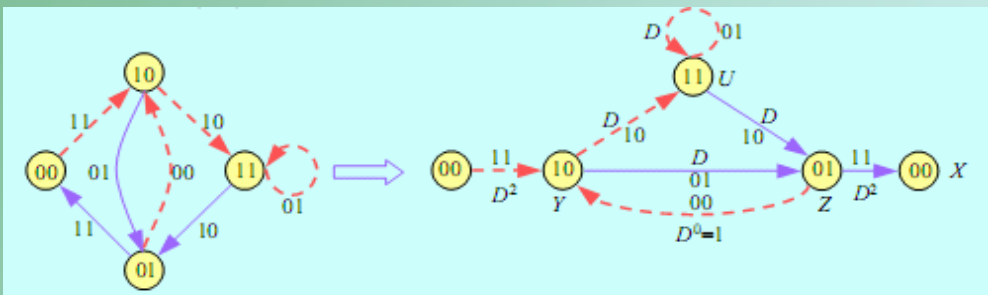
Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение высшего образования

Новгородский государственный университет
имени Ярослава Мудрого

Быстров Н.Е.

РАДИОТЕХНИЧЕСКИЕ СИСТЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ



Великий Новгород
2017 г.

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

Новгородский государственный университет
имени Ярослава Мудрого

Быстров Н.Е.

РАДИОТЕХНИЧЕСКИЕ СИСТЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

УЧЕБНОЕ ПОСОБИЕ

Великий Новгород
2017 г.

ББК

УДК 621.396; 681.6.07

Быстров Н.Е.. Радиотехнические системы передачи информации: Учебное пособие /
ФГБОУ «Новгородский государственный университет им. Ярослава Мудрого», Великий Новгород, 2017 г. - 195 с.

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 «Системы и устройства передачи, приема и обработки сигналов».

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

ПРЕДИСЛОВИЕ

Главным побудительным мотивом написания 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 автор 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. Использованы 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 56. 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.

Появление работ Найквиста (H.Nyquist) и Хартли (R.Hartley), посвященных исследованию канала связи и скорости передачи информации (1924, 1928 гг.) положило начало современной теории цифровой связи. Шеннон (C.Shannon) в 1947-48 гг. сформулировал основную проблему надежной передачи информации в статистических терминах, используя вероятностные модели источников и каналов связи.

Шеннон также установил существование для любого канала основополагающей константы – пропускной способности, т.е. предельной скорости безошибочной передачи, зависящей от физических ресурсов, доступных передатчику, и интенсивности помех в канале. В работах Хэмминга (R.Hamming) в 1949-50 гг. впервые были разработаны методы кодирования сообщений (т.е. отображения их в последовательности некоторых символов), позволяющие обнаруживать и исправлять ошибки, вызванные канальными помехами.

Определим основные понятия и определения, наиболее часто используемые в современных системах цифровой связи.

Отличительной особенностью систем цифровой связи – СЦС (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 ПРЕОБРАЗОВАНИЕ СИГНАЛОВ В СИСТЕМЕ ПЕРЕДАЧИ ИНФОРМАЦИИ

1.1 Структура системы цифровой передачи информации

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

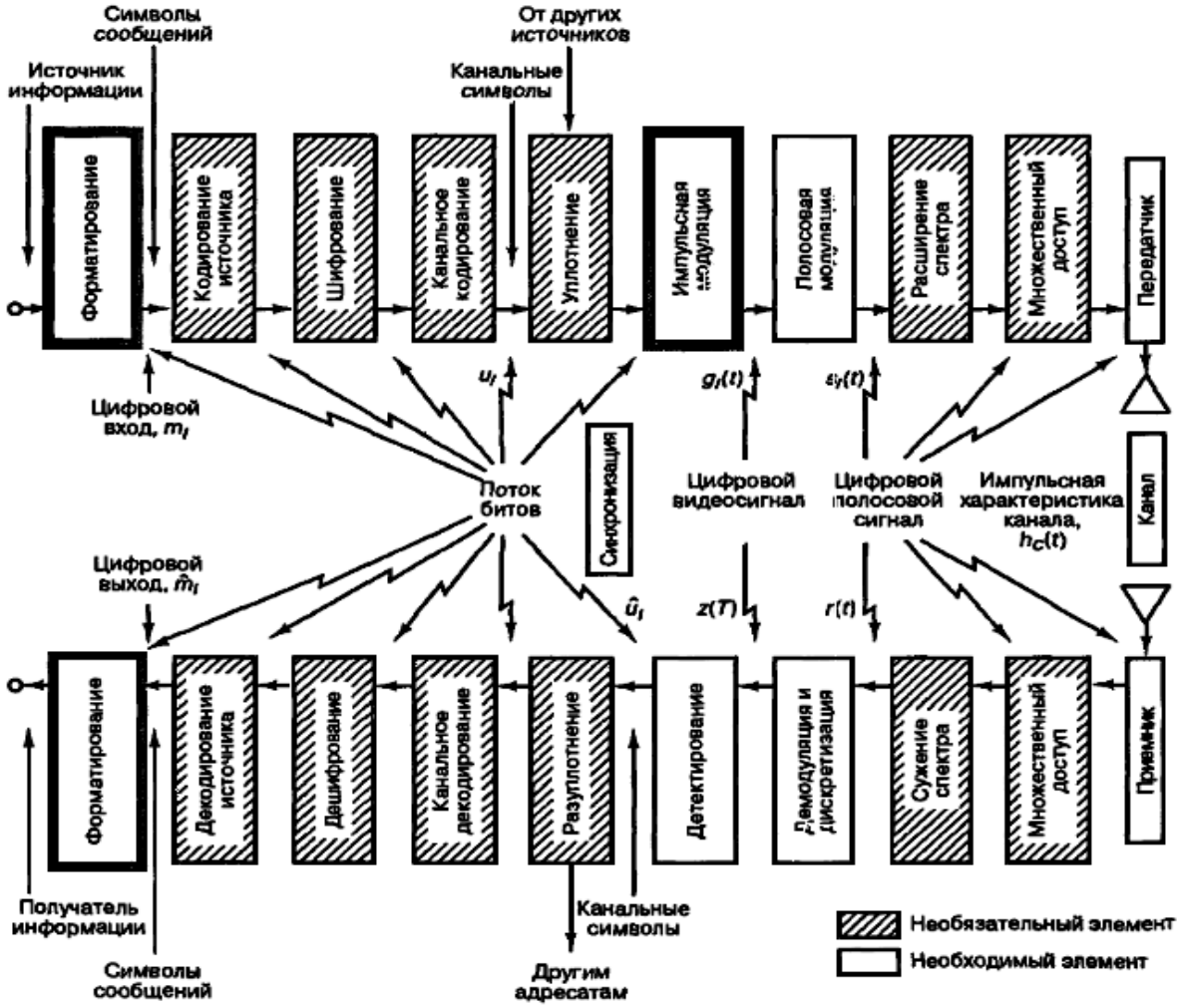


Рис. 1.1. Функциональная схема системы цифровой передачи информации.

В верхних блоках осуществляется преобразование сигнала на пути от источника к передатчику 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.

Рассмотрим ключевые блоки обработки сигналов.

Форматирование преобразовывает исходную информацию в последовательность битового потока символов, образуя цифровое сообщение.

Кодирование источника осуществляется, как правило, с целью уменьшения избыточности его сообщений; в этом случае оно называется ***статистическим*** или ***эффективным***.

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

Канальное 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.

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). Специфические вопросы многоканальной передачи здесь рассматриваться не будут; дальнейшее описание предполагает наличие одного источника (канала).

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).

Модуляция - это процесс, посредством которого символы сообщений или канальные символы (если используется канальное 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.

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 недвоичным символам результирующий сигнал именуется 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 канальный символ принимает форму видеосигнала $g_i(t)$, где i = 1, ..., *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 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, ..., *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 $s_i(t)$: $y(t) = s_i(t) * h_c(t) + n(t), \quad i = 1,...,M$, 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. 1.2, 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. 1.2; 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. 1.2 показан безотносительно к чему-либо, хотя фактически он участвует в регулировании операций практически в каждом блоке, приведенном на рисунке.

1.2 Модели источников сообщений

Источники сообщений можно рассматривать как некие генераторы случайных процессов. Математическая модель источника при этом сводится к описанию вероятностных свойств этих процессов. По типу генерируемых сообщений источники подразделяются на *дискретные* и *непрерывные*.

Дискретным называется источник, множество или алфавит X возможных сообщений которого конечно или счетно: $X = x_1, x_2, \ldots, x_L$ Каждое состояние x_i характеризуется вероятностью $p_i = p(x_i)$, причем $\sum_{i=1}^{L} p_i = 1$. Для полного описания источника необходимо знание вероятностей не только отдельных сообщений в какой-то фиксированный момент времени, но и любых последовательностей m сообщений в произвольные моменты времени. Таким образом, дискретный источник, выбирающий сообщения из множества X, задан исчерпывающим образом, если для любых значений m, i и последовательности сообщений $\mathbf{X} = x^{(i+1)}, x^{(i+2)}, \ldots, x^{(i+m)}$ определена вероятность последовательности сообщений $p(\mathbf{X}) = p\ x^{(i+1)}, \ldots, x^{(i+m)}$. Отметим, что использованный выше надстрочный индекс (j) есть номер элемента в последовательности, т. е. аргумент дискретного времени.

В общем случае вероятность последовательности сообщений может меняться со временем i. Однако для важного класса *стационарных* источников поведение во времени инвариантно, что можно формально выразить как

$$p\ x^{(i+1)}, x^{(i+2)}, \ldots, x^{(i+m)} = p\ x^{(1)}, x^{(2)}, \ldots, x^{(m)} . \qquad (1.1)$$

Приведенное равенство (В.1) означает, что вероятность любой последовательности сообщений источника не зависит от того, когда эта последовательность появляется.

Важнейший класс источников образуют *дискретные источники без памяти* (ДИБП). Сообщения, последовательно выдаваемые ДИБП, являются независимыми и таким образом, для ДИБП справедливо равенство

$$p\ x^{(1)},\ x^{(2)},\ldots,\ x^{(m)}\ = \prod_{j=1}^{m} p_j\ x^{(j)}\ . \qquad (1.2)$$

В случае стационарных источников вероятность любого элементарного сообщения неизменна со временем: $p_j(x) = p(x)$, так что

$$p\ x^{(1)},\ x^{(2)},\ldots,\ x^{(m)}\ = \prod_{j=1}^{m} p\ x^{(j)}\ . \qquad (1.3)$$

Непрерывные (по состоянию) *источники сообщений* имеют несчетный (континуальный) ансамбль состояний. В случае, когда элементы такого ансамбля X можно отождествить с точками числовой оси, для его описания пригодна плотность вероятности $W(x),\ x \in X$, причем $\int_{-\infty}^{\infty} W(x)\,dx = 1.$

1.3 Модели канала связи

Канал связи можно определить как совокупность устройств, обеспечивающих передачу сигналов между любыми двумя точками, лежащими по разные стороны среды распространения. В зависимости от положения точек, а, следовательно, и вида входных и выходных сигналов, канал связи может оказаться *непрерывным, дискретным или дискретно-непрерывным*.

Канал называется *дискретным по времени*, если входные и выходные сигналы доступны для наблюдения только в дискретные моменты времени. Если же входные и выходные сигналы наблюдаются непрерывно, канал называется *непрерывным по времени*. Канал называется *дискретным по состоянию*, если входные и выходные сигналы принадлежат дискретным множествам. Если же сигналы принимают значения из континуального множества, канал называется *непрерывным по состоянию*. Канал можно назвать *дискретно-непрерывным*, если один из сигналов (входной или выходной) дискретный, а другой – непрерывный.

Рассмотрим *модели непрерывных каналов* (по времени и состоянию). Пусть $c(t)$ – входной сигнал, а $y(t)$ – выходной. Первое предположение состоит в том, что $y(t) = c_p(t) + n(t)$, где $c_p(t)$ – принятый сигнал, соответствующий переданному $c(t)$, $n(t)$ – помеха, действующая в канале. Тогда имеем множество так называемых аддитивных каналов.

Если считается, что помехой является только нормальный белый шум, то получаем модели каналов с аддитивным белым гауссовским шумом (АБГШ).

Таким образом, при сделанных предположениях непрерывные каналы классифицируются по виду искажений входного сигнала $c(t)$. К самой простой модели непрерывного канала, так называемому гауссовскому каналу, приводит предположение об отсутствии искажений входного сигнала. Тогда выходной сигнал $y(t) = c\ (t) + n(t)$.

Рассмотрим *модель дискретных каналов.* Пусть на вход канала поступает последовательность сообщений $\mathbf{X} = x^{(1)}, x^{(2)}, \ldots, x^{(n)}$. Тогда множество X^n всех последовательностей $\mathbf{X}$ имеет мощность L^n. На выходе канала наблюдается последовательность $\mathbf{Y} = y^{(1)}, y^{(2)}, \ldots, y^{(n)}$, $y^{(i)} \in Y$, причем мощность выходного алфавита $|Y|$ может и не совпадать с мощностью входного $|X|$.

Для полного задания канала недостаточно описания входных и выходных сигналов – необходимо и статистическое описание процесса передачи сообщений. Наличие шума в канале может привести к трансформации одного и того же входного сигнала в различные выходные сигналы.

Дискретный канал оказывается полностью заданным, если для любой пары последовательностей $\mathbf{X}$ и $\mathbf{Y}$ с элементами из дискретных множеств X и Y известны *переходные вероятности* $p\ \mathbf{Y}|\mathbf{X}$.

В зависимости от свойств переходных вероятностей модели каналов допускают дальнейшую систематизацию.

Дискретный канал называется *каналом без памяти*, если

$$p\ \mathbf{Y}|\mathbf{X} = \prod_{i=1}^{n} p\ y^{(i)} \Big| x^{(i)}, \qquad (1.4)$$

т.е. вероятности тех или иных значений текущего выходного символа определяются лишь текущим входным и не зависят от прошлых входных символов.

Дискретный канал называется *стационарным*, если переходные вероятности $p\ \mathbf{Y}|\mathbf{X}$ не зависят от момента начала передачи и сохраняют неизменными свои значения на протяжении всего времени передачи.

Рассмотрим *модель двоичного канала связи*. Множества входных и выходных сигналов двоичного канала состоят из двух элементов: $X = Y = \{0,1\}$. Введем обозначения переходных вероятностей $p\,(y=1|x=0) = p$ и $p\,(y=0|x=1) = q$, которые определяют вероятности искажения в канале 0 и 1. Наглядным является описание рассматриваемого канала графом на рис. 1.3.

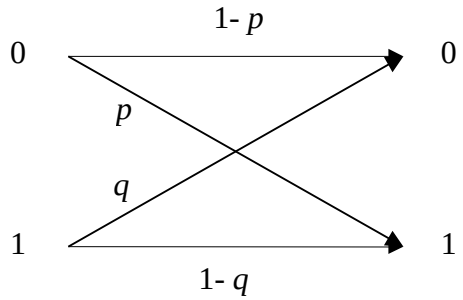


Рис. 1.3. Граф двоичного канала связи

Двоичный канал без памяти с равными вероятностями ошибочных переходов $(p = q)$ называется *двоичным симметричным каналом* (ДСК).

Пусть входной сигнал $X = x_1, x_2, x_3, ..., x_n$. Выходной сигнал ДСК можно представить в виде

$$Y = X \oplus E = y_1, y_2, y_3, \cdots y_n, \tag{1.5}$$

где $E = e_1, e_2, e_3, \cdots, e_n$ – двоичный вектор искажений или вектором ошибок, знак $\oplus$ означает поразрядное сложение символов $y_j = x_j \oplus e_j$ по модулю 2. Единичные компоненты вектора ошибок $e_j = 1$ указывают номера искаженных символов $(y_j \neq x_j)$. Число единиц в векторе $t = \sum_1^n e_j$ называется кратностью ошибки или весом. Так как компоненты вектора ошибки в ДСК не зависят ни от входных символов, ни от друг друга, ни от номера позиции, то вероятность появления ошибки кратности t определяется выражением

$$P_n(E) = C_n^t p^t (1-p)^{n-t}, \tag{1.6}$$

где C_n^t – число сочетаний из n по t. Так, вероятность однократной ошибки равна $np\ (1-p)^{n-1}$, а вероятность появления конкретного вектора ошибки, например $E = 0100100$ равна $p^2(1-p)^5$.

При малых значениях p распределение $P_n(E)$ резко убывает с увеличением кратности ошибки. Наиболее вероятно (с вероятностью $(1-p)^n$), что при передаче по ДСК искажений вообще не будет. Со значительно меньшей вероятностью произойдет искажение одного бита, с ещё меньшей вероятностью – двух бит и т. д.

1.4. Характеристики системы передачи информации

Рассмотрим основные понятия и определения.

Сообщения, знаки и символы. Источник *сообщений* может быть *аналоговым* или *дискретным.*

При цифровой передаче данных *сообщение* представляет собой последовательность текстовых *знаков*, принадлежащих конечному алфавиту.

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 др.

При передаче текстовые знаки вначале кодируются в последовательность битов, которая называется *битовым потоком.* После этого формируются группы из k бит, именуемые *символами* (*цифровое сообщение)* m_i, $i = 1...M$. Символ - это группа из k бит, рассматриваемых как единое целое, причем число всех символов конечно и равно $M = 2^k$, где k -число битов в символе. Система, использующая символьный набор размера M, называется *M-арной.* При k=1 система является *бинарной,* размер набора символов M=2, используется один из двух различных сигналов для представления двоичного значения «1», а другой – для представления двоичного значения «0». В этом частном случае символ и бит – это одно и тоже. При k=2 система именуется *четверичной* или 4-х *уровневой* (M=4).

Если информация является аналоговой, её знаковое кодирование (как в случае текстовой информации) невозможно. Вначале информацию следует перевести в цифровой формат. Для этого над сигналом осуществляется в аналого-цифровое преобразование. Исходный непрерывный сигнал дискретизируется во времени и квантуется в один из L уровней. После этого каждая квантованная выборка проходит цифровое кодирование для превращения в l-битовое ($l = \log_2 L$) кодовое слово.

При *низкочастотной* передаче каждый из сигналов m_i, $i = 1...M$ будет представлен одним из набора видеосигналов g_i, $i = 1...M$. Характеристики видеосигнала - амплитуда, длительность и положение. Для типичной *полосовой* передачи каждый видеосигнал g_i, $i = 1...M$ будет представляться одним из набора полосовых радиосигналов s_i, $i = 1...M$. Характеристики радиосигнала (амплитуда, частота и фаза) позволяют его идентифицировать как один из символов конечного алфавита.

Пример. На рис. П.1 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.

Таблица 1. Кодировка Windows 1251

128 Ъ	144 ђ	160	176 °	192 А	208 Р	224 а	240 р
129 Ѓ	145 ‘	161 Ў	177 ±	193 Б	209 С	225 б	241 с
130 ‚	146 ’	162 ў	178 І	194 В	210 Т	226 в	242 т
131 ѓ	147 “	163 Ј	179 і	195 Г	211 У	227 г	243 у
132 „	148 ”	164 ¤	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 ¦	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 Є	186 є	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 Ќ	157 ќ	173 -	189 Ѕ	205 Н	221 Э	237 н	253 э
142 Ћ	158 ћ	174 ®	190 ѕ	206 О	222 Ю	238 о	254 ю
143 Џ	159 џ	175 Ї	191 ї	207 П	223 Я	239 п	255 я

Сообщение (текст): «ДУМАЙ!»

Знаковое кодирование Д У М А Й !

Кодировка Win 1251 110001001101001111001100110000001100100110000100

8-ричные цифры 6 1 1 5 1 7 1 4 6 0 1 4 4 6 0 4

8-ричные сигналы s_7 s_2 s_2 s_6 s_2 s_8 s_2 s_5 s_7 s_1 s_2 s_5 s_5 s_7 s_1 s_6

а)

Знаковое кодирование Д У М А Й !

Кодировка Win 1251 110001001101001111001100110000001100100110000100

32-ричные числа 24 19 9 28 25 16 6 9 16 16

32-ричные сигналы s_{25} s_{20} s_{10} s_{29} s_{26} s_{17} s_7 s_{10} s_{17} s_{17}

б)

Рис. П.1. Сообщения, знаки и символы:
а) 8-ричный пример; б) 32-ричный пример.

На рис. П.1, а) размер набора символов, *M,* был выбран равным 8 (каждый символ представляет восьмеричное число). Таким образом, биты группируются по три (k=Iog$_2$8); полученные в результате 16 чисел представляют 16 готовых к передаче восьмеричных символов. Передатчик должен иметь набор из восьми сигналов $s_i\ t$, где $i = 1,...,8$, сопоставляемых со всеми возможными символами, причем передача каждого сигнала возможна в течение времени символа. В последней строке рис. П.1, а) указаны 16 сигналов, передаваемых

восьмеричной системой модуляции для представления текстового сообщения "ДУМАЙ!".

На рис. П.1, 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 эффектное 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 $i = 1,...,32$, сопоставляемых со всеми возможными символами. На рис. В.1, б) указаны десять сигналов, передаваемых 32-ричной системой модуляции для представления текстового сообщения “ДУМАЙ!”.

Конец примера.

Скорость передачи данных. Любая система связи характеризуется скоростью передачи информации. Эта величина в битах в секунду (бит/с) дается формулой

$$R = k/T = \ 1/T\ \log_2\ M\ \ , \tag{1.7}$$

где *k* бит определяют символ из $M = 2^k$ символьного алфавита,

T - длительность *k*-битового символа.

Любая система связи характеризуется скоростью передачи и

Удельные затраты энергии и полосы частот. 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 передачи информации наиболее широко применяются удельные критерии качества.

Удельные затраты энергии будем характеризовать величиной

$$\beta_E = E_b / N_0\,, \tag{1.8}$$

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

$$\beta_W = W/R, \tag{1.9}$$

где W - эквивалентная (эффективная) полоса частот спектра сигнала, которой соответствует полоса пропускания приемного устройства,

R - информационная скорость передачи (бит/с).

Пропускная способность канала связи. Для гауссовского канала с белым шумом и ограниченной средней мощностью сигнала К.Шеннон получил формулу, определяющую максимум средней скорости передачи информации по такому каналу

$$C = W \log_2\left(1 + \frac{S}{N}\right), \tag{1.10}$$

где C - пропускная способность канала связи, бит/сек;

W - ширина полосы пропускания системы, Гц;

S - средняя мощность принятого сигнала, Вт;

N - средняя мощность шума, Вт.

Эта формула характеризует предельные возможности канала, в котором реализованы оптимальные кодирование и декодирование, обеспечивающие согласование производительности источника с пропускной способностью канала. Теоретически информация по каналу связи может быть передана со сколь угодно малой ошибкой при любой скорости передачи данных R, удовлетворяющей условию $R \le C$, что достигается при помощи использования сложных методов кодирования. В случае, если $R > C$, передача данных со сколь угодно малой вероятностью ошибки невозможна.

При сравнении различных видов модуляции обычно оперируют не отношением сигнал/шум, а отношением энергии бита к плотности мощности шума, которое определяется следующим образом

$$\frac{E_b}{N_0} = \frac{S}{N}\left(\frac{W}{R}\right), \tag{1.11}$$

где E_b - энергия бита, Вт/бит/сек; N_0 - плотность мощности шума, Вт/Гц.

Таким образом, выражение верхней границы пропускной способности канала связи может быть модифицировано

$$\frac{E_b}{N_0} = \frac{W}{C} 2^{C/W} - 1. \tag{1.12}$$

На рисунке В.1 показан график зависимости отношения нормированной полосы пропускания сигнала *W/C* от отношения энергии бита к плотности мощности шума E_b/N_0. Как видно из рисунка, существует нижнее предельное значение E_b/N_0, при котором ни при какой скорости передачи нельзя осуществить безошибочную передачу информации. Это значение E_b/N_0 называется *пределом Шеннона*

$$\beta_E = \frac{E_b}{N_0} = \ln\ 2\ \approx 0.7\ \ при\ \ \beta_W \to \infty . \qquad (1.13)$$

.

Выражение (1.13) 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 $\ln 2$. Зависимость (1.13) изображена на рис. 1.4. Эта зависимость называется границей Шеннона для идеального гауссовского канала связи. Она показывает, что существует бесконечное множество оптимальных систем, каждая из которых характеризуется своей парой β_E и β_W.

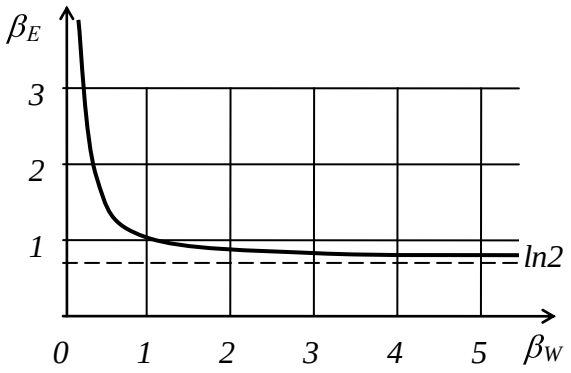


Рис. 1.4. Зависимость границы Шеннона для идеального гауссовского канала связи.

Из приведенного графика 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.

Шеннон теоретически доказал существование кодов, которые могут понизить вероятность битовой ошибки или снизить требуемое значение E_b/N_0 от уровней некодированных двоичных систем модуляции до ееровней, приближающихся к предельной кривой. Одной из основных задач разработчика современных систем цифровой связи является повышение спектральной эффективности модуляции, т. е. уменьшение отношения W/R до теоретического предела, что возможно при использовании современных методов канального кодирования.

2 МЕТОДЫ ЗАЩИТЫ ИНФОРМАЦИИ

Изучение путей передачи сообщений, которые не допускали бы постороннего вмешательства, называется *криптографией*. Термин криптография означает скрытный, и определяет теорию и технику специального кодирования, направленную на защиту информации от несанкционированного доступа и использования. Криптография как наука сформировалась только в середине 20-го века в значительной степени благодаря работам К. Шеннона.

Термин криптография, или шифрование обозначает преобразование исходного *открытого текста* $x \in X$, где X – множество всех возможных сообщений, в *криптограмму* (или *шифротекст*) y посредством ряда специальных взаимно обратимых преобразований F_z с целью скрыть смысл передаваемого открытого сообщения, т.е. осуществить операцию отображения вида

$$y = F_z(x),$$

где все возможные шифротексты образуют множество Y, т.е. $y \in Y$. Специфическое взаимно обратимое преобразование F_z (алгоритм преобразования), называемое *шифром*, однозначно определяется параметром z, который в свою очередь называется *ключом* шифрования. Параметры ключа известны только законным абонентам и неизвестны (секретны) для злоумышленников. Законный пользователь легко может осуществить обратное преобразование шифровки в исходный текст, а *криптоаналитик*, т.е. субъект, пытающийся несанкционированно получить открытый текст, будет вынужден угадывать значение ключа.

Схемы (или алгоритмы) шифрования можно разбить на две категории: *блочное* и *поточное*. При блочном шифровании открытый текст делится на блоки фиксированного размера m, после чего каждый блок шифруется независимо друг от друга. Следовательно, одинаковые блоки открытого текста с помощью определенного ключа преобразуются в одинаковые блоки шифрованного текста. При поточном шифровании разбиение на блоки фиксированного объема не производится. Каждый символ (бит) открытого текста x_i шифруется помощью i–го элемента z_i последовательности символов (ключевого потока), генерируемого ключом.

При блочном шифровании единственный бит ошибки на входе дешифратора может изменить значения многих выходных битов в блоке. Данный эффект, известный как *накопление ошибок*, часто рассматривается желаемым криптографическим свойством, поскольку создает для несанкционированных пользователей дополнительные сложности при расшифровке сообщений. Тогда как при канальном кодировании данный эффект считается нежелательным.

2.1 Система шифрования

На рис. 2.1 представлена блок-схема системы секретной связи, иллюстрирующая пример установления канала секретной связи между отправителем (абонент *A*) и получателем (абонент *B*).

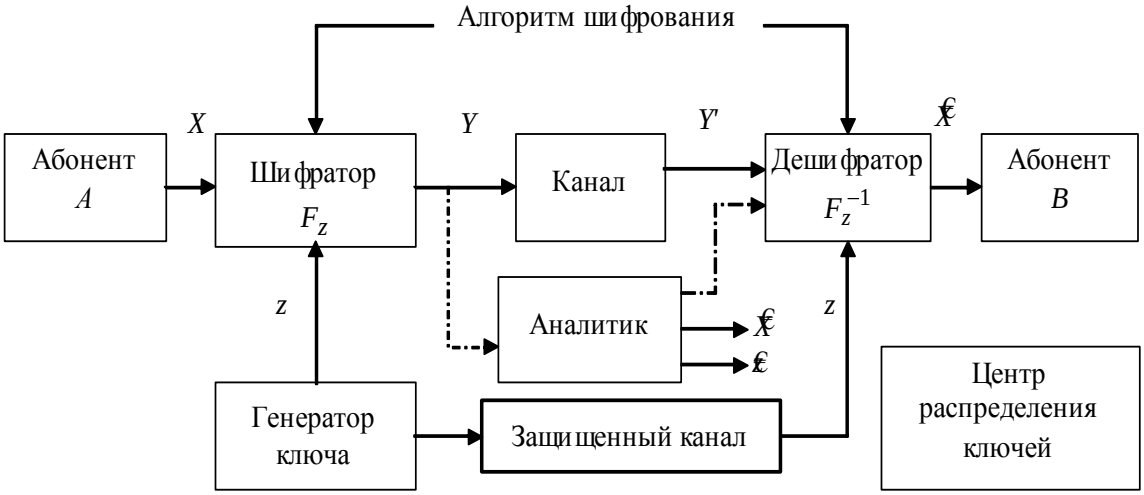


Рис. 2.1. Схема секретной связи.

На передающей стороне исходное сообщение x с помощью известного алгоритма F_z и секретного ключа z преобразуется в шифротекст $y = F_z(x)$. В качестве компонентов этой криптограммы могут выступать биты, символы, взятые из одного или нескольких алфавитов разных объемов. На приемной стороне получатель, располагая ключом z, доставляемым по защищенному каналу, осуществляет обратное преобразование F_z^{-1} над принятым шифротекстом, т.е. осуществляет расшифровку и восстанавливает исходный открытый текст

$$F_z^{-1}(y) = F_z^{-1}\ F_z(x)\ = x\,.$$

Помимо отправителя и получателя в системе передачи присутствует и третий участник, называемый криптоаналитиком, или злоумышленником, который, перехватив шифрограмму, пытается понять исходный текст или определить ключ шифрования. Кроме того, в случае нахождения ключа, он может внедрить в систему собственную, фальшивую криптограмму с целью обмана абонента *B*.

Из приведенной схемы видно, что для оперативной доставки ключа шифрования на приемную сторону должен быть предусмотрен специальный, защищенный канал. Возникает задача распределения ключей шифрования между пользователями или задача управления ключами, гарантирующая от риска перехвата ключа злоумышленником. Особое значение проблема распределения ключей приобретает в случае большого числа абонентов (например, мобильной цифровой связи или Интернета), когда многочисленные отправители и получатели обмениваются между собой сообщениями с

минимумом риска быть подслушанными неуполномоченными пользователями той же системы.

В зависимости от способа решения указанной задачи различают два больших класса криптосистем: *симметричные (одноключевые)* и *асимметричные (двухключевые)*. В классических симметричных системах для шифрования и расшифровки сообщений используется один и тот же ключ. Распределение ключей в симметричных системах шифрования является чрезвычайно серьезной задачей. В асимметричных системах для формирования шифрограммы используется открытый, известный всем (и злоумышленнику) ключ, а для получения по шифровке исходного текста - другой, секретный, известный только законному получателю сообщения.

Для обоих типов систем стойкость системы шифрования, т. е. свойство противостоять действиям криптоаналитика по вскрытию исходного текста, во многом определяется трудностью разгадывания ключа.

Соответствие между открытыми сообщениями и криптограммами иллюстрирует рис. 2.6.

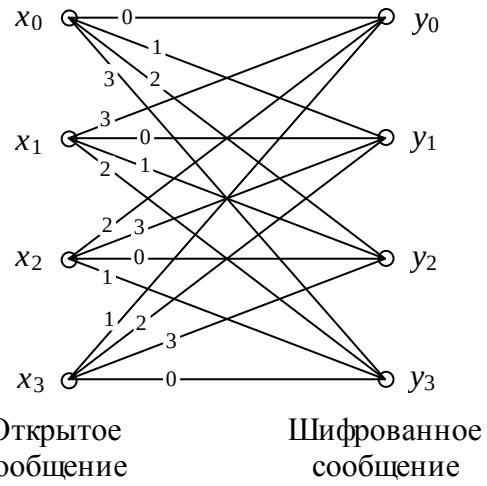


Рис. 2.6.

Криптоаналитик, перехвативший один из шифротекстов y_s, $s = 0,1,2,3$, не сможет определить, какой из четырех ключей использовался и, следовательно, какое из x_i, $i = 0,1,2,3$ является верным.

Общепринятой считается практика разделения потока данных источника на блоки длины m символов и шифрования каждого блока с использованием определенного секретного ключа. В этом случае речь может идти только об условной или т.н. *вычислительной* (*практической*) секретности. Она подразумевает выбор такого ансамбля ключей, который гарантирует возможность взлома криптосистемы только ценой значительных (может быть даже нереальных) вычислительных и временных затрат. Очевидно, что одним из факторов, обуславливающих вычислительную стойкость, является число потенциальных ключей $N_z = 2^m$, выбираемых для обеспечения совершенной секретности в рамках отдельного m–символьного блока.

Однако имеется и фактор противоположной направленности: присущая всем без исключения языкам избыточность. В явном виде избыточность языка проявляет себя следующим образом: только незначительная часть общего числа комбинаций из *m* букв имеет осмысленное значение, что позволяет аналитику при вскрытии шифра отбрасывать потенциальные ключи, приводящие к лишенным смысла комбинациям.

При раскрытии многих схем шифрования может применяться статистический анализ, использующий частоту появления отдельных символов или их комбинаций.

2.2 Простейшие методы шифрования

Рассмотрим элементарные методы шифрования такие, как метод подстановки, перестановки и скремблирования.

2.2.1 Метод подстановки

Простейшим и имеющим наибольшую историю является метод подстановки, суть которого заключается в том, что символ исходного текста заменяется другим, выбранным из этого или другого алфавита по правилу, задаваемому ключом шифрования. Местоположение символа в тексте при этом не изменяется. В нем каждая буква открытого текста заменялась другой, взятой из того же алфавита, но циклически сдвинутого на определенное количество символов. Применение данного метода шифрования иллюстрирует пример, представленный в табл. 2.1, в которой шифрующее преобразование основано на использовании алфавита с циклическим сдвигом на пять позиций.

Таблица 2.1, а.

А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.
Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ь	Ы	Ь	Э	Ю	Я		.	А	Б	В	Г	Д

Таблица 2.1, в.

Исходный текст	Г	А	Й		Ю	Л	И	Й		Ц	Е	З	А	Р	Ь
Криптограмма	И	Е	О	Г	Б	Р	Н	О	Г	Ы	К	М	Е	Х	.

Очевидно, что ключом шифра служит величина циклического сдвига. При выборе другого ключа, чем указано в примере, шифр будет изменяться.

Другим примером классической схемы, основанной на методе подстановки, может служить система шифрования, называемая *квадратом Полибиуса*. Применительно к русскому алфавиту данная схема может быть описана следующим образом. Первоначально объединяются в одну буквы Е, Ё; И, Й и Ъ, Ь, истинное значение которых в дешифрованном тексте легко восстанавливается из контекста. Затем 30 символов алфавита размещаются в таблицу размером 6×5, пример заполнения которой представлен в таблице 2.6.

Таблица 2.6.

	1	2	3	4	5
1	А	Б	В	Г	Д
2	Е	Ж	З	И	К
3	Л	М	Н	О	П
4	Р	С	Т	У	Ф
5	Х	Ц	Ч	Ш	Щ
6	Ь	Ы	Э	Ю	Я

Шифрование любой буквы открытого текста осуществляется заданием ее адреса (т.е. номера строки и столбца или наоборот) в приведенной таблице. Так, например, слово ЦЕЗАРЬ шифруется с помощью квадрата Полибиуса как 52 21 23 11 41 61. Совершенно ясно, что изменение кода может быть осуществлено в результате перестановок букв в таблице.

Примером *полиалфавитного* шифра может служить схема, основанная на т.н. *прогрессивном ключе Тритемиуса*. Основой данного способа шифрования служит таблица 2.3, строки которой представляют собой циклически сдвинутые на одну позицию копии исходного алфавита. Так, первая строка имеет нулевой сдвиг, вторая циклически сдвинута на одну позицию влево, третья - на две позиции относительно первой строки и т.д.

Таблица 2.3.

А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.
Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А
В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А	Б
Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А	Б	В
Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А	Б	В	Г
Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А	Б	В	Г	Д
Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А	Б	В	Г	Д	Е
З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А	Б	В	Г	Д	Е	Ж
И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А	Б	В	Г	Д	Е	Ж	З
Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А	Б	В	Г	Д	Е	Ж	З	И
К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А	Б	В	Г	Д	Е	Ж	З	И	Й
Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А	Б	В	Г	Д	Е	Ж	З	И	Й	К
М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л
Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М
О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н
П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О
Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П
С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р
Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С
У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т
Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У

Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф
Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х
Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц
Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч
Щ	Ъ	Ы	Ь	Э	Ю	Я		.	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш
Ъ	Ы	Ь	Э	Ю	Я		.	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ
Ы	Ь	Э	Ю	Я		.	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ
Ь	Э	Ю	Я		.	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы
Э	Ю	Я		.	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь
Ю	Я		.	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ь	Ы	Ь	Э
Я		.	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю
	.	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я
.	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	

Один из методов шифрования с помощью подобной таблицы состоит в использовании вместо первого символа открытого текста символа из первого циклического сдвига исходного алфавита, стоящего под шифруемым символом, второго символа открытого текста - из строки, соответствующей второму циклическому сдвигу и т.д. Пример шифрования сообщения подобным образом представлен в таблице 16.6.

Таблица 2.6.												
Открытый текст	К	Р	И	П	Т	О	Г	Р	А	М	М	А
Шифрованный текст	Л	Т	Л	У	Ч	Ф	К	Ш	Й	Ц	Ч	М

Известны несколько интересных вариантов шифров, основанных на прогрессивном ключе Тритемиуса. В одном из них, называемом *методом ключа Вижинера*, применяется ключевое слово, которое указывает строки для шифрования и расшифрования каждого последующего символа открытого текста. Первая буква ключа указывает строку таблицы 2.3, с помощью которой шифруется первый символ сообщения, вторая буква ключа определяет строку таблицы, шифрующей второй символ открытого текста и т.д. Пусть в качестве ключа выбрано слово «ТРОМБ», тогда сообщение, зашифрованное с помощью ключа Вижинера, может быть представлено следующим образом (см. таблицу 2.5). Очевидно, что вскрытие ключа возможно осуществить на основе статистического анализа шифрограммы.

Таблица 2.56.												
Открытый текст	К	Р	И	П	Т	О	Г	Р	А	М	М	А
Ключ	Т	Р	О	М	Б	Т	Р	О	М	Б	Т	Р
Шифрованный текст	Ь		Ц	Ы	У		У	Ю	М	Н	Ю	Р

Разновидностью этого метода является т.н. *метод автоматического* (*открытого*) *ключа Вижинера*, в котором в качестве *образующего ключа*

используется единственная буква или слово. Этот ключ дает начальную строку или строки для шифрования первого или нескольких первых символов открытого текста аналогично ранее рассмотренному примеру. Затем в качестве ключа для выбора шифрующей строки используются символы открытого текста. В приведенном ниже примере в качестве образующего ключа использована буква «И» (см. таблицу 2.6):

Таблица 2.4.												
Открытый текст	К	Р	И	П	Т	О	Г	Р	А	М	М	А
Ключ	И	К	Р	И	П	Т	О	Г	Р	А	М	М
Шифрованный текст	Т	Ъ	Ш	Ч	.		С	У	Р	М	Ш	М

Как показывает пример, выбор строк шифрования полностью определяется содержанием открытого текста, т.е. в процесс шифрования вводится обратная связь по открытому тексту.

Еще одной разновидностью метода Вижинера служит *метод автоматического* (*шифрованного*) *ключа Вижинера*. В нем, подобно шифрованию с открытым ключом, также используется образующий ключ и обратная связь. Отличие состоит в том, что после шифрования с помощью образующего ключа, каждый последующий символ ключа в последовательности берется не из открытого текста, а из получаемой криптограммы. Ниже представлен пример, поясняющий принцип применения данного метода шифрования, в котором, как и ранее, в качестве образующего ключа использована буква «И» (см. таблицу 16.7).

Как видно из приведенного примера, хотя каждый последующий символ ключа определяется предшествующим ему символом криптограммы, функционально он зависит от всех предшествующих символов открытого сообщения и образующего ключа. Следовательно, наблюдается эффект рассеивания статистических свойств исходного текста, что затрудняет применение статистического анализа криптоаналитиком. Слабым звеном данного метода является то, что шифрованный текст содержит символы ключа.

Таблица 2.6.												
Открытый текст	К	Р	И	П	Т	О	Г	Р	А	М	М	А
Ключ	И	Т	А	И	Ч	З	Х	Ш	Ж	Ж	Т	Ю
Шифрованный текст	Т	А	И	Ч	З	Х	Ш	Ж	Ж	Т	Ю	Ю

Вариантом реализации подстановочной технологии, который в достаточной степени реализует концепцию смешения, служит следующий пример, базирующийся на нелинейном преобразовании. Поток информационных бит предварительно разбивается на блоки длиной m, причем каждый блок представляется одним из 2^m различных символов. Затем множество из 2^m символов перемешивается таким образом, чтобы каждый символ заменялся другим символом из этого множества. После операции перемешивания символ вновь превращается в m–битовый блок. Устройство,

реализующее описанный алгоритм при $m=3$, представлено на рис. 2.3, где в таблице 2.8 задано правило перемешивания символов множества из 2^m элементов.

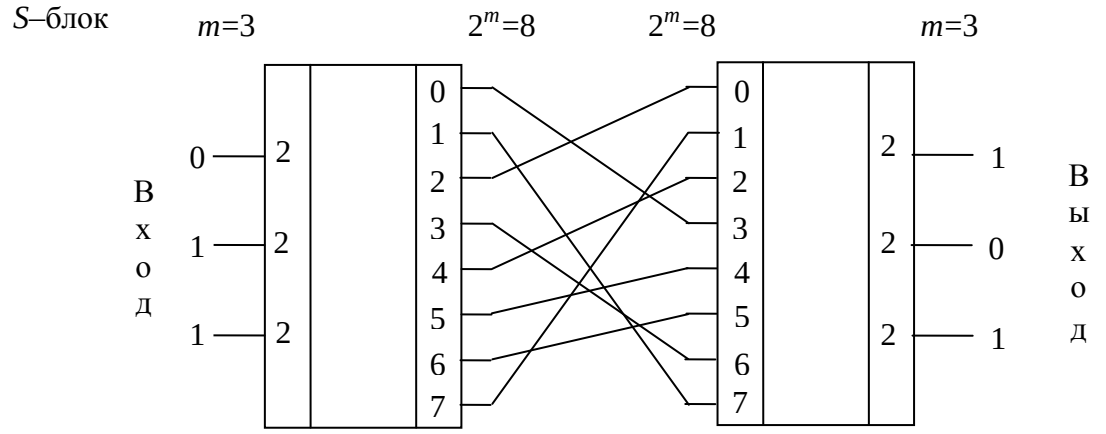


Рис. 2.3.

Таблица 2.4.								
Вход	000	001	010	011	100	101	110	111
Выход	011	111	000	110	010	100	101	001

Не составляет труда показать, что существует $(2^m)!$ различных подстановок или связанных с ними возможных моделей. В связи, с чем при больших значениях m задача криптоаналитика становится в вычислительном плане практически невозможной. Например, при $m=128$ число возможных подстановок определяется как $(2^{128})!\approx(10^{38})!$, т.е. представляет собой астрономическое число. Очевидно, что при подобном значении m данное преобразование с помощью *блока подстановки* (*S*-блок) можно считать обладающим вычислительной секретностью. Однако его практическая реализация вряд ли возможна, поскольку предполагает существование $2^{128}\approx10^{38}$ соединений.

Легко убедиться, что *S*–блок, представленный на рис. 2.3, действительно осуществляет нелинейное преобразование, для чего воспользуемся принципом суперпозиций. Преобразование $F_z(\cdot)$ является линейным, если выполняется условие, что $F_z(x_1+x_2)=F_z(x_1)+F_z(x_2)$. Предположим, что $x_1=101$, а $x_2=010$. Тогда $F_z(101\oplus010)=F_z(111)=001$, а $F_z(101)+F_z(010)=100\oplus000=100$, откуда и следует, что *S*-блок является нелинейным.

2.2.2 Метод перестановки

При *перестановке* (или *транспозиции*) в соответствии с ключом изменяется порядок следования символов открытого текста, а значение символа при этом сохраняется. Шифры перестановки являются блочными, т. е.

исходный текст предварительно разбивается на блоки, в которых и осуществляется заданная ключом перестановка.

Простейшим вариантом реализации данного метода шифрования может служить рассмотренный ранее алгоритм перемежения, суть которого заключается в разбиении потока информационных символов на блоки длиной $m = k \times n$, построчной записи его в матрицу памяти размером k строк и n столбцов и считывании по столбцам. Иллюстрацией данному алгоритму служит пример с $m = 5 \times 5$ (см. таблицу 16.9), в ходе которого производится запись фразы X=«СКОРО НАЧНЕТСЯ ЭКЗАМЕНАЦИ|ОННАЯ ПОРА». Тогда на выходе устройства перестановки будет получена криптограмма вида $Y =$ С ЕЭЕКНТКНОАСЗАРЧЯАЦОН МИ.

Таблица 2.6.					
Z	1	2	3	4	5
1	С	К	О	Р	О
2		Н	А	Ч	Н
3	Е	Т	С	Я	
4	Э	К	З	А	М
5	Е	Н	А	Ц	И

Рассмотренный вариант метода перестановки может быть усложнен введением ключей Z_X и Z_Y, определяющих порядок записи строк и считывания столбцов соответственно, иллюстрацией чему служит таблица 2.6. Результат преобразования будет иметь следующий вид: $Y =$ ЯРЦАЧСОАЗАТКНКНЕСЕЭ ОИМН.

Таблица 2.6.					
Z	4	3	2	1	5
3	Е	Т	С	Я	
1	С	К	О	Р	О
5	Е	Н	А	Ц	И
4	Э	К	З	А	М
2		Н	А	Ч	Н

На рис. 2.4 приведен пример бинарной перестановки данных (линейная операция), из которого видно, что данные просто перемешиваются или переставляются. Преобразование осуществляется с помощью блока перестановки (*P*–блок). Технология перестановки, реализуемая этим блоком,

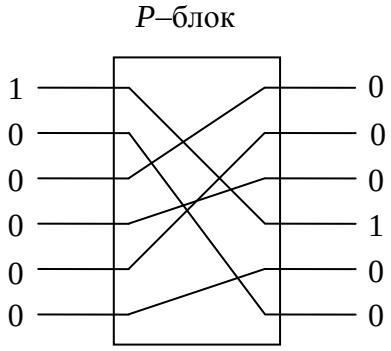


Рис. 2.6.

имеет один основной недостаток: она уязвима по отношению к обманным сообщениям. Обманное сообщение изображено на рис. 2.4 и заключается в подаче на вход одной единственной единицы при остальных нулях, что позволяет обнаружить одну из внутренних связей. Если криптоаналитику необходимо осуществить анализ подобной схемы с помощью атаки открытого текста, то он отправит последовательность подобных обманных сообщений, смещая при каждой передаче единственную единицу на одну позицию. В результате подобной атаки будут установлены все связи входа и выхода. Данный пример демонстрирует, почему защищенность схемы не должна зависеть от ее архитектуры.

2.2.3 Метод гаммирования

Гаммирование относится к поточным методам шифрования, когда следующие друг за другом символы открытого текста последовательно превращаются в символы шифрограммы, что повышает скорость преобразования.

Попытки приблизиться к совершенной секретности демонстрируют многие современные телекоммуникационные системы, использующие операцию скремблирования. 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). Так, например, поток информационных бит $U_X = \ldots, u_{-1}, u_0, u_1, \ldots$ поступает на один вход сумматора по модулю 2, изображенного на рис. 2.5, тогда как на второй - скремблирующая двоичная последовательность

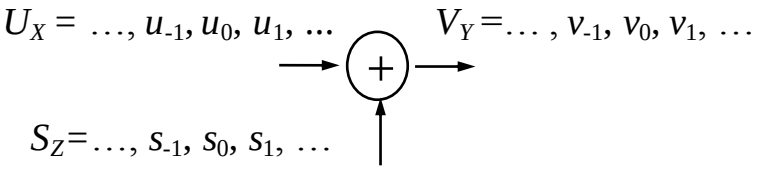


Рис. 2.6.

$S_Z = \ldots, s_{-1}, s_0, s_1, \ldots$.

В идеале последовательность S_Z должна быть случайной последовательностью с равновероятными значениями нулей и единиц. Тогда выходной шифрованный поток $V_Y = \ldots, v_{-1}, v_0, v_1, \ldots$ будет статистически независимым от информационной последовательности U_X, а значит, будет выполняться достаточное условие совершенной секретности.

В действительности абсолютная случайность S_Z не является необходимой, поскольку в противном случае получатель не сможет восстановить открытый текст. Действительно, восстановление открытого текста на приемной стороне должно производиться по правилу $U_X = V_Y \oplus S_Z$, так что на приемной стороне должна генерироваться точно такая же скремблирующая последовательность и с той же фазой. Однако вследствие абсолютной случайности S_Z данная процедура становится невозможной.

На практике в качестве скремблирующих широкое применение нашли псевдослучайные последовательности (ПСП), которые могут быть воспроизведены на приемной стороне. Так, например, в сотовой системе мобильной связи стандарта IS-95 в качестве скремблирующей используется ПСП периода 2^{42} в числе элементарных чипов. При чиповой скорости $1.2288 \cdot 10^6$ симв/сек ее период составляет:

$$\frac{2^{42}}{1.2288 \cdot 10^6} = \frac{2^{10} 2^{10} 2^{10} 2^{10} 2^2}{1.23 \cdot 10^6} = \frac{4 \cdot 10^{12}}{1.23 \cdot 10^6} \approx 3.6 \cdot 10^6 \text{ сек} \approx 6 \cdot 10^5 \text{ мин} \approx 994 \text{ час} \approx 41.4 \text{ суток}.$$

Следовательно, можно предполагать, что поскольку последовательность не повторяется в течение такого длительного периода, то она может рассматриваться случайной и обеспечивать совершенную секретность.

2.3. Стандарт шифрования данных

В настоящее время любая система шифрования с секретным ключом, основана на использовании длинной последовательности элементарных операций шифрования таких, как подстановка и перестановка. Примером может служить разработанная в 1977 г. Национальным бюро стандартов США система, используемая в качестве Национального стандарта шифрования данных (DES) (см. рис. 2.9).

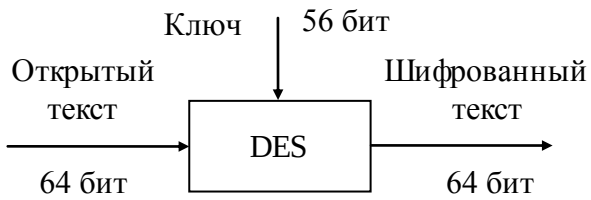


Рис. 2.6.

Система DES является блочной системой шифрования с алфавитом из 2^{64} символов Входной блок из 64 бит заменяется новым символом шифрованного текста. Алгоритм шифрования начинается с начальной перестановки 64 бит открытого текста, описываемой таблицей 2.16.

Таблица 2.16. Таблица начальной перестановки.

58	50	42	34	26	18	10	2
60	52	44	36	28	20	12	4
62	54	46	38	30	22	14	6
64	56	48	40	32	24	16	8
57	49	41	33	25	17	9	1
59	51	43	35	27	19	11	3
61	53	45	37	29	21	13	5
63	55	47	39	31	23	15	7

В результате данной операции исходный блок $\{x_1, x_2, x_3, \ldots, x_{62}, x_{63}, x_{64}\}$ преобразуется в блок $\{x_{58}, x_{50}, x_{42}, \ldots, x_{23}, x_{15}, x_7\}$. После этой начальной перестановки начинается основная часть алгоритма шифрования, представляющая собой 16-ти этапную процедуру операций подстановки и перестановки, определяемую 48-битовым ключом. Стандартный блок, описывающий данную процедуру, представлен на рис. 2.6.

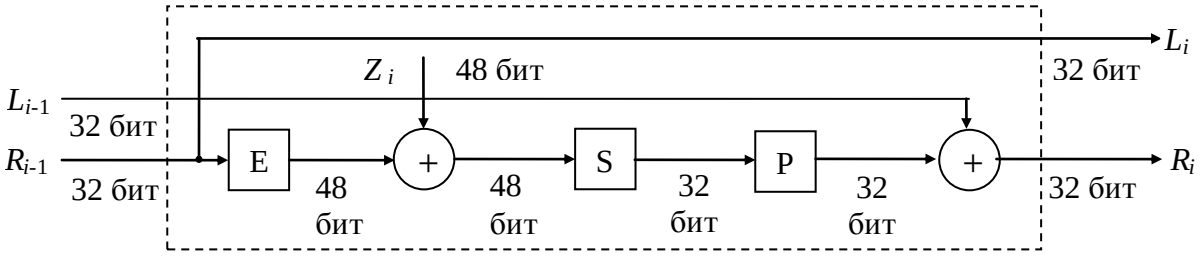


Рис2.6.

Перед поступлением в стандартный блок входной блок данных в 64 бита разбивается на две половины: правую (R) и левую (L), а функционирование стандартного блока осуществляется под управлением 48-битового ключа. Выход каждого стандартного блока становится входом следующего. На i–1 этапе входные 32 бит правой половины R_{i-1} без изменений подаются на выход и становятся 32 битами левой половины L_i на следующем i–м этапе. Эти же 32 бита R_{i-1} расширяются в блоке E до 48 бит согласно таблице 2.7. Так, например, если $R_{i-1} = \{x_1, x_2, x_3, \ldots, x_{30}, x_{31}, x_{32}\}$, то с выхода блока E снимается последовательность бит вида: $(R_{i-1})_E = \{x_{32}, x_1, x_2, \ldots, x_{31}, x_{32}, x_1\}$.

Отметим, что биты расположенные в первом и последнем столбцах таблицы расширения отвечают битам, используемым дважды для расширения с 32 до 48 бит.

Таблица 2.7. Таблица расширения.					
32	1	2	3	4	5
4	5	6	7	8	9
8	9	10	11	12	13
12	13	14	15	16	17
16	17	18	19	20	21
20	21	22	23	24	25
24	25	26	27	28	29
28	29	30	31	32	1

Далее блок $(R_{i-1})_E$ суммируется по модулю 2 с 48-битовым Z_i ключом, выбор которого будет описан несколько позднее, а результат суммирования представляет собой восемь 6-битовых блоков, т.е. $(R_{i-1})_E \oplus Z_i = B_1, B_2, \ldots, B_8$.
В свою очередь каждый из восьми 6-битовых блоков $B_j, j = 1,\ldots,8$ выступают в качестве входных данных для *S*–блока, преобразующего 6-ти битовые блоки B_j в 4-битовые $S_j(B_j)$, так что на выходе *S*–блока будет сформирован 32-битовый блок. Отображение входных данных *S*–блоком описывается таблицей 2.14, для краткости представленной только своей 1/8 частью.

Таблица 2.16. Таблица отображения *S*–блоком.																	
Строка	Столбец																
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	S_1
0	14	4	13	1	2	15	11	8	3	10	6	12	5	9	0	7	
1	0	15	7	4	14	2	13	1	10	6	12	11	9	5	3	8	
2	4	1	14	8	13	6	2	11	15	12	9	7	3	10	5	0	
3	15	12	8	2	4	9	1	7	5	11	3	14	10	0	6	13	

Преобразование блока $B_j = b_1, b_2, b_3, b_4, b_5, b_6$ в блок $S_j(B_j)$ осуществляется по следующему алгоритму. Значения бит, отвечающие b_1, b_6, определяют номер строки таблицы, а b_2, b_3, b_4, b_5 - номер столбца. Например, если $B_1 = 110001$, то в результате преобразования *S*–блоком получаем значение $S_1(B_1) = 5$, определяемое третьей строкой и 8-м столбцом таблицы, которое в итоге представляется в двоичной форме: 0101. Аналогично происходит преобразование и остальных B_j, $j = \overline{2,8}$.32-битовый поток, сформированный *S*–блоком, подвергается операции перестановки в *P*–блоке согласно правилу, задаваемому таблицей 2.156. В результате данной перестановки из 32-битового блока $\{x_1, x_2, x_3, \ldots, x_{30}, x_{31}, x_{32}\}$ формируется блок вида $\{x_{16}, x_7, x_{20}, \ldots, x_{11}, x_4, x_{25}\}$. Блок в 32 бита с выхода *P*–блока суммируется по модулю два с 32 битами левой половины (L_{i-1}), образуя выходные 32 бита, использующиеся в качестве правой

половины (R_i). Таким образом, алгоритм функционирования стандартного блока может быть описан, как

$$L_i = R_{i-1},$$
$$R_i = L_{i-1} \oplus F_{z_i}(R_{i-1})'$$

где функциональное преобразование $F_{z_i}(R_{i-1})$ обозначает описанные выше операции.

Таблица 2.6.
Таблица перестановки.

16	7	20	21
29	12	28	17
1	15	23	26
5	18	31	10
2	8	24	14
32	27	3	9
19	13	30	6
22	11	4	25

В результате данной перестановки из 32-битового блока $\{x_1, x_2, x_3, \ldots, x_{30}, x_{31}, x_{32}\}$ формируется блок вида $\{x_{16}, x_7, x_{20}, \ldots, x_{11}, x_4, x_{25}\}$. Блок в 32 бита с выхода P–блока суммируется по модулю два с 32 битами левой половины (L_{i-1}), образуя выходные 32 бита, использующиеся в качестве правой половины (R_i). Таким образом, алгоритм функционирования стандартного блока может быть описан, как

$$L_i = R_{i-1},$$
$$R_i = L_{i-1} \oplus F_{z_i}(R_{i-1})'$$

где функциональное преобразование $F_{z_i}(R_{i-1})$ обозначает описанные выше операции.

После 16 этапов реализации стандартного блока операций данные размещаются согласно окончательной обратной перестановке, задаваемой таблицей 2.16:

Таблица 2.14. Таблица
окончательной перестановки.

40	8	48	16	56	24	64	32
39	7	47	15	55	23	63	31
38	6	46	14	54	22	62	30
37	5	45	13	53	21	61	29
36	4	44	12	52	20	60	28
35	3	43	11	51	19	59	27
34	2	42	10	50	18	58	26
33	1	41	9	49	17	57	25

Для расшифрования применяется тот же алгоритм, однако ключевая

последовательность, используемая в стандартном блоке, берется в обратном порядке.

В ходе выполнения шифрования в стандарте DES выбор ключа также претерпевает 16-этапную процедуру. Входной ключ представляет собой 64-битовый блок с 8 битами четности в разрядах 8,16,…, 6. В ходе первой процедуры перестановки производится отбрасывание бит четности и размещение оставшихся 56 бит согласно правилу, описываемому таблицей 2.17 размером 8×7, так что вместо исходной последовательности $\{z_1, z_2, z_3, \ldots, z_{62}, z_{63}, z_{64}\}$ на выходе P-блока имеем $\{z_{57}, z_{49}, z_{41}, \ldots, z_{20}, z_{12}, z_4\}$.

Таблица 2.2. Таблица первой перестановки.

57	49	41	33	25	17	9
1	58	50	42	14	26	18
10	2	59	51	43	35	27
19	11	3	60	52	44	36
63	55	47	39	31	23	15
7	62	54	46	38	30	22
14	6	61	53	45	37	29
21	13	5	28	20	12	4

Полученный в результате перестановки блок делится на два подблока C и D, каждый длиной 28 бит. Выбор 48-битового ключа, используемого на каждом из этапов шифрования, осуществляется следующим образом. Первоначально блоки C и D подвергаются сдвигу согласно следующим выражениям

$$C_i = LS_i(C_{i-1}) \quad \text{и} \quad D_i = LS(D_{i-1}),$$

где LS_i - левый циклический сдвиг соответствующего блока C (или D), полученного на $i-1$ этапе, на число позиций, указанных в таблице 2.18:

Таблица 2.14. Таблица величины циклического сдвига.

Номер этапа	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Величина сдвига	1	1	2	2	2	2	2	2	1	2	2	2	2	2	2	1

Затем блоки C_i, D_i подвергаются операции перестановки в соответствии с порядком, указанным в таблице 2.16.

Таблица 2.16. Таблица второй перестановки.

14	17	11	24	1	5
3	28	15	6	21	10
23	19	12	4	26	8
16	7	27	20	13	2
41	52	31	37	47	55
30	40	51	45	33	48
44	49	39	56	34	53
46	42	50	36	29	32

из которой видно, что отбрасываются биты, стоящие на позициях 9, 18, 22, 25, 35, 38, 43 и 54, т.е. по 4 из каждого подблока. В результате на выходе второго

P-блока образуется ключевая последовательность Z_i, которая используется на i-м этапе шифрования.

Согласно некоторым опубликованным оценкам, раскрытие ключа стандарта DES с помощью гипотетической для нашего времени ЭВМ со скоростью 10^{12} длинных операций в секунду потребует 10^{10} лет. За более чем двадцатилетнее существование этого стандарта не сообщалось о хотя бы минимальном успехе в его раскрытии.

3 ПОМЕХОУСТОЙЧИВОЕ КОДИРОВАНИЕ ИНФОРМАЦИИ

Повышение требований к достоверности передачи информации приводит к необходимости принятия специальных мер, направленных на уменьшение вероятности возникновения ошибок в процессе передачи. Одним из возможных решений указанной задачи служит *помехоустойчивое кодирование*. Под помехоустойчивыми понимаются коды, позволяющие обнаруживать и исправлять ошибки, возникающие при передаче из-за воздействия помех.

При кодировании информации исходные данные делятся на блоки из k бит, которые называют информационными битами, или битами сообщения.

В процессе кодирования каждый k - битовый блок данных $m = m_1, m_2, ..., m_k$ преобразуется в больший блок из n бит, который называется кодовым словом $U = u_1, u_2 ..., u_n$.

К каждому блоку данных кодирующее устройство прибавляет $r = n - k$ бит, которые называются *избыточными битами* или *контрольными битам.* Для обозначения описанного кодового слова используется запись n, k .

Расстояние между двумя любыми кодовыми комбинациями определяется числом позиций, в которых эти комбинации имеют различные символы. Например, расстояние между комбинациями 0001101 и 1001010 равно $d = 4$. Минимальное расстояние между разрешаемыми комбинациями данного кода $d_{\min}$ называется расстоянием Хэмминга. Минимальное расстояние линейного кода равно наименьшему из весов ненулевых слов кода:

$$d = \min_{\mathbf{U}_i \neq \mathbf{0}} w\ \mathbf{U}_i \ .$$

Если код используется только для обнаружения ошибок кратностью α, то необходимо и остаточно, чтобы $d_{\min} \geq \alpha + 1$. Для исправления всех ошибок кратности α и менее необходимо иметь $d_{\min} \geq 2\alpha + 1$.

Правила преобразования исходного потока символов в новый поток определяет метод помехоустойчивого кодирования.

3.1. Систематические линейные коды

Систематический линейный код n, k - это такое отображение k-мерного вектора сообщения $m = m_1, m_2, ..., m_k$ в n-мерное кодовое слово $U = u_1, u_2 ..., u_n$, в котором часть формируемой кодовой последовательности совмещается с k символами сообщения. Остальные $n - k$ бит - это биты четности: $p_0, p_1, ..., p_{n-k}, m_1, m_2, ..., m_k$. Систематические кодовые слова иногда записываются так, чтобы биты сообщения занимали левую часть кодового слова, а биты четности - правую: $m_1, m_2, ..., m_k, p_0, p_1, ..., p_{n-k}$. Такая

перестановка не влияет на свойства кода, связанные с процедурами обнаружения и исправления ошибок.

Введем в рассмотрение *порождающую* (*производящую*) матрицу **G** систематического линейного блочного кода

$$\mathbf{G} = \left[\mathbf{P} \vdots \mathbf{I}_k\right] = \begin{bmatrix} p_{11} & p_{12} & \dots p_{1\,n-k} & 1 & 0 & \dots & 0 \\ p_{21} & p_{22} & \dots p_{2\,n-k} & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ p_{k1} & p_{k2} & \dots p_{k\,n-k} & 0 & 0 & \dots & 1 \end{bmatrix}. \quad (3.1)$$

Здесь **P** - массив четности, содержащий элементы $p_{ij} = 0{,}1$, $i = 1,...,k$, $j = 1,...,\ n-k$; $\mathbf{I}_k$ - единичная матрица размерностью $k \times k$.

Рассмотрим передаваемое сообщение из k бит: $m = m_1, m_2, ..., m_k$. Тогда используя формулу (3.1) , формируемое кодовое слово можно представить в следующем виде

$$U = u_1, u_2, ... u_n = \left[m_1, m_2, ..., m_k\right] \cdot \begin{bmatrix} p_{11} & p_{12} & \dots p_{1\,n-k} & 1 & 0 & \dots & 0 \\ p_{21} & p_{22} & \dots p_{2\,n-k} & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ p_{k1} & p_{k2} & \dots p_{k\,n-k} & 0 & 0 & \dots & 1 \end{bmatrix}, \quad (3.2)$$

где

$$u_i = \begin{cases} m_1 p_{1i} + m_2 p_{2i} + \ldots + m_k p_{ki}, & \text{для } i = 1, \ldots, \ n-k \\ m_{i-\ n-k}, & \text{для } i = \ n-k+1\ , \ldots, n \end{cases}$$

Систематический кодовый вектор можно записать в следующем виде

$$U = \underbrace{p_1, p_2, ..., p_{n-k}}_{\text{биты четности}}, \underbrace{m_1, m_2, ..., m_k}_{\text{биты сообщения}},$$

где

$$\begin{aligned} p_1 &= m_1 p_{11} + m_2 p_{21} + \ldots + m_k p_{k1} \\ p_2 &= m_1 p_{12} + m_2 p_{22} + \ldots + m_k p_{k\,2} \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ p_{n-k} &= m_1 p_{1\ n-k} + m_2 p_{2\ n-k} + \ldots + m_k p_{k\ n-k} \end{aligned} \quad (3.3)$$

В качестве иллюстрации воспользуемся кодом (6,3) и рассмотрим структуру кодового слова

$$\mathrm{U}=\left[m_1,m_2,m_3\right]\cdot\left[\begin{array}{c}\\ 1\ 1\ 0\ .1\ 0\ 0\\ 0\ 1\ 1\ .0\ 1\ 0\\ \underbrace{1\ 0\ 1}_{\mathrm{P}}\ .0\ 0\ 1\\ \end{array}\right]=\underbrace{m_1+m_3}_{u_1},\underbrace{m_1+m_2}_{u_2},\underbrace{m_2+m_3}_{u_3},m_1,m_2,m_3\ \ u_4\ u_5\ u_6$$

Соответствие векторов сообщения m и кодовых слов U представлены в таблице 3.1

Табл. 3.1.

Вектор сообщения	Кодовое слово	Вектор сообщения	Кодовое слово
000	000000	001	101001
100	110100	101	011101
010	011010	011	110011
110	101110	111	000111

Далее определим матрицу **H**, именуемую *проверочной*. Проверочная матрица **H** позволит нам декодировать полученные вектора. Для каждой порождающей матрицы **G** размерности $k\times n$ существует проверочная матрица **H** размером $k\times\ n-k$, такая, что строки матрицы **G** ортогональны к строкам матрицы **H**. Чтобы матрица **H** удовлетворяла требованиям ортогональности $\mathbf{U}\cdot\mathbf{H}^{\mathrm{T}}=0$, ее компоненты записываются в следующем виде

$$\mathbf{H}=\left[\mathbf{I}_{\mathrm{n\text{-}k}}:\ \mathbf{P}^{\mathrm{T}}\right]. \qquad (3.5)$$

Проверочная матрица имеет следующий вид

$$\mathbf{H}^{\mathrm{T}}=\left[\begin{array}{c}\mathbf{I}_{n-k}\\ .....\\ \mathbf{P}\end{array}\right]=\left[\begin{array}{l}1\ \ 0\ \ ...\ \ 0\\ 0\ \ 1\ \ ...\ \ 0\\ .....\\ 0\ \ 0\ \ ...\ \ 1\\ p_{11}p_{12}\cdots p_{1\ n-k}\\ p_{21}p_{22}\cdots p_{2\ n-k}\\ .....\\ p_{k1}p_{k1}\cdots p_{k\ n-k}\end{array}\right].$$

Рассматривая структуру матриц **H** и **G** можно вынести следующее заключение. Обе матрицы состоят из множества линейно независимых векторов, поскольку наличие в их структуре единичной матрицы делает невозможным существование линейной комбинации строк с нулевой суммой. Учитывая сказанное выше, можно их «поменять ролями» и использовать **H** в качестве порождающей, а **G** – проверочной матрицы некоторого другого кода. Коды, связанные с таким преобразованием, называются *дуальными* друг другу. Таким образом, если исходным являлся (n,k) кодом, то дуальным ему будет $(n,n-k)$ код.

3.1.1. Построение порождающей матрицы.

Рассмотрим принципы построение порождающей матрицы G систематического линейного блочного кода n,k . На первом этапе необходимо из известных таблиц выбрать некоторый примитивный полином степени r

$$g\ x = g_r x^r + g_{r-1} x^{r-1} + ... + g_1 x + g_0 .$$

Полином $g\ x$ степени r называют порождающим. Степень полинома r должна удовлетворять условию $2^r - 1 \geq n$.

Построение порождающей матрицы рассмотрим на конкретном примере линейного кода $6,3$. В этом случае $n = 6$ и принимаем $r = 3$, $2^r - 1 = 7 > 6$.

Пусть для кода $6,3$ выбран порождающий полином $g\ x = x^3 + x^2 + 1 \Rightarrow 1101$.

Для получения порождающей матрицы G достаточно в двоичной системе счисления выполнить операции *умножения позиционных двоичных кодов* ρ длины $r+1$ (содержащих *единицу в одной позиции)* на 2^r и *деления на порождающий полином* $g\ x$. Описанные действия можно представить уравнением $\rho \cdot 2^r \Big/ g = C \oplus R/g$.

Пусть, например, $\rho = 0\ 0\ 0\ 1$. Тогда, умножив ρ на $2^r = 2^3 = 8$ в двоичном счислении, т.е. на $1\ 0\ 0\ 0$, получим двоичное слово: $0\ 0\ 0\ 1\ 0\ 0\ 0$. Разделим его на $g = 1\ 1\ 0\ 1$:

$$\frac{0\ 0\ 0\ 1\ 0\ 0\ 0}{1\ 1\ 0\ 1} \Rightarrow C = 0\ 0\ 0\ 1,\ \ R = 1\ 0\ 1.$$

Аналогично можно получить остатки от деления для всех кодов $\rho_i \Rightarrow R_i,\ \ i = 1,...,r+1$.

$$\rho = \begin{Bmatrix} 0\ 0\ 0\ 1 \\ 0\ 0\ 1\ 0 \\ 0\ 1\ 0\ 0 \\ 1\ 0\ 0\ 0 \end{Bmatrix} \Rightarrow R = \begin{Bmatrix} 1\ 0\ 1 \\ 1\ 1\ 1 \\ 0\ 1\ 1 \\ 1\ 1\ 0 \end{Bmatrix}.$$

Поскольку $2^r - 1 = 7 > n = 6$, поэтому для кода $n,k = 6,3$ целесообразно остаток нечетного веса $R = 111$ исключить из построения порождающей матрицы.

Таким образом, порождающая матрица G представляется в виде единичной матрицы I и дополняющей ее матрицы P, строки которой представляют остатки R. Если длина кодового слова n равна $2^r - 1$, то код будет иметь минимальную избыточность.

3.1.2. Контроль с помощью синдромов.

Определим понятия *синдрома*. Пусть $\mathbf{e} = e_1, e_2, ..., e_n$ - вектор ошибки или модель ошибки, внесенная каналом. Синдром вектора ошибки определяется следующим образом

$$\mathbf{S} = \mathbf{e} \cdot \mathbf{H}. \qquad (3.6)$$

При воздействии канальной помехи вектор принятого сигнала можно представить в следующем виде

$$\mathbf{Ur} = \mathbf{U} + \mathbf{e}. \qquad (3.7)$$

Используя приведенные уравнения, мы можем представить синдром вектор принятого сигнала в следующем виде

$$\mathbf{Sr} = \mathbf{Ur} \cdot \mathbf{H}^{\mathrm{T}} = \mathbf{U} \cdot \mathbf{H}^{\mathrm{T}} + \mathbf{e} \cdot \mathbf{H}^{\mathrm{T}}.$$

Для всех элементов набора кодовых слов имеем $\mathbf{U} \cdot \mathbf{H}^{\mathrm{T}} = 0$ и поэтому синдром равен

$$\mathbf{Sr} = \mathbf{e} \cdot \mathbf{H}^{\mathrm{T}} = \mathbf{S}.$$

Из сказанного следует, что по вектору принятого сигнала **Ur** можно обнаружить ошибку и установить синдром канальной помехи **S**.

Затем, используя таблицу соответствия синдромов (табл. 6.3.2), находим соответствующую модель ошибки, которая является оценкой ошибки (далее будем обозначать ее через $\hat{\mathbf{e}}$). Затем декодер прибавляет оценку $\hat{\mathbf{e}}$ к сигнальному вектору **Ur** и оценивает переданное кодовое слово **Us**

$$\mathbf{Us} = \mathbf{Ur} + \hat{\mathbf{e}} = \mathbf{U} + \mathbf{e} + \hat{\mathbf{e}} = \mathbf{U} + \mathbf{e} + \hat{\mathbf{e}} . \qquad (3.8)$$

Если правильно вычислили ошибку: $\mathbf{e} = \hat{\mathbf{e}}$, тогда оценка **Us** совпадает с переданным кодовым словом **U**. С другой стороны, если оценка ошибки неверна, декодер неверно определит переданное кодовое слово, и мы получим необнаружимую ошибку декодирования.

Пример 3.1. Пусть передано кодовое слово $U = 101110$ и принят вектор $Ur = 001110$, т.е. крайний левый бит был принят с ошибкой.

Определим синдром ошибки $\mathbf{S} = \mathbf{e}\,\mathbf{H}^{\mathrm{T}}$.

Решение.

$$\mathbf{S} = \begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1, & 1+1, & 1+1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}.$$

Определим синдром $\mathbf{S} = \mathbf{e}\,\mathbf{H}^{\mathrm{T}}$, соответствующий каждой последовательности исправимых ошибок.

Результаты соответствия синдромов приводятся в табл. 3.2.

Табл. 3.2

Модель ошибки	Синдром	Модель ошибки	Синдром
000000	000	001000	001
000001	101	010000	010
000010	011	100000	100
000100	110	010001	111

3.1.3. Декодирование с исправлением ошибок.

Итак, мы показали, что контроль с помощью синдромов, выполняемый как на искаженном кодовом слове, так и на соответствующей модели ошибки, дает один и тот же синдром. Этот момент является ключевым, поскольку мы имеем возможность не только определить ошибку, но и (поскольку существует взаимно однозначное соответствие между исправимой моделью ошибки и синдромом) исправить подобные модели ошибки.

Процедура декодирования с исправлением ошибок состоит из следующих этапов:

1. Вычисляется синдром канальной ошибки $\mathbf{S} = \mathbf{e}\ \mathbf{H}^{\mathrm{T}}$.

2. Определяются модели ошибки $\mathbf{e}$, синдром которых равен $\mathbf{S} = \mathbf{e}\ \mathbf{H}^{\mathrm{T}}$.

3. Полученный исправленный вектор, или кодовое слово, определяется как: $\mathbf{Us} = \mathbf{Ur} + \hat{\mathbf{e}}$.

Пример 3.2. Пусть передано кодовое слово $\mathrm{U} = 101110$ и принят вектор наблюдения $\mathrm{Ur} = 001110$. Нужно показать, как декодер, используя таблицу соответствия синдромов (табл. 3.2), может исправить ошибку. Синдрому $\mathbf{S} = 1\,0\,0$ соответствует вектор ошибки $\mathrm{e} = 100000$. Исправленный вектор равен следующему $\hat{\mathrm{U}}\mathrm{s} = \mathrm{Ur} + \hat{\mathrm{e}} = 0\,0\,1\,1\,1\,0 + 1\,0\,0\,0\,0\,0 = 1\,0\,1\,1\,1\,0$.

Процедура исправления ошибки дает $\hat{\mathrm{U}}\mathrm{s} = \mathrm{U}$.

3.2 Циклические коды

Линейный код n, k называется *циклическим,* если он обладает следующим свойством. Если $\mathrm{u} = \mathrm{u}_0,\mathrm{u}_1,\mathrm{u}_2,...,\mathrm{u}_{n-1}$ является кодовым словом, то код $\mathrm{u\ i} = \mathrm{u}_{n-i},\mathrm{u}_{n-i+1},...,\mathrm{u}_{n-1},\mathrm{u}_0,\mathrm{u}_1,...,\mathrm{u}_{n-i-1}$, полученный i циклическими сдвигами, является также кодовым словом.

Компоненты кодового слова $\mathrm{u} = \mathrm{u}_0,\mathrm{u}_1,\mathrm{u}_2,...,\mathrm{u}_{n-1}$ можно рассматривать как коэффициенты полинома $u\ x$:

$$u\ x = u_0 + u_1 x + u_2 x^2 + ... + u_{n-1} x^{n-1}. \qquad (3.9)$$

Циклический код (n, k) полностью задаётся *порождающим многочленом*

$$g\ x = g_0 + g_1 x + g_2 x^2 + ... + g_{r-1} x^{r-1} + g_r x^r \qquad (3.10)$$

Степени $r = n - k$. Здесь g_0 и g_r должны быть равны 1.

Порождающий полином $g\ x$ циклического кода n,k является множителем бинома $p\ x = x^n + 1,$ т.е. $x^n + 1 = g\ x \cdot h\ x$.

Полином $h(x)$, получаемый как результат деления бинома $p\ x = x^n + 1$ на порождающий многочлен $g(x)$, называется *проверочным многочленом*

$$h(x) = \frac{x^n + 1}{g(x)}. \tag{3.11}$$

Легко убедиться, что $p\ x = x^7 + 1 = x+1 \cdot x^3 + x + 1 \cdot x^3 + x^2 + 1$, где все множители - неприводимые над полем $GF(2)$ многочлены. Следовательно, порождающими многочленами $g(x)$ могут быть либо множители приведенного бинома, либо их произведение.

Например, многочлен $g\ x = x^3 + x^2 + 1$ или $g\ x = x^3 + x\ + 1$ может быть порождающим для циклического кода $(7,4)$. Многочлен вида $g(x) = x+1\ \ x^3 + x + 1\ = x^4 + x^3 + x^2 + 1$ может быть порождающим для циклического кода $(7,3)$.

Учитывая сказанное, можно порождающий многочлен и проверочный многочлен «поменять ролями». Коды, связанные с таким преобразованием, называются *дуальными* друг другу. Таким образом, если исходным являлся (n,k) код, то дуальным ему будет $(n,n-k)$ код.

3.2.1. Кодирование циклических кодов.

Пусть $\mathrm{m}(x) = m_k x^k + m_{k-1} x^{k-1} + \ldots + m_1 x + m_0$ - информационный полином и $g\ x = g_r x^r + g_{r-1} x^{r-1} + \ldots + g_1 x^1 + g_0$ - порождающий полином для циклического кода n,k . Тогда полином кодового слова

$$u\ x = u_n x^n + u_{n-1} x^{n-1} + \ldots + u_1 x + u_0$$

можно определить произведением:

$$u(x) = m\ x \cdot g\ x =$$

$$= m_0 + m_1 x + m_2 x^2 + \ldots + m_{k-1} x^{k-1} \cdot g_0 + g_1 x + g_2 x^2 + \ldots + g_r x^r =$$

$$= m_k g_r x^{m+r} + m_k g_{r-1} + m_{k-1} g_r\ x^{k+r-1} + \ldots + (m_1 g_0 + m_0 b_1) x + m_0 g_0 =$$

$$= \sum_{i=0}^{k+r} u_i \cdot x^i . \qquad (3.12)$$

где

$$\mathrm{u} = \mathrm{m} * \mathrm{g} \Rightarrow u_i = \sum_c m_c \cdot g_{i-c} .$$

В данном выражении для коэффициентов u_i, определяющихся как свертка последовательностей коэффициентов сомножителей значения m_c и g_{i-c} считаются равными нулю, если их подстрочный индекс не принадлежит соответствующим допустимым областям: $c \in \{0,1,\ldots,k\}$, $i-c \in \{0,1,\ldots,r\}$.

Принципы построения циклических кодеров. Как отмечалось выше компоненты соответствующего кодового вектора $u = u_0, u_1, ..., u_{n-1}$ могут быть найдены, как свертка информационной последовательности $m = m_0, m_1, \ldots, m_{k-1}$ с последовательностью коэффициентов порождающего полинома $g = g_0, g_1, \ldots, g_r$. Данная операция реализуется с помощью фильтра с конечным импульсным откликом (КИО–фильтра), структура которого представлена на рис. 3.1.

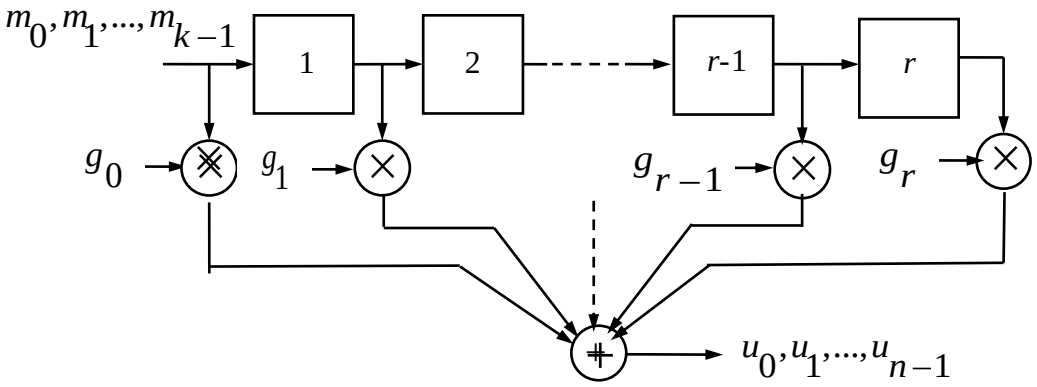


Рис. 3.1.Кодирующее устройство циклических кодов

Отметим, что построение кодового полинома путем перемножения информационного и порождающего полиномов приводит к не систематическому кодовому слову, в котором отсутствует явное разделение информационных и проверочных символов.

3.2.2. Декодирование циклических кодов.

Поскольку любой полином циклического кода (и только он) делится на порождающий полином $g\ x$, т.е. дает нулевой остаток от деления, то декодирование циклических кодов может быть организовано с ориентацией на данное свойство. Пусть

$$y(x) = y_{n-1}x^{n-1} + y_{n-2}x^{n-2} + \ldots + y_1 x + y_0 \qquad (3.13)$$

- *полином наблюдения*, коэффициентами которого являются элементы вектора наблюдения $y = y_0, y_1, \ldots, y_{n-1}$.

Если принятое сообщение принято с ошибками, то полином наблюдения (3.13) можно представить в виде

$$y(x) = u(x) + e(x) = m(x) \cdot g(x) + e(x), \qquad (3.14)$$

где вектор наблюдения $\mathbf{y} = \mathbf{u} + \mathbf{e}$ и $\mathbf{e} = e_1, e_2, \ldots, e_n$ - вектор ошибок.

Назовем *синдромным полиномом* $s(x)$ (или *полиномом синдрома*) остаток от деления полинома наблюдения $y(x)$ на $g(x)$.

Из (3.14) следует, что синдром $s(x)$ определяется только полиномом $e(x)$. Поскольку $\deg s(x) < n - k$, то возможны $2^{n-k} = 2^r$ различных полиномов $s(x)$, тогда как число векторов ошибки - 2^n.

Декодер производит обнаружение ошибки и исправление ошибки. При положительном результате проверки синдром $s = 0$. Если принятый вектор $y = y_0, y_1, \ldots, y_{n-1}$ содержит ошибки, которые можно исправить, то синдром $s = s_0, s_1, \ldots, s_r$ имеет определенное ненулевое значение, что позволяет отметить конкретную модель ошибки. Соответствующую модель ошибки, которая является оценкой ошибки, будем обозначать ее через $\hat{e}$. Декодер прибавляет модель ошибки $\hat{e}$ к принятому вектору y и оценивает переданное кодовое слово $\hat{u}$:

$$\hat{u} = y + \hat{e} = u + e + \hat{e} = u + \left(e + \hat{e} \right) \qquad (3.15)$$

Если правильно вычислили ошибку: $e = \hat{e}$, тогда оценка $\hat{u}$ совпадает с переданным кодовым словом U. С другой стороны, если оценка ошибки неверна, декодер неверно определит переданное кодовое слово, и мы получим *необнаружимую ошибку декодирования.*

Прямое вычисление синдрома связано с нахождением остатка от деления полиномов. Пусть даны два некоторых полинома $y(x)$ и $g(x)$

$$y(x) = y_0 + y_1 \cdot x + y_2 \cdot x^2 + \ldots + y_n \cdot x^n$$

$$g(x) = g_0 + g_1 \cdot x + g_2 \cdot x^2 + \ldots + g_r \cdot x^r$$

причем $n \geq r$.

Схема деления, приведенная на рис. 3.2, выполняет полиномиальное деление $y\ x$ на $g\ x$, определяя частное $q\ x$ и остаток $p\ x$.

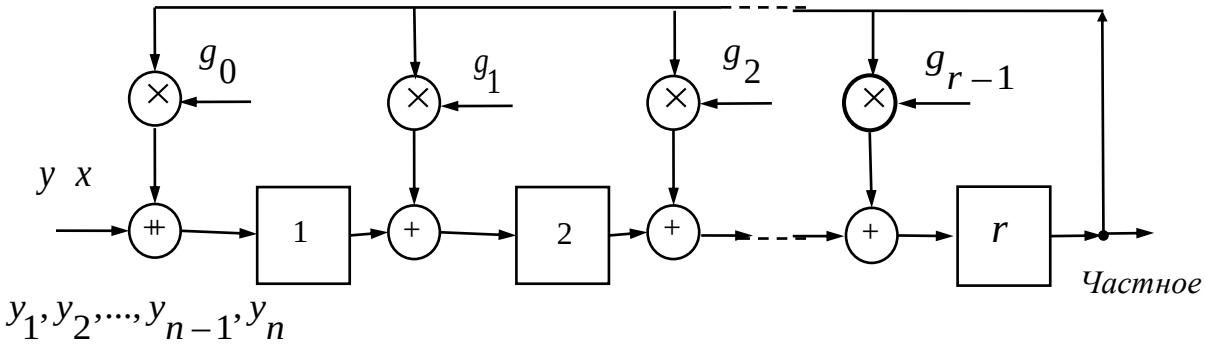


Рис. 3.2. Схема деления двух полиномов

Мы рассмотрели случай вычисление полинома синдрома $s(x)$ посредством деления полинома наблюдения $y(x)$ на порождающий полином $g(x)$. Однако, как отмечалось ранее, *можно выполнить и операцию умножения на проверочный полином* $h(x)$. Данная операция может быть выполнена с помощью фильтра, представленного на рис. 3.1.

Таким образом, синдромное декодирование циклических кодов и исправление ошибок может быть реализовано в 3 этапа:

1. Производится вычисление полинома синдрома $s(x)$ посредством деления полинома наблюдения $y(x)$ на порождающий полином $g(x)$;

2. По вычисленному полиному $s(x)$ находится возможный полином ошибки $\hat{e}\ x$;

6. Вычитая только что найденную оценку $\hat{e}\ x$ из полинома наблюдения $y(x)$, получаем декодированное значение кодового полинома $u(x)$.

Если вектор синдрома нулевой, считается, что принятый вектор является правильным кодовым словом. Если синдром отличен от нуля, значит, обнаружена ошибка и принятый вектор - это искаженное кодовое слово; данная ошибка исправляется путем прибавления к принятому кодовому слову вектора ошибки (указанной синдромом).

3.3 Сверточные коды

Наряду с блоковыми кодами существует большой и эффективный класс древовидных или решетчатых кодов, среди которых особый интерес представляют *сверточные* коды. Отличительной чертой сверточных кодов (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.

Причины чрезвычайной популярности сверточных кодов заложены в их рекуррентной форме, позволяющей реализовать значительно более практичные по сравнению с другими процедуры декодирования с исправлением ошибок. При сверточном (древовидном, в общем случае) кодировании используется отличный от предыдущего метод отображения информационных данных в кодовые символы. Суть метода может быть наглядно пояснена на примере использования «скользящего окна» шириной m информационных символов. Указанное окно вырезает в потоке данных m–символьный информационный блок, который отображается в n–символьную кодовую группу, передаваемую в течение интервала, равного одному информационному символу, и называемую иногда *кадром кодовых символов*. Затем окно сдвигается на один информационный символ вправо и обновленный m–символьный информационный блок отображается в следующую n–символьную кодовую группу и т.д. Теперь вместо последовательности индивидуальных кодовых слов, отвечающих не пересекающимся во времени блокам данных, имеет место последовательный поток кодовых символов, в котором каждая группа из n кодовых символов отвечает за защиту не только текущего символа данных, но и $m-1$ предшествующих.

Иллюстрацией описанного выше алгоритма служит рис. 3.3, на котором изображено формирование сверточного кода с m = 3 и n = 3.

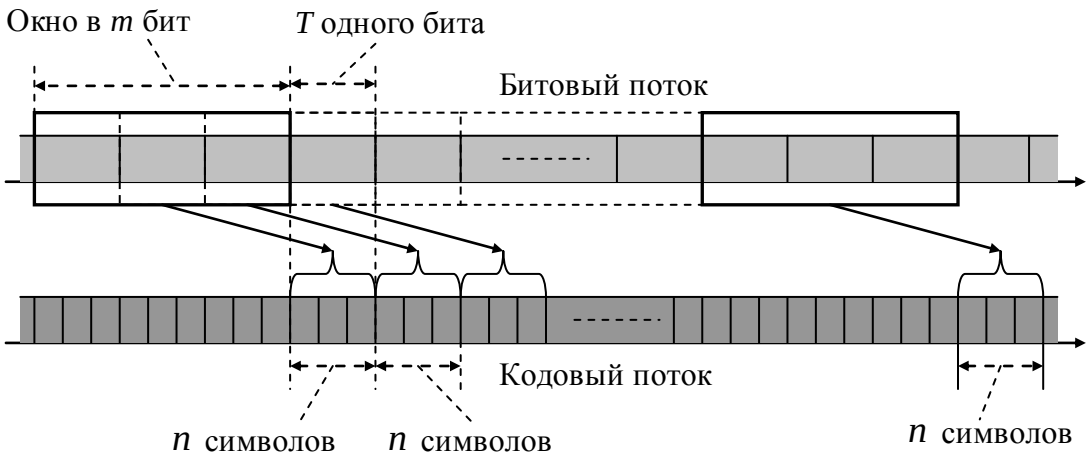


Рис. 3.3. Формирование сверточного кода.

Параметр m сверточного кода, определяющий число информационных символов, за которое отвечает текущая n–символьная группа, получил название *длины кодового ограничения*.

3.3.1 Сверточный кодер.

Как правило, аппаратная реализация описанного алгоритма кодирования осуществляется на основе регистра сдвига. В случае формирования двоичного

кода структура кодера, изображенная на рис. 3.4, содержит $m-1$ двоичных ячеек памяти (элементов задержки), единственное предназначение которых состоит в запоминании $m-1$ предшествующих битов данных с тем, чтобы совместно с текущим битом образовать m битовый сегмент, кодируемый в текущую n –символьную кодовую группу. 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.

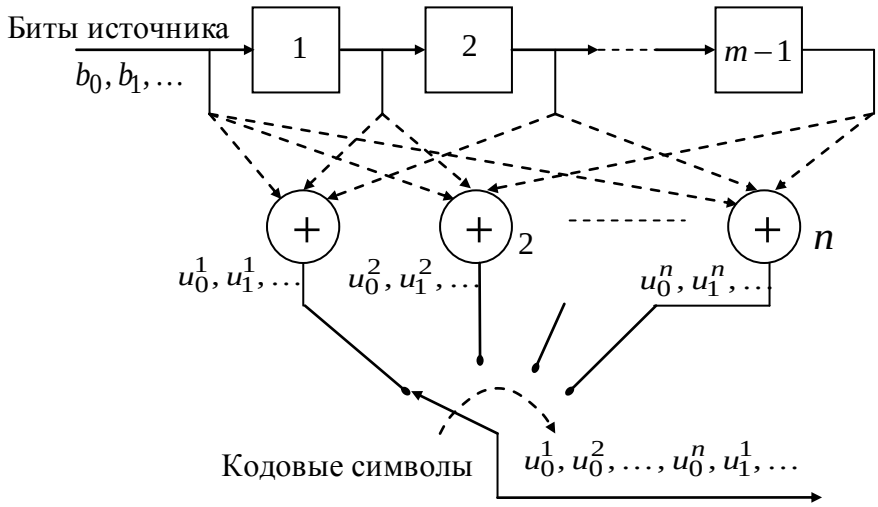


Рис. 3.4. Структура сверточного кодера.

Сама операция кодирования осуществляется посредством n сумматоров по модулю 2. Ключ s производит последовательное соединение выходов всех сумматоров таким образом, чтобы за длительность одного бита данных сформировать n –символьный кодовый кадр. После кодирования текущих m бит содержимое регистра сдвигается вправо и на его вход поступает новый бит, а «старейший» выводится и больше не участвует в формировании кодовых символов.

При поступлении на вход текущего бита источника b_i на выходах сумматоров параллельно появляются n кодовых символов $u_i^1, u_i^2, \ldots, u_i^n$. 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 $u_0^1, u_0^2, \ldots, u_0^n, u_1^1, u_1^2, \ldots, u_1^n, \ldots$. Отметим, что если на вход кодера поступает «нулевой» поток бит, то и на выходе будет также «нулевой» поток бит.

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 $k >> 1$, а $m << k$, то $R = 1/n$. Очевидно, что рассмотренной версии сверточных кодов отвечают скорости из множества $1/2, 1/3, 1/4, \ldots$.

Из рассмотренного очевидным образом следует, что любой сверточный код может быть полностью задан длиной кодового ограничения и схемой соединения сумматоров с ячейками памяти регистра сдвига. Как правило, эти соединения описываются порождающими полиномами $g_i(z), i = 1, 2, \ldots, n$. Из определения порождающего полинома следует, что его максимальная степень всегда на единицу меньше длины кодового ограничения. Другими словами, длина регистра сдвига (т.е. число ячеек памяти) совпадает с максимальной степенью порождающего полинома.

Степени $0, 1, \ldots, m-1$ формальной переменной z в порождающих полиномах соответствуют порядковому номеру ячейки памяти регистра сдвига, а коэффициенты g_{ij} в полиноме $g_i(z)$ при z^j принимают (для двоичных кодов) значения, равные единице только в том случае, если есть соединение между i–м сумматором по модулю два и j–й ячейкой регистра сдвига.

Описание сверточного кода может быть осуществлено и в полиномиальной форме. В этом случае входному информационному потоку сопоставляется многочлен $a(z) = a_0 + a_1 z + a_2 z^2 + \ldots$, коэффициенты a_i которого определяются значениями двоичных символов информационных данных.

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).

Сигнал на выходе фильтра 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

$$u_i^l = \sum_{t=0}^{\nu_C - 1} b_{i-t} g_t^l, \ i = 0,1,\ldots; \ l = 1,2,\ldots,n, \qquad (3.16)$$

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 $i < 0$.

Подходящим инструментом для описания дискретных систем является z–преобразование. В z–области свертке соответствует произведение z–преобразований, так что соотношение (3.16) может быть представлено эквивалентной формой

$$u^l(z) = \sum_{i=0}^{\infty} u_i^l z^i = b(z) g_l(z),\ l = 1,2,\ldots,n, \qquad (3.17)$$

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

$$g_l(z) = g_0^l + g_1^l z + \ldots + g_{\nu_c - 1}^l z^{\nu_c - 1},\ l = 1,2,\ldots,n \qquad (3.18)$$

- передаточная функция l–го КИО–фильтра, называемая также l–м *порождающим полиномом* сверточного кода.

Нижеприведенный пример поможет лучше понять принцип сверточного кодирования.

Пример 3.3.1. Рассмотрим устройство формирования двоичного сверточного кода, изображенное на рис. 3.6. Логика сверточного кодера описывается порождающими полиномами вида $g_1(z) = 1 + z^2$, $g_2(z) = 1 + z + z^2$.

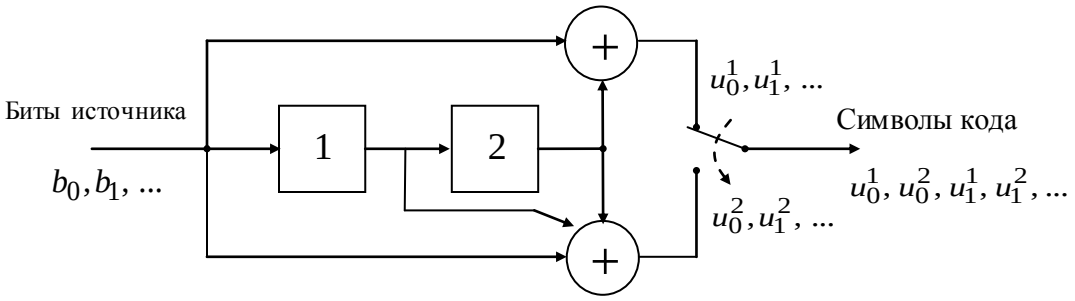


Рис. 3.6. Сверточный кодер.

Кодер содержит двухразрядный регистр сдвига и, следовательно, отвечает длине кодового ограничения $m = 3$. Скорость формируемого им кода $R = 1/2$, поскольку каждый информационный бит отображается двумя кодовыми символами.

Входные биты $b_0, b_1, \ldots$ порождают два потока кодовых символов $u_0^1, u_1^1, \ldots$ и $u_0^2, u_1^2, \ldots$, которые затем мультиплексируются таким образом, что u_i^1 и u_i^2 занимают соответственно четные и нечетные позиции в общем кодовом потоке. Пусть битовый поток определен последовательностью $b_i = 010100\ldots$. Приведенная таблица 3.1 состояний иллюстрирует процесс кодирования.

Таблица 3.2.1

Такт i	b_i	Регистр 1 2	u_i^1	u_i^2
0	0	0 0	0	0
1	1	0 0	1	1
2	0	1 0	0	1
3	1	0 1	0	0
4	0	1 0	0	1
5	0	0 1	1	1
6	0	0 0	0	0
7	0	0 0	0	0

Как можно видеть, битовый поток $b_i = 010100...$ порождает последовательности $u_i^1 = 01000100...$ и $u_i^2 = 01101100...$, которые затем мультиплексируются в кодовый поток $u_i^1\, u_i^2 = 0011010001110000...$.

Следует отметить, что построение хорошего сверточного кода непосредственно связано с нахождением соответствующих порождающих полиномов. Вместе с тем, общая тенденция к улучшению характеристик кода непосредственно связана с увеличением памяти кода $m-1$, а значит и максимальной степени порождающих полиномов. Однако, в отличие от циклических кодов построение порождающих полиномов сверточных кодов не находит адекватной теоретической поддержки и в большей степени ориентировано на компьютерный поиск.

3.3.2 Диаграмма состояний и решетка сверточного кода.

Любой двоичный сверточный кодер может трактоваться как *автомат*, или *устройство с конечным числом внутренних состояний*, под которыми понимается текущее содержимое регистра сдвига. Регистр сдвига сверточного кодера обладает 2^m возможными состояниями. В зависимости от того, какое из двух значений принимает входной бит информационного потока, кодер при поступлении тактового импульса переходит в одно из двух возможных состояний. С другой стороны, в данный момент времени кодер может оказаться в некотором конкретном состоянии только в результате перехода в него из одного из двух предшествующих возможных состояний. Подобный двузначный характер переходов между состояниями объясняется поступлением на вход кодера только одного бита информации за один такт работы.

Данную закономерность отражает *диаграмма состояний* сверточного кода, описывающая все состояния и возможные переходы из текущего состояния в последующее в виде узлов и ветвей со стрелками. Построение подобное диаграмм поясняется следующим примером.

Пример 3.2.2. Обратимся к кодеру из *примера 3.2.1*. Данное устройство может находиться в одном из следующих состояний: 00, 10, 11 и 01 (первый символ отражает содержимое крайней левой ячейки памяти). Очевидно, что из состояния 00 кодер может либо остаться в состоянии 00 (при входном бите, равном 0), либо перейти в 10 (при подаче на вход единицы).

Аналогичным образом представлены и остальные переходы в диаграмме состояний на рис. 3.6, в которой узлы изображены в виде кружков, внутреннее содержимое которых отражает текущее состояние регистра сдвига. Сплошные стрелки соответствуют переходам, которые обусловлены поступлением нулевого входного бита, тогда как пунктирные - битом, равным единице. Легко заметить, что каждому узлу отвечают два исходящих и столько же входящих переходов. Что же касается кодовых символов, то их значения приведены над соответствующими ветвями диаграммы. Так например, кодер, находящийся в состоянии 00, вырабатывает кодовую комбинацию 11 при поступлении на вход единичного бита. Поэтому ветвь, отражающая переход $00 \to 10$, помечена символами 11 и т.д.

Диаграмму состояний можно изобразить в виде решетки, представленной на рис. 3.7. Она является аналогичной графической иллюстрацией состояния кодера для двух последовательных временных интервалов. Однако эта версия диаграммы состояний указывает путь к построению модели, учитывающей всю

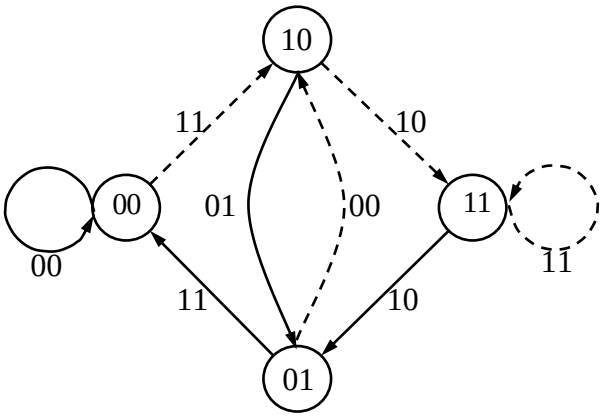


Рис. 3.6. Диаграмма состояний.

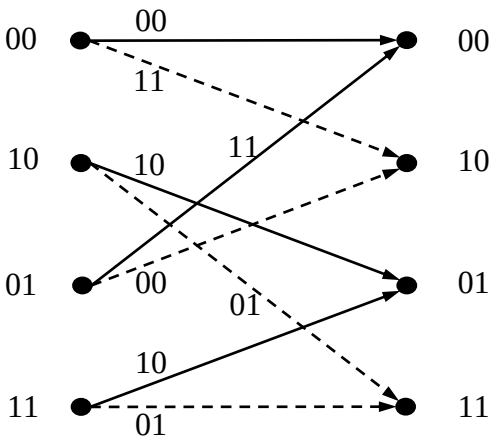


Рис. 3.7. Решетка кодера.

динамику процесса кодирования. Как и ранее, текущее состояние изменяется на последующее в зависимости от значения поступающего входного бита.

Пусть кодер находится в начальном состоянии 00. После первого такта работы кодер переходит либо в состояние 00, либо 10. Если его состояние остается прежним, т.е. 00, то в следующий такт он снова переходит в одно из двух ранее указанных состояний. Если же после первого такта работы кодер находится в состоянии 10, то возможен переход либо в состояние 01, либо в 11. При поступлении третьего такта алгоритм работы кодер, находившегося в состоянии 00 или 10, остается прежним. Однако из состояния 01 он переходит либо в 00, либо 10, а из состояния 11 - либо в 01, либо в 11. После третьего такта работа кодера становится устойчивой, повторяя ранее описанные

переходы, пока не прекратится поток входных бит. При отсутствии входной информации кодеру потребуется еще $m-1$ тактов для обнуления, в течение которых будет продолжено формирование кодовых символов.

Модель поведения кодера, построенная на основе вышеприведенных рассуждений, представлена на рис. 3.8 и называется *решетчатой диаграммой*, или просто *решеткой*. Любой путь по решетке представляет собой ничто иное, как *кодовое слово* сверточного кода.

Для иллюстрации сказанного обратимся к рис. 3.8, на котором представлена решетчатая диаграмма кода из примера 3.3.1.

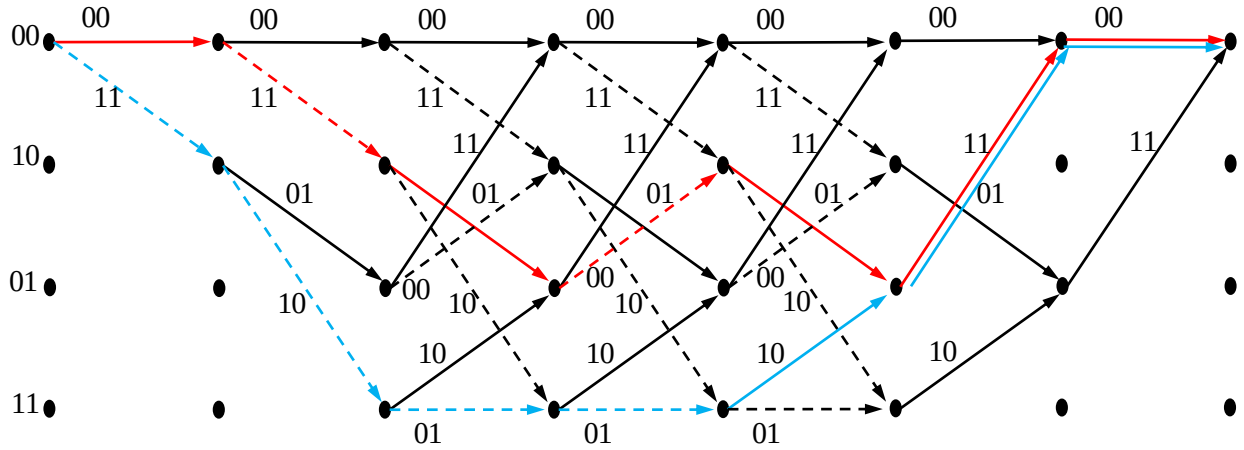


Рис. 3.7. Решетчатая диаграмма кодера.

Каждая последовательность - кандидат входных битов данных выбирает определенный путь по решетчатой диаграмме, который можно отследить, например, для последовательности $b_i = 010100\ldots$. Ее первый бит равен 0, что направляет кодер из узла (00) в узел (00), выдавая выходные кодовые символы 00. Второй бит, равен 1, переводит кодер из узла (00) в (10), генерируя кодовые символы 11. Третий бит 0 изменяет состояние кодера с (10) на (01), выдавая на выход кодовые символы 01. Четвертый бит 1 изменяет состояние кодера с (01) на (10), выдавая на выход кодовые символы 00 и т.д. Красной (жирной) чертой выделен результирующий путь, отвечающий кодовому слову $y = 00\ 11\ 01\ 00\ 01\ 11\ 00\ 00\ldots$.Если входная последовательность имеет вид $b_i = 111100\ldots$, то путь по решетчатой диаграмме, выделенный синей чертой, образует кодовое слово $y = 11\ 10\ 01\ 01\ 10\ 11\ 00\ 00\ldots$.

Для определения исправляющей способности сверточного кода необходимо установить минимальное хэммингово расстояние между различными путями решетки. В случае сверточных кодов минимальное расстояние часто называют *свободным расстоянием*. Данное название объясняется тем фактом, что вследствие линейности кода свободное рас стояние есть ничто иное, как минимальный вес ненулевого пути, т.е. пути, не совпадающего с верхней траекторией решетчатой диаграммы. Свободное

расстояние d_f гарантирует исправление любых $\left[\frac{d_f - 1}{2}\right]$ ошибок в символах, однако, как правило, исправляются также и многие образцы с большим числом ошибок.

В отличие от блоковых свободное расстояние для сверточных кодов может быть найдено достаточно просто с помощью диаграммы состояний, что является предметом рассмотрения следующего параграфа.

3.3.3 Передаточная функция сверточного кода.

Анализируя решетчатую диаграмму сверточного кода, можно легко установить, что для нахождения минимального расстояния (минимального веса) следует рассматривать только те пути, которые отходят в некоторой точке от нулевого, а затем вновь с ним сливаются и более не расходятся. Поскольку движение по нулевому пути не вносит вклада в значения веса рассматриваемого пути, то допустимо пренебречь всеми ребрами решетки, соединяющими нулевые состояния кодера. Принимая во внимание последнее, определим минимальный вес ненулевого участка пути, т.е. участка от момента выхода из нулевого состояния до слияния с ним через некоторое количество шагов.

Пометим каждую ветвь диаграммы состояния степенью формальной переменной D, величина которой определяется весом ветви. Так для примера 3.2.1 вес ветви, ведущей из состояния 00 в состояние 10, равен 2 (соответствующая кодовая комбинация 11), а значит, ей присваивается метка D^2. Ветвь из состояния 10 в 01 обозначается через D (кодовая комбинация 01), а ветвь из 01 в 10 - $D^0 = 1$ (00) и т.д. Диаграмма состояний с присвоенными подобным образом отметками изображена на рис. 3.9, причем ветвь, замкнутая

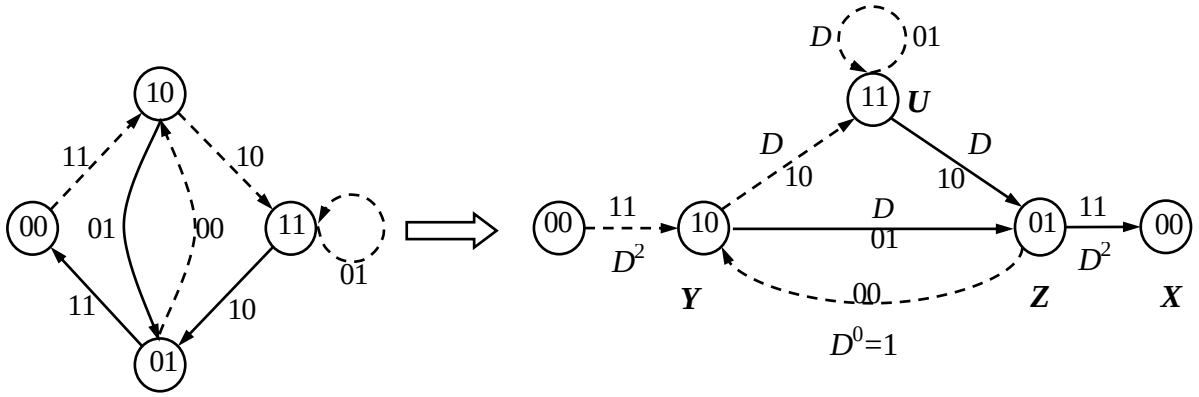


Рис. 3.9.

на состоянии 00, как указывалось ранее, исключается из рассмотрения, а состояние 00 разделяется на два: начальное (отвечающее моменту отклонения от нулевого пути) и конечное (связанное с моментом слияния).

Если теперь, двигаясь по некоторому пути, перемножать введенные метки всех его ветвей, то степень D произведения в точности будет соответствовать

весу пути. Если для некоторого узла диаграммы состояний существует два (и более) сходящихся путей с весами w и v соответственно, естественно формально записать результат как $D^w + D^v$, не опасаясь неопределенности.

Поскольку движение по ветвям диаграммы состояний представляет собой рекуррентный пошаговый процесс, выразим вес пути на текущем шаге как его вес на предшествующем шаге и вес ветви. Обозначим вес пути при текущем состоянии 10, 01 и 11 через Y, Z, U соответственно, а окончательный вес пути в точке слияния через X. Начальное значение веса пути в точке отклонения положим равным нулю $(D^0 = 1)$. С учетом введенных обозначений можно записать следующие линейные уравнения

$$X = D^2 Z, \qquad Y = D^2 + Z, \qquad Z = DY + DU, \qquad U = DY + DU\ .$$

Решение этой системы уравнений приводит к следующему результату. Поскольку

$$U = DY + DU = D(D^2 + Z) + DU \qquad \text{или} \qquad U(1 - D) = D^3 + DZ,$$

откуда
$$U = \frac{D^3 + DZ}{1 - D}.$$

С другой стороны

$$Z = DY + DU = D^3 + DZ + \frac{D(D^3 + DZ)}{1 - D} = \frac{D^3 + DZ}{1 - D},$$

а значит
$$Z = \frac{D^3}{1 - 2D}.$$

Следовательно
$$X = T(D) = D^2 Z = \frac{D^5}{1 - 2D}.$$

Функцию $T(D)$ называют *передаточной функцией* сверточного кода. Она содержит информацию о весах всех ненулевых путей решетчатой диаграммы, начинающихся с нулевого состояния. Используя результат деления вида $1/(1 - x) = 1 + x + x^2 + \ldots$, выражение для $T(D)$ преобразуется к виду

$$T(D) = D^5 + 2D^6 + 4D^7 + \ldots,$$

означающему, что рассматриваемый код имеет минимальное расстояние $d = 5$: единственный путь из нулевого состояния веса 5, два пути веса 6 и т.д.

Более того, передаточную функцию можно преобразовать таким образом, что ее новый вариант позволит определить число единиц во входной информационной последовательности и длину пути. Модернизация заключается в дополнительной маркировке ветвей диаграммы состояний формальными переменными N, степень которой равна нулю при входном нулевом бите и единице при подаче на вход единичного бита, и L, степень которой всегда равна единице. С учетом вновь введенных обозначений диаграмма состояний принимает вид, представленный на рис. 3.10.

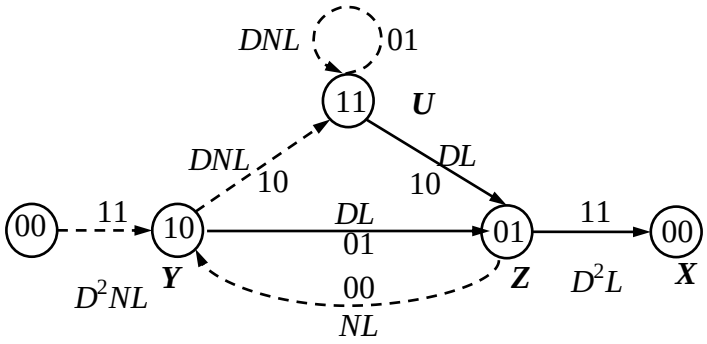


Рис. 3.10.

Составив систему линейных уравнений с учетом дополнительных обозначений

$$X = D^2LZ, \qquad Y = D^2NL + NLZ, \qquad Z = DLY + DLU, \qquad U = DNLY + DNLU .$$

и разрешив ее, получим передаточную функцию $T(D,N,L)$, в которой степени при D, N, L определяют соответственно вес, число единиц во входной последовательности и длину пути в решетчатой диаграмме. Для рассмотренного выше примера 3.2.1 получаем

$$T(D,N,L) = \frac{D^5NL^3}{1-DNL(1+L)} = D^5NL^3 + D^6N^2L^4 + D^6N^2L^5 + \ldots,$$

откуда очевидным образом следует, что в сверточном коде имеется один путь длины 3 и веса 5, соответствующий кодированию информационной последовательности с одним ненулевым битом, путь длиной 4 веса 6, кодирующий последовательность с двумя единичными битами и т.п.

Ясно, что для более сложных сверточных кодов, чем рассмотренный в примере 3.1, нахождение передаточной функции в ручную предполагает выполнение достаточно объемной и утомительной работы, однако привлечение вычислительных средств значительно облегчает решение подобной задачи.

3.3.4 Алгоритм декодирования Витерби.

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). Следует подчеркнуть, что алгоритм Витерби является оптимальным, т.е. максимально правдоподобным, и заключается в нахождении кодового слова, наиболее близкого к принятому наблюдению.

Рассмотрим его применение для двоичного симметричного канала, когда наблюдение $\mathbf{Y}$ декодируется в такое кодовое слово, которое находится на наименьшем расстоянии Хэмминга относительно принятого $\mathbf{Y}$.

Основная идея декодирования по Витерби заключается в поэтапном сравнении всех путей решетчатой диаграммы (которые в точности представляют собой кодовые слова) с наблюдением $\mathbf{Y}$ и отбрасывании тех из них, которые находятся на большем расстоянии от $\mathbf{Y}$, 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 $\mathbf{Y}$. 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. Сравниваются 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 3.2.3. Рассмотрим декодирование наблюдения из *примера 3.2.2* $y = 00\ 11\ 01\ 00\ 01\ 11\ 00\ 00...$. Допустим, что при передаче сообщения произошла ошибка в 4-ой группе бит и принятым оказалось сообщение $y = 00\ 11\ 01_{*}11_{*}01\ 11\ 00\ 00...$

Рис. 3.11 иллюстрирует процесс декодирования с указанием метрик узлов, расположенных непосредственно около них, и кадров, содержащих пары принятых символов наблюдения на текущем шаге.

Первые два этапа процедуры декодирования являются тривиальными. На первом шаге возможны два перехода, описываемые ветвями, которые выходят из состояния 00 и заканчиваются соответственно в состоянии 00 и 10 с метрикой, равной 0 и 2. (см. рис.3.11.1).

На втором этапе метрика пути, приходящего в состояние 00, становится равным 2, так как расстояние между ветвью $(00 \to 00)$ и наблюдаемой 2–х символьной группой **Y** (в текущий момент 11) равно 2. 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 2 в состояние 11 отличается от наблюдения **Y** в одной позиции (см. рис.3.11.2).

Критическим для понимания алгоритма является третий этап данного примера (см. рис.3.11.3). Рассмотрим состояние 00. Достижение его возможно двумя путями: из состояния 00 или 01. Значение накопленной метрики первого пути составит 3, поскольку данный путь начинается из состояния 00 с метрикой 2, а расстояние Хэмминга между ветвью (00→00) и 2–х символьной группой наблюдения **Y** (01) равно 1. С другой стороны, путь, приходящий из состояния 01, характеризуется аккумулированной метрикой, равной 4, так как метрика предшествующего состояния 01 равнялась 3, а расстояние Хэмминга между переходом (01→00) и наблюдением (01) – 1. Результирующие накопленные метрики состояния 00 обозначим (3/4).

Далее рассмотрим состояние 10. Достижение его возможно двумя путями: из состояния 00 или 01. Значение накопленной метрики первого пути составит 3, поскольку данный путь начинается из состояния 00 с метрикой 2, а расстояние Хэмминга между ветвью (00→10) и 2–х символьной группой наблюдения **Y** (01) равно 1. С другой стороны, путь, приходящий из состояния

01, характеризуется аккумулированной метрикой, равной 4, так как метрика предшествующего состояния 01 равнялась 3, а расстояние Хэмминга между переходом (01→10) и наблюдением **Y** (01) – 1. Результирующие накопленные метрики состояния 10 обоих путей будут равны (3/4).

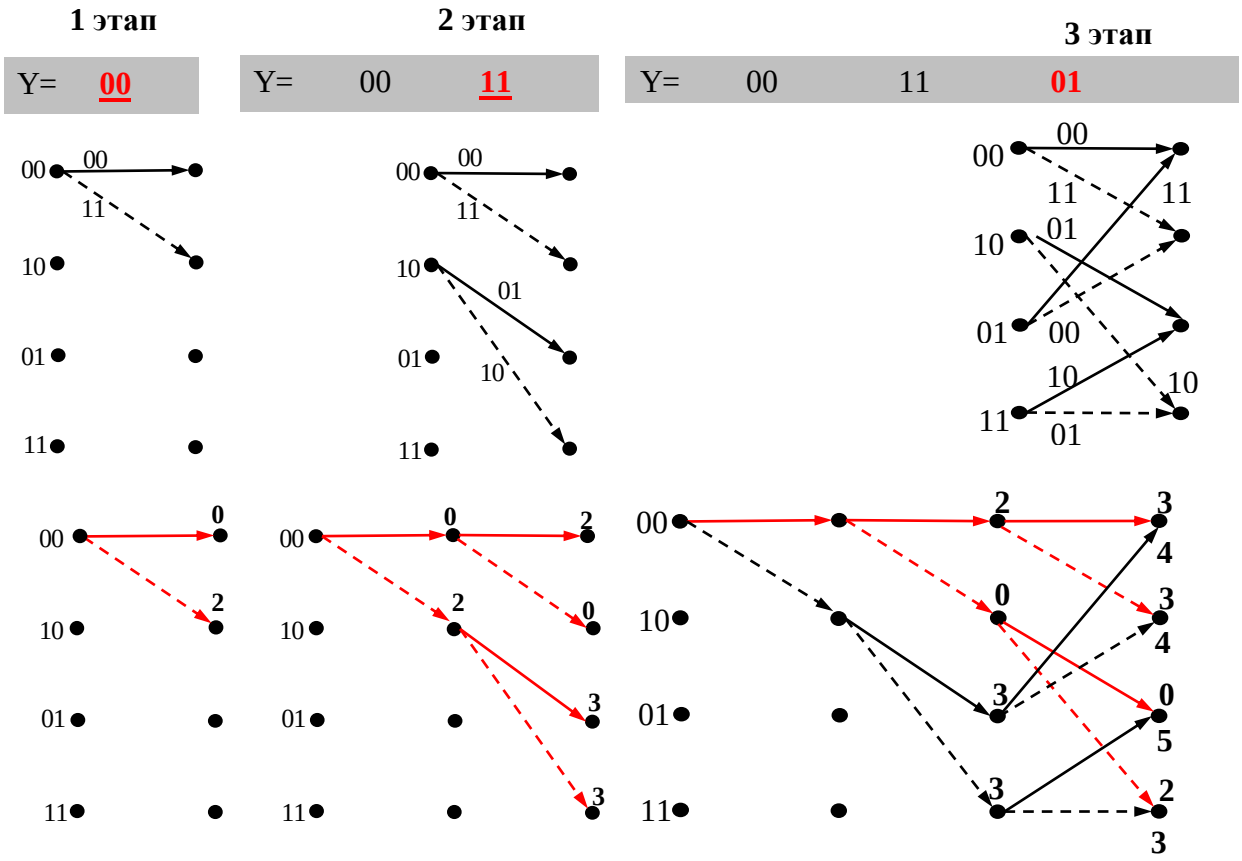


Рис. 3.11.1. Рис. 3.11.2. Рис. 3.11.3.

Дальнейшее рассмотрение состояний 01 и 11 приведет к результирующим накопленным метрикам (0/5) и (2/3).

Декодирование по минимуму расстояния предполагает нахождение кодового слова, ближайшего к наблюдению, значит, совершенно оправдано отбрасывание того пути, который в большей степени отличается от **Y**, чем альтернативный. Следовательно, ветви, дающие большие значения результирующих накопленных метрик (0/5) и (2/3), на предшествующем этапе, отбрасываются (помечены черными стрелками на рис. 3.11.3), и сохраняется только пути (помечены красными стрелками на рис. 3.11.3), которые являются *выжившими путями*.

Этапы декодирования 4, 5, 6 и 7 изображенные на рис. 3.11.4–3.11.7, абсолютно аналогичны третьему и не нуждаются в дополнительных комментариях.

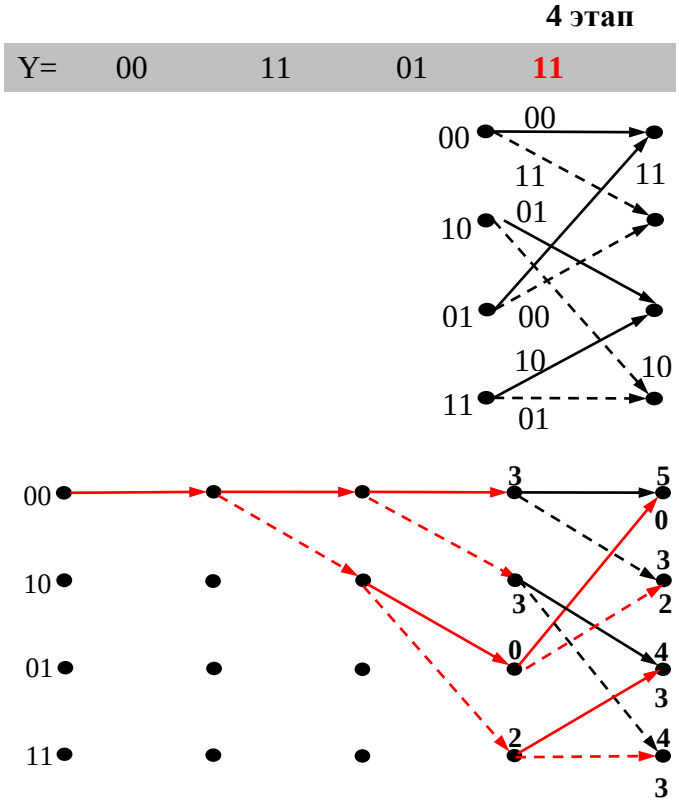


Рис. 3.11.4.

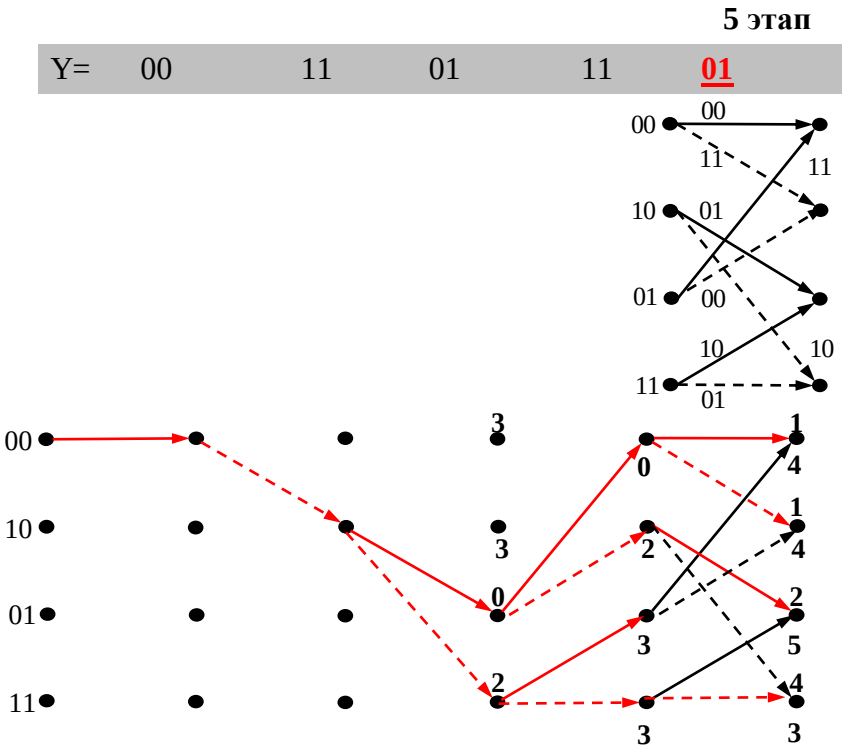


Рис. 3.11.5

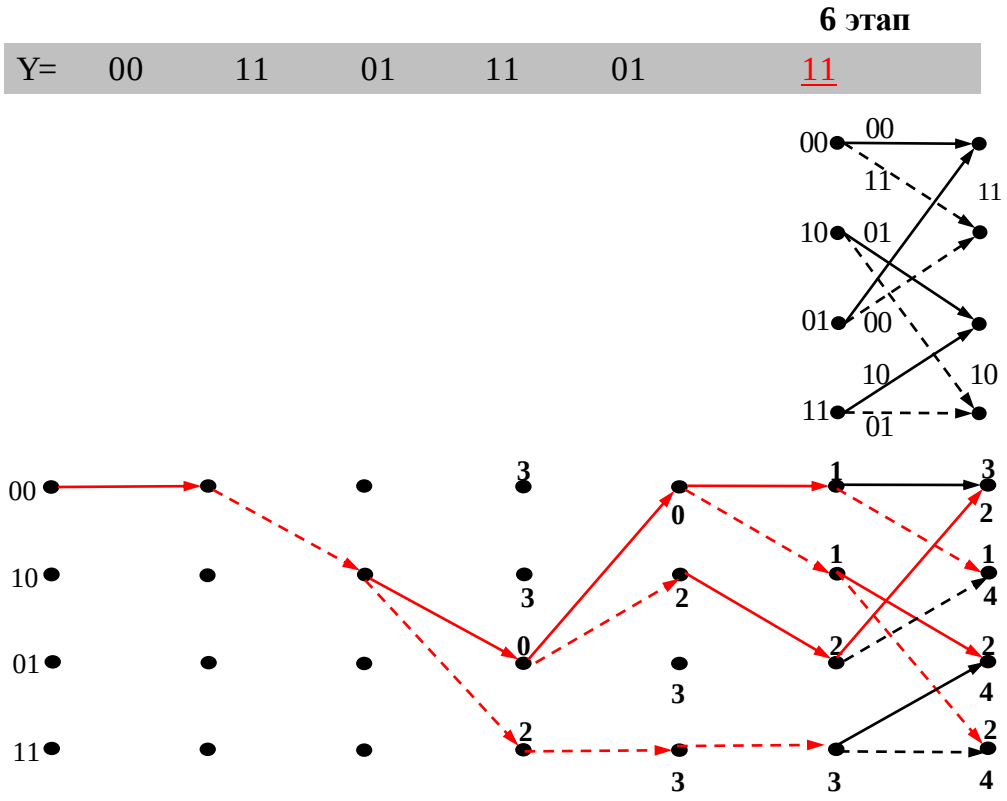


Рис. 3.11.7.

7 этап

Y=	00	11	01	11	01	11	00

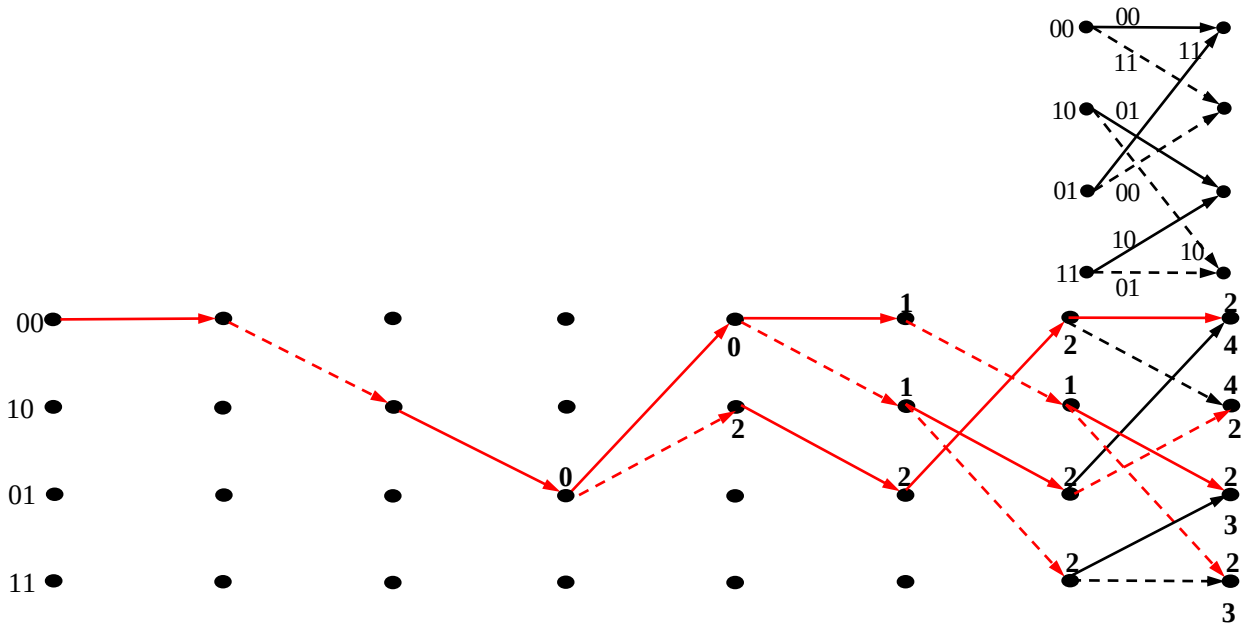


Рис. 3.11.7.

Особенностью 8-го этапа, изображенного на рис. 3.11.8 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 3, точно так же, как и пути, сходящиеся в состоянии 11.

8 этап

Y=	00	11	01	11	01	11	00	00

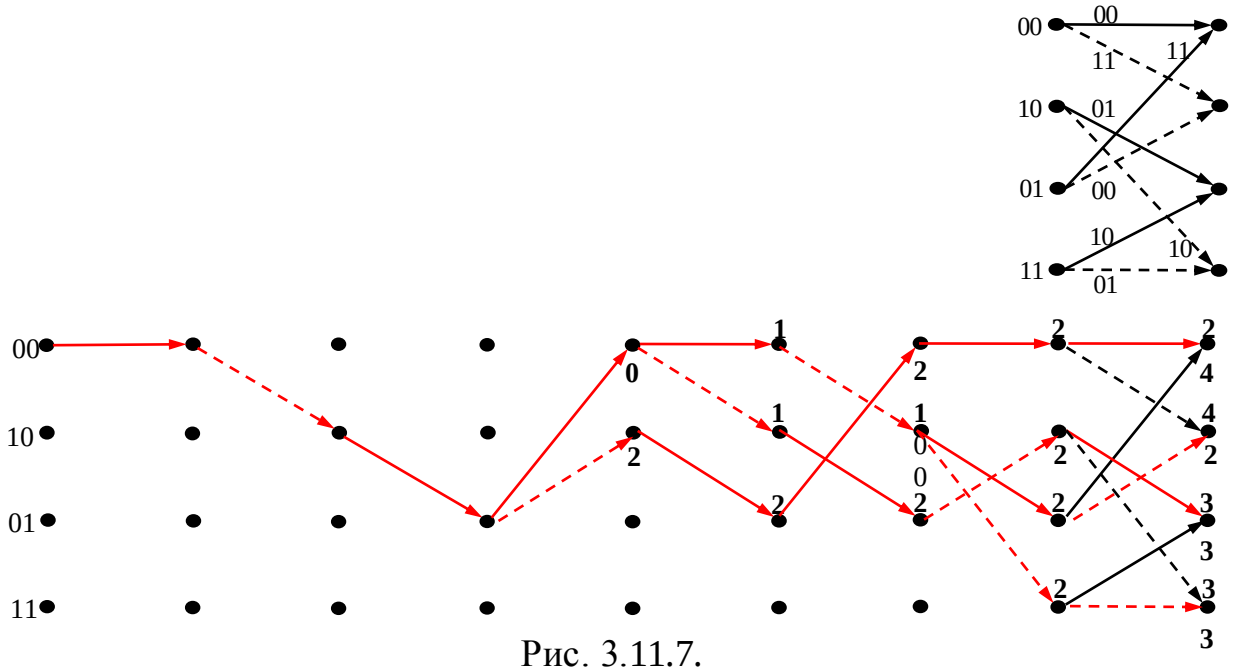


Рис. 3.11.7.

Для разрешения подобной неопределенности могут быть использованы различные стратегии. Простейшей является стратегия, основанная на случайном выбор одного из двух путей с вероятностью 0,6. Применяя указанную стратегию, предположим, что выжившими остаются пути, изображенные на рис. 3.11.8 (красные стрелки). Следует отметить тот факт, что оба «случайных» пути, прошедшие на предыдущем этапе, в дальнейшем отбраковываются, т.е. они не остаются в числе выживших. Подтверждением сказанного является диаграмма 9-го этапа, приведенная на рис. 3.11.9. С учетом этого становится очевидным, что случайный выбор пути на предыдущем этапе не имел последствий, поскольку отвечающие ему пути исключаются из дальнейшего рассмотрения.

На девятом этапе получается модель переходов, имеющая чрезвычайно определенный вид, которую рассмотрим более подробно (рис. 3.11.9).

Во-первых, маршруты с предшествующими неопределенными решениями отбраковываются, и значит, возможное влияние не очевидных решений на более ранних этапах полностью исключается.

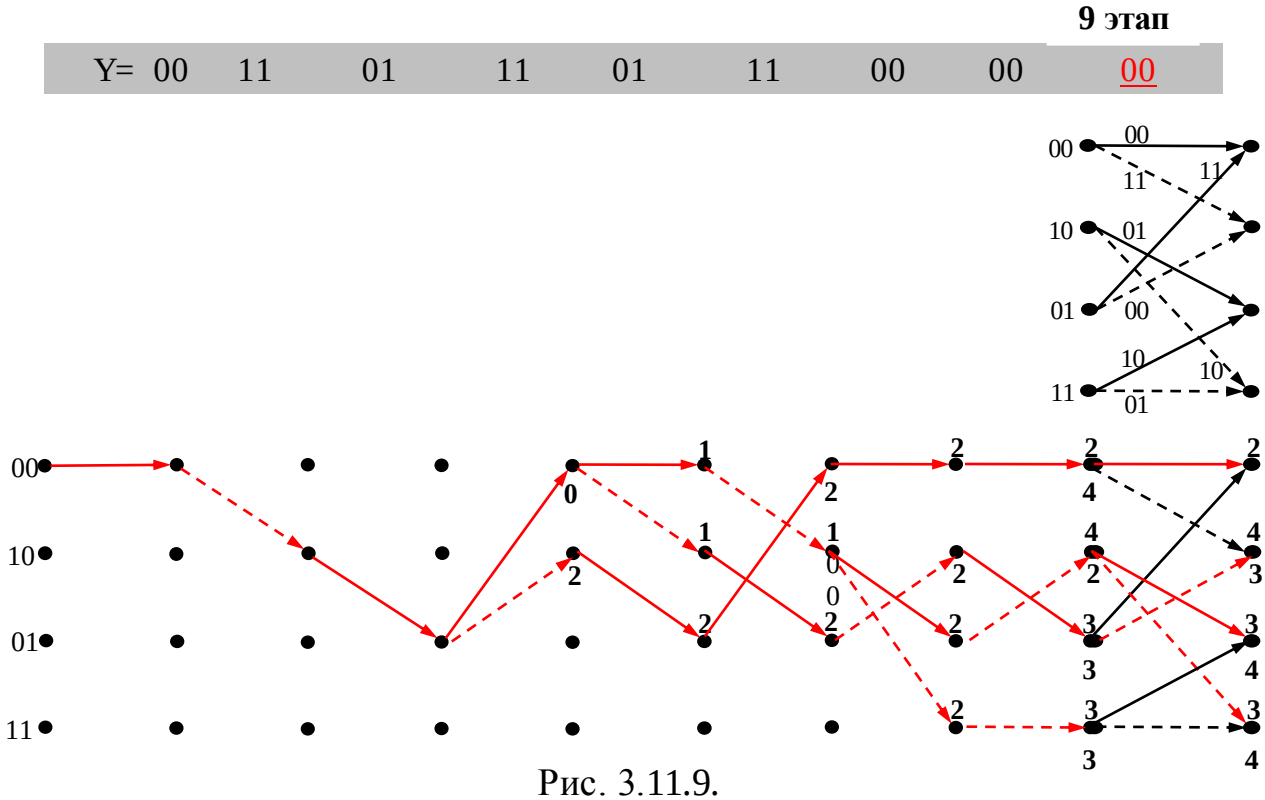


Рис. 3.11.9.

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 принятой кодовой последовательности. Диаграмма на рис. 3.11.10 отражает процесс декодирования принятого информационного сообщения.

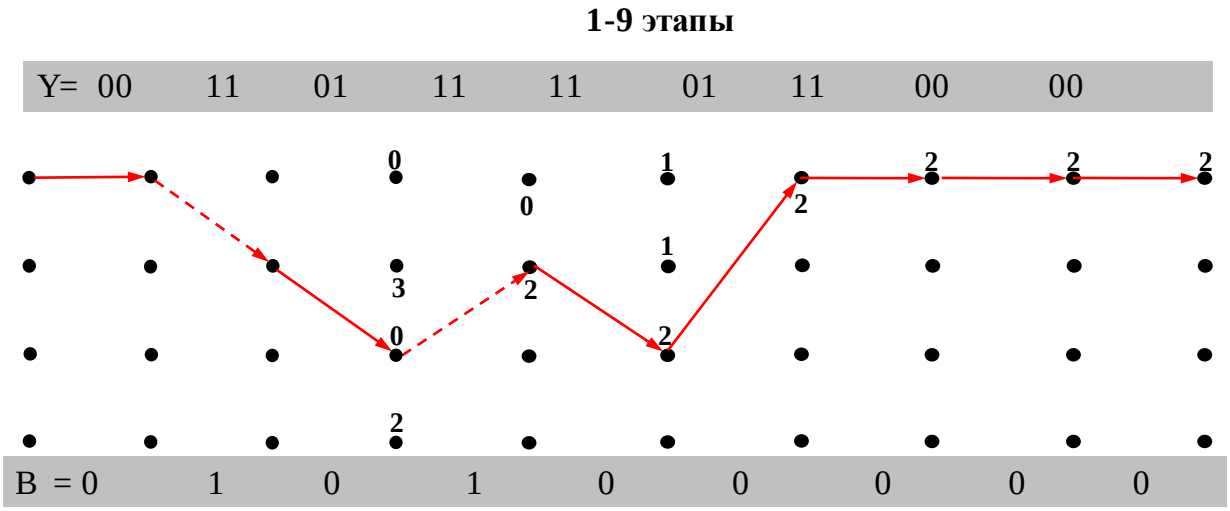


Рис. 3.11.10.

Обсудим теперь, что происходит после прекращения потока передаваемых битовых символов. 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.

Поскольку декодеру известно, что, начиная с 10-го шага, идет процесс обнуления кодера, то возможными являются только переходы в состояния 00 или 01, что нашло отражения в диаграмме, представленной на рис. 3.11.10. В результате остаются выжившими только два пути и декодируется еще один (10–й) бит. На окончательном шаге декодирования (см. рис. 3.11.11) осуществляется выбор между двумя путями, сливающимися в состоянии 00, устанавливая окончательный результат декодирования.

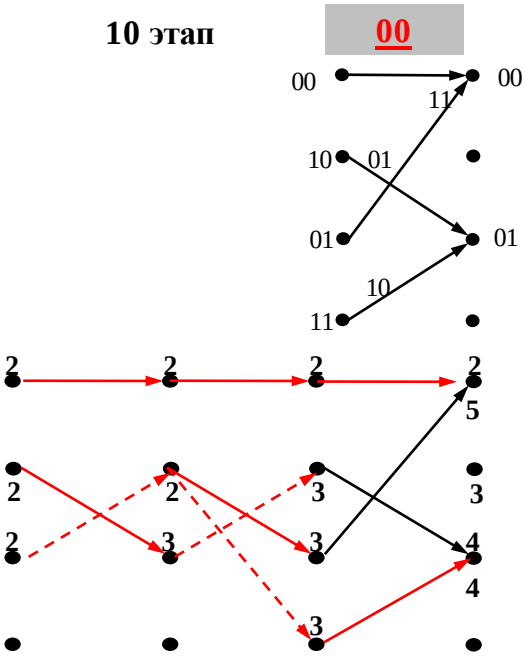


Рис. 3.11.10.

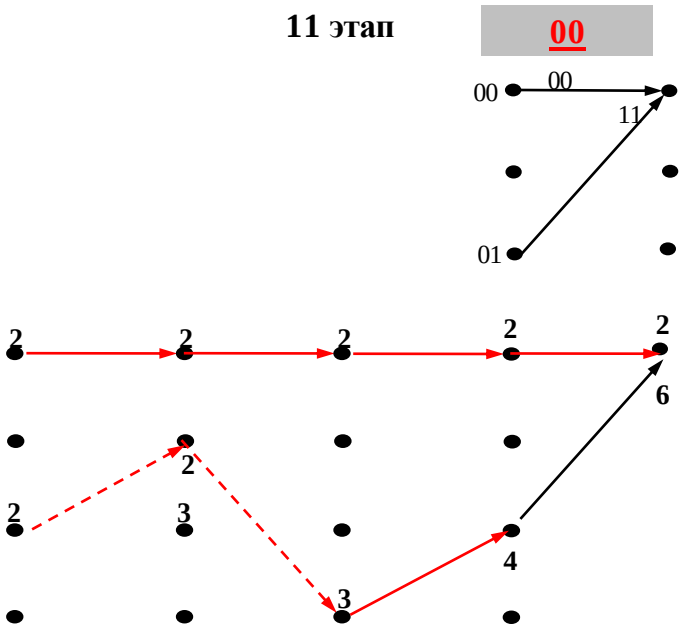


Рис. 3.11.11.

Как отмечалось в рассматриваемом примере 3.2.3, при передаче сообщения произошла ошибка в 4-ой группе бит, и принятым оказалось сообщение $y = 00\ 11\ 01_{*}11_{*}01\ 11\ 00\ 00\ldots$. После процесса декодирования принято сообщение $B = 0101000\ldots$. Таким образом, произошло исправление двукратной ошибки в полном соответствии с корректирующей способностью кода (свободное расстояние рассматриваемого кода $d = 5$).

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 завершенный пример может служить частичным подтверждением данного утверждения.

3.4. Общие понятия о турбо–кодах

Хотя теоретическая возможность достижения скоростей безошибочной передачи данных, близких к пропускной способности канала, была доказана Шенноном еще в 1948 г, в течение десятилетий вплоть до 1993 г она рассматривалась скорее как нереалистический идеал, чем практическое руководство. Проблема заключалась в том, что при стремлении к пропускной способности длина соответствующих кодов растет и, как следствие, экспоненциально увеличивается сложность эффективных методов декодирования, быстро становящихся технологически нереализуемыми. Огромное количество усилий и времени было затрачено на поиски кодов, обеспечивающих близкое к пропускной способности качество и полиномиальную зависимость сложности декодера от длины, однако в течение долго времени разрыв между теоретически предсказанным и практически реализуемым выигрышем от кодирования оставался значительным.

С открытием турбо-кодов ситуация изменилась коренным образом. Название *турбо-коды* присвоено авторами (Berrou и др.), открывшими их в 1993 году. В настоящее время турбо–коды широко признаны в качестве эффективного средства обеспечения высококачественной связи, особенно в условиях малого отношения сигнал-шум на бит. *Точная и компактная теория этого класса кодов остается скорее разрабатываемой, чем действительно уже существующей.* До сих пор интуиция в сочетании с обширными компьютерными поисками играют значительную роль в получении многих результатов.

3.4.1 Построение турбо-кодов.

Другим их наименованием турбо-кодов является *параллельно-формируемые коды*. Основная идея формирования состоит в параллельном кодировании одного и того же блока бит источника несколькими кодерами, называемых *образующими* или *компонентными*. Как оказалось, уже двух образующих кодеров достаточно для достижения скорости, близкой к пропускной способности канала.

В дальнейшем акцентируем внимание именно на данном варианте.

Схема кодирования турбо-кодов представлена на рис. 3.12.

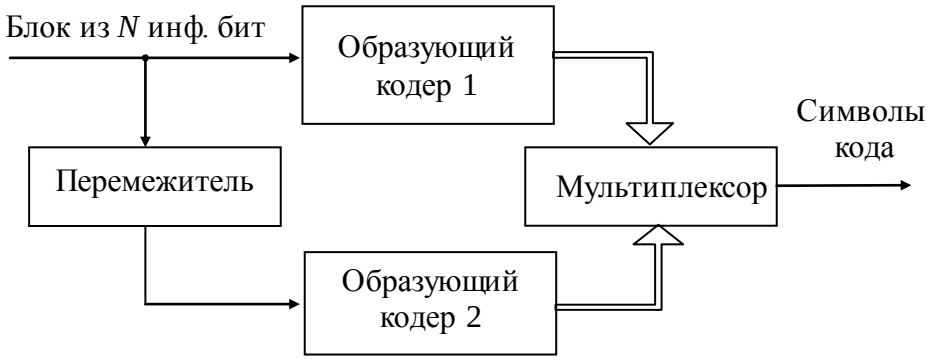


Рис. 3.12. Схема кодирования турбо-кодов.

Кодеры компонентов, как правило, обладают одинаковыми длинами кодового ограничения и множествами порождающих полиномов. Первый из них кодирует данные непосредственно, тогда как перед поступлением на второй кодер битовый поток данных подвергается перемежению. Данная операция состоит в перестановке по псевдослучайному закону бит данных в пределах блока фиксированной длины *I*. Действие перемежителя направлено на формирование образующих кодовых слов максимально независимыми с тем, чтобы минимизировать риск одновременного возникновения слов минимального веса. Отметим, что в случае линейности обоих образующих кодов результирующий турбо-код также будет линейным, поскольку преобразование, выполняемое перемежителем, также линейно. Таким образом, максимизация веса кодовых слов эквивалентна улучшению дистанционных свойств кода.

Полученные после кодирования обоими образующими кодерами слова объединяются в одно путем обычного мультиплексирования (т.е. преобразования из параллельной формы в последовательную).

3.4.2 Турбо-коды на основе сверточных кодов.

Рассмотренные ранее сверточные коды основаны на структуре с *конечным импульсным откликом* (КИО, англ. FIR). Применение структуры с *бесконечным импульсным откликом* (БИО, англ. IIR) вместо КИО открывает

путь к формированию систематического сверточного кода, обладающего такими же дистанционными свойствами, как и несистематический.

Целью применения рекурсивной схемы (на основе IIR фильтров) служит стремление избежать одновременного генерирования слов с малым весом обоими образующими кодерами. Дополнительным достоинством IIR-кодера может рассматриваться и то, что он способен генерировать такие кодовые слова, в которых биты источника будут вложены в слово в систематической форме. Систематическое компонентное кодирование в некоторой степени упрощает процедуру турбо декодирования.

Для приведения рекурсивного сверточного IIR кодера к эквивалентному FIR кодеру с фиксированным множеством порождающих полиномов $g_i(z),\ i=1,2,\ldots,n$ достаточно поделить входной полином данных $a(z)$ (z–преобразование) на первый порождающий полином $g_1(z)$ $(b(z)=a(z)/g_1(z))$ перед подачей на вход FIR кодера.

Обобщенная структура кодера турбо-кодов представлена на рис. 3.13.

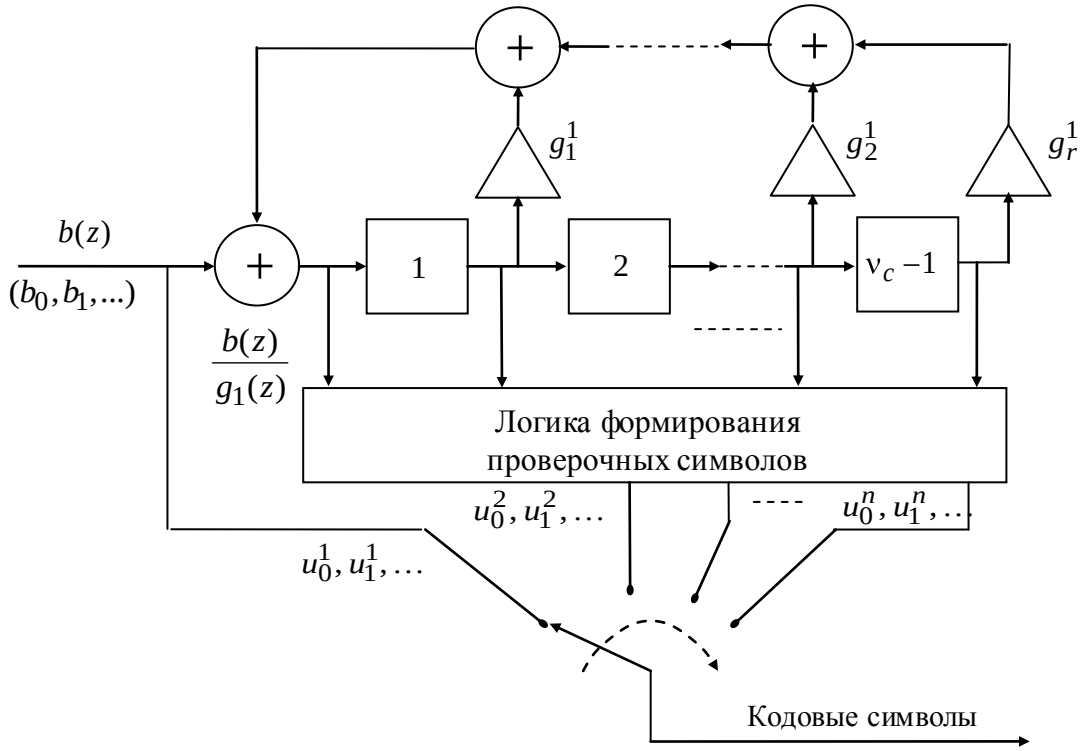


Рис. 3.13. Систематический сверточный кодер.

Регистр с обратной связью, представленный на рис. 3.13, реализует деление на полином $g_1(z)=g_r^1 z^r + g_{r-1}^1 z^{r-1} + \ldots + 1$, где $r=\nu_c - 1$. Согласно общему правилу (применимому к любой линейной системе) в z–области передаточная функция $h_f(z)$ замкнутой петли обратной связи определяется выражением

$$h_f(z) = \frac{h(z)}{1+\beta(z)h(z)},$$

$$h_f(z) = \frac{h(z)}{1+\beta(z)h(z)},$$

где $h(z)$ и $\beta(z)$ – передаточные функции разомкнутой петли и схемы обратной связи соответственно. Рассматривая выход крайнего слева сумматора на рис. 3.13 как выход замкнутой петли, имеем

$$h(z)=1,\ \beta(z)=g_1^1 z+g_2^1 z^2+\ldots+g_r^1 z^r,$$

так что

$$h_f(z)=1/g_1(z). \qquad (3.12)$$

Поделенный битовый поток $b(z)/g_1(z)$ непосредственно подается на вход регистра. Логическая схема (система сумматоров по модулю 2 на рис. 3.4) формирования проверочных символов, определяемая полиномами $g_l(z), l=2,\ldots,n$, может быть подсоединена к тому же регистру для дальнейшего выполнения обычного сверточного кодирования.

Таким образом, в систематическом кодере с обратной связью битовый поток источника непосредственно определяет поток бит данных $u_1(z)$, тогда как потоки проверочных символов $u_l(z), l=2,\ldots,n$ формируются схемой мультиплексирования, но подсоединенной к регистру с обратной связью.

Пример 3.3.1. Рассмотрим сверточный код с длиной кодового ограничения $m=3$. Его порождающие полиномы $g_1(z)=1+z^2$, $g_2(z)=1+z+z^2$ определяют структуру FIR кодера, представленного на рис. 3.14.

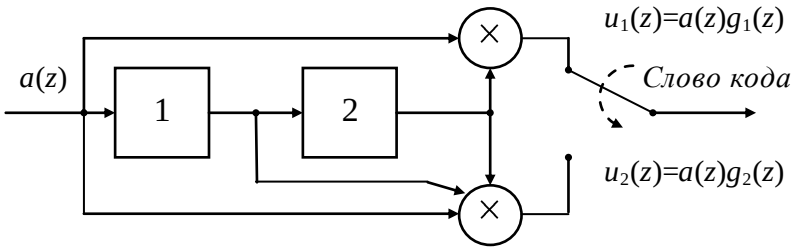


Рис. 3.14.

Поскольку $g_1(z)+1=z^2$, то единственным соединением в обратной связи IIR кодера остается только то, которое идет с крайне правой ячейки регистра. Очевидно, что не имеет смысла первоначально делить $a(z)$ на $g_1(z)$, а затем снова на него умножать, чтобы в результате получить $u_1(z)=a(z)$. Вместо этого входной поток бит непосредственно подается на выход, формируя, тем самым,

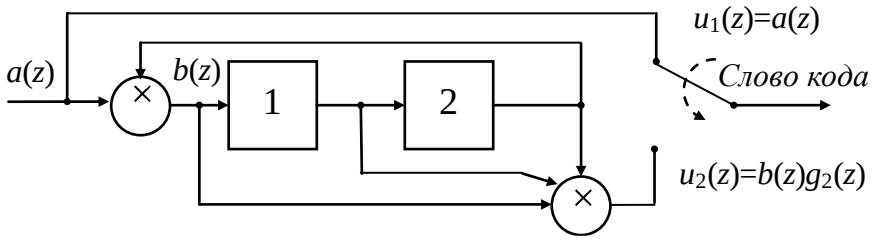


Рис. 3.16.

систематический код. В результате, приходим к структуре IIR кодера, представленного на рис. 3.16.

В **о**кончательной структуре турбо кодера биты данных со второго образующего сверточного кодера отбрасываются и в кодовом слове используются только избыточные кодовые символы. Поэтому при образующих кодах со скоростью $R=1/n$ окончательная скорость составит величину $R=1/(2n-1)$. Как правило, работа первого образующего кодера завершается дополнением входного битового потока соответствующими $m-1$ (главным образом ненулевыми из-за IIR) хвостовыми битами. Для того чтобы вернуться к исходной скорости образующих кодов, удобным инструментом служит процедура выкалывания. Очевидно, что выкалываются только избыточные символы, причем таким образом, чтобы обеспечить баланс удаляемых символов в обоих образующих кодах.

Пример 3.3.2. Опираясь на образующий кодер из предыдущего примера, приходим к турбо коду со скоростью $R=1/3$, генерируемому кодером на рис. 3.16. Для возвращения к образующей скорости $R=1/2$ можно отбрасывать каждый четный и каждый нечетный проверочные символы первого и второго образующих кодов соответственно.

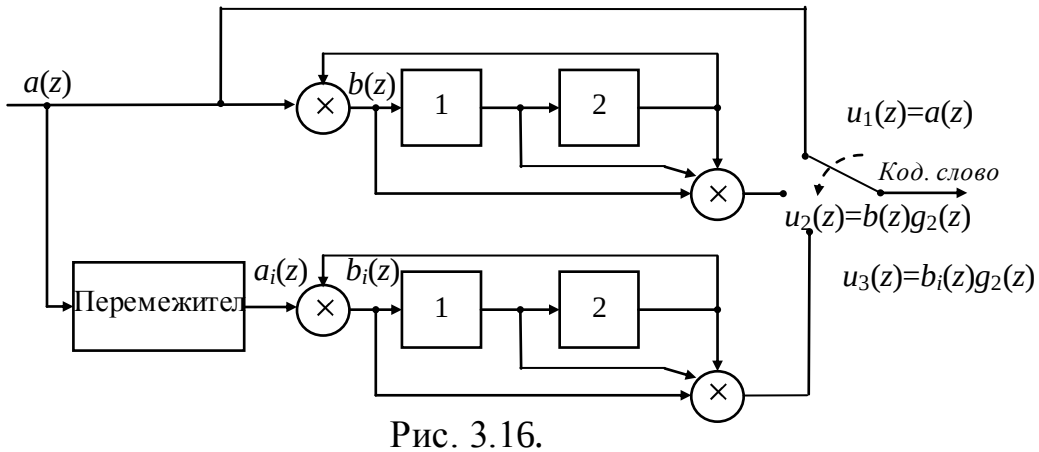


Рис. 3.16.

Предположим, что входной блок бит представляет собой $(1,0,1,0,0,0,0)$, так что $a(z)=1+z^2$. Тогда $b(z)=a(z)/g_1(z)=1$ и последовательность проверочных символов первого компонентного кода определяется как $u_2(z)=g_2(z)=1+z+z^2$,

формируя первое образующее кодовое слово с общим весом пять (минимальным для (5,7) кода). Пусть псевдослучайный перемежитель перегруппирует входные биты по правилу $(7,4,1,2,5,3,6)$, т.е. первый бит становится седьмым, второй – четвертым, третий – первым и т.д. Тогда перемеженные биты представимы как $(1,0,0,0,\ 0,0,1)$ или в полиномиальном представлении $a_i(z)=1+z^6$. После деления на $g_1(z)$ приходим к результату $b_i(z)=(1+z^6)/(1+z^2)=1+z^2+z^4$, на основании которого получаем выражение для проверочных символов второго образующего кода в виде $u_3(z)=b_i(z)g_2(z)=(1+z+z^2)(1+z^2+z^4)=1+z+z^3+z^5+z^6$. В итоге, слово турбо-кода скорости $1/3$ представимо как $(1,1,1,0,1,1,1,1,0,0,0,1,0,0,0,0,0,1,0,0,1)$, т.е. имеет вес равный десяти. Выкалывание каждого четного символа в $u_2(z)$ и каждого нечетного символа в $u_3(z)$ приводит к слову турбо-кода скорости $1/2$ вида $(1,1,0,1,1,1,0,1,0,0,0,1,0,0)$ веса семь.

3.4.3 Процедура итеративного декодирования.

Несмотря на то, что турбо–код состоит из двух сверточных кодов, его оптимальное декодирование не может быть реализовано в виде двух независимых процедур Витерби, поскольку пути на решетчатых диаграммах его компонентов связаны друг с другом вследствие кодирования одного и того же (хотя и с перемежением) потока данных. Открыватели турбо–кодирования предложили использовать итеративную версию правила *максимума апостериорной вероятности* (МАП) (MAP). Представленное ниже описание касается только основной ее идеи.

Апостериорные вероятности $p\ b_i=0|\mathbf{y}$ и $p\ b_i=1|\mathbf{y}$ характеризуют правдоподобность одного или другого значения i–го бита данных, вычисленную на основе вектора наблюдения $\mathbf{y}$. Они содержат в себе полную информацию о b_i, доставляемую $\mathbf{y}$ и извлекаемую из него. В случае достоверной связи одна их этих вероятностей близка к единице, тогда как вторая почти равна нулю. Действительно, МАП декодер будет выдавать на выход в качестве оценки $\hat{b}_i$ i–го бита данных такое значение, которому отвечает большая апостериорная вероятность, следуя правилу

$$\Lambda_i=\frac{p\ b_i=0|\mathbf{y}}{p\ b_i=1|\mathbf{y}}\ \begin{matrix}\hat{b}_i=0\\ >\\ <\\ \hat{b}_i=1\end{matrix}\ 1. \qquad (3.13)$$

В рассматриваемой ситуации бит данных b_i фактически присутствует в наблюдениях вследствие применения систематического кодирования. При непосредственной реализации рекуррентный МАП алгоритм вычисляет

апостериорную вероятность b_i, используя все ранее полученные наблюдения, включающие данный бит. После принятия всех отсчетов наблюдения, он возвращается назад и пересматривает результаты, содержащие информацию, извлеченную из наблюдений, поступивших после b_i. Таким образом, после прохождения вперед и возвращения назад могут быть найдены апостериорные вероятности для всех бит данных. Для реализации этого алгоритма необходимо знать решетку кода, канальные переходные вероятности и распределение априорной вероятности $q(b_i)$ для каждого бита данных. Очевидно, что не возникнет никаких проблем при реализации алгоритма в отношении каждого из двух компонент сверточных кодов, образующих турбо–код. Однако основная идея турбо–кода состоит в кодировании одних и тех же перемеженных данных двухкомпонентным кодом, так что информация о b_i может быть восстановлена совместно по обоим из них. Итерационный процесс, организованный, как показано на рис. 3.17, удовлетворяет данному условию.

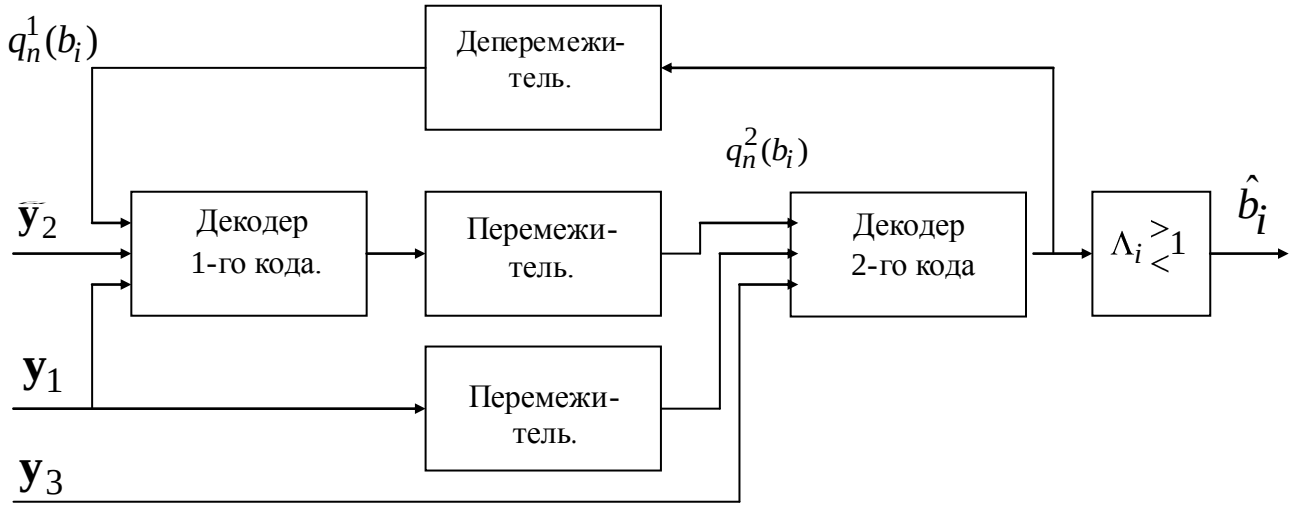


Рис.3.17. Итерактивный турбо-кодер

Полный вектор наблюдения на приемной стороне может быть разложен на три составляющих: $\mathbf{y}_l = \mathbf{u}_l + \mathbf{n}_l$, $l = 1,2,3$, где $\mathbf{u}_1 = (u_0^1, u_1^1, \ldots, u_{M-1}^1)$, $u_i^1 = b_i$ – вектор бит данных b_i непосредственно представлен в любом кодовом слове вследствие систематического характера турбо–кодов, $\mathbf{u}_2, \mathbf{u}_3$ – вектора избыточных символов первого и второго кодеров соответственно, а $\mathbf{n}_l$ – вектор независимых выборок шума.

На первом шаге декодер первого компонента кода вычисляет апостериорные вероятности $p_1(b_i | \mathbf{y}_1, \mathbf{y}_2)$ всех I бит данных b_i, использую наблюдения $\mathbf{y}_1, \mathbf{y}_2$, связанные с проверочными символами и символами данных этого кода. В качестве исходной информации используется а приори равномерное распределение $q_1^1(b_i) = 1/2, b_i = 0,1$, поскольку естественно считать равновероятными все образцы бит данных. После этого декодер

второго кода, вычисляющий апостериорные вероятности $p_1(b_i|\mathbf{y}_1,\mathbf{y}_3)$, может основываться не только на соответствующих наблюдениях $\mathbf{y}_1,\mathbf{y}_3$, но также на информации, доставляемой первым декодером, используя его апостериорное распределение $p_1(b_i|\mathbf{y}_1,\mathbf{y}_2)$ в качестве априорного: $q_1^2(b_i)=p_1(b_i|\mathbf{y}_1,\mathbf{y}_2)$. В результате имеем первое приближение $p_1(b_i|\mathbf{y})$ апостериорного распределения $p(b_i|\mathbf{y})$. Поскольку на этом шаге первый декодер не снабжается информацией от второго, то это положение исправляется на второй итерации, когда первый декодер снова декодирует первый компонент кода, но при использовании априорного распределения $q_2^1(b_i)=p_1(b_i|\mathbf{y})$. Продолжая подобным образом, после *n*–го шага следующая аппроксимация $p_n(b_i|\mathbf{y})$ распределения $p(b_i|\mathbf{y})$ формируется вторым декодером и используется первым декодером в качестве следующего априорного распределения $q_{n+1}^1(b_i)$ для выработки $p_{n+1}(b_i|\mathbf{y}_1,\mathbf{y}_2)$. Последнее в свою очередь, используется вторым декодером как следующее априорное распределение $q_{n+1}^2(b_i)$ для выработки следующей аппроксимации ожидаемой апостериорной вероятности $p_{n+1}(b_i|\mathbf{y})$ и т.п. Поскольку перемежитель меняет порядок следования бит данных перед подачей их на вход второго кодера, перемежители аналогичным образом воздействуют на наблюдения $\mathbf{y}_1$ и априорное распределение $q_n^2=p_n(b_i|\mathbf{y}_1,\mathbf{y}_2)$, подающиеся на второй декодер. Аналогично, деперемежитель восстанавливает первоначальный порядок бит в схеме обратной связи, передавая $p_n(b_i|\mathbf{y})$ с выхода второго декодера на вход первого. С учетом этих преобразований все данные, обрабатываемые обоими декодерами, соединяются должным образом.

Обширное моделирование экспериментально подтвердило сходимость этих итераций, однако теоретическое обоснование до сих пор остается под вопросом.

4 ПОСТРОЕНИЯ СИСТЕМ МНОЖЕСТВЕННОГО ДОСТУПА

Большое число современных радиосистем связи относятся к *многопользовательским*. В многопользовательской системе множество линий связи размещаются в пределах общего частотно-временного ресурса, так что любому индивидуальному абоненту разрешалось передавать или принимать свою собственную информацию параллельно с другими пользователями и независимо от них.

4.1 Принципы построения сотовых систем 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, разделения 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 разделения 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. 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 Fl – F3 представлен на рис. 4.1. Группа сот с различными наборами частот называется кластером. Определяющим его параметром является количество используемых в соседних сотах частот. На рис. 4.1, например, размерность кластера равна трем. Но на практике это число может достигать пятнадцати.

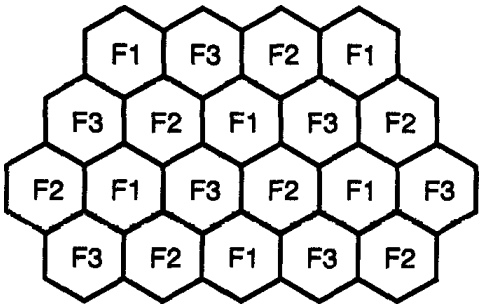


Рис. 4.1. Построение сот для трех частот
F1. F2. F3 – частоты базовых станций.

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. Он предполагает передачу сигнала одинаковой мощности по всем направлениям, что для абонентских станций эквивалентно приему помех от всех базовых станций со всех направлений.

На рис. 4.2 базовые станции, на которых допускается повторное использование выделенного набора частот, удалены друг от друга на расстояние D, называемое «защитным интервалом». Именно возможность повторного применения одних и тех же частот определяет высокую эффективность использования частотного спектра в сотовых системах связи.

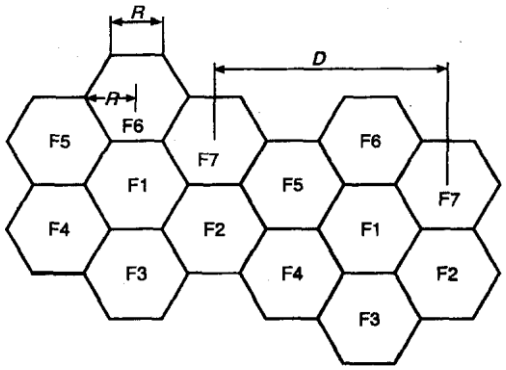


Рис. 4.2. Повторное использование частот в несмежных сотах.

Смежные базовые станции, использующие различные наборы частотных каналов, образуют группу из C станций, образуя кластер. Если каждой базовой

станции выделяется набор из m каналов с шириной полосы каждого F_k, то общая ширина полосы, занимаемая системой сотовой связи, составит $F_c = F_k mC$. Величина C определяет минимально возможное число каналов в системе, поэтому ее часто называют частотным параметром системы, или коэффициентом повторения частот. Коэффициент C 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 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.

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. На рис. 4.3 приведен способ повторного использования частот в организованных таким образом сотах. Модель основана на применении 3-секторных антенн для каждой базовой станции и трех соседних базовых станций с формированием ими девяти групп частот. В этом случае используются антенны с шириной диаграммы направленности 120°.

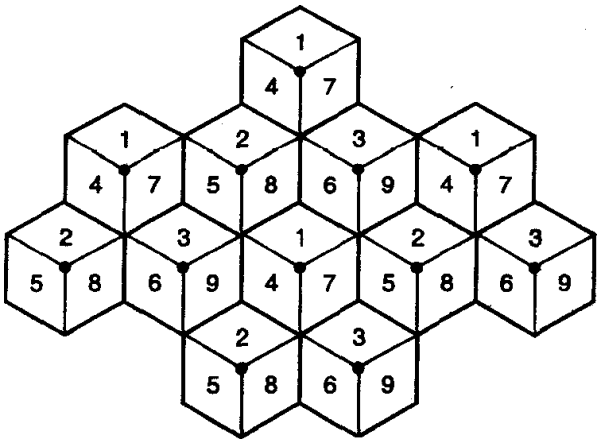


Рис. 4.3. Модель повторного использования частот в 3-секторных сотах

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 (см. рис. 4.4) 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°.

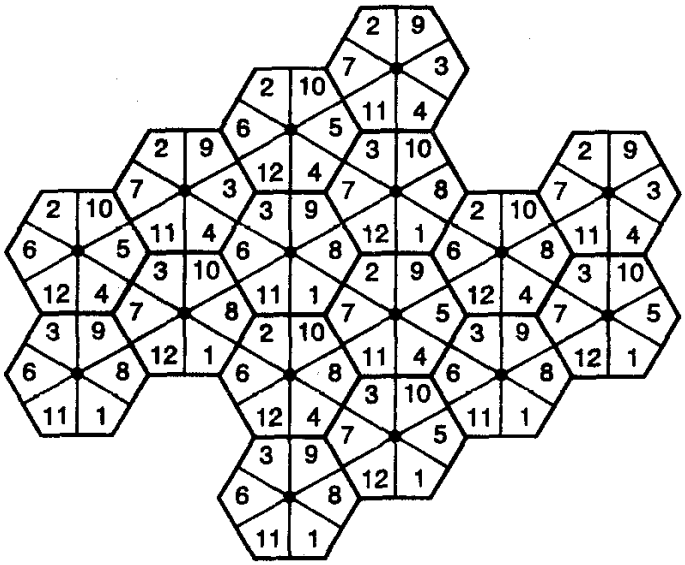


Рис. 4.4. Модель повторного использования частот в двух соседних

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, 36... Один из каналов является управляющим. В некоторых ситуациях он может называться также каналом вызова. 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. 4.5).

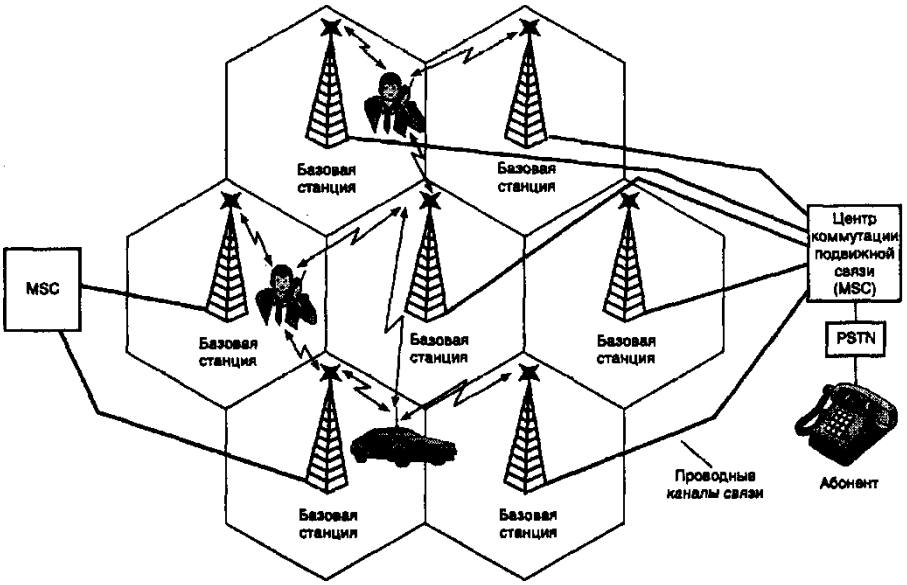


Рис. 4.5. Основные составляющие систем сотовой связи

Центр коммутации MSC - это автоматическая телефонная станция системы сотовой связи, обеспечивающая все функции управления сетью. Она осуществляет постоянное слежение за подвижными станциями, организует их эстафетную передачу, в процессе которой достигается непрерывность связи при перемещении подвижной станции из соты в соту и переключение рабочих каналов в соте при появлении помех или неисправностей, 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). Когда же радиотелефон находится в режиме ожидания, 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, 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 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. В этом случае 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. 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. Перспективы развития этой сферы услуг зависят уже от распространенности стандартов.

4.2 Понятие о модели взаимодействия открытых систем

Под *открытой* системой понимается система, удовлетворяющая требованиям международных организаций или требованиям производителя. 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.

Таблица. 4.1 иллюстрирует схему взаимодействия элементов и уровней базовой модели, представленной в общем виде. Рассмотрим используемые термины и определения.

Табл. 4.1. Модель взаимодействия открытых систем

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.

Прикладной процесс выполняет обработку данных для нужд пользователя либо с помощью некоторых аппаратно-программных средств.

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.

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 в табл. 4.1.

Такое деление позволяет разбить сложную систему на ряд более простых, 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).

В модели 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.

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. Под доступностью сервиса понимается возможность использования всеми верхними уровнями услуг (сжатие данных, защита от помех, от несанкционированного использования информации), выполняемых на нижних уровнях.

В табл. 4.1 линиями, соединяющими одноименные уровни, обозначены протоколы. Названия протоколов совпадают с названиями соответствующих уровней, например, существуют сеансовые, транспортные, сетевые протоколы.

Протоколы должны удовлетворять определенным требованиям. Протокол - это последовательность шагов, выполняемых взаимодействующими объектами, в порядке строгой очередности для достижения поставленной цели. Протокол должен содержать описание действий его участников при любых возникающих ситуациях (должны быть исключены тупиковые ситуации).

Следует упомянуть о существовании криптографических протоколов, обеспечивающих достижение цели в условиях возможного мошенничества, как со стороны законных пользователей, так и внешнего нарушителя.

4.3. Перспективы и проблема множественного доступа

Каждый пользовательский приемник в системе множественного доступа должен обеспечить выделение информации, адресованной непосредственно данному абоненту, из наблюдаемого сигнала, который содержит данные, направленные множеству пользователей.

Как частотный метод (FDMA), так и временной метод (TDMA)

множественного доступа предполагают распределение общего доступного частотно-временного ресурса между различными пользователями, так что каждый из них использует только свою собственную, специфическую для пользователя часть, и отсутствуют абоненты, совместно использующие некоторые части ресурса. При FDMA подобная фрагментация осуществляется в частотной области (рис. 4.6, а). Так, в распоряжении k–го пользователя находится весь временной ресурс $T = T_t$ и только часть частотного ресурса. Когда достижение максимального числа абонентов является приоритетной задачей, то $W = 1/T \approx W_t / K$.

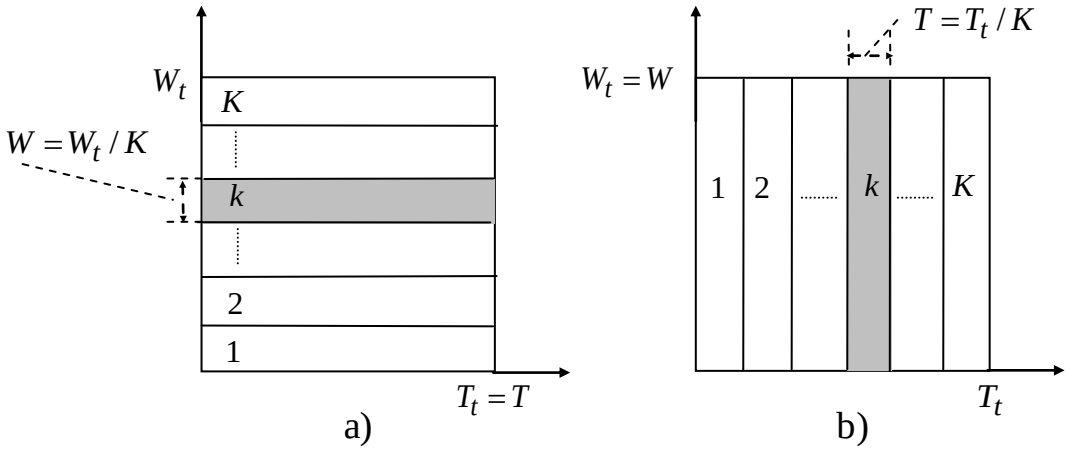


Рис. 4.6. Распределение ресурса при ортогональном частотном (*a*) и временном (*b*) кодировании.

Деление на части временной области при TDMA (рис. 4.6, b) предоставляет возможность одиночному пользователю занимать весь доступный частотный ресурс $W = W_t$, но только часть полного временного кадра $T = 1/W \approx T_t / K$. Если необходимо максимизировать число пользователей, то в обоих этих ортогональных схемах множественного доступа фрагментация ресурса заставляет использовать каждого абонента простой сигнал, поскольку для последнего частотно-временное произведение $WT = 1$.

С другой стороны, при необходимости обслуживания большого числа пользователей K общий частотно-временной ресурс также должен быть значительным $(WT >> 1)$, и если каждый пользовательский сигнал занимает как всю доступную полосу $W = W_t$, так и временной интервал $T = T_t$, то приходим к необходимости применения ортогональной схемы множественного доступа, в которой все пользовательские сигналы широкополосны.

Индивидуальный широкополосный сигнал, назначенный k-му пользователю, называется k-й *сигнатурой*. Каждая сигнатура занимает всю полосу W_t и весь временной кадр T_t (см. рис. 4.7), передавая $\log_2 M$ бит данных за интервал T_t. Если $M > 2$, то данный метод множественного доступа может обслуживать до $K = WT = W_t \log_2 M \ / R$ абонентов.

В рассмотренном методе множественного доступа соответствующее кодирование сигнатур обеспечивает ортогональность сигналов пользователей вместо фрагментации временной или частотной области. Именно поэтому

данному методу присвоили название *множественного доступа с кодовым разделением* (CDMA). Достоинства CDMA в сравнении с классическими методами FDMA и TDMA (помехоустойчивость, низкая вероятность обнаружения и др.) автоматически следуют из широкополосной природы CDMA сигнатур.

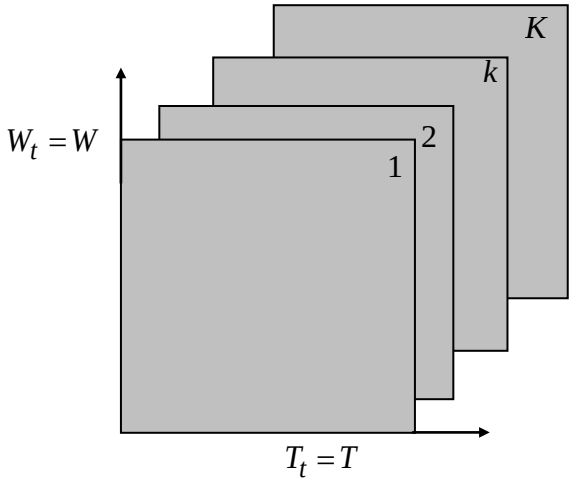


Рис. 4.7. Распределение ресурса в широкополосном методе множественного доступа.

Стратегической целью развития современных 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.

4.4 Широкополосная передача данных

В многопользовательской сети каждый из K абонентов передает или получает свои индивидуальные данные посредством использования некоторой специфической для пользователя сигнатуры, причем тщательно выбранный ансамбль из K сигнатур обеспечивает максимально возможную совместимость абонентов. С целью обеспечения транспортировки k–й сигнатурой потока данных необходимо применение некоторого вида модуляции, которая, учитывая широкополосную природу CDMA сигнатур, часто называется широкополосной модуляцией.

Наиболее типична для современных многопользовательских систем классическая версия широкополосной модуляции: *прямой кодовой последовательностью*.

Рассмотрим случай передачи бинарных ФМ данных с расширением спектра бинарной ФМ прямой последовательностью.

Начнем рассмотрение с простейшего случая передачи данных с помощью бинарной модуляции. Пусть $B_k(t)$ информационный сигнал k–го пользователя,

в котором импульсы положительной и отрицательной полярности длительности T_b отвечают передаче бита информации, равного 0 или 1 соответственно. Передача $B_k(t) = b_{k,i} = \pm 1,\ (i-1)T_b < t \le iT_b$ с использованием бинарной ФМ просто означает умножение ее на непрерывную несущую частоту f_0, позволяя прийти к отсылаемому модулированному сигналу

$$s_k(t;\mathbf{b}_k) = B_k(t)\cos(2\pi f_0 t). \qquad (4.1)$$

Рассмотрим теперь изменения, которые необходимо осуществить в случае передачи бинарных ФМ данных с расширением спектра бинарной ФМ прямой последовательностью.

Пусть $s_k(t)$ – k–я пользовательская сигнатура, т.е. дискретный сигнал, состоящий из чипов длительности Δ, манипулированная некоторой специфической для каждого пользователя бинарной последовательностью. Пусть на интервале одного бита данных содержится N чипов сигнатуры. Тогда расширение прямой последовательностью бинарного ФМ сигнала описывается выражением

$$s_k(t;\mathbf{b}_k) = s_k(t)B_k(t)\cos(2\pi f_0 t). \qquad (4.2)$$

Поскольку ширина полосы сигналов (4.1) и (4.2) есть величина обратная длительности бита $T_b = 1/R$ и длительности чипа $\Delta = T_b / N = 1/RN$ соответственно, то распределение прямой последовательностью расширяет спектр в N раз.

После распространения по каналу, в ходе которого приобретается задержка τ_k и начальная фаза φ_k, сигнал принимает вид

$$s_{kr}(t;\mathbf{b}_k) = s_k(t-\tau_k)B_k(t-\tau_k)\cos(2\pi f_0 t + \varphi_k). \qquad (4.3)$$

При полном знании о параметрах τ_k, φ_k, приемник для восстановления текущего (i–го) бита должен лишь сделать выбор между сигналом $s_k(t-\tau_k)\cos(2\pi f_0 t + \varphi_k)$ и ее противоположной копией. Для выполнения этой операции оптимальным образом вначале может быть вычислена корреляция

$$z_k = \int_{(i-1)T_b+\tau_k}^{iT_b+\tau_k} y(t)\, s_k(t-\tau_k)\cos(2\pi f_0 t + \varphi_k)dt \qquad (4.4)$$

между наблюдаемым колебанием $y(t)$ и формируемой на месте полосной копией сигнатуры $s_k(t-\tau_k)\cos(2\pi f_0 t + \varphi_k)$, а ее полярность использована при вынесении решения.

Следует отметить, что та же самая оптимальная процедура может быть

реализована в два этапа: на первом снимается расширение, а на втором осуществляется демодуляция данных точно так, как и в случае отсутствия расширения.

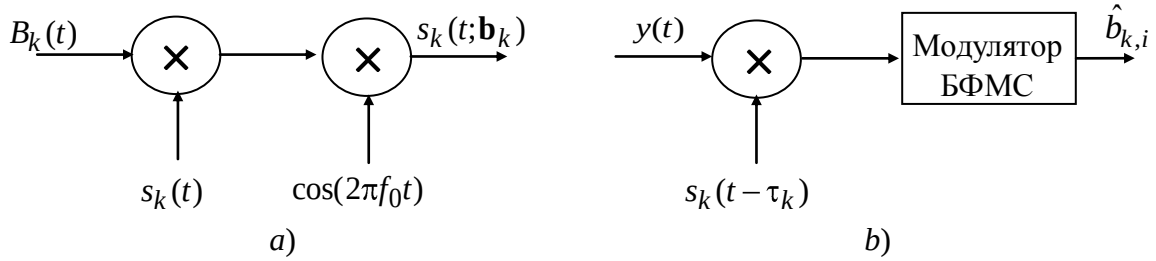
Пусть наблюдение $y(t)$ умножается на формируемую в приемнике низкочастотную копию $s_k(t-\tau_k)$ сигнатуры, в точности синхронизированную с принимаемым сигналом. Полезная составляющая наблюдения после этой операции примет вид

$$s_{kr}(t;\mathbf{b}_k)s_k(t-\tau_k) = s_k^2(t-\tau_k)B_k(t-\tau_k)\cos(2\pi f_0 t+\varphi_k) = B_k(t-\tau_k)\cos(2\pi f_0 t+\varphi_k)$$

,

где учтена бинарная природа сигнатур $(s_k(t)=\pm 1)$, вследствие чего $s_k^2(t)=1$. Как видно, после данного шага принятый сигнал не обладает никакими чертами широкополосного, полностью совпадая с простым сигналом с бинарной манипуляцией, определяемой потоком данных. Вследствие этого процедура умножения наблюдения на копию сигнатуры называется *снятием расширения*.

Полный цикл расширения и сжатия демонстрирует рис. 4.8.



a) *b*)

Рис. 4.8. Расширение (*a*) и сужение (*b*) сигнала с бинарной ФМ.

Для проведения обсуждения в терминах частотной области обратимся к свойствам спектральной плотности мощности $\tilde{S}_b(f)$, $\tilde{S}_{bs}(f)$ исходного потока данных $B_k(t)$ и его широкополосной версии $s_k(t)B_k(t)$ соответственно.

Для последовательности $B_k(t)$ битовых импульсов длительности T_b, полярности которых случайны и независимы, спектр мощности определяется как

$$\tilde{S}_b(f) = T_b \mathrm{sinc}^2(\pi f T_b).$$

Рассмотрение расширенного потока данных снова как случайную последовательность импульсов с независимыми полярностями (на этот раз длительности Δ) приводит к спектру мощности той же формы, но занимающему в N раз большую полосу

$$\tilde{S}_{bs}(f) = \Delta\,\mathrm{sinc}^2(\pi f \Delta) = (T_b / N)\,\mathrm{sinc}^2(\pi f T_b / N).$$

Идея прямого расширения спектра, рассмотренная выше для случая передачи данных с помощью бинарной ФМ и применения бинарных сигнатур,

легко может быть обобщена на более широкий диапазон методов модуляции данных и сигнатур.

Пусть $\dot{B}_k(t)$ обозначает комплексное колебание, отвечающее потоку данных $\mathbf{b}_k$ k–го пользователя, передаваемых в некотором цифровом формате.

В случае обычной (не широкополосной) M-ичной модуляции передаваемый сигнал, переносящий поток данных $\mathbf{b}_k$ (теперь удобно полагать, что компоненты $b_{k,i}$ потока $\mathbf{b}_k$ представляют собой M –ичные комплексные символы $\dot{B}_k(t) = b_{k,i}$, $(i-1)T < t \le iT$), имеет вид

$$s_k(t;\mathbf{b}_k) = \mathrm{Re}\left[\dot{B}_k(t) \exp(j2\pi f_0 t) \right].$$

Тогда принятый сигнал

$$s_{kr}(t;\mathbf{b}_k) = \mathrm{Re}\ \ \dot{B}_k(t-\tau_k) \exp[\, j(2\pi f_0 t + \varphi_k)]$$

обладает комплексной огибающей вида

$$\dot{S}_{kr}(t;\mathbf{b}_k) = \dot{B}_k(t-\tau_k) \exp(j\varphi_k). \qquad (4.5)$$

Передаваемый i–й символ данных $b_{k,i}$ есть ничто иное, как комплексная амплитуда непрерывной несущей, постоянной на интервале $((i-1)T, iT]$. Следовательно, для восстановления этого символа приемник должен принять решение относительно M копий прямоугольного импульса, существующего на интервале $((i-1)T + \tau_k, iT + \tau_k]$ и обладающего различными конкурирующими значениями комплексной амплитуды. Для выполнения указанной задачи необходимо вычислить корреляцию между наблюдаемой комплексной огибающей $\dot{Y}(t)$, выделенной из зашумленного наблюдения $y(t)$, и $\exp(j\varphi_k)$, представимую в виде

$$\dot{z}_k = \int\limits_{(i-1)T+\tau_k}^{iT+\tau_k} \dot{Y}(t) \exp(-j\varphi_k) dt,$$

которая (после нормировки к энергии символа опорного образца) используется для вынесения необходимой оценки $\hat{b}_{k,i}$ о символе $b_{k,i}$.

Демодулятор сигналов представлен на рис. 4.9, где двойной круг символизирует операцию комплексного умножения, реализует указанную процедуру. Данная схема является обобщением корреляционной структуры, представленной на рис. 4.8, b, на случай произвольной цифровой модуляции данных.

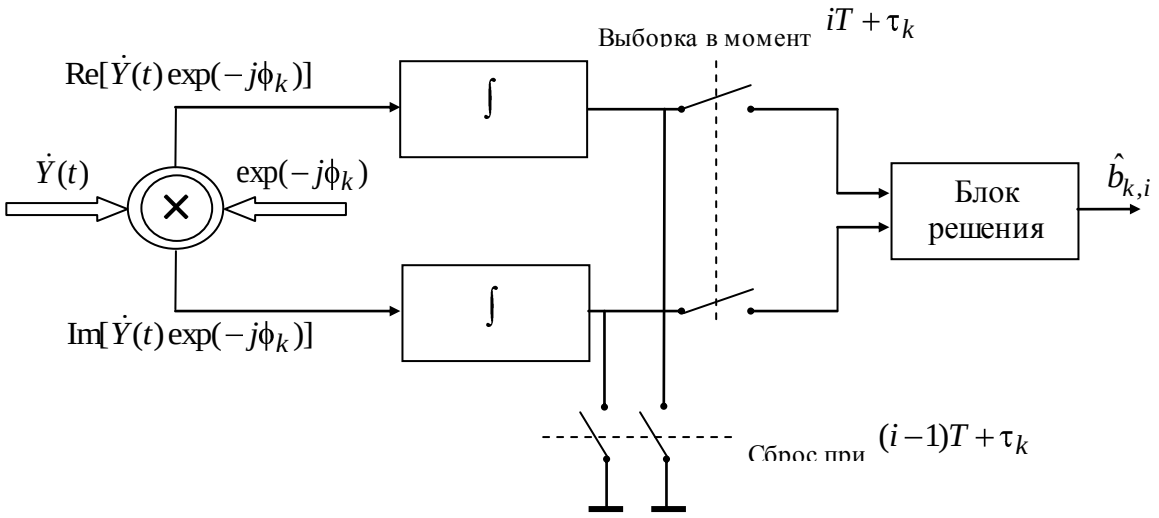


Рис. 4.9. *M*-ичный демодулятор.

Первоначально наблюдаемая комплексная огибающая $\dot{Y}(t)$ умножается на $\exp(-j\varphi_k)$ с целью компенсации канального фазового сдвига φ_k. После этой операции полезная компонента наблюдения $\dot{Y}(t)$ становится $\dot{B}_k(t-\tau_k)$. Любое комплексное колебание эквивалентно двум вещественным (реальной и мнимой частям), так что выход умножителя на рис. 4.9 трактуется в терминах реальной и мнимой частей произведения $\dot{Y}(t)\exp(-j\varphi_k)$. Полезными компонентами произведения служат реальная и мнимая части $\dot{B}_k(t-\tau_k)$, а последующее их интегрирование, как и ранее, служит выделению из шума. Отсчеты с выхода интегратора в моменты $iT+\tau_k$ являются оценками реальной и мнимой частей принятого *M*–ичного символа данных, которые используются в блоке решения для выдачи демодулированного символа.

Рассмотрим теперь, как прямое расширение спектра может быть включено в данную схему модуляции – демодуляции. Пусть $\dot{S}_k(t)$ будет комплексной огибающей k–й пользовательской CDMA сигнатурой. Тогда прямое расширение означает умножение колебаний модулированных данных $\dot{B}_k(t)$ на сигнатуру $\dot{S}_k(t)$ с целью использования их произведения $\dot{S}_k(t)\dot{B}_k(t)$ как комплексной огибающей передаваемого сигнала

$$s_k(t;\mathbf{b}_k)=\mathrm{Re}\left[\dot{S}_k(t)\dot{B}_k(t)\exp(j2\pi f_0 t)\right]. \qquad (4.6)$$

Принятый полезный сигнал представляет собой задержанную и сдвинутую по фазе копию сигнала (4.6)

$$s_{kr}(t;\mathbf{b}_k)=\mathrm{Re}\left[\dot{S}_k(t-\tau_k)\dot{B}_k(t-\tau_k)\exp[j(2\pi f_0 t+\varphi_k)]\right]$$

с комплексной огибающей

$$\dot{S}_{kr}(t;\mathbf{b}_k) = \dot{S}_k(t-\tau_k)\dot{B}_k(t-\tau_k)\exp(j\varphi_k)\,. \qquad (4.7)$$

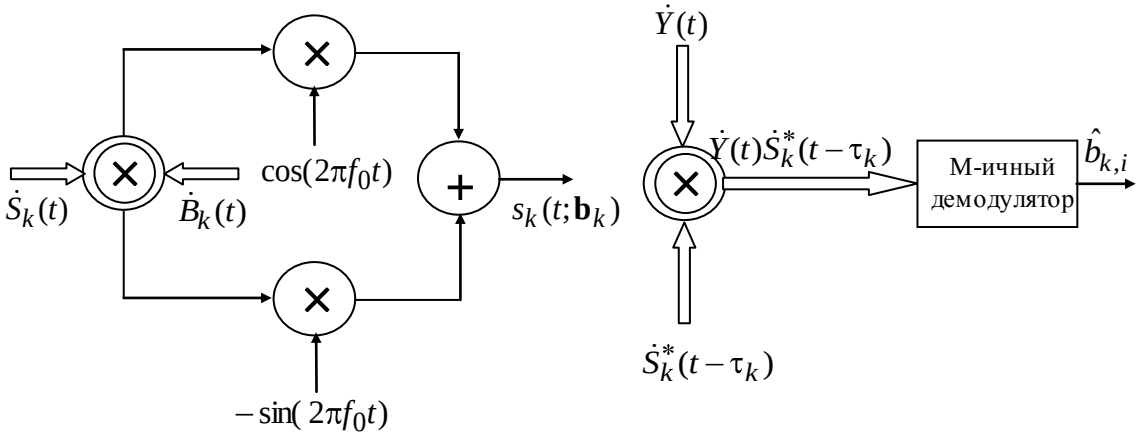
Постоянство $\dot{B}_k(t-\tau_k)$ на интервале $((i-1)T+\tau_k, iT+\tau_k]$ снова означает, что для выделения i–го символа приемник должен принять решение об истинности одной из M конкурирующих копий сигнатуры $\dot{S}_k(t-\tau_k)\exp(-j\varphi_k)$, умноженной на различные символы данных $b_{k,i}$. Тогда корреляция вида

$$\dot{z}_k = \int_{(i-1)T+\tau_k}^{iT+\tau_k} \dot{Y}(t)\dot{S}_k^*(t-\tau_k)\exp(-j\varphi_k)dt \qquad (4.8)$$

послужит (после соответствующей нормировки) для получения необходимой оценки $b_{k,i}$ и снова может трактоваться, как пара отсчетов с выходов интеграторов демодулятора, изображенного на рис. 4.9, при изменении опорного сигнала в комплексном умножителе с $\exp(-j\varphi_k)$ на $\dot{S}_k^*(t-\tau_k)\exp(-j\varphi_k)$. После умножения на подобный опорный сигнал полезная компонента наблюдаемой комплексной огибающей

$$\dot{S}_{kr}(t;\mathbf{b}_k)\dot{S}_k^*(t-\tau_k) = \left|\dot{S}_k(t-\tau_k)\right|^2 \dot{B}_k(t-\tau_k)\exp(j\varphi_k) \qquad (4.9)$$

на интервале i-го символа данных становится просто одной из M возможных копий видеосигнала $\left|\dot{S}_k(t-\tau_k)\right|^2$, умноженного на различные комплексные коэффициенты $b_{k,i}$. Если чипы сигнатуры подвергаются только фазовой манипуляции, то $\left|\dot{S}_k(t-\tau_k)\right|^2 = 1$ и, как показывает предыдущее соотношение, умножение на $\dot{S}_k^*(t-\tau_k)$ превращает комплексную огибающую принятого сигнала в одну из характеристик обычной (не широкополосной) M–ичной модуляции данных, т.е. осуществляет снятие расширения. Благодаря этому приемник снова может рассматриваться как двухэтапный: первоначально осуществляется снятие расширения, а затем обычная M–ичная демодуляция.
Схема на рис. 4.10 иллюстрирует операции, выполняемые на передающей и приемной сторонах обычной широкополосной системы с прямым расширением спектра. Модулятор (рис.4.10, a) реализует алгоритм, определяемый соотношением (4.6), используя только вещественную часть комплексного произведения. В демодуляторе (рис. 4.10, b) комплексная огибающая наблюдения в дальнейшем подается на вход обычного M–ичного демодулятора (см. рис. 4.9) для принятия решения о принятых символах.



a) b)

Рис. 4.10. Модуляция (*a*) и демодуляция (*b*) при прямом расширении спектра.

Отметим, что только что описанный способ реализации расширения–сжатия не является единственным и что существует многообразие конкретных схемных решений этих операций.

4.5 Синхронный метод множественного доступа

Рассмотрим K–пользовательскую CDMA систему с прямым расширением спектра, в которой все пользовательские потоки данных и все сигнатуры жестко синхронизованы, т.е. характеризуются нулевым взаимным временным сдвигом на входе приемника.

Очевидно, что групповой сигнал достигает приемника абонента с сохранением начальной синхронизации между сигналами, адресованными различным индивидуальным пользователям. В приводимом здесь анализе будем считать, что имеет место идеализированная модель канала, в котором многолучевая задержка на распространение $\tau_{\max}$ меньше периода следования чипов Δ пользовательских сигнатур.

Согласно концепции прямого расширения спектра, комплексная огибающая принятого группового сигнала $\dot{S}(t;\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_K)$ представляет собой сумму по k комплексных огибающих сигнатур, манипулированных пользовательскими потоками данных. В общем случае каждый из пользовательских сигналов может обладать собственной амплитудой, однако ограничимся простейшим случаем равной интенсивности сигналов. Поскольку предположение об идеальном синхронизме позволяет положить все задержки τ_k и начальные фазы φ_k равными нулю, получаем

$$\dot{S}(t;\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_K)=\sum_{k=1}^{K}\dot{B}_k(t)\dot{S}_k(t). \qquad (4.12)$$

Сосредоточим наше внимание на интервале длительности T одиночного символа данных. Снова, вследствие полного синхронизма, начало и конец текущих символов данных всех пользователей строго совпадают. Если через b_k обозначить $k-$ й текущий пользовательский символ данных, то на одиночном символьном интервале выражение (4.12) может быть записано в следующем виде

$$\dot{S}(t;\mathbf{b}) = \sum_{k=1}^{K} b_k(t)\dot{S}_k(t), \qquad (4.13)$$

где $\mathbf{b} = (b_1, b_2, \ldots, b_K) -$ $K-$мерный вектор текущих символов данных всех пользователей. Каждая сигнатура в CDMA системе с прямым расширением является сигналом, описываемым моделью

$$\dot{S}(t) = \sum_{i=0}^{N-1} a_{k,i}\dot{S}_0(t - i\Delta),$$

где $\{a_{k,0}, a_{k,1}, \ldots, a_{k,N-1}\}$ – кодовая последовательность, манипулирующая чипы $k-$ й сигнатуры, а N – коэффициент расширения, т.е. число чипов на один символ данных.

Диаграмма на рис. 4.11 демонстрирует строгое взаимное расположение между чипами сигнатуры и границами передаваемых символов данных при синхронном варианте CDMA.

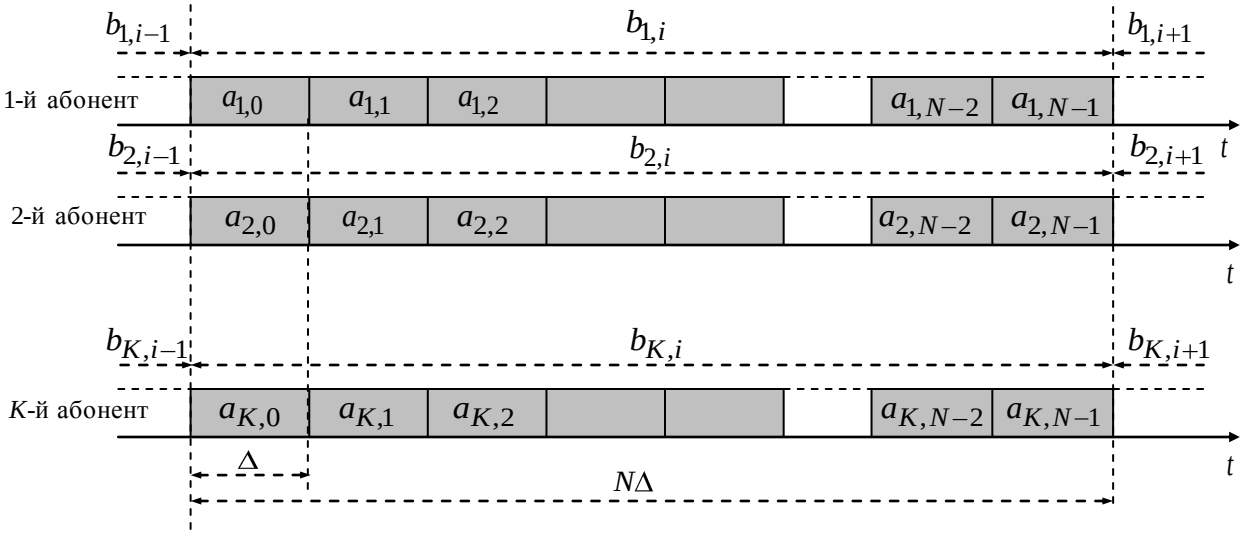


Рис. 4.11 Положение символа данных и чипа в синхронной CDMA.

Можно сформулировать несколько подходов к выбору ансамблей сигнатур для сетей с синхронным CDMA и прямым расширением спектра.

Среди основных факторов, оказывающих влияние на процедуру и результаты оптимизации множества сигнатур, следует отметить связь между числом пользователей K и показателем расширения N.

Для математического описания подобной задачи предположим, что данные k-го пользователя образуют последовательность $\mathbf{b}_k = b_{k,0}, b_{k,1}, \ldots$), где $b_{k,i}$ отвечает i-му символу потока данных k-го пользователя. Данная последовательность тем или иным способом модулирует специфический сигнал k-го пользователя $s_k\ t$, образуя модулированный сигнал $s_k\ t, \mathrm{b}_K$. При распространении по каналу каждый такой сигнал может приобрести амплитуду A_k и временное запаздывание τ_k и после суммирования с сигналами других пользователей образует общий или *групповой* сигнал, достигающий приемника, вида

$$s_k\ t; \mathrm{b}_1, \mathrm{b}_2, \ldots, \mathrm{b}_K\ = \sum_{k=1}^{K} A_k s_k\ t - \tau_k; \mathrm{b}_k\ ,$$

где K - число активных пользователей.

Очевидно, что групповой сигнал сопровождается канальным шумом $n\ t$, так что результирующее наблюдение представимо как

$$y\ t\ = s_k\ t; \mathrm{b}_1, \mathrm{b}_2, \ldots, \mathrm{b}_K\ + n\ t\ = \sum_{k=1}^{K} A_k s_k\ t - \tau_k; \mathrm{b}_k\ + n\ t\ . \qquad (4.14)$$

Приемник должен выделить данные пользователя из наблюдения $y(t)$. Ключевая роль в решении о принимаемых данных $\mathbf{b}_k\ , k = 1,2,\ldots,K$ в случае гауссовского канала принадлежит евклидову расстоянию между наблюдением $y\ t$ и различными копиями группового сигнала $s\ t; \mathrm{b}_1, \mathrm{b}_2, \ldots, \mathrm{b}_K$, соответствующими всем возможным комбинациям данных K пользователей

$$d^2(\mathbf{s}, \mathbf{y}) = \int_0^T y(t) - s(t; \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_K)\ ^2 dt\ . \qquad (4.15)$$

Подстановка (4.14) в (4.15) и раскрытие скобок приводит к соотношению

$$d^2(\mathbf{s}, \mathbf{y}) = \|\mathbf{y}\|^2 - 2\sum_{k=1}^{K} A_k z_k(\mathbf{b}_k) + \sum_{k=1}^{K} \sum_{l=1}^{K} A_k A_l \int_0^T s_k(t - \tau_k; \mathbf{b}_k) s_l(t - \tau_l; \mathbf{b}_l) dt\ ,$$

где $z_k(\mathbf{b}_k)$ - корреляция наблюдения $y\ t$ и задержанного на τ_k пользовательского сигнала, промодулированного последовательностью данных $\mathbf{b}_k$

$$z_k(\mathbf{b}_k) = \int_0^T y(t) s_k(t - \tau_k; \mathbf{b}_k) dt\ \ . \qquad (4.16)$$

Как правило, принятию решения о последовательностях данных предшествует операция оценивания интенсивности сигналов и их задержек, так что параметры $A_k, \tau_k, k = 1,2,\ldots,K$ в соотношении (4.16) можно полагать известными с достаточной точностью. Тогда оптимальная (максимально правдоподобная или по минимуму расстояния) стратегия восстановления пользовательских данных состоит в подстановке всех возможных реализаций последовательностей $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_K$ в (4.16) и выборе тех из них, которые совместно минимизируют квадрат расстояния (4.15). Данное решающее правило, называемое *многопользовательским алгоритмом*, может оказаться совершенно не практичным в ситуации, когда число пользователей K измеряется десятками или более.

Так называемый *обычный* или *однопользовательский* приемник осуществляет альтернативное решающее правило, заключающееся в раздельной оценке каждой из последовательностей данных $\mathbf{b}_k$ на основании максимизации корреляции (4.16). Очевидно, что данная стратегия совпадает с оптимальной (многопользовательской) при использовании схемы модуляции, обладающей следующими свойствами:

а) энергия пользовательских сигналов не зависит от передаваемых данных;

б) все сигналы пользователей ортогональны вне зависимости от передаваемых данных.

Поэтому этот метод множественного доступа называется ортогональным.

Для определения максимально достижимого числа пользователей в ортогональной схеме множественного доступа ограничимся M-ичной цифровой передачей данных с фиксированной скоростью R бит/сек. Учитывая, что все пользовательские сигналы должны быть ортогональны на интервале времени, равном длительности M-ичного символа, приходим к соотношению $T_t = \lfloor \log_2 M \rfloor R$. Следовательно, максимальная размерность сигнального пространства составляет $2W_t T_t = 2W_t \log_2 M \,/R$.

Приведенное соотношение дает верхнюю границу максимального числа пользователей в ортогональной схеме множественного доступа

$$K = \begin{cases} \dfrac{2W_t}{R}, & M = 2, \\ \dfrac{W_t \log_2 M}{R}, & M > 2 \end{cases} \qquad (4.17)$$

Синхронный метод CDMA (S–CDMA) достаточно легко реализуется в системах с единственным передатчиком одновременно излучающим потоки индивидуальных данных, каждый из которых адресован определенному пользователю.

4.6 Построения сигнатур для синхронных систем множественного доступа

Рассмотрим построение ансамбля бинарных сигнатур длины $N = 16 = 4^2$.

Для построения бинарных сигнатур воспользуемся матрицей Адамара вида

$$\mathbf{H}_4 = \begin{bmatrix} + & + & + & + \\ - & - & + & + \\ + & - & + & - \\ + & - & - & + \end{bmatrix},$$

содержащей одну или три +1 в своих столбцах и образуем ее кронекеровский квадрат в виде

$$\mathbf{H}_{16} = \mathbf{H}_4 \otimes \mathbf{H}_4 = \begin{bmatrix} \mathbf{H}_4 & \mathbf{H}_4 & \mathbf{H}_4 & \mathbf{H}_4 \\ -\mathbf{H}_4 & -\mathbf{H}_4 & \mathbf{H}_4 & \mathbf{H}_4 \\ \mathbf{H}_4 & -\mathbf{H}_4 & \mathbf{H}_4 & -\mathbf{H}_4 \\ \mathbf{H}_4 & -\mathbf{H}_4 & -\mathbf{H}_4 & \mathbf{H}_4 \end{bmatrix}.$$

Первичные сигнатуры представляют собой строки этой матрицы, т.е. в нормированной форме:

$$\left(\mathbf{a}_0^0, \mathbf{a}_1^0, \mathbf{a}_2^0, \mathbf{a}_3^0, \mathbf{a}_4^0, \mathbf{a}_5^0, \mathbf{a}_6^0, \mathbf{a}_7^0, \mathbf{a}_8^0, \mathbf{a}_9^0, \mathbf{a}_{10}^0, \mathbf{a}_{11}^0, \mathbf{a}_{12}^0, \mathbf{a}_{13}^0, \mathbf{a}_{14}^0, \mathbf{a}_{15}^0\right)^T =$$

$$= \frac{1}{4}\begin{pmatrix} + & + & + & + & + & + & + & + & + & + & + & + & + & + & + & + \\ - & - & + & + & - & - & + & + & - & - & + & + & - & - & + & + \\ + & - & + & - & + & - & + & - & + & - & + & - & + & - & + & - \\ + & - & - & + & + & - & - & + & + & - & - & + & + & - & - & + \\ - & - & - & - & - & - & - & - & + & + & + & + & + & + & + & + \\ + & + & - & - & + & + & - & - & - & - & + & + & - & - & + & + \\ - & + & - & + & - & + & - & + & + & - & + & - & + & - & + & - \\ - & + & + & - & - & + & + & - & + & - & - & + & + & - & - & + \\ + & + & + & + & - & - & - & - & + & + & + & + & - & - & - & - \\ - & - & + & + & + & + & - & - & - & - & + & + & + & + & - & - \\ + & - & + & - & - & + & - & + & + & - & + & - & - & + & - & + \\ + & - & - & + & - & + & + & - & + & - & - & + & - & + & + & - \\ + & + & + & + & - & - & - & - & - & - & - & - & + & + & + & + \\ - & - & + & + & + & + & - & - & + & + & - & - & - & - & + & + \\ + & - & + & - & - & + & - & + & - & + & - & + & + & - & + & - \\ + & - & - & + & - & + & + & - & - & + & + & - & + & - & - & + \end{pmatrix}.$$

Используя приведенное правило можно построить ортогональные сигнатуры больших размерностей.

4.7 Асинхронный метод множественного доступа

Во многих приложениях типичной является ситуация, когда задержка τ_k может изменяться в широком диапазоне, делая процедуру синхронизации сигнатур на входе приемника проблематичной или вообще невозможной.

Распространим задачу синтеза сигнатур на случай асинхронного CDMA с прямым расширением спектра, в котором временная задержка и сдвиг начальной фазы между отдельными пользовательскими сигналами случайны.

Можно допустить, что границы чипов всех K сигнатур синхронизированы, т.е. взаимные задержки кратны Δ: $\tau_l = n_l \Delta$, где n_l – целое число, такое что $0 \le n_l < N$. Тогда ситуация хорошо поясняется рис. 4.12, подчеркивающим, что в асинхронном варианте CDMA в отличие от синхронного (см. рис. 4.11) символы данных других пользователей могут изменяться в течение приема текущего символа k–го пользователя.

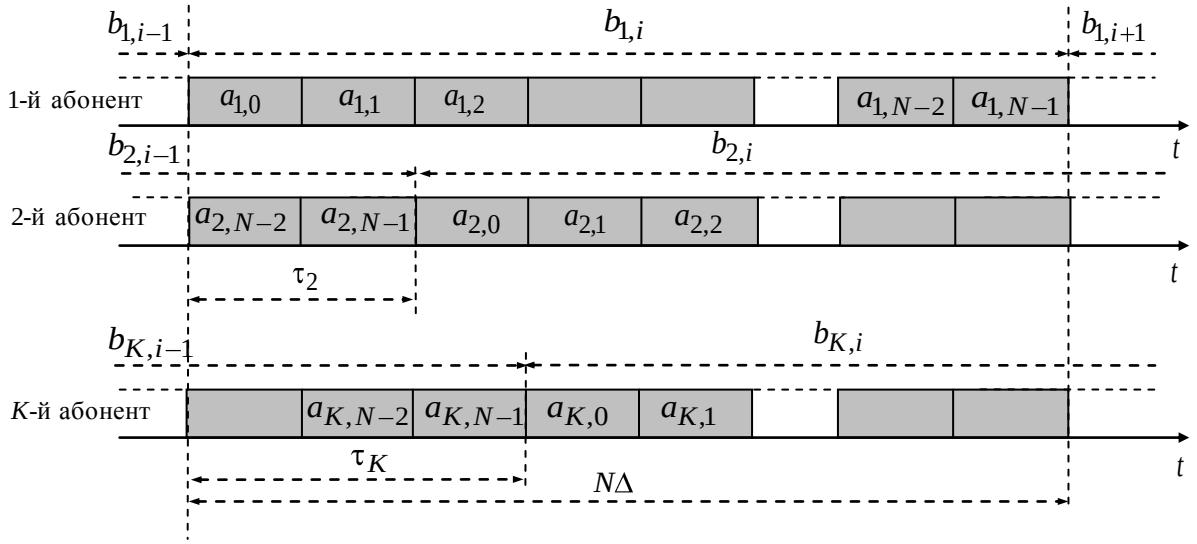


Рис. 4.12. Поток данных и сигнатуры в асинхронном CDMA.

Еще одним фактором, затрудняющим синтез множества асинхронных сигнатур, является необходимость различения каждой сигнатуры относительно всех возможных сдвинутых копий других сигнатур, что излишне при синхронном CDMA.

Проанализируем последствия асинхронного характера принимаемых сигналов пользователей.

Рассмотрим два сигнала $u\,t$ и $v\,t$, и вычислим их взаимную корреляционную функцию (ВКФ)

$$R_{uv}(\tau) = \int_{-\infty}^{\infty} u(t)v(t-\tau)dt\,.$$

Применение теоремы Парсеваля дает

$$R_{uv}(\tau) = \int_{-\infty}^{\infty} \tilde{u}(f)\tilde{v}^*(f)\exp(-j2\pi f\tau)df\,. \qquad (4.18)$$

Если целью является достижение ортогональности сигналов вне зависимости от взаимной задержки, то при всех значениях τ должно выполняться равенство $R_{uv}(\tau) = 0$, что, вследствие линейности преобразования Фурье, возможно, если и только если $\tilde{u}(f)\tilde{v}(f) = 0$ во всей частотной области.

Последнее утверждение свидетельствует о том, что два сигнала

ортогональны при любом произвольном временном сдвиге, если и только если их спектры не перекрываются. Однако схема множественного доступа с не перекрывающимися спектрами есть FDMA! Следовательно, асинхронный вариант ортогонального множественного доступа реализуем только в виде FDMA.

Однако какова плата за попытку реализации CDMA при отсутствии синхронизации сигнатур на входе приемника? Поскольку сигнатуры различных абонентов в рамках CDMA обладают перекрывающимися спектрами, то они не могут оставаться ортогональными в широком диапазоне взаимных задержек. Следствием этого является возникновение межпользовательского мешающего воздействия, проявлением которого служит ненулевой отклик приемника, настроенного на k -го пользователя, от сигналов других абонентов.

Рассмотрим обычный приемник k -го пользователя. Без потери общности можно считать, что $\tau_k = 0$, переписав корреляцию наблюдения (4.16) как

$$z_k(\mathbf{b}_k) = \int_0^T y(t) s_k(t;\mathbf{b}_k)dt . \qquad (4.19)$$

Согласно правилу одиночного наблюдателя оценка $\hat{\mathbf{b}}_k$ данных $\mathbf{b}_k$ должна максимизировать решающую статистику $z_k(\mathbf{b}_k)$ как функцию $\mathbf{b}_k$. Подставив (4.14) в (4.19), выразим $z_k(\mathbf{b}_k)$ в виде

$$z_k(\mathbf{b}_k) = A_k \int_0^T s_k(t;\mathbf{b}'_k) s_k(t;\mathbf{b}_k)dt + \sum_{\substack{l=1 \\ l\neq k}}^{K} A_l \int_0^T s_l(t-\tau_l;\mathbf{b}'_l) s_k(t;\mathbf{b}_k)dt + \int_0^T n(t) s_k(t;\mathbf{b}_k)dt , \quad (4.20)$$

где осуществлена замена $\mathbf{b}_k$ в (4.14) на $\mathbf{b}'_k$ с тем, чтобы отличить подлинно передаваемые данные $\mathbf{b}'_k$ от предполагаемых в процессе принятия решения относительно $\mathbf{b}_k$.

Первый и последний члены в (4.20) отвечают вкладу собственного, т.е. k-го сигнала пользователя и аддитивного шума в отклик приемника k-го пользователя. В случае отсутствия сторонних абонентов $(K=1)$ второе слагаемое равнялось бы нулю. При $K>1$ и произвольных запаздываниях сигнатур указанное слагаемое отлично от нуля, определяя вклад других пользовательских сигналов в выходной эффект k-го приемника, т.е. взаимную помеху или *помеху множественного доступа* (MAI).

Простейшим путем оценки влияния MAI является представление всех чужих сигналов в виде случайных шумоподобных процессов.

В любой реальной асинхронной CDMA системе должны быть предприняты меры по выравниванию уровней всех пользовательских сигналов на входе приемника с тем, чтобы избежать проблемы *близости–дальности*. Поэтому можно полагать, что в результате эффективной *регулировки мощностью* все амплитуды $A_k, k=1,2,\ldots,K$ одинаковы или мощности всех

сигналов идентичны и равны P. Тогда шумоподобный сигнал l-го стороннего пользователя, энергия которого полагается равномерно распределенной в полосе W, создаст избыточную спектральную плотность мощности шума $N_l = P/W$, которая добавится к спектру шума. Полная спектральная плотность MAI составит $N_I = (K-1)N_l = (K-1)P/W$. Теперь можно записать отношение сигнал-помеха по мощности (SIR), учитывающее как MAI, так и тепловой шум

$$q_I^2 = \frac{2E}{N_0 + N_I} = \frac{2E}{N_0 + (K-1)\dfrac{P}{W}}. \qquad (4.21)$$

Если пренебречь тепловым шумом, то можно записать

$$q_{If}^2 = \frac{2E}{(K-1)\dfrac{P}{W}} = \frac{2PT}{(K-1)\dfrac{P}{W}} = \frac{2WT}{K-1}. \quad . \qquad (4.22)$$

Последний результат показывает, что пороговое отношение SIR и, следовательно, пороговая достоверность приема полностью определяется частотно-временным произведением, т.е. выигрышем от широкополосной обработки WT, и числом пользователей. Соотношение (4.22) дает возможность оценить максимальное число пользователей, которое может обслужить асинхронная CDMA система при заданном частотно-временном ресурсе WT

$$K \le \frac{2WT}{q_I^2} + 1, \qquad (4.23)$$

где q_I^2 отвечает необходимому отношению SIR, гарантирующему требуемую достоверность приема в анализируемой системе.

4.8 Построения сигнатур для асинхронных систем множественного доступа

Детерминированный сигнал, для которого произведение длительности на полосу частот $B = WT >> 1$ и полосой которого можно управлять независимо от длительности, называется сигналом с *расширенным спектром* или *широкополосным* (ШПС). Другими словами, энергия любого широкополосного сигнала распределена в прямоугольнике на плоскости время-частота, площадь которого значительно больше единицы. Величина B, называемая *базой сигнала*, определяет важнейшие характеристики радиотехнических систем.

Распределение спектра сигнала осуществляется посредством соответствующего управления комплексной огибающей сигнала $\dot{S}(t)$. 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 сигналов, учитывая их доминирующую роль в большинстве современных и проектируемых цифровых системах.

Дискретный сигнал представляет собой последовательность элементарных символов (импульсов) фиксированной формы, повторяющихся с некоторым фиксированным временным интервалом Δ. Комплексная огибающая $\dot{S}_0(t)$, определяющая форму элементарного импульса, и внутренняя его угловая модуляция могут быть любыми. Элементарный импульс называют чипом.

Модуляция всего сигнала заключается в манипулировании амплитудами, фазами и частотами отдельных чипов. Тогда формальное представление комплексной огибающей дискретного сигнала дается соотношением

$$\dot{S}\ t\ =\ \sum_{i=-\infty}^{\infty} a_i \dot{S}_0\ t-i\Delta\ \exp\ j2\pi F_i t\ , \qquad (4.24)$$

где a_i и F_i - соответственно комплексная амплитуда и частота (в значениях сдвига относительно фиксированной центральной частоты) i–го чипа. Очевидно, что последовательность $\left|a_i\right|, i=\ldots,-1,0,1,\ldots$ определяет вещественные амплитуды чипов, т.е. их амплитудную модуляцию. Аналогично, последовательности $\varphi_i=\arg a_i, i=\ldots,-1,0,1,\ldots$ и $\{F_i, i=\ldots,-1,0,1,..\}$ задают законы модуляции чипов по фазе и частоте.

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

1. Если манипуляции подвергаются только комплексные амплитуды чипов, а все частоты остаются одинаковыми $(F_i=0, i=0,1,\ldots,N-1)$, то сигнал называется *амплитудно–фазоманипулированным* (АФМ или APSK) Последовательность комплексных амплитуд чипов $\{a_i, i=0,1,\ldots,N-1\}$ называется *кодовой последовательностью* или просто *кодом*.

2. Если манипуляция осуществляется только над фазами φ_i чипов сигнала, а амплитуды остаются неизменными, то сигнал является

фазоманипулированным (ФМ или PSK).

3. Если применяются только бинарные комплексные амплитуды $a_i = \pm 1,\ |a_i| = 1,\ \varphi_i \in \{0, \pi\}$, то сигнал называется *бинарным фазоманипулированным* (BPSK).

4. При четверичном алфавите вида $a_i = \pm 1, \pm j$, или, что эквивалентно, $|a_i| = 1, \varphi_i \in \{0, \pi, \pm \pi / 2\}, i = 0,1,\ldots, N-1$ сигнал является квадратурным ФМ (QPSK), и т.д.

6. 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 $F_i, i = 0,1,\ldots, N-1$. Сигналы этого типа используются, в частности, в системах с прыгающей частотой.

Другой важной версией является сигнал, у которого закон модуляции чипов повторяется с периодом N. Дискретный сигнал подобного типа называется *периодическим*. Фактический период сигнала составляет $T = N \cdot \Delta$.

4.8.1. Бинарные кодовые М-последовательности.

В настоящем разделе рассмотрим основные подходы к синтезу кодовых последовательностей, основанных на теории конечных полей Галуа.

Ограничимся рассмотрением метода построения двоичных кодов, основанного на свойствах рекуррентных последовательностей с элементами из конечного поля $GF(2)$. Обратимся к рис. 4.13, демонстрирующему структуру типичного генератора линейной рекуррентной последовательности.

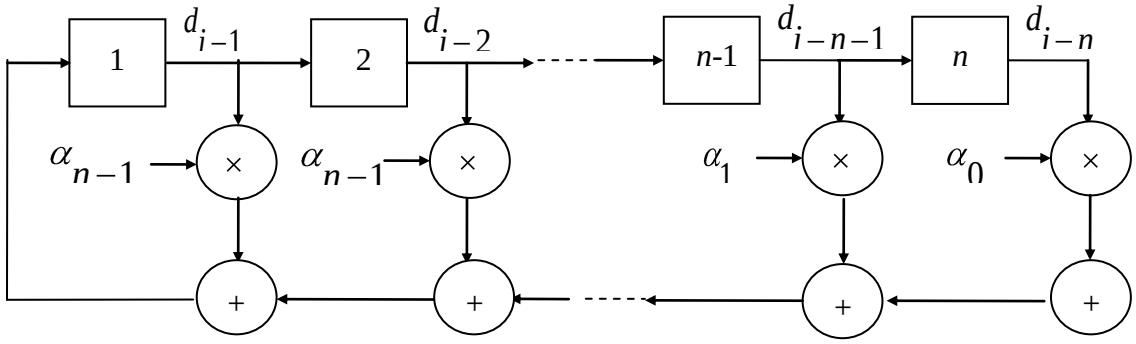


Рис. 4.13. Генератора линейной рекуррентной последовательности

В соответствие со своей структурой данная схема называется *регистром сдвига с линейной обратной связью*. Регистр состоит из n 2-ичных элементов задержки, изображенных в виде квадратов. Каждый элемент задержки принимает 2 возможных состояний и хранит некоторый элемент поля $GF(2)$ в течение тактового интервала времени.

Пусть первоначально в регистр занесен некоторый n-разрядный

ненулевой двоичный код d_i, $i = 0 \ldots n-1$ принимающий значения из множества $0,1$. Под действием каждого тактового импульса состояние любого разряда передается следующему слева направо.

Схема обратной связи включает умножители элементов на константы α_i и сумматоры. Обе арифметические операции выполняются по правилам сложения и умножения конечного поля $GF(2)$.

Линейная рекуррентная последовательность d_i, может быть непосредственно считана с крайнего правого разряда, начиная с самого первого символа d_0 или с любого другого разряда, но с соответствующим сдвигом.

Правило формирования двоичной последовательности d_i подчиняются линейной рекурсии вид

$$d_i = \alpha_{n-1} \cdot d_{i-1} \oplus \alpha_{n-2} \cdot d_{i-2} \oplus \ldots \oplus \alpha_0 \cdot d_{i-n}, \quad i = n, \; n+1, \ldots, \qquad (4.25)$$

где коэффициенты $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$ - фиксированные константы, принадлежащие полю $GF(2)$. Подобная последовательность называется *линейной рекуррентной последовательностью.*

Поскольку число различных состояний регистра конечно и не более 2^n, то неизбежна ситуация, когда после некоторого числа тактов состояние повторится в виде одного из ранее случившихся. В общем случае значение периода $N \le 2^n - 1$ зависит от *начального состояния регистра и его коэффициентов.*

Линейные рекуррентные последовательности, обладающие наибольшим периодом $N = 2^n - 1$, представляют особый интерес и называются последовательностями *максимальной длины* или просто M-последовательностями.

Для формирования M-последовательностей структура обратных связей генератора линейной рекуррентной последовательности задается примитивным полиномом степени n

$$d\ x = \alpha_n x^n + \alpha_{n-1} x^{n-1} + \ldots + \alpha_1 x + \alpha_0 . \qquad (4.26)$$

Примитивные полиномы можно взять из таблиц, приведенных в научно-технической литературе. В таблице 4.1 приведены коэффициенты некоторых примитивных полиномов степени $n = 5 \ldots 10$.

Табл. 4.1

Полиномы построения М-последовательностей

$n := 5$	$\alpha := (0\ 0\ 1\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0)^T$	$n := 8$	$\alpha := (0\ 1\ 1\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0)^T$
$n := 6$	$\alpha := (0\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0)^T$	$n := 9$	$\alpha := (0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0)^T$
$n := 7$	$\alpha := (0\ 0\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0)^T$	$n := 10$	$\alpha := (0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0)^T$

Далее перейдем к рассмотрению бинарных M - последовательностей. Пусть определена двоичная M - последовательность $\{d_i \in 0,1\}$ памяти n, т.е. длины $N = 2^n - 1$. Отобразим ее символы 0 и 1 в бинарный алфавит ± 1 согласно правилу

$$a_i = (-1)^{d_i} = \begin{cases} +1, & d_i = 0 \\ -1, & d_i = 1 \end{cases}. \qquad (4.27)$$

Полученная в результате подобного преобразования последовательность $\{a_i \in \pm 1\}$ вещественных бинарных символов ± 1 называется бинарной M - последовательностью. Очевидно бинарная M - последовательность также обладает периодом $N = 2^n - 1$.

В качестве примера построим бинарную M -последовательность памяти $n = 3$ при условии, что коэффициентами рекурсии являются $\alpha_2 = 0,\ \alpha_1 = 1,\ \alpha_0 = 1$. Задав исходные элементы $d_0 = 1,\ d_1 = 0,\ d_2 = 0$, получаем исходную двоичную последовательность d_i вида

$$1,0,0,1,0,1,1,1,0,0,1,0,1,1,\ldots.$$

Выполнив над элементами d_i преобразование (4.27), получаем бинарную M - последовательность a_i:

$$-1,1,1,-1,1,-1,-1,-1,1,1,1,-1,1,-1,-1,-1,..$$

Легко проверить, что ненормированная периодическая АКФ бинарной M - последовательности является одноуровневой

$$R_p(m) = \sum_{i=0}^{N-1} a_i \cdot a_{i-m} = \begin{cases} N, & m = 0 \bmod N, \\ -1, & m \neq 0 \bmod N. \end{cases} \qquad (4.28)$$

Бинарные M - последовательности относятся к числу наиболее популярных дискретных сигналов, применяемых в современных системах множественного доступа.

Рассмотрим основные структурные свойства М последовательностей.

Свойство 1. Произведение двух бинарных M -последовательностей $\{u_i\}$ и $\{v_i\}$ периода $N = 2^n - 1$ дает снова M -последовательность $\{z_i\} = \{u_i\} \cdot \{v_i\}$.

Свойство 2. Если бинарную M -последовательность $\{u_i\}$ периода $N = 2^n - 1$ подвергнуть операции *децимирования* с *индексом децимации* d, где d взаимно простое с N, то результирующая последовательность $v_i = u_{d\,i}$ снова окажется бинарной M -последовательностью такого же периода. Операция децимирования означает выбор каждого d-го символа последовательности $\{u_i\}$. Назовем полученную в результате указанной процедуры последовательность $\{v_i\}$ децимацией последовательности $\{u_i\}$.

Свойство 3. Пусть память n бинарной M -последовательности $\{u_i\}$ является *нечетным числом*, а индекс децимации $d = 2^s + 1$, где s взаимно просто с n. Тогда d будет взаимно простое число с длиной $N = 2^n - 1$ последовательности $\{u_i\}$. Децимированная последовательность $\{v_i\}$ является M - последовательностью того же периода N.

Ненормированная периодическая $R_{p,uv}(m)$, $m = 0,1,\ldots,N-1$ последовательностей $\{u_i\}$ и $\{v_i\}$ принимает только три следующих значения

$$R_{p,uv}(m) \in \pm\sqrt{2(N+1)} - 1, -1 = \{\pm 2^{n+1/2} - 1, -1\} . \qquad (4.29)$$

Свойство 4. Пусть память n бинарной M - последовательности $\{u_i\}$ является *четным числом*, но не кратно четырем, а индекс децимации $d = 2^s + 1$, где s - четное и взаимно простое с $n/2$. Тогда d взаимно просто с длиной $N = 2^n - 1$ последовательности. Децимированная последовательность является M - последовательностью того же периода N.

Ненормированная периодическая ВКФ $R_{p,uv}(m)$, $m = 0,1,\ldots,L-1$ последовательностей $\{u_i\}$ и $\{v_i\}$ принимает только три следующих значения

$$R_{p,uv}(m) \in \pm 2\sqrt{(L+1)} - 1, -1 = \{\pm 2^{n+2/2} - 1, -1\} . \qquad (4.30)$$

К достоинствам *М*-последовательностей относятся хорошие корреляционные свойства и простота генерирования. Поэтому *М*-последовательности используются как исходные для получения больших систем квазиортогональных сигналов.

4.8.2 Ансамбли сигнатур, получаемых на основе М-последовательностей.

Как отмечалось ранее, в многопользовательской сети каждый из K абонентов передает или получает свои индивидуальные данные посредством использования некоторой специфической для пользователя сигнатуры, причем тщательно выбранный ансамбль из K сигнатур обеспечивает максимально возможную совместимость абонентов (минимальную взаимную помеху множественного доступа).

Ансамбли сигнатур, получаемые временным сдвигом М-последовательностей.

Во многих реальных ситуациях взаимные временные сдвиги асинхронных сигнатур могут изменяться только в пределах ограниченного диапазона задержек $\tau_{\max} << T_c$, где T_c - длительность кодового сигнала.

Выразим величину максимальной задержки $\tau_{\max}$ в округленном до целого числе $s_{\max}$ чипов: $s_{\max} = \left[\tau_{\max}/\Delta\right]$. Тогда диапазон значений s, определяется как $[-s_{\max}, s_{\max}]$.

Возьмем теперь кодовую последовательность $\{z_i\}$ периода $N \geq K(s_{\max}+1)$ и используем в качестве K сигнатур ее циклически сдвинутые одна относительно другой копии на величину $s_{\max}+1$ позиций

$$a_{k,i} = z_{i-(k-1)(s_{\max}+1)},\ k = 1,2,\ldots,K;\ i = \ldots,-1,0,1,\ldots. \qquad (4.31)$$

Очевидно, что все корреляции между образованными таким образом сигнатурами будут выражены в терминах АКФ $\rho_z(m)$ исходной последовательности $\{z_i\}$. Оценка ВКФ $k-$й и $l-$й сигнатур приводит к следующему результату

$$\rho_{kl}(m) = \frac{1}{N} = \sum_{i=0}^{N-1} a_{k,i} a^*_{l,i-m} = \frac{1}{N}\sum_{i=0}^{N-1} a_{1,i-(k-1)(s_{\max}+1)} a^*_{1,i-(l-1)(s_{\max}+1)-m},$$

или

$$\rho_{kl}(m) = \rho_z[(k-l)(s_{\max}+1)+m]. \qquad (4.32)$$

Таким образом, показано, что соответствующие сдвиги копий исходной последовательности с хорошей периодической АКФ могут образовывать

кодовый ансамбль.

Бинарная M- последовательность чрезвычайно большой длины $N = 2^{42} - 1$, расширенная еще на один символ, используется в качестве исходной, а конкретные пользовательские сигнатуры всех мобильных станций представляют собой ее соответствующие циклические копии. Так обратные каналы стандартов 2-го поколения cdmaOne (IS-95) и 3-го поколения cdma2000 представляют хороший пример реализации данной версии асинхронного варианта CDMA.

Ансамбли сигнатур, получаемые частотным сдвигом М-последовательностей.

В качестве исходной кодовой последовательности возьмем бинарную M-последовательность $\{z_i \in \pm 1\}$ периода $N = 2^n - 1$ и используем ее для построения ансамбля из K сигнатур. Примем, что первая кодовая последовательность $\{a_{1,i}\} = z_i$, $i = 0,\ldots,N-1$. Остальные $K-1$ кодовых последовательностей формируются посимвольным умножением $\{a_{1,i}\}$ на дискретные гармоники частот $(k-1)/N$, $k = 1,2,3\ldots,K$:

$$a_{k,i} = a_{1,i} \cdot \exp\left(j\frac{2\pi(k-1)i}{N} \right),\ i = \ldots,-1,0,1,\ldots;\ k = 1,2,\ldots,K. \qquad (4.33)$$

Произведем анализ характеристик взаимной корреляции кодовых последовательностей

Квадрат модуля периодической ВКФ *k*–й и *l*–й последовательностей запишется как

$$\left|R_{p,kl}(m)\right|^2 = \left|\sum_{i=0}^{N-1} a_{k,i} \cdot a^*_{l,i-m}\right|^2 = \left|\sum_{i=0}^{N-1} a_{1,i} \cdot a_{1,i-m} \exp\left(j\frac{2\pi(k-l)i}{N} \right)\right|^2. \qquad (4.34)$$

Первоначально рассмотрим случай $m = 0 \mod N$, т.е. $a_{1,i} \cdot a_{1,i-m} = \left|a_{1,i}\right|^2 = 1$. Тогда, если $k = l$, соотношение (4.21) определяет основной лепесток АКФ *k*–й сигнатуры, т.е. $\left|R_{p,kk}(0)\right|^2 = N^2$. Если же $k \neq l$, то сумма в (4.21) есть сумма всех корней из единицы степени *N*, которая равняется нулю. Теперь пусть $m \neq 0 \bmod N$. Сумма в (4.21) отличается от нуля и равна *N* только при $k = l$.

Таким образом, перейдя к нормированным корреляциям, получаем

$$\left|\rho_{p,kl}(m)\right|^2 = \begin{cases} 1,\ k=l,\ m=0 \bmod N, \\ \dfrac{1}{N},\ k=l,\ m\neq 0 \bmod N, \\ 0,\ k\neq l,\ m=0 \bmod N, \\ \dfrac{N+1}{N^2},\ k\neq l,\ m\neq 0 \bmod N. \end{cases} \quad (4.35)$$

Откуда видно, что квадрат максимального корреляционного пика ансамбля кодовых последовательностей определяется соотношением

$$\rho^2_{\max} = N+1 \,/ N^2 \approx 1/N, \quad (4.36)$$

т.е. практически совпадает с границей Велча.

Таким образом, рассматриваемый ансамбль кодовых последовательностей является минимаксным, оптимальным образом реализуя кодовое разделение сигналов.

Ансамбли кодовых последовательностей Голда.

Рассмотрим правило построения кодовых последовательностей Голда. Возьмем *M*-последовательность $\{u_i\}$ периода N, и ее децимацию, удовлетворяющую вышеприведенным свойствам *3* или *4*.

Обладая парой кодовых последовательностей $\{u_i\}$ и $\{v_i\}$, сформируем ансамбль из K кодов по следующему правилу

$$\begin{aligned} a_{k,i} &= u_i \cdot v_{i-k}, \quad k = 1,2,\ldots,N \\ a_{N+1,\ i} &= u_i \\ a_{N+2,i} &= v_i \end{aligned} \quad (4.37)$$

Таким образом, построение ансамбля кодов производится путем посимвольного умножения $\{u_i\}$ на циклические копии $\{v_i\}$, а еще два кода образуют непосредственно сами исходные M - последовательности. В итоге имеем всего $K = N+2 = 2^n+1$ кодовых последовательностей.

Можно показать, что корреляционный пик множества кодов Голда определяется максимальным по модулю значением ВКФ. После нормировки к длине N приходим к следующей оценке

$$\rho_{\max}^2 \le \begin{cases} \dfrac{\sqrt{2(N+1)}+1}{N^2}^2, & n \neq 0 \mod 2, \\ \dfrac{2\sqrt{(N+1)}+1}{N^2}^2, & n = 2 \mod 4 \end{cases} \approx \begin{cases} \dfrac{2}{N}, & n \neq 0 \mod 2 \\ \dfrac{4}{N}, & n = 2 \mod 4 \end{cases}. \qquad (4.38)$$

Рис. 4.14 иллюстрирует практическую реализацию алгоритма Голда в соответствие с вышеприведенным правилом.

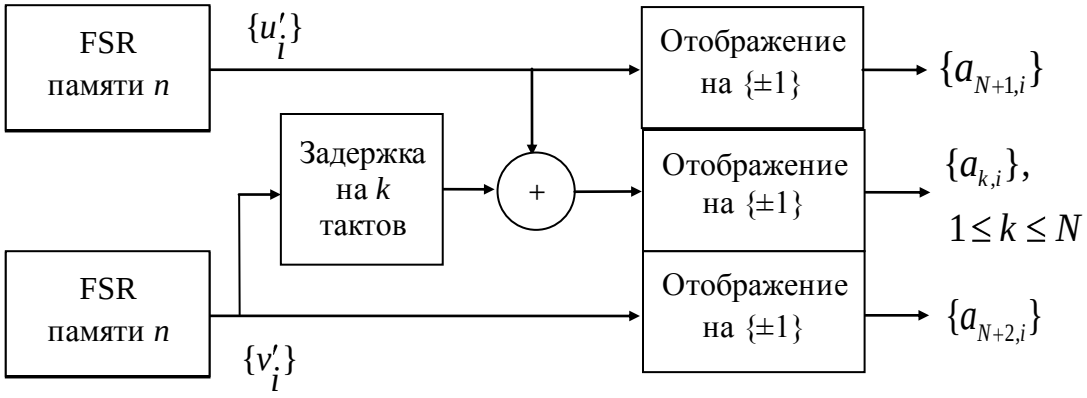


Рис. 4.14. Схема формирования кодов Голда.

В качестве примера построим последовательности Голда длины $N = 2^3 - 1 = 7$. Начнем с двоичной $\{0,1\}$ M - последовательности: $\{u_i'\} = \{1,0,0,1,0,1,1\}$. Индекс децимации $d = 3$ удовлетворяет вышеприведенному ограничению в свойстве 3. Тогда получаемая децимированная последовательность $\{v_i\} = \{1,1,1,0,1,0,0\}$.

Посимвольное суммирование $\{u_i'\}$ и $\{v_i'\}$ по модулю два дает последовательность $\{0,1,1,1,1,1,1\}$, которая после отображения на алфавит $\{\pm 1\}$ дает первую последовательность Голда $\{a_{1,i}\} = \{+------\}$. Сдвиг $\{v_i'\}$ вправо на одну позицию и сложение по модулю 2 с $\{u_i'\}$ дает последовательность $\{1,1,1,0,0,0,1\}$, которая после перехода к символам $\{\pm 1\}$ дает вторую последовательность Голда $\{---+++-\}$. Еще шесть последовательностей Голда получаются в результате дальнейших сдвигов $\{v_i'\}$, сложения по модулю 2 с $\{u_i'\}$ и изменения символов на $\{\pm 1\}$. Совместно с $\{u_i'\}$ и $\{v_i'\}$, преобразованных в $\{\pm 1\}$ последовательности, получаем всего $K = 2^3 + 1 = 9$ последовательностей.

Ансамбли Голда пользуются большой популярностью в современных CDMA системах. Достаточно сказать, что они используются в глобальной спутниковой навигационной системе GPS NAVSTAR для разделения сигналов. спутников и т.п.

5 ЭФФЕКТИВНЫЕ МОДУЛЯЦИОННЫЕ ФОРМАТЫ СИГНАЛОВ

В технике связи при регламентации использования радиоспектра, измеряют полосу сигнала шириной окна Δf_{99}, в котором удерживается не менее 99% излучаемой мощности.

В простейшем случае бинарной фазовой модуляции - БФМ реализуется наиболее помехоустойчивый способ передачи двоичной информации. Однако в отношении расходования частотного ресурса БФМ оказывается чрезвычайно неэффективной. 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 $1/f_0^2$. Для БФМ получится цифра $\Delta f_{99} \approx 18.5/T$, многократно превосходящая традиционный ориентир 1/T. По этой причине БФМ с прямоугольными посылками не применяется в цифровой связи.

Для повышения спектральной эффективности существует несколько путей модуляции сигналов. Следует упомянуть, что в русской терминологии радиосвязи существует традиция использовать для модуляции цифровым сигналом термин «манипуляция».

Существуют три фундаментальных типа цифровой манипуляции и один гибридный:

ASK – Amplitude shift keying (АМ, АМн – амплитудная манипуляция)
FSK – Frequency shift keying (ЧМ, ЧМн – частотная манипуляция).
PSK – Phase shift keying (ФМ, ФМн – фазовая манипуляция).
ASK/PSK – комбинированный способ модуляции.

5.1 Квадратурная фазовая манипуляция

Простейший путь повышения эффективности модуляции сигналов состоит в увеличении длительности прямоугольной посылки T с сохранением прежней скорости передачи

$$R = k/T = 1/T \ \log_2 M \ , \qquad (5.1)$$

где k бит определяют символ из $M = 2^k$ символьного алфавита,

T - длительность k-битового символа.

При бинарной фазовой манипуляции (БФМ) один бит передается за время $T_b = T$, так что $R = 1/T_b$. 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. Это может быть достигнуто за счет использования четырех разрешенных значений фазы вместо двух, например *0, π, π/2, –π/6*. Такой способ манипуляции и называют *квадратурной ФМ* (КФМ, англ..(QPSK).

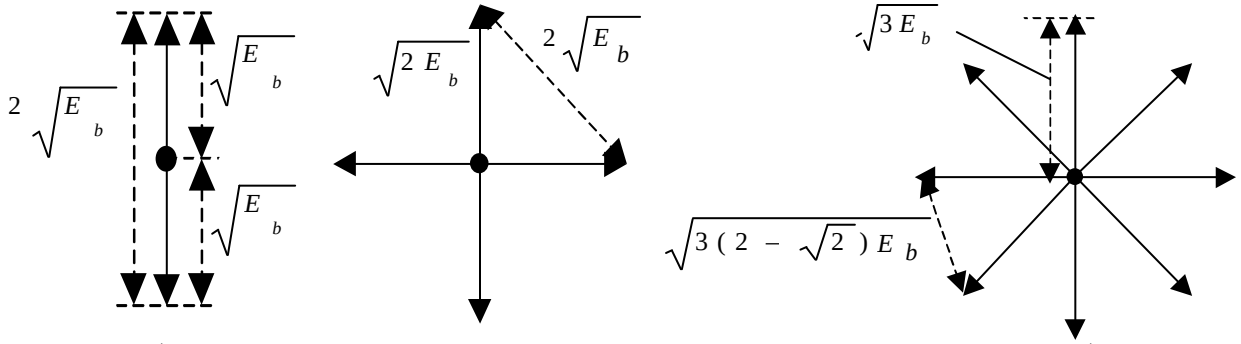
Общая модель сигнала описывается последовательностью одинаковых по форме импульсов (посылок), повторяющихся с постоянным интервалом T

$$s(t) = \sum_{i=-\infty}^{\infty} S_0\ t - iT\ \cos\ 2\pi f_0 t + \varphi_i\ , \qquad (5.2)$$

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

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. 5.1, а).



а) б) в)

Рис. 5.1. Геометрическая интерпретация фазовой манипуляции.

При КФМ четырем сообщениям отвечает четверка биортогональных векторов длины $\sqrt{2E_b}$ (рис. 5.1, 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. 5.1, 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).

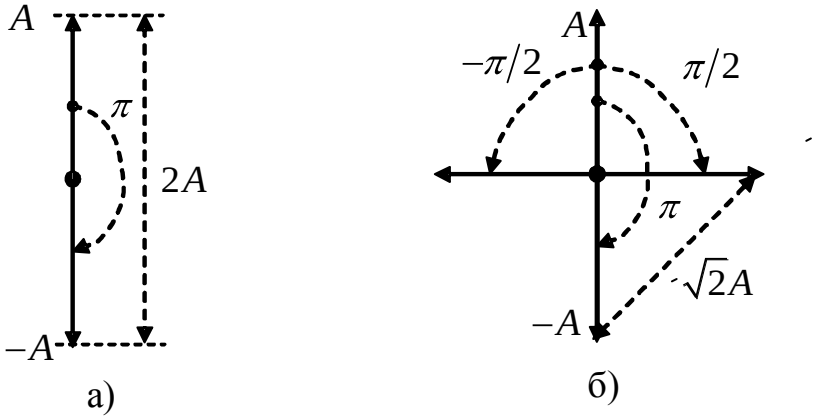
Дальнейшие повышение спектральной эффективности подобным способом окажется тем более невыгодным в плане энергозатрат: M - кратный выигрыш в полосе при $2M$-ичной ФМ будет сопровождаться энергетическим проигрышем в

$$\gamma = \frac{2}{M\left(1-\cos\frac{\pi}{2^M-1}\right)} \qquad (5.3)$$

раз.

5.2 Квадратурная фазовая манипуляция со сдвигом

Обратимся к рис. 5.2. Когда значения передаваемых символов 0 и 1 при БФМ (рис. 5.2, *а*) равновероятны, в среднем в половине переходов между последовательными посылками наблюдается скачок фазы на 180°, т.е. смена полярности на противоположную. Переход от БФМ к стандартной КФМ (рис 6.2, *б*) с сохранением прежней мощности не ослабляет требований к динамическому диапазону усилителя, поскольку максимальные скачки фазы в 180° сохраняются, хотя и случаются примерно вдвое реже.



а) б)

Рис. 5.2. Перепады мгновенной амплитуды при различных видах ФМ.

Смягчение требований к динамическому диапазону передатчика достигается в формате КФМ *со сдвигом* (КФМС, англ. *offset* QPSK).

Для понимания смысла которого разумно сначала ввести в выражение (5.2) начальную фазу в π/4 (разумеется, это никак не повлияет на свойства КФМ и лишь повернет целиком диаграмму рис. 5.2, *б* на 45°), а затем переписать результат в форме

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

$$= \frac{1}{\sqrt{2}}\left[\sum_{i=-\infty}^{\infty} a_i S_0\ t - iT\ \cos\ 2\pi f_0 t\ + \sum_{i=-\infty}^{\infty} b_i S_0\ t - iT\ \sin\ 2\pi f_0 t\ \right], \qquad (5.4)$$

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

$$s(t) = \frac{1}{\sqrt{2}}\left[\sum_{i=-\infty}^{\infty} a_i S_0\ t - iT\ \cos\ 2\pi f_0 t\ + \sum_{i=-\infty}^{\infty} b_i S_0\left(t - iT - \frac{T}{2}\right)\sin\ 2\pi f_0 t\ \right], \qquad (5.5)$$

сохраняющему ортогональность квадратурных компонент и, следовательно, все характеристики обычной КФМ в части качества передачи данных.

Теперь, однако, изменения бинарных символов, манипулирующих квадратурные несущие, происходят не одновременно: 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. 5.2, 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.

Разумеется, на приемной стороне этот сдвиг легко учитывается, так что демодуляция выполняется практически так же, как при обычной КФМ.

Подобный вид КФМ получил название π/4-КФМ (π/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 π/4-КФМ не столь эффективна: переходам на угол ±3π/4 соответствует перепад комплексной огибающей,

равный $\sqrt{2+\sqrt{2}}A$ и примерно в 1,31 раза превышающий аналогичный показатель КФМС.

5.3 Квадратурная амплитудная манипуляция

В квадратурной амплитудной манипуляции (КАМ) для одновременной передачи двух отдельных *k*-битовых информационных блоков используется две несущих, находящихся в квадратуре $(\cos(\omega_c t);\ \sin(\omega_c t))$.

Временные диаграммы, иллюстрирующие квадратурную амплитудную манипуляцию, приведены на рис. 5.3. Сигнал исходных передаваемых данных разделяется на два потока - s_I и s_Q (рис. 5.3, а) и 5.3, б). Сигнал s_Q модулирует синусоидальное несущее колебание, а сигнал s_I - косинусоидальное (рис. 5.3, в и 5.3, г). Затем два модулированных колебания складываются, образуя единое квадратурно-модулированное колебание (рис. 5.3, д).

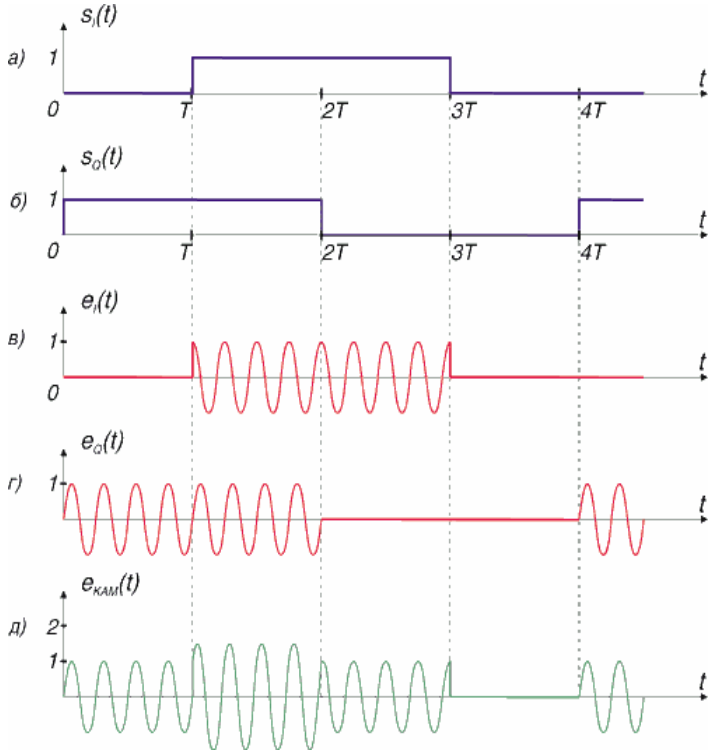


Рис. 5.3. Диаграммы сигналов при квадратурной амплитудной манипуляции.

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.5.4).

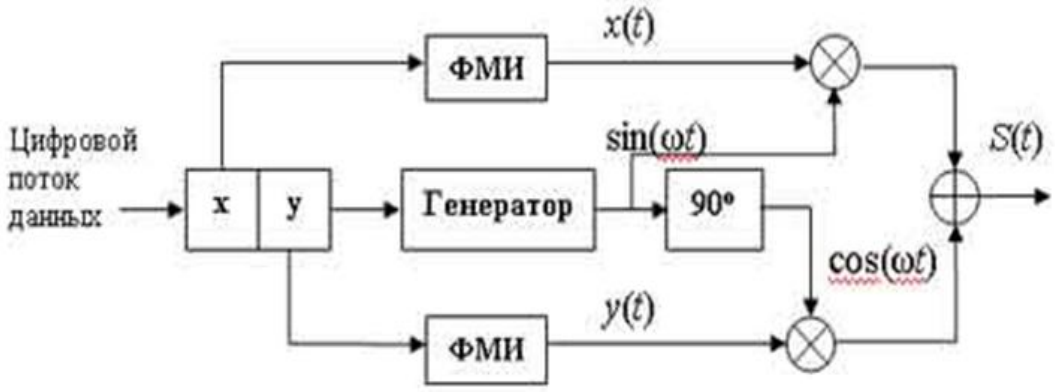
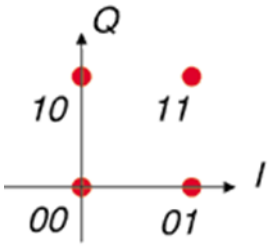


Рис. 5.4. Схема квадратурного модулятора.

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 ($\cos(\omega_c t)$) и синфазного ($\sin(\omega_c t)$) 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 канальных перемножителей, где они независимо друг от друга модулируют по амплитуде два одинаковых несущих колебания, сдвинутых по фазе на 90^o, т.е. находящихся в квадратуре. В результате, на их выходах формируются двухфазные (0, π) ФМ колебания. После суммирования они образуют сигнал ФМ-4 или квадратурный ФМ-сигнал (QPSK). Поскольку в каждом канале осуществляется амплитудная манипуляция, этот вид модуляции называют еще квадратурной амплитудной манипуляцией (QASK) 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. 5.3, представлено в векторной форме на рис. 5.6.



а)

s_Q \ s_I	0	1
0	00	01
1	10	11

б)

Рис. 5.6. Сигнальная диаграмма КАМ.

Как видно, пространство комплексной плоскости используется не слишком эффективно - занят только один квадрант. На рис. 5.7 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 6.7. 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.

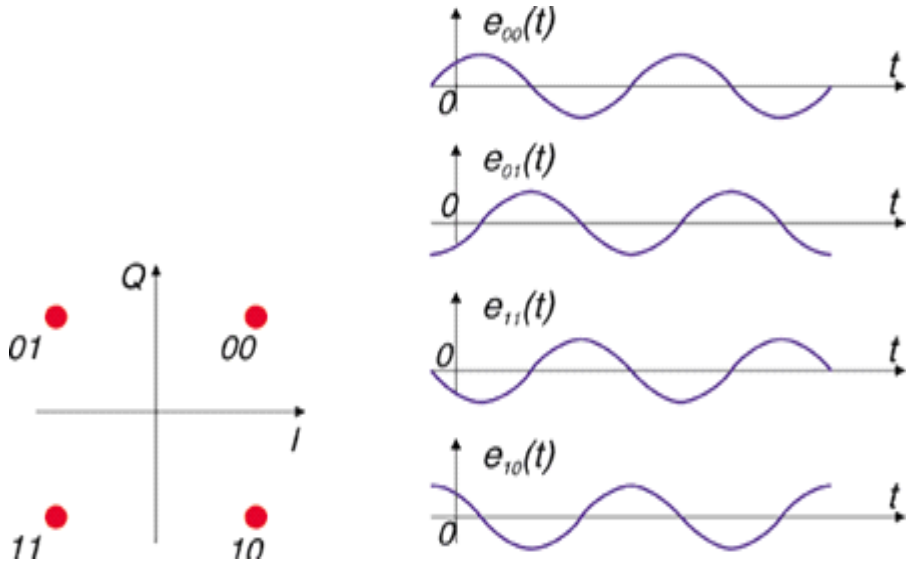


Рис. 5.7. КАМ с более высокой помехоустойчивостью.

Таким образом, скачки мгновенной фазы, сопровождающие переход от данной посылки к последующей, должны быть по возможности минимизированы.

Сигналы КАМ можно выразить так

$$s_m\ t\ = \mathrm{Re}\left[\ A_{mc} + jA_{ms}\ \ g\ t\ \exp\ j\omega_c t\ \right] =$$

$$= A_{mc} g\ t\ \cos\ \omega_c t\ - A_{ms} g\ t\ \sin\ \omega_c t\ , \qquad (5.6)$$

где A_{mc} и A_{ms} - информационные амплитуды сигнала для квадратурных несущих, а $g\ t$ - форма импульса, $m = 1,2,...,M$, $0 \le t \le T$.

Сигнал КАМ можно также выразить и так

$$s_m\ t\ = \mathrm{Re}\left[V_m \exp\ j\theta_m\ \ g\ t\ \exp\ j\omega_c t\ \right] = V_m g\ t\ \cos\ \omega_c t + \theta_m\ , \quad (5.7)$$

где

$$V_m = \sqrt{A_{mc}^2 + A_{ms}^2}\ \text{ и }\ \theta_m = arctg\ \ A_{ms} / A_{mc}\ .$$

Из этой формы представления видно, что сигнал КАМ можно рассматривать как комбинацию амплитудной и фазовой модуляции.

5.4 Амплитудно-фазовая квадратурная манипуляция

Для максимизации минимального расстояния между сигнальными векторами фазовая манипуляция может дополняться 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.

Действительно, мы можем образовать определенную комбинацию M_1-уровневой АМ и M_2-позиционной ФМ, чтобы сконструировать комбинированное АМ-ФМ сигнальное созвездие, содержащее $M = M_1M_2$ точек пространства сигналов. Если $M_1 = 2^n$ и $M_2 = 2^m$, то сигнальное созвездие комбинированной АМ-ФМ сводится к мгновенной передаче $m+n=\log M_1M_2$ двоичных символов, возникающих со скоростью $R/(m+n)$.

Примеры сигнальных пространственных диаграмм для комбинированной АМ-ФМ показаны на рис. 5.7 для M=8 и M=16.

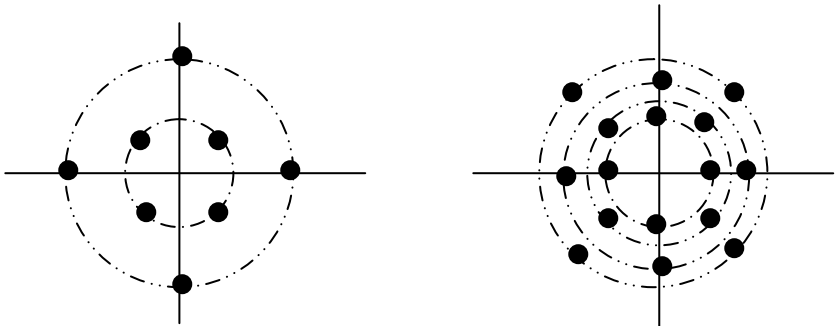


Рис. 5.7. Примеры пространственных диаграмм
для комбинированной АМ-ФМ.

Таким образом, 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.

5.5 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 констелляции (constellation – созвездие) или диаграммой совокупности состояний.

На рис. 5.8 показана структурная схема модулятора и диаграмма состояний (сигнальное созвездие) системы КАМ-16, в которой $x(t)$ и $y(t)$ принимают значения ±1, ±3 (4-х уровневая КАМ).

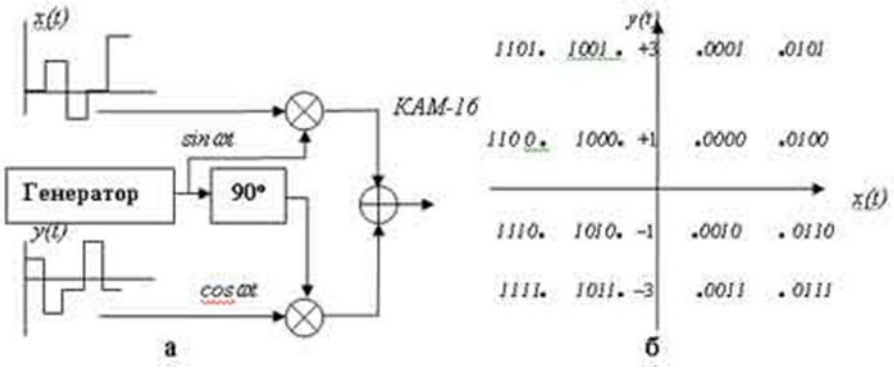


Рис. 5.8. Схема модулятора и сигнальная диаграмма КАМ-17.

Как видно из рис. 6.8, созвездие содержит 16 сигнальных точек, каждая из которых соответствует четырем передаваемым информационным битам.

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

Структурная схема модулятора SPM и диаграммы, поясняющие его работу, приведены на рис. 5.9. В схеме, реализующей данный способ, используются два одинаковых 4-фазных модулятора.

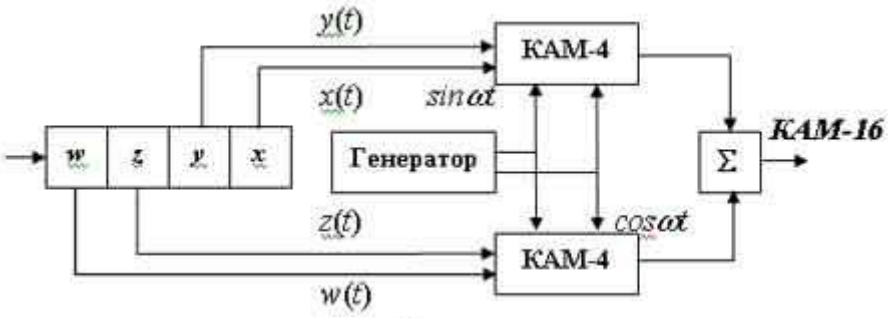


Рис. 5.9. Схема модулятора SPM.

Если задается шесть амплитудных уровней (±1, ±3, ±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. 5.10).

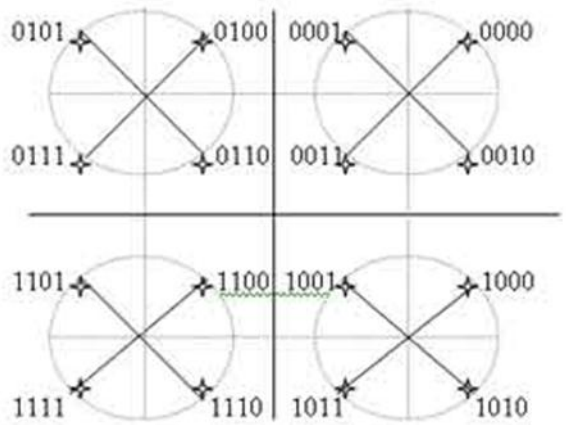


Рис. 5.10. Сигнальная диаграмма модулятора КАМ-17.

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) $d_{кам}$=0,471 и $d_{фм}$=0,390, а при M=32 (L=6) $d_{кам}$=0,283; $d_{фм}$=0,196.

5.6 Минимальная частотная манипуляция - МЧМ

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 спадания спектра мощности сигналов с частотой является разрывность прямоугольной посылки, радикального сужения спектра можно добиться только за счет сглаживания ее формы. Если, в частности, принять в (5.5) за посылку импульс с огибающей в виде положительной полуволны косинуса (см. рис 5.11) с амплитудой $\sqrt{2}A$ КФМ преобразуется в свой подвид, известный как *минимальная частотная манипуляция* – (МЧМ, англ. MSK).

$$S_0(t)=\begin{cases}\sqrt{2}A\cos\left(\dfrac{\pi t}{T}\right), & |t|\le\dfrac{T}{2},\\ 0, & |t|>\dfrac{T}{2},\end{cases} \tag{5.8}$$

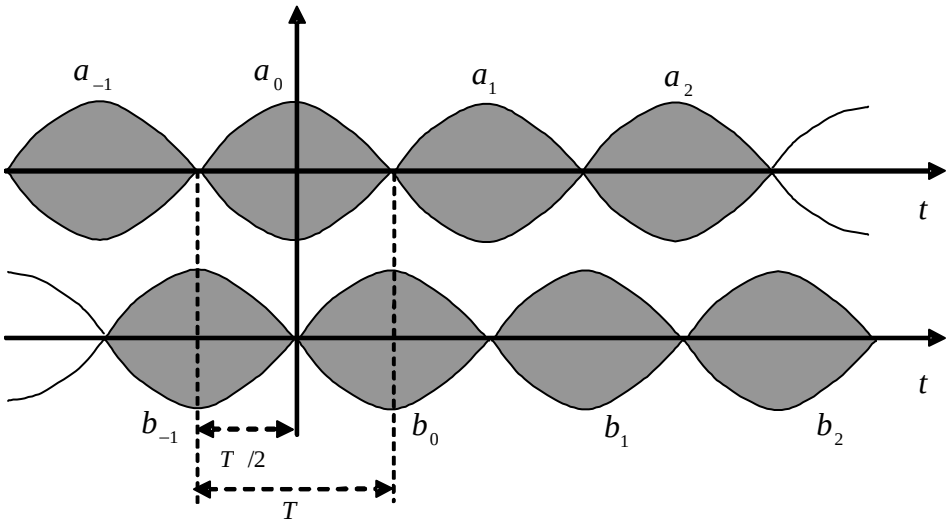


Рис. 5.11. Квадратурные составляющие МЧМ сигнала.

Для того, чтобы убедиться в преимуществах МЧМ и попутно выяснить корни ее наименования, найдем комплексную огибающую $\dot{S}(t)$ сигнала (6.5) с формой посылки, задаваемой (5.8). При этом в силу полной однородности поведения сигнала во времени достаточно рассмотреть лишь два смежных отрезка $[-T/2,0]$ и $[0,T/2]$. Используя тригонометрию, формулу Эйлера и бинарность символов a_i, b_i, в силу которой $a_i/b_i=a_ib_i=\pm1$, получаем

$$\dot{S}(t)=\begin{cases}a_0A\cos\left(\dfrac{\pi t}{\Delta}\right)+jb_{-1}A\cos\left[\dfrac{\pi}{T}\left(t+T-\dfrac{T}{2}\right)\right]=a_0A\exp\left(-j\dfrac{\pi a_0b_{-1}t}{T}\right)\\ a_0A\cos\left(\dfrac{\pi t}{\Delta}\right)+jb_0A\cos\left[\dfrac{\pi}{T}\left(t-\dfrac{T}{2}\right)\right]=a_0A\exp\left(j\dfrac{\pi a_0b_0t}{T}\right)\end{cases} \tag{5.9}$$

Как показывает (6.9), действительная огибающая $S(t)=\left|\dot{S}(t)\right|=A$ 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 (5.9) 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 $\pm1/2T$. Принципиальным является то, что переключение частоты между посылками каждые интервалы

$T/2$ происходит без скачков фазы, т.е. МЧМ является модуляцией с непрерывной фазой.

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 5.12.

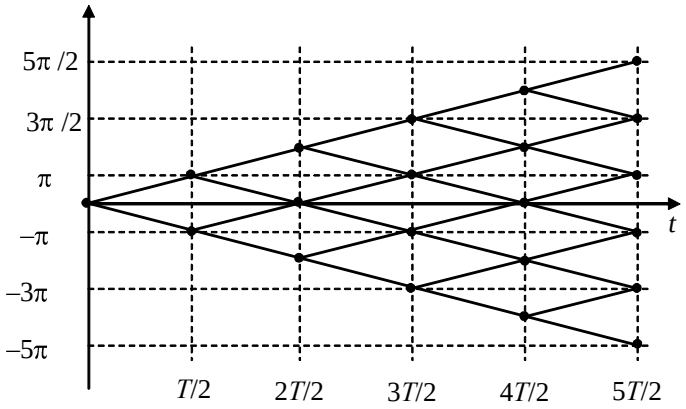


Рис. 5.12. Траектории мгновенной фазы МЧМ сигнала.

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 (6.5) 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.

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. 6.12 могут быть заменены более гладкими. В стандарте GSM применен именно такой вариант модуляции – *гауссовская* МЧМ (GMSK), 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. 5.13.

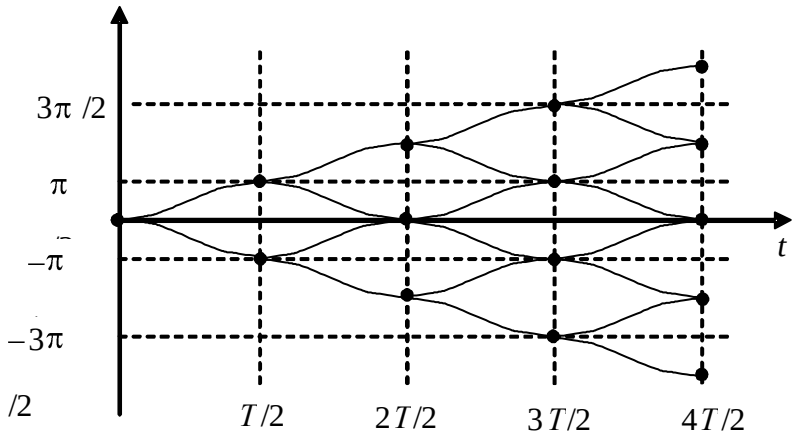


Рис. 5.13. Траектории мгновенной фазы при гауссовской МЧМ

В литературе можно встретить упоминание об энергетическом проигрыше гауссовской МЧМ относительно БФМ в 0,46 дБ.

Приведем итоговую таблицу 5.1, связывающую ширину полосы Δf_{99} модулированного сигнала со скоростью передачи данных R (бит/с) для рассмотренных форматов.

Таблица 5.1.

Вид модуляции	БФМ	КФМ, КФМС, $\pi/4-$КФМ,	МЧМ	Гауссовская МЧМ
$\Delta f_{99} / R$	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.

6 СИСТЕМЫ МОБИЛЬНОЙ СВЯЗИ СТАНДАРТА IS-95

6.1 Эволюция систем сотовой связи

Можно говорить о трех поколениях сотовых систем мобильной связи (ССМС), различия между которыми устанавливаются следующими критериями.

Все ССМС первого поколения являются аналоговыми. В их числе:

- Усовершенствованная мобильная телефонная служба (AMPS). Диапазон рабочих частот: 869…894 МГц для базовых станций (БС) и 824…849 МГц для мобильных станций (МС), ширина полосы канала связи - 30 кГц. Начало коммерческого применения - 1983 г. Широко использовалась в США, Канаде, Центральной и Южной Америке, Австралии.
- Общедоступная система связи (TACS). Частотный диапазон: 935…950 МГц для БС и 890…905 МГц для МС, ширина полосы канала связи - 25 кГц. Начало коммерческого применения - 1985 г. Наибольшее распространение система TACS получила в европейских странах - Англия, Италия, Испания, Австрия и др.
- Скандинавская система мобильной телефонной связи (NMT). Существует в двух основных вариантах NMT 450 и NMT 900, отличаясь только диапазоном используемых частот. В системе NMT 450 - 463…467,5 МГц для БС и 453…457,5 для МС. В системе 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 использовалась 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 символьной информации - 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.

В значительной степени указанных недостатков лишены цифровые ССМС второго поколения, среди которых наибольшее распространение получили следующие:

- D-AMPS (цифровая AMPS) 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, заключается в наличии полностью цифровых каналов управления, начала применяться с 1996 г. Версия IS-136 используется в диапазонах 800 и 1900 МГц.

- GSM (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 удается строить сотовые сети значительно большей емкости. Первоначально данная версия именовалась PCN (Сеть персональной связи), затем 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 (Персональная цифровая сотовая связь). Цифровая ССМС, разработанная в Японии в 1993 г. и первоначально называвшаяся JDC (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.

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 стандарт ССМС на основе метода множественного доступа с кодовым разделением (CDMA), получивший название IS-96. Первые промежуточные спецификации CDMA сотового телефона были опубликованы в 1993-95 года. В настоящее время сети стандарта IS-95 (называемого сейчас как cdmaOne) охватывают обширные территории, обслуживая десятки миллионов потребителей.

Документы по стандарту IS-95 определяют только частотный разнос прямого (869-894 МГц) и обратного (824-849) каналов и не накладывают никаких ограничений на повторное использование частот в соседних сотах или секторах. Стандартная полоса сигнала IS-95 составляет 1.25 МГц, так что в пределах выделенной полосы в 25 МГц оператор обладает значительной свободой в выборе несущих и частотном планировании сети. Все БС, входящие в сеть, жестко синхронизированы с помощью сигналов GPS для работы по единой шкале времени, позволяющей МС легко переключаться с одной БС на другую.

6.2 Характеристика мобильных системы связи стандарта IS-95

Стандарт IS-95 и его эволюционный наследник cdma2000 являются типичными широкополосными системами с прямым расширением спектра, которые явным образом проявляют все преимущества этой технологии. Современные мобильные системы связи обеспечивают передачу сигнала со скоростью 9,6 кбит/с (с кодированием) и 14,4 кбит/с (без кодирования).

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.

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). Системы на базе CDMA 2000 совместимы с сетями на основе IS-95A и IS-95B.

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.

В системе IS-95 реализовано прямое расширение спектра (DSSS) с использованием функций Уолша длины 64 и двух типов псевдослучайных последовательностей (ПСП): короткой и длинной. Линия "вниз" организована на основе синхронного варианта CDMA, использующего для разделения абонентских каналов ансамбль сигнатур в виде функций Уолша. В обратном канале осуществлен асинхронный режим CDMA приписыванием различным абонентам одной соты специфических сдвигов длинной ПСП.

Временные шкалы отдельных БС сети IS-95 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.

Требуемое качество передачи данных в системе достигается с помощью довольно мощного канального кодирования, выполняемого в несколько этапов.

На первом этапе цифровой речевой сигнал структурирован в кадры длительностью в 20 мс и кодируется блоковым циклическим кодом (CRC).

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 мощностью передатчика, имеет вдвое бóльшую избыточность, т.е. скорость 1/3 при той же длине кодового ограничения.

На третьем этапе выполняется перемежение информационного потока в кадре для нейтрализации эффекта пакетирования ошибок.

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.

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.

6.3 Коды расширения спектра

Расширяющие последовательности, используемые в стандарте IS-95, конструировались для обеспечения:

- кодового разделения каналов,
- различения сигналов различных БС, поступающих на вход приемника МС,
- секретности передачи данных.

Синхронное кодовое мультиплексирование физических каналов прямой линий связи, обслуживаемое данной БС, реализовано на основе последовательностей Уолша длины N=64. Ортогональность последовательностей Уолша позволяет разделить 64 физических канала без взаимных помех. Длительность чипа последовательностей Уолша приблизительно равна 0.81 мксек, а скорость изменения чипов - 1.2288 Мчип/сек, что приводит к упоминавшейся ранее полосе 1.25 МГц. Очевидно, что число образованных таким способом прямых физических каналов составляет 64, и в соответствие с CDMA принципом они занимают одну и ту же общую полосу без какого-либо частотного или временного смещения.

Все базовые станции используют одно и то же множество из 64 функций Уолша, а расширение спектра осуществляется т.н. *короткими кодами*, которые позволяют разделить сигналы различных базовых станций друг от друга на входе приемника МС. Используются две бинарные псевдослучайные последовательности короткого кода PN-I и PN-Q, применяемые соответственно

в синфазной и квадратурной ветвях модулятора БС. Первоначально они формируются как *m*-последовательности генераторами на основе линейного регистра сдвига с обратной связью, содержащего15 ячеек памяти, и задаются своими примитивными полиномами.

Полученные последовательности обладают периодом $L = 2^{15} - 1$, но для получения коротких кодов PN-I и PN-Q они расширяются еще на один нулевой символ, добавляемый после 14 последовательных нулей. В итоге получаем последовательности длиной $N = L + 1 = 2^{15} = 32768$ чипов, и при той же чиповой скорости, что и кодов Уолша.

Для распознавания различных базовых станций каждой из них присваиваются специфические копии основных последовательностей коротких кодов, получаемые за счет сдвига во времени. Всего имеется 512 подобных пар копий, каждая последующая сдвинута относительно предыдущей пары на 64 чипа или примерно 52 мксек. При сетевом планировании присвоение пар коротких кодов базовым станциям осуществляется таким образом, чтобы исключить вероятность приема любой МС сигнала случайной БС, имеющего задержку на распространение, как и сигнал полезной БС. Следует также отметить, что относительные временные сдвиги между базовыми станциями, входящими в сеть, устанавливаются одномоментно и остаются постоянными навсегда, поскольку все БС используют GPS приемники для синхронизации работы своих опорных генераторов.

Еще одним расширяющим кодом является *длинный код*, первоначально генерируемый как бинарная *m*-последовательность памяти 42. Полученная *m*-последовательность снова расширяется вставкой одного дополнительного нуля после серии из 41 последовательного нуля, образуя период равный $N = 2^{42}$. При прежней чиповой скорости длинный код обладает периодом в 42 суток.

Различные, специфические для каждого пользователя временные сдвиги длинного кода используются в прямом канале для скремблирования потока данных, а в обратном канале - для организации CDMA разделения сигналов МС и одновременно для защиты передаваемых данных.

6.4 Каналы прямой линии связи

Существуют 4 типа логических, т.е. отражающих содержание передаваемых данных, прямых каналов линии связи:

- пилотный канал,
- канал синхронизации,
- каналы вызова,
- каналы трафика.

6.4.1 Пилотный канал.

Для конкретной БС пилотный канал является единственным, и по нему в

«чистом» виде передается короткий код без модуляции данными. Он может трактоваться как канал, в котором последовательность, состоящая из одних нулей, расширяется с помощью короткого кода. Короткий код, принятый МС, обеспечивает синхронизацию местного генератора с системным (т.е. БС) с точностью, достаточной для формирования приемного когерентного образца, необходимого для снятия расширения и демодуляции принятого сигнала. Поскольку период короткого кода содержит целое число (512) периодов последовательностей Уолша, то последние, сформированные в приемном устройстве, автоматически оказываются согласованными с расширяющими кодами Уолша принятого сигнала после завершения поиска приемником и захвата пилот сигнала схемой слежения за задержкой.

Функция Уолша, являющаяся первой строкой матрицы Адамара размера 64 и состоящая только из единиц, фактически выделяет (канализирует) пилотный канал от остальных.

6.4.2 Канал синхронизации.

Канал синхронизации также является единственным для любой конкретной БС. Наряду с некоторой другой информацией по нему передаются данные, позволяющие приемнику получить специфическую для каждого пользователя маску длинного кода и, следовательно, синхронизовать свою копию длинного кода с используемым БС для шифрования и размещения бит контроля мощности. Исходные данные передаются со скоростью 1.2 кбит/с и структурированы в кадры, длительность которых совпадает с периодом короткого кода (26.67 мсек), причем три последовательных кадра объединяются в суперкадр, содержащий 96 бит. Суперкадры организованы в капсулы сообщения, включающие 30 CRC символов для образования индикатора качества сообщения.

Сверточный код со скоростью 1/2 и длиной кодового ограничения 9 обеспечивает мощную защиту данных от воздействия канальных помех. Скорость кодированного потока составляет 2.4 кбит/с, однако каждый символ представляет собой удвоенный более короткий символ (символьное повторение), так что скорость на входе последующего перемежителя составляет 4.8 кбит/с. Перемежитель используется для декорреляции пакетов ошибок, в результате чего улучшается исправляющая способность сверточного кода в отношении длительных уменьшений интенсивности сигнала из-за замираний или других коррелированных искажений сигнала. Перемежение охватывает один кадр из 128 кодовых символов (26.67 мс) и использует матрицу памяти размерности 16×6. Кодированный поток записывается в нее по колонкам, а затем считывается согласно стандартно определенному образцу.

Последовательность Уолша, являющаяся 33-й строкой матрицы Адамара размерности 64, фактически канализирует данные синхронизации.

6.4.3 Каналы вызова.

В стандарте могут быть до 7 каналов вызова, предназначенных для уведомления абонента о вызове, пришедшем из сети, ответа на запрос пользователя об установлении соединения и передачи другой информации, связанной с предоставлением доступа в сеть. Исходные данные поделены на слоты длительностью 80 мс, состоящие из четырех кадров по 20 мс. Кадры или их части образуют капсулы сообщений, каждая из которых содержит до 1184 бит, включая 30 CRC символов. Возможны две скорости передачи данных вызова: 6.6 или 4.8 кбит/с. При входной скорости в 4.8 кбит/с скорость кодированного потока составляет 6.6 кбит/с, однако каждый символ трактуется как удвоение более короткого (символьное повторение), так что независимо от исходной скорости данных кодированный поток обладает скоростью 16.2 кбит/с.

Перемежитель оперирует с 20 мс кадрами, используя матрицу памяти размерности 24×16, в которую кодированные символы записываются колонками. Порядок считывания из памяти устанавливается стандартом.

Следующей операцией, которой подвергается кодированный и перемеженный поток, является шифрование, реализуемое в форме скремблирования, т.е. его посимвольное суммирование по модулю два с псевдослучайной двоичной последовательностью. Последняя образуется как результат децимирования с индексом децимации 64 специфической для каждого пользователя сдвинутой копии длинного кода, т.е. выборе каждого 64-го символа. Таким образом, скорость 1.2288 Мчип/с длинного кода делится на 64, превращаясь в 1.2288/64=16.2 Кчип/с. Как видно, шифрование является индивидуальным для любого пользователя, предотвращая несанкционированный перехват при мониторинге данных доступа, посылаемых БС.

Функции Уолша с номерами от 1 до 7 (со второй по восьмую строки матрицы Адамара) используются в качестве образующих каналов вызова. Первичный канал вызова, образованный функцией Уолша W_1, всегда доступен, тогда как остальные могут оказаться либо не активизированными, либо используемыми как каналы трафика.

6.4.4 Каналы трафика.

Каналы, рассмотренные до настоящего момента, выполняют служебные функции, необходимые для начала и поддержания передачи основного сообщения пользователю. Каналы трафика ответственны за доставку основной информации: оцифрованной речи, компьютерных или мультимедийных данных и т.п. Для конкретности ограничимся передачей речи. Тогда предварительной операцией является кодирование речи, осуществляемое блоком, называемым *вокодером* (кодером речи).

Оставляя в стороне обсуждение о достаточно сложных принципах работы

этого устройства, отметим только, что в настоящее время в оборудовании стандарта IS-95 используется несколько типов вокодеров, выдающих речевой поток данных с номинальными скоростями 6.6 и 7.3 кбит/с. Номинальная скорость отвечает наивысшей речевой активности (когда речь говорящего непрерывна), тогда как существуют три меньших скорости, отвечающих более низкой активности. Для наибольшей скорости в 6.6 кбит/с они составляют соответственно 4.0 кбит/с, 2.0 кбит/с и 0.8 кбит/с. Вокодер постоянно отслеживает энергию аналоговой речи на интервале 20 мс кадров и сравнивает ее с тремя адаптивными порогами для установления соответствующей скорости оцифрованного выхода. Уменьшение скорости в периоды низкой активности речи сопровождается пропорциональным уменьшением излучаемой мощности сигнала, так что значение энергии, приходящейся на бит, остается постоянной. В свою очередь, более низкий уровень мощности означает уменьшенный уровень помех множественного доступа MAI для других пользователей как внутри так и вне соты и, возможно, большее число абонентов, обслуживаемых данной БС.

Поток оцифрованной речи упаковывается в 20 мс кадры, содержащие, наряду с информационными битами, также CRC символы и 8 хвостовых бит, устанавливающих в нуль сверточный кодер. В результате, множество скоростей, отвечающих номиналу в 6.6 кбит/с, трансформируется на входе канала во множество RC1 «исходных» скоростей 6.6, 4.8, 2.4 и 1.2 кбит/с.

Входные данные на одной из четырех ранее упомянутых скоростей поступают на сверточный кодер скорости $1/2$ и длиной кодового ограничения 6. Не зависимо от входной скорости выходная скорость вследствие символьного повторения всегда равна 16.2 кбит/с, например, поток со скоростью 1.2 кбит/с кодируется в поток с истинной скоростью 2.4 кбит/с, однако в выходном потоке каждый символ интерпретируется как 8 последовательных символов, укороченных в 8 раз. Перемежитель оперирует с блоками в 20 мс (384 кодовых бита), перемешивая символы кодового потока в соответствие с образцом, определяемым спецификацией. Затем сигнал с выхода перемежителя скремблируется аналогично операции в канале вызова для обеспечения секретности передаваемого сообщения.

Ранее было установлено, что мощности сигналов мобильных станций должны эффективным образом контролироваться с целью преодоления проблемы близости-дальности и удержания уровня помех MAI на входе БС ниже критического уровня. Замкнутая петля регулировки мощности является одним из инструментов, решающих упомянутую задачу в стандарте IS-96. Базовая станция постоянно отслеживает интенсивность каждого принятого от МС сигнала и посылает управляющую команду МС либо на увеличение, либо на снижение излучаемой ею мощности. Команда представляет собой просто *биты регулировки мощности*, значение которых нулю или единице диктует соответственно увеличение или уменьшение мощности МС. Для вставки команд в сигнал прямого канала любой 20 мс кадр после перемежителя делится на 16 *групп регулировки мощности* (PCG), каждая из которых занимает интервал в 1.25 мс или $19.2 \cdot 10^3 \times 1.25 \cdot 10^{-3} = 24$ кодовых символа 16.2 кбит/с

потока. В каждой группе единственный бит контроля мощности заменяет два кодовых символа. Очевидно, что приемник МС, зная положение замененных символов (т.е. бит контроля мощности), исключает их из процедуры декодирования, как не имеющих отношения к содержанию сообщения.

Псевдослучайная последовательность на выходе первого дециматора на рис. 6.4 имеет такую же скорость, как и кодовый поток, т.е. 16.2 кбит/с. На интервале длительностью 1.25 мс, располагаются 24 чипа этой последовательности. Последние четыре из них образуют двоичное число, причем 24-й чип отвечает наиболее значимому биту. Данное число, принимающее значение из диапазона от 0 до 15, используется в качестве указателя позиции бита контроля мощности в следующей после текущей группы регулировки мощности. Таким образом, бит контроля мощности может случайно занимать любую позицию из первых 16 в каждой группе PCG. На рис. 6.4 блоки, реализующие позиционирование и вставку бит контроля мощности, обозначены как «Положение бит контроля мощности» и «Мультиплексор».

Каждая МС периодически сообщает БС о достоверности принятых данных, и БС соответствующим образом регулирует уровень мощности сигнала в канале трафика, назначенном данной МС, для поддержания качества приема данных выше заданного порога. Затем взвешенные канальные сигналы складываются в сумматоре и подаются параллельно в синфазную и квадратурную ветви модулятора для умножения с описанными выше бинарными PN-I и PN-Q кодами и формирования в частотной области полосовыми фильтрами. Умножение синфазного и квадратурного сигналов на косинусную и синусную CW компоненты частоты f_0 с их последующим суммированием завершает процедуру формирования сигнала БС и заканчивает процесс модулирования.

Очевидно, что входной модулирующий сигнал в обеих ветвях является одним и тем же. Кроме того, он образован суммой множества бинарных напряжений, т.е. является многоуровневым вещественным сигналом. Временное предположение, что имеется только один физический модулирующий канал, немедленно приводит к ветвям модулятора без суммирования с другими каналами. Можно постулировать, что каждый физический канал обрабатывается именно таким образом, т.е. имеется такое множество пар ветвей модулятора, сколько присутствует каналов, и выходы всех этих параллельных модуляторов когерентно складываются. Поскольку схема на рис. 6.5 линейна, а, значит, действует принцип суперпозиции, то выходной эффект аналогичен тому, который имеет место в гипотетической схеме, т.е. с индивидуальными модулирующими каналами. По этой причине можно говорить, что в прямом канале связи стандарта IS-95 используется прямое расширение спектра, при котором бинарный поток данных (канализированный с помощью функций Уолша) модулирует расширяющий код с QPSK. Поскольку скорость кодового потока на входе модулятора составляет 16.2 кбит/с, то каждый символ обладает длительностью, равной 64 чипам короткого кода. Следовательно, коэффициент расширения прямого канала равен 64.

Стоит отметить, что длинный код не играет роли в прямом расширении сигнала прямого канала, участвуя только в шифровании данных и определении местоположения бит контроля мощности. Часто говорят, что расширения прямого канала осуществляется как кодами Уолша, так и короткими PN-кодами. Более подходящим выглядит классификация функций Уолша как каналообразующих, а PN-кодов - как расширяющих.

Обработка сигнала в приемнике МС основывается на классических процедурах. При успешном поиске пилотного сигнала схема слежения за задержкой входит в захват и непрерывно отслеживает короткий код БС, находящейся в соединении. Локальная копия короткого кода используется для снятия расширения принятого сигнала. На выходе суженного пилотного канала наблюдается «чистая» CW несущая, перенесенная на промежуточную частоту. Схема ФАПЧ согласовывает местный генератор на когерентную работу с принятым CW сигналом, вырабатывая тем самым опорное колебание для когерентной демодуляции данных.

После демодуляции и деперемежения данные, передаваемые по каналам синхронизации, вызова и трафика, с помощью корреляторов, использующих в качестве опор последовательности Уолша, разделяются, декодируются с помощью алгоритма Витерби и используются в соответствии со своим назначением.

Любой приемник МС содержит четыре и более параллельных каналов, способных найти и отслеживать пилотный сигнал. Одной из целей подобного построения является организация RAKE приемника, материализующего преимущества многолучевого разнесения широкополосных сигналов. На практике, по меньшей мере, три подобных канала используются для реализации RAKE алгоритма. Другой процедурой, требующей независимых пилотных каналов в приемнике МС, является процедура передачи обслуживания. Резервный коррелятор (или несколько) осуществляет непрерывное сканирование во временной области, стараясь обнаружить сигналы других базовых станций, характеризующихся большей интенсивностью сигнала и более пригодных для контакта. В последнем случае сеть может подать команду МС на переключение БС, которое легко выполняется, поскольку приемник уже отслеживает ее сигнал (мягкая эстафетная передача).

6.5 Обратный каналы линии связи

В соответствие с логическим содержанием данных, передаваемых по обратному каналу, любая МС работает по одному из двух возможных типов каналов:

- канал связи обратного трафика,
- канал связи обратного доступа.

6.6.1 Канал связи обратного трафика.

Поток передаваемых МС данных (оцифрованная речь от вокодера и т.п.) с вставленными CRC символами делится на 20 мс кадры. Номинальная скорость поступления данных на входе кодера составляет 6.6 кбит/с, однако для пониженной речевой активности используются также три более низких скорости (4.8, 2.4 и 1.2 кбит/с).

Отдельный 20 мс кадр кодового потока (576 бит) делится на 16 групп регулировки мощности по 36 бит (1.25 мс) каждая. Перемежитель, оперирующий в рамках кадра, использует 32×18 матрицу, в которую кодовый поток записывается по колонкам. Считывание осуществляется построчно, Такой порядок является наиболее подходящим для процедуры последующего снижения средней излучаемой мощности пропорциональной уменьшению скорости, которая осуществлена в обратном канале несколько отлично по сравнению с реализацией в прямом канале.

Кодовый поток от перемежителя подается на 64-ичный ортогональный модулятор, в котором каждый его 6-символьный блок, рассматриваемый как 6-ти разрядной двоичное число, определяет один из 64 ортогональных сигналов (номер функции Уолша). Данная операция дает дополнительный выигрыш от кодирования до трех раз к тому, который обеспечивается сверточным кодированием. Поскольку каждые 6 входных двоичных символов теперь заменяются 64 двоичными символами, то поток чипов с выхода ортогонального модулятора становится в $64/6 = 32/3$ раз быстрее (304.2 кбит/с). Отметим, что функции Уолша в обратном канале не имеют отношения к каналообразованию, а только осуществляют широкополосное ортогональное кодирование передаваемых данных.

Следующим шагом по формированию канала трафика является расширение спектра, т.е. суммирование по модулю два потока бинарных символов после ортогонального модулятора со сдвинутой последовательностью длинного кода. Поскольку длинный код представляет собой поток чипов со скоростью 1.2288 Мчип/с, то 4 чипа длинного кода приходятся на один символ последовательности Уолша, или 256 чипов кода на один сигнал Уолша на выходе ортогонального модулятора.

Приемник БС после снятия расширения использует 64-канальный банк корреляторов, каждый из которых настроен на одну из функций Уолша, и принимает решение в пользу такого опорного сигнала Уолша, при которой отклик на выходе коррелятора максимален. В этом случае интервал обработки охватывает полную длительность функции Уолша, т.е. 256 чипов длинного кода, и, значит, коэффициент расширения в обратном канале оказывается равным 256. Длинный код является единственным, жестко определяемым стандартом, а CDMA разделение различных пользователей обеспечивается специфическими для пользователей масками (т.е. временными сдвигами кода). Следовательно, в данном построении мы встречаемся со стратегией

асинхронного варианта CDMA.

Величина сдвига определяется текущим идентификационным номером МС, назначаемым сетью. Параллельно с CDMA канализацией специфическая маска пользователя обеспечивает шифрование (скремблирование) потока после модулятора. Благодаря огромной длине длинного кода задача синхронизации локального генератора длинного кода, необходимая для перехвата сигнала и снятия расширения (т.е. дескремблирования и дешифрации) данных, не представляется простой для несанкционированного пользователя, который не обладает знанием пользовательской маски.

6.6.2 Канал связи обратного доступа.

МС использует канал доступа при ответе на уведомление о входящем вызове в состоянии покоя и либо при необходимости регистрации в сети, либо инициализации вызова. Процедуры деления на кадры, сверточного кодирования, ортогональной модуляции, перемежения и расширения длинным кодом в канале доступа в основном аналогичны процедурам в обратном канале трафика. Очевидно, что по данному каналу не передаются речевые данные, так что не производится регулировка скорости/мощности в соответствие с речевой активностью. Одна из специфичных черт канала доступа связана с инициализацией доступа МС. Не обладая точными сведениями об условиях распространения в обратном канале, МС начинает соединение посылкой пробных сигналов низкой интенсивности, постепенно увеличивая уровень сигнала с каждой последующей попыткой до получения подтверждения со стороны БС об установлении соединения. Пробные сигналы посылаются в пакетном режиме со случайными интервалами для уменьшения вероятности совпадения запросов от нескольких пользователей, поскольку существует вероятность, что в режиме доступа различным МС отвечает одно и тоже значение маски длинного кода.

Поскольку МС никогда не использует каналы трафика и доступа одновременно, то отсутствует необходимость суммирования канальных сигналов на входе модулятора, как это имело место в прямом канале. Выход канала трафика или доступа непосредственно параллельно подается в синфазную и квадратурную ветви, отличающиеся от изображенных на рис. 6.5 в следующих деталях. Во-первых, теперь отсутствует необходимость в идентификации БС сдвинутыми короткими кодами PN-1 и PN-2 (каждая мобильная станция обладает своим уникальным идентификатором – маской длинного кода – во всей зоне покрытия), и даже не пригодны для одновременного приема сигнала одной и той же МС несколькими базовыми во время мягкой эстафетной передачи. Таким образом, все МС используют нулевой сдвиг короткого кода. Во-вторых, имеется элемент задержки на половину длительности чипа, вставленный в квадратурную ветвь после умножителя на PN-Q. Это сделано с целью превращения QPSK в ее версию, называемую QPSK со сдвигом (O-QPSK). Последнее снова считается

предпочтительным для продления работоспособности батареи. Обычный QPSK модулятор, изображенный на рис.6.4, может трактоваться как два BPSK модулятора, независимо работающие с косинусной и синусной составляющими CW несущими. В случае, когда бинарные символы в обеих ветвях одновременно изменяются на противоположные, фаза QPSK сигнала изменяется на 180°, и усилитель мощности передатчика должен обладать линейным динамическим диапазоном, равным примерно удвоенной амплитуде сигнала. В случае O-QPSK при изменении одного из двух модулирующих бинарных потоков другой остается неизменным благодаря сдвигу границ символов на половину длительности чипа. Следовательно, максимальный скачок фазы результирующего сигнала составит только 90°, смягчая тем самым требования к линейности динамического диапазона усилителя и способствуя большей работоспособности батареи.

6.7 Развитие эфирного интерфейса от cdmaOne к cdma2000

Для проектов cdma2000 характерно то, что они рассматривают стандарт IS-95 в качестве отправной точки и принимают в расчет критерий обратной совместимости с ним.

Одним из основных стимулов продвижения стандартов cdmaOne было стремление к расширению функциональности мобильного телефона от простой трубки к терминалу, способному к высокоскоростному обмену данными с сетью, приему и отправке мультимедийной информации, доступу к Интернету и пр. Все эти новации требуют значительно более высокой скорости передачи по сравнению с системами 2-го поколения и основной скоростью, ассоциируемой с философией 3G является 2 Мбит/с. Подобное критическое увеличение скорости без компромисса с качеством услуг и числом обслуживаемых пользователей требует более широкой полосы системы.

При трех несущих (в будущем это число может возрасти) прямой канал cdma2000 просто повторяет три раза спектр IS-95, занимая общую полосу порядка 3.75 МГц и используя в каждой из трех 1.25 МГц подполос методы расширения спектра и модуляции в чем-то отличные от применяемых в IS-96. В отличие от IS-95, где данные с BPSK манипулируют расширяющую последовательность с QPSK, в cdma2000 и данные, и расширяющий код манипулированы с помощью QPSK. Последнее позволяет при той же скорости данных удвоить длительность символов кодового потока и, следовательно, увеличить в два раза (до 128) коэффициент расширения на одну подполосу при той же чиповой скорости. Очевидно, что больший коэффициент расширения означает удвоение числа прямых физических каналов и, следовательно, потенциально большее число обслуживаемых пользователей. Но, что является даже более важным, число параллельных каналов, предоставляемых сетью некоторому пользователю, может способствовать передаче данных в число каналов раз с большей полной скоростью. Данный *многокодовый* метод, наследованный от спецификации IS-95B, является одним из основных ресурсов приближения скоростей к значениям, утверждаемым 3G концепцией.

Другим достоинством прямого канала cdma2000 является использование разнесения на передаче. Поскольку сигналы от различных передающих антенн распространяются по различным путям, то организовано несколько дополнительных пилотных каналов помимо одного, существующего в стандарте IS-95, передающих специфические для антенн пилотные сигналы от двух антенн БС, участвующих в разнесении. Это предоставляет возможность МС разделить пилотные сигналы различных антенн и строго синхронизовать опоры для сигналов разнесения, чтобы демодулировать, декодировать и соответствующим образом скомбинировать ветви разнесения.

Более сложная организация стандарта cdma2000 требует более строгой систематизации прямых логических каналов. Вместо четырех градаций в IS-95 иерархия логических каналов в cdma2000 насчитывает 10 только для первого уровня, включая общий канал управления, общий канал регулировки мощности, канал управления сообщениями, каналы быстрого вызова и др.

Что касается каналов трафика, то они помимо прочих включают основные (первоначально используемые) и дополнительные каналы, которые привлекаются к многокодовой передаче в случае, когда высокая скорость передачи не достигается с использованием только основных каналов. Дополнительные каналы могут использовать либо сверточное, либо турбо-кодирование.

Обратный канал cdma2000 реализует обычный DS-CDMA с полосой в три раза более широкой (3.75 МГц). Этому значению отвечает в три раза более высокая чиповая скорость (3.6864 Мчип/с) длинного расширяющего кода. При столь большой скорости прямого расширения ряд решений стандарта IS-95 было пересмотрено и заменено более уместными. Прежде всего, используется сверточное кодирование со скоростью $1/4$ вместо $1/3$, означающее, что исходная скорость данных, равная например, 6.6 кбит/с становится 36.4 кбит/с.

Использование модуляции данных с помощью QPSK и расширения такого кодового потока длинной последовательностью приводит к коэффициенту расширения, равного $2\cdot(3.6864/38.4)\cdot10^3=192$, что незначительно меньше 256 в IS-96. В то же время уменьшение скорости сверточного кода с длиной кодового ограничения 9 до величины $1/4$ приводит к некоторому дополнительному выигрышу от кодирования. На основании этого факта проектировщики решили отказаться от ортогональной 64-ичной модуляции и сделать структуру обратного канала трафика более близкой к структуре аналогичного назначения в прямом канале.

Другим отличием служит широкий перечень логических каналов, включая обратный пилотный канал, который становится теперь обязательным, поскольку в отсутствии ортогональной модуляции БС необходимо иметь локальную когерентную опору для демодуляции QPSK данных МС.

Имеются и другие новые логические каналы, часть из которых работают одновременно. Будучи образованными с помощью сигналов Уолша, упомянутые каналы затем линейно суммируются, означая, что усилитель мощности передатчика должен быть линейным. Следовательно, используется

обычный QPSK модулятор данными, при котором кодовый поток, отображенный QPSK, первоначально умножается на расширяющий код с QPSK, а затем вещественная и мнимая части произведения модулируют косинусную и синусную компоненты CW несущей.

Отметим, что аналогично прямому каналу обратные дополнительные каналы могут поддерживать обратный основной (трафика) канал (один на МС), обеспечивая многокодовую передачу на высоких скоростях, не достижимых с помощью только одного основного канала. Дополнительные каналы опять могут использовать либо сверточные, либо турбо-коды.

7 СИСТЕМЫ РАДИОСВЯЗИ СТАНДАРТА UMTS.

Система *Всемирной Мобильной Связи* (UMTS) представляет собой широкополосный CDMA стандарт третьего поколения, проект которой был разработан Европейским телекоммуникационным сообществом. В настоящее время UMTS наряду с cdma2000 входят в т.н. IMT-2000 семейство, т.е. перечень стандартов, принятых Международным союзом связи (ITU) в качестве основы для 3G (третьего поколения) систем.

Существуют две версии UMTS:

- с частотным дуплексом (FDD),
- временным дуплексом (TDD).

Непосредственно из названий видно, что эти системы отличаются друг от друга способом разделения прямого («вниз») и обратного («вверх») каналов. В FDD-UMTS прямой и обратный каналы занимают неперекрывающиеся частотные полосы, тогда как в TDD-UMTS они используют различные временные слоты. Для краткости ограничимся обсуждением только FDD-UMTS, учитывая, что многие принятые решения являются общими для обеих систем.

В Европе для FDD-UMTS Международными регламентирующими документами выделены частотные полосы 1920-1980 МГц (канал «вверх») и 2010-2170 МГц (канал «вниз») с ограничением в 5 МГц на ширину полосы канала. Прямое расширение спектра (DS spreading) является основной технологией, обеспечивающей разделение физических каналов и, в частности, множественный доступ (DS-CDMA). Стандарт определяет единую и постоянную чиповую скорость в 3.84 Мчип/с в полном соответствие с ограничением на полосу канала. UMTS, как и cdma2000, является системой, в которой скорость передачи может варьироваться в очень широких пределах. Следствием этого, наряду с неизменностью чиповой скорости, является изменение коэффициента расширения при изменении скорости передачи.

В дальнейшем снова сконцентрируем наше внимание на физическом уровне системы, т.е. решениях, касающихся расширения, канализации и модуляции.

Подчеркнем, что БС системы UMTS не пользуются поддержкой GPS и функционируют с собственными автономными не синхронизированными опорными генераторами. Выгодная с точки зрения экономии оборудования, подобная архитектура, однако, превращает поиск и эстафетную передачу в более сложные процедуры и определяет многие отличия в физических уровнях UMTS и cdma2000.

Данный материал есть не более чем очень краткий обзор, посвященный стандарту UMTS.

7.1 Типы каналов стандарта UMTS

В соответствие с терминологией спецификаций UMTS различают:

- логические,
- транспортные,
- физические каналы.

На уровнях, выше физического, данные распределяются между логическими каналами на основе информационного содержания, но перед переходом на физический уровень они реструктурируются в транспортные каналы.

Критерием различения транспортных каналов служит способ и формат представления данных.

Физические каналы есть просто сигналы, переносящие информацию и различаются своими специфическими кодами.

Физический уровень состоит из двух подуровней. Передаваемые данные доставляются на первый физический подуровень с верхних уровней, упакованными в транспортные каналы в соответствии с информационным содержанием. Первый подуровень включает среди других присоединение CRC символов для защиты блоков данных, канальное кодирование и перемежение. Канальное кодирование выполняется либо как сверточное (длина кодового ограничения 9, скорость $1/3$ или $1/2$) или как турбо-кодирование (образованное кодером памяти 3 и скорости $1/3$). Для очень высокой скорости передачи данных возможна также не кодированная передача. Второй подуровень, т.е. линия радиосвязи, охватывает отображение транспортных каналов на физические (т.е. сигналы), передачу сигналов через среду распространения.

Сигналы, принятые МС (канал «вниз») или БС (канал «вверх»), подвергаются затем всем необходимым взаимным операциям двух подуровней (демодуляции, деперемежению, декодированию и т.п.), обратному отображению, если это необходимо, и передаче данных на более высокий уровень.

Другая классификация, применимая как к транспортным, так и физическим каналам, различает *общие* и *выделенные* каналы. Первые из них содержат данные, относящиеся ко всей системы и используемые совместно всеми пользователями, тогда как вторые служат только для установления соединения между БС и конкретной МС.

Временная структура всех каналов строго определена стандартом. Все они состоят из кадров длительностью 10 мс (38 400 чипов), а каждый кадр, в свою очередь, поделен на 15 слотов длительностью 665.66 мкс (2560 чипов).

7.2 Выделенные физические каналы «вверх».

Выделенный физический канал является каналом, который сеть предоставляет МС для эксклюзивного использования, т.е. два мобильных абонента в пределах одной и той же соты никогда не используют одинаковый выделенный канал.

Существуют два типа выделенных каналов «вверх»:

- выделенный физический канал данных (DPDCH),

- выделенный физический канал управления (DPCCH).

Канал DPCCH используется для передачи БС служебной информации:

- пилотный сигнал, необходимый для оценки параметров распространения
- синхронизации когерентной опоры БС,
- сообщение о формате данных в DPDCH,
- информацию обратной связи, используемую при эстафетной передаче, и команды регулировки мощностью.

На одну МС всегда приходится только один DPCCH точно также, как передаваемые пользовательские данные попадают на физический уровень, упакованные в единственный транспортный канал, приходящийся на МС. Однако этот единственный транспортный канал может быть отражен на несколько (вплоть до 6) DPDCH каналов, если необходимая скорость превосходит максимальную скорость, обеспечиваемую единственным DPDCH каналом.

Канал DPCCH передается через квадратурную ветвь QPSK модулятора, тогда как первый DPDCH канал поступает в синфазную ветвь. Остальные DPDCH каналы при их использовании по возможности равномерно распределяются между ветвями. Таким образом, несмотря на применение QPSK модулятора в канале «вверх» используется BPSK модуляция данных. Скорость передачи по DPCCH всегда постоянна и равна 15 кбит/с или 10 бит/слот. Тогда каждый бит занимает интервал, равный 2560/10=256 бит, и коэффициент расширения в DPCCH всегда равен 255. В отличие от этого коэффициент расширения в DPDCH канале варьируется в зависимости от требуемой скорости данных и может изменяться от 256 (минимальная скорость) до 4 (максимальная скорость). Таким образом минимальная скорость передачи данных равна 15 кбит/с, тогда как максимальная скорость на один DPDCH канал в 64 раза выше, т.е. 960 кбит/с. Использование 6 параллельных DPDCH каналов в многокодовом режиме передачи, описанном в предыдущем разделе, в принципе позволяет достичь максимальной скорости, равной 5760 кбит/с. Очевидно, что т.н. гросс-скорость (т.е. скорость кодового потока после канального кодера) является желаемой, а исходная скорость потока данных пропорционально меньше скорости канального кода.

Каждый физический канал «вверх» обладает собственным каналообразующим кодом, умноженным на кодированные данные. После умножения на канализирующий код все физические каналы взвешиваются своими коэффициентами усиления, реализуя регулировку мощности обратного канала.

7.3 Общие физические каналы «вверх».

Общие физические каналы представляют собой ресурс, которые находится в общем распоряжении всех мобильных абонентов. Существуют два типа

общих каналов «вверх»:

- канал случайного доступа (RACH),
- общий канал пакетирования (CPCH).

Физический RACH (PRACH) канал содержит преамбулу из 4096 чипов, которая представляет собой 16-чиповый идентификатор МС, повторенный 256 раз. МС может начать передачу по PRACH в начале любого из 15 слотов доступа, занимающих совместно два кадра, т.е. 5120 чипов или 20 мс. Диаграмма слотов доступа задается в формате передачи БС, использующей специальный транспортный канал «вниз» (BCH). В начальном состоянии у МС нет достоверных сведений об ослаблении сигнала в канале «вверх», и первая передача преамбулы реализуется при низком уровне мощности. В случае отсутствия подтверждения от БС об установлении соединения МС случайным образом выбирает новые слоты доступа и предпринимает новые попытки соединения, каждый раз увеличивая мощность сигнала. После получения от БС подтверждения МС передает собственно сообщение, занимающее один или два кадра (10 или 20 мс). В канале PRACH отсутствует замкнутая петля регулировки мощности, поскольку сеансы соединения по нему чрезвычайно коротки.

Структуры физических каналов CPCH (PCPCH) и PRACH в основном подобны, однако сегмент сообщения по каналу PCPCH может занимать большее число кадров, а сегмент преамбулы сопровождается еще одной частью: преамбулой коллизии обнаружения, помогающей выявить одновременные попытки нескольких МС воспользоваться каналом PCPCH. Вследствие большей длительности пакетов PCPCH представляется оправданным присутствие в канале замкнутой петли регулировки мощности в сравнении с PRACH.

Как и в выделенных, в общих каналах PCPCH и PRACH используется квадратурное мультиплексирование для комбинирования информационных и служебных компонентов передаваемых данных.

Разделение общих и выделенных физических каналов осуществляется с помощью канализирующих кодов линии «вверх», рассматриваемых в следующем пункте.

7.4 Канализирующие коды линии «вверх»

Индивидуальная МС использует несколько типов каналов, которые должны быть разделены в приемнике БС. Поскольку все физические каналы МС привязаны к одной и той же временной шкале, то наилучшим способом организации разделения физических каналов (нескольких DPDCH, DPCCH, PRACH, PCPCH) случит синхронное кодовое разделение.

Отметим, что речь идет только о разделении каналов отдельной МС, тогда как сигналы других мобильных абонентов разделяются с помощью асинхронного варианта CDMA, аналогично реализованному в cdmaOne и cdma2000.

Формат канализирующего кодирования описывается с помощью двоичного графа, определяемого итерационной процедурой. На каждом шаге итерации любое кодовое слово, полученное на предыдущей итерации, порождает два новых слова удвоенной длины путем добавления либо исходного слова, либо его инверсии. Следует отметить, что данный алгоритм отличается от конструирования матрицы Адамара по правилу Сильвестра только переупорядочиванием строк результирующей матрицы, так что полученные кодовые слова есть ничто иное, как функции Уолша. Несмотря на это в спецификациях по UMTS они фигурируют под специальным названием *ортогональных кодов с переменной величиной расширения* (OVSF).

Граф OVSF кодов канала «вверх» построен с использованием k=8 итераций, так что кодовые слова имеют максимальную длину, равную $n = 256$ чипов. Состоящее из одних единиц слово указанной длины определено для выделенного канала управления DPCCH, в результате чего последующее скремблирование обеспечивает DPCCH каналу коэффициент расширения спектра 255.

Выбор каналообразующих кодов аналогичной длины для DPDCH (коэффициент расширения 256) позволяет вести передачу данных со скоростью 15 кбит/с. Если требуется более высокая скорость, то символы данных укорачиваются, т.е. уменьшается коэффициент расширения (чиповая скорость никогда не изменяется, оставаясь равной 3.84 Мбит/с).

Подобный выбор всегда сохраняет ортогональность DPDCH и DPCCH каналов вне зависимости от коэффициента расширения в DPDCH канале. Если единственный DPDCH канал (с коэффициентом расширения 4) не может обеспечить требуемую скорость данных, то привлекается метод многокодовой передачи по нескольким DPDCH каналам, обладающих одним и тем же минимальным коэффициентом расширения 4.

Способ назначения кодовых слов PRACH и PCPCH каналам, устанавливаемый спецификацией, также обеспечивает их ортогональность выделенным каналам во всем диапазоне скоростей передачи данных по DPDCH каналам.

7.5 Скремблирование линии «вверх»

Завершающим этапом формирования сигнала МС является его прямое расширения с помощью специфической для каждого пользователя сигнатурой, реализующей асинхронный вариант CDMA разделения сигналов различных МС в приемнике БС. Согласно терминологии UMTS, данная операция, как и подобная в линии «вниз», называется *скремблированием*, употребление термина которой несколько отлично от используемого в документах по cdmaOne и cdma2000.

Существует два типа скремблирующих кодов линии «вверх», определенных стандартом: длинный или короткий коды, применяемые при скремблировании выделенных каналов.

Последовательности Голда длиной $L = 2^{25} - 1$, усеченные до

длительности одного кадра, т.е. 38 400 чипов, используются для построения длинных скремблирующих кодов. Последовательности Голда затем отображается на {±1} алфавит. «Истинная» усеченная последовательность Голда c_i определяет вещественный компонент QPSK сигнатуры МС. Для получения мнимой части исходная (до усечения) последовательность Голда вначале сдвигается на 777 232 чипа, а затем подвергается усечению и отображению с целью построения бинарной {±1} последовательности c'_i. Далее, инверсией каждого четного элемента последней последовательности заменяется следующий нечетный элемент, а полученный результат посимвольно перемножается с c_i. Последняя операция вдвое уменьшает число элементов в результирующей QPSK последовательности, полярность которых противоположна предшествующему.

Скремблирующая последовательность точно синхронизирована с временной шкалой МС, начинаясь в любом кадре с одного и того же символа (чипа). Короткие скремблирующие последовательности обладают периодом 256 и планируются к использованию в случае, когда приемник БС готов к выполнению многокодового алгоритма. В случае асинхронного CDMA сложность этого типа приемника обычно возрастает с ростом длины расширяющего кода. Правило формирования коротких скремблирующих кодов, устанавливаемое спецификацией, включает операции генерирования четвертичной линейной рекуррентной последовательности длиной 255, ее суммирования по модулю четыре с двумя двоичными рекуррентными последовательностями длин 51 и 85, расширения результирующей последовательности на один элемент до длины 255.

7.6 Отображение транспортных каналов «вниз» на физические каналы

Информация, передаваемая сетью некоторой конкретной МС на транспортном уровне, организована в виде единственного выделенного канала, который затем отображается двумя физическими каналами «вниз», подобными в линии «вверх»: *выделенные каналы данных* (DPDCH) и *управления* (DPCCH).

Перечень общих транспортных каналов линии «вниз» много шире по сравнению с линией «вверх». В частности, он содержит уже упоминавшийся канал оповещения BCH, по которому передаются параметры сети или соты, используемые всеми МС; *прямой канал доступа* (FACH), используемый БС для отсылки командных данных МС, положение которой известно; *канал вызова* (PCH), посредством которого БС передает команды МС с неизвестным местоположением и т.п. На физическом уровне *первичный общий физический канал управления* (P-CCPCH) передает данные BCH канала, канал FACH отображается во *вторичный общий физический канал управления* (S-CCPCH), передача данных PCH частично осуществляется снова S-CCPCH каналом и частично *каналом синхронизации* (SCH). Еще имеется *общий пилотный канал*

(CPICH), транслирующий сигнал немодулированных данных, которые используются приемником для оценки параметров канала. Полная схема соответствия между транспортными и физическими каналами линии «вниз» устанавливается соответствующей спецификацией.

7.7 Формат физических каналов линии «вниз»

Механизм мультиплексирования выделенных каналов данных и управления линии «вниз» DPDCH и DPCCH отличается от используемого в линии «вверх»: каждый слот разбит на части и каждая часть отведена либо данным DPDCH канала, либо DPCCH канала. При подобной организации передатчик БС работает в прерывистом режиме, что обусловлено паузами в потоке данных во время низкой речевой активности. Прерывистость излучения в работе передатчика МС может вызвать помехи, сильно затрагивающие соседние электронные устройства (например, слуховой аппарат). О данном явлении неоднократно сообщалось при использовании телефона стандарта GSM. Однако для БС прерывистость излучения не является столь значимым фактором, поскольку маловероятно, чтобы приборы, о которых говорилось выше, находились достаточно близко от антенны БС.

Другим отличием формата физического канала линии «вниз» служит применение QPSK модуляция данных (в отличие от BPSK в линии «вверх»).

Завершающим шагом является умножение данного QPSK сигнала данных на QPSK расширяющий (скремблирующий) код и QPSK модуляция несущей, аналогично осуществляемой в линии «вверх».

7.8 Канализирующие коды линии «вниз»

Мультиплексирование во времени каналов данных и управления, рассмотренное в предыдущем пункте, означает, что для передачи обоих этих информационных потоков необходим только один *выделенный физический канал* (DPCH). Если единственный данный канал удовлетворяет требованиям скорости передачи данных, то БС формирует его с помощью специфической для данного пользователя кодовой последовательности, которая, конечно, является единственной в соте (или секторе) и не может быть использована для соединения с любой другой МС. Сценарий многокодовой передачи возникает тогда, когда единственный физический канал не способен обеспечить передачу данных с требуемой скоростью. Тогда БС организует несколько параллельных физических каналов для соединения с той же МС. Эти каналы всегда строятся с одинаковым коэффициентом расширения спектра, но, очевидно, с различными канализирующими кодами, которые не могут использоваться БС при соединении с другими мобильными. Поскольку отсутствует необходимость в повторении командных сообщений по всем многокодовым каналам, то информация управления пересылается МС только по одному из них.

Некоторые канализирующие последовательности не доступны для DPCH канала, поскольку предназначены для общих каналов, как CPICH.

При работе многих пользователей с разными скоростями их канализирующие коды должны сохранять ортогональность, несмотря на различие в величине коэффициента расширения спектра или, что то же самое, в длине кода.

` Действительно, разным МС следует назначать различные подмножества последовательностей Уолша, не содержащие потомков последовательностей, обслуживающих в текущее время другие МС. Данная проблема динамического распределения ресурса решается на более высоком уровне протоколом сети.

7.8. Скремблирующие коды линии «вниз»

Как указывалось ранее, скремблирующие коды линии «вниз» обеспечивают разделение сигналов различных БС. Каждая скремблирующая последовательность базируется на последовательности Голда длины $L = 2^{18} - 1 = 262\,143$. Хотя имеется $2^{18} + 1 = 262\,145$ последовательностей Голда указанной длины, спецификация ограничивает число используемых величиной $2^{13} = 8\,192$. Затем от каждой разрешенной последовательности отрезаются два сегмента длиной 38 400: исходный и сдвинутый на $2^{17} = 131\,072$ чипа, отображаемые далее в $\{\pm 1\}$ алфавит по обычному правилу. Результирующие бинарные последовательности представляют собой вещественную и мнимую части QPSK скремблирующего кода линии «вниз».

Спецификацией устанавливается строгая иерархия скремблирующих последовательностей. Множество всех скремблирующих кодов делится на 512 подмножеств, каждое из которых содержит один первичный и 15 вторичных кодов. В свою очередь, все 512 первичных кодов поделены на 64 группы по 8 последовательностей в каждой. Конкретной БС назначается только один уникальный первичный код. Для некоторых физических каналов разрешено использовать только первичный код, тогда как для других можно применять как первичный, так и вторичный коды.

7.10 Канал синхронизации

Канал синхронизации (SCH) играет исключительно важную роль в структуре сети, обеспечивая начальный поиск соты и согласование временной шкалы МС с границами кадров и слотов принятого сигнала БС. Ни канализирующие, ни скремблирующие коды не принимают участия в формировании канала синхронизации, поскольку его сигнал должен быть обнаружен и обработан, прежде чем МС будет осведомлена о скремблирующем коде контактирующей с ней БС. Подчеркнем еще раз, что временные шкалы базовых станций сети UMTS не синхронизированы между собой и что при переключении МС на другую базовую (например, в ходе эстафетной передачи) некоторое время должен осуществляться начальный поиск соты. Последнее служит обоснованием одного из ключевых положений проектирования

архитектуры SCH канала, заключающегося в проведении двухэтапной процедуры поиска, которая потенциально способна обеспечить экономию времени поиска по сравнению с обычным последовательным поиском. Действительно, SCH канал образован парой из первичного и вторичного каналов синхронизации, сигналы которых - первичный и вторичный коды синхронизации - используются на первом и втором этапах поиска соответственно. Оба сигнала занимают начальные 256 чипов в каждом слоте. Последовательность первичного кода синхронизации идентична не только во всех слотах, но и для всех базовых станций сети. Указанный факт делает невозможным установление синхронизации с первичным сигналом заранее определенной БС; определение БС, с которой установлен контакт МС после завершения первого этапа синхронизации, становится ясным только по завершению второго этапа. После нахождения первичного сигнала синхронизации МС знает границы слотов, но не кадров. На втором этапе МС должна снять эту неопределенность, проверяя все 15 (по числу слотов в кадре) возможных значений рассогласования между шкалой МС и принятым сигналом БС для всех возможных версий вторичного кода синхронизации. Для способствования решению этой задачи вторичный код синхронизации имеет период, равный одному кадру (15 слотам или 38 400 чипам), и кодовую структуру, специфичную для данной БС. Кроме того, вторичный код синхронизации жестко связан с группой скремблирующих кодов (одной из 64), назначенных базовой станции. Таким образом, после завершения второго этапа поиска МС опознает группу скремблирующих кодов захваченного сигнала БС и, тестируя 8 гипотез относительно возможной последовательности первичного скремблирующего кода, окончательно определяет, какая из них используется контактируемой БС.

Первичный код синхронизации.

Первичный код синхронизации (PSC) определен спецификацией как 16-элементная последовательность

$$\mathbf{a} = (1,1,1,1,1,1,-1,-1,1,-1,1,-1,1,-1,-1,1),$$

повторяющаяся 16 раз с прямой или инвертированной полярностью, образуя 256-элементную бинарную последовательность

$$\mathbf{c}_{ps} = (\mathbf{a},\mathbf{a},\mathbf{a},-\mathbf{a},-\mathbf{a},\mathbf{a},-\mathbf{a},-\mathbf{a},\mathbf{a},\mathbf{a},\mathbf{a},-\mathbf{a},\mathbf{a},-\mathbf{a},\mathbf{a},\mathbf{a}).$$

Каждая БС непосредственно передает одну и ту же последовательность $\mathbf{c}_{ps}$, реализуя PSC и предоставляя МС возможность осуществить первый этап поиска соты. Поскольку сигнал PSC предназначен для измерения времени в обстановке многолучевости, то он должен обладать хорошими автокорреляционными свойствами (см. параграф 5.1). Несмотря на периодичность PSC большая часть его периода свободна, так что уровень боковых лепестков апериодической АКФ является адекватным критерием качества PSC. Как указывалось в примере 5.5.2 (см. также рис. 5.16), апериодическая АКФ PSC системы UMTS достаточно далека от идеала, что,

возможно, является неизбежной платой за простоту аппаратной реализации, к которой стремились разработчики.

Вторичный код синхронизации.

Вторичные коды синхронизации (SSC) строятся на основе 16-элементной последовательности **b**, повторяющей последовательность **a** в первых 8-ми элементах, и с инверсией **a** (т.е. –**a**) - в остальных:

$$\mathbf{b} = (1,1,1,1,1,1,-1,-1,-1,1,-1,1,-1,1,1,-1).$$

Данная последовательность повторяется 16 раз с изменением полярности или без него, образуя 256-элементную последовательность вида

$$\mathbf{z} = (\mathbf{b},\mathbf{b},\mathbf{b},-\mathbf{b},\mathbf{b},\mathbf{b},-\mathbf{b},-\mathbf{b},\mathbf{b},-\mathbf{b},\mathbf{b},-\mathbf{b},-\mathbf{b},-\mathbf{b},-\mathbf{b},-\mathbf{b}).$$

Затем последовательность **z** посимвольно перемножается с каждой 16-й строкой построенной по алгоритму Сильвестра матрицы Адамара (см. пункт 2.4.3) размерности 256 для образования 16 ортогональных символов 256-ричного алфавита, которые используются в дальнейшем для построения кодовых слов SSC. Алгоритм кодирования SSC должен обеспечивать низкий уровень всех неправильных корреляций, т.е. между всеми 15 циклическими сдвигами любого одного слова, а также между любыми временными сдвигами различных слов. В спецификации UMTS все 64 кодовые слова SSC представлены в виде таблицы, однако простая проверка устанавливает, что эти слова взяты из (15,3) 256-ричного кода Рида-Соломона. Минимальное хэммингово расстояние любого кода Рида-Соломона определено как на единицу большее числа проверочных символов [31,33], т.е. в нашем случае равно 7. Оно гарантирует, что возможно не более двух совпадений 256-ричных символов между любым SSC словом и любой из 14-ти циклически сдвинутых копий, разнесенных друг относительно друга на целое число слотов. Другими словами, нормированная периодическая АКФ любого SSC кодового слова при указанных сдвигах имеет уровень боковых лепестков не выше $2/15$. Аналогичное утверждение справедливо для любых двух кодовых слов, назначенных различным БС, при взаимных сдвигах, равных целому числу слотов. В то же время, поскольку БС сети не привязаны к общей шкале, передаваемые ими коды синхронизации скользят друг относительно друга и, значит, низкий уровень корреляции должен сохраняться при произвольном временном сдвиге, а не только равном целому числу слотов. Оптимальность расстояния кода Рида-Соломона не может обеспечить низкий уровень корреляции при произвольных сдвигах SSC последовательностей, и данная проблема заслуживает дальнейших исследований.

В качестве общего резюме снова подчеркнем, что данный краткий обзор 3G систем и стандартов касался только решений практической реализации широкополосности и идей CDMA.

8 ЧЕТВЕРТОЕ ПОКОЛЕНИЯ МОБИЛЬНОЙ СВЯЗИ

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 службами 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.

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.

8.1 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.

8.2 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).

8.3 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].

8.4 Характеристика 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 4.5. 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», 2006. – 1104 с.
2. Ипатов В. Широкополосные системы и кодовое разделение сигналов. Принципы и приложения. Москва: Техносфера, 2006. – 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, 1986. – 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, 2006. – 472 с.: ил.
6. Основы теории кодирования и шифрования в современных РТС: Конспект лекций / Самойлов И.М. – Санкт-Петербург, СПбГЭТУ «ЛЭТИ», 2006. – 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 1986. С. 41 – 56.
9. Стиффлер Дж. Дж. Теория синхронной связи: Пер. с англ. /Под ред. Э.М. Габидуллина.-М.: Связь. 1976. – 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.
13. Варакин Л.Е. Системы связи с широкополосными сигналами. – М.: Радио и связь. 1986. – 384 с.