на 0.5, так что $u_d = Y^2/2$, откуда $D\{u_d\} = D\{Y^2\}/4$. В свою очередь, дисперсия любой случайной величины может быть вычислена как среднее значение ее квадрата минус квадрат ее среднего значения:

$$D\{Y^2\} = \overline{Y^4} - \left(\overline{Y^2}\right)^2. \quad (5.4)$$

Теперь, используя тот факт, что огибающая Y гауссовского случайного полосного процесса с дисперсией σ^2 имеет релеевский закон распределения [3]

$$W(Y) = \begin{cases} \frac{Y}{\sigma^2} \exp\left(-\frac{Y^2}{2\sigma^2}\right), & Y \geq 0, \\ 0, & Y < 0 \end{cases}$$

а ее четные моменты находятся посредством интегрирования [1,3]

$$\overline{Y^{2n}} = \int_0^\infty Y^{2n} W(Y) dY = (2\sigma^2)^n \int_0^\infty \left(\frac{Y^2}{2\sigma^2}\right)^n \exp\left(-\frac{Y^2}{2\sigma^2}\right) d\left(\frac{Y^2}{2\sigma^2}\right) = n! (2\sigma^2)^n.$$

Подстановка полученного результата в (5.4) дает $D\{Y^2\} = 8\sigma^4 - 4\sigma^4 = 4\sigma^4$. В случае отсутствия сигнала на входе приемник перехватчика в данном выражении следует положить $\sigma^2 = \sigma_n^2$. Строго говоря, при появлении сигнала отфильтрованное наблюдение может отличаться от гауссовского процесса, делая применимость только что полученного результата сомнительной. Данное замечание, однако, не имеет особой важности в рассматриваемом вопросе, поскольку скрытная (intended) система спроектирована наилучшим образом для сокрытия своего сигнала под тепловым шумом и имеются веские основания полагать, что сигнал оказывает пренебрежимо малый эффект на дисперсию мгновенной мощности и, таким образом, на дисперсию отклика детектора. Поэтому, независимо от присутствия сигнала, значение дисперсии u_d можно считать неизменной, т.е.

$$D\{u_d\} = D\{Y^2\}/4 = \sigma_n^4 = (N_0 W_i)^2. \quad (5.5)$$

С целью получения оценки среднего значения $\bar{u}_d$ и регистрации его увеличения, обусловленного присутствием сигнала, интегратор в схеме на рис. 5.3 осуществляет усреднение во времени величины отклика детектора на интервале наблюдения T . Для того, чтобы сделать постоянную составляющую на выходе детектора заметной на фоне случайных флуктуаций необходимо их сгладить посредством интегрирования. Данная процедура возможна только тогда, когда флуктуации отклика детектора около $\bar{u}_d$ достаточно быстры и изменяют свою полярность много раз за период T , что и позволяет обеспечить их компенсацию проведением усредняющего действия. Другими словами, число статистически независимых отсчетов n_s отклика детектора на интервале T должно быть достаточно большим. Протяженность во времени (время корреляции τ_c) автокорреляционной функции случайного процесса служит достоверной начальной аппроксимацией минимального интервала времени между отсчетами, начиная с которого отсчеты могут считаться независимыми. Поскольку отфильтрованное наблюдение занимает полосу W_i , время корреляции может быть оценено как $\tau_c \approx 1/W_i$, что обеспечивает $n_s \approx W_i T$ независимых отсчетов.

Хотя фактически процедура интегрирования может быть выполнена как непрерывная операция, ее результат оказывается достаточно близким с получаемым путем суммирования n_s независимых отсчетов, что реализуется значительно проще особенно средствами цифровой техники. Для более точного выполнения анализа необходимо найти плотность распределения выходного значения интегратора $\hat{E}$ для обеих гипотез (отсутствия и наличия сигнала) и затем, интегрируя по решающим областям, получить две вероятности: *ложной тревоги* и *правильного обнаружения*. Эта плотность вероятности подчиняется закону распределения хи-квадрат, который достаточно сложен и не поддается прозрачному физическому толкованию. Однако снова можно использовать тот факт, что сигнал является слабым и для его надежного обнаружения требуется интегрирование большого числа отсчетов n_s . Тогда центральная предельная теорема

позволяет полагать, что плотность распределения величины $\hat{E}$ на выходе интегратора является гауссовской и, следовательно, нахождение среднего и дисперсии $\hat{E}$ достаточно для выполнения вычислений ранее упомянутых вероятностей.

Когда на входе присутствует только АБГШ, среднее и дисперсия $\hat{E}$ определяются, как $\bar{\hat{E}} = n_s \bar{u}_{d0}$, $D\{\hat{E}\} = n_s D\{u_d\} = n_s \sigma_n^4$, где второй результат следует из (5.5) и статистической независимости интегрируемых отсчетов. Аналогично, когда наблюдаются АБГШ и сигнал, $\bar{\hat{E}} = n_s \bar{u}_{d1}$, однако дисперсия остается неизменной, поскольку сигнал слаб. Таким образом, плотности распределения на выходе интегратора, отвечающие гипотезам H_0 (сигнал отсутствует) и H_1 (сигнал присутствует) представимы в виде

$$W(\hat{E}|H_i) = \frac{1}{\sqrt{2\pi n_s \sigma_n^2}} \exp \left[-\frac{(\hat{E} - n_s \bar{u}_{di})^2}{2n_s \sigma_n^4} \right], i = 0, 1.$$

Когда $\hat{E}$ превосходит порог E_t , ложная тревога происходит тогда, когда сигнал в действительности отсутствует, а правильное обнаружение, если сигнал достиг приемника. Следовательно, вероятности P_f и P_d указанных событий определяются соответственно как

$$p_f = P(\hat{E} \geq E_t | H_0) = \int_{E_t}^{\infty} W(\hat{E}|H_0) d\hat{E} = Q \left(\frac{E_t - n_s \bar{u}_{d0}}{\sqrt{n_s \sigma_n^2}} \right)$$

и

$$p_d = P(\hat{E} \geq E_t | H_1) = \int_{E_t}^{\infty} W(\hat{E}|H_1) d\hat{E} = Q \left(\frac{E_t - n_s \bar{u}_{d1}}{\sqrt{n_s \sigma_n^2}} \right).$$

Перепишав второе соотношение в виде

$$p_d = Q \left(\frac{E_t - n_s \bar{u}_{d0}}{\sqrt{n_s \sigma_n^2}} - \frac{n_s (\bar{u}_{d1} - \bar{u}_{d0})}{\sqrt{n_s \sigma_n^2}} \right), \quad (5.6)$$

можно увидеть, что если допустимый уровень вероятности ложной тревоги задан заранее, то первая дробь в скобках (5.6) фиксирована, и вероятность обнаружения полностью определяется отношением

$$q_i = \frac{n_s (\bar{u}_{d1} - \bar{u}_{d0})}{\sqrt{n_s \sigma_n^2}} = \frac{n_s \Delta u_d}{\sqrt{n_s D\{u_d\}}}. \quad (5.7)$$

Фактическое содержание последнего соотношения очевидно: это отношение сигнал-шум по напряжению на выходе интегратора, показывающее пропорцию между полезным (увеличением математического ожидания, обусловленное появлением сигнала) и мешающим (стандартное отклонение случайных флуктуаций) компонентами $\hat{E}$.

Применяя (5.3) и (5.5) совместно с соотношением $n_s = W_i T$ в (5.7), получаем

$$q_i = \sqrt{n_s} \frac{P}{WN_0} = \sqrt{W_i T} \frac{P}{WN_0}.$$

Это соотношение позволяет понять, что с точки зрения перехватчика максимально возможная полоса фильтра, т.е. равная полосе сигнала ($W_i = W$), является оптимальной, обеспечивая наибольшее отношение сигнал-шум

$$q_i = \frac{P\sqrt{T}}{\sqrt{W} N_0} = \frac{q^2}{2\sqrt{WT}}, \quad (5.8)$$

где $q^2 = 2E/N_0 = 2PT/N_0$, как и ранее, отношение сигнал-шум на выходе согласованного фильтра приемника скрытной системы.

Очевидно, что значение q^2 должно быть достаточно большим, поскольку в противном случае скрытная система не сможет нормально функционировать. Совершенно ясно также, что у скрытной системы имеется только единственная возможность предотвратить обнаружение своего сигнала потенциальным перехватчиком: *использовать сигнала с распределенным спектром, обладающие максимально возможным значением выигрыша от обработки WT*. Обращение к рис. 5.4 позволяет дать физическое обоснование данного заключения. Расширение спектра сигнала с постоянной энергией и длительностью уменьшает уровень его спектральной плотности мощности, скрывая ее под спектром естественного теплового шума.

Завершая этот раздел, отметим, что рассмотренные преимущества распределенного спектра в настоящее время широко используются не только военными потребителями и спецслужбами. То обстоятельство, что широкополосный сигнал практически незаметен для оборудования систем радиоконтроля, серьезным образом влияет на лицензионную политику. В частности, число коммерческих систем, активно работающих в эфире, не прибегая к получению лицензии, становится все шире, и в некоторых регионах выделяются специальные участки частотного диапазона для подобного безлицензионного использования.

5.4. Структурная скрытность сигнала

Продолжая линию предыдущего параграфа, напомним еще раз, что единственной причиной, вынуждающей перехватчик прибегнуть к такому неэффективному инструменту как энергетический приемник, является отсутствие информации о структуре обнаруживаемого сигнала, т.е. его закона модуляции. По этой причине перехватчик не может обрабатывать сигнал аналогично приемнику скрытной системы (т.е. осуществлять согласованную фильтрацию). Очевидно, что в случае недостаточной структурной сложности сигнала и осведомленности перехватчика о его возможных альтернативных вариантах, последний может попытаться их все реализовать. Соответствующим оборудованием для этого может служить набор параллельных согласованных фильтров либо единый перестраиваемый фильтр (несколько фильтров), пригодный для обработки сигналов различных по структуре последовательно во времени, если сигнал, который необходимо обнаружить, принимается достаточно долго. Поэтому другая сторона стратегии скрытной системы в борьбе с перехватчиком состоит в применении сигналов с практически не раскрываемой структурой.

Подобная задача характерна для военных и коммерческих систем, которые не стремятся полностью скрыть факт своего присутствия в эфире. Одним из основных требований, стоящих перед подобными системами, является стремление избежать несанкционированного доступа к обслуживанию, адресованного только разрешенным пользователям, или искажения передаваемой информации. Глобальная спутниковая навигационная система GPS Navstar служит наглядным примером подобного рода. В ней используется два дальномерных канала: открытый (обозначаемый как C/A) и специальный (P). Сигнал, применяемый во втором канале, позволяет с высокой точностью определять свое местоположение, и правительство США, руководящее системой, не считает рациональным разрешать безусловный доступ к этому каналу. Для его защиты от несанкционированного использования предприняты специальные меры, касающиеся закона модуляции сигнала.

В дисциплинах, имеющих дело с безопасностью информации, степень защиты данных определяется числом равновероятных конкурирующих *ключей*, с помощью которых противостоящий криптоаналитик старается взломать шифротекст, т.е. зашифрованные данные. В применении к структуре сигнала каждый из таких ключей есть

ничто иное, как закон модуляции, который, как правило, повторяется с периодом T . Предположим, что сигнал построен на основе M -ичного алфавита, т.е. возможны M реализаций индивидуального сигнального элемента (чипа). Если полоса, отводимая системе, равна W , то общее сигнальное пространство имеет размерность, определяемое как WT (игнорируя возможность его удвоения при переходе к полосным сигналам, см. 2.3-2.5), т.е. закон модуляции может быть сконструирован посредством WT чипов. Очевидно, что величина M^{WT} определяет общее число различных законов модуляции, т.е. число конкурирующих ключей, и, значит, системный конструктор, отвечающий за секретность модуляционного формата разрабатываемой системы, должен использовать сигналы с достаточно большим выигрышем от обработки.

Пример 5.1. Сигнал Р-канала (Р-код) в системе GPS Navstar является бинарным ($M = 2$) с полосой $W \approx 10$ МГц. Его структура имеет регулярный характер и повторяется с периодом $T = 7$ суткам. Будучи скрытым под тепловым шумом, этот сигнал не может быть восстановлен путем посимвольного приема и только знание его тонкой структуры позволяет эффективно очистить его от АБГШ. Для раскрытия структуры данного сигнала необходимо проанализировать 2^{WT} альтернатив, а поскольку $WT = 7 \times 86400 \times 10^7 > 10^{12}$, то число ключей больше, чем 2 в степени 10 с двенадцатью нулями, что невозможно и вообразить. По этой причине Р-код считается не раскрываемым, и за 25 лет истории GPS не появилось ни одной статьи с сообщениями об успешных попытках его взлома.

Заклучим параграф еще одним заявлением о достоинствах широкополосности: данная технология в значительной степени способствует криптографической защите структуры сигнала.

5.5. Дискретные широкополосные сигналы

Обратимся вновь к общей модели полосного сигнала

$$s(t) = \operatorname{Re}[\dot{S}(t) \exp(j2\pi f_0 t)], \quad \dot{S}(t) = S(t) \exp(j\gamma(t)).$$

Совершенно очевидно, что распределение спектра сигнала осуществляется посредством соответствующего управления комплексной огибающей сигнала $\dot{S}(t)$, т.е. модуляцией мгновенных значений амплитуды $S(t)$ и начальной фазы $\gamma(t)$. Как уже было отмечено, «чистая» амплитудная модуляция не может служить эффективным инструментом для распределения спектра, поскольку она значительно расширяет полосу только ценой концентрации энергии сигнала в пределах коротких временных интервалов. Действительно, данный тип модуляции подразумевает работу с короткими простыми сигналами. Напротив, угловая (фазовая или частотная) модуляция способна безгранично (по крайней мере, теоретически) расширить спектр без изменения распределения энергии сигнала во времени, т.е. длительности сигнала, благодаря чему ее роль в технологии распределенного спектра является фундаментальной. Амплитудная же модуляция служит только вспомогательным инструментом, который иногда проявляет себя продуктивно в комбинации с угловой модуляцией.

В зависимости от типа привлеченной модуляции все широкополосные сигналы могут быть подразделены на *непрерывные* и *дискретные*. Для первых закон модуляции, т.е. комплексная огибающая $\dot{S}(t)$, является непрерывной функцией времени, тогда как модулируемые параметры вторых (амплитуда, частота, начальная фаза) – кусочно-постоянной, скачкообразно изменяющей свои значения только в дискретные моменты времени. В дальнейшем основное внимание будет сфокусировано на дискретных сигналах, учитывая их доминирующую роль в большинстве современных и проектируемых коммерческих системах.

Дискретный сигнал представляет собой последовательность элементарных символов (импульсов) фиксированной формы, повторяющихся с некоторым

фиксированным временным интервалом. Элементарный импульс, называемый *чипом*, комплексная огибающая $\dot{S}_0(t)$, определяющая его форму, и внутренняя угловая модуляция могут быть любыми. Как правило (но не обязательно), временной интервал Δ между последовательными чипами равен или превосходит длительность чипа Δ_c . Модуляция всего сигнала заключается в манипулировании амплитудами, фазами и, возможно, частотами отдельных чипов. Тогда формальное представление комплексной огибающей дискретного сигнала дается соотношением

$$\dot{S}(t) = \sum_{i=-\infty}^{\infty} a_i \dot{S}_0(t - i\Delta) \exp(j2\pi F_i t), \quad (5.9)$$

где, в дополнение к уже объясненным обозначениям, a_i и F_i – соответственно

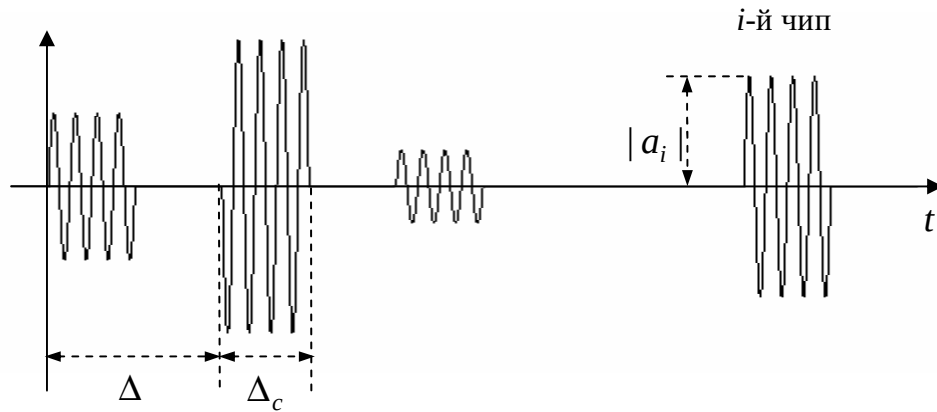


Рис. 5.4. Пример дискретного сигнала.

комплексная амплитуда и частота (в значениях сдвига относительно фиксированной центральной частоты) i -го чипа. Очевидно, что последовательность $\{|a_i|, i = \dots, -1, 0, 1, \dots\}$ определяет вещественные амплитуды чипов, т.е. их амплитудную модуляцию. Аналогично, последовательности $\{\phi_i = \arg a_i, i = \dots, -1, 0, 1, \dots\}$ и $\{F_i, i = \dots, -1, 0, 1, \dots\}$ задают законы модуляции чипов по фазе и частоте. Рис. 5.4 может быть полезен для понимания некоторых вышеприведенных определений.

Предположим, что в модели (5.9) вещественные амплитуды $|a_i|$ могут принимать ненулевые значения только при попадании i в диапазон $0 \leq i \leq N-1$, а при $i < 0$ и $i > N$ значения амплитуд $|a_i| = 0$. В этом случае сигнал представляет собой пакет конечного числа N манипулированных чипов. Подобный сигнал будем называть *импульсным* или *апериодическим*. Длительность апериодического сигнала определится, как $T = (N-1)\Delta + \Delta_c$. Другой важной версией является сигнал, у которого закон модуляции повторяется с периодом N чипов: $a_i = a_{i+N}$, $F_i = F_{i+N}$, $i = \dots, -1, 0, 1, \dots$. Естественно назвать дискретный сигнал подобного типа *периодическим*. Фактический период сигнала составляет $T = N\Delta$, и любой периодический сигнал представляет собой по сути повторение с периодом $N\Delta$ апериодического, который, в свою очередь, является однопериодным сегментом периодического сигнала. В обоих случаях параметр N назовем *длиной* дискретного сигнала.

В рамках описанной обобщенной модели различают несколько категорий дискретного сигнала в зависимости от конкретного способа модуляции чипа.

1. Если манипуляции подвергаются только комплексные амплитуды чипов, а все

частоты остаются одинаковыми ($F_i = 0, i = 0, 1, \dots, N-1$), то сигнал называется *амплитудно-фазоманипулированным* (АФМ) (*amplitude-phase shift keying* (APSK)). Последовательность комплексных амплитуд чипов $\{a_i, i = 0, 1, \dots, N-1\}$ называется *кодовой последовательностью* или просто *кодом*.

2. Если манипуляция осуществляется только над фазами чипов АФМ сигнала, а амплитуды остаются неизменными ($|a_i| = 1, i = 0, 1, \dots, N-1$), то сигнал является ФМ (PSK). ФМ сигналы типичны для т.н. широкополосных систем с прямым расширением спектра.

3. Дальнейшая типизация в рамках ФМ сигналов может производиться с учетом модуляционного алфавита. Если применяются только бинарные комплексные амплитуды ($a_i = \pm 1$, или, что эквивалентно, $|a_i| = 1, \phi_i \in \{0, \pi\}, i = 0, 1, \dots, N-1$), то сигнал является бинарным ФМ (BPSK), при четверичном алфавите вида $a_i = \pm 1, \pm j$, или, что эквивалентно, $|a_i| = 1, \phi_i \in \{0, \pi, \pm \pi/2\}, i = 0, 1, \dots, N-1$ сигнал является квадратурным ФМ (QPSK), и т.д.

4. Если регулированию подлежат только частоты чипов, а комплексные амплитуда остаются постоянными, то сигнал является частотно-манипулированным (ЧМ или FSK). Кодовая последовательность подобного сигнала представляет собой последовательность частот $\{F_i, i = 0, 1, \dots, N-1\}$. Сигналы этого типа используются, в частности, в системах с прыгающей частотой.

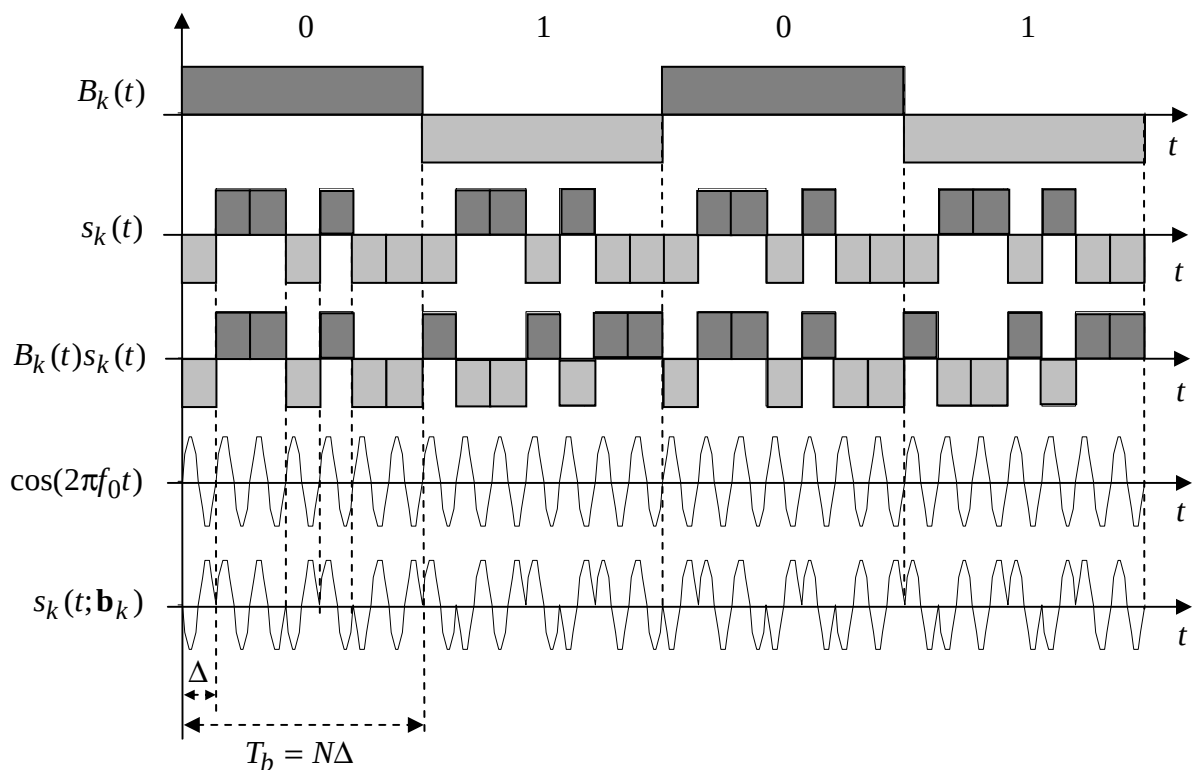


Рис.5.4. Прямое расширение бинарных ФМ данных бинарной сигнатурой.

Прямое расширение используется во всех CDMA стандартах 2-го и 3-го поколений: IS-95 (cdmaOne), UMTS и cdma2000. В них используются различные комбинации алфавитов манипулированных данных и расширяющих последовательностей.

Частотно-манипулированные сигналы. На рис. 5.6 с помощью решетки задан закон модуляции ЧМ сигнала длины $N=8$ с параметрами $N=8, M=5$, N – число временных дискретов, M – частотных.

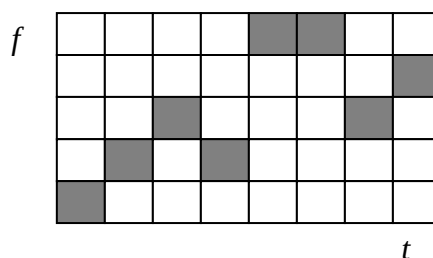


Рис. 5.6. ЧМ сигнал с параметрами $N=8, M=5$.

Возьмем ЧМ сигнатуру рисунка 5.6 и используем ее для расширения спектра с

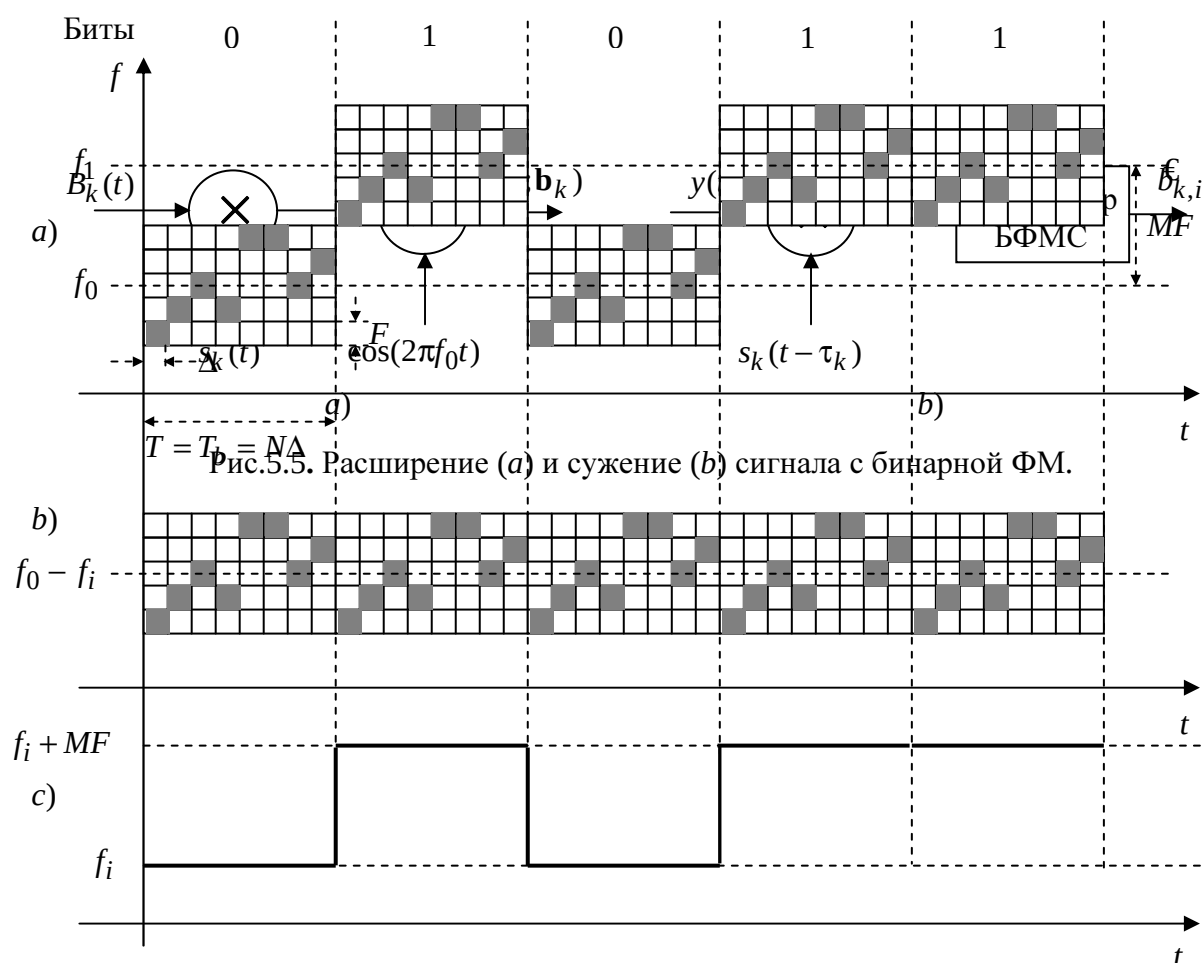


Рис.5.7. Расширение-сужение с помощью быстрой FH (fast FH).

помощью быстрой FH в комбинации с бинарной модуляцией данных. В этом случае число различных частот в сигнатуре $M=5$, длина сигнатуры $N=8$, а один символ данных

передает один бит информации, так что $T = T_b$. Предположим, что на один бит данных приходится 8 скачков частоты. Тогда полная последовательность ЧМ чипов, представленная на рис. 5.6, передается в течение одного бита. Если бит данных равен нулю, то частотный образец излучается на несущей частоте, равной f_0 , тогда как для бита, равного единице, несущая частота изменяется на f_1 . Очевидно, что разность между частотами f_1 и f_0 должна быть не меньше полосы, занимаемой сигнатурой, т.е. MF . На рис. 5.7, а изображен передаваемый частотный образец, соответствующий потоку битовых данных вида 01011. Как можно видеть, спектр одиночного бита данных, полоса которого перед расширением была примерно $1/T_b$, расширяется до полосы, равной $MF \approx M/\Delta = MN/T_b$, т.е. становится в MN раз шире.

На приемной стороне снятие расширения заключается в обратном преобразовании наблюдаемого колебания на промежуточную частоту f_i . С этой целью используется опорное колебание несущей частоты $f_0 - f_i$, модулированное согласно ЧМ образца сигнатуры с необходимой задержкой во времени (рис. 5.7, б). В результате сигнал со снятым расширением представляет собой обычное узкополосное колебание, частотно-манипулированное в соответствии с передаваемыми данными, где нулевой бит данных передается более низкой частотой, чем бит, отвечающий единице. Теперь спектр отдельного символа данных возвращается к полосе $1/T_b$ и может быть использован обычный бинарный ФМ демодулятор для восстановления принятых данных

5.6. Псевдослучайные кодовые последовательности

5.6.1. Основные характеристики

Из перечисленных фазоманипулированных сигналов (ФМ), дискретные частотные (ДЧ), дискретные составные частотные с внутриимпульсной ФМ (ДСЧ - ФМ) и сигналы с псевдослучайной перестройкой рабочей частоты (ППРЧ), наибольший интерес для использования в системах связи и управления военного назначения представляют сигналы ФМ и ППРЧ и их комбинации.

Основой для формирования этих сигналов являются псевдослучайные кодовые последовательности (ПСП), используемые либо для фазовой модуляции ($0/\pi$) сигнала несущей частоты в случае ФМ ШПС, либо для задания закона переключения рабочих частот при ППРЧ. К таким ПСП в общем случае предъявляются следующие требования:

большой размер ансамбля последовательностей (сигналов), формируемых на единой алгоритмической основе;

сбалансированность структуры последовательностей;

• оптимальность авто- и взаимокорреляционных функций последовательностей сигналов в ансамбле.

Рассмотрим их более подробно.

Системы сигналов по величине объема L можно разделить на три вида:

- малые, когда $L = \sqrt{B} \ll B$;
- нормальные (ортогональные или квазиортогональные), когда $L = B$;
- большие, когда $L \gg B$,

причем для ПСП база B определяется как $B = N$, где N - длина ПСП. Большинство известных систем сигналов - малые или нормальные.

Для современных систем связи и управления необходимо иметь системы ШПС, объем которых экспоненциально зависит от базы, т.е. $L = A \exp(\gamma B)$, где A и γ - некоторые постоянные. Если такой закон реализовать нельзя (таких систем в настоящее время нет),

то необходимо реализовать большие системы, объем которых растет по степенному закону, т.е. $L = A B^n$, где A, n - постоянные, причем $n > 1$.

Сигналы, входящие в систему, должны обеспечить минимально возможный уровень взаимных помех, который в основном определяется допустимым уровнем максимальных пиков взаимокорреляционных функций (ВКФ). Периодическая ВКФ определяется как

$$R = \frac{1}{N} \sum_{j=1}^N a_{rj} a_{sj} \quad (5.9)$$

где a_{rj} и a_{sj} - символы двух кодовых последовательностей, соответствующих сигналам с номерами r и s (при периодическом режиме две последовательности могут быть сдвинуты относительно друг друга, но в формуле (8.21-5.9) всегда будет N слагаемых); аperiodическая ВКФ имеет следующий вид:

$$R_i = \frac{1}{N} \sum_{j=i-1}^N a_{rj} a_{sj-i}$$

т.е. здесь последовательности перекрываются частично.

Известны пределы любой большой системы кодовых последовательностей - так называемые полные коды. Полный код состоит из всех последовательностей, отличающихся хотя бы одним элементом, т.е. содержит $L = p^N$ кодовых последовательностей (здесь p - объем алфавита символов). При $p = 2$ имеем $L = 2^N$.

Если значения символов a_{ri} и a_{sj} составляют +1 или -1, то любую кодовую последовательность длиной N можно охарактеризовать ее весом, определяемым как разность числа символов 1 и -1 на длине N

$$W = \sum_{j=1}^N a_j \quad (5.10)$$

Количество всех весов двоичного полного кода равно 2^N , но число разных весов будет гораздо меньше. При этом веса обладают групповыми свойствами. Количество кодовых последовательностей, имеющих данный вес, находится как число сочетаний из N элементов по Q и равно C_N^Q , где Q - число +1 в сумме (5.10).

Вес W и значение корреляционной функции R последовательности связаны соотношением $R = W/N$, поэтому знание распределения весов полного кода позволяет определить статистические характеристики корреляционных функций и оценить объемы ансамблей последовательностей с заданными корреляционными характеристиками.

В [23] приведены результаты, на основании которых можно оценить объемы больших систем кодовых последовательностей, если заданы допустимые значения корреляционных функций КФ в ансамбле. Так, в случае малых допустимых значений КФ среднее значение объема большой системы

$$\bar{L} > 3\sqrt{\pi} \exp(R_0^2 N) / 4R_0^3 N^{5/2}$$

где R_0 - допустимый уровень (максимум модуля) КФ.

Поскольку при увеличении N уменьшается дисперсия КФ, возникает вопрос, как правильно задавать совместно допустимый уровень R_0 и длину последовательности N . Исследование этого вопроса показало, что КФ обладают «пороговым свойством»: при $R_0 < R_{пор}$ КФ произвольного сигнала превышает порог с вероятностью, близкой к 1, при $R_0 > R_{пор}$ это событие происходит с малой вероятностью. Пороговое значение

$$R_{пор} = \sqrt{\ln(aN) / N},$$

где $a = 1,6$. Если положить, что допустимый уровень $R_0 = \sqrt{\alpha} R_{пор}$, где $1 < \alpha < 10$ (α зависит от N), то среднее значение объема большой системы

$$\bar{L} > C(\alpha) N^{\alpha-1} / [\ln(aN)]^{3/2}, \quad (5.11)$$

где $C(\alpha) = 3\pi^{1/2} a^{-\alpha} 2^2 \alpha^{3/2}$. Из (8.23-5.11) следует, что среднее значение объема больших систем с допустимыми корреляционными свойствами растет по степенному закону, который существенно отличается от экспоненциального закона. Введение «относительной единицы» измерения уровня КФ в виде $R_{\text{пор}}$ позволило корректно определить объем большой системы сигналов. Например, если необходимо построить систему сигналов с объемом, равным базе, т. е. длине последовательности N , то надо положить $\alpha = 2$.

При этом допустимый уровень КФ должен в $2\sqrt{\ln(aN)}$ раз превышать среднеквадратическое значение, равное $1/\sqrt{2N}$. Если необходимо построить системы с $L = N^2$, то множитель $\alpha = 3$ и т.д.

Таким образом, более реальная задача, которую возможно решить, - построение системы сигналов с объемом, который определяется степенным законом $L = AN^{\alpha-1}$, где A - некоторая постоянная, зависящая от N и α .

Кодовые последовательности являются псевдослучайными в том смысле, что удовлетворяют следующим трем критериям случайности:

- критерий 1 (свойство уравнищенности): в каждом периоде последовательности число «1» отличается от числа «-1» не более, чем на единицу;
- критерий 2 (свойство серий): в течение периода последовательности половина серий единиц и «-1» имеет длину 1, одна четверть - 2, одна восьмая - 3 и т. д. до тех пор, пока это продолжение имеет смысл;
- критерий 3 (свойство корреляции): если последовательность поэлементно сравнивать с любым ее циклическим сдвигом в течение периода этой последовательности, то число совпадений отличается от числа несовпадений не больше чем на единицу.

5.6.2. Классы псевдослучайных последовательностей

По виду алгоритма формирования ПСП можно разделить на два класса:

- линейные ПСП;
- нелинейные ПСП.

К линейным ПСП относятся М-последовательности, последовательности Голда, последовательности Касами и ряд других. Их свойства и методы формирования представлены в [22-24, 28].

Рассмотрим особенности некоторых из перечисленных линейных последовательностей.

5.6.2.1. М-последовательности

Это линейные рекуррентные последовательности (ЛРП) максимальной длины, формирующими полиномами которых являются примитивные неприводимые полиномы $f_n(x)$ памяти n над полем чисел $GF(q)$. Символы последовательности рассчитываются по рекуррентному уравнению вида

$x_i = (a_1 x_{i-1} + a_2 x_{i-2} + \dots + a_n x_{i-n}) \bmod q, i > n$, где $\{a_k\}$ - коэффициенты порождающего полинома

$$f_n(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_{n-1} x + a_n. \quad (5.12)$$

Порождающий полином называется примитивным, если он *неприводим*, т.е. не допускает разложение на множители, и делит без остатка двучлен $2^n - 1$, корни которого описывают все состояния регистра.

Формируются М-последовательности с помощью генераторов, содержащих регистры сдвига, охваченных цепью обратной связи. Известно, что с помощью такой схемы можно получить последовательность двоичных элементов любого периода в пределах от 1 до 2^n . М-последовательность имеет наибольший период $L = 2^n - 1$ благодаря

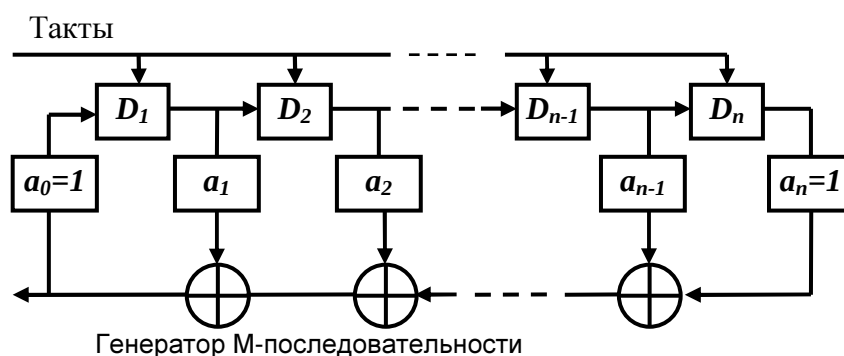
тому, что она включает все, кроме нулевого, состояния регистра. Задача синтеза структуры логической обратной связи решается с помощью примитивных полиномов.

Последние записываются, как обычно, в виде (5.12), либо двоичным числом $\{a_k\}_{k=0,1,\dots,n}$ полинома (5.12), либо соответствующим ему восьмеричным числом. Различные формы записи для некоторых полиномов представлены в таблице 5.1.

Таблица 5.1

Восьмеричный код	Двоичный код $a_0, a_1, \dots, a_n$	Уравнение $f(x)$
$n = 5$		
45	100101	$x^5 + x^2 + 1$
75	111101	$x^5 + x^4 + x^3 + x^2 + 1$
		$x^5 + x^4 + x^2 + x^1 + 1$
$n = 6$		
103	1 000 011	$x^6 + x + 1$
147	1 100 111	$x^6 + x^5 + x^2 + x + 1$
155	1 101 101	$x^6 + x^5 + x^3 + x^2 + 1$
141	1 100 001	$x^6 + x^5 + 1$

Структура логической обратной связи определяется значениями коэффициентов $\{a_k\}$, а длина регистра сдвига – величиной n .



К достоинствам M -последовательностей относятся хорошие корреляционные свойства и простота генерирования. Однако количество сигналов в ансамбле весьма ограничено, поэтому M -последовательности используются только как исходные для получения больших систем квазиортогональных сигналов (путем их логического комбинирования).

5.6.2.2. Последовательности Голда

Эти последовательности представляют собой результат сложения по модулю 2 символов двух исходных M -последовательностей с различными формирующими полиномами $f_{n1}(x)$ и $f_{n2}(x)$, степени которых не кратны четырем. Формирующие полиномы таких сигналов определяются как произведение исходных полиномов:

$$f_{\text{гр}}(x) = f_{n1}(x) f_{n2}(x).$$

Достоинством последовательностей Голда является большой ансамбль ($Q = L+1$) сигналов заданной длины $L = 2^n - 1$ с хорошими корреляционными свойствами. При этом пара последовательностей имеет трехуровневую взаимокорреляционную функцию, боковые выбросы которой могут принимать значения $1; -1; |2^{(n+2)/2}|$, где n - степень исходных формирующих полиномов. Достоинством последовательностей Голда является простота их генерирования. Эти последовательности являются ЛРП не максимальной длины, их можно рассматривать в качестве комбинированных, когда функцией комбинирования является сложение по модулю 2.

5.6.2.3. Последовательности Касами

Также представляют собой ЛРП не максимальной длины. Отыскание формирующих полиномов основано на следующих доказанных теоремах:

1. Если $a \in GF(2^n)$ является корнем примитивного полинома $f_n(x)$ степени n , n - четное и a является корнем неприводимого полинома $f_s(x)$ степени s , $s = 2^{n/2} + 1$, то можно сформировать 2 последовательностей периода $L = 2^n - 1$ в соответствии с полиномом $g(x) = f_n(x) f_s(x)$. При этом боковые выбросы автокорреляционной и взаимокорреляционной функций имеют величины: -1 ; $-s$; $|s-2|$.

2. Полиномы $f_n(x)$ и $f_s(x)$ удовлетворяют условиям теоремы 1, при этом n - четное число. Пусть также при $r = 2^{(n+2)/2} + 1$ a^r является корнем неприводимого полинома $f_r(x)$ степени r . Тогда в соответствии с полиномом $g(x) = f_n(x) f_s(x) f_r(x)$ можно сформировать $2^{n/2}(2^n + 1)$ последовательностей, боковые выбросы АКФ и ВКФ которых равны: -1 ; $-s$; $-r$, $|s-2|$; $|r-2|$.

3. Полиномы $f_n(x)$ и $f_s(x)$ удовлетворяют условию теоремы 2, причем $n \equiv 0 \pmod{4}$. Тогда в соответствии с полиномом $g(x) = f_n(x) f_r(x)$ можно сформировать 2^n последовательностей, корреляционные свойства которых такие же, как в условиях теоремы 2.

Линейные ПСП не обеспечивают структурной скрытности ШПС из-за их легкой предсказуемости. Действительно, достаточно осуществить посимвольный прием $2n$ следующих друг за другом символов ПСП (где n - длина эквивалентного сдвигового регистра с линейной обратной связью (СРЛОС), причем n много меньше длины последовательности N), чтобы раскрыть закон формирования ПСП и сформировать имитационную помеху, весьма опасную для радиосистемы связи и управления с ШПС.

5.6.2.4. Нелинейные последовательности

Нелинейные последовательности обладают более высокой непредсказуемостью. Используя длину эквивалентного СРЛОС в качестве оценки непредсказуемости нелинейных ПСП, можно классифицировать ПСП как практически непредсказуемую, если n достигает периода последовательности N , т.е. $n = N$.

Нелинейные последовательности могут быть сформированы следующими способами.

1. Генераторами с нелинейной внешней логикой (рис. 5.6, а).
2. Генераторами составных ПСП (рис. 5.6, б).
3. Генераторами с нелинейной функцией обратной связи (рис. 5.6, в).

Нелинейные ПСП обладают следующими существенными недостатками [28].

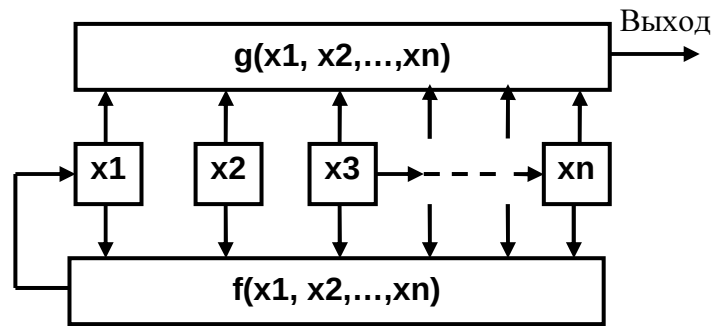
1. Они принадлежат к категории последовательностей, образующих ансамбли с малыми объемами.

2. Номенклатура длин последовательностей в пределах фиксированных числовых интервалов невелика.

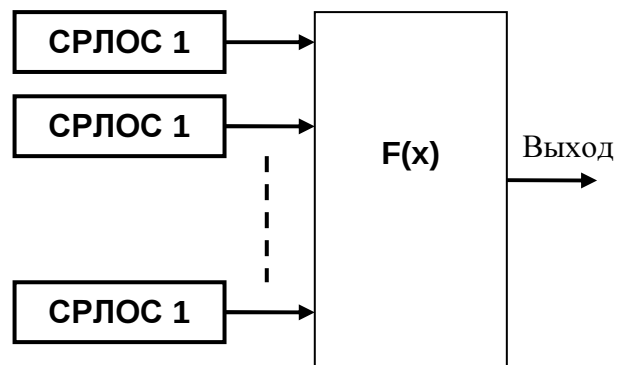
3. Усложнение алгоритма формирования, как правило, приводит к нарушению сбалансированности структуры ПСП.

4. Взаимокорреляционные свойства ПСП с увеличением объема ансамбля начинают резко ухудшаться.

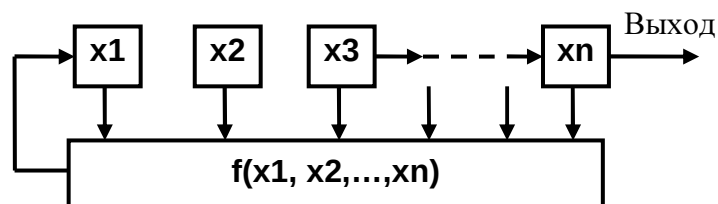
В настоящее время существует нерешенная проблема – разработка алгоритмов построения больших систем нелинейных ПСП с хорошими взаимокорреляционными и структурными свойствами. Алгоритмы построения систем ПСП должны быть детерминированными, поскольку сигналы должны быть известны в точке приема. Регулярных методов синтеза больших ансамблей нелинейных ПСП не существует, поэтому задачу формирования больших ансамблей таких ПСП можно решать методом статистического синтеза, т.е. случайным выбором кодовых последовательностей.



а)



б)



в)

Рис. 5.6. Генераторы нелинейных ПСП

Исследование численными методами корреляционных свойств случайных последовательностей показало, что их апериодические АКФ существенно хуже АКФ М-последовательностей, в то время как ВКФ и тех и других примерно одинаковы. Для улучшения автокорреляционных свойств случайных последовательностей статистический синтез целесообразно дополнять отбором сигналов по автокорреляционным функциям в соответствии с заданным критерием. Простейший критерий отбора предполагает отбрасывание тех сигналов, у которых наибольший боковой выброс АКФ превышает заданный уровень. При жестких требованиях к параметрам отбора и при значительных базах сигналов процедуры отборов становятся весьма трудоемкими. В связи с этим представляется желаемым повышение производительности отбора случайных последовательностей по АКФ.

Известно, что сигналы, у которых АКФ обладают малыми боковыми пиками, например М-последовательности, содержат число блоков $M_0 = (N + 1)/2$. Блоки, представляющие собой последовательности одинаковых символов, могут быть единичными (состоять из одного символа), двойными (из двух символов) и т.д. Более того, не только для М-последовательностей, но вообще для двоичных последовательностей было показано, что при числе блоков $M = M_0$ вероятность больших пиков АКФ минимальна. В соответствии с этим целесообразно статистический синтез

последовательностей дополнить отбором по числу блоков, т. е. считать пригодными те сигналы, у которых число блоков M близко к M_0 .

В типичной или «средней» случайной последовательности число символов должно удовлетворять соотношению

$$N = \sum_{k=1}^{k_{\max}} k V_k, \quad (5.12)$$

где $\sum_{k=1}^{k_{\max}} V_k = M$, общее число блоков должно равняться $M = M_0 = (N + 1)/2$,

а число блоков длины k удовлетворять равенству

$$V_k = M_0 \rho_k = M_0 2^{-k} \quad (5.13)$$

Здесь V_k - число блоков одинаковой длины k , $k_{\max}$ - длина максимального блока, $\rho_k = 1 / 2^k$ - вероятность появления блока длиной k .

Число таких последовательностей определяется полиномиальным законом

$$L(M_0, V) = M_0 ! / \prod_{k=1}^{k_{\max}} V_k !$$

где аргумент характеризует блоковую структуру последовательности. Например, при $N = 16$, $M_0 = 8$, $V_1 = 4$, $V_2 = 2$, $V_4 = 0$ и $V_3 = V_5 = 1$ имеем $L(M_0, V) = 840$. То есть число последовательностей, удовлетворяющих (5.13), составляет примерно 1/80 часть от общего числа последовательностей длины $N = 16$ ($2^{16} = 6,4 \cdot 10^4$).

Для последовательности, удовлетворяющей равенствам (5.12) и (5.13), можно утверждать, что статистические характеристики их АКФ и ВКФ близки к статистическим характеристикам М-последовательностей, что иллюстрируется табл. 5.2.

Таблица 5.2

Корреляционные функции	$\sigma_R \sqrt{N}$	$m_{ R } \sqrt{N}$	$\sigma_{ R } \sqrt{N}$	$R_{\max} \sqrt{N}$
АКФ М-последовательностей	0,40	0,32	0,26	0,7...1,25
ВКФ М-последовательностей	0,73	0,54	0,48	1,4...5
КФ случайных последовательностей	0,70	0,56	0,43	2,1...3,5

В качестве характеристик взяты следующие: среднеквадратичное значение боковых пиков R_i , определяемое через дисперсию $\sigma_R^2 = \frac{1}{2N} \sum_{i=-(N-1)}^{M-1} R_i^2$, среднее значение модулей

боковых пиков $m_{|R|} = \frac{1}{2N} \sum_{i=-(N-1)}^{N-1} |R_i|$, среднеквадратичное значение модулей боковых пиков,

определяемое через дисперсию $\sigma_{|R|}^2 = \sigma_R^2 - m_{|R|}^2$, а также значение максимального бокового пика $R_{\max}$.

. В табл. 5.2 все характеристики приведены в ненормированном виде, т. е. умножены на $\sqrt{N}$. В результате цифры, приведенные в табл. 5.2, характеризуют превышение $\sigma_R, m_{|R|}, \sigma_{|R|}, R_{\max}$ уровня $\sqrt{N}$. Отметим, что среднее значение боковых пиков $m_R = 0$.

5.6.3. Практическое применение ШПС в системах связи и управления

В настоящее время широкополосные шумоподобные сигналы благодаря своим специфическим особенностям:

- низкой спектральной плотности,
- высокой устойчивости по отношению к организованным помехам,
- высокой помехозащищенности по отношению к узкополосным помехам большой мощности,
- возможности разделения по кодовому признаку,
- высокой стойкости в условиях многолучевого распространения,
- высокой разрешающей способности при измерении расстояния находят все более широкое применение в различных радиотехнических системах.

ШПС используются при построении спутниковых систем связи и навигации, сотовых систем подвижной радиосвязи, локальных радиосетей, систем связи внутри зданий и в ряде других систем.

Спутниковые системы связи

Наиболее полно специфические свойства ШПС проявляются при создании спутниковых систем связи. Организация многостанционного доступа с кодовым разделением сигналов, создание локальных сетей малых станций для связи с удаленными и труднодоступными районами, создание сетей малых земных станций для распределения базы данных, решение проблемы использования линий связи с напряженной энергетикой, организация дополнительных связей в загруженных стволах ретранслятора - вот далеко не полный перечень задач, эффективное решение которых связано с применением ШПС. Система многостанционного доступа с кодовым разделением сигналов (МДКР), использующая ШПС, позволяет реализовать такие потребительские характеристики, которые невозможно получить в системах с временным либо частотным разделением сигналов. В системе МДКР постоянно используется полная выходная мощность и полоса частот спутникового ретранслятора независимо от числа одновременно активных абонентов. Система обладает свойством эластичности, т.е. способностью плавно изменять качество связи в зависимости от ее загрузки.

Система МДКР может быть построена так, что в ней для передачи каждого сообщения не потребуется выделения особого сигнала, организованного путем назначения специальной временной позиции, либо частотной полосы. Каждый абонент использует один и тот же широкополосный канал со случайным доступом без предварительной заявки, которую необходимо делать в системах многостанционного доступа по требованию.

Система не требует взаимной синхронизации всех передатчиков, имеющих доступ к спутниковому ретранслятору, что является серьезной проблемой для систем многостанционного доступа с временным разделением сигналов. Примером наиболее полного использования преимуществ кодового разделения сигналов может служить спутниковая система связи, разработанная в центре перспективных исследований Великобритании для большого числа земных станций, работающих в условиях высокой связности, широкого диапазона пропускных способностей каналов связи и неоднородного характера нагрузки. В разработанной системе КР - МДКР кодовое разделение (КР) сигналов используется на двух этапах: при объединении передаваемых сообщений на каждой земной станции в единый поток и для реализации многостанционного доступа (МДКР) к спутниковому ретранслятору. Использование КР вместо широко применяемого временного разделения позволило организовать прозрачные каналы с широким диапазоном пропускных способностей - от 100 бит/с до 10 000 бит/с, оперативно изменять объем трафика и требуемую пропускную способность каналов, использовать различные виды передаваемых сообщений - стартстопную и синхронную телеграфию, аналоговую и

цифровую телефонию и т. п., предъявлять различные требования к их качественным показателям.

Многостанционный доступ в этой системе был реализован путем использования фазоманипулированных ШПС. Тактовая частота этих сигналов могла быть установлена равной 5, 10 либо 20 МГц. Выигрыш при обработке по отношению к синусоидальной помехе, расположенной в центре спектра сигнала, оказался близким к расчетной величине и для сигнала, ограниченного в полосе частот 20 МГц, составил 40,5 дБ.

Считается, что наиболее широкое применение системы связи с многостанционным доступом на основе кодового разделения найдут при организации сетей с использованием малых земных станций, поскольку кодовое разделение обеспечивает возможность свободного доступа к спутниковому ретранслятору практически без каких-либо требований на координацию станций по частоте и времени и хорошую электромагнитную совместимость с другими спутниковыми и наземными радиосетями.

Как известно, ШПС могут быть переданы по каналу связи при отрицательных (в децибелах) отношениях сигнал/помеха. Эта особенность иногда используется для организации дополнительных каналов со сравнительно небольшой пропускной способностью в загруженных стволах спутниковых ретрансляторов, либо для передачи данных совместно с телевидением. Интерес к передаче низкоскоростных сигналов по загруженным стволам определяется не столько перспективой увеличения их пропускной способности, сколько желанием организовать независимую сеть передачи сообщений на базе имеющегося оборудования. Так, например, сообщается об экспериментальной проверке возможности одновременной работы системы связи с ШПС с телевизионным ЧМ сигналом и многоканальным телефонным ЧМ сигналом с частотным разделением каналов. Показано, что при использовании помехоустойчивого кодирования допустимое для совместной работы входное отношение мощности ШПС к мощности основного сигнала, загружающего ствол ретранслятора, находится в достаточно широких границах - от -15,6 дБ до -20 дБ для телевидения и от -14,5 дБ до -20,5 дБ для ЧМ телефонии и, таким образом, позволяет организовать совместную передачу с телевидением нескольких десятков, а с ЧМ телефонией до сотни ШПС.

5.6.3.1. Сотовые системы подвижной радиосвязи

Еще одна большая область использования ШПС связана с созданием наземных систем подвижной радиосвязи. В настоящее время в связи с проблемой острого дефицита спектра широкое распространение получили сотовые системы подвижной радиосвязи, обеспечивающие возможность многократного использования одних и тех же частотных каналов благодаря их географическому разнесению. При этом в крупных городах с высокой плотностью подвижных абонентов размеры ячеек приходится уменьшать до нескольких километров в диаметре, и пропускная способность сотовой системы многостанционного доступа подвижных абонентов к базовой станции оказывается ограниченной взаимными помехами от ячеек с повторным использованием частотных каналов. В этих условиях более эффективными с точки зрения обеспечения максимальной пропускной способности в выделенном для системы диапазоне частот оказываются ШПС, структура которых как раз и предназначена для обеспечения предельно возможной устойчивости в условиях взаимных помех. При этом защита от взаимного влияния между различными ячейками при использовании ШПС определяется не наличием буферных ячеек с другими частотными каналами, как это имеет место в сотовых системах с узкополосными ЧМ сигналами, а величиной запаса помехоустойчивости, создаваемой большой базой ШПС.

Существенным преимуществом использования шумоподобных сигналов является уменьшение глубины замираний, вызванных многолучевым распространением радиоволн,

что, в свою очередь, может обеспечить значительный (10 дБ и более) энергетический выигрыш в радиолиниях с такими сигналами по сравнению с обычными (узкополосными) методами передачи.

Важный шаг в направлении создания сотовых систем с МДКР был сделан фирмой QUALCOMM, разработавшей сотовую систему подвижной связи CDMA (работающую в диапазоне 800 МГц при полосе сигнала 1,25 МГц), в которой была реализована структура сотовой сети с использованием в каждой из рабочих зон одного и того же участка радиоспектра. Эта система сочетает высокую помехоустойчивость, обеспечиваемую использованием ШПС, с большой пропускной способностью. В 1993 г. в США на основе CDMA был принят федеральный стандарт IS-95

5.6.3.2. Локальные радиосети

В последнее время широко обсуждается целесообразность использования ШПС в различного рода локальных сетях. Пристальный интерес к локальным сетям обусловлен их важной ролью как инструмента автоматизации учрежденческой деятельности. Поскольку в процессе своей деятельности все учреждения имеют тенденцию к структурным изменениям и различного рода реформированиям, стационарные кабельные сети стали заменяться мобильными радиосетями. Последние удобно строить с использованием ШПС, обладающих стойкостью к искажениям из-за многолучевого распространения и способностью к разделению по кодовому признаку.

Фирма «Хьюлетт Паккард» (США) разработала два типа экспериментальных локальных радиосетей. В первой ШПС образуется псевдослучайной последовательностью с тактовой частотой 25,5 МГц и излучается передатчиком мощностью 5 Вт в диапазоне 1,5 ГГц. Приемник, характеризующийся выигрышем при обработке в 24 дБ, выполнен на приборе с ПАВ. Эта система используется в оборудовании, предназначенном для проведения голосования. Вторая система построена на основе использования существующей сети распределения учебных ТВ программ в диапазоне 2500...2650 МГц и имеет те же технические параметры.

Разновидностью локальных сетей являются системы связи внутри зданий, предназначенные для связи ЭВМ с удаленными терминалами и для реализации бесшнурового телефона. В обоих случаях применение ШПС обусловлено наличием внутри зданий многолучевого распространения сигналов.

5.6.3.3. Спутниковые радионавигационные системы (СРНС)

К основным преимуществам СРНС по сравнению с другими навигационными системами следует отнести:

- высокую точность навигационных определений с учетом собственной динамики движения потребителя (П);
- глобальность действия;
- устойчивость к помехам.

В качестве примера СРНС остановимся на системе Navstar (Navigation Satellite Time and Ranging - навигационной спутниковой системе измерения времени и координат). Для измерения пространственных координат в этой системе используется псевдодальномерный метод измерения, а для измерения скорости - доплеровский метод.

В системе Navstar выбрана двухчастотная передача, причем на каждой из частот излучается сложный ФМ сигнал. Сигнал формируется из одной опорной частоты 10,23 МГц следующим образом: $f_1 = 10,23 \text{ МГц} \times 154 = 1572,42 \text{ МГц}$; $f_2 = 1023 \text{ МГц} \times 120 = 12276 \text{ МГц}$. Сигнал на частоте f_1 является основным и состоит из двух квадратурных составляющих

$$s_1(t) = A_1 [p(t) \oplus b(t)] \cos \omega_1 t + 2A_1 [(C/A)t \oplus b(t)] \sin \omega_1 t,$$

где p - код высокой точности с тактовой частотой 10,23 МГц; $b(t)$ - двоичный

служебный код данных, имеющих скорость 50 Бод; C/A - код, обеспечивающий быстрый захват с тактовой частотой 1,023 МГц; A , - амплитуда сигнала.

Сигнал на частоте f_2 служит для устранения ионосферных ошибок

$$s_2(t) = A_2 [p(t) \oplus b(t)] \cos \omega_2 t.$$

Отметим, что коды p и C/A сфазированы между собой.

Код C/A является кодом Голда, состоящим из $N = 1023$ элементов, с периодом $T - N \tau = 1023 / 1,023 \text{ МГц} = 1 \text{ мс}$.

Код p образуется в результате перемножения двух укороченных псевдошумовых кодов и G_2 . При этом код G_1 имеет число элементов $N_1 = 15,345 \cdot 10^6$, а код $G_2 - N_2 = 15,345037 \cdot 10^6$, период этих кодов близок 1,5 с, а период производного кода равен 265 суткам. Однако каждые 7 суток происходит смена (обновление) всех кодов, что в сочетании с большой длиной кода p и отсутствием описания образующих полиномов G_1 и G_2 обеспечивает высокую структурную скрытность и защиту системы Navstar от несанкционированного использования.

Таким образом, выбранные характеристики сигнала обеспечивают выполнение требований к высокоточной навигации потребителей, которые были сформулированы ранее. Кроме того, выбор пары кодов (p и C/A) является своеобразным компромиссом для совмещения в одной системе требований, предъявляемых военными и гражданскими потребителями.

Военные системы радиосвязи

Применение ШПС в зарубежных военных системах радиосвязи имеет достаточно длинную историю. Уже в начале 60-х годов в литературе появилось упоминание о системе Фантом, а затем о системе спутниковой связи Таксатком. Далее системы с ШПС создавались весьма быстрыми темпами как в интересах стратегического, так и тактического звена управления для различных родов войск.

В этих системах применялись разные типы ШПС: сложные ФМ сигналы, сигналы в виде частотно-временных матриц (ДЧ), а также сигналы с псевдослучайной перестройкой рабочей частоты (ППРЧ).

В отличие от систем гражданского применения важнейшим требованием к таким системам является обеспечение сигнальной (структурной) скрытности используемых в них ШПС. Это требование приводит к необходимости уделять особое внимание созданию ансамблей ШПС, обладающих высокой стойкостью к разгадыванию их структуры, т. е. в наибольшей степени приближающихся к случайным (шумовым) сигналам. Достижение результатов в развитии этого направления позволит создавать системы связи и управления с принципиально новыми (в смысле обеспечения их помехозащищенности от действия средств радиоэлектронного подавления) возможностями.

5.6.4. Широкополосные методы борьбы с многолучевостью

Сигнал многолучевого канала (индекс “ x ” обозначает принадлежность к случайной величине) можно представить в виде

$$s_x(t) = \sum_k s_{ik}(t) = \sum_k S_{ik}(t) f_i[t - \tau_{zk}(t)] \quad (5.14)$$

где k – номер луча, $k = (1, K)$; K – общее число лучей;

S_{ik} – огибающая i -й посылки, принятой по k -му лучу;

τ_{zk} – запаздывание посылки k -го луча;

f_i – посылка единичной амплитуды, соответствующая выбранному виду манипуляции и передаваемому символу. Например, для БФМ:

$$f_i(t) = \begin{cases} \sin(\omega_0 t + \varphi_i), & 0 \leq t \leq T \\ 0, & \text{для других } t \end{cases}$$

Принятый сигнал (5.14) имеет две компоненты:

$$s_x(t) = s_p(t) + s_{\phi l}(t), \quad (5.15)$$

$s_p(t)$ – регулярная, $s_{\phi l}(t)$ – случайная.

Мгновенные значения сигнала $s_x(t)$ имеют нормальное распределение. Для огибающей S_x используют законы:

Релея-Райса

$$W(S_x) = \frac{S_x}{\sigma_{\phi l}^2} \exp\left(-\frac{S_p^2 + S_x^2}{2\sigma_{\phi l}^2}\right) I_0\left(\frac{S_p S_x}{\sigma_{\phi l}^2}\right), \quad \{S_x^2 = 2\sigma_{\phi l}^2\} \quad (5.16)$$

и Накагами (m -распределение) – более общий закон

$$W(S_x) = \frac{2}{\Gamma(m)} \left(\frac{m}{S_x^2}\right)^m I_0\left(-\frac{m S_x^2}{S_x^2}\right), \quad S_x \geq 0, \quad m \geq 0,5 \quad (5.17)$$

Он позволяет описать более широкий класс распределений, нежели закон Релея-Райса в зависимости от параметра m :

$m = 0,5$ – нормальный закон

$m = 1$ – распределение Релея

$m > 1$ – распределение Релея-Райса

Если ни одна составляющая сигнала не имеет регулярной части, то описание производится законом Релея. Такой канал, как отмечалось выше, называется релеевским и характеризуется наиболее глубокими замираниями.

Полную информацию о наличии многолучевого распространения радиоволн и его параметрах (количество лучей K , интенсивность a_k , запаздывание t_k , фазовый сдвиг огибающей θ_k), дает импульсная характеристика канала связи h_c . Для ее получения необходимо, чтобы передатчик излучал настолько короткий одиночный импульс, что на

приемной стороне все многолучевые сигналы были бы разрешимы. На практике, однако, достаточно иметь длительность импульса передатчика T_p не более обратной величины полосы частот, занимаемой системой связи: $T_p \leq 1/W$.

На рис. 5.7 представлена осциллограмма принятого многолучевого сигнала в деловом квартале Сан-Франциско. Ее обычно называют *профилем многолучевости*.



Рис. 5.7. Пример профиля многолучевости

Передатчик излучал радиоимпульсы длительностью 0,1 мкс.

Видно, что при этом не все лучи разрешаются; часть из них, отраженная от протяженных объектов, сливается в импульсы длительностью более 0,1 мкс. Запаздывания сигналов отдельных лучей измеряются относительно сигнала, приходящего по линии визирования (ЛВ): передатчик – приемник (прямой луч). Его запаздывание обозначают как ЛВЗ – запаздывание вдоль линии визирования. Профиль многолучевости есть не что иное, как действительная часть импульсной характеристики канала $\text{Re}\{h_c\}$.

Применительно к широкополосным системам, которым будет уделено основное внимание, целесообразно положить, что все лучи разрешимы, т. е. выполняется условие

$$|t_k - t_l| > 1/W \quad \text{для всех } k \neq l. \quad (5.18)$$

Любые два луча, например, k_1 и k_2 , для которых $|t_{k1} - t_{k2}| < 1/W$, будут рассматриваться как один луч с общим запаздыванием $t_k \approx t_{k1} \approx t_{k2}$, интенсивность и фазовый сдвиг которого определяются выражением

$$a_k \exp(j\theta_k) \equiv a_{k1} \exp(j\theta_{k1}) + a_{k2} \exp(j\theta_{k2}).$$

Прием сигналов в условиях многолучевости распространения представляет собой одну из форм разнесенного приема, где информация от передатчика к приемнику передается по нескольким естественно разнесенным лучам, а не по нескольким заранее разнесенным частотным, временным, пространственным и т. п. каналам. Следовательно, вместо того, чтобы рассматривать многолучевость распространения как некое вредное явление, влияние которого необходимо так или иначе устранить, его можно рассматривать как некое благоприятное явление, которое можно использовать для улучшения характеристик системы.

Опубликованные в литературе работы по расчету приемников многолучевых сигналов можно разбить на две основные группы. В наиболее ранних из них (см. библиографию в обзоре [7]) главное внимание уделялось структуре приемника при наличии явно разрешимых лучей; использование такой схемы приема обосновывалось возможностью оптимального объединения вкладов отдельных лучей. В своей простейшей форме этот подход пренебрегает наличием межсимвольной интерференции, которая может появиться в том случае, когда из-за запаздываний сигнала, обусловленных многолучевостью распространения, отклик приемника на переданный символ будет попадать на временные интервалы, занятые последующими переданными символами. Поэтому такой подход оправдан лишь в том случае, когда длительность переданного символа велика по сравнению с длительностью профиля многолучевости (длительностью многолучевого сигнала).

В более поздних работах было показано, что методы, разработанные для коррекции данных, передаваемых по телефонным линиям, применимы и для решения проблемы многолучевости в системах радиосвязи. Здесь при конструировании приемника основное внимание уделяется уменьшению эффектов межсимвольной интерференции, а объединение сигналов, пришедших по различным лучам, подразумевается лишь в неявной форме. Такой подход, по-видимому, наиболее целесообразен в том случае, когда отдельные лучи распространения сигнала не разрешаются, а длительность символа много меньше времени многолучевого растяжения сигнала.

Под временем многолучевого растяжения сигнала τ_p , которое выше определялось как длительность профиля многолучевости, будем понимать разницу между максимальным и минимальным запаздыванием лучей:

$$\tau_p = \max \tau_{зан} - \min \tau_{зан}$$

5.6.4.1. Оптимальный приемник при многолучевом распространении сигнала

Начнем с простейшего случая, когда канал связи образован только одним лучом, т.е. положим в (5.14) $K = 1$. Предположим сначала, что интенсивность луча a_k и его запаздывание t_k известны (для упрощения положим $t_0 = 0$), а фаза несущей θ_0 не известна и является случайной переменной, равномерно распределенной на интервале $[0, 2\pi]$.

Поскольку отсутствие многолучевости подразумевает отсутствие межсимвольной интерференции, то основное внимание мы уделим приему отдельного символа, скажем, на интервале $0 \leq t < T$. Знание этого интервала подразумевает, конечно, наличие некоторой процедуры синхронизации приемника.

Оптимальный приемник, т. е. приемник, который с минимальной вероятностью ошибки принимает решение о том, какой из сигналов s_i , $i = (1, M)$ передан, имеет, как известно, структуру, показанную на рис. 5.8. Это некогерентный приемник: фаза любого луча полагается случайной, так как при приеме сигналов подвижной станцией она изменяется слишком быстро, для того чтобы можно было применить когерентные методы приема.

Из рисунка видно, что принятый сигнал $y(t)$ пропускается через набор из M фильтров, каждый из которых согласован с одним из M сигналов, т. е. через фильтры с импульсными характеристиками $s_i(T-t)$, $0 \leq t < T$. К выходам фильтров подключены детекторы огибающей, выходные напряжения которых сравниваются в момент времени $t = T$. Индекс $i=1, \dots, M$ наибольшего из них является выходом приемника.

Изображенный на рис. 5.8 случай соответствует переданному сигналу с индексом “ j ”, шумовая составляющая принятого сигнала полагается пренебрежимо малой.

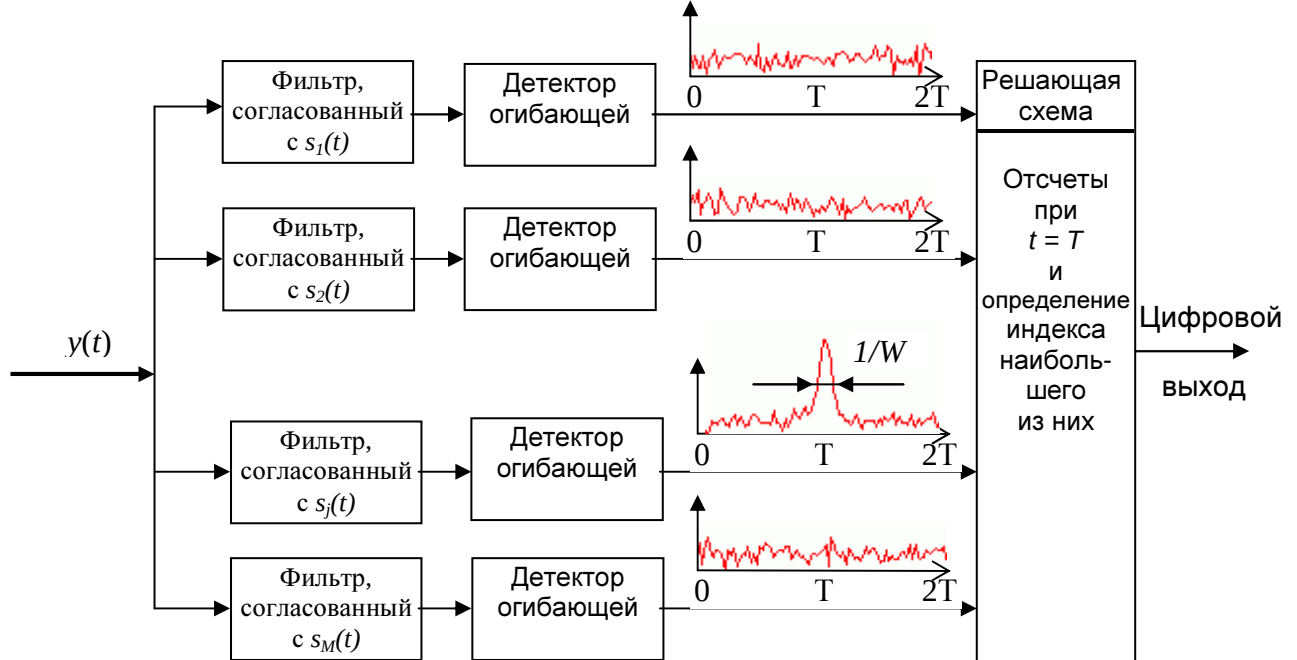


Рис. 5.8. Оптимальный некогерентный приемник для однолучевого канала (равновероятные сигналы равной энергии; гауссовский шум).

Главный пик на выходе детектора j -го канала имеет амплитуду, равную энергии сигнала E в отсутствии шумов; сам сигнал полагается широкополосным (сложным) с базой $B_c = WT \gg 1$, так что согласованный фильтр сжимает его во времени до длительности $1/W$. В окрестности $|t| > 1/W$ j -го канала, а также в интервале $0 \leq t < 2T$ остальных каналов сигнал представляет собой взаимокорреляционные функции принятого сигнала с импульсными характеристиками согласованных фильтров

$$R_{ik} \triangleq \int_0^T S_i^*(\tau) S_k(\tau-t) d\tau, \quad i, k = (1, M), \quad |t| \leq T, \quad (5.19)$$

показанные на рисунке в виде шумовых дорожек. Напомним, что S_i – огибающая i -го сигнала. Для реальных сигналов должны выполняться

$$\begin{aligned} |R_{ik}| &\ll E, & k \neq i; \\ |R_{ii}| &= E, & |t| < 1/W. \end{aligned} \quad (5.20)$$

Эти условия гарантируют, что огибающая j -го канала будет иметь острый главный пик, окруженный низкими «боковыми лепестками», а остальные выходные сигналы будут содержать только боковые лепестки низкого уровня. При тщательном выборе сигналов выполнение условий (5.20) означает, что максимальный уровень боковых лепестков у всех сигналов будет примерно в $2\sqrt{WT}$ раз ниже уровня главного лепестка.

Наконец, отметим еще одну важную особенность рис. 5.8. Так, выходные сигналы на нем изображены для одного отдельного символа, переданного в течение временного интервала $0 \leq t < T$. Если другой символ, скажем, $s_i(t)$, был передан сразу после него на временном интервале $T \leq t < 2T$, то, очевидно, отклик на него появится на временном

интервале $T \leq t < 3T$. Пик главного лепестка i -го выходного сигнала будет точно соответствовать моменту времени $t=2T$, т. е. появится тогда, когда все отклики на первый символ уже затухнут. Следовательно, беря отсчеты выходных сигналов в момент времени $t=2T$, можно вынести решение относительно одного второго символа, что и подтверждает наше утверждение об отсутствии межсимвольной интерференции в однолучевом случае.

5.6.4.2. Оптимальный приемник многолучевого сигнала. Случай известных запаздываний

Если мы попытаемся использовать приемник, показанный на рис. 5.8, в случае, когда имеется несколько лучей ($K>1$), то вследствие линейности среды и наличия согласованных фильтров можно ожидать, что

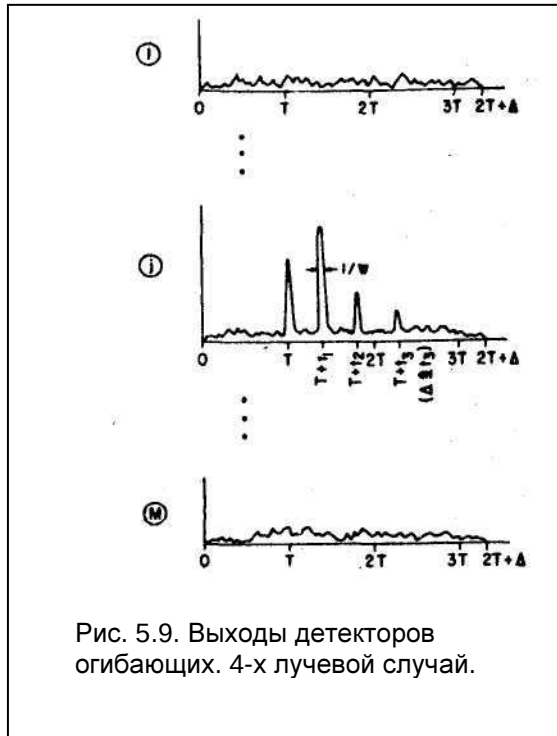


Рис. 5.9. Выходы детекторов огибающих. 4-х лучевой случай.

напряжения на выходах детекторов огибающей будут выглядеть примерно так, как показано на рис. 5.9, на котором они изображены для случая $K = 4$. Здесь через Δ обозначено время многолучевого растяжения: $\Delta = \tau_p$.

Огибающая сигнала в l -м канале, согласованном с пришедшим сигналом, в отсутствии шума равна

$$e_l(t) = \left| \sum_{k=0}^{K-1} a_k \exp(j\theta_k) \int_0^T S_j^*(\tau) S_l(\tau + T + t_k - t) d\tau \right|,$$

$$l = (1, M), \quad 0 \leq t < 2T + \Delta. \quad (5.21)$$

При выполнении условия разрешимости (5.18) главные пики огибающей $e_l(t)$ различны и соответствуют моментам времени $t_0=0, t_1, t_2$, и $t_3 = \Delta$. Высоты этих пиков пропорциональны интенсивностям лучей a_k . Боковые лепестки и сигнала $e_j(t)$, и других выходных сигналов (ни один из последних не имеет главного пика) представляют собой смесь боковых лепестков

сигналов, пришедших по нескольким лучам. Еще раз напомним, что рис. 5.9 соответствует передаче только одного сигнала $s_j(t)$, $0 \leq t < T$.

Сигналы, показанные на рис. 5.8 и 5.9, имеют несколько важных различий.

1) Огибающая $e_l(t)$ имеет несколько сильных пиков. Если решающей схеме приемника, показанного на рис. 5.8, известны времена запаздываний $t_0, \dots, t_{K-1}$, то она может объединить отсчеты всех выходных сигналов, тем самым обеспечивая приемнику преимущества разнесенного приема, о котором говорилось выше. Возможность разрешения лучей по сигналам, показанным на рис. 5.9, составляет суть широкополосного подхода. Если бы мы имели $TW \approx 1$, то пики на рис. 5.9 слились бы и идея разнесенного приема (т. е. объединение сигналов) в явном виде была бы больше не применима.

2) Уровни боковых лепестков всех выходных сигналов возрастают, что, согласно (5.21), обусловлено дополнительными вкладами из-за многолучевости.

3) Отклики на символ, переданный на интервале $0 \leq t < T$, теперь делятся более времени $t=2T$ и, следовательно, перекрываются откликами на следующий символ, переданный на интервале $T \leq t < 2T$. Иными словами, из-за наличия многолучевости мы будем иметь дело с межсимвольной интерференцией.

Эффекты 2 и 3 оказывают вредное действие, эффект 1 является благоприятным. Однако, как будет показано ниже, выгоды, обеспечиваемые эффектом 1, значительно превосходят вредные последствия, вызванные эффектами 2 и 3.

Пока же мы будем пренебрегать эффектами межсимвольной интерференции и рассмотрим структуру приемника, для оптимального приема отдельного символа в

условиях многолучевости распространения. Предположим сначала, что времена запаздывания сигналов $\{t_k\}_0^{K-1}$ для всех лучей известны. Кроме того, будем снова полагать, что фазы сигналов случайны, независимы и равномерно распределены на интервале $[0, 2\pi]$, а интенсивности лучей $\{a_k\}_0^{K-1}$ — также случайны и, возможно, имеют различные распределения.

Интуитивно ясно, что при наложенных условиях оптимальный приемник все еще будет иметь форму, показанную на рис. 5.8, однако решающая схема должна теперь брать отсчеты каждого своего входного сигнала в моменты времени $T+t_k$, $k=0, \dots, K-1$, объединять их для каждого входного сигнала и затем сравнивать полученные величины для определения индекса наибольшей из них. Именно так и обстоит дело при выполнении условий (5.18), когда амплитуды импульсов на выходе j -го канала (рис. 5.9) различны. Однако закон оптимального объединения отсчетов несколько усложняется, так как теперь он зависит от статистик лучевых интенсивностей.

Пусть x_{ik} — отсчет i -й огибающей, взятый в момент времени $T+t_k$. (Согласно (5.21), при отсутствии шума $x_{ik} = e_i(t_k)$). Тогда, если интенсивности лучей известны, оптимальный закон объединения отсчетов будет определяться выражением [8]

$$w_i = \sum_{k=0}^{K-1} \ln I_0 \left(\frac{2a_k x_{ik}}{N_0} \right), \quad (5.22)$$

где I_0 — функция Бесселя, а N_0 — плотность мощности шума в канале. Если интенсивность k -го луча имеет релеевское распределение со среднеквадратичным значением $\psi_k = E[a_k]^2$, а интенсивности всех лучей независимы, то оптимальный закон объединения отсчетов будет иметь следующий вид [8]:

$$w_i = \sum_{k=0}^{K-1} \frac{\psi_k x_{ik}^2}{N_0 + \psi_k E}, \quad (5.23)$$

где E — энергия сигналов. Возможны другие, более сложные законы объединения отсчетов для других законов распределения интенсивностей. Однако во всех случаях решение принимается на основе сравнения величин w_i и определения наибольшего из них.

Заметим, что в предложенных законах объединения отсчеты берутся с различными весами. Так, например, в (5.22), отсчеты объединяются приблизительно по линейному закону, так как при больших значениях аргумента функция $I_0(\cdot)$ растет примерно по экспоненциальному закону; однако отсчеты, соответствующие лучам с большой интенсивностью сигнала, берутся с большим весом. В (5.23) отсчеты объединяются по среднеквадратичному закону, т. е. больший вес придается здесь отсчетам, имеющим большие значения, однако отсчеты всех сигналов с большой среднеквадратичной интенсивностью ($\psi_k E \gg N_0$) берутся с равными весами, а отсчеты слабых сигналов подавляются. Таким образом, любой оптимальный закон объединения состоит в относительном подчеркивании более надежных данных и в относительном подавлении менее надежных.

5.6.4.3. Оптимальный приемник многолучевого сигнала. Случай неизвестных запаздываний

Как отмечалось ранее, времена запаздывания лучей $\{t_k\}_0^{K-1}$ и число лучей K часто являются случайными и заранее не известными переменными. Для того чтобы определить структуру оптимального приемника в этом случае, мы несколько упростим модель многолучевости. Теперь мы будем полагать, что времена t_k независимы и описываются одним общим распределением плотности вероятностей $p(t_k)$, $0 < t_k < \Delta$, а их индексы переупорядочены в соответствии с ростом величины t_k . Такая модель противоречит принятым допущениям, потому что условие разрешимости лучей (5.18) в данном случае уже не выполняется, так как возможны два луча, для которых $|t_k - t_l| < 1/W$. Но вероятность

этого события будет мала, если мала суммарная вероятность интервалов длительностью, равной или меньшей $1/W$; следовательно, распределение $p(t_k)$ должно иметь размытый («диффузный») характер, не иметь высоких пиков и удовлетворять условию $W\Delta \gg 1$.

Однако для отыскания структуры нашего квазиоптимального приемника указанные отличия от реальности несущественны. Поэтому для случая неизвестных запаздываний интенсивности любых двух лучей a_k и a_l для всех $k \neq l$ будут полагаться независимыми, что отличается от реальности, так как лучи с незначительно различающимися запаздываниями имеют, как правило, коррелированные интенсивности. Однако интуитивно ясно, что, как и в случае с упрощенной моделью запаздываний, приемник, полученный на основе этой упрощенной модели интенсивностей лучей, будет близок к оптимальному.

Такой приемник, с учетом сделанных допущений, должен вычислять величины

$$w_i = \int_T^{T+\Delta} p(t-T) F[X_i(t)] dt, \quad i=(1, M) \quad (5.24)$$

где $p(\cdot)$ — определенная выше плотность распределения времен запаздывания лучей, $X_i(\cdot)$ — огибающая выхода i -го фильтра, а $F[\cdot]$ — некоторая нелинейная функция. Затем значения этих величин сравниваются и индекс наибольшей из них считается принятым сигналом. Таким образом, величину w_i можно реализовать, пропуская $X_i(t)$ сначала через нелинейность F , а затем через фильтр с импульсной характеристикой $h_p(t) = p(\Delta - t)$ и беря отсчет сигнала на выходе этого фильтра в момент времени $t = T + \Delta$. Заметим, что фильтр $h_p(t)$ в идеале должен быть согласован с импульсной характеристикой канала $h_c(t)$: $h_p(t) = h_c(\Delta - t)$.

Решающая схема, показанная на рис. 5.8, в этом случае заменяется схемой, показанной на рис. 5.10.

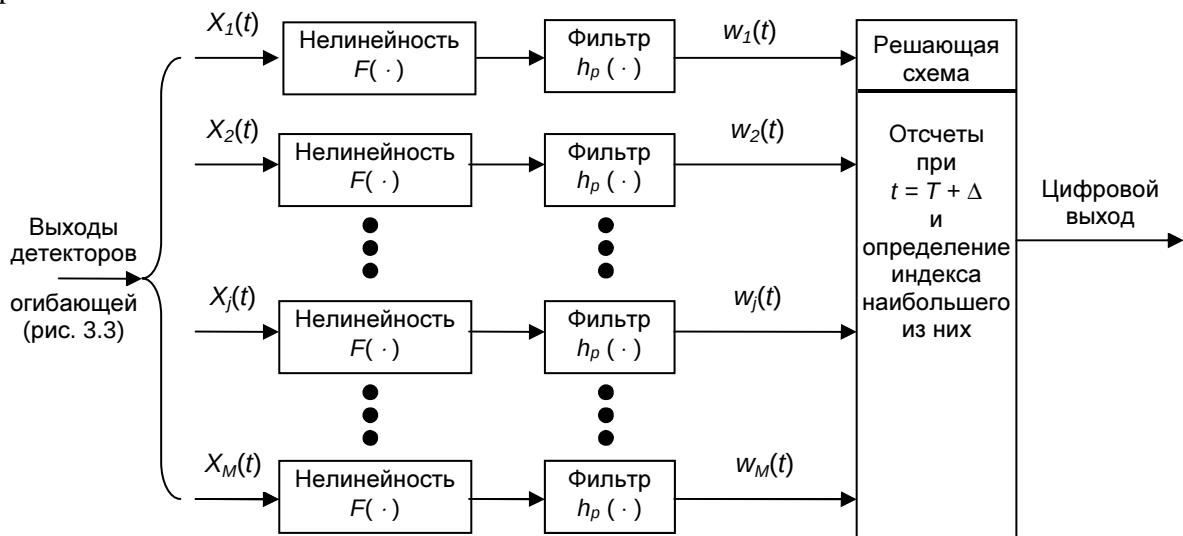


Рис. 5.10. Оптимальный некогерентный приемник для многолучевого канала с неизвестными временами

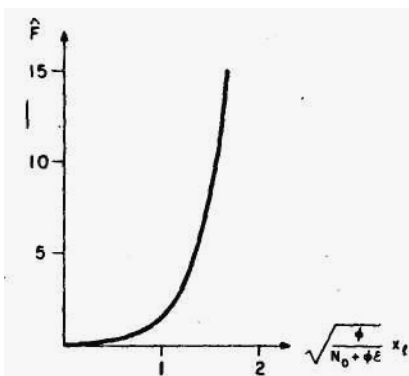


Рис. 5.11. Используемый тип нелинейности $F(\cdot)$

Следует отметить, что в отличие от выражений (5.22) и (5.23), где отсчеты x_{ik} , $k=0, \dots, K-1$, объединяются по линейному и квадратичному законам, для объединения отсчетов в (5.24) используются экстремальные нелинейности экспоненциального типа (рис. 5.11).

$$F(X_i) = \exp \left[\frac{\psi x_i^2}{N_0 + \psi E} \right] - 1 \quad (5.25)$$

Такая нелинейность подчеркивает сильные сигналы и подавляет слабые. В этом проявляется определенная

«самоадаптивность» приемника, в неявном виде присутствующая в (5.24). Следовательно, сигналы на выходе нелинейностей несут в себе информацию, непосредственно характеризующую апостериорные вероятности существования лучей. Согласно (5.24), эти сигналы далее взвешиваются в соответствии с априорными вероятностями размещения лучей, в неявном виде входящими в функцию $p(\bullet)$. Следует подчеркнуть, что изложенная позитивная роль нелинейности проявляется при относительно большой средней величине ОСШ, когда пики сигнала имеют большую величину по сравнению с шумом и уровнем боковых лепестков. При малом ОСШ эффект подавления шума нелинейностями будет выражаться слабее – тем слабее, чем меньше ОСШ. Можно показать, что в данном случае приемник сигналов с известными запаздываниями (например, приемник, реализующий соотношение (5.22) или (5.23)), в котором отсчеты выходных сигналов детекторов огибающей берутся только в моменты времени, соответствующие пикам сигналов, а вклады шума не учитываются, будет иметь более высокие характеристики, чем приемник сигналов с неизвестными запаздываниями.

5.6.4.4. Приемник ОФМ-сигналов

Как отмечалось выше, фазово-когерентные методы не рассматриваются из-за сложности когерентных приемников и из-за быстрых временных изменений фаз лучевых сигналов, принимаемых подвижной станцией. Однако разностно-когерентные методы оказываются вполне применимыми, когда на интервале между двумя последовательно переданными сигналами фазы лучей существенно не изменяются. Это — обычный для практики случай. И хотя структура оптимального разностно-когерентного приемника многолучевого сигнала нам не известна, она, по-видимому, должна быть весьма сходной со структурой приемника при однолучевом распространении сигнала (рис. 5.12).

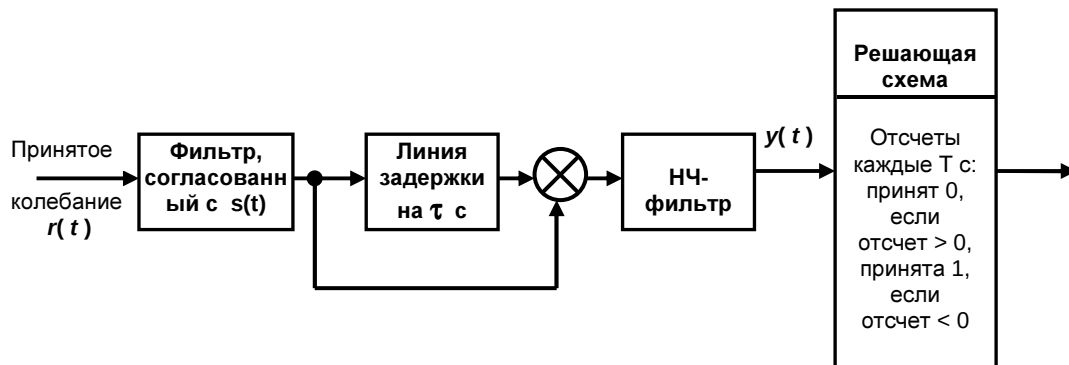


Рис. 5.12. Оптимальный приемник однолучевого ОФМ-сигнала

Предположим теперь, что в многолучевом случае, удовлетворяющем условиям упрощенной модели (распределенные по пуассоновскому закону запаздывания разрешимых лучей и независимые интенсивности), оптимальный приемник ОФМ-сигнала также связан с приемником, показанным на рис. 5.12, как и приемник, показанный на рис. 5.10, с приемником, показанным на рис. 5.8. Иными словами, мы будем полагать, что этот оптимальный приемник имеет вид, показанный на рис. 5.13.

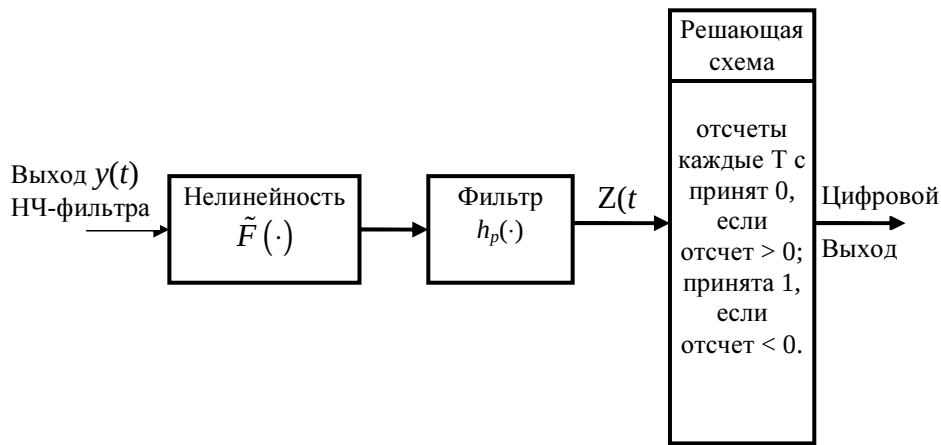


Рис. 5.13. Оптимальный приемник ОФМ-сигналов для многолучевого канала с неизвестными запаздываниями лучей.

На этом рисунке нелинейность

$$\tilde{F}(y) = F(\sqrt{|y|}) \text{sign}(y) \quad (5.26)$$

является биполярным вариантом нелинейности F , что обусловлено тем фактом, что входом этой нелинейности является квадрат огибающей выхода согласованного фильтра, а не сама огибающая, как на рис. 5.10. График функции $\tilde{F}$ для случая, когда F определяется выражением (5.26), показан на рис. 5.15 (ср. с рис. 5.11,б).

Решающая схема на рис. 5.13 с периодом T берет отсчеты сигнала на выходе интегрирующего фильтра $h_p(\bullet)$, с тем, чтобы определить экстремумы этого сигнала. Форма колебаний в различных точках приемника

показана на рис. 5.14. В отсутствие шума выход $y(t)$ детектора произведения в схеме на рис. 5.12 и, следовательно, вход нелинейности в схеме на рис. 5.13, будут приближенно иметь вид

$$y(t) = (-1)^d \sum_{k=0}^{K-1} a_k^2 |\gamma(t - 2T - t_k)|^2, \quad 2T \leq t \leq 2T + \Delta, \quad (5.27)$$

при передаче символа $d = 0$ или 1 , принятого на интервале $[T, 2T]$.

Пример такого «незашумленного» выхода $y(t)$ показан на рис. 5.14, $г$ для входной двоичной последовательности 101100. Выходы нелинейности F и фильтра $h_p(\bullet)$

$$h_p(t) = \begin{cases} 1, & 0 \leq t \leq \Delta; \\ 0 & \text{при других } t \end{cases} \quad (5.27)$$

показаны на рис. 5.14, $г$, $д$. Отсчеты выхода фильтра $h_p(\bullet)$ берутся в моменты времени $nT + \Delta$, $n=2, 3, \dots$, и используются для формирования выходных сигналов приемника, задержанных на $(T+\Delta)$ с

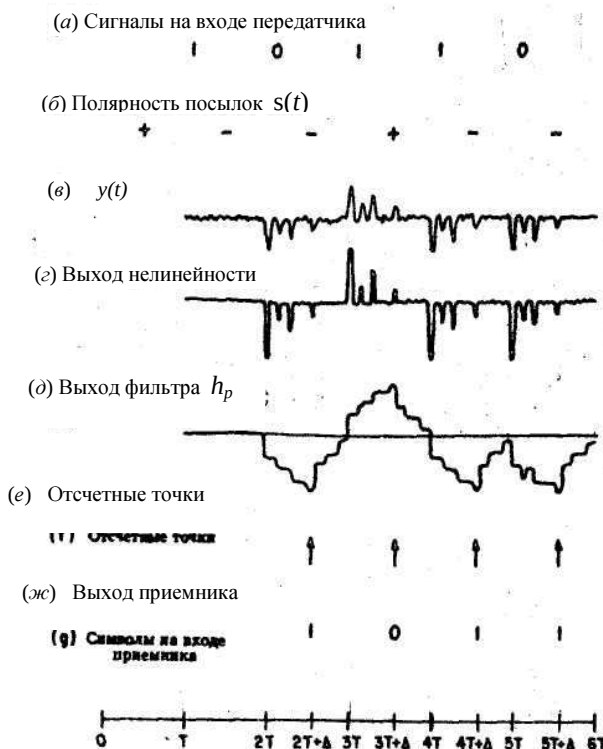


Рис. 5.14. Пример приема ОФМ-сигналов.

относительно соответствующих входных символов.

5.6.4.5. Приемники, зондирующие канал

До сих пор, рассматривая структуру приемника, мы исходили из априорных сведений о статистиках канала. Однако во многих случаях характеристики канала можно измерить с помощью зондирующих сигналов, что позволяет определить апостериорные статистики. В качестве зондирующих сигналов могут использоваться либо специальные сигналы, либо сами информационные сигналы; в последнем случае зондирование и передача данных осуществляются одновременно.

На первый взгляд может показаться, что для зондирования можно непосредственно использовать приемники, которые обсуждались выше при рассмотрении случая неизвестных запаздываний, а характеристики функций F и фильтра $h_p(\cdot)$ определять не по

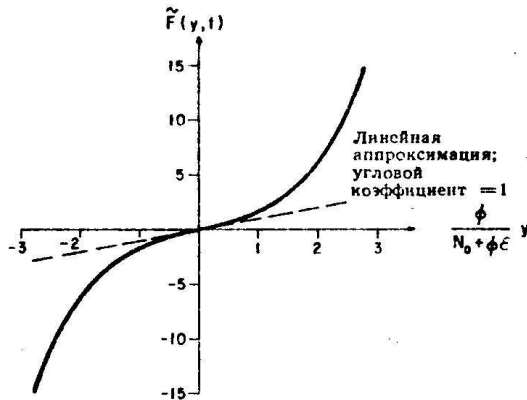


Рис. 5.15. Характеристика нелинейности

априорным, а по апостериорным статистикам. Действительно, такой подход уместен в отношении распределений интенсивностей лучей и нелинейностей F , которые они определяют.

Обычно используется другой подход, в котором предполагается, что в результате зондирования получаются очень точные оценки всех трех переменных $\{\tau_k\}$, $\{a_k\}$ и $\{\theta_k\}$ или хотя бы двух первых из них. Эти оценки полагаются точными и используются в приемнике в качестве параметров.

Ниже мы ограничимся рассмотрением некогерентных приемников с зондированием канала, в которых используются только оценки интенсивностей и запаздываний лучей.

«Оптимальный» приемник M -лучевого сигнала, в котором используются оценки интенсивностей и запаздываний, можно, очевидно, реализовать на основе выражения (5.22). Такой приемник должен вычислять величины

$$w_l = \sum_{k=0}^{K-1} \ln I_0 \left(\frac{2\hat{a}_k \hat{x}_{lk}}{N_0} \right) \quad (5.28)$$

где $\hat{a}_k$ — оценка интенсивности k -го луча, $\hat{x}_{lk}$ — отсчет огибающей сигнала на выходе l -го согласованного фильтра в момент времени $t = T + t_k$, а $\hat{t}_k$ — оценка запаздывания k -го луча. Индекс $l = 1, \dots, M$, для которого величина w_l максимальна, является цифровым выходом приемника. Если ОСШ для каждого луча достаточно велико, то (5.28) можно аппроксимировать выражением

$$w_l \cong \frac{2}{N_0} \sum_{k=0}^{K-1} \hat{a}_k \hat{x}_{lk}, \quad (5.29)$$

так как при больших x $\ln I_0(x) \approx x$. Выражение (5.29) описывает оптимальный линейный сумматор Бреннана [8], используемый при разнесенном приеме сигналов; мы также будем его использовать, поэтому остановимся на нем несколько подробнее.

Линейный сумматор, описываемый выражением (5.29), можно реализовать с помощью трансверсального фильтра, как показано на рис. 5.16. Этот фильтр состоит из линии задержки на время Δ , у которой сделаны отводы с интервалом, равным по крайней

мере $1/W$, так что всего имеется $W\Delta$ отводов. На вход трансверсального фильтра подается напряжение с выхода детектора огибающей, например $x_j(t)$ (см. рис. 5.10). Выход трансверсального фильтра представляет собой взвешенную сумму сигналов с определенных отводов, номера которых зависят от оценок запаздываний лучей.

Оценки t_k ($k=0, \dots, K-1$) используются для включения усилителей, подсоединенных к отводам, задержки которых (измеряемые с правого конца линии) наиболее точно аппроксимируют оценки $\hat{t}_k$. Иными словами, активируется K усилителей, коэффициенты усиления которых устанавливаются пропорционально величинам соответствующих оценок интенсивностей $\hat{a}_k$. На рис. 5.16 указаны коэффициенты усиления усилителей при приеме четырех лучевого сигнала.

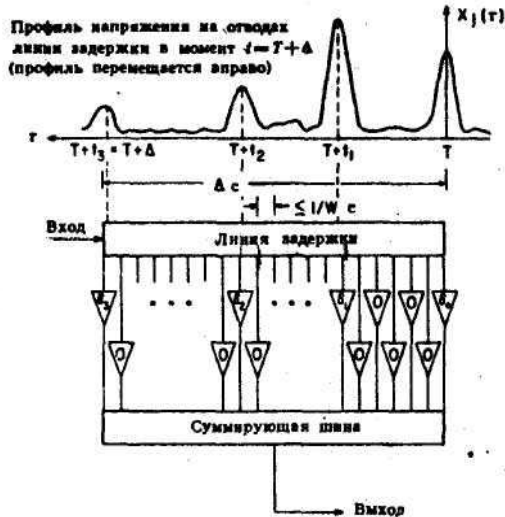


Рис. 5.16. Трансверсальный фильтр.
Число $t_0 = \Delta$, $1 \text{ дс } \Delta$

Для того чтобы лучше понять, как работает трансверсальный фильтр, на 5.16 показан профиль изменения напряжения вдоль линии задержки для случая, когда на ее вход поступает сигнал $x_j(t)$ с выхода детектора огибающей, включенного после j -го согласованного фильтра (см. рис. 5.9). Этот профиль постепенно сдвигается вправо, и на рис. 5.16 его положение соответствует моменту времени $t=T+\Delta$. В этот момент пики сигнала соответствуют задержкам $t_0, t_1, \dots, t_{K-1}$, измеряемым с правого конца линии. На рис. 5.16 изображен случай, когда $t_0=0$ и $K=4$, причем предполагается,

— максимальное относительное

запаздывание (т. е. превышение над t_0 , ожидаемое в

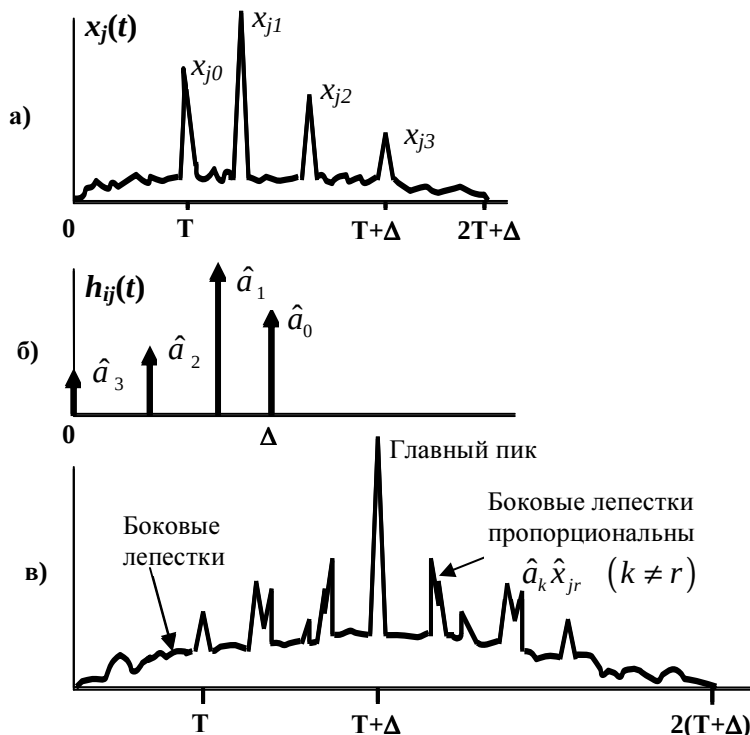


Рис. 5.17. (а) – вход фильтра; (б) – импульсная характеристика фильтра; (в) – Выход фильтра (свертка входа и импульсной характеристики фильтра)

показанных соответственно на рис. 5.17, а и 5.17, б, в результате чего формируется выходной сигнал, показанный на рис. 5.17, в. Главный пик этого сигнала соответствует

данном канале). Если $\hat{t}_k$ — точные оценки времен запаздывания, то в момент времени $t=T+\Delta$, напряжение на выходе трансверсального фильтра будет приблизительно равно величине:

$$\sum_{k=0}^{K-1} \hat{a}_k \hat{x}_j(T + \hat{t}_k) \cong \sum_{k=0}^{K-1} \hat{a}_k \hat{x}_{jk}, \quad (5.30)$$

т. е. оказывается пропорциональной величине w_l , определяемой (5.29) при $l=j$.

Таким образом, трансверсальный фильтр является фильтром, согласованным с оценкой значения низкочастотного эквивалента импульсной характеристики канала $h_c(t)$.

В процессе перемещения $x_j(t)$ на рис. 5.16 вправо трансверсальный фильтр осуществляет свертку $x_j(t)$ с $h_{ij}(t)$,

моменту времени $t=T+\Delta$, т. е. тому моменту, когда входной сигнал $x_j(t)$ занял такое положение, при котором его пики расположены точно над активированными отводами линии задержки. Высота главного лика пропорциональна величине (5.20). Выходной сигнал имеет также небольшие побочные пики, пропорциональные величинам $\hat{a}_k \hat{x}_{jr}$ ($k=0, 1, 2, 3; r \neq k$), и боковые лепестки, накопленные в результате свертки $h_{ij}(t)$ с боковыми лепестками сигнала $x_j(t)$. Ясно, что для получения (5.20) необходимо оставить только один главный пик.

Таким образом, полный приемник M -лучевого сигнала, в котором используются оценки характеристик канала, будет иметь вид, показанный на рис. 5.18. Выход $x_l(t)$, $l=1, \dots, M$ каждого детектора огибающей подается на вход трансверсального фильтра типа показанного на рис. 5.16, параметры которого изменяются в соответствии с оценками $\{\hat{t}_k\}$ и $\{\hat{a}_k\}$, полученными от зондирующего приемника. В момент времени $t= T+ \Delta$ выходы трансверсальных фильтров считываются и сравниваются, в результате чего определяется индекс наибольшего из них. Такой приемник получил название «Рейк» из-за внешнего сходства профиля напряжения на отводах трансверсального фильтра с зубьями гребня или грабель [гаке (рейк) — грабли].

Зондирующий приемник, показанный на рис. 5.18, также может иметь структуру, образованную последовательным соединением согласованного фильтра, детектора огибающей и линии задержки с отводами, причем согласованный фильтр согласуется с известным колебанием, которое может иметь большую энергию и длительность, чем информационные сигналы s_j . Перед передачей данных (или во время передач, если характеристики канала изменяются достаточно быстро) зондирующий приемник «слушает» свой сигнал. Когда сигнал принят, огибающая напряжения на выходе согласованного фильтра будет иметь множество пиков, т. е. будет иметь вид, аналогичный показанному на рис. 5.17, а. Затем это колебание будет перемещаться по линии задержки так же, как показано на рис. 5.16. Когда первый пик сигнала достигает правого конца линии, срабатывает пороговый элемент и берутся отсчеты напряжений на всех отводах линии задержки, и их значения, которые пропорциональны интенсивностям лучей,

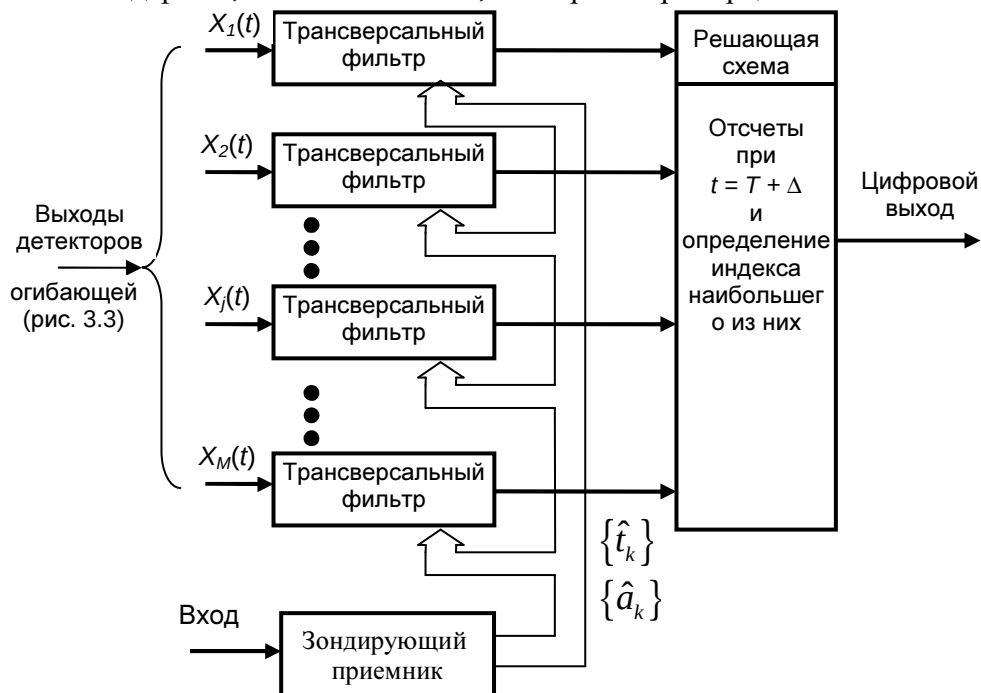


Рис. 5.18. Приемник многолучевого сигнала с зондированием канала (приемник системы «Рейк»).

запоминаются, если они превышают некоторую пороговую величину, в противном случае они полагаются равными нулю. Далее эти отсчеты используются для установки

коэффициентов усиления усилителей, подключенных к отводам трансверсальных фильтров приемника информационных сигналов.

В самонастраивающемся варианте этого приемника для зондирования канала используются сами информационные сигналы. В момент времени, когда принято решение о том, что принят, скажем, j -й сигнал, напряжения на отводах j -го трансверсального фильтра фиксируются точно так же, как это было описано выше для зондирующего приемника. Затем значения этих напряжений используются для установки коэффициентов усиления соответствующих усилителей во всех трансверсальных фильтрах.

Следует отметить, что механизм зондирования выполняет также и функцию синхронизации приемника, которую мы пока не рассматривали. И в приемнике с раздельным зондированием, и в самонастраивающемся приемнике при срабатывании порогового устройства, подключенного к крайнему правому отводу линии задержки, запускаются часы (генератор синхронизирующих импульсов), с помощью которых обеспечивается соответствующая синхронизация отсчетных импульсов, поступающих на входы решающей схемы.

Наконец, отметим, что приемник ОФМ-сигнала можно построить, используя те же принципы, которые были заложены в приемнике системы «Рейк».

Комбинирование сигналов на основе RAKE-приемника реализовано в стандарта CDMA-2000. Приемные устройства мобильной станции (МС) и базовой станции (БС) включают несколько (3 на МС и 4 на БС) параллельно работающих корреляторов, которые выделяют наиболее сильные сигналы. Выходы корреляторов сводятся к одному и тому же моменту времени и суммируются. Тем самым эффект многолучевого распространения используется для повышения качества связи.

5.6.4.6. Межсимвольная интерференция

До сих пор мы пренебрегали эффектами межсимвольной интерференции, поскольку основное внимание уделялось приемнику отдельных изолированных посылок (см., например, рис. 5.9). Даже в случае ОФМ-сигнала, когда информация передается посредством изменения или сохранения полярностей двух последовательных посылок, эти посылки полагались достаточно длинными или достаточно разнесенными по времени, благодаря чему временная дисперсия, обусловленная многолучевостью канала, приводила лишь к незначительному ухудшению рабочих характеристик или же вообще на них не влияла (см. рис. 5.17).

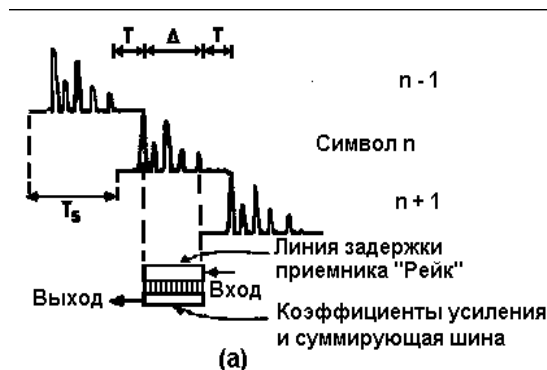


Рис. 5.19,а. Межсимвольная интерференция отсутствует

широкополосном случае ($TW \gg 1$), на рис. 5.19 показана последовательность огибающих откликов отдельного согласованного фильтра на периодически повторяемый входной сигнал при двух значениях T_s .

На рис. 5.19 а показаны огибающие откликов согласованного фильтра на три последовательные посылки для случая, когда $T_s \geq T + \Delta$. Для наглядности показаны также

Предположим теперь, что каждая посылка имеет длительность T , а последовательная передача любых посылок осуществляется с интервалом T_s . В соответствии с (5.21) отклики согласованных фильтров в рассмотренных выше приемниках на каждую посылку длиной T будут длиться $2T + \Delta$, что может привести к межсимвольной интерференции.

Для того чтобы наглядно проиллюстрировать возникновение межсимвольной интерференции в

линия задержки приемника «Рейк» длиной Δ с и сумматор (см. рис. 5.16). Межсимвольная интерференция в данном случае полностью отсутствует. Указанная последовательность огибающих выходных напряжений фильтра «заморожена» в тот момент времени, когда пики центрального (n -го) многолучевого входного сигнала точно соответствуют отводам линии задержки и «готовы» для взятия отсчетов. Заметим, что в данный момент ($n-1$)-й и ($n+1$)-й отклики «находятся» вне линии задержки, и, следовательно, не могут интерферировать с n -м откликом.

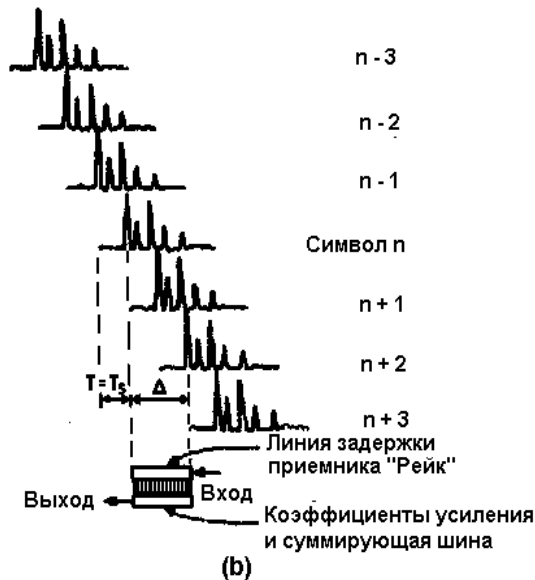


Рис. 5.19, б. Умеренная межсимвольная интерференция $T_s = T$; $\Delta = 2T$

На рис. 5.19, б показан случай умеренной межсимвольной интерференции. Здесь $T_s = T$, а $\Delta = 2T$, так что $T_s = (T + \Delta)/3$. Огибающие откликов показаны для семи последовательных посылок, которые зафиксированы в тот момент времени, когда пики n -го многолучевого входного сигнала точно соответствуют отводам линии задержки. Однако в данном случае два предшествующих ($(n-2)$ -й и $(n-1)$ -й) сигнала еще частично не вышли из линии задержки, а два последующих ($(n+1)$ -й и $(n+2)$ -й) уже частично вошли в нее, и это обстоятельство приводит к возникновению межсимвольной интерференции. В общем случае с каждым символом будет интерферировать $2[\Delta/T]$ предшествующих и последующих символов, где $[x]$ — наименьшее целое число, большее или равное x .

Для того чтобы полностью устранить межсимвольную интерференцию, необходимо выполнить условие $T_s \geq T + \Delta$. Поскольку величина Δ для данного канала фиксирована, то остается лишь несколько уменьшить величину T (однако выбор T зависит от требования $TW \gg 1$ и ограничений, налагаемых на ширину полосы W). Следовательно, в бинарных системах мы не можем полностью устранить межсимвольную интерференцию, если скорость передачи данных превышает величину, равную примерно $1/\Delta$ бит/с. Так, например, при работе такой системы в городских условиях (эффективное значение $\Delta = 5$ мкс) со скоростью, большей примерно 200 кбит/с, межсимвольная интерференция будет возникать даже в том случае, когда используются широкополосные сигналы.

Используемый в данном случае подход основан на том факте, что, хотя в момент принятия решения о принятом символе, многолучевые сигналы, соответствующие предыдущему и последующему символам, частично находятся в линии задержки (см. рис. 5.19 б), вероятность того, что их пики будут точно соответствовать активированным отводам линии, весьма мала. Напомним, что отводы у линии задержки сделаны через каждые $1/W$ с. Так как $TW \gg 1$ и так как мы теперь полагаем $\Delta/T > 1$, то $W\Delta \gg 1$ и это означает, что линия задержки должна иметь большое

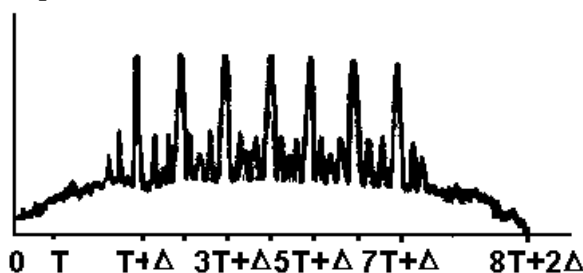


Рис. 5.20. Свертка семи лучей.

число отводов. Например, при $TW = 100$ и $\Delta/T = 5$ линия задержки будет иметь $W\Delta = 500$ отводов. Однако при приеме сигнала активируются только те отводы, на которых ожидаются пики этого сигнала, поэтому в рассмотренном выше примере даже при наличии, скажем, 20 лучей будет активироваться всего $20/500 = 2,5\%$ от общего числа отводов. Эти отводы будут соответствовать пикам центрального n -го отклика на рис. 5.19 б; при

этом весьма маловероятно, что пики смежных с ним откликов также попадут на эти активированные отводы.

Последовательность импульсов после семи таких сверток и соответствующего их объединения будет иметь примерно такой вид, как показано на рис. 5.20. Каждый из главных пиков соответствует здесь точному согласованию одной последовательности импульсов с активированными отводами линии задержки, и именно по ним берутся далее отсчеты. Пьедестал, на котором расположены эти пики, состоит из таких же небольших побочных пиков, как и на рис. 5.17, и смеси боковых лепестков. Этот пьедестал мы будем называть «помехой, обусловленной многолучевостью». Способность приемника системы «Рейк» выделять именно ту часть отклика согласованного фильтра, которая соответствует принимаемому в данный момент символу, обусловлена структурой применяемой в нем линии задержки с отводами.

Суммируя изложенное выше, нетрудно видеть, что приемники системы «Рейк» могут работать со значительно более высокими, чем $1/\Delta$ бит/с, скоростями передачи данных без заметного ухудшения характеристик.

Ниже будут рассмотрены другие приемники многолучевых ОФМ-сигналов, структура которых проще приемника «Рейк».

5.6.4.7. Приемник с последетекторным интегрированием сигналов

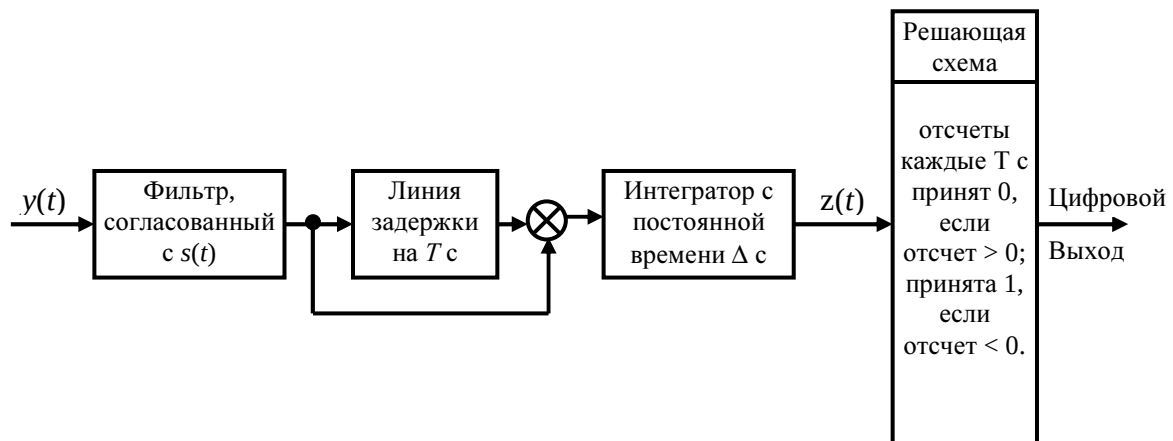


Рис. 5.21. ПДИ – квазиоптимальный приемник ОФМ-сигнала для неизвестных времен запаздывания лучей.

Роль фильтра, согласованного с каналом, в приемнике с последетекторным интегрированием (ПДИ-приемник, или просто ПДИ), представленным рис. 5.21, играет интегратор с постоянной времени интегрирования Δ , импульсная характеристика которого имеет вид (5.27а). Эта характеристика предполагает, что запаздывания лучей равномерно распределены на интервале $[0, \Delta]$ и $p(t_k)=1/\Delta$, $0 \leq t_k \leq \Delta$. В действительности запаздывания лучей распределены неравномерно, поэтому характеристика $h_p(t)$, определяемая выражением (5.26а), будет субоптимальной. Так как НЧ-фильтр на рис. 5.12 необходим лишь для устранения составляющих удвоенной частоты на выходе умножителя, то на рис. 5.21 его функция возложена на интегрирующий фильтр, который также является фильтром низких частот.

Вход решающей схемы на рис. 5.21 будет все еще иметь вид, сходный с показанным на рис. 5.14 д, но теперь он будет получаться в результате интегрирования напряжения $u(t)$ (рис. 5.14 в), а не напряжения, показанного на рис. 5.14 г. Отсчеты выхода интегрирующего фильтра будут, как и прежде, браться в моменты времени, указанные на рис. 5.14 е.

Расчет характеристик ПДИ проводился в предположении, что:

- передаваемые сигналы ортогональны и имеют нулевые боковые лепестки;

- нелинейность на рис. 5.13 имеет квадратичный характер;
- все лучи независимы и их интенсивности имеют одинаковые релеевские распределения;
- все лучи разрешимы;
- $\Delta \ll T$, МСИ отсутствует;
- в канале присутствует аддитивный белый шум с односторонней плотностью мощности N_0 .

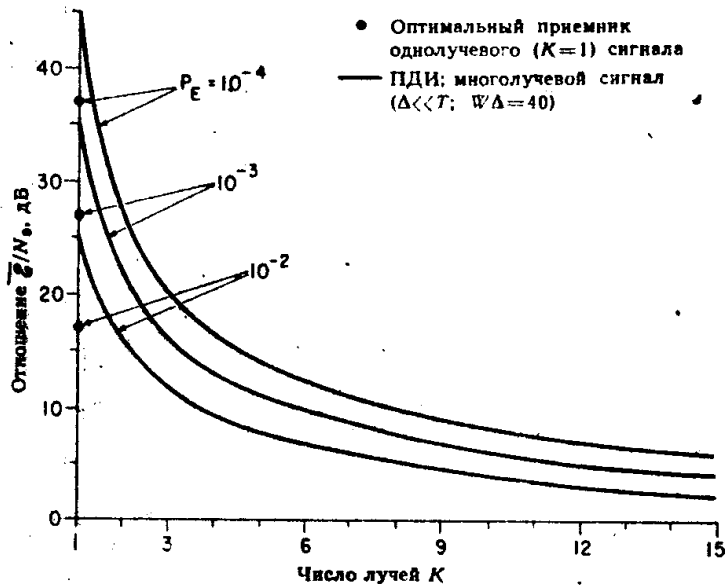


Рис. 5.22 Зависимость среднего на 1 луч отношения $\bar{S}/N_0$ для ОФМ-приемника с последетекторным интегрированием сигнала, требуемого для обеспечения заданной вероятности ошибки на 1 бит, от числа лучей K (независимые лучи с релеевскими замираниями).

Результаты расчета для $W\Delta = 40$ представлены на рис. 5.22.

Поскольку ранее принималось, что для ПДИ $\Delta \ll T$, то это, в свою очередь означает, что коэффициент расширения спектра сигнала $TW \gg 40$.

Для сравнения на рис. 5.22 указаны значения вероятности ошибки для оптимального однолучевого ОФМ-приемника. Нетрудно видеть, что при наличии ряда релеевских лучей с близкими среднеквадратичными интенсивностями, ПДИ значительно (на 5 – 20 дБ) превосходит приемник, который принимает сигнал по одному лучу (например, по ЛВЗ) и пренебрегает другими

лучами. Это сравнение весьма убедительно иллюстрирует улучшение характеристик при разнесенном приеме, которое может обеспечить многолучевой канал.

5.6.4.8. Цифровой вариант приемника «Рейк»

В приемниках с зондированием канала (рис. 5.18 и 5.17) установка коэффициентов усиления усилителей, подключенных к отводам трансверсального фильтра (или фильтров) осуществляется в соответствии с оценками интенсивностей лучей. Однако для реализации значительно проще приемник (правда, он будет иметь несколько худшие характеристики), в котором отводы трансверсального фильтра просто подключаются к суммирующей шине в соответствии с текущими оценками времен запаздывания лучей. В этом случае необходимость получения оценок интенсивностей лучей отпадает, и зондирующий приемник должен лишь определять оценки времен запаздывания лучей $\{\hat{t}_k\}$. Такой упрощенный вариант приемника мы будем называть «цифровым Рейк». Структура этого приемника при использовании ОФМ-сигналов показана на рис. 5.23, где в целях дальнейшего упрощения исключен также и вычислитель квадратного корня.

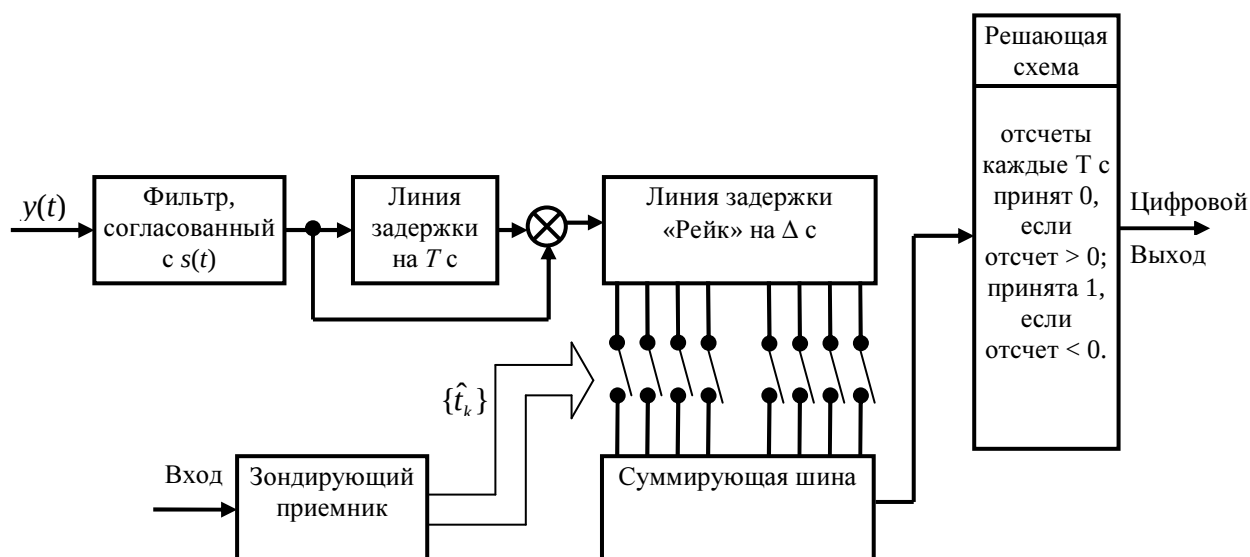


Рис. 5.23. Цифровой вариант приемника «Рейк» – квазиоптимальный приемник ОФМ-сигнала, в котором используются оценки времен запаздывания лучей.

Оценки запаздываний лучей, поступающие от зондирующего приемника, используются здесь для замыкания соответствующих ключей в цепях отводов трансверсального фильтра, в результате чего эти отводы подключаются к суммирующей шине. Заметим, что при замыкании всех ключей приемник, показанный из рис. 5.23, оказывается примерно аналогичным ФМ-приемнику, показанному на рис. 5.21, у которого операция интегрирования на интервале Δ с

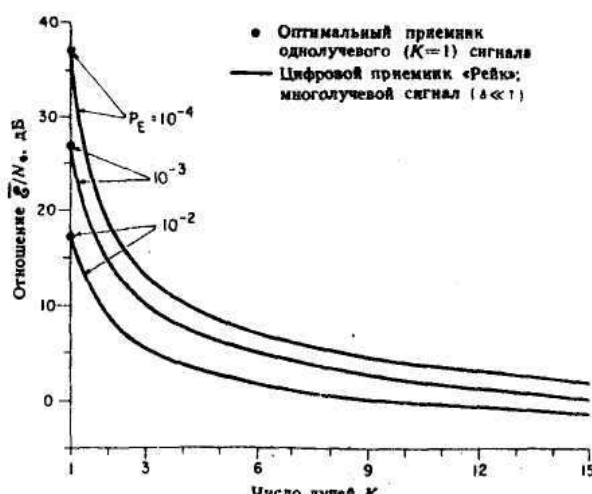


Рис. 5.24. Зависимость среднего на 1 луч ОСШ для цифрового Рейк от числа лучей, требуемого для обеспечения заданной вероятности ошибки на 1 бит.

заменена операцией суммирования в дискретных точках этого интервала. Характеристики этого приемника, аналогичные кривым рис. 5.22, представлены на рис. 5.24 для случая релеевских замираний без учета межлучевой и межсимвольной интерференции, наличие которой можно учесть введением дополнительного эквивалентного шума.

5.6.4.9. Установка пакетной радиосвязи (УПР)

Пакетные радиосети (ПРС), начавшие развиваться в 1960-е годы, получили широкое распространение в коммерческой, военной и научной областях. В 1972 году Управление перспективных научно-исследовательских работ Министерства обороны США (DARPA) – то самое агентство, которое стояло у истоков Internet – приступило к проведению исследований с целью создания собственной пакетной радиосети. В их задачу входил поиск эффективных методов коллективного использования выделенного радиоканала в условиях неполной или непостоянной связности, характерной для связи с подвижными объектами. В ходе разработки потребовалось решить целый ряд технических проблем, начиная с объединения высокоскоростных модемов и высокопроизводительных микропроцессоров в малогабаритном недорогом блоке с целью реализации протоколов коллективного использования канала и кончая созданием алгоритмов маршрутизации пакетов в сетях с быстроменяющимися характеристиками.

Ниже будет рассмотрен приемопередатчик УПР [8], созданной в рамках этого проекта, а точнее – обработку в нем принимаемых широкополосных многолучевых сигналов. В четвертой главе было показано, что когерентный прием сигналов с абсолютной фазовой манипуляцией наталкивается на большие, практически непреодолимые, трудности, связанные с формированием опорного колебания, фаза которого однозначно связана с фазой принимаемого сигнала. Поэтому в контексте настоящего пособия весьма уместно познакомиться с решением этой проблемы, найденной в [8].

Не касаясь других, подчас довольно интересных, технических решений и характеристик указанной УПР, детально описанных в [8], рассмотрим демодулятор этой установки (рис. 5.25).

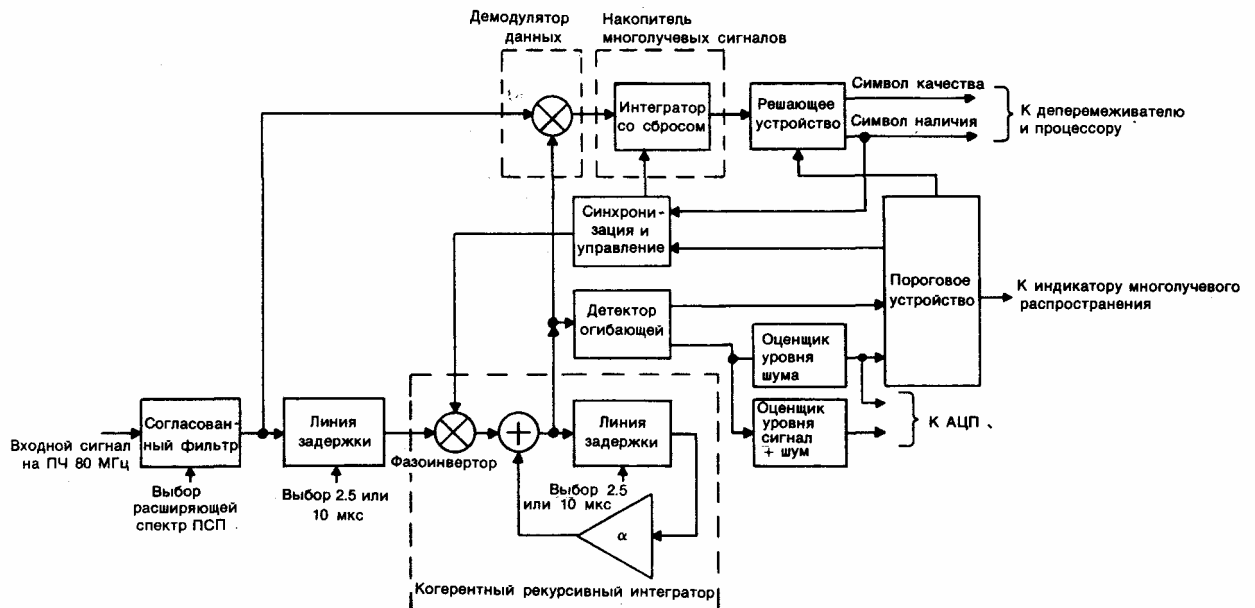


Рис. 5.25. Функциональная схема демодулятора УПР.

Приведем некоторые характеристики УПР, необходимые для понимания работы демодулятора. УПР использует классические сигналы БФМ, спектр которых расширен кодовой последовательностью (ПСП) с тактовой частотой 12,8 МГц. Кодовая последовательность может меняться от символа к символу; информация о текущем коде содержится в преамбуле пакета. Передача осуществляется на двух скоростях: 400 Кбит/с (высокая, длительность символа 2,5 мкс) и 100 Кбит/с (низкая, длительность символа 10 мкс). Приемник с двойным преобразованием частоты имеет вторую ПЧ, равную 80 МГц.

В демодуляторе с помощью программируемого согласованного фильтра (СФ) снимают модуляцию элементарными символами ПСП и получают сжатый импульс с фазовой манипуляцией информационными символами пакета (преамбула, заголовок, данные и т. д.).

1) *Согласованный фильтр НУПР.* Структурная схема реализованного программируемого согласованного фильтра НУПР приведена на рис. 5.26. Входной фильтр для приема многоэлементной ПСП, согласованный с длительностью элементарного символа, соединен с многоотводной линией задержки, на соседних выходах которой сигналы появляются с задержкой, равной длительности элемента ПСП

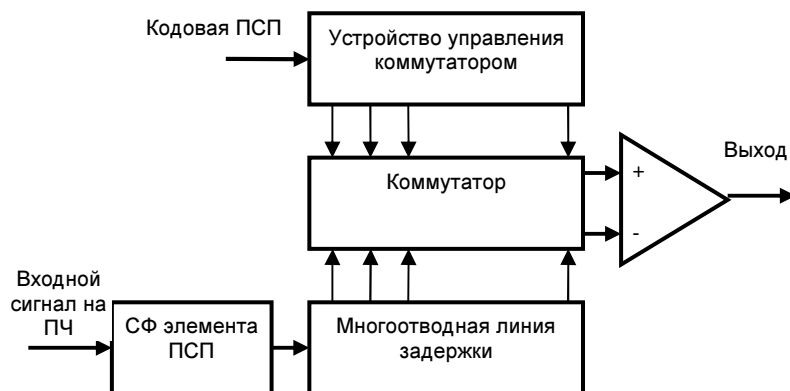


Рис. 5.26. Программируемый согласованный фильтр

(78.125 нс). Для формирования выходного сигнала СФ сигналы, поступающие с отводов линии задержки, суммируют в шинах с весами +1 или -1 в соответствии со структурой ожидаемой кодовой ПСП, а выходы двух шин вычитаются. Многоотводная линия задержки согласованного фильтра реализована с использованием техники ПАВ на четырех

интегральных схемах, каждая из которых имеет 32 отвода. При скорости передачи данных 100 тыс. символов в секунду задействованы все 128 отводов линии задержки и выходы четырех интегральных схем объединяются вместе. При скорости передачи данных 400 тыс. символов в секунду обрабатываются сигналы с 32 отводов линии задержки и используется выходной сигнал лишь первой интегральной схемы. Выходной сигнал СФ поступает на демодулятор данных и — после задержки на длительность одного символа — на когерентный рекурсивный интегратор (КРИ) с решающей обратной связью.

2) *Когерентный рекурсивный интегратор УПР.* КРИ обеспечивает формирование опорного колебания несущей частоты и сигналы профиля многолучевости (зависящие от структуры принимаемого многолучевого сигнала). Опорное колебание необходимо для когерентной демодуляции принимаемого радиосигнала в демодуляторе данных и после детектирования огибающей для определения факта захвата сигнала. Требуемое качество выполнения этих функций обеспечивается за счет увеличения ОСШ на выходе КРИ по сравнению с ОСШ на его входе при наличии широкополосных помех. Увеличение ОСШ зависит от коэффициента усиления α контура КРИ и равно $(1+\alpha)/(1-\alpha)$. При $\alpha=0.85$ ОСШ повышается на 10.9 дБ.

3) *Накопление многолучевых сигналов и демодуляция данных.* Процессор накопления многолучевых сигналов содержит КРИ, демодулятор данных, интегратор со сбросом и следующее за ним устройство принятия решения об информационном символе. Линия задержки перед фазоинвертором КРИ необходима в связи с тем, что сигнал решения формируется в конце принятого символа. Выходной сигнал КРИ соответствует профилю многолучевости канала на радиочастоте (с точностью до фазы) и используется как когерентное опорное колебание для демодуляции данных. В демодуляторе данных происходит умножение этого опорного колебания на выходной сигнал СФ. Применение сигнала с расширенным спектром в условиях многолучевого распространения радиоволн позволяет выделить с помощью СФ сигналы отдельных лучей, взаимная задержка прихода которых приблизительно равна длительности одного элемента ПСП или около 78 нс. Взвешенные сигналы отдельных лучей суммируются в интеграторе со сбросом. Выполнение операций перемножения и интегрирования позволяет реализовать

почти оптимальный коррелятор для приема входного сигнала, искаженного многолучевостью. Такая обработка сигнала обеспечивает повышение помехоустойчивости в условиях замираний при многолучевом распространении радиоволн и взаимной задержки прихода сигналов отдельных лучей до 6 мкс при скорости передачи данных 100 тыс. символов в секунду и до 1.6 мкс при скорости передачи данных 400 тыс. символов в секунду.

В каждый момент принятия решения с выхода решающего устройства на вход микропроцессора УПР поступают два двоичных символа: символ наличия (т. е. решение о переданном сигнале) и символ качества. Символ наличия также поступает по цепи обратной связи на вход фазового инвертора (перемножителя) КРИ, в результате на выходе инвертора получают немодулированный импульсный сигнал, который подается на вход интегратора. Символ качества несет информацию об амплитуде принятого информационного символа и используется при работе в режиме последовательного декодирования с мягким (многоальтернативным) решением. Формируемый устройством оценки шума показатель уровня шума служит для выбора порога при формировании символа качества, значение которого используется сетью.

Таким образом, КРИ формирует опорное напряжение для фазового детектора демодулятора данных в виде текущего радиочастотного профиля многолучевости. Темп обновления опорного сигнала (профиля многолучевости) зависит от коэффициента передачи цепи обратной связи КРИ α . Так, при $\alpha = 0,85$ уровень предыдущего профиля многолучевости спадает до $\sim 0,2$ за 10 принятых символов. Демодулятор УПР обеспечивает почти оптимальный прием многолучевого сигнала БФМ; потери возникают из-за остаточной зашумленности опорного сигнала, но они невелики при выигрыше 10,9 дБ в схеме КРИ.

Ошибка в приеме символа не влечет за собой фатальных последствий, а лишь уменьшает уровень накопленного профиля многолучевости. Критическим является пакет ошибок кратностью 10 и более, но благодаря перемежению символов в передатчике и деперемежению в приемнике этой ситуации удастся избежать.

6. ОРГАНИЗАЦИЯ МНОЖЕСТВЕННОГО ДОСТУПА

Понятие множественного доступа подразумевает совокупность мер по обеспечению возможности параллельной работы многих пользователей в рамках частотно-временного ресурса, выделенного данной системе. Отождествляя каждого абонента с некоторым физическим каналом, можно сказать, что конкретная технология множественного доступа есть способ распределения ограниченного частотно-временного ресурса между каналами.

По способу организации множественного доступа, т. е. технологии распределения между отдельными каналами связи частотно-временного ресурса, выделяют СМР на основе одной из трех конкурирующих технологий:

- множественный доступ с частотным разделением каналов (МДЧР, англоязычная аббревиатура FDMA – *frequency division multiple access*);
- множественный доступ с временным разделением каналов (МДВР или TDMA – *time division multiple access*);
- множественный доступ с кодовым разделением каналов (МДКР или CDMA – *code division multiple access*);

6.1. Множественный доступ с частотным разделением каналов

Системы множественного доступа с частотным разделением каналов строятся таким образом, что каждому абоненту на время сеанса связи выделяется частотный канал с полосой Δf_a в пределах общего частотного диапазона системы Δf_p , не совпадающий ни с одним из каналов, уже предоставленных активным абонентам. При этом спектры канальных сигналов не перекрываются (рис. 6.1, а), чем и обеспечивается выполнение условия ортогональности. Прямоугольник со сторонами Δf_p и T_p на рис. 6.1, б характеризует общий частотно-временной ресурс системы. Как видно, при FDMA этот общий ресурс "нарезается" на K (по числу абонентов) горизонтальных полосок, занимающих каждая весь доступный временной ресурс и только K -ю часть частотного.

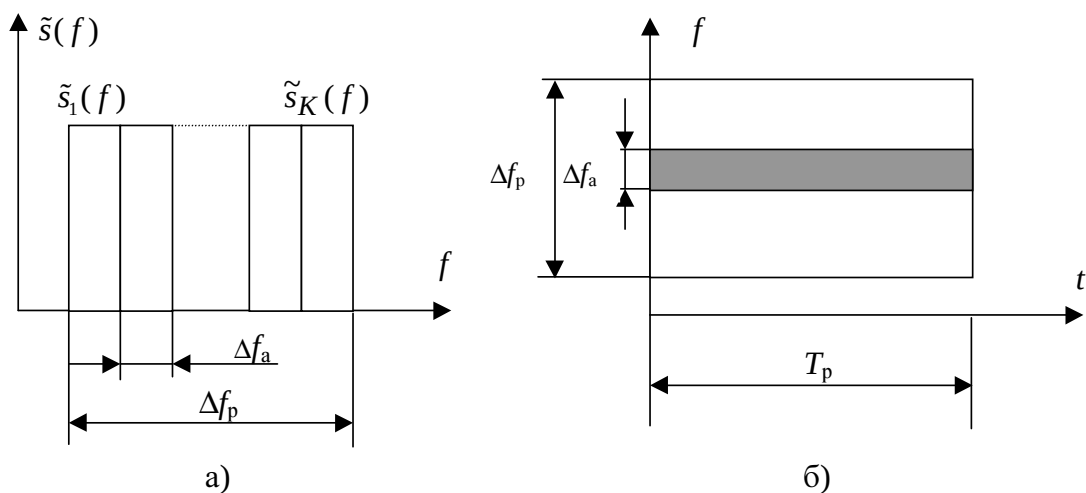


Рис. 6.1. Спектры канальных сигналов (а) и распределение частотно-временного ресурса между абонентами (б)

Спектральный разнос абонентских каналов полностью исключает влияние каналов друг на друга лишь теоретически. На практике же избежать возникновения межканальных

(внутрисистемных) помех невозможно, например, из-за неидеальности разделительных фильтров в приемнике, в результате которой часть энергии сигнала одного канала просачивается в соседний. Ослабить влияние межканальных помех можно соответствующим выбором манипуляции сигналов (уменьшением "внеполосных" излучений) и фильтров (улучшением подавления в соседнем канале). Еще одним способом снижения уровня взаимных помех является введение защитных интервалов между частотными каналами, что, однако, приводит к уменьшению полосы частот, используемой для связи, т.е. снижению эффективности использования спектра.

Если положить $\Delta f_a = 1/T_b = R_t$, где по-прежнему T_b - длительность одного бита передаваемой информации, а R_t - скорость передачи информации, то возможное число каналов связи для FDMA систем определится соотношением

$$K = \frac{\Delta f_p}{\Delta f_a} = \frac{\Delta f_p}{R_t}. \quad (6.1)$$

6.2. Множественный доступ с временным разделением каналов

Множественный доступ с временным разделением в традиционном понимании заключается в том, что каждому абоненту системы на время сеанса связи выделяется временной интервал T_a (временной канал) в пределах общего временного ресурса системы T_p (цикла или кадра системы), не совпадающий ни с одним из интервалов, уже предоставленных другим активным абонентам. Тем самым каждый канальный сигнал размещается в своем индивидуальном окне (*сlotte*) без перекрытия с другими (см. рис. 6.2, а). Вместе с тем спектры сигналов абонентов могут занимать всю выделенную системе полосу частот Δf_p и полностью перекрываться. Иллюстрацией подобного ресурсного распределения служит рис. 6.2, б, из которого видно, что суммарный частотно-временной ресурс "нарезан" в виде K вертикальных полос, каждая из которых занимает весь доступный частотный диапазон и лишь K -ю часть отведенного времени.

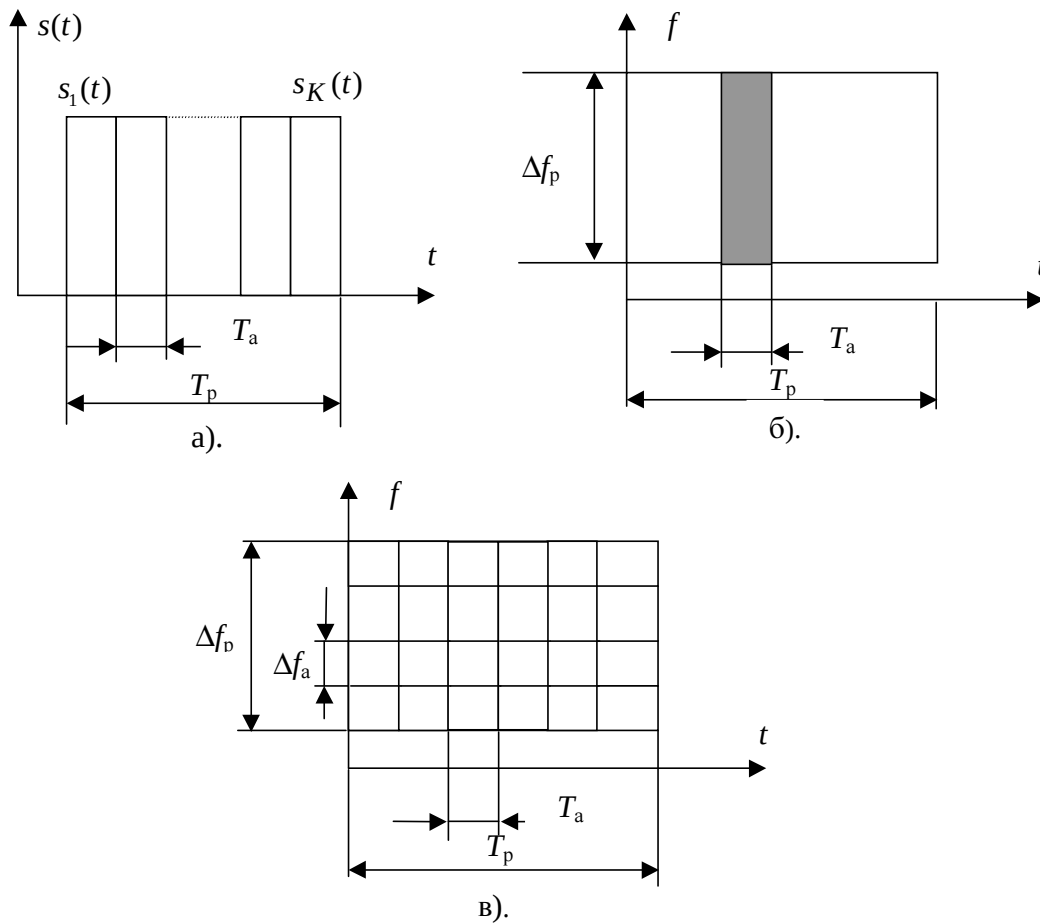


Рис. 6.2. Расположение канальных сигналов во времени (а),
распределение частотно-временного ресурса между абонентами (б) и
комбинация частотного и временного разделения FDMA/TDMA (в).

В идеале несовпадение канальных сигналов во времени обеспечивает их ортогональность, а значит и исключает влияние друг на друга. На деле из-за ограниченности полосы системы переходные процессы от сигналов предыдущих каналов к началу появления последующих могут не закончиться и, складываясь с последними, создавать перекрестные (межканальные) помехи. Уменьшить влияние соседних каналов, т. е. уровень межканальных помех, удастся способом, аналогичным описанному в предыдущем подразделе – введением защитных временных интервалов, что, в свою очередь, приводит к уменьшению временного диапазона, в течение которого возможна передача информации, т.е. фактическому снижению скорости передачи.

Оценим возможное число TDMA-каналов. При скорости передачи информации по одному каналу R_t за время одного цикла работы системы может быть передано $R_t T_p$ бит информации. Тогда общее количество информации, переданной во всех абонентских каналах, составит величину $K R_t T_p$, где $K = \frac{T_p}{T_a}$. Отсюда длительность одного бита информации определится как

$$T_6 = \frac{T_p}{K R_t T_p} = \frac{1}{K R_t} = \frac{1}{\Delta f_p},$$

так что $K = \frac{\Delta f_p}{R_t}$, что полностью совпадает с оценкой (6.1) числа каналов для систем с FDMA.

6.3. Множественный доступ с кодовым разделением каналов

6.3.1. Варианты реализации кодового разделения

В основе множественного доступа с кодовым разделением (МДКР или CDMA) лежит ориентация на широкополосную (*spread spectrum*) идеологию построения систем передачи информации, предусматривающую сознательное и многократное расширение полосы передаваемого сообщения по сравнению с той, которая характерна для традиционных узкополосных систем. Искусственное расширение спектра в подобных системах, как правило, реализуется одним из двух основных способов (см. главу 5):

- прямое расширение – *direct sequence spread spectrum* (DSSS);
- скачкообразное изменение несущей частоты – *frequency hop spread spectrum* (FHSS).

Процедура прямого расширения спектра не ухудшает помехоустойчивости двоичной передачи в гауссовском канале, оставляя противоположными сигналы, отвечающие значениям 0 и 1 передаваемого бита.

При использовании второго способа расширения спектра каждый символ информационного сообщения должен передаваться с помощью набора дискретных частот, задаваемого определенной последовательностью.

В существующих и разрабатываемых на перспективу системах многоадресной связи преимущественно применяется прямое расширение спектра, реализуемое либо в синхронном, либо в асинхронном варианте. Различия этих двух модификаций DSSS весьма значительны. Первая может быть применена тогда, когда есть возможность синхронизировать между собой все индивидуальные адресные последовательности (*сигнатуры*), присвоенные отдельным абонентам, так чтобы на приемной стороне сигналы разных абонентов не имели взаимных временных сдвигов.

Преимущества CDMA по отношению к FDMA и TDMA можно условно разбить на две группы. Первую из них составляют те, которые отличают любые широкополосные (*spread spectrum*) системы: высокая помехоустойчивость к сосредоточенным и широкополосным помехам (в том числе преднамеренным), возможность эффективной работы в условиях многолучевого распространения, широкий диапазон доступных мер криптозащиты, высокая точность измерения частотно-временных параметров, хорошая электромагнитная совместимость с системами радиосвязи и вещания и др. Вторая группа непосредственно связана с аспектами множественного доступа: большая абонентская емкость на соту (сектор) в системах сотовой связи, "мягкий" характер снижения качества связи при возрастании интенсивности трафика, простота реализации режима "мягкой" эстафетной передачи.

7. Системы мобильной связи

Стратегической целью развития современных общедоступных телекоммуникаций является предоставление высококачественной связи любому потребителю в любое время в любой точке земного шара. В соответствии с этим Международной организацией по стандартизации (ISO – *International Standard Organisation*) разработана концепция глобальной сети связи общего пользования. Под сетью понимается множество взаимодействующих друг с другом систем, осуществляющих хранение, обработку и передачу разнообразной информации.

Глобальная сеть объединяет телефонные сети общего пользования, сети передачи данных, системы мобильной и спутниковой связи. Ею охватывается все многообразие пользовательских терминалов (проводные и беспроводные телефонные аппараты, персональные компьютеры, сотовые и спутниковые телефоны, аппараты факсимильной связи и пр.), линий связи (кабельных, оптических, беспроводных и т.д.), узлов коммутации (учрежденческих, районных, междугородних), систем управления, контроля и технического обслуживания. Для того, чтобы составить адекватное представление о наиболее общих чертах и архитектуре столь сложного и разветвленного образования, разумно вписать его в рамки некоторой абстрактной модели, классифицирующей элементы сети по их функциональному назначению и механизмам взаимодействия.

7.1. Структура сетей мобильной связи

Основные, наиболее общие элементы сети связи показаны на рис. 7.1. Они содержат мобильные станции (МС), базовые станции (БС), центры управления и коммутации (ЦК), радиоинтерфейсы (РИ). МС используются подвижными абонентами. БС – стационарные приемопередающие станции, обеспечивающие радиосвязь с МС в определенной зоне (ячейке, соте, секторе). В ЦК выбирается направление передачи данных (коммутация),

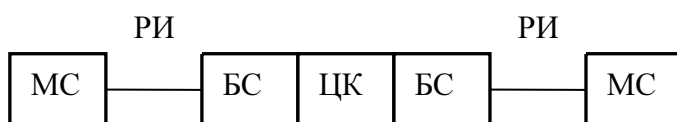


Рис. 7.1. Элементы сетей связи

производится управление БС и осуществляется доступ к внешним системам.

МС могут обмениваться информацией друг с другом через ЦК (рис. 7.1). Используя

основные элементы, можно построить региональные и национальные сети (подсети) мобильной связи.

Отдельные *сети и подсети* объединяются в единую структуру с помощью межсетевых интерфейсов (СИ) – кабельных или радиолиний. Контроллеры базовых станций (КБС) выполняют функции коммутации, управления БС, обеспечивают доступ к внешним сетям (телефонным, передачи данных). Они могут подразделяться на главные (с расширенными возможностями) и подчиненные (с ограниченными функциями). Подсеть может содержать собственную станцию управления (СУ). Таким образом, функции управления и коммутации распределяются по всей сети, что обеспечивает быструю передачу вызовов и сохранение работоспособности сети при неисправностях отдельного оборудования.

7.2. Классификация и эволюция систем мобильной радиосвязи

7.2.1. Основные типы систем мобильной связи

Приведем в кратком изложении следующий основной набор классификационных признаков систем мобильной связи (СМР):

- способ управления системой, иначе способ объединения абонентов – *централизованный* (координированный) или *автономный* (некоординированный). При централизованном объединении связь между абонентами производится через центральные (или базовые) станции. В противном случае связь между пользователями устанавливается непосредственно, без участия базовых станций;
- зона обслуживания – *радиальная* (в пределах радиуса действия радиостанции), *линейная* (для линейно протяженных зон), *территориальная* (для определенных конфигураций территории);
- направленность связи – *односторонняя* или *двусторонняя* связь между абонентом и базовой станцией;
- вид работы системы – *симплекс* (поочередная передача от абонента к базовой станции и обратно) или *дуплекс* - одновременная передача и прием в каждом из двух названных направлений;
- метод разделения каналов в системе радиосвязи, или метод множественного доступа – *частотный*, *временной* или *кодировый*;
- способ использования частотного ресурса, выделенного системе связи – жесткое закрепление каналов за абонентами, возможность доступа абонентов к общему частотному ресурсу (транкинговые системы), повторное использование частот за счет пространственного разнесения передатчиков (сотовые системы);
- категория обслуживаемых системой связи абонентов – профессиональные (служебные, корпоративные) абоненты, частные лица;
- вид передаваемой информации – речь, кодированное сообщение и др.

В зависимости от назначения системы, объема предоставляемых услуг и размеров зоны обслуживания можно выделить следующие четыре типа СМР:

- транкинговые системы связи (ТСС);
- системы персонального радиовызова (СПРВ);
- системы персональной спутниковой связи (СПСС);
- сотовые системы мобильной связи (ССМС).

7.2.2. Транкинговые системы связи

Термин "*транкинговая*" (или *транковая*) связь происходит от английского слова *trunk* (ствол) и отражает то обстоятельство, что "ствол связи" содержит несколько каналов, причем жесткое закрепление каналов за абонентами отсутствует. В литературе можно найти различные определения транкинговых систем, общим для которых является именно предоставление в распоряжение абонента одного из свободных на данный момент каналов.

Типичной сферой применения транкинговых систем являются государственные, ведомственные, корпоративные организации и институты, такие как служба скорой помощи, пожарная служба, охрана правопорядка, органы безопасности, различные коммерческие структуры и др. По большей части транкинговые системы используются как средства оперативной связи с жестко лимитированным и постоянно контролируемым контингентом абонентов в пределах ограниченной территориальной зоны.

Транкинговые системы связи могут строиться как системы с однозоновой или многозоновой структурой. Принимая во внимание специфический характер ТСС, т.е. ограниченность числа пользователей системы, переход от однозоновой к многозоновой структуре объясняется в первую очередь расширением географической зоны действия системы, а не стремлением к повышению числа абонентов (абонентской емкости) системы. При пересечении границ радиопокрытия ТСС отслеживают перемещение абонентов, обеспечивают их регистрацию, и назначение им нового частотного канала. Однако, как правило, подобный переход происходит с прерыванием связи, для восстановления которой абонентам необходимо произвести повторный вызов.

Транкинговые системы могут использовать как симплексные, так и дуплексные каналы радиосвязи, однако с целью упрощения и удешевления в них нередко применяется полудуплексный режим работы, при котором один и тот же канал поочередно используется для связи от центра управления (базовой станции) к абоненту и в обратном направлении.

Реализация принципа равного доступа к каналу связи может быть осуществлена децентрализованно либо при централизованном управлении. В первом случае функция нахождения свободного канала возлагается на абонентскую станцию, которая проводит последовательный поиск незанятого частотного канала во всем выделенном системе диапазоне. Во втором случае анализ занятости каналов связи осуществляет базовая станция либо непосредственно центр коммутации мобильной связи. Как правило, установление связи при последовательном сканировании частотного диапазона занимает достаточно большой интервал времени. Для обеспечения оперативности управления в современных ТСС предусматривается существование специального канала, посредством которого производится управление транкинговой системой, в том числе выполнение процедур установления и прекращения связи.

По способу организации канала управления различают ТСС с выделенным и распределенным каналом управления. В первом случае, как следует из названия, выделенный канал используется исключительно для управления работой системы. Во втором – в процессе сеанса связи сигналы управления передаются одновременно с речевым сигналом.

С учетом сказанного транкинговая система связи может быть представлена обобщенной структурной схемой (см. рис.7.2), где использованы следующие обозначения:

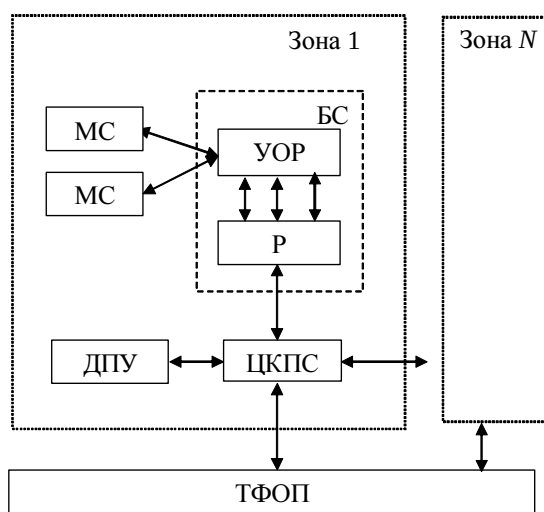


Рис. 7.2. Структурная схема ТСС

МС – мобильная станция (мобильный абонент);

БС – базовая станция (центр управления);

УОР – устройство объединения радиосигналов;

Р – ретрансляторы;

ЦКПС – центр коммутации мобильной связи;

ТФОП – телефонная сеть общего пользования;

ДПУ – диспетчерский пункт управления.

Следует отметить, что для ТСС наиболее характерно разделение каналов связи по частоте с индивидуальными ретрансляторами на разных частотах.

Возможен и вариант ТСС с использованием широкополосных ретрансляторов, обслуживающих сразу все каналы. Назначение остальных блоков структурной схемы является очевидным и не требует дополнительных комментариев.

Стандарт MPT 1327, разработанный министерством почт и телекоммуникаций Великобритании (*Ministry of Post and Telecommunication* (MPT)), определяет в основном протокол передачи информации управления и контроля состояния аппаратуры (иначе - информации сигнализации) для транкинговых систем наземной мобильной радиосвязи, причем информационные сообщения передаются по аналоговому радиоканалу. На его основе разработаны радиоинтерфейс МС (абонента), определяемый протоколом MPT 1343, и радиоинтерфейс БС – MPT 1347. Стандартами предусматривается передача информации со скоростью 1,2 кбит/с по каждому из 500 каналов связи в диапазоне частот 201,2125...207,4875 МГц (MPT 1347) и 193,2125...199,4875 МГц (MPT 1343), причем каждый дуплексный канал занимает две полосы шириной 12,5 кГц с разносом каналов приема и передачи в 8 МГц.

Фирмой Ericsson разработана система транкинговой радиосвязи, получившая название EDACS (*Enhanced Digital Access Communications System* – Усовершенствованная система связи с цифровым доступом). Системы EDACS выпускаются в различных модификациях, причем различают "системы EDACS", "сети EDACS" и "расширенные сети EDACS". Системы EDACS, объединенные между собой посредством контроллеров узлов связи и диспетчерских пунктов управления, образуют сети EDACS, которые, в свою очередь, с помощью некоторых интегрированных узлов связи могут объединяться в расширенную сеть для покрытия значительных территорий.

В системе EDACS применяются два типа радиоканалов – рабочий канал и канал управления. Канал управления служит для обмена цифровой информацией сигнализации между мобильными станциями и устройствами управления работой всей системы. Рабочие каналы используются для обмена собственно информацией (разговорной или данными) между мобильными станциями. Системы и сети EDACS рассчитаны на использование как аналоговых, так и цифровых станций, обеспечивающих передачу речевых сигналов в цифровой форме. Стандартная скорость передачи данных составляет 9,6 кбит/с по каждому из 20 каналов системы EDACS в диапазонах частот 30...300 МГц, 800 МГц или 900 МГц с разносом каналов связи 25, 30 и 12,5 кГц.

Общие тенденции, связанные с унификацией и интеграцией СМР идентичного назначения, привели к разработке под эгидой ETSI (*European Telecommunications Standards Institute* – Европейский институт стандартов связи) общеевропейского стандарта TETRA (*Trans-European Trunked Radio* – Общеевропейская система транкинговой связи), изменившего свое название с апреля 1997 г. на *Terrestrial Trunked Radio* (Сухопутная система транкинговой связи) ввиду своего широкого распространения. ТСС на основе стандарта TETRA представляют собой новое поколение систем этого типа, следующее за аналоговым. В отличие от предыдущих в стандарте TETRA осуществлен полный переход к цифровому представлению передаваемой информации и использовано не частотное, а временное разделение каналов. В результате этих и ряда других мер скорость передачи в системе достигла 36 кбит/с. Для системы TETRA выделены два дуплексных участка спектра в полосе частот 380...400 МГц при разносе радиоканалов для приема и передачи в 10 МГц и разносе соседних каналов в 25 кГц.

7.2.3. Системы персонального радиовызова

Системы персонального радиовызова (СПРВ), или *пейджинговые* системы (от английского термина *paging* – вызов) представляют собой системы мобильной радиосвязи, обеспечивающие одностороннюю передачу коротких сообщений из центра системы (терминала персонального радиовызова или пейджингового терминала) на миниатюрные абонентские приемники (пейджеры), причем передаваемые сообщения могут быть четырех типов – тональные, цифровые, буквенно-цифровые и речевые. В последнее время стали появляться публикации о разработке пейджеров для двусторонней связи, способных осуществлять передачу подтверждений о приеме сообщений.

По своему назначению СПРВ можно подразделить на частные (ведомственные) и общего пользования. Частные СПРВ предназначены для передачи сообщений на ограниченной территории в интересах отдельных групп пользователей, причем, как правило, передача осуществляется с диспетчерского терминала радиовызова без взаимодействия с ТФОП. Под системами общего пользования понимают системы мобильной радиосвязи, обеспечивающие передачу сообщений ограниченного объема от пользователей ТФОП (или с персонального компьютера) через пейджинговый терминал на абонентские приемники СПРВ. Потенциальными пользователями СПРВ в основном являются различные экстренные службы (скорая помощь, пожарная и аварийно-спасательная службы, полиция и пр.), а также лица, занятые в сфере оказания различного рода услуг (консультационных, бытового технического обслуживания и др.).

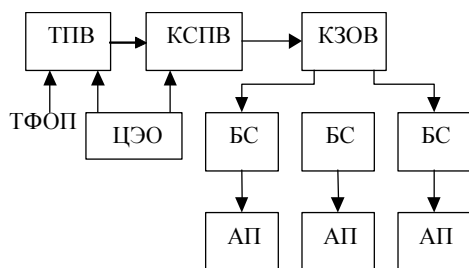
Основной отличительной чертой СПРВ служит ее работа вне реального времени, т.е. сообщения передаются не в момент его выдачи отправителем, а в порядке очереди, устанавливаемой центром управления, хотя на практике величина задержки между моментом отправления и получения не превышает нескольких минут. В сочетании с краткостью сообщений и односторонней направленностью СПРВ оказываются технически очень простыми, а значит и экономически выгодными потребителю, поэтому системы пейджинговой связи нередко называют "мобильной связью для бедных".

В самой общей версии структурная схема СПРВ может быть представлена рис.7.2, где используются сокращения:

ТПВ – терминал персонального вызова (*paging terminal – PT*),

КСПВ – контроллер сети персонального вызова (*paging network controller – PNC*);

ЦЭиО – центр эксплуатации и обслуживания (*operations and maintenance center – OMC*);



КЗОВ – контроллер зоны обслуживания вызовов (*paging area controller – PAC*);

БС – базовая станция (*base station – BS*);

АП – абонентский приемник (*pager – P*).

Терминал персонального вызова принимает и анализирует входящие данные вызовов из ТФОП (или от других источников). Контроллер сети персонального вызова во взаимодействии с ЦЭиО обеспечивает все функции управления системой, в том числе такие, как определение типа вызова абонента – индивидуальный, общий или групповой.

Контроллер зоны обслуживания осуществляет распределение данных к передатчикам базовых станций, а также может выполнять некоторые статистические вычисления. Базовые станции контролируют и передают сигналы радиовызова на соответствующие абонентские приемники. Применение нескольких БС в пределах зоны действия СПРВ позволяет обеспечить более надежную связь с абонентами.

Примером дальнейшего развития СПРВ может служить система протоколов FLEX (включая ReFLEX и InFLEXion), разработанная фирмой Motorola. Достоинством последних является повышенная скорость передачи сообщений, возможность двустороннего пейджинга, а также передача голосовых сообщений.

На ранних этапах развития массовой мобильной связи системы персонального радиовызова получили достаточно широкое распространение, причем рост количества обслуживаемых абонентов доходил до 20% ежегодно. Так, если в 1990 г. в мире насчитывалось 20 млн. пользователей, то в 1997 г. их число превысило 100 млн.. В России на начало 21-го столетия пейджинговая абонентская сеть исчислялась многими сотнями тысяч. Вместе с тем для стран с высоким уровнем проникновения мобильного телекоммуникационного сервиса (ЕС, Скандинавия, США и др.) характерно постепенное

свертывание сетей персонального радиовызова и замещение их услуг более прогрессивными и постоянно дешевеющими системами сотового телефона.

7.3. Сотовые системы мобильной связи

В 70-е годы был предложен новый принцип организации связи, который позволил увеличить число абонентов и повысить качество связи. Было предложено разбивать обслуживаемую территорию на небольшие участки, называемые сотами, или ячейками.

Разделить обслуживаемую территорию на ячейки (соты) можно двумя способами: либо основанным на измерении статистических характеристик распространения сигналов в системах связи, либо основанным на измерении или расчете параметров распространения сигнала для конкретного района.

При реализации первого способа вся обслуживаемая территория разделяется на одинаковые по форме зоны и с помощью расчетов определяются их допустимые размеры и расстояния до других зон, в пределах которых выполняются условия допустимого взаимного влияния.

Для оптимального, т. е. без перекрытия или пропусков участков, деления территории на соты могут быть использованы только три геометрические фигуры: треугольник, квадрат и шестиугольник. Наиболее подходящей фигурой является шестиугольник, так как, если антенну с круговой диаграммой направленности устанавливать в его центре, то будет обеспечен доступ почти ко всем участкам соты.

При использовании первого способа интервал между зонами, в которых используются одинаковые рабочие каналы, обычно получается больше требуемого для поддержания взаимных помех на допустимом уровне.

Более приемлем второй способ деления на зоны. В этом случае тщательно измеряют или рассчитывают параметры системы для определения минимального числа базовых станций, обеспечивающих удовлетворительное обслуживание абонентов по всей территории, определяют оптимальное место расположения базовой станции с учетом рельефа местности, рассматривают возможность использования направленных антенн, пассивных ретрансляторов и смежных центральных станций в момент пиковой нагрузки и т. д.

Каждая из ячеек обслуживается своим передатчиком с невысокой выходной мощностью и ограниченным числом каналов связи. Это позволяет без помех использовать повторно частоты каналов этого передатчика в другой, удаленной на значительное расстояние, ячейке. Теоретически такие передатчики можно использовать и в соседних ячейках. Но на практике зоны обслуживания сот могут перекрываться под действием различных факторов, например, вследствие изменения условий распространения радиоволн. Поэтому в соседних ячейках используются различные частоты. Пример построения сот при использовании трех частот F_1 – F_3 представлен на рис. 7.3.

Группа сот с различными наборами частот называется кластером. Определяющим его параметром является количество используемых в соседних сотах частот. На рис. 7.3, например, размерность кластера равна трем. Но на практике это число может достигать пятнадцати.

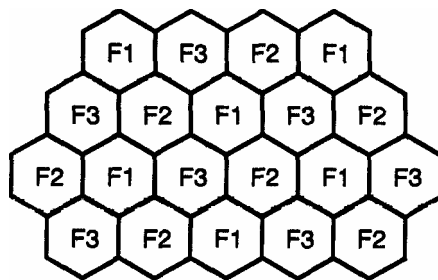


Рис. 7.3. Построение сот для трех частот

F1, F2, F3 – частоты базовых станций.

Основной идеей, на которой базируется принцип сотовой связи, является повторное использование частот в несмежных сотах. Первым способом организации повторного использования частот, который применялся в аналоговых системах сотовой подвижной связи первого поколения, был способ, использующий антенны базовых станций с круговыми диаграммами направленности (рис. 7.4). Он предполагает передачу сигнала одинаковой мощности по всем направлениям, что для абонентских станций эквивалентно приему помех от всех базовых станций со всех направлений.

Базовые станции, на которых допускается повторное использование выделенного набора частот, удалены друг от друга на расстояние D , называемое «защитным интервалом» (см. рис. 7.4). Именно возможность повторного применения одних и тех же частот определяет высокую эффективность использования частотного спектра в сотовых системах связи.

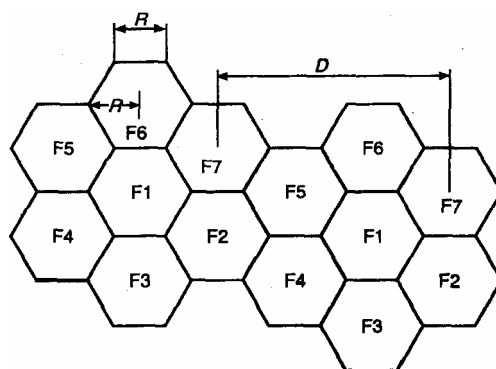


Рис. 7.4. Повторное использование частот в несмежных сотах

Смежные базовые станции, использующие различные наборы частотных каналов, образуют группу из C станций, образуя, так называемый, кластер (см. рис. 7.4). Если каждой базовой станции выделяется набор из m каналов с шириной полосы каждого F_k , то общая ширина полосы, занимаемая системой сотовой связи, составит $F_c = F_k m C$.

Таким образом, величина C определяет минимально возможное число каналов в системе, поэтому ее часто называют частотным параметром системы, или коэффициентом повторения частот. Коэффициент C не зависит от числа каналов в наборе и увеличивается по мере уменьшения радиуса ячейки. Таким образом, при использовании ячеек меньших радиусов имеется возможность увеличения повторяемости частот.

Применение шестиугольных ячеек позволяет минимизировать ширину необходимого частотного диапазона, поскольку такая форма обеспечивает оптимальное соотношение между величинами C и D . Кроме того, шестиугольная форма наилучшим образом вписывается в круговую диаграмму направленности антенны базовой станции, установленной в центре ячейки.

Остановимся более подробно на вопросе выбора размера ячейки (радиуса R). Эти размеры определяют защитный интервал (см. рис. 7.4) между ячейками, в которых одни и те же частоты могут быть использованы повторно. Заметим, что величина защитного интервала D , кроме уже перечисленных факторов, зависит также от допустимого уровня помех и условий распространения радиоволн. В предположении, что интенсивность вызовов в пределах всей зоны одинакова, ячейки выбираются одного размера. Размер зоны обслуживания базовой станции, выражаемый через радиус ячейки R , определяет также число абонентов N , способных одновременно вести переговоры на всей территории обслуживания. Следовательно, уменьшение радиуса ячейки позволяет не только повысить эффективность использования выделенной полосы частот и увеличить абонентскую емкость системы, но и уменьшить мощность передатчиков и чувствительность

приемников базовых и подвижных станций. Это, в свою очередь, улучшает условия электромагнитной совместимости средств сотовой связи с другими радиоэлектронными средствами и системами.

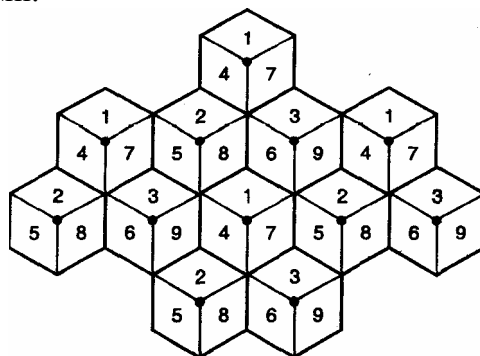


Рис. 7.5. Модель повторного использования частот в 3-секторных сотах

Эффективным способом снижения уровня помех может быть использование направленных секторных антенн с узкими диаграммами направленности. В секторе такой направленной антенны сигнал излучается преимущественно в одну сторону, а уровень излучения в противоположном направлении сокращается до минимума. Деление сот на секторы позволяет чаще применять частоты в сотах повторно. Общеизвестный способ повторного использования частот в организованных таким образом сотах основан на применении 3-секторных антенн для каждой базовой станции и трех соседних базовых станций с формированием ими девяти групп частот (рис. 7.5). В этом случае используются антенны с шириной диаграммы направленности 120° .

Самую высокую эффективность использования полосы частот и, следовательно, наибольшее число абонентов сети, работающих в этой полосе, обеспечивает разработанный фирмой Motorola (США) способ повторного использования частот, при котором задействуются две базовые станции. При реализации этого способа (рис. 7.6) каждая частота используется дважды в пределах кластера, состоящего из 4 ячеек; базовая станция каждой из них может работать на 12 частотах, используя антенны с диаграммой направленности шириной 60° .

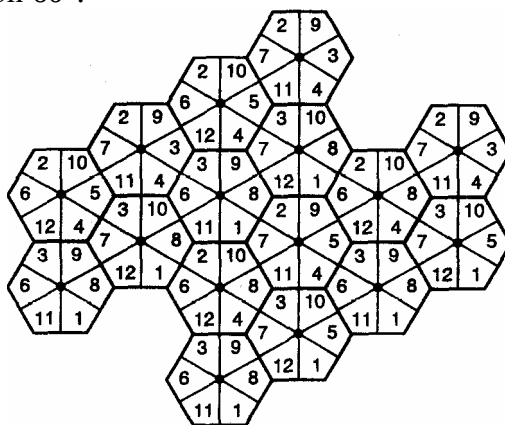


Рис. 7.6. Модель повторного использования частот в двух соседних сотах

Каждая из сот обслуживается многоканальным приемопередатчиком, называемым базовой станцией. Она служит своеобразным интерфейсом между сотовым телефоном и центром коммутации подвижной связи, где роль проводов обычной телефонной сети выполняют радиоволны. Число каналов базовой станции обычно кратно 8, например, 8, 16, 32... Один из каналов является управляющим (*control channel*). В некоторых ситуациях он может называться также каналом вызова (*calling channel*). На этом канале происходит непосредственное установление соединения при вызове подвижного абонента сети, а сам разговор начинается только после того, как будет найден свободный в данный момент

канал и произойдет переключение на него. Все эти процессы происходят очень быстро и потому незаметно для абонента. Он лишь набирает нужный ему телефонный номер и разговаривает, как по обычному телефону.

Любой из каналов сотовой связи представляет собой пару частот для дуплексной связи, т. е. частоты базовой и подвижной станций разнесены. Это делается для того, чтобы улучшить фильтрацию сигналов и исключить взаимное влияние передатчика на приемник одного и того же устройства при их одновременной работе.

Все базовые станции соединены с центром коммутации подвижной связи (коммутатором) по выделенным проводным или радиорелейным каналам связи (рис. 7.7). *Центр коммутации MSC* – это автоматическая телефонная станция системы сотовой связи, обеспечивающая все функции управления сетью.

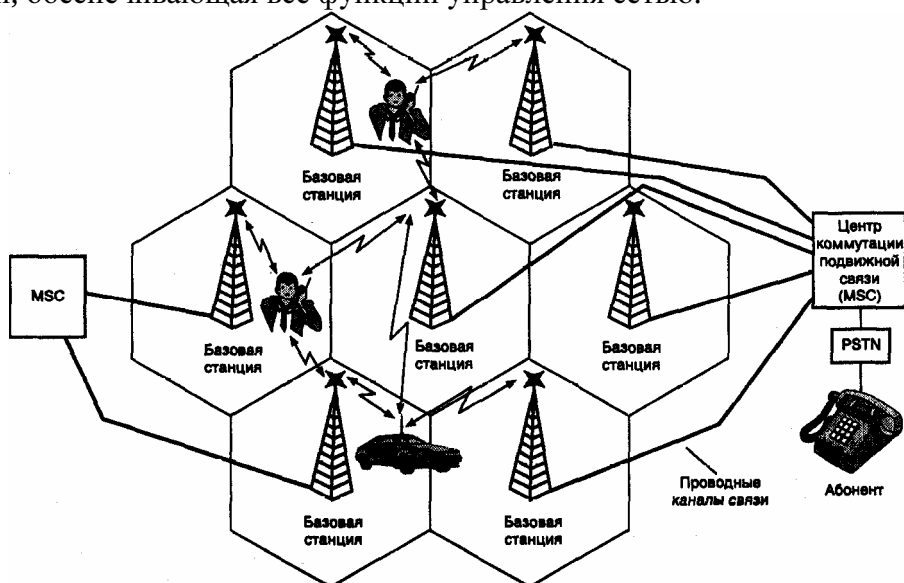


Рис. 7.7. Основные составляющие систем сотовой связи

Она осуществляет постоянное слежение за подвижными станциями, организует их эстафетную передачу, в процессе которой достигается непрерывность связи при перемещении подвижной станции из соты в соту и переключение рабочих каналов в соте при появлении помех или неисправностей, производит соединение подвижного абонента с тем, кто ему необходим в обычной телефонной сети и др.

Несмотря на разнообразие стандартов сотовой связи, алгоритмы их функционирования, независимо от имеющихся особенностей, в основном сходны. Для абонента практически нет никакой разницы, в каком стандарте осуществляется связь. Если ему нужно позвонить, то он просто нажимает клавишу на своем радиотелефоне (это может быть любой сотовый радиотелефон), что соответствует снятию трубки обычного телефона. Когда же радиотелефон находится в режиме ожидания (состояние «трубка положена» обычного телефона), его приемное устройство постоянно сканирует (просматривает) либо все каналы системы, либо только управляющие. Для вызова соответствующего абонента всеми базовыми станциями сотовой системы связи по управляющим каналам передается сигнал вызова. Сотовый телефон вызываемого абонента при получении этого сигнала отвечает по одному из свободных каналов управления. Базовые станции, принявшие ответный сигнал, передают информацию о его параметрах в центр коммутации, который, в свою очередь, переключает разговор на ту базовую станцию, где зафиксирован максимальный уровень сигнала сотового радиотелефона вызываемого абонента.

Во время набора номера радиотелефон занимает один из свободных каналов, уровень сигнала базовой станции в котором в данный момент максимален. По мере удаления абонента от базовой станции или в связи с ухудшением условий распространения радиоволн уровень сигнала уменьшается, что ведет к ухудшению качества связи.

Улучшение качества разговора достигается путем автоматического переключения абонента на другой канал связи. Это происходит следующим образом. Специальная процедура, называемая передачей управления вызовом или эстафетной передачей (в иностранной технической литературе – *handover*, или *handoff*), позволяет переключить разговор на свободный канал другой базовой станции, в зоне действия которой оказался в это время абонент. Аналогичные действия предпринимаются при снижении качества связи из-за влияния помех или при возникновении неисправностей коммутационного оборудования. Для контроля таких ситуаций базовая станция снабжена специальным приемником, периодически измеряющим уровень сигнала сотового телефона разговаривающего абонента и сравнивающим его с допустимым пределом. Если уровень сигнала меньше этого предела, то информация об этом автоматически передается в центр коммутации по служебному каналу связи. Центр коммутации выдает команду об измерении уровня сигнала сотового радиотелефона абонента на ближайшие к нему базовые станции. После получения информации от базовых станций об уровне этого сигнала центр коммутации переключает радиотелефон на ту из них, где уровень сигнала оказался наибольшим. Это происходит так быстро, что абонент совершенно не замечает этих переключений.

Иногда возникает ситуация, когда поток заявок на обслуживание, поступающий от абонентов сотовой сети, превышает количество каналов, имеющихся на всех близко расположенных базовых станциях. Это происходит тогда, когда все каналы станций заняты обслуживанием абонентов и нет ни одного свободного и поступает очередная заявка на обслуживание от подвижного абонента. В этом случае как временная мера (до освобождения одного из каналов) используется принцип эстафетной передачи внутри соты. При этом происходит поочередное переключение каналов в пределах одной и той же базовой станции для обеспечения связью всех абонентов.

Одна из важных услуг сети сотовой связи – предоставление возможности использования одного и того же радиотелефона при поездке в другой город, область или даже страну, причем сотовая сеть позволяет не только самому абоненту звонить из другого города или страны, но и получать звонки от тех, кто не успел застать его дома. В сотовой радиосвязи такая возможность называется роуминг (от англ. *room* – скитаться, блуждать). Для организации роуминга сотовые сети должны быть одного стандарта (телефон стандарта GSM не будет работать в сети стандарта CDMA и т. п.), а центры коммутации подвижной связи этого стандарта должны быть соединены специальными каналами связи для обмена данными о местонахождении абонента. Иными словами, применительно к сотовым системам для обеспечения роуминга необходимо выполнение трех условий:

- Наличие в требуемых регионах сотовых систем стандарта, совместимого со стандартом компании, у которой был приобретен радиотелефон
- Наличие соответствующих организационных и экономических соглашений о роуминговом обслуживании абонентов
- Наличие каналов связи между системами, обеспечивающих передачу звуковой и другой информации для роуминговых абонентов

При перемещении абонента в другую сеть ее центр коммутации запрашивает информацию в первоначальной сети и при наличии подтверждения полномочий абонента регистрирует его. Данные о местоположении абонента постоянно обновляются в центре коммутации первоначальной сети, и все поступающие туда вызовы автоматически переадресовываются в ту сеть, где в данный момент находится абонент.

При организации роуминга недостаточно провести только технические мероприятия по соединению различных сетей сотовой связи. Очень важно еще решить проблему взаиморасчетов между операторами этих сетей.

Различают три вида роуминга:

- Автоматический (именно с этой формой за рубежом обычно и связывают понятие роуминга), т. е. предоставление абоненту возможности выйти на связь «в любое время в любом месте»;
- Полуавтоматический, когда абоненту для пользования данной услугой в каком-либо регионе необходимо предварительно поставить об этом в известность своего оператора
- Ручной, по сути, простой обмен одного радиотелефона на другой, подключенный к сотовой системе другого оператора

Существующий объем услуг роуминга во многом определяется активностью деятельности конкретных компаний, так как возникающие при этом технические проблемы у всех приблизительно одинаковы (хотя здесь и можно отметить стандарт GSM, в который возможность роуминга была заложена изначально). Перспективы развития этой сферы услуг зависят уже от распространенности стандартов.

7.3.1. Понятие о модели взаимодействия открытых систем

Эта эталонная модель (OSI – *Open Systems Interconnection*), разработанная ISO, обычно обозначается как модель ISO/OSI и является базовой в том смысле, что для описания некоторых систем достаточна лишь часть ее компонентов. Несмотря на существование конкурентоспособных альтернатив, модель OSI получила повсеместное распространение и широко используется как концептуальная основа в информационной системотехнике.

Под *открытой* системой понимается система, удовлетворяющая требованиям международных организаций или требованиям производителя, контролирующего большую часть рынка сбыта. Открытость означает, что система не является чьей-то собственностью, т.е. любой производитель вправе создавать и выставлять на продажу собственный продукт (аппаратуру или программное обеспечение), предназначенный для практического доступа к услугам системы. Между производителями возникает конкуренция, что способствует снижению цены оборудования. Пользователи при этом не зависят от единственного производителя и могут менять поставщиков аппаратуры.

Диаграмма рис. 7.8 иллюстрирует схему взаимодействия элементов и уровней базовой модели, представленной в общем виде. Модель некоторых систем содержит лишь часть этих элементов. Так, системы с коммутацией каналов описываются одноуровневой моделью. Расшифруем используемые термины.

Прикладной процесс выполняет обработку данных для нужд пользователя либо с помощью некоторых аппаратно-программных средств, либо непосредственно в мозгу человека.

Под *интерфейсом* (стыком) понимается определенная стандартами граница между взаимодействующими объектами. Для систем мобильной связи главенствующее значение имеет радиоинтерфейс, охватывающий средства соединения между мобильной и базовой станциями.

Физические средства соединения - это совокупность аппаратуры и среды распространения, обеспечивающих передачу сигналов между системами.

Абонент через интерфейс пользователя дает задание прикладному процессу и получает результат его выполнения.

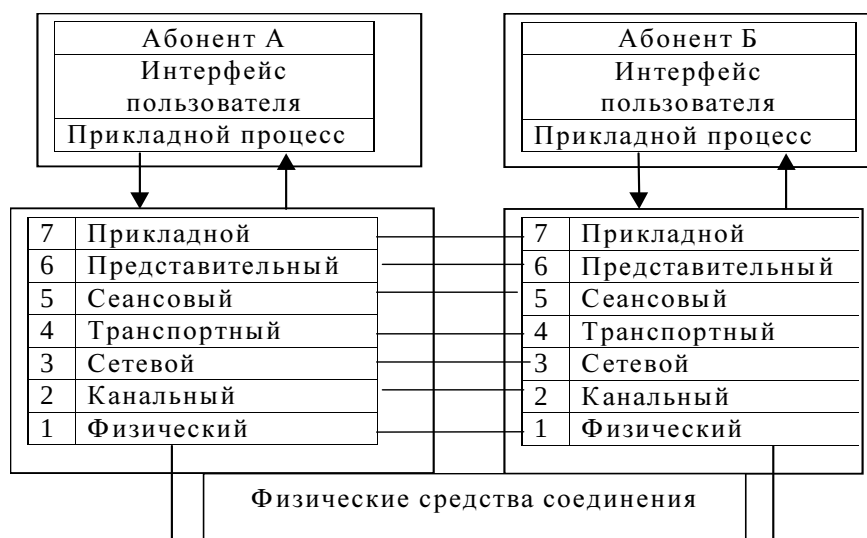


Рис.7.8. Модель взаимодействия открытых систем

Областью взаимодействия называется иерархическая группа функциональных блоков, предназначенных для обмена данными между системами. В базовой модели эти функции поделены на семь расположенных друг над другом слоев, именуемых *уровнями*, как показано на рис. 7.8.

Такое деление позволяет разбить сложную систему на ряд более простых (осуществить декомпозицию), образовать стандартные модули и интерфейсы между уровнями. Опишем кратко основные функции, закрепленные за каждым уровнем.

Верхний, седьмой прикладной уровень обеспечивает различные формы взаимодействия прикладных процессов, т. е. является прикладным интерфейсом. На этом уровне выполняется управление заданиями, идентификация пользователей по паролям, адресам, электронным подписям.

На представительном уровне осуществляется кодирование сообщений, поступающих с седьмого уровня, в нужную форму. Здесь же при необходимости осуществляется сжатие данных и шифрование.

Сеансовый уровень определяет процедуру проведения сеансов между прикладными процессами, т. е. обеспечивает установку, поддержание и разрыв соединения.

Транспортный уровень отвечает за точность доставки пакетов (блоков данных по 500 – 2000 бит), по виртуальным каналам, проложенным между портами.

Сетевой уровень создает виртуальные каналы – пути, по которым данные передаются от одного порта к другому.

На канальном уровне производится формирование и передача кадров (блоков данных), обнаружение и исправление ошибок.

Физический уровень создает интерфейсы для подключения систем к физическим средствам соединения (электрическим и оптическим кабелям, радиоэфиру вместе с аппаратурой передачи данных). Он отвечает за передачу последовательности бит, оповещает о появлении неисправностей и отказов (столкновении кадров, обрывов, отключении питания).

Хотя все сообщения передаются через первый – физический уровень, в модели OSI определены также *логические* каналы, понимаемые как пути, по которым передаются сигналы или данные между одноименными уровнями. Так, путь, связывающий четвертые уровни, называется транспортным каналом, путь между третьими – сетевым, вторыми – информационным и, наконец, собственно физическим каналом, называется путь, соединяющий первые уровни. Особое место занимает логический канал *трафика* (пользовательский канал), обеспечивающий обмен данными между абонентами.

При технической реализации систем логические каналы проектируются на физические, т. е. указывается, какой физический канал в данный момент времени выделяется для организации конкретного логического канала.

Эталонная модель обладает свойствами *прозрачности* и *доступности* сервиса. Прозрачность означает, что данные, переданные с какого-то уровня на нижние, на приемной стороне возвращаются на этот же уровень без изменений. Создается видимость, что для рассматриваемых одноименных уровней нижних не существует, они прозрачны. Ярким примером прозрачного объекта может служить телефонный тракт. Большинство людей, пользующихся телефоном, даже не имеют представления о преобразованиях, которым подвергается их речь в телефонном канале.

Под доступностью сервиса понимается возможность использования всеми верхними уровнями услуг (сжатие данных, защита от помех, от несанкционированного использования информации), выполняемых на нижних уровнях.

На рис. 7.8 линиями, соединяющими одноименные уровни, обозначены *протоколы* – правила взаимодействия двух (или более) объектов. Названия протоколов совпадают с названиями соответствующих уровней, например, существуют сеансовые, транспортные, сетевые протоколы.

Протоколы должны удовлетворять определенным требованиям. Прежде всего, протокол – это последовательность шагов, выполняемых взаимодействующими объектами добровольно, в порядке строгой очередности для достижения поставленной цели. Далее, протокол должен содержать описание действий его участников при любых возникающих ситуациях (должны быть исключены тупиковые ситуации). Правила игры в шахматы, в шашки являются примерами протоколов.

Следует упомянуть о существовании криптографических протоколов, обеспечивающих достижение цели в условиях возможного мошенничества, как со стороны законных пользователей, так и внешнего нарушителя. Примером такого протокола может служить аутентификация абонента.

Для описания систем мобильной связи базовая семиуровневая модель трансформируется путем деления на два укрупненных слоя. В стандартах GSM и TETRA верхние уровни с четвертого по седьмой образуют один слой, нижние (с первого по третий) – другой. При передаче речи функции седьмого и шестого уровней выполняет человек. В его мозгу зарождается сообщение, он выбирает язык общения. Задачи, решаемые на пятом и четвертом уровнях, важны для соединительных линий, объединяющих отдельные системы связи в единую сеть. Поэтому указанные системы описываются трехуровневой моделью, включающей лишь сетевой, канальный и физический уровни.

7.3.2. Эволюция систем и стандартов сотовой связи

История развития ССМС насчитывает немногим более 30 лет, однако, этот короткий период был отмечен рядом поворотных моментов и существенной эволюцией воззрений на роль и философию рассматриваемых систем. Можно говорить о трех поколениях ССМС, различия между которыми – с известной долей условности – устанавливаются следующими критериями.

Все ССМС или стандарты первого поколения являются аналоговыми. В их числе:

- AMPS (*Advanced Mobile Phone Service*) – Усовершенствованная мобильная телефонная служба. Диапазон рабочих частот – 869...894 МГц для БС и 824...849 МГц для МС, ширина полосы канала связи – 30 кГц. Начало коммерческого применения – 1983 г. Широко использовалась в США, Канаде, Центральной и Южной Америке, Австралии.

- TACS (*Total Access Communications System*) – Общедоступная система связи. Частотный диапазон: 935...950 МГц для БС, 890...905 МГц для МС, ширина полосы канала связи – 25 кГц. Начало коммерческого применения – 1985 г. Наибольшее

распространение стандарт TACS получил в европейских странах – Англия, Италия, Испания, Австрия и др.

- NMT (*Nordic Mobile Telephone System*) – Скандинавская система мобильной телефонной связи. Существует в двух основных вариантах NMT 450 и NMT 900, отличаясь только диапазоном используемых частот: NMT 450 – 463...467,5 МГц для БС и 453...457,5 для МС; NMT 900 – 935...960 МГц для БС и 890...915 МГц для МС. Ширина полосы канала – 25 кГц. Начало коммерческого использования – 1981 г. (NMT 450) и 1986 г. (NMT 900). Помимо скандинавских стран эти стандарты широко использовались во многих странах Западной и Восточной Европы и ряде других регионов мира.

Во всех перечисленных аналоговых стандартах применяется метод множественного доступа с частотным разделением каналов, для передачи речевой информации используется частотная модуляция, а для передачи информации сигнализации - частотная манипуляция. Системам первого поколения присущ ряд недостатков, основными из которых являются относительно низкая абонентская емкость, несовместимость различных стандартов, отсутствие засекречивания передаваемых сообщений, невозможность взаимодействия с цифровыми системами с интеграцией служб (ISDN) и пакетной передачей данных (PDN).

В значительной степени указанных недостатков лишены цифровые ССМС второго поколения, среди которых наибольшее распространение получили следующие:

- D-AMPS (*Digital-AMPS* – цифровая AMPS), или IS-54 (IS – сокращение от *Interim Standard*, т.е. промежуточный стандарт) представляет двухрежимную аналого-цифровую систему, совмещающую работу в аналоговом и цифровом режимах в том же диапазоне, что и AMPS. Начало практического использования относится к 1992 г. Усовершенствованная версия данного стандарта IS-136, отличие которой от IS-54 заключается в наличии полностью цифровых каналов управления, начала применяться с 1996 г. Версия IS-136 используется в диапазонах 800 и 1900 МГц.

- GSM (*Global System for Mobile Communications*) – Глобальная система мобильной связи. Данным стандартом предусматривается работа в диапазоне 935...960 МГц для БС и 890...915 МГц для МС при ширине полосы канала связи 200 кГц. Практическое применение общеевропейского стандарта GSM 900 началось в 1991 г. Совершенствование данного стандарта привело к освоению нового частотного диапазона 1800 МГц, в котором благодаря более широкой рабочей полосе частот в сочетании с меньшими размерами сот удастся строить сотовые сети значительно большей емкости. Первоначально данная версия именовалась *Personal Communication Network (PCN)* – Сеть персональной связи, затем *Digital Cellular System (DCS)* – Цифровая система сотовой связи, а спустя три года после начала эксплуатации (1993 г.) была переименована в GSM 1800. Диапазон работы БС – 1710...1785 МГц, МС – 1805...1880 МГц при ширине полосы канала связи 200 кГц. Стандарт GSM нашел применение и в США, однако, из-за того, что диапазон 1800 МГц занят системой D-AMPS в версии IS-136, ему была выделена полоса частот в диапазоне 1900 МГц. Соответствующая версия стандарта GSM получила наименование "американский" GSM или IS-661. На начало 1999 г. стандарт GSM в различных версиях использовался в 129 странах мира, а объем абонентской базы достиг величины 137 млн. чел., что составляет 45% от общего числа пользователей ССМС.

- PDC (*Personal Digital Cellular*) – Персональная цифровая сотовая связь. Цифровая ССМС, разработанная в Японии в 1993 г. и первоначально называвшаяся JDC (*Japan Digital Cellular* – Японская цифровая сотовая связь), по своим техническим характеристикам подобна D-AMPS и отличается от последней возможностью работы в нескольких диапазонах частот – 800, 1400 и 1500 МГц. Хотя стандарт PDC используется только в Японии, его абонентская база достигает 39,5 млн. пользователей или 13% от общемирового числа абонентов ССМС.

- IS-95 (*Interim Standard-95*), альтернативное наименование *cdmaOne*. Все перечисленные ранее цифровые ССМС второго поколения используют метод

множественного доступа с частотно-временным разделением каналов связи. Критическими явились 1992-93 г.г., когда в США был разработан первый стандарт ССМС на основе метода множественного доступа с кодовым разделением (CDMA), получивший название IS-95. Диапазон рабочих частот –824...848 МГц для МС и 869...894 МГц – для БС при ширине спектра излучаемых сигналов 1,25 МГц. Практическое применение указанного стандарта началось в 1995-96 г.г. в Гонконге, США и Южной Корее, причем в США используется версия этого стандарта для диапазона 1900 МГц.

Что касается ССМС третьего поколения, то процесс формирования их облика еще продолжается, изобилуя коллизиями, конфликтами интересов и компромиссами. Принципиальные критерии, отличающие ССМС третьего поколения, а также некоторые детали конкретных спецификаций, имеющих фактический статус стандартов и весьма близких к практическому воплощению, рассмотрены в разделе 7.6.

7.3.3. Описание ССМС на примере GSM

Рассмотрим основные элементы ССМС и характер их взаимодействия. Хотя сотовые системы разных стандартов имеют значительные отличия в деталях, интегральное их описание с помощью некоторой обобщенной модели возможно и полезно. В качестве подобного обобщения выберем схему ССМС, приближенную к стандарту GSM, как наиболее отчетливо структурированную. Функциональное построение ССМС иллюстрируется схемой на рис. 4.6. В приведенной схеме условно можно выделить четыре основных компонента – центр управления и обслуживания (ЦУО) (иначе *operations and maintenance center* (ОМС)) и три подсистемы:

- подсистема мобильных станций ПМС (*mobile station subsystem* MSS);
- подсистема базовых станций ПБС (*base station subsystem* BSS);
- подсистема коммутации ПК (*switching subsystem* SSS),

функциональное сопряжение которых описывается рядом интерфейсов.

ЦУО является центральным элементом ССМС, который обеспечивает управление другими компонентами системы, а также контроль качества функционирования. Подсистема мобильных станций объединяет оборудование, обеспечивающее доступ абонентов в систему. Главным звеном в архитектуре ССМС является подсистема коммутации, которая включает в себя центр коммутации мобильной связи ЦКМС (*mobile switching center* MSC), визитный (гостевой) регистр местоположения ВРМ (*visited location register* VLR), домашний регистр местоположения ДРМ (*home location register* HLR), центр аутентификации ЦА (*authentication center* AUC) и регистр идентификации оборудования РИО (*equipment identity register* EIR). В подсистему базовых станций входят базовые приемо-передающие станции БС (*base transceiver station* BTS) и контроллеры базовых станций КБС (*base station controller* BSC).

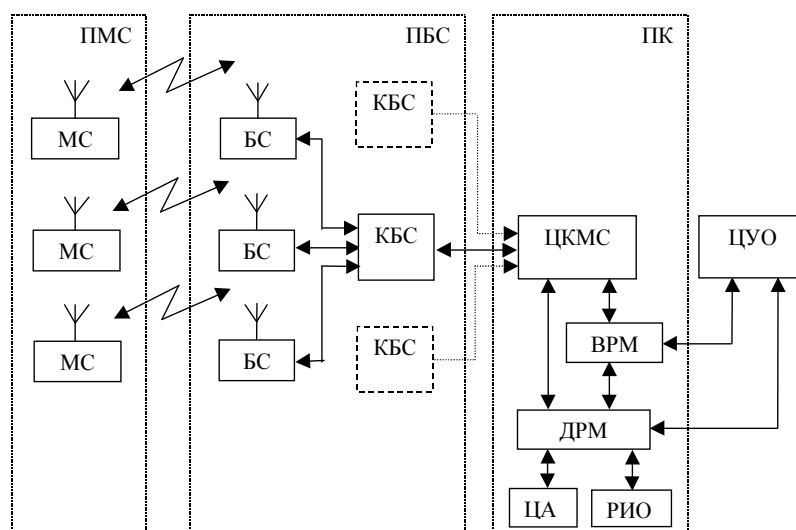


Рис. 7.9. Обобщенная структурная схема сотовой системы мобильной связи

Центр коммутации мобильной связи обслуживает группу сот и обеспечивает все виды соединений, в которых нуждается в процессе работы мобильная станция, т.е. коммутацию мобильных абонентов друг с другом, с абонентом ТФОП и др. На ЦКМС возлагаются также функции коммутации каналов, к которым относятся "передача обслуживания" (или "эстафетная передача") и переключение каналов в соте при появлении сильных помех и неисправностей, если только это не является обязанностью КБС. Помимо коммутационных задач ЦКМС управляет процедурами слежения за мобильными станциями с помощью домашнего и визитного регистров местоположения для обеспечения доставки вызова, а также процедурами аутентификации и идентификации абонентов с помощью ЦА и РИО.

Когда передача вызовов осуществляется между двумя сетями, управляемыми разными КБС (BSC), то первичное управление осуществляется в ЦКМС (MSC). В стандарте GSM также предусмотрены процедуры передачи вызова между сетями (контроллерами), относящимися к разным ЦКМС (MSC). Центр коммутации осуществляет постоянное слежение за подвижными станциями, используя регистры положения ДРМ (HLR) и перемещения ВРМ (VLR). В HLR хранится та часть информации о местоположении какой-либо подвижной станции, которая позволяет центру коммутации доставить вызов станции. Регистр HLR содержит международный идентификационный номер подвижного абонента (IMSI). Он используется для опознавания подвижной станции в центре аутентификации ЦА (AUC).

Регистры ДРМ (HLR) и ВРМ (VLR) по своей сути представляют собой базы данных. Первый содержит сведения о постоянно приписанных к данному ЦКМС абонентах и о видах услуг, которые им могут быть оказаны, второй содержит информацию об абонентах, временно находящихся в зоне обслуживания данного ЦКМС (MSC). Центр аутентификации обеспечивает возможность проведения процедуры аутентификации абонентов и шифрования передаваемых сообщений, РИО содержит сведения об эксплуатируемых мобильных станциях на предмет их исправности и санкционированного использования.

К данным, содержащимся в HLR, имеют дистанционный доступ все MSC и VLR сети и, если в сети имеются несколько HLR, в базе данных содержится только одна запись об абоненте, поэтому каждый HLR представляет собой определенную часть общей базы данных сети об абонентах. Доступ к базе данных об абонентах осуществляется по номеру IMSI или MSISDN (номеру подвижного абонента в сети ISDN). К базе данных могут получить доступ MSC или VLR, относящиеся к другим сетям, в рамках обеспечения межсетевого роуминга абонентов.

Второе основное устройство, обеспечивающее контроль за передвижением подвижной станции из зоны в зону, - регистр перемещения VLR. С его помощью достигается функционирование подвижной станции за пределами зоны, контролируемой HLR. Когда в процессе перемещения подвижная станция переходит из зоны действия одного контроллера базовой станции BSC, объединяющего группу базовых станций, в зону действия другого BSC, она регистрируется новым BSC, и в VLR заносится информация о номере области связи, которая обеспечит доставку вызовов подвижной станции. Для сохранности данных, находящихся в HLR и VLR, в случае сбоев предусмотрена защита устройств памяти этих регистров.

В сети подвижной связи GSM соты группируются в географические зоны (LA), которым присваивается свой идентификационный номер (LAC). Каждый VLR содержит данные об абонентах в нескольких LA. Когда подвижный абонент перемещается из одной LA в другую, данные о его местоположении автоматически обновляются в VLR. Если старая и новая LA находятся под управлением различных VLR, то данные на старом VLR стираются после их копирования в новый VLR. Текущий адрес VLR абонента, содержащийся в HLR, также обновляется.

VLR обеспечивает также присвоение номера "блуждающей" подвижной станции (MSRN). Когда подвижная станция принимает входящий вызов, VLR выбирает его MSRN и передает его на MSC, который осуществляет маршрутизацию этого вызова к базовым станциям, находящимся рядом с подвижным абонентом.

VLR также распределяет номера передачи управления при передаче соединений от одного MSC к другому. Кроме того, VLR управляет распределением новых временных международных идентификационных номеров (TMSI) и передает их в HLR. Он также управляет процедурами установления подлинности во время обработки вызова. По решению оператора TMSI может периодически изменяться для усложнения процедуры идентификации абонентов. Доступ к базе данных VLR может обеспечиваться через IMSI, TMSI или MSRN. В целом VLR представляет собой локальную базу данных о подвижном абоненте для той зоны, где находится абонент, что позволяет исключить постоянные запросы в HLR и сократить время на обслуживание вызовов.

Для исключения несанкционированного использования ресурсов системы связи вводятся механизмы аутентификации - удостоверения подлинности абонента. Центр аутентификации состоит из нескольких блоков и формирует ключи и алгоритмы аутентификации. С его помощью проверяются полномочия абонента и осуществляется его доступ к сети связи. AUC принимает решения о параметрах процесса аутентификации и определяет ключи шифрования абонентских станций на основе базы данных, сосредоточенной в регистре идентификации оборудования (EIR - Equipment Identification Register).

Каждый подвижный абонент на время пользования системой связи получает стандартный модуль подлинности абонента (SIM), который содержит: международный идентификационный номер (IMSI), свой индивидуальный ключ аутентификации (Ki), алгоритм аутентификации (A3).

С помощью записанной в SIM информации в результате взаимного обмена данными между подвижной станцией и сетью осуществляется полный цикл аутентификации и разрешается доступ абонента к сети.

Процедура проверки сетью подлинности абонента реализуется следующим образом. Сеть передает случайный номер (RAND) на подвижную станцию. На ней с помощью Ki и алгоритма аутентификации A3 определяется значение отклика (SRES), т.е.

$$SRES = Ki * [RAND]$$

Подвижная станция посылает вычисленное значение SRES в сеть, которая сверяет значение принятого SRES со значением SRES, вычисленным сетью. Если оба значения совпадают, подвижная станция приступает к передаче сообщений. В противном случае

связь прерывается, и индикатор подвижной станции показывает, что опознавание не состоялось. Для обеспечения секретности вычисление SRES происходит в рамках SIM. Несекретная информация (например, Ki) не подвергается обработке в модуле SIM.

EIR - регистр идентификации оборудования, содержит централизованную базу данных для подтверждения подлинности международного идентификационного номера оборудования подвижной станции (IMEI). Эта база данных относится исключительно к оборудованию подвижной станции. База данных EIR состоит из списков номеров IMEI, организованных следующим образом:

БЕЛЫЙ СПИСОК - содержит номера IMSI, о которых есть сведения, что они закреплены за санкционированными подвижными станциями.

ЧЕРНЫЙ СПИСОК - содержит номера IMSI подвижных станций, которые украдены или которым отказано в обслуживании по другой причине.

СЕРЫЙ СПИСОК - содержит номера IMSI подвижных станций, у которых существуют проблемы, выявленные по данным программного обеспечения, что не является основанием для внесения в "черный список".

ПБС (BSS) - оборудование базовой станции, состоит из контроллера базовой станции КБС (BSC) и приемо-передающих базовых станций БС (BTS). Контроллер базовой станции может управлять несколькими приемо-передающими блоками. BSS управляет распределением радиоканалов, контролирует соединения, регулирует их очередность, обеспечивает режим работы с прыгающей частотой, модуляцию и демодуляцию сигналов, кодирование и декодирование сообщений, кодирование речи, адаптацию скорости передачи для речи, данных и вызова, определяет очередность передачи сообщений персонального вызова.

К числу операций, производимых КБС, относятся передача обслуживания при переходе МС между сотами, контролируемые одним и тем же КБС, передача вызовов между МС, находящимися в зонах действия разных БС, но одного и того же контроллера (в противном случае, первичное управление возлагается на ЦКМС). Под управлением КБС осуществляются радиоизмерения в каналах связи, регулируется мощность передатчиков мобильных и базовых станций.

Внутренние GSM – интерфейсы

Интерфейс между MSC и HLR совмещен с VLR (B-интерфейс). Когда MSC необходимо определить местоположение подвижной станции, он обращается к VLR. Если подвижная станция инициирует процедуру местоопределения с MSC, он информирует свой VLR, который заносит всю изменяющуюся информацию в свои регистры. Эта процедура происходит всегда, когда MS переходит из одной области местоопределения в другую. В случае, если абонент запрашивает специальные дополнительные услуги или изменяет некоторые свои данные, MSC также информирует VLR, который регистрирует изменения и при необходимости сообщает о них HLR.

Интерфейс между ЦКМС (MSC) и HLR (C-интерфейс) используется для обеспечения взаимодействия между MSC и HLR. MSC может послать указание (сообщение) HLR в конце сеанса связи для того, чтобы абонент мог оплатить разговор. Когда сеть фиксированной телефонной связи не способна исполнить процедуру установления вызова подвижного абонента, MSC может запросить HLR с целью определения местоположения абонента для того, чтобы послать вызов MS.

Интерфейс между HLR и VLR (D-интерфейс) используется для расширения обмена данными о положении подвижной станции, управления процессом связи. Основные услуги, предоставляемые подвижному абоненту, заключаются в возможности передавать или принимать сообщения независимо от местоположения. Для этого HLR должен пополнять свои данные. VLR сообщает HLR о положении MS, управляя ею и переприсваивая ей номера в процессе блуждания, посылает все необходимые данные для обеспечения обслуживания подвижной станции.

Интерфейс между MSC (Е-интерфейс) обеспечивает взаимодействие между разными MSC при осуществлении процедуры HANDOVER - "передачи" абонента из зоны в зону при его движении в процессе сеанса связи без ее перерыва.

7.3.4. Структура TDMA кадров и формирование сигналов в стандарте GSM

В результате анализа различных вариантов построения цифровых сотовых систем подвижной связи (ССПС) в стандарте GSM принят многостанционный доступ с временным разделением каналов (TDMA). Общая структура временных кадров показана на рис. 7.10. Длина периода последовательности в этой структуре, которая называется гиперкадром, равна $T_g = 3 \text{ ч } 28 \text{ мин } 53 \text{ с } 760 \text{ мс}$ ($12533,76 \text{ с}$). Гиперкадр делится на 2048 суперкадров, каждый из которых имеет длительность $T_e = 12533,76/2048 = 6,12 \text{ с}$.

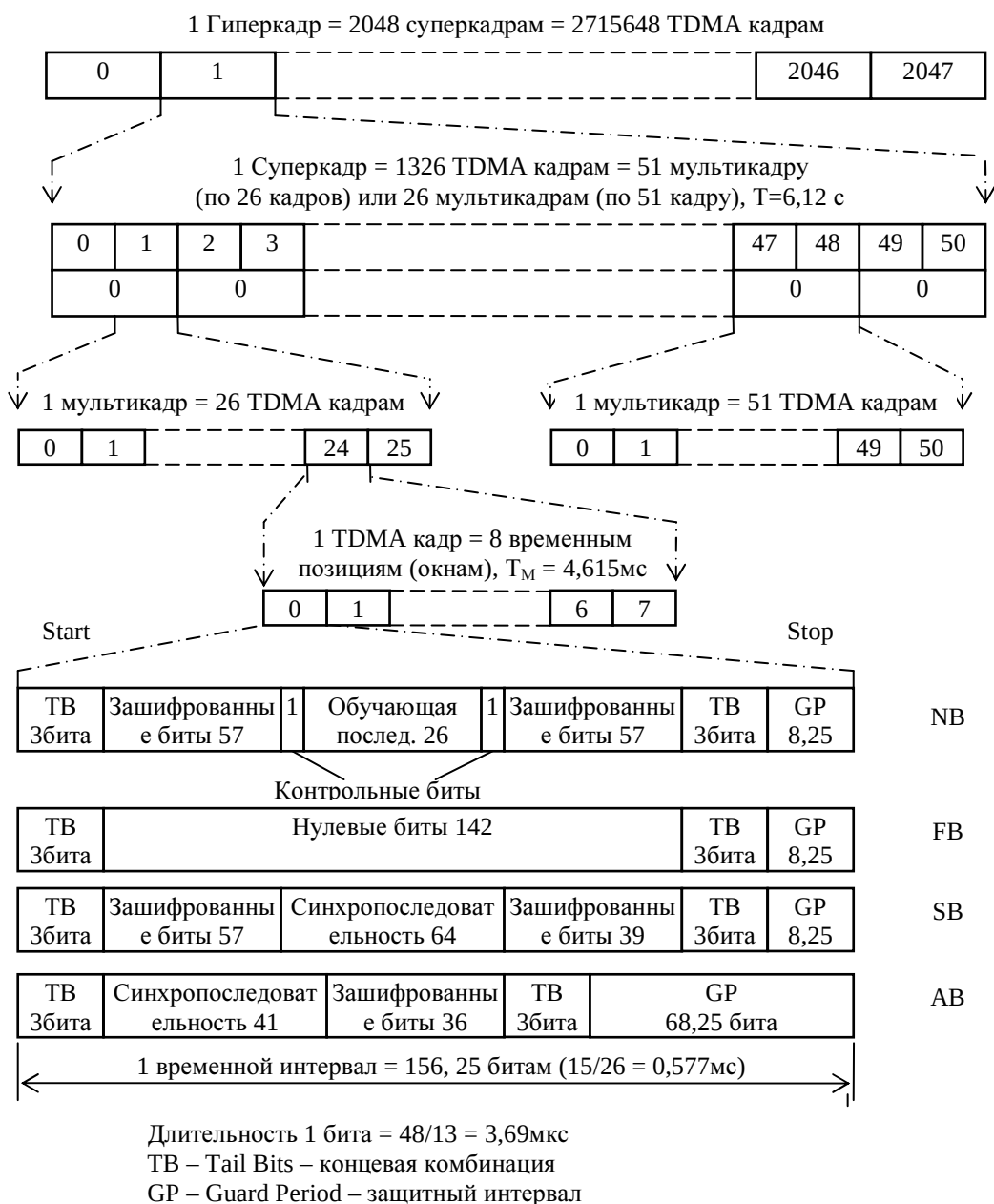


Рис. 7.10. Временная структура кадров стандарта GSM

Таким образом, гиперкадр состоит из 2715647 TDMA кадров. Необходимость такого большого периода гиперкадра объясняется требованиями применяемого процесса криптографической защиты, в котором номер кадра NF используется как входной параметр. TDMA кадр делится на восемь временных позиций с периодом

$$T_0 = 60/13:8 = 576,9 \text{ мкс (15/26 мс)}$$

Каждая временная позиция обозначается TN с номером от 0 до 7. Физический смысл временных позиций, которые иначе называются окнами, - время, в течение которого осуществляется модуляция несущей цифровым информационным потоком, соответствующим речевому сообщению или данным.

Цифровой информационный поток представляет собой последовательность пакетов, размещаемых в этих временных интервалах (окнах). Пакеты формируются немного короче, чем интервалы, их длительность составляет 0,546 мс, что необходимо для приема сообщения при наличии временной дисперсии в канале распространения.

Информационное сообщение передается по радиоканалу со скоростью 270,833 кбит/с.

Это означает, что временной интервал TDMA кадра содержит 156,25 бит.

Длительность одного информационного бита $576,9 \text{ мкс} / 156,25 = 3,69 \text{ мкс}$.

Каждый временной интервал, соответствующий длительности бита, обозначается VN с номером от 0 до 155; последнему интервалу длительностью 1/4 бита присвоен номер 156.

Для передачи информации по каналам связи и управления, подстройки несущих частот, обеспечения временной синхронизации и доступа к каналу связи в структуре TDMA кадра используются пять видов временных интервалов (окон):

NB используется для передачи информации по каналам связи и управления, за исключением канала доступа RACH. Он состоит из 114 бит зашифрованного сообщения и включает защитный интервал (GP) в 8,25 бит длительностью 30,46 мкс. Информационный блок 114 бит разбит на два самостоятельных блока по 57 бит, разделенных между собой обучающей последовательностью в 26 бит, которая используется для установки эквалайзера в приемнике в соответствии с характеристиками канала связи в данный момент времени. Под эквалайзером здесь понимается трансверсальный фильтр, согласованный с каналом (см. раздел 3.2.1.5).

В состав NB включены два контрольных бита (Stealing Flag), которые служат признаком того, содержит ли передаваемая группа речевую информацию или информацию сигнализации. В последнем случае информационный канал (Traffic Channel) "уведен" для обеспечения сигнализации.

Между двумя группами зашифрованных бит в составе NB находится обучающая последовательность из 26 бит, известная в приемнике. С помощью этой последовательности обеспечивается:

- оценка частоты появления ошибок в двоичных разрядах по результатам сравнения принятой и эталонной последовательностей. В процессе сравнения вычисляется параметр RXQUAL, принятый для оценки качества связи. Конечно, речь идет только об оценке связи, а не о точных измерениях, так как проверяется только часть передаваемой информации. Параметр RXQUAL используется при вхождении в связь, при выполнении процедуры "эстафетной передачи" (Handover) и при оценке зоны покрытия радиосвязью;

- оценка импульсной характеристики радиоканала на интервале передачи NB для последующей коррекции тракта приема сигнала за счет использования адаптивного эквалайзера в тракте приема;

- определение задержек распространения сигнала между базовой и подвижной станциями для оценки дальности связи. Эта информация необходима для того, чтобы пакеты данных от разных подвижных станций не накладывались при приеме на базовой станции. Поэтому удаленные на большее расстояние подвижные станции должны

передавать свои пакеты раньше станций, находящихся в непосредственной близости от базовой станции. FB предназначен для синхронизации по частоте подвижной станции. Все 142 бита в этом временном интервале - нулевые, что соответствует немодулированной несущей со сдвигом 1625/24 кГц выше номинального значения частоты несущей. Это необходимо для проверки работы своего передатчика и приемника при небольшом частотном разnose каналов (200 кГц), что составляет около 0,022% от номинального значения диапазона частот 900 МГц. FB содержит защитный интервал 8,25 бит так же, как и нормальный временной интервал. Повторяющиеся временные интервалы подстройки частоты (FB) образуют канал установки частоты (FCSN).

SB используется для синхронизации по времени базовой и подвижной станций. Он состоит из синхропоследовательности длительностью 64 бита, несет информацию о номере TDMA кадра и идентификационный код базовой станции. Этот интервал передается вместе с интервалом установки частоты. Повторяющиеся интервалы синхронизации образуют так называемый канал синхронизации (SCH).

DB обеспечивает установление и тестирование канала связи. По своей структуре DB совпадает с NB (рис. 1.6) и содержит установочную последовательность длиной 26 бит. В DB отсутствуют контрольные биты и не передается никакой информации. DB лишь информирует о том, что передатчик функционирует.

AB обеспечивает разрешение доступа подвижной станции к новой базовой станции. AB передается подвижной станцией при запросе канала сигнализации. Это первый передаваемый подвижной станцией пакет, следовательно, время прохождения сигнала еще не измерено. Поэтому пакет имеет специфическую структуру. Сначала передается концевая комбинация 8 бит, затем - последовательность синхронизации для базовой станции (41 бит), что позволяет базовой станции обеспечить правильный прием последующих 36 зашифрованных бит. Интервал содержит большой защитный интервал (68,25 бит, длительностью 252 мкс), что обеспечивает (независимо от времени прохождения сигнала) достаточное временное разнесение от пакетов других подвижных станций,

Этот защитный интервал соответствует двойному значению наибольшей возможной задержки сигнала в рамках одной соты и тем самым устанавливает максимально допустимые размеры соты. Особенность стандарта GSM - возможность обеспечения связи подвижных абонентов в сотах с радиусом около 35 км. Время распространения радиосигнала в прямом и обратном направлениях составляет при этом 233,3 мкс.

Одна из особенностей формирования сигналов в стандарте GSM - использование медленных скачков по частоте в процессе сеанса связи. Главное назначение таких скачков (SFH - Slow Frequency Hopping) - обеспечение частотного разнесения в радиоканалах, функционирующих в условиях многолучевого распространения радиоволн. SFH используется во всех подвижных сетях, что повышает эффективность кодирования и перемежения при медленном движении абонентских станций. Принцип формирования медленных скачков по частоте состоит в том, что сообщение, передаваемое в выделенном абоненту временном интервале TDMA кадра (577 мкс), в каждом последующем кадре передается (принимается) на новой фиксированной частоте. В соответствии со структурой кадров время для перестройки частоты составляет около 1 мс.

В процессе скачков по частоте постоянно сохраняется дуплексный разнос 45 МГц между каналами приема и передачи. Всем активным абонентам, находящимся в одной соте, ставятся в соответствие ортогональные формирующие последовательности, что исключает взаимные помехи при приеме сообщений абонентами в соте. Параметры последовательности переключения частот (частотно-временная матрица и начальная

частота) назначаются каждой подвижной станции в процессе установления канала. Ортогональность последовательностей переключения частот в соте обеспечивается начальным частотным сдвигом одной и той же (по алгоритму формирования) последовательности. В смежных сотах используются различные формирующие последовательности.

Принятая структура TDMA кадров и принципы формирования сигналов в стандарте GSM в совокупности с методами канального кодирования позволили снизить требуемое для приема отношение сигнал/помеха до 9 дБ, тогда как в стандартах аналоговых сотовых сетей связи оно составляет 17-18 дБ.

7.4. Радиоинтерфейс системы cdmaOne (IS-95)

7.4.1. Общая характеристика системы CDMAOne

С опорой на преимущества, предоставляемые технологией CDMA, компанией Qualcomm (США) была разработана ССМС общего пользования с кодовым разделением каналов. При этом основными стимулами явились увеличение абонентской емкости по сравнению с действующими ССМС, основанными на традиционном частотно/временном доступе, улучшение качества обслуживания и повышение информационной безопасности. Технические требования к системе сформулированы в ряде стандартов Ассоциации Промышленности связи (*Telecommunications Industry Association – TIA*), определяющих характер и порядок взаимодействия различных функциональных узлов ССМС и условия совместимости аппаратуры различных фирм–производителей.

Современные сети, базирующиеся на стандарте IS-95, обеспечивают передачу сигнала со скоростью 9,6 кбит/с (с кодированием) и 14,4 кбит/с (без кодирования), тогда как исходные спецификации cdmaOne предполагали скорость передачи 8 кбит/с, 13 кбит/с и 8 кбит/с EVRC (Enhanced Variable Rate Vocoder). В настоящее время широко применяется версия стандарта **IS-95A**.

Версия IS-95B основана на объединении нескольких каналов CDMA, организуемых в прямом направлении (от базовой станции к мобильной). Скорость может увеличиваться до 28,8 кбит/с (при объединении двух каналов по 14,4 кбит/с) или до 115,2 кбит/с (8 каналов по 14,4 кбит/с). Сети на основе IS-95B смогут обеспечивать доступ в Internet до появления систем третьего поколения. Однако для того, чтобы предоставлять услуги пакетной передачи, контроллер БС нужно дооснастить маршрутизатором. В спецификациях стандарта предусмотрено качественное улучшение характеристик обслуживания за счет снижения потерь при переходе абонента от одной БС к другой, а также повышение точности контроля мощности до 0,25 дБ, организация каналов приоритетного доступа и другие усовершенствования.

В версии IS-95C (CDMA 2000) модификации коснулись повышения частотной эффективности и увеличения емкости телефонной сети в два раза. Спецификациями предусматривается дополнительный канал с ортогональным сдвигом несущей, по которому может передаваться полный кодовый ансамбль сигналов (т. е. 64 кода Уолша), такой же, как и по синфазному каналу. Системы на базе **CDMA 2000** обратно совместимы с сетями на основе **IS-95A** и **IS-95B** и сохраняют прежнюю полосу частот- 1,25 МГц. По сравнению с предыдущими версиями скорость передачи в системе возрастает до 144 кбит/с; при этом сокращается энергопотребление терминала.

К числу основных стандартов рассматриваемой ССМС относятся:

IS-95 – эфирный интерфейс (радиоинтерфейс);

IS-96 – интерфейс речевых служб;

IS-97 – интерфейс мобильной станции;

IS-98 – интерфейс базовой станции;

IS-99 – интерфейс службы передачи данных.

Учитывая особую значимость радиоинтерфейса, рассматриваемую систему называют ССМС стандарта IS-95, либо, принимая во внимание технологию организации множественного доступа, системой cdmaOne.

Система IS-95 рассчитана на работу в диапазоне частот 800 МГц, причем для прямого канала (линия "вниз") выделен участок спектра 869,04...893,97 МГц, а для обратного (линия "вверх") – 824,04...848,96 МГц. Ширина полосы канала связи составляет 1,25 МГц, поэтому при развертывании ССМС IS-95 операторы могут осуществлять частотное планирование исходя из указанных полос. Однако согласно решению Федеральной Комиссии связи США, одному оператору может быть выделен максимальный диапазон частот, равный 12,5 МГц, как в прямом (от БС к МС), так и в обратном (от МС к БС) направлениях, что соответствует 10 физическим частотным радиоканалам с полосой в 1,25 МГц.

Как указывалось в разделе 6.3, технология организации множественного доступа с кодовым разделением каналов основана на применении сложных (*spread spectrum*) сигналов, полоса которых значительно превышает ширину спектра информационного сообщения. В системе IS-95 реализовано прямое расширение спектра (DSSS) с использованием функций Уолша длины 64 и двух типов псевдослучайных последовательностей (ПСП): короткой и длинной (таблица 4.3). Линия "вниз" организована на основе синхронного варианта CDMA, использующего для разделения физических абонентских каналов ансамбль сигнатур в виде функций Уолша. Асинхронный режим CDMA в обратном канале осуществлен приписыванием различным абонентам одной соты специфических сдвигов длинной ПСП.

Согласованные фильтры приемников БС квазиоптимальны в условиях взаимной интерференции между абонентами одной соты и весьма чувствительны к неодинаковым уровням сигналов от МС, находящимся на различных расстояниях от БС (эффект «далеко-близко»). Для максимизации абонентской емкости системы необходимо, чтобы терминалы всех абонентов излучали сигнал такой мощности, которая обеспечила бы одинаковый уровень принимаемых БС сигналов. Чем точнее управление мощностью, тем больше абонентская емкость системы.

Выравнивание мощностей сигналов МС на входе приемника БС является непременным условием работоспособности линии "вверх" системы CDMA. В стандарте IS-95 применена быстродействующая петля автоматической регулировки мощности сигналов МС, принцип действия которой описан в разделе 4.6.5. Регулировка осуществляется в динамическом диапазоне 84 дБ с шагом $\pm 0,5$ дБ и с периодичностью от нескольких микросекунд до 1,25 мс. Резерв фактора речевой активности также утилизирован в системе посредством использования речепреобразующего устройства (вокодера) с переменной частотой преобразования аналогового речевого сигнала в цифровой. В зависимости от активности абонента вокодер формирует потоки данных со скоростями 8,6; 4; 2 и 0,8 кбит/с.

Требуемое качество передачи данных в системе достигается с помощью довольно мощного канального кодирования, выполняемого в несколько этапов. На первом, предварительном этапе цифровой речевой сигнал с выхода вокодера, структурированный в кадры длительностью в 20 мс, кодируется блоковым циклическим кодом (CRC) для формирования индикатора качества кадра и дополняется "хвостом", необходимым для последующего сверточного кодирования. Введение подобной избыточности увеличивает фактические скорости данных, заменяя первоначальный их набор на 9,6; 4,8; 2,4 и 1,2 кбит/с. Следующим этапом является сверточное кодирование. В прямом канале используется код с длиной кодового ограничения 9 и скоростью $1/2$, тогда как код в обратном канале, обладающем меньшей помехоустойчивостью (мощность передатчика МС весьма мала), имеет вдвое большую избыточность, т.е. скорость $1/3$ при той же длине кодового ограничения. Наконец, на третьем этапе выполняется перемежение информационного потока в кадре для нейтрализации эффекта пакетирования ошибок.

Помимо перемежения для борьбы с быстрыми замираниями в системе использовано и многолучевое разнесение, т.е. приемники на основе алгоритма RAKE. Для этого на БС используется минимум четыре, а на МС – три параллельно работающих коррелятора.

Помимо этих корреляторов, настраиваемых на определенную задержку, в каждом приемнике имеется еще и сканирующий по задержке канал, позволяющий осуществлять настройку RAKE-каналов на сигналы с наибольшей интенсивностью.

Наличие нескольких параллельных каналов корреляционной обработки позволяет осуществить *мягкую эстафетную передачу*, (*soft handoff* или *soft handover*) при переходе МС из одной соты в другую. В процессе мягкой эстафетной передачи МС может поддерживать соединение с двумя и более БС, выбирая сигнал с большей интенсивностью. Указанная процедура позволяет поддерживать высокое качество связи при переключении МС с одной БС на другую и делает эстафетную передачу практически незаметной для пользователя.

Стандарт IS-95 обеспечивает высокую степень безопасности передаваемых данных за счет их скремблирования выборками из вышеупомянутой длинной псевдослучайной последовательности. Ключ (маска) скремблирования индивидуален для каждой МС и формируется по секретному правилу на базе ее идентификационного номера.

Основные характеристики стандарта IS-95 и технические параметры приемо-передающего оборудования представлены в табл. 7.1.

Для систем с CDMA, и, в частности для CCMC IS-95, характерна определенная асимметрия линий "вниз" и "вверх": в первой из них (от БС) все абонентские сигналы передаются из одной пространственной точки и потому могут быть легко синхронизированы (синхронный вариант CDMA). В направлении же от БС к МС приходится применять асинхронное кодовое разделение. Отметим, что временные шкалы отдельных БС сети IS-95 также синхронизированы. Для этого каждая БС оборудована приемником радионавигационной системы GPS NAVSTAR. Благодаря единому системному времени, кодовое разделение сигналов различных БС, занимающих один и тот же частотный канал, осуществлено за счет применения различных сдвигов одной и той же короткой ПСП.

Таблица 7.1. Основные характеристики стандарта IS-95

Параметр стандарта	Значение
Диапазон частот передачи, МГц:	
МС	824,040...848,860
БС	869,040...893,970
Относит. нестабильность несущей частоты:	
МС	$\pm 2,5 \cdot 10^{-6}$
БС	$\pm 5 \cdot 10^{-8}$
Вид модуляции несущей частоты:	
МС	OQPSK
БС	QPSK
Ширина спектра излучаемого сигнала, МГц	
по уровню –3 дБ	1,23
по уровню –40 дБ	1,5
Тактовая частота ПСП, МГц	1,2288
Количество каналов на одной несущей:	
МС: канал доступа	1
канал обратного трафика	1
БС: пилотный канал	1
канал синхронизации	1
канал персонального вызова	1...7
канал прямого трафика	61...55
Скорость передачи данных в канале, бит/с:	
канал синхронизации	1200
канал персонального вызова и доступа	4800,9600

канал трафика	1200, 2400, 4800, 9600
Кодирование информации в каналах: МС: канал доступа и обратного трафика БС: канал синхронизации, персонального вызова и прямого трафика.	Сверточ. код с длиной кодового ограничения 9 и скоростью 1/3. Сверт. код с длиной кодового ограничения 9 и скоростью 1/2.
Требуемое отношение сигнал/шум, дБ	6...7
Макс. эффектив. излучаемая мощность, Вт: МС: (3-го, 2-го и 1-го класса) БС	1,0; 2,5; 6,3 до 50
Точность управления мощностью передатчика, дБ	$\pm 0,5$
Чувствительность приемника, дБм: МС БС	-105 -117

Для модуляции сигнала используется три вида функций: «короткая» и «длинная» ПСП и функции Уолша порядков от 0 до 63. Все они являются общими для базовых и мобильных станций, однако реализуют разные функции (табл. 7.2).

Таблица 7.2. Параметры кодовых последовательностей в стандарте IS-95

Тип сигнала	Длина кода	Выполняемые функции	
		Базовая станция	Мобильная станция
Код Уолша	64	Кодовое уплотнение или разделение 64 каналов CDMA	Помехоустойчивое кодирование
Короткий код	32768	Разделение сигналов базовых станций по величине циклического сдвига	Код с одинаковым фиксированным циклическим сдвигом - как опорный сигнал скремблера
Длинный код	$2^{42}-1$ (4.4×10^{12})	Прореженный длинный код - как опорная последовательность скремблера	Длинный код с разными циклическими сдвигами - как адресная последовательность

7.4.2. . Архитектура линии "вниз"

Логические каналы линии "вниз" включают:

- *пилотный канал* (*pilot channel* – PI);
- *канал синхронизации* (*synchronization channel* – SYNC);
- *канал персонального вызова* (*paging channel* – PCH);
- *канал прямого трафика* (*forward traffic channel* – TCH).

Отображение логических каналов на физические в прямом направлении осуществляется с помощью системы ортогональных функций Уолша длины 64: w_i , $i = 0, 1 \dots 63$, где i – номер функции Уолша. Стандартом IS-95 предусматривается организация одного пилотного канала, одного канала синхронизации, от одного до семи каналов вызова (в зависимости от абонентской нагрузки на БС) и от 55 до 62 каналов прямого трафика, поскольку часть каналов вызова допускается использовать в качестве каналов трафика. Соответствие между логическими и физическими каналами показано на рис. 7.11.

Прямой канал связи										
Пилот канал	Синхро канал	Канал вызова	...	Канал вызова	Канал прям. траф.	...	Канал прям. траф.	Канал прям. траф.	...	Канал прям. траф.
w_0	w_{32}	w_1	...	w_7	w_8	...	w_{31}	w_{33}	...	w_{63}

Рис. 7.11. Структура прямого канала CCMC стандарта IS-95

Укрупненная схема преобразований сигнала в передающем тракте базовой станции показана на рис. 7.12:

1. Для различения разных физических каналов данной БС в прямом канале от БС к подвижной – кодирование сигнала функциями Уолша и бинарная фазовая манипуляция.
2. Для шифрования сообщений используется кодирование длинной ПСП и бинарная фазовая манипуляция.
3. Для расширения полосы и различения сигналов разных БС используется кодирование короткой ПСП и квадратурная фазовая манипуляция двумя ПСП одинакового периода.

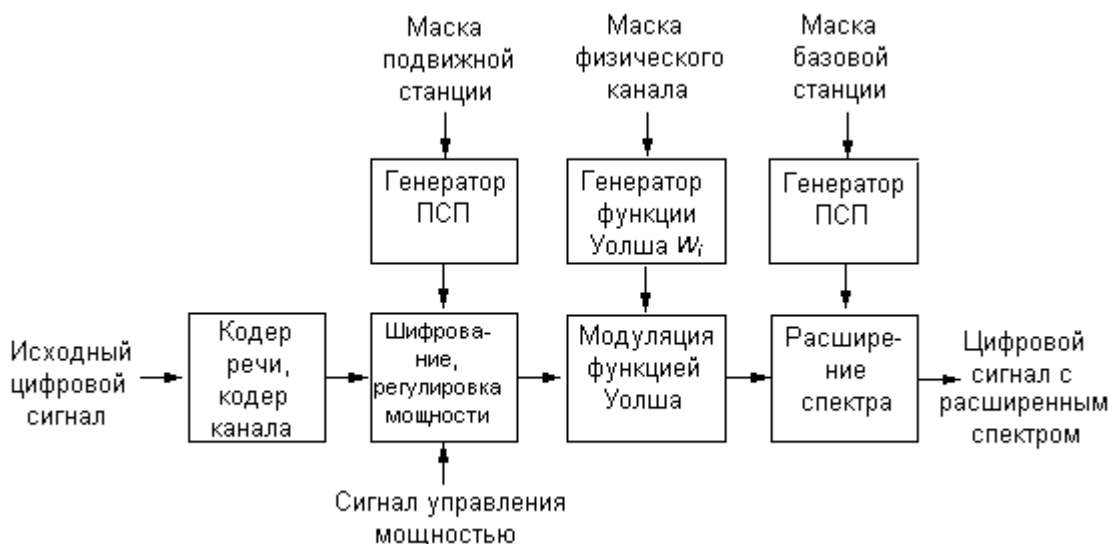


Рис. 7.12. Обработка сигналов в передающем тракте базовой станции

4. Различение сигналов разных станций обеспечивается тем, что все БС используют одну и ту же пару коротких ПСП, но со сдвигом на 64 чипа (дискрета ПСП частотой 1,2288 МГц) между разными станциями. Т.е. всего в сети имеется 511 кодов; при этом все физические каналы одной БС имеют одну и ту же фазу последовательности.

7.4.3. Пилотный канал

Пилотный канал предназначен для начальной синхронизации МС с сетью и оценки параметров прямого канала приемником МС. Мощность, отведенная пилотному каналу, обычно на 4...6 дБ превышает мощность в канале трафика. Тем самым обеспечиваются, с одной стороны, условия для надежного ввода и последующей устойчивой работы петель автоподстройки фазы и задержки когерентного приемника МС, а с другой – точность измерений параметров сигнала БС, достаточная для процедур RAKE, эстафетной передачи и управления мощностью БС.

Согласно рис. 7.11, пилотному каналу присвоена нулевая функция Уолша w_0 , т.е. последовательность из одних нулей. Информационные данные по этому каналу не

передаются, т.е. могут также трактоваться как последовательность из одних нулей. После сложения по модулю два этих двух последовательностей результирующий поток (также состоящий из одних нулей)² поступает на преобразователь, превращающий булевы значения символов (0,1) в двуполярные отсчеты (+1,-1) соответственно. После объединения с сигналами других физических каналов (см. ниже) суммарный сигнал поступает на КФМ модулятор, в котором подвергается прямому расширению спектра, перемножаясь с двумя псевдослучайными скремблирующими последовательностями ПСП-I и ПСП-Q (символы I и Q отвечают синфазной и квадратурной составляющим). Период каждой из названных ПСП содержит 2^{15} чипов, частота следования которых согласно стандарту составляет 1,2288 Мчип/с. Прямой подсчет показывает, что на одном двухсекундном отрезке уместается в точности 75 периодов коротких ПСП. Структурно короткие ПСП представляют собой M -последовательности длины $N = 2^{15} - 1$ с характеристическими полиномами

$$f_I = x^{15} + x^{13} + x^9 + x^8 + x^7 + x^5 + 1$$

и

$$f_Q = x^{15} + x^{12} + x^{11} + x^{10} + x^6 + x^5 + x^4 + x^3 + 1,$$

расширенные добавлением нулевого символа к цепочке из 14 последовательных нулей на каждом периоде

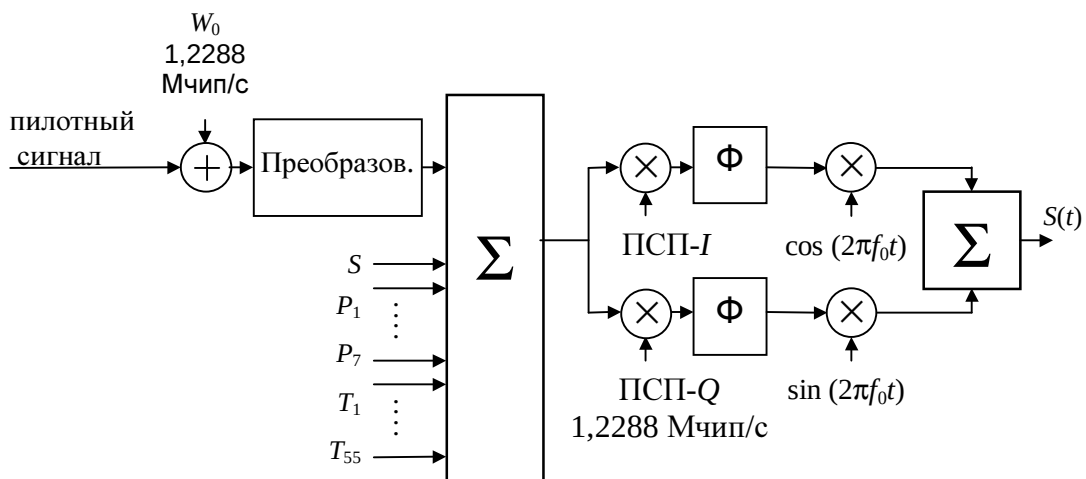


Рис. 7.13. Структурная схема пилотного канала

Как можно видеть, фактически в пилотном канале передается только пара ПСП-I и ПСП-Q или, что эквивалентно комплексная ПСП. Данная комплексная короткая ПСП одинакова для всех 64 CDMA-каналов и используется всеми БС системы, но с разными циклическими сдвигами. Разница в циклических сдвигах позволяет МС разделять сигналы, излучаемые БС разных сот или секторов, т.е. позволяет идентифицировать номер БС либо сектора. Для различных БС сдвиг изменяется с постоянным шагом, равным 64 чип $\times$ PILOT_INC, где системный параметр PILOT_INC принимает значения от 1 до 4. Таким образом, при минимальном шаге доступны $2^{15} / 2^6 = 512$ сдвигов коротких ПСП, т.е. возможно бесконфликтное существование сети, состоящей из 512 БС. Если же необходимо, чтобы сеть состояла из большего числа БС, то при территориальном планировании сети легко можно добиться, чтобы БС с одинаковыми циклическими сдвигами коротких ПСП не могли одновременно находиться в зоне радиовидимости МС.

С другой стороны, шаг сдвига ПСП однозначно определяет размер соты (или сектора), при котором МС с гарантией различает ПСП, имеющие минимальный

² Тривиальные нулевые последовательности данных и W_0 явно показаны на рис. 4.15, чтобы сохранить единообразие с последующими схемами.

временной сдвиг. Нетрудно убедиться, что при минимальном сдвиге в 64 чипа радиус соты составит порядка 15,5 км.

После перемножения с квадратурными ПСП сигнал в каждом из квадратурных плеч КФМ-модулятора фильтруется (фильтры Φ на рис. 7.13) для формирования приемлемого спектра и перемножается со сдвинутыми на 90° гармоническими колебаниями центральной частоты. Суммирование выходов квадратурных плеч дает модулированный сигнал, в котором пилотная компонента совпадает с произведением пилот-сигнала (т.е. константы) с четырехфазно манипулированной несущей.

7.4.4. Канал синхронизации

После вхождения в синхронизм с пилот-сигналом МС располагает информацией о фазе принимаемой несущей, временных границах чипов и периодов короткой ПСП (один период занимает $2^{15}/1,2288 \cdot 10^6 = 26,666... \text{ мс}$), тогда как для приема сообщений необходимо знать границы кадров (каждый кадр трафика занимает 20 мс). Соответствующие сведения наряду с другими, нужными для установления и поддержания соединения, МС получает по каналу синхронизации. В формате сообщения этого канала содержатся данные о точном времени в системе (SYS_TIME), значении циклического сдвига короткой ПСП данной БС (PILOT_PN), идентификаторы БС и ЦКПС (MSC), значение мощности сигнала в пилотном канале, параметры "длинной" ПСП (LC_STATE), скорость передачи данных в канале персонального вызова (PRAT). Структура физического канала синхронизации поясняется рис. 7.14.

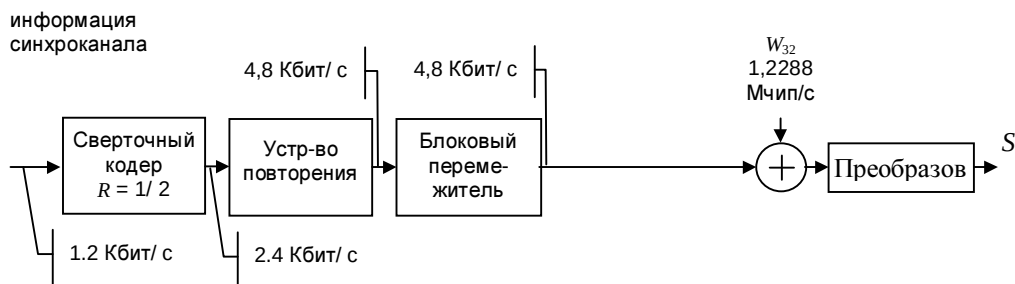


Рис. 7.14. Структурная схема канала синхронизации

Данные канала синхронизации, поступающие со скоростью 1200 бит/с, подаются на вход сверточного кодера. После осуществления операции кодирования (длина кодового ограничения – 9, скорость кода $1/2$) с выхода снимается двоичная последовательность со скоростью 2400 бит/с, поступающая далее на устройство повторения, после которого скорость потока данных удваивается до 4800 бит/с. Информация, передаваемая по каналу, структурирована в кадры длительностью в 26,666... мс. Каждый кадр совпадает с одним периодом короткой ПСП и содержит 32 бита исходных данных (128 символов с учетом сверточного кодирования и повторения). Как видно из рис. 7.14, следующей операцией является блочное перемежение в пределах кадра синхроканала, служащее для борьбы с пакетными ошибками. После блочного перемежителя поток данных подвергается прямому расширению спектра путем сложения по модулю два с присвоенной каналу синхронизации функцией Уолша w_{32} и преобразования булевых переменных в действительные $+1$ и -1 . Как можно видеть, каждому биту информационного потока с выхода перемежителя (скорость 4,8 кбит/с) сопоставляется $1,2288 \cdot 10^6 / 4,8 \cdot 10^3 = 256$ чипов, т.е. четыре периода последовательности Уолша. Иными словами, каждый бит кодированного потока в зависимости от значения заменяется четырьмя периодами прямой или инвертированной функции Уолша w_{32} .

Последующие операции в канале повторяют рассмотренные ранее: сигнал канала синхронизации объединяется с сигналами других каналов, поступая на вход S на рис. 7.12,

после чего в КФМ модуляторе перемножается с комплексной короткой ПСП (скремблируется) и переносится на несущую.

Каждые три кадра (96 информационных бит) в канале синхронизации образуют один суперкадр длительности 80 мс, соответствующий четырем кадрам канала трафика (см. ниже). Сообщение, передаваемое по каналу синхронизации, может занимать более одного суперкадра, вследствие чего передаваемые данные подвергаются определенной структурной организации, называемой *капсулированием*. Полученная в результате подобного преобразования капсула состоит из самого сообщения синхроканала и добавочных бит, заполняющих оставшееся битовое пространство до начала следующего суперкадра. Необходимость подобной упаковки данных объясняется тем, что начало любого сообщения всегда должно совпадать с началом суперкадра.

Каждый кадр синхроканала начинается битом-заголовком, называемым *битом начала сообщения* (*start-of-message* – SOM). Значение этого бита, равное единице, указывает на начало сообщения, передаваемого по синхроканалу, а значение, равное нулю, свидетельствует о том, что текущий кадр содержит сообщение, начавшееся в некотором предыдущем кадре. В результате БС может передавать синхросообщение, занимающее несколько кадров подряд. Следует также отметить, что значение SOM-бита, равное 1, всегда должно совпадать с началом суперкадра, поскольку, как уже указывалось ранее, начало любого сообщения синхроканала всегда отвечает началу суперкадра. Ситуацию, соответствующую передаче по синхроканалу сообщения, занимающего два последовательных суперкадра, иллюстрирует рис. 7.15.



Рис. 7.15. Структура кадров канала синхронизации

Состав оборудования сетей стандарта CDMA IS-95 во многом сходен с составом оборудования сетей стандарта GSM и включает в себя ПС и БС, цифровые коммутаторы, центр управления и обслуживания, различные дополнительные системы и устройства, функциональное сопряжение элементов системы осуществляется с помощью ряда

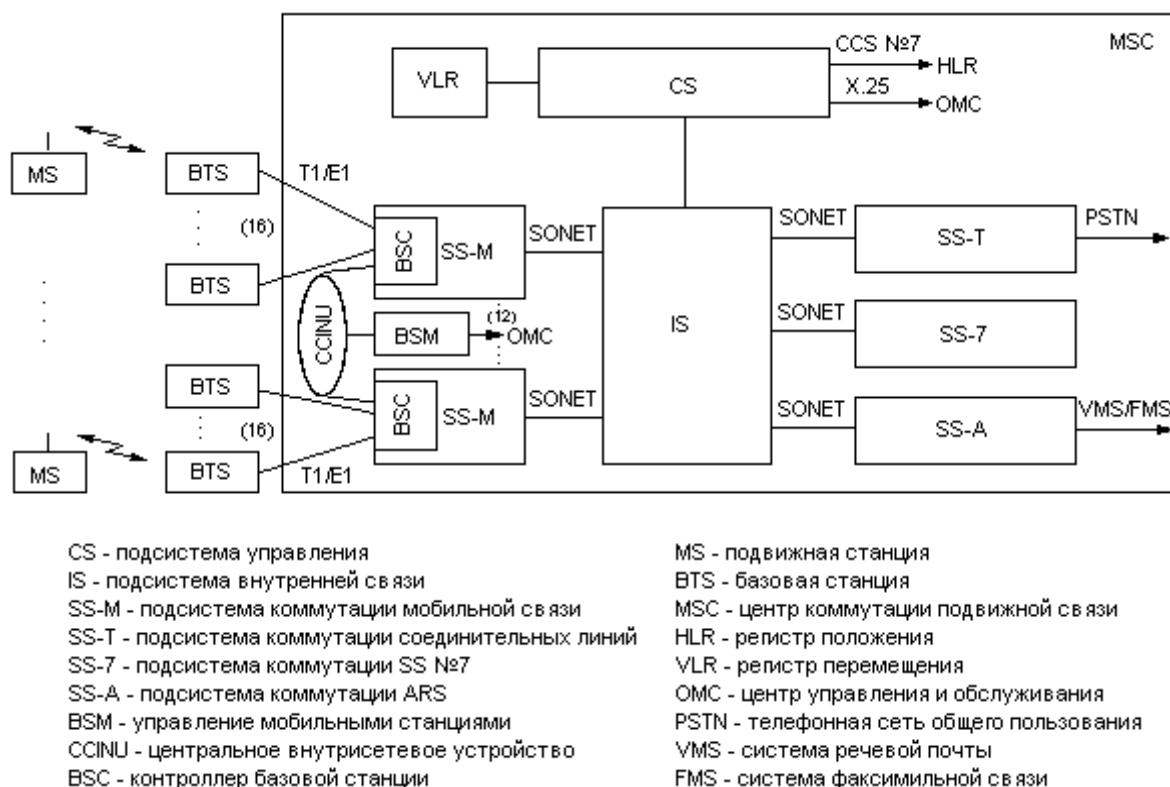


Рис. 7.16. Конфигурация системы стандарта IS-95

интерфейсов. Конфигурация системы стандарта CDMA IS-95 представлена на рис. 7.16 (SONET – линия оптической связи).

Для обеспечения требуемой помехоустойчивости в каналах системы CDMA применяется *сверточное кодирование* со скоростью 1/2 (в прямом канале) и 1/3 (в обратном канале), *декодер Витерби* с мягким решением и *перемежение* передаваемых сообщений.

Передача речи и данных по стандарту IS-95 осуществляется кадрами длительностью 20 мс. При этом скорость передачи в пределах сеанса связи может изменяться от 1,2 до 9,6 кбит/с, но в течение одного кадра она остается неизменной. Если количество ошибок в кадре превышает допустимую норму, то искаженный кадр удаляется.

7.4.5. Регулировка мощности в стандарте IS-95

Процесс регулирования мощности передающих устройств в ОК (от абонента к БС) заключается в следующем. Каждая МС непрерывно передает информацию об уровне ошибок в принимаемом сигнале. На основании этой информации БС распределяет излучаемую мощность между абонентами таким образом, чтобы в каждом случае обеспечить приемлемое качество речи. Абоненты, на пути к которым радиосигнал испытывает большее затухание, получают возможность излучать сигнал большей мощности. Основная цель регулировки мощности в ОК - оптимизация площади соты.

В процессе регулирования мощности в прямом канале (от БС к абоненту) возможны два варианта регулирования: по открытому циклу (разомкнутая петля) и по замкнутому циклу (замкнутая петля). Схема управления мощностью в прямом канале изображена на рис. 7.17.

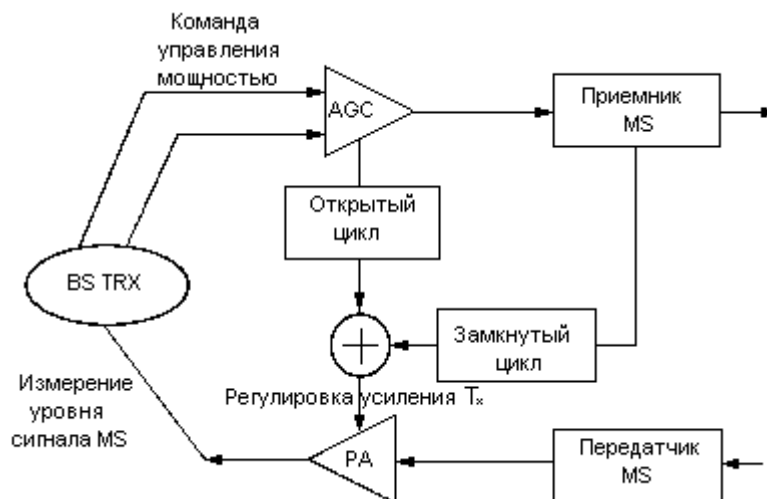


Рис. 7.17. Схема управления мощностью в прямом канале

При открытом цикле МС после включения ищет сигнал БС. После синхронизации МС по этому сигналу производится замер его мощности и вычисляется мощность передаваемого сигнала, необходимая для обеспечения соединения с БС. Вычисления основываются на том, что сумма уровней предполагаемой мощности излучаемого сигнала и мощности принятого сигнала должна быть постоянна и равна 73 дБ. Этот процесс повторяется каждые 20 мс, но он все же не обеспечивает желаемой точности регулировки мощности, так как прямой и обратный каналы работают в разных частотных диапазонах (разнос частот 45 МГц) и, следовательно, имеют различные уровни затухания при распространении и по-разному подвержены воздействию помех.

При замкнутом цикле возможно точно отрегулировать мощность передаваемого сигнала. БС постоянно оценивает вероятность ошибки в каждом принимаемом сигнале. Если она превышает программно заданный порог, то БС дает команду соответствующей МС увеличить мощность излучения. Регулировка осуществляется с шагом 1 дБ. Этот процесс повторяется каждые 1,25 мс. Цель такого процесса регулирования заключается в том, чтобы каждая МС излучала сигнал минимальной мощности, которая достаточна для обеспечения приемлемого качества речи. За счет того, что все МС излучают сигналы необходимой для нормальной работы мощности, и не более, их взаимное влияние минимизируется, и абонентская емкость системы возрастает. МС должны обеспечивать регулирование выходной мощности в широком динамическом диапазоне - до 85 дБ.

При процедуре мягкой эстафетной передачи (переходе абонента из зоны обслуживания одной БС в зону другой) схема регулирования мощности несколько иная. МС принимает одновременно несколько команд управления мощностью от разных БС (обычно двух) и сравнивает их между собой. Если все команды указывают на необходимость увеличения мощности, то МС последовательно увеличивает свою мощность с шагом 1 дБ.

Регулирование мощности как в прямом, так и в обратном канале влияет на срок службы аккумуляторов МС. Средняя излучаемая мощность МС в CDMA меньше, чем в системах, использующих другие методы доступа. Это непосредственно связано с такими параметрами радиотелефона, как длительность непрерывного занятия канала и время нахождения в режиме ожидания.

7.5. Деятельность МСЭ по развитию сетей UMTS

Общая Международная организация электросвязи / International Telecommunications Union (МСЭ/ITU) начал исследования вопросов глобализации мобильной связи в 1986 г. и определил долгосрочные требования к частотному спектру для будущего третьего поколения систем мобильной связи. В 1992 г. МСЭ на Всемирной административной конференции радиосвязи (ВАКР-92) распределил 230 МГц спектра в диапазоне 2 ГГц глобальной системе мобильной связи третьего поколения для обоих сегментов — спутникового и наземного.

Ожидалось, что появление на телекоммуникационном рынке (более 20 лет назад) сетей мобильной связи 3G завершит процесс глобализации мобильной связи в мире, начатый созданием и внедрением сетей связи второго поколения (сетей 2G), несмотря на национальные и региональные интересы и трудности при развертывании сетей.

Вначале МСЭ предполагал разработать единый глобальный стандарт на технологию радиointерфейса системы 3G. Этот проект назывался «Future Public Land Mobile Telephony System» (FPLMTS) и позднее был переименован в international Mobile Telephone System» (IMT-2000). Однако техническая несовместимость радиointерфейсов и сетевых платформ, а также конкурентная борьба между производителями оборудования за технологическое лидерство не позволили МСЭ установить единый гармонизированный всемирный стандарт. В связи с этим рабочими органами МСЭ было принято решение об одобрении семейства стандартов радиодоступа IMT-2000, в которое после завершения процесса гармонизации вошли пять радиointерфейсов. Использование технологии модуляции OFDM существенно повысило спектральную эффективность систем беспроводного доступа WiMAX (IEEE 802.16e). В связи с этим МСЭ в 2007 г. ввел в состав семейства радиointерфейсов новый интерфейс IMT Advanced (с технологией OFDM). Полное семейство интерфейсов и их основные характеристики приведены в табл. 7.3. В табл. 7.3 использованы следующие обозначения радиointерфейсов: - IMT-DS (IMT-2000 Direct Spread) — радиointерфейс широкополосной системы связи с прямым расширением спектра (DS-CDMA) и частотным дуплексным разнесом (FDD) для применения в парных полосах частот;

Таблица 7.3. Типы и значения характеристик семейства радиointерфейсов стандарта IMT

Характеристика радиointерфейса	Семейство радиointерфейсов					
	IMT-DS	IMT-MC	IMT-TC	IMT-SC	IMT-FT	IMT Advanced
Орган по разработке спецификаций	3GPP, ARIB, ETSI	3GPP2, TTA, TR-45.3	3GPP, ETSI, CWTS	3GPP2, UWCC, TR-45.3, TTA	ETSI	ETSI, 3GPP
Базовая технология	W-CDMA, UTRA FDD	cdma2000	UTRA TDD, TD-SCDMA	UWC-136	DECT EP	OFDMA/ SC-FDMA
Метод доступа	DS-CDMA	MC-CDMA	TDMA/CDMA	TDMA	MC-TDMA	OFDMA
Метод	FDD	FDD	TDD	FDD	FDD/TDD	FDD

дуплексного разнеса						
Канальная скорость манипуляции, Мбод (Мсимв/с)	3,84		3,841 ¹⁾ 1,282 ²⁾			(0,5x3,84)... (8x3,84)
Для технологии	UTRA TDD. 2)	Для технологии	TD-SCDMA.			

- IMT-MC (IMT-2000 Multi Carrier) — радиоинтерфейс многочастотной системы связи с одновременной передачей нескольких несущих MS-CDMA и частотным дуплексным разнесом FDD для применения в парных полосах частот;

- IMT-TC (IMT-2000 Time-Code) — радиоинтерфейс комбинированной системы связи TDMA/CDMA с временным дуплексным разнесом TDD для применения в непарных полосах частот;

- IMT-SC (IMT-2000 Single Carrier) — радиоинтерфейс одночастотной системы связи TDMA с временным дуплексным разнесом TDD для применения в парных полосах частот;

- IMT-FT (IMT-2000 Frequency Time) — радиоинтерфейс микросотовой системы связи DECT с комбинированным частотно-временным дуплексным разнесом FDD/TDD для применения как в парных, так и в непарных полосах частот;

- IMT Advanced — радиоинтерфейс системы связи с одновременной передачей нескольких ортогональных несущих OFDMA и частотным дуплексным разнесом FDD.

Семейство из пяти радиоинтерфейсов было представлено в Рекомендации МСЭ-Р М.1457 «Детальные спецификации радиоинтерфейсов международной подвижной электросвязи — 2000 (IMT-2000)». Ожидается, что в ближайшее время эта рекомендация будет дополнена описанием радиоинтерфейса IMT Advanced.

Распределение полос частот для сетей UMTS/IMT-2000 было закреплено в Регламенте радиосвязи на Всемирной конференции радиосвязи в 1997 г. (ВКР-97) в соответствии с Резолюцией 212 в виде специального примечания S5.388. Это распределение частотных полос по различным районам и отдельным государствам показано на рис. 7.18. При этом полосы частот 1885... 1980, 2010...2025 и 2110...2170 МГц были определены для наземного сегмента, а полосы частот 1980...2010 и 2170...2200 МГц — для спутникового сегмента системы IMT-2000.

В соответствии со своими Резолюциями ВКР-2000 внесла ряд примечаний в Регламент радиосвязи и определила для системы IMT-2000 возможные дополнительные полосы частот: 806...960, 1710... 1885 и 2500...2690 МГц (или их части), которые могут назначаться национальными Администрациями связи как полосы развития системы IMT-2000.

За последнее десятилетие МСЭ, учитывая решения ВКР-2000 и ВКР-2003, разработал значительное число рекомендаций по использованию полос частот для системы IMT-2000. При этом наиболее важной является Рекомендация МЛ 036-3 («Частотные планы для внедрения наземного сегмента системы IMT-2000 в полосах частот 806...960 МГц; 1710...2025 МГц; 2110...2200 МГц и 2500...2690 МГц»). Эта

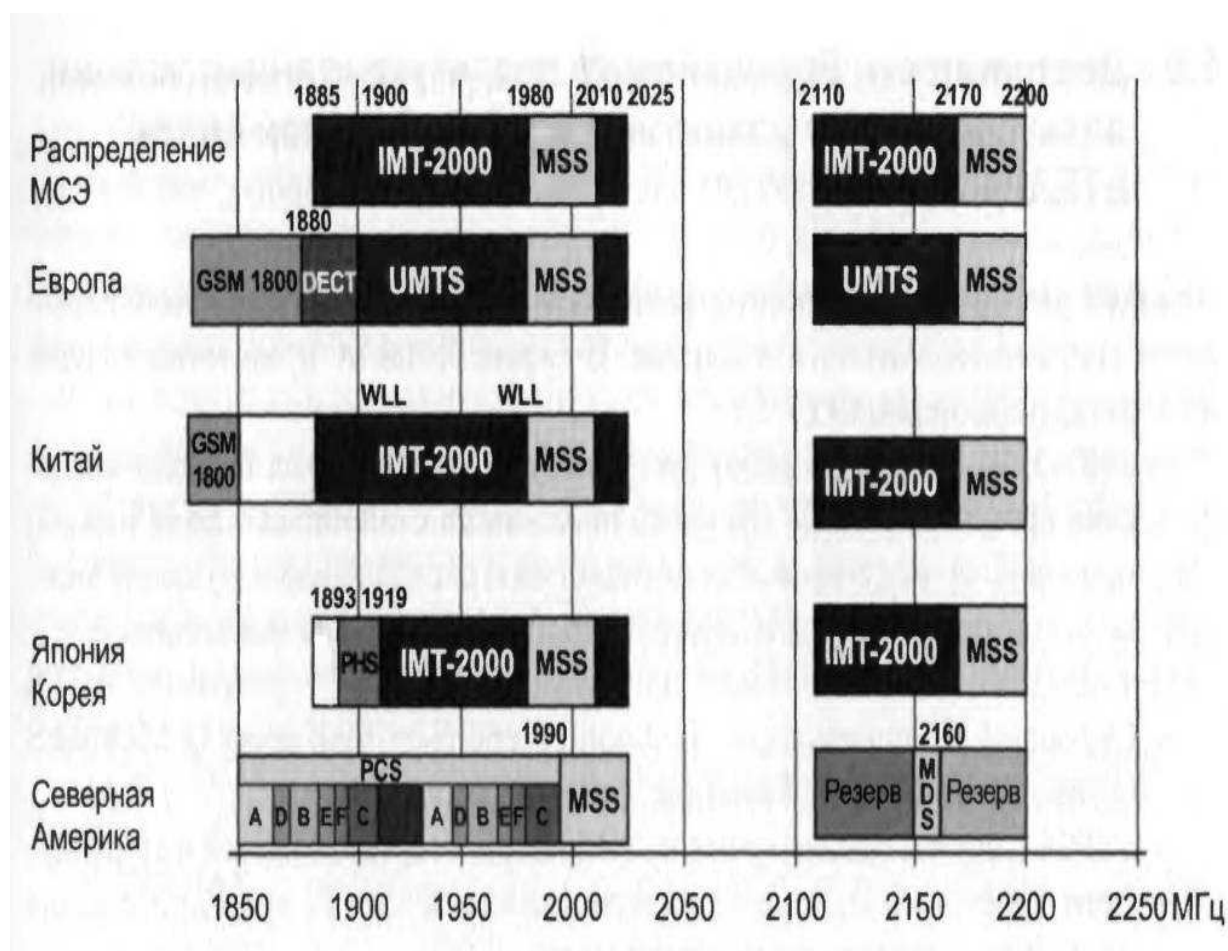


Рис. 7.18. Распределение полос частот для сетей 3G:

MDS (Mobile Communication Service) — подвижная служба передачи данных;

MSS (Mobile Satellite Service) — подвижная спутниковая служба;

PCS (Personal Communication System) — система персональной связи;

PHS (Personal Handyphone System) — система персональной портативной связи;

рекомендация предусматривает шесть вариантов использования полосы частот 1710...2200 МГц для систем IMT-2000 и позволяет Администрациям связи максимально гибко подходить к разработке частотных планов для сетей IMT-2000 национальных операторов с учетом реальной загрузки планируемых к использованию диапазонов частот. Кроме того, в МСЭ был подготовлен Отчет ITU-RM.2039, содержащий технические характеристики наземного сегмента системы IMT-2000, как для оборудования, использующего радиоинтерфейс IMT-DS, так и для оборудования, использующего интерфейс IMT-MS.

При разработке системы 3G в Европе в основном использовались технические решения, полученные для сетей GSM, по следующим двум причинам: во-первых, технология GSM была доминирующей на европейском рынке мобильной связи, во-вторых, в развитие сетей GSM были вложены гигантские инвестиции, требовавшие скорейшей окупаемости. Кроме того, будущая система 3G должна была удовлетворять следующим требованиям:

- иметь полное описание и технические требования в виде открытых стандартов, так же как и сети GSM. Технические требования к системе 3G разработаны и приняты как документы международных организаций стандартизации;
- превосходить сети GSM по всем аспектам. В начале разработки и эксплуатации система 3G должна быть совместима как минимум с сетями GSM и ISDN (Integrated Services Digital Network);
- поддерживать мультимедийные и другие услуги во всех подсистемах сети;
- обеспечивать высокую пропускную способность сети радиодоступа и получить распространение во всем мире. Требования к пропускной способности сети 3G должны превышать соответствующие требования к узкополосным мобильным сетям GSM и широкополосным стационарным мультимедийным сетям;
- услуги, предоставляемые конечным пользователям сетей 3G, не должны зависеть от особенностей технологии радиодоступа, а выбранная архитектура — ограничивать внедрение новых услуг связи. Технологическая платформа и услуги должны быть взаимонезависимы, иметь открытую структуру.

В рамках европейской исследовательской программы ACTS («Перспективные технологии и услуги связи») по проекту FRAMES («Перспективные системы широкополосного множественного радиодоступа») была проведена всесторонняя оценка различных технологий множественного доступа с целью выбора технологии радиоинтерфейса для европейского стандарта системы 3G — системы UMTS (Universal Mobile Telecommunication System). Основной задачей проекта FRAMES была разработка концепции радиоинтерфейса и предложений для стандартизации системы UMTS. В процессе работы над радиоинтерфейсом для системы UMTS европейские производители оборудования связи предложили рабочей группе SMG2 Технического комитета по мобильной связи ETSI пять концепций:

- 1) W-CDMA — широкополосный множественный доступ с кодовым разделением каналов;
- 2) OFDMA — ортогональный множественный доступ с частотным разделением каналов;
- 3) W-TDMA — широкополосный множественный доступ с временным разделением каналов;
- 4) TDMA/CDMA — широкополосный множественный доступ с кодово-временным разделением каналов;
- 5) ODMA — множественный доступ с гибкими возможностями.

Отобранные варианты были представлены для окончательного рассмотрения группе SMG2 (концепции W-TDMA и TDMA/CDMA), а также подгруппе W-CDMA японской ассоциации ARIB, занимающейся стандартизацией системы 3G (концепция W-CDMA).

В течение 1998 г. Европейский институт стандартизации систем электросвязи и японские региональные органы стандартизации (TTC и ARIB) пришли к согласию о выпуске единого стандарта на систему UMTS. Кроме того, свои усилия по разработке единого европейского стандарт 3G объединили шесть региональных организаций стандартизации в области электросвязи, которые вошли в Партнерский проект 3GPP.

7.6. характеристика 3G – 4G поколений ССМС

Таблица 7.4 дает представление о характеристиках 1G – 4G поколений ССМС.
Таблица 7.4

Поколение	1G	2G	2.5G	3G	3.5G	4G
Начало разработок	1970	1980	1985	1990	<2000	2000
Реализация	1984	1991	1999	2002	2006-2007	2010
Сервисы	аналоговый стандарт, синхронная передача данных со скоростью до 9,6 кбит/с	цифровой стандарт, поддержка коротких сообщений (SMS)	большая ёмкость, пакетная передача данных	ещё большая ёмкость, скорости до 2 Мбит/с	увеличение скорости сетей третьего поколения	большая ёмкость, ориентированная на мультимедийные сервисы, скорости до 100 Мбит/с
Стандарты	AMPS, TACS, NMT и др.	TDMA, CDMA, GSM, PDC	GPRS, EDGE, 1xRTT	WCDMA, CDMA2000, UMTS	HSDPA	есть стандарты
Скорость передачи	1,9 кбит/с	14,4 кбит/с	384 кбит/с	2 Мбит/с	3-14 Мбит/с	1 Гбит/с
Сеть	PSTN*	PSTN*	PSTN*, сеть пакетной передачи данных	сеть пакетной передачи данных	сеть пакетной передачи данных	Инт.

*PSTN – Public Switched Telephone Network – ТФОП

В сотовой связи смена поколений выражена намного более ярко, чем, скажем, в индустрии персональных компьютеров или другой подобной техники. В мобильном мире все четко: 1G (первое поколение) — это аналоговая связь (стандарт NMT). 2G — поколение цифровой связи с коммутацией каналов (стандарты GSM и CDMA). Третье поколение — 3G (стандарт UMTS) — предусматривает наряду с коммутацией каналов и пакетную передачу данных. Появились и промежуточные стандарты, призванные повысить скорость передачи сетей GSM; так, 2,5G определяет три основных дополнения к базовому стандарту:

1. **HSCSD** (*high-speed circuit-switched data*) – предоставление одному и тому же пользователю нескольких слотов в GSM кадре;
2. **GPRS** (*general packet radio service*) – предусматривает организацию специального канала пакетной передачи со скоростями до 115 кбит/с (или даже выше, но без помехоустойчивого кодирования);
3. **EDGE** (*enhanced data rates for global evolution*) подразумевает введение наряду с гауссовской МЧМ добавочного модуляционного формата: восьмеричной ФМ, утраивающей скорость передачи в рамках фиксированного частотного ресурса.

Совокупность всех мероприятий позволяет повысить скорость передачи до 384 кбит/с.

Субъектами разработки стандартов третьего поколения (3G) явились Международный Телекоммуникационный Союз **ITU** (*International Telecommunication Union*), Европейский институт телекоммуникационных стандартов **ETSI** (*European*

Telecommunication Standards Institute), ассоциации радиопромышленников США (**TIA** – *Telecommunication Industry Association*) и Японии (**ARIB** – *Association of Radio Industries and Businesses*), органы, представляющие интересы телекоммуникационных сообществ Китая и Кореи, а также крупнейшие корпорации, определяющие конъюнктуру мирового рынка (Nokia, Eriksson, Qualcomm, Motorola, Lucent, Nortel и др.).

Стандарт 3G, разработанный ITU, носит название IMT 2000 (*International Mobile Telecommunications 2000*). Под этой аббревиатурой объединены пять стандартов, и только некоторые из них обеспечивают полное покрытие в различных диапазонах, поэтому фактически только они могут рассматриваться в качестве полноценных 3G решений.

Используются три основных стандарта 3G: UMTS (*Universal Mobile Telecommunications Service*), CDMA2000 и WCDMA (*Wide CDMA*). Все они настроены на пакетную передачу данных и, соответственно, на работу с цифровыми компьютерными сетями, включая Интернет.

Для реализации систем третьего поколения разработаны рекомендации по глобальным унифицированным стандартам мобильной связи: обеспечение качества передачи речи, сравнимого с качеством передачи в проводных сетях связи; обеспечение безопасности, сравнимой с безопасностью в проводных сетях; обеспечение национального и международного роуминга; поддержка нескольких местных и международных операторов; эффективное использование спектра частот; пакетная и канальная коммутация; поддержка многоуровневых сотовых структур; взаимодействие с системами спутниковой связи; поэтапное наращивание скорости передачи данных вплоть до 2 Мбит/с.

Согласованные базовые требования, определяющие философию 3G, в укрупненной формулировке таковы:

- Скорость передачи данных до 2 Мбит/с в пределах полосы не более 5 МГц;
- Возможность варьирования скорости в широком диапазоне в зависимости от характера передаваемых данных;
- Возможность мультиплексной передачи в рамках одного контакта данных разнообразного содержания и разного уровня требований к качеству (речь, мультимедиа, пакеты и т.п.);
- Сосуществование систем 2G и 3G и возможность эстафетной передачи между ними с целью расширения зон покрытия и выравнивания трафика;
- Поддержка асимметричного режима работы, когда линия "вниз" имеет значительно более напряженный трафик (например, при поиске информации в Интернете) по сравнению с линией "вверх";
- Возможность реализации дуплексной передачи в вариантах как частотного, так и временного разнеса линий "вниз" и "вверх";
- Максимальная гибкость сетевого оборудования и возможность построения его на основе "набора инструментов" (*toolbox*);
- Доступность мобилизации наиболее современных резервов улучшения качества связи (адаптация к условиям распространения, пространственная селекция терминала – *smart* антенны, многопользовательские – *multiuser* – приемники и т.д.).

К середине 2002 г. лидирующие позиции закрепились за двумя проектами, общей платформой которых является технология кодового разделения (**CDMA**). Первый из них – **UMTS** (*Universal Mobile Telecommunication System*) является продуктом кооперации Евросоюза (в лице ETSI) и Японии (ARIB) и имеет действенную поддержку со стороны Кореи, Китая, а также многих корпораций Европы и США.

Хотя радиоинтерфейс UMTS, получивший название UTRAN (*Universal Terrestrial Radio Access Network*), основан на кодовом разделении и практически не имеет ничего общего с физическим уровнем GSM, на сетевом уровне предлагаемый стандарт целиком опирается на "готовую" инфраструктуру GSM.

Второй проект, известный под именем **cdma2000**, является максимально приемственным по отношению к системе **cdmaOne (IS-95)** и настойчиво продвигается компаниями Qualcomm, Lucent и Motorola – основными поставщиками и операторами действующих CDMA систем.

CDMA2000 является стандартом 3G в эволюционном развитии сетей cdmaOne. При сохранении основных принципов, заложенных версией IS-95A, технология стандарта CDMA непрерывно развивается.

При построении системы мобильной связи на основе технологии CDMA2000 1X первая фаза обеспечивает передачу данных со скоростью до 153 кбит/с, что позволяет предоставлять услуги голосовой связи, передачу коротких сообщений, работу с электронной почтой, Интернетом, базами данных, передачу данных и неподвижных изображений. Переход к следующей фазе CDMA2000 1xEV-DO происходит при использовании той же полосы частот 1,23 МГц, скорость передачи — до 2,4 Мбит/с в прямом канале и до 153 кбит/с в обратном, что делает эту систему связи отвечающей требованиям 3G и даёт возможность предоставлять самый широкий спектр услуг, вплоть до передачи видео в режиме реального времени. Следующей фазой развития стандарта в направлении увеличения сетевой ёмкости и передачи данных является 1xEV-DO Rev A: передача данных со скоростью до 3,1 Мбит/с по направлению к абоненту и до 1,8 Мбит/с — от абонента. Операторы смогут предоставлять те же услуги, что и на базе Rev. 0, а, кроме того, передавать голос, данные и осуществлять ширококовещание по IP сетям. Следующая фаза — 1xEV-DO Rev B, — разрабатывается с целью достигнуть более высоких скоростей на одном частотном канале: 4,9 Мбит/с к абоненту и 2,4 Мбит/с от абонента. К тому же будет обеспечиваться возможность объединения нескольких частотных каналов для увеличения скорости. Например, объединение 15-ти частотных каналов (максимально возможное количество) позволит достигать скоростей 73,5 Мбит/с к абоненту и 27 Мбит/с от абонента. Применение таких сетей — улучшенная работа чувствительных к временным задержкам приложений типа VoIP, Push to Talk, видеотелефония, сетевые игры и т.п.

Основными компонентами коммерческого успеха системы CDMA2000 являются более широкая зона обслуживания, высокое качество речи (практически эквивалентное проводным системам), гибкость и дешевизна внедрения новых услуг, высокая помехозащищённость, устойчивость канала связи от перехвата и прослушивания.

WCDMA (англ. Wideband Code Division Multiple Access — широкополосный CDMA) — технология радиointерфейса, избранная большинством операторов сотовой связи для обеспечения широкополосного радиодоступа с целью поддержки услуг 3G.

Технология оптимизирована для предоставления высокоскоростных мультимедийных услуг типа видео, доступа в Интернет и видеоконференций; обеспечивает скорости доступа вплоть до 2 Мбит/с на коротких расстояниях и 384 Кбит/с на больших с полной мобильностью. Такие величины скорости передачи данных требуют широкую полосу частот, поэтому ширина полосы WCDMA составляет 5 МГц.

Технология может быть добавлена к существующим сетям GSM и PDC, что делает стандарт WCDMA наиболее перспективным с точки зрения использования сетевых ресурсов и глобальной совместимости.

Использование выделенных ресурсов различными технологиями показано на рис. 7.19, который дает также наглядное представление об их отличиях.

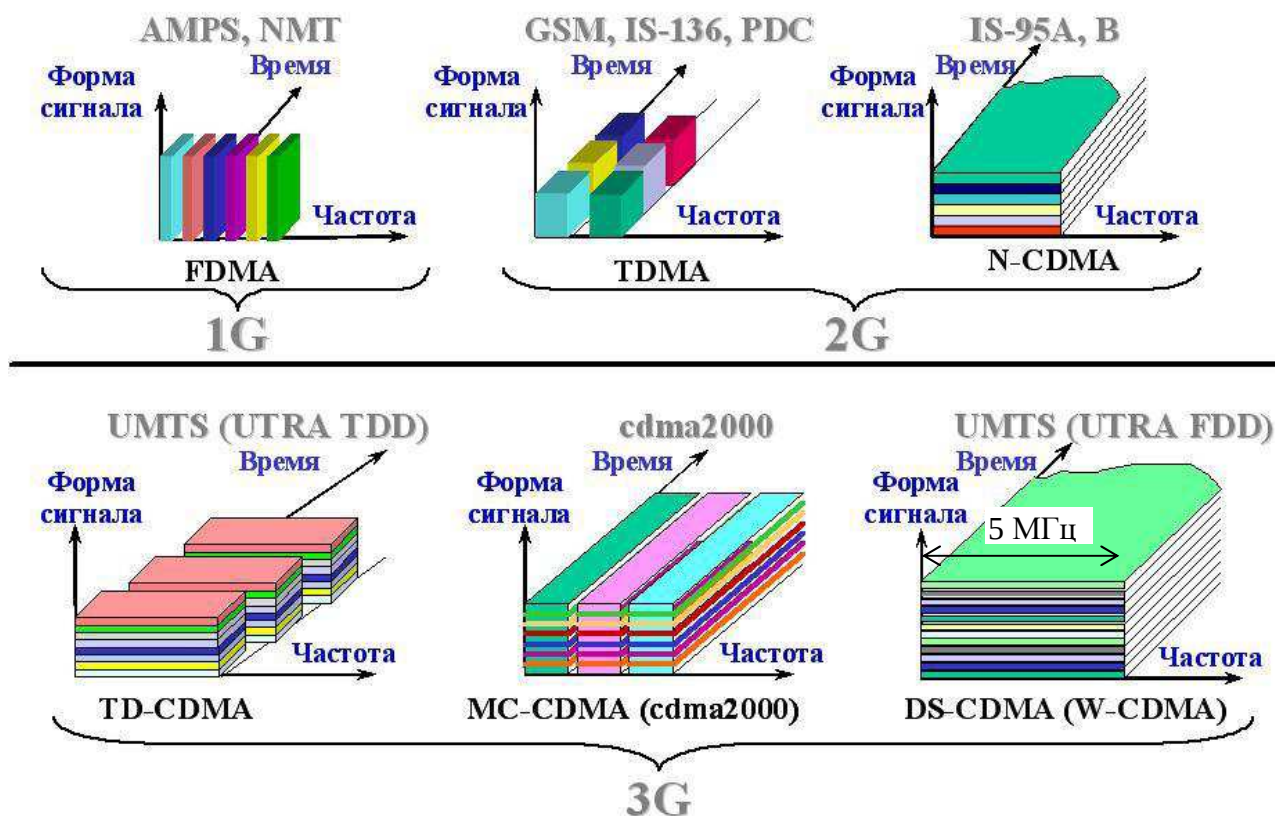


Рис. 7.19. Распределение ресурсов в мобильных сотовых сетях 1G – 3G.

Главным отличием 3G от сетей второго поколения является индивидуализация, то есть присвоение каждому абоненту IP-адреса, подобно Интернету. Еще один плюс – абоненты могут находиться в сети постоянно, не беспокоясь о материальных средствах, так как оплата насчитывается не за время, а за трафик.

Для завершения работы над спецификацией стандарта создано организационное ядро **3GPP** (*3G Partnership Project*), наименование которого часто используется как синоним названия самого проекта UMTS.

Ключевой технологией для UMTS является WCDMA.

Спецификация 3GPP определяет UMTS как широкополосную (*spread spectrum*) систему с кодовым разделением (CDMA) и прямым расширением спектра, жестко оговаривая структуры **расширяющих**, **канализирующих** и **скремблирующих** кодов, а также устанавливая единую и неизменную частоту следования чипов названных кодов в 3,84 Мчип/с ($3,84 \cdot 10^6$ чипов в секунду).

Под **скремблированием** понимается процесс наложения на коды символов открытого текста кодов случайной последовательности чисел, которую называют также гаммой (по названию буквы γ греческого алфавита, используемой в математических формулах для обозначения случайного процесса). Гаммирование относится к поточным методам шифрования, когда следующие друг за другом символы открытого текста последовательно превращаются в символы шифрограммы, что повышает скорость преобразования.

В качестве гаммы широко используются *псевдослучайные последовательности* (ПСП). Коды символов криптограммы Y_i есть поразрядная сумма по модулю 2 кодов X_i и Z_i . Расшифровка производится по правилу: $\hat{X} = Y + Z = X + Z + Z = X$, так что на приемной стороне должна генерироваться такая же ПСП и с той же фазой. Стойкость скремблирования определяется параметрами ПСП, главным из которых является период последовательности. На практике используются генераторы ПСП с периодом в несколько недель.

Начиная с 2006 года, в сетях UMTS повсеместно распространяется технология высокоскоростной пакетной передачи данных от базовой станции к мобильному терминалу HSDPA, которую принято относить к сетям поколения «3,5G». К началу 2008 года HSDPA поддерживала скорость передачи данных в режиме «от базовой станции к мобильному терминалу» до 7,2 Мбит/сек. Также ведутся разработки по повышению скорости передачи данных в режиме от мобильного терминала к базовой станции HSUPA. В долгосрочной перспективе, согласно проектам 3GPP, планируется эволюция UMTS в сети четвёртого поколения 4G, позволяющие базовым станциям передавать и принимать информацию на скоростях 100 Мбит/сек и 50 Мбит/сек соответственно, благодаря усовершенствованному использованию воздушной среды посредством мультиплексирования с ортогональным частотным разделением сигналов OFDM.

UMTS позволяет пользователям проводить сеансы видеоконференций посредством мобильного терминала, однако опыт работы операторов связи Японии и некоторых других стран показал невысокий интерес абонентов к данной услуге. Гораздо более перспективным представляется развитие сервисов, предлагающих загрузку музыкального и видео контента: высокий спрос на услуги такого рода был продемонстрирован в сетях 2,5G.

HSDPA (англ. High-Speed Downlink Packet Access — высокоскоростная пакетная передача данных от базовой станции к мобильному телефону) — стандарт мобильной связи, рассматривается специалистами как один из переходных этапов миграции к технологиям мобильной связи четвёртого поколения (4G). Максимальная теоретическая скорость передачи данных по стандарту составляет 14,4 Мбит/сек, практически же достижимая скорость в существующих сетях обычно не превышает 6 Мбит/сек.

Стандартом предусмотрено 20 категорий с различными значениями максимальной скорости передачи данных в пакетном режиме от базовой станции к мобильному терминалу. Каждая категория характеризуется максимальным числом одновременно используемых кодов (до 15), и типом модуляции в радиоканале (QPSK/16 QAM).

Категории 1-5 используют до 5 кодов и имеют максимальную скорость 1,2-3,6 Мб/с, категория 1 — до 14 Мб/с.

HSURA — аналогично HSDPA, технология высокоскоростной пакетной передачи данных по направлению «вверх» (англ. High-Speed Uplink Packet Access) представляет собой стандарт мобильной связи, позволяющий ускорить передачу данных от W-CDMA устройств конечного пользователя до базовой станции за счёт применения более совершенных методов модуляции.

Теоретически стандарт HSUPA рассчитан на максимальную скорость передачи данных по направлению «вверх» до 5,7 Мбит/с, позволяя, таким образом, использовать приложения третьего поколения, требующие обработки огромных потоков данных от мобильного устройства к базовой станции, например, видеоконференцсвязь.

7.6.1. Четвертое поколение MC

К связи четвертого поколения (4G), как правило, относят технологии, которые позволяют передавать данные в сотовых сетях со скоростью выше 100 Мбит/сек.

Технология LTE (Long-Term Evolution) — это основное направление эволюции сетей сотовой связи третьего поколения (3G). В январе 2008 года международное объединение Third Generation Partnership Project (3GPP), разрабатывающее перспективные стандарты мобильной связи, утвердило LTE в качестве следующего после UMTS стандарта широкополосной сети мобильной связи.

Сети 4G на основе стандарта LTE способны работать практически по всей ширине спектра частот от 700 МГц до 2,7 ГГц.

LTE обеспечивает теоретическую пиковую скорость передачи данных до 326,4 Мбит/с от базовой станции к пользователю и до 172,8 Мбит/с в обратном направлении.

Технология Long Term Evolution, как ожидается, приведет к появлению качественно новых мобильных сервисов: пользователи смогут в режиме реального времени получать видео высокого качества, работать с интерактивными услугами и пр.

В широком понимании к 4G относят также технологии беспроводной передачи интернет-данных Wi-Fi (скоростные варианты этого стандарта) и WiMAX (в теории скорость может превышать 1 Гбит/сек).

Wi-Fi (Wireless Fidelity) – современная беспроводная технология соединения компьютеров в сеть или подключения их к Интернету. Wi-Fi разработан консорциумом Wi-Fi Alliance. Стандарт Wi-Fi позволяет предоставить высокоскоростной доступ ко всем ресурсам сети Интернет (электронная почта, Интернет-серфинг, ICQ и т.д.) с ноутбука, смартфона или КПК в зоне покрытия сети Wi-Fi. Технология обеспечивает одновременную работу в сети нескольких десятков активных пользователей, скорость передачи информации для конечного абонента может достигать 54 Мбит/с.

Стандарт WiMAX (аббревиатура от Worldwide Interoperability for Microwave Access) – технология широкополосной беспроводной связи, которая в отличие от других технологий радиодоступа, обеспечивает высокоскоростные соединения на больших расстояниях даже при отсутствии прямой видимости объекта, на отраженном сигнале.

Основана на стандарте IEEE 802.16, который также называют Wireless MAN. Название «WiMAX» было создано WiMAX Forum — организацией, которая была основана в июне 2001 года с целью продвижения и развития технологии WiMAX. Форум описывает WiMAX как «основанную на стандарте технологию, предоставляющую высокоскоростной беспроводной доступ к сети, альтернативный выделенным линиям и DSL».

WiMAX позволяет осуществлять доступ в Интернет на высоких скоростях, с гораздо большим покрытием, чем у Wi-Fi сетей. Это позволяет использовать технологию в качестве «магистральных каналов», продолжением которых выступают традиционные DSL- и выделенные линии, а также локальные сети. В результате подобный подход позволяет создавать масштабируемые высокоскоростные сети в рамках целых городов.

Фиксированный и мобильный вариант WiMAX

Набор преимуществ присущ всему семейству WiMAX, однако его версии существенно отличаются друг от друга. Разработчики стандарта искали оптимальные решения как для фиксированного, так и для мобильного применения, но совместить все требования в рамках одного стандарта не удалось. Хотя ряд базовых требований совпадает, нацеленность технологий на разные рыночные ниши привела к созданию двух отдельных версий стандарта (вернее, их можно считать двумя разными стандартами). Каждая из спецификаций WiMAX определяет свои рабочие диапазоны частот, ширину полосы пропускания, мощность излучения, методы передачи и доступа, способы кодирования и модуляции сигнала, принципы повторного использования радиочастот и прочие показатели. А потому WiMAX-системы, основанные на версиях стандарта IEEE 802.16 e и d, практически несовместимы. Краткие характеристики каждой из версий приведены ниже.

802.16-2004 (известен также как 802.16d и фиксированный WiMAX). Спецификация утверждена в 2004 году. Используется ортогональное частотное мультиплексирование (OFDM), поддерживается фиксированный доступ в зонах с наличием либо отсутствием прямой видимости. Пользовательские устройства представляют собой стационарные модемы для установки вне и внутри помещений, а также PCMCIA-карты для ноутбуков. В большинстве стран под эту технологию отведены диапазоны 3,5 и 5 ГГц. По сведениям

WiMAX Forum, насчитывается уже порядка 175 внедрений фиксированной версии. Многие аналитики видят в ней конкурирующую или взаимодополняющую технологию проводного широкополосного доступа DSL.

802.16-2005 (известен также как 802.16e и мобильный WiMAX). Спецификация утверждена в 2005 году. Это — новый виток развития технологии фиксированного доступа (802.16d). Оптимизированная для поддержки мобильных пользователей версия поддерживает ряд специфических функций, таких как хэндовер(англ.), idle mode и роуминг. Применяется масштабируемый OFDM-доступ (SOFDMA), возможна работа при наличии либо отсутствии прямой видимости. Планируемые частотные диапазоны для сетей Mobile WiMAX таковы: 2,3-2,5; 2,5-2,7; 3,4-3,8 ГГц. В мире реализованы несколько пилотных проектов, в том числе первым в России свою сеть развернул «Скартел». Конкурентами 802.16e являются все мобильные технологии третьего поколения (например, EV-DO, HSDPA).

Основное различие двух технологий состоит в том, что фиксированный WiMAX позволяет обслуживать только «статичных» абонентов, а мобильный ориентирован на работу с пользователями, передвигающимися со скоростью до 120 км/ч. Мобильность означает наличие функций роуминга и «бесшовного» переключения между базовыми станциями при передвижении абонента (как происходит в сетях сотовой связи). В частном случае мобильный WiMAX может применяться и для обслуживания фиксированных пользователей.

Пользовательское оборудование

Оборудование для использования сетей WiMAX поставляется несколькими производителями и может быть установлено как в помещении (устройства размером с обычный DSL модем), так и вне него (устройства размером с флэш-накопитель). Следует заметить что оборудование, рассчитанное на размещение внутри помещений и не требующее профессиональных навыков при установке, конечно, более удобно, однако способно работать на значительно меньших расстояниях от базовой станции, чем профессионально установленные внешние устройства. Поэтому оборудование, установленное внутри помещений требует намного больших инвестиций в развитие инфраструктуры сети, так как подразумевает использование намного большего числа точек доступа.

С изобретением мобильного WiMAX все больший акцент делается на разработке мобильных устройств. В том числе специальных телефонных трубок (похожи на обычный мобильный смартфон), и компьютерной периферии (USB радио модулей и PC card).

Принцип работы Основные понятия

В общем виде WiMAX сети состоят из следующих основных частей: базовых и абонентских станций, а также оборудования, связывающего базовые станции между собой, с поставщиком сервисов и с Интернетом.

Для соединения базовой станции с абонентской используется высокочастотный диапазон радиоволн от 1,5 до 11 ГГц. В идеальных условиях скорость обмена данными может достигать 70 Мбит/с, при этом не требуется обеспечения прямой видимости между базовой станцией и приемником.

Как уже говорилось выше, WiMAX применяется как для решения проблемы «последней мили», так и для предоставления доступа в сеть офисным и районным сетям[1].

Между базовыми станциями устанавливаются соединения (прямой видимости), использующие диапазон частот от 10 до 66 ГГц, скорость обмена данными может

достигать 120 Мбит/с. При этом, по крайней мере одна базовая станция подключается к сети провайдера с использованием классических проводных соединений. Однако, чем большее число БС подключено к сетям провайдера, тем выше скорость передачи данных и надёжность сети в целом.

Структура сетей семейства стандартов IEEE 802.16 схожа с традиционными GSM сетями (базовые станции действуют на расстояниях до десятков километров, для их установки не обязательно строить вышки — допускается установка на крышах домов при соблюдении условия прямой видимости между станциями)[1].

7.6.2. Сравнение Wi-Fi и WiMAX

Сопоставления WiMAX и Wi-Fi далеко не редкость, возможно, потому, что эти термины созвучны, название стандартов, на которых основаны эти технологии, похожи (стандарты IEEE, оба начинаются с «802.»), а также обе технологии используют беспроводное соединение и используются для подключения к Интернету (каналу обмена данными). Но несмотря на это, эти технологии направлены на решение совершенно различных задач.

Таблица 7.5. Сравнительная таблица стандартов беспроводной связи

Технология	Стандарт	Использование	Пропускная способность	Радиус действия	Частоты
Wi-Fi	802.11a	WLAN	до 54 Мбит/с	до 100 метров	5,0 ГГц
Wi-Fi	802.11b	WLAN	до 11 Мбит/с	до 100 метров	2,4 ГГц
Wi-Fi	802.11g	WLAN	до 108 Мбит/с	до 100 метров	2,4 ГГц
Wi-Fi	802.11n	WLAN	до 300 Мбит/с (в перспективе до 450, а затем до 600 Мбит/с)	до 100 метров	2,4 — 2,5 или 5,0 ГГц
WiMax	802.16d	WMAN	до 75 Мбит/с	6-10 км	1,5-11 ГГц
WiMax	802.16e	Mobile WMAN	до 30 Мбит/с	1-5 км	2-6 ГГц
WiMax	802.16m	WMAN, Mobile WMAN	до 1 Гбит/с (WMAN), до 100 Мбит/с (Mobile WMAN)	н/д (стандарт в разработке)	н/д (стандарт в разработке)
Bluetooth v. 1.1.	802.15.1	WPAN	до 1 Мбит/с	до 10 метров	2,4 ГГц
Bluetooth v. 1.3.	802.15.3	WPAN	от 11 до 55 Мбит/с	до 100 метров	2,4 ГГц

WiMAX это система дальнего действия, покрывающая километры пространства, которая обычно использует лицензированные спектры частот (хотя возможно и использование нелицензированных частот) для предоставления соединения с Интернетом типа точка-точка провайдером конечному пользователю. Разные стандарты семейства 802.16 обеспечивают разные виды доступа, от мобильного (схож с передачей данных с

мобильных телефонов) до фиксированного (альтернатива проводному доступу, при котором беспроводное оборудование пользователя привязано к местоположению).

Wi-Fi это система более короткого действия, обычно покрывающая десятки метров, которая использует нелицензированные диапазоны частот для обеспечения доступа к сети. Обычно Wi-Fi используется пользователями для доступа к их собственной локальной сети, которая может быть и не подключена к Интернету. Если WiMAX можно сравнить с мобильной связью, то Wi-Fi скорее похож на стационарный беспроводной телефон.

Bluetooth – технология беспроводной ближней коротковолновой радиосвязи (до 30 м), позволяющая объединять устройства разных типов для передачи речи и данных. Её разработкой и развитием занимается ассоциация Bluetooth SIG . Стандарт получил обозначение IEEE 802.15 . Он определяет работу на частоте 2,4 ГГц, со скоростями передачи 722-784 Кбит/с и расстояниями до 10 м. Свое название технология получила в честь датского короля Гаральда Синий Зуб, правившего Данией и Норвегией в X веке

WiMAX и Wi-Fi имеют совершенно разный механизм Quality of Service (QoS). WiMAX использует механизм, основанный на установлении соединения между базовой станцией и устройством пользователя. Каждое соединение основано на специальном алгоритме планирования, который может гарантировать параметр QoS для каждого соединения. Wi-Fi, в свою очередь, использует механизм QoS подобный тому, что используется в Ethernet, при котором пакеты получают различный приоритет. Такой подход не гарантирует одинаковый QoS для каждого соединения.

Из-за дешевизны и простоты установки, Wi-Fi часто используется для предоставления клиентам быстрого доступа в Интернет различными организациями. Например, в некоторых кафе, отелях, вокзалах и аэропортах можно обнаружить бесплатную точку доступа Wi-Fi.

Стандарт WiMAX был разработан корпорацией Intel, крупнейшим мировым производителем микрочипов. Соответственно, WiMAX-чипами будут в первую очередь комплектовать ноутбуки. Скорей всего, со временем WiMAX вытеснит Wi-Fi, так как Wi-Fi действует в радиусе нескольких метров от точки доступа, у мобильного WiMAX покрытие существенно больше. И кроме того, он позволяет абоненту, если тот перемещается со скоростью до 120 км/ч, переключаться между станциями.

Список литературы

1. Скляр Б. Цифровая связь. Теоретические основы и практическое применение. Изд. 2-е, испр.: Пер. с англ. – М.: Издательский дом «Вильямс», 2003. – 1104 с.
2. Ипатов В. Широкополосные системы и кодовое разделение сигналов. Принципы и приложения. Москва: Техносфера, 2007. – 488 с.
3. Прокис Дж. Цифровая связь. Пер. с англ. / Под ред. Д.Д. Кловского – М.: Радио и связь, 2000. – 800 с.
4. Пенин П.И. Системы передачи цифровой информации: Учебное пособие для вузов.- М.: Сов. радио, 1983. – 368с.
5. Радиосистемы передачи информации: Учебное пособие для вузов / В.А. Васин, В.В. Калмыков, Ю.Н. Себекин, А.И. Сенин, И.Б. Федоров; под ред. И.Б. Федорова и В.В. Калмыкова. – М.: Горячая линия – Телеком, 2005. – 472 с.: ил.
6. Основы теории кодирования и шифрования в современных РТС: Конспект лекций / Самойлов И.М. – Санкт-Петербург, СПбГЭТУ «ЛЭТИ», 2007. – 179 с.
7. Турин Дж.Л. Введение в широкополосные методы борьбы с многолучевостью распространения радиосигналов и их применение в городских системах цифровой связи. ТИИЭР, т. 68, № 2, март 1980. С. 30 – 60.
8. Файфер У.С., Бруно Ф.Дж. Недорогая приемопередающая установка пакетной радиосвязи. ТИИЭР, т. 75, № 1, январь 1987. С. 41 – 53.
9. Стиффлер Дж. Дж. Теория синхронной связи: Пер. с англ. /Под ред. Э.М. Габидуллина.-М.: Связь. 1975. – 488 с.
10. Системы мобильной связи: Учебное пособие для вузов/В. П. Ипатов, В. И. Орлов, И. М. Самойлов, В. Н. Смирнов; под. ред. В. П. Ипатова. -М.: Горячая линия – Телеком, 2003, -272с.
11. Невдяев Л.М. Мобильная связь 3-го поколения. М.: «Связь и бизнес», МЦНТИ, 2000 – 208 с.
12. Градштейн И.С., Рыжик И.М.. Таблицы интегралов, сумм, рядов и произведений. Москва: - Наука, 1971. – 1109 с.