МИНИСТЕРСТВО ЦИФРОВОГО РАЗВИТИЯ, СВЯЗИ И МАССОВЫХ КОММУНИКАЦИЙ РОССИЙСКОЙ ФЕДЕРАЦИИ
Северо-Кавказский филиал
ордена Трудового Красного Знамени федерального государственного бюджетного образовательного учреждения высшего образования
«Московский технический университет связи и информатики»

Кафедра «Инфокоммуникационные технологии и системы связи»

Методические указания по выполнению курсовой работы по дисциплине: «Общая теория связи» по теме: «Разработка кодека и модема»

Часть 1(кодек)

(направление подготовки 11.03.02)

Ростов-на-Дону
2022

Методические указания по выполнению курсовой работы
по дисциплине: «Общая теория связи» по теме: «Разработка кодека и модема»
Часть 1(кодек)

Составители: Заведующий кафедрой ИТСС к.т.н. доцент Юхнов В.И.
Доцент кафедры ОНП к.ф-м.н. доцент Бородин А.В.

Рецензент: доцент кафедры ИТСС доцент Ершов В.В.

Методическое пособие обсуждено и одобрено на заседании
кафедры ИТСС.
Протокол от «19» декабря 2022 г. № 5.

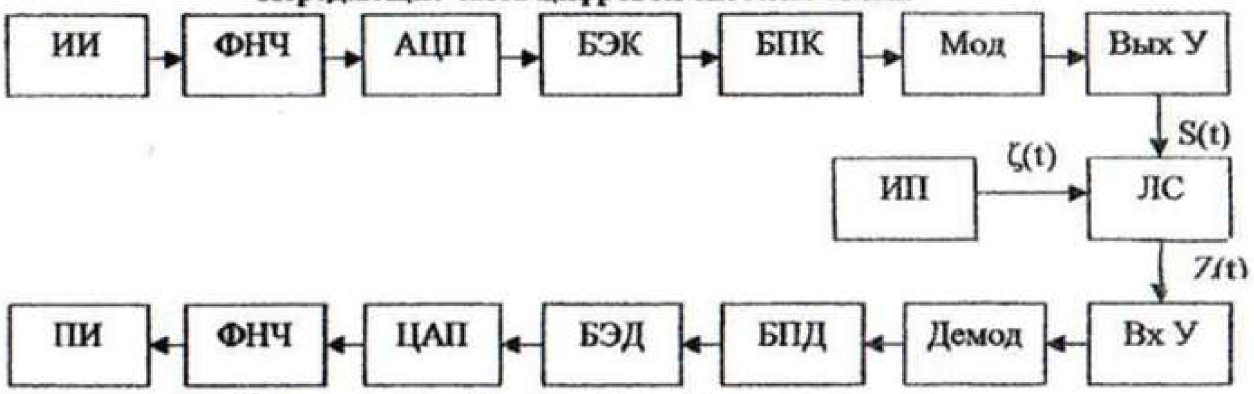
МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ДИСЦИПЛИНЕ ОТС
(часть 2.1. «КОДЕК»)

1. Структурная схема системы связи

Современный этап развития телекоммуникационной техники характеризуется интенсивным внедрением оптимальных способов обработки сигналов. Успехи микропроцессорной техники сделали возможной реализацию самых сложных алгоритмов. Пожалуй, наиболее впечатляющие результаты получены в области кодирования и оптимальной обработки принятых сигналов, что позволило вплотную приблизиться к параметрам идеальной системы связи.

На рис. 1 показана обобщенная структурная схема цифровой системы связи.

Передающая часть цифровой системы связи.



Приемное устройство цифровой системы связи.

Рис. 1

Назначение отдельных блоков будет рассматриваться по мере изучения очередного преобразования сигнала. Сокращенные обозначения на схеме расшифровываются так:

ИИ - источник информации,
ФНЧ – фильтр низких частот с верхней частотой пропускания $F_в$,
АЦП — аналогово-цифровой преобразователь,
БЭК – блок эффективного кодирования,
БПК – блок помехоустойчивого кодирования,
Мод – модулятор несущей,
Вых У – выходное устройство (выходные усилитель и фильтр),
ИП – источник помех $\zeta(t)$,
ЛС – линия (канал) связи,
Вх У – входное устройство (входной фильтр и усилитель приемника),
Демод – демодулятор входного сигнала,

СК – сверточный кодер, ДСК – декодер сверточного кода;
ПМ – перемежитель, ДПМ – деперемежитель;
ЦАП – цифро-аналоговый преобразователь,
ПИ – приемник информации.
Анализ этой схемы показывает, что подавляющая часть преобразований сигнала – это разнообразные виды кодирования.

Кодирование - это замена передаваемого сообщения сочетанием кодовых символов. Общее количество различимых символов, образующих все кодовые комбинации,называется **основанием кода m**.

Количество символов в одной кодовой комбинации называется **длиной кодовой комбинации n.**

Общее количество кодовых комбинаций равно $N=m^n$.
Например, комбинации 00,01,10,11 образуют код, у которого m=2, n=2, N=4; комбинации 0, 1 образуют код, у которого m=2, n=1, N= 2.
Физическая сущность символов может быть разная: 1 и 0 могут быть импульсами разной амплитуды, частоты, фазы, формы.

2.Импульсно-кодовая модуляция (ИКМ)

Из рис.1 следует, что исходный информационный сигнал от источника информации ИИ поступает на АЦП, который превращает аналоговый сигнал в сигнал ИКМ.

ИКМ – это двоичный сигнал (цифровой сигнал), который однозначно, с заданной точностью, соответствует исходному непрерывному сигналу.

Все сигналы делятся на непрерывные (аналоговые) и дискретные (цифровые).

Аналоговые сигналы – это сигналы,которые могут принимать в любой момент времени любые, сколь угодно близкие друг к другу значения. Пример аналогового сигнала - гармоническое колебание.

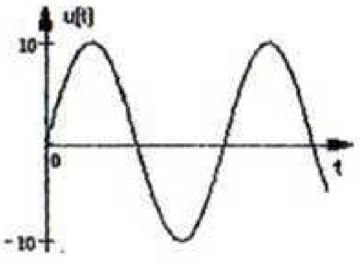


Рис.2

Этот сигнал принимает значения, например, 6 в, 6.01в, 6.0001в и т.д. Таким образом, он может принимать сколь угодно близкие значения.
Дискретный сигнал – это сигнал, который в тактовые моменты времени принимает определенные значения, отличающиеся одно от другого на заданную величину, например, двоичный или троичный сигналы (Рис.3):

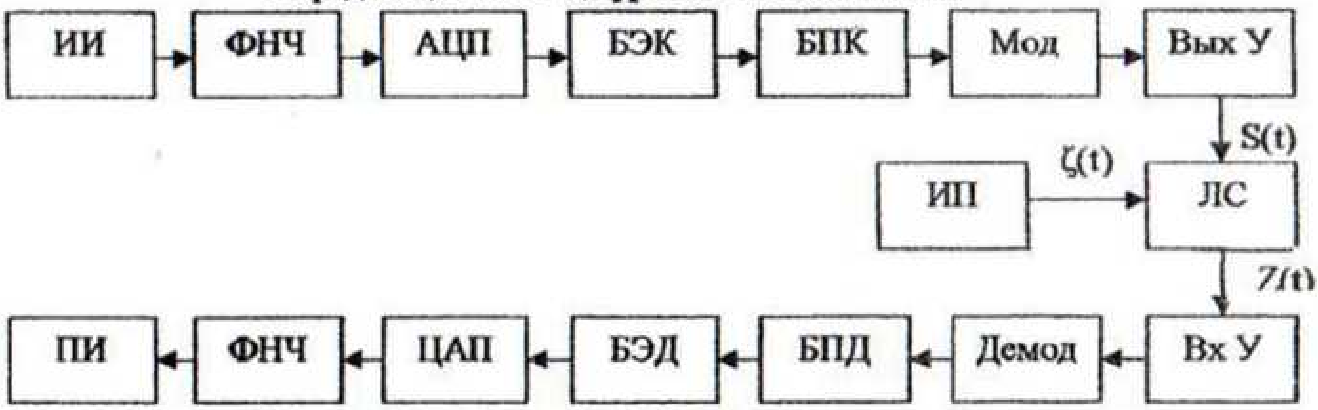
МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ДИСЦИПЛИНЕ ТЭС
(часть 2.1. «КОДЕК»)

1. Структурная схема системы связи

Современный этап развития телекоммуникационной техники характеризуется интенсивным внедрением оптимальных способов обработки сигналов. Успехи микропроцессорной техники сделали возможной реализацию самых сложных алгоритмов. Пожалуй, наиболее впечатляющие результаты получены в области кодирования и оптимальной обработки принятых сигналов, что позволило вплотную приблизиться к параметрам идеальной системы связи.

На рис. 1 показана обобщенная структурная схема цифровой системы связи.

Передающая часть цифровой системы связи.



Приемное устройство цифровой системы связи.

Рис. 1

Назначение отдельных блоков будет рассматриваться по мере изучения очередного преобразования сигнала. Сокращенные обозначения на схеме расшифровываются так:

ИИ - источник информации,
ФНЧ – фильтр низких частот с верхней частотой пропускания $F_в$,
АЦП — аналогово-цифровой преобразователь,
БЭК – блок эффективного кодирования,
БПК – блок помехоустойчивого кодирования,
Мод – модулятор несущей,
Вых У – выходное устройство (выходные усилитель и фильтр),
ИП – источник помех ζ(t),
ЛС – линия (канал) связи,
Вх У – входное устройство (входной фильтр и усилитель приемника),
Демод – демодулятор входного сигнала,

СК – сверточный кодер, ДСК – декодер сверточного кода;
ПМ – перемежитель, ДПМ – деперемежитель;
ЦАП – цифро-аналоговый преобразователь,
ПИ – приемник информации.
Анализ этой схемы показывает, что подавляющая часть преобразований сигнала – это разнообразные виды кодирования.

Кодирование - это замена передаваемого сообщения сочетанием кодовых символов. Общее количество различимых символов, образующих все кодовые комбинации,называется **основанием кода m**.

Количество символов в одной кодовой комбинации называется **длиной кодовой комбинации n.**

Общее количество кодовых комбинаций равно $N=m^n$.
Например, комбинации 00,01,10,11 образуют код, у которого m=2, n=2, N=4; комбинации 0, 1 образуют код, у которого m=2, n=1, N= 2.
Физическая сущность символов может быть разная: 1 и 0 могут быть импульсами разной амплитуды, частоты, фазы, формы.

2.Импульсно-кодовая модуляция (ИКМ)

Из рис.1 следует, что исходный информационный сигнал от источника информации ИИ поступает на АЦП, который превращает аналоговый сигнал в сигнал ИКМ.

ИКМ – это двоичный сигнал (цифровой сигнал), который однозначно, с заданной точностью, соответствует исходному непрерывному сигналу.

Все сигналы делятся на непрерывные (аналоговые) и дискретные (цифровые).

Аналоговые сигналы – это сигналы,которые могут принимать в любой момент времени любые, сколь угодно близкие друг к другу значения. Пример аналогового сигнала - гармоническое колебание.

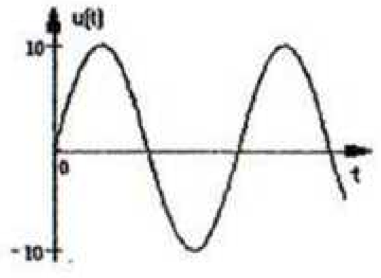
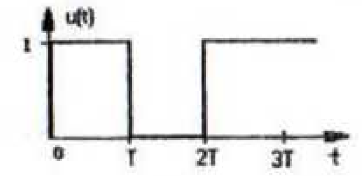


Рис.2

Этот сигнал принимает значения, например, 6 в, 6.01в, 6.0001в и т.д. Таким образом, он может принимать сколь угодно близкие значения.
Дискретный сигнал – это сигнал, который в тактовые моменты времени принимает определенные значения, отличающиеся одно от другого на заданную величину, например, двоичный или троичный сигналы (Рис.3):



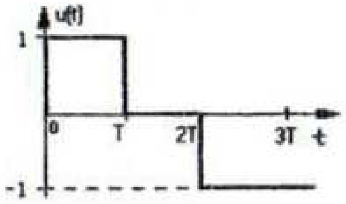


Рис.3

У двоичного сигнала только 2 значения: 0 и 1. Никаких промежуточных значений нет. У троичного сигнала – три значения: 1,0,-1 .

Для перехода от непрерывного сигнала к сигналу ИКМ необходимо выполнить следующие операции (Рис.4):

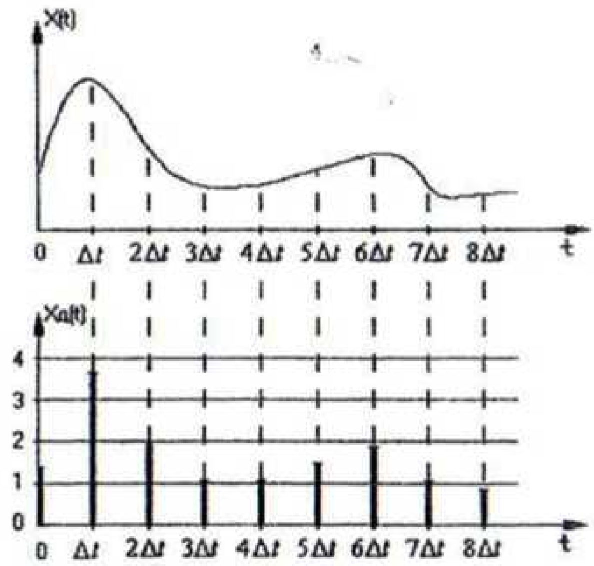


Рис.4

1. Необходимо осуществить дискретизацию аналогового сигнала x(t) в соответствии с теоремой Котельникова. Интервал дискретизации определяется максимальной частотой $F_в$ или $\omega_в$ в спектре x(t):

$$\Delta t = \frac{\pi}{\omega_в} = \frac{1}{2F_в}, \omega_в\left[\frac{рад}{с}\right], F_в[Гц] \quad . \tag{1}$$

В результате дискретизации получим дискретизированный сигнал $x_д(t)$.

2. Необходимо осуществить квантование импульсов - отсчетов по уровню.

Для осуществления квантования диапазон допустимых значений сигнала разбивается на разрешенные уровни квантования, т.о. вместо истинной амплитуды импульса передаем ближайший уровень квантования (Рис.5).

В результате получим квантованный сигнал $x_{кв}(t)$.

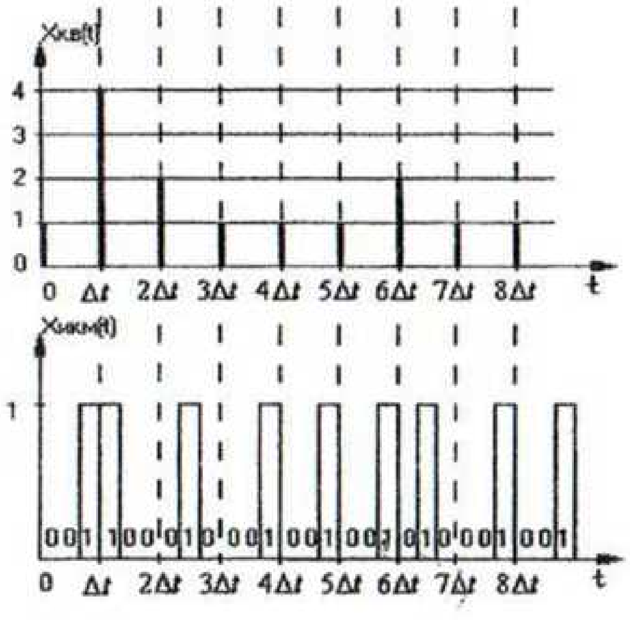


Рис.5

3. Необходимо квантованные импульсы - отсчёты закодировать двоичным кодом.

В простейшем случае номер уровня квантования записываем двоичным числом:

0 В = 000	4 В = 100
1 В = 001	5 В = 101
2 В = 010	6 В = 110
3 В = 011	7 В = 111

Таким образом мы получили двоичный сигнал ИКМ $x_{икм}(t)$, который однозначно соответствует, но с определенной точностью, исходному непрерывному сигналу x(t).

Сигнал ИКМ поступает на вход следующего блока структурной схемы. В соответствии с рис.1 это цифровой фильтр, основные характеристики которого были рассмотрены в первой части курса ТЭС.

Если никаких искажений в сигнале нет, то на выходе демодулятора мы должны получить такой же сигнал ИКМ, который изображен на нижнем рисунке рис.5. Восстановление из сигнала ИКМ исходного непрерывного сигнала осуществляется цифро-аналоговым преобразователем ЦАП, который выполняет следующие операции:

1. Декодирование принятых кодовых комбинаций сигнала ИКМ, т.е. преобразование комбинаций в импульсы-отсчеты соответствующей амплитуды;
2. Восстановление аналогового сигнала путем подачи импульсов-отсчетов на вход ФНЧ, с характеристиками, близкими к характеристикам идеального ФНЧ.

Задача №1

Задана некоторая последовательность букв (например, фамилия студента): Д О Б Р О Л Ю Б О В. Пусть каждая буква фамилии представляет собой импульс-отсчет, амплитуда которого равна порядковому номеру данной буквы в русском алфавите, начиная с нулевого номера: А – 0, Б - 1, В - 2, Г - 3, Д - 4, Ё,Е - 5, Ж - 6, З - 7, И - 8, Й - 9, К - 10, Л - 11, М - 12, Н - 13, О - 14, П - 15, Р - 16, С - 17, Т - 18, У - 19, Ф - 20, Х - 21, Ц - 22, Ч - 23, Ш - 24, Щ - 25, Ъ - 26, Ы - 27, Ь - 28, Э - 29, Ю - 30, Я –31.

Необходимо преобразовать эти отсчеты в сигнал ИКМ. Длина кодовой комбинации равна n=5.

Решение

Заданная фамилия: Д О Б Р О Л Ю Б О В, соответствует некоторому гипотетическому непрерывному сигналу, отсчеты которого в тактовые моменты времени равны: 4в, 14в, 1в, 16в, 14в, 11в, 30в, 1в, 14в, 2в (очевидно, что 4в - это буква Д, 14в – это буква О и т.д.).

Запишем последовательность отсчетов двоичными числами. Для двоичного кода основание m=2, длина кодовой комбинации по заданию равна n=5. Общее количество уровней, которое мы можем закодировать равно $N = m^n = 2^5 = 32$. Получим кодовые комбинации:

4в – 00100, 14в – 01110, 1в – 00001, 16в – 10000, 14в – 01110, 11в – 01011, 30в – 11110, 1в – 00001, 14в – 01110, 2в – 00010.

ДОСТОИНСТВА ИКМ

1. Сигнал ИКМ – цифровой сигнал и поэтому использование ИКМ позволяет реализовать преимущества цифровой аппаратуры по сравнению с аналоговой: большая степень интеграции, унификации и стандартизации, меньше объем аппаратуры, больше точность и стабильность параметров.
2. Сигнал ИКМ имеет более высокую помехоустойчивость благодаря тому, что его можно регенерировать.

НЕДОСТАТКИ ИКМ

1. Сигнал ИКМ имеет ширину спектра больше, чем исходный аналоговый сигнал. Определим ширину спектра сигнала ИКМ, если ширина спектра аналогового сигнала равна F_B. За время $\Delta t = 1/2F_B$ необходимо передать комбинацию сигнала ИКМ из n импульсов. Следовательно, длительность одного импульса равна $T = 1/2nF_B$. Ширина спектра прямоугольного импульса, т.е. ширина спектра ИКМ, примерно, равна:

$$\Pi_{ИКМ} = 1/T = 2nF_B \quad (2)$$

Обычно используются сигналы ИКМ с длиной комбинации 5 – 8. Следовательно, ширина спектра ИКМ в 10-16 раз больше ширины спектра исходного непрерывного сигнала.

2. Квантование импульсов-отсчетов по уровню эквивалентно наложению на сигнал ИКМ помехи, которая называется «шум квантования». Рассчитаем дисперсию шума квантования σ^2, т.е. среднюю мощность шума квантования на единичном сопротивлении.

Мгновенные значения шума квантования равномерно распределены на интервале от $-\Delta/2$ до $+\Delta/2$, т.е. его функция плотности вероятности равна $w(x) = 1/\Delta$.

$\Delta = (U_{max} - U_{min})/(N-1)$ - шаг квантования, т.е. разность между соседними уровнями квантования, где U_{max} и U_{min} – максимальное и минимальное значения непрерывного сигнала, N - количество уровней квантования. Следовательно:

$$\sigma^2 = \int_{-\Delta/2}^{\Delta/2} x^2\, w(x)\, dx = \int_{-\Delta/2}^{\Delta/2} \frac{x^2}{\Delta}\, dx = \frac{\Delta^2}{12} \;. \qquad (3)$$

3. Энтропия

В соответствии со структурной схемой рис. 1, цифровой сигнал после обработки в цифровом фильтре поступает на блок эффективного кодирования. Цель эффективного кодирования - увеличение скорости передачи информации, т.е. увеличение энтропии передаваемого сигнала H.

Энтропия - это среднее количество информации, приходящееся на один символ, посылку.

Пусть вероятность появления некоторого кодового символа равна p, тогда количество информации I, содержащееся в этом символе равно:

$I = -\log_2 p$ двоичных единиц. Например, $p = 0.25$, тогда $I = -\log_2 0.25 = 2$ дв. ед. (бит). Т.к. источник информации производит различные символы, имеющие разную вероятность, то для характеристики информационной содержательности сообщения источника используют среднее количество информации на символ, т.е. энтропию.

Рассмотрим дискретный источник информации, который производит последовательность кодовых символов, соответствующих передаваемой информации. Объем алфавита дискретного источника - n. Вероятность появления i-го символа из этого набора – p_i. Количество информации, содержащееся в i-ом символе равно $-\log_2 p_i$.

Чтобы найти среднее количество информации в одном символе, надо усреднить $-\log_2 p_i$ с учетом вероятности появления символов p_i.

Если символы источника независимы, то энтропия дискретного источника H равна

$$H = (1/n)\Sigma_n p_i \log_2 p_i \qquad (4)$$

Энтропия максимальна, если символы источника независимы и равновероятны, т.е. $p_1 = ... = p_n = 1/n$.

Максимальное значение энтропии равно

$$H_{max} = \log_2 n, \text{ дв. ед./символ, (бит/символ).} \qquad (5)$$

При $n = 2$; $H_{max} = 1$дв. ед./символ; при $n = 16$; $H_{max} = 4$ дв. ед./символ.

Для основания кода $n = 2$ энтропия вычисляется по формуле

$H_2 = - p_0 \log_2 p_0 - p_1 \log_2 p_1 = (1 - p_1)\log_2(1 - p_1) - p_1 \log_2 p_1.$

Здесь p_1 - вероятность передачи 1, p_0 - вероятность передачи 0.

Если $p_0 = 1$, $p_1 = 0$ (передается только 0) или $p_0 = 0$, $p_1 = 1$ (передается только 1), то H = 0, т.к. ситуация детерминированная.

Чем больше неопределенность, тем больше энтропия - Н. Когда вероятности 1 и 0 равны 0.5, то неопределенность максимальна, энтропия также максимальна и равна 1 дв. ед./символ (Рис.6).

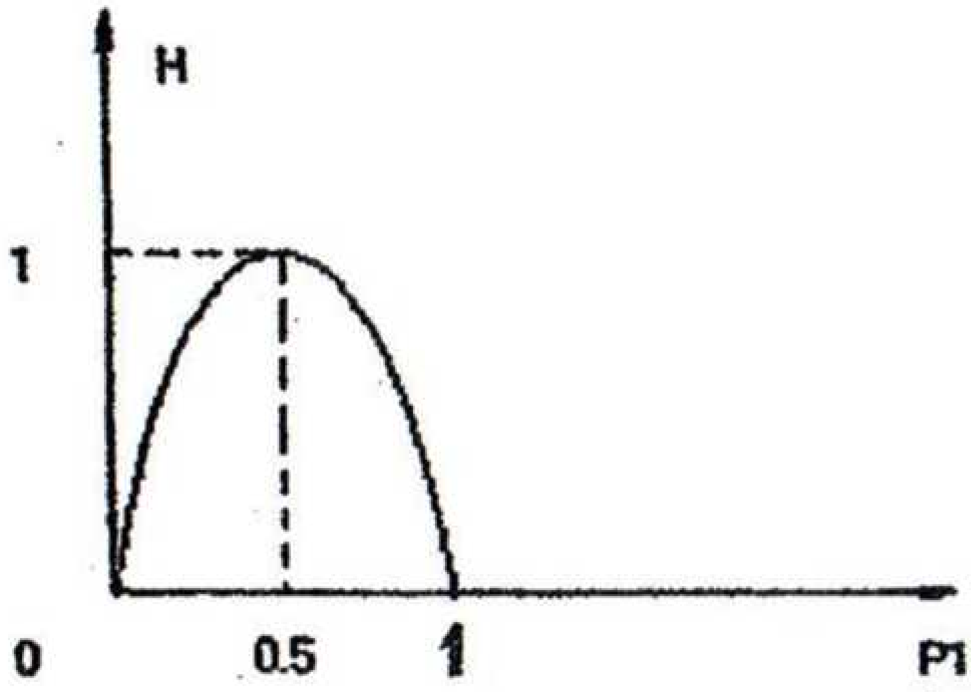


Рис. 6

Наша задача получить максимальную энтропию. Чем больше энтропия, тем быстрее можно передать заданное количество информации. Например, надо передать количество информации I = 1000 дв. ед. Если энтропия H = 0.5 дв. ед./символ, то для передачи заданного количества информации нужно передать N = 1000/0.5 = 2000 символов. Если длительность символа T = 1мс, то время передачи заданного количества информации $T_N = NT = 2000$ мс.

Если H = 1 дв. ед./символ, то N = 1000 символов и $T_N = 1000$ мс. Увеличив энтропию в два раза, мы в два раза уменьшили время передачи сообщения.

Способы увеличения энтропии

1) Наличие корреляционных связей между символами уменьшает энтропию. Для увеличения энтропии используют различные способы предсказания или укрупняют алфавит, кодируя не символы, а целые слова.
2) Неравновероятность символов уменьшает энтропию. Для увеличения энтропии сообщения перекодируют так, чтобы символы нового кода были равновероятны.
   Для этого наиболее вероятные сообщения кодируются наиболее короткими кодовыми комбинациями.
3) Дальнейшее увеличение энтропии достигается увеличением основания кода m.

4.Эффективное (статистическое) кодирование. Кодирование с предсказанием

Между отдельными буквами в словах, символами в кодовых комбинациях существуют корреляционные связи (отметим, что кодовые комбинации также называют кодовыми словами). Это значит, что вероятность появления следующей буквы, символа зависит от того, какие буквы, символы были переданы до этого.

Для устранения корреляционных связей между символами используют линейное кодирование с предсказанием (LPC-кодер).

Слово или сегмент в последовательности из n букв, символов описывается корреляционной матрицей.

$$M = \begin{vmatrix} R_{11} & R_{12} & R_{13} \ldots R_{1n} \\ R_{21} & R_{22} & R_{23} \ldots R_{2n} \\ . & . & . \quad . \\ . & . & . \quad . \\ R_{n1} & . & . \quad R_{nn} \end{vmatrix} \qquad (6)$$

R_{ik} - коэффициент корреляции между i-ой и k-ой буквами.

Зная предыдущие буквы, можно предсказать последующие. Предсказанное значение x_k запишем в виде:

$$\widetilde{x}_k = c_1x_1 + c_2x_2 + \ldots\ldots + c_{k-1}x_{k-1} \quad , \qquad (7)$$

x_j - буквы, символы, которые мы передаём.

Ошибка предсказания равна:

$$x_k - \widetilde{x}_k = x_k - (c_1x_1 + c_2x_2 + \ldots\ldots + c_{k-1}x_{k-1}) \; .$$

Коэффициенты c_j подбираются таким образом, чтобы обеспечить минимум среднеквадратической ошибки предсказания.

Оптимальное значение c_j для предсказания x_k имеет вид:

$$c_{j\,опт} = -A_{jk}/D_{k-1}\,, \qquad (8)$$

A_{jk} - алгебраическое дополнение матрицы для элемента R_{jk}.

D_{k-1} - определитель матрицы (k - 1)-го порядка.

Рассмотрим простейший случай предсказания к-го символа по (к - 1)-му

$$x_k = c_{k-1}x_{k-1}.$$

В линию связи передаётся ошибка предсказания

$$\Delta x = x_k - x_k.$$

Это позволяет уменьшить динамический диапазон передаваемого сигнала, т.е. уменьшить необходимое количество уровней квантования.

5. Эффективное кодирование. Укрупнение алфавита источника

Уменьшить корреляционные связи между буквами, символами сообщения можно путем укрупнения алфавита источника. Как уже указывалось выше, это означает, что мы будем кодировать не буквы, а целые слова, сегменты информационной двоичной последовательности.

Удобно рассмотреть эту операцию на конкретном примере.

Пусть задан двоичный источник дискретных сообщений, который производит слова из двух букв А и М: МА, АМ, АА, ММ.

Между буквами существуют статистические связи. Между словами корреляционных связей нет. Рассчитаем энтропию такого источника, если его вероятностные характеристики следующие:

p(A) = 0.7;p(M) = 0.3; p(A/A) = 0.8; p(M/A) = 0.2; p(A/M) = 0.7;
p(M/M) = 0.3.

Определим вероятность каждого слова:

p(AM) = p(A)p(M/A) = 0.7*0.2 = 0.14
p(MA) = p(M)p(A/M) = 0.33*0.7 = 0.21
p(MM) = p(M)p(M/M) = 0.3*0.3 =0.09
p(AA) = p(A)p(A/A) = 0.7*0.8 = 0.56

Т.к. буквы коррелированны, то энтропия источника коррелированных букв рассчитывается следующим образом:

1) Определяем количество независимых слов: N=4.

2) Рассчитаем вероятность каждого слова (см. выше).

3) Т.к. слова независимы, то можно рассчитать энтропию $H_{сл}$ на слово по простейшей формуле:

$H_{сл} = -0.14*\log_2 0.14 - 0.21*\log_2 0.21 - 0.09*\log_2 0{,}09 - 0.56*\log_2 0.56 =$
$= -1.656$ дв. ед./слово

4) Для определения энтропии H на букву надо разделить полученное значение $H_{сл}$ на количество букв в слове (в данном случае на 2):

$H = 0.828$ дв. ед./символ.

Т.к. наш источник двоичный (производит только 2 буквы), то его максимальная энтропия равна

$H_{max} = \log_2 2 = 1$ дв. ед./символ.

Энтропия нашего источника меньше максимальной.

Степень отличия H от H_{max} характеризуется избыточностью источника:

$$K = 1 - H/H_{max} \ . \qquad (9)$$

В данном случае $R = 1 - 0.828/1 = 0.172$.

Наша цель - увеличить энтропию этого источника до $H_{max} = \log_2 n$. Для устранения корреляционных связей между буквами будем кодировать целые слова исходного источника.

Т.к. разных слов - 4, то нужно использовать код с основанием $n = 4$. АМ - первое слово S_1 кодируем символом 0, МА - второе слово S_2 кодируем символом 1, ММ - третье слово S_3 кодируем символом 2, АА - четвертое слово S_4 кодируем символом 3.

Получили новый код с основанием $n = 4$, длиной комбинации $m = 1$, общее число комбинаций $N = n^m = 4$.

Сообщения этого кода независимы, т.к. независимы слова (между словами нет корреляции). Поэтому энтропия рассчитывается по формуле для дискретного источника независимых сообщений:

$$H = -P(S_1)\log_2 P(S_1) - P(S_2)\log_2 P(S_2) - P(S_3)\log_2 P(S_3) - P(S_4)\log_2 P(S_4) =$$
$$= -0.14*\log_2 0.14 - 0.21*\log_2 0.21 - 0.09*\log_2 0.09 - 0.56*\log_2 0.56 = 1.656$$
(дв. ед./сообщение).

Вместо «сообщение» можно употребить термин «буква», т.к. буквами нового источника являются слова первичного источника.

Ранее для исходного первичного источника было $H = 0.828$ дв. ед./символ.

Таким образом, мы увеличили энтропию в 2 раза.

Определим избыточность нового источника, для которого $H_{max} = \log_2 4 = 2$ дв. ед./символ

$R = 1 - H/H_{max} = 1 - 1.656/2 = 0.172$.

Избыточность осталась та же, т.е. есть ещё возможности увеличения энтропии.

6. Эффективное кодирование.
Кодирование источника неравновероятных независимых сообщений

Энтропия нового четверичного кода, полученного в предыдущей главе, не максимальна, т.к. сообщения неравновероятны. Закодируем 4 наших сообщения новым двоичным кодом таким, чтобы символы этого нового двоичного кода были равновероятными. Для этого построим код с префиксными свойствами по алгоритму Хаффмана:

никакая короткая кодовая комбинация не должна являться началом более длинной комбинации, тогда не нужно разделительных символов.

Алгоритм Хаффмена предполагает построение «кодового дерева».

Алгоритм построения кодового дерева:

1. Располагаем сообщения в порядке убывания вероятностей.
2. Объединяем два наименее вероятных сообщения в одно с суммарной вероятностью появления (т.е. складываем вероятности объединяемых сообщений).
3. Вновь располагаем сообщения в порядке убывания вероятностей и т.д., пока не получим в сумме единицу.

Построим кодовое дерево для нашего примера:

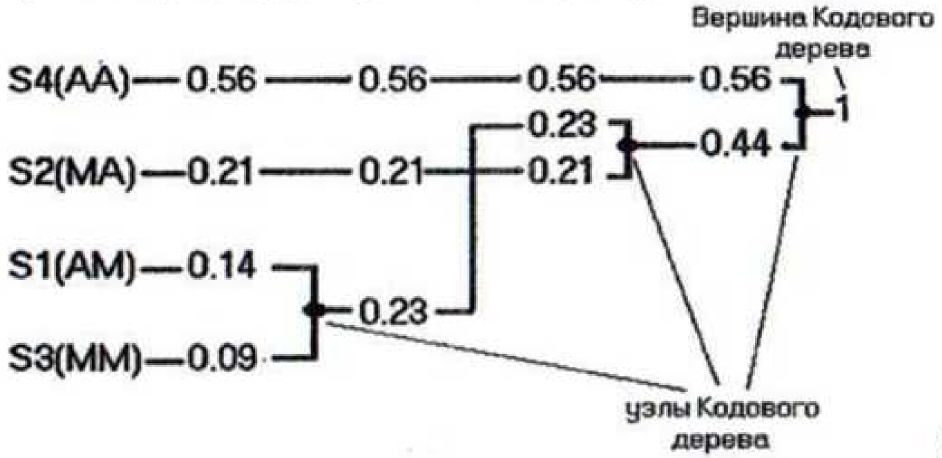


Рис.7

Алгоритм кодирования слов новым двоичным кодом следующий:

-идём от вершины кодового дерева к сообщению,

-если в узле мы идём вверх, то в кодовую комбинацию записывается единица, если вниз – ноль. В результате получим:

S4 => "1" ; S2 => "00" ; S1 => "011" ; S3 => "010" .

Мы получили код с префиксными свойствами.

Рассчитаем энтропию нового двоичного кода. Для этого надо определить вероятности нулей и единиц в новом коде.

Из 100 среднестатистических сообщений мы имеем сообщений:

S1 => 14 штук, т.е. 14 комбинаций «011»;

S2 => 21 штука, т.е. 21 комбинация «00»;

S3 => 9 штук, т.е. 9 комбинаций «010»;

S4 => 56 штук, т.е. 56 комбинаций «1».

Таким образом, в 100 сообщениях содержится «единиц»:

$N1 = 2*14+9*1+56*1 = 93$.

Содержится «нулей»:

$N_0 = 1*14 + 21*2 + 9*2 = 74.$

Вероятность появления единиц и нулей:

$p(1) = N_1/(N_1+N_0) = 93/167 = 0.557; \quad p(0) = 1 - p(1) = 0.443.$

Энтропия нового двоичного источника H:

$H = -p(1)*\log_2 p(1) - p(0)*\log_2 p(0) = -0.557*\log_2 0.557 - 0.443*\log_2 0.443 =$
$= 0.994$ (дв. ед./символ).

Избыточность нового двоичного источника существенно уменьшилась $R = 1 - 0.994 = 0.006$. Определим среднюю длину кодовой комбинации:

$$n_{cp} = \sum_{k=1}^{N} p_k n_k , \qquad (10)$$

p_k - вероятность k-го сообщения (слова); n_k - длина кодовой комбинации k-го сообщения.

Для нашей задачи получим

$n_{cp} = 0.56*1 + 0.14*3 + 0.21*2 + 0.09*3 = 1.67$ двоичных символов.

Задача № 2

Последовательность двоичных импульсов сигнала ИКМ, полученная в задаче № 1, состоит из «дибитов» («дибит» - это комбинации из 2-х символов: 00, 01, 10, 11). Необходимо определить вероятность каждого «дибита» и закодировать их безызбыточным кодом с префиксными свойствами.

Решение.

В задаче №1 получили последовательность, которую разделим на «дибиты»

00100011 10000011000001 11001011 11 1100000101 11 000010.

Всего 25 «дибитов». Комбинация 00 встречается 11 раз, т.е. вероятность $p(00) = 11/25 = 0.44$; аналогично получим: $p(01) = 3/25 = 0.12$; $p(10) = 4/25 = 0.16$; $p(11) = 7/25 = 0.28$. Дальнейшее решение задачи соответствует примеру, рассмотренному выше в разделе 6.

7. Эффективное кодирование. Увеличение энтропии путем увеличения основания кода

Для дальнейшего увеличения энтропии снова увеличим основание кода m. Технически эта операция реализуется следующим образом. Поток двоичных посылок разбивается на комбинации из n бит. Общее количество разных комбинаций равно 2^n. Каждой комбинации ставится в соответствие один из символов кода с основанием $m = 2^n$. Так как все символы благодаря вышеописанным операциям равновероятны и независимы, то каждая посылка m-ичного кода несет n двоичных единиц информации.

«Дибиты» двоичного кода 00,01,10,11 будем кодировать четверичным кодом: 00 - закодируем символом 0; 01 - закодируем символом 1;

10 - закодируем символом 2; 11 - закодируем символом 3.

Каждый символ нового четверичного кода несёт уже не 1 бит, а 2 бита, т.к. при m=4 H_{max} = log 4 =2бит/символ.

Два символа двоичного кода длительностью 2Т несут, максимум, 2 бита информации (m=2; H_{max}=1бит/символ; n=2; I = n * H_{max} = 2бита). Один символ четверичного кода длительностью Т несет тоже 2 бита информации (m=4; H_{max}=2бит/символ ; n=1; I = n * H_{max} = 2 бита). Следовательно, мы в 2 раза увеличили скорость передачи информации. При этом помехоустойчивость приёма уменьшается.

Более подробно этот вопрос рассматривается в разделе «Модуляция».

8.Эффективные способы передачи

Описанные методы нашли широкое применение в современных модификациях ИКМ.

Блочная ИКМ

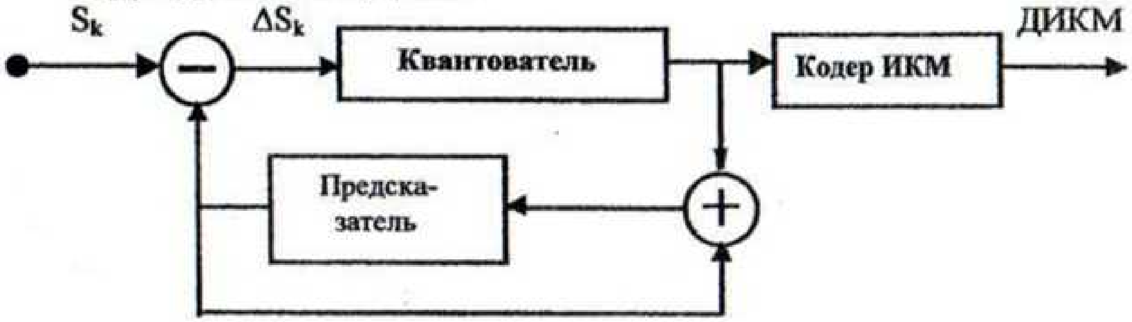
Блочная ИКМ - система с почти мгновенным компандированием.

Основная идея системы состоит в том, что в каждом сегменте цифрового сигнала определяется самая большая амплитуда отсчета, и этот отсчёт кодируется хорошим помехозащищённым кодом. Он определяет масштаб для этого сегмента. Остальные уровни кодируются исходя из этого масштаба. Это позволяет обеспечить высокое отношение сигнал / шум квантования во всём диапазоне изменения сигнала.

Дифференциальная ИКМ (ДИКМ)

Этот способ передачи состоит в вычислении ошибки предсказания. Ошибка предсказания кодируется при меньшем числе уровней квантования.

Структурная схема ДИКМ.



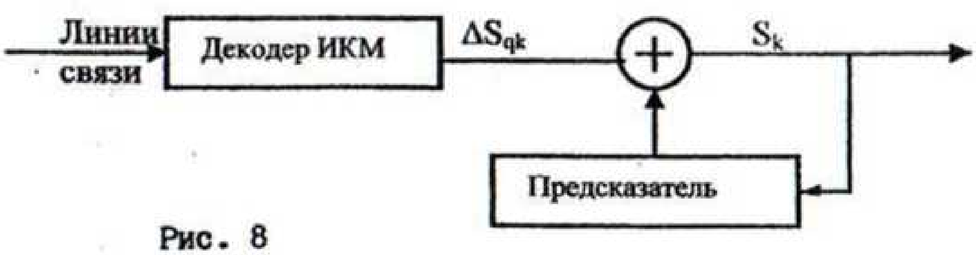


Рис. 8

Предсказатель - это трансверсальный фильтр. Его структурная схема показана на рис.9.

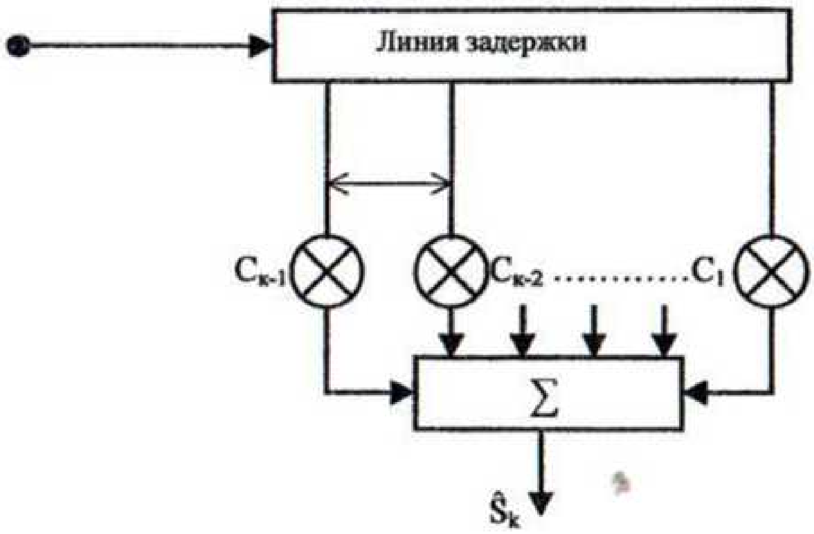


Рис.9

Дельта - модуляция (ДМ)

При ДМ в тактовый момент времени передаётся только знак изменения функции по сравнению с предыдущим отсчётом (Рис.10).

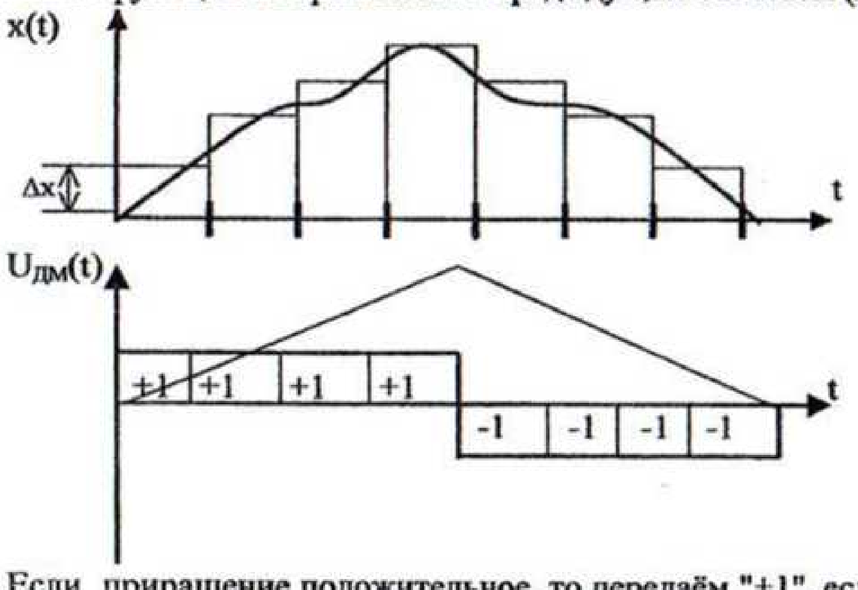


Рис.10

Если приращение положительное, то передаём "+1", если приращение отрицательное, то передаём "-1".

переданной комбинации 000 (d=1), чем к другой возможной 111 (d=2). Таким образом, мы декодируем комбинацию 010 как 000, т.е. исправляем ошибку. Выигрыш в помехоустойчивости достигается за счет еще большего проигрыша в скорости передачи, т.к. два сообщения мы могли бы передавать с помощью двух комбинаций безызбыточного кода с m=2, n=1: т.е. 0 и 1.

Проигрыш по скорости передачи данного кода, исправляющего одиночные ошибки, по сравнению с безызбыточным кодом равен 3. Для исправления ошибок с кратностью k следует использовать коды с минимальным кодовым расстоянием $d_{min}=2k+1$.

Для обнаружения одиночных ошибок одним из наиболее совершенных способов кодирования является «проверка на четность»: к кодовой комбинации из k информационных символов добавляется один проверочный такой, чтобы количество единиц в кодовой комбинации было четным. Например, к комбинации 0100110 добавляем проверочный символ 1, и передаем комбинацию 01001101. Одиночная ошибка делает число единиц в принятой кодовой комбинации нечетным (3 или 5), что и обнаруживается на приеме.

10.Линейный двоичный блочный код

Широко используются в технике связи линейные блочные коды. Блочный код состоит из кодовых комбинаций, называемых также кодовыми словами. Кодовые комбинации линейного кода образуют линейное пространство относительно поразрядного сложения по модулю 2: сумма двух комбинаций дает тоже комбинацию этого кода.

Длина каждого кодового слова равна n, т.е. общее количество кодовых комбинаций равно $N=m^n$. Если код –двоичный, то $N=2^n$.

Код - систематический, т.е. k символов являются информационными, а (n - k) являются корректирующими. Блочный код обозначается, как код (n,k).

Разрещёнными являются только 2^k комбинаций, остальные – запрещённые. Величина R=k/n называется скоростью кода. Каждое кодовое слово имеет свой вес w, т.е. количество ненулевых элементов. Для двоичного кода вес w равен количеству единиц в слове. Операции кодирования и декодирования выполняются в соответствии с правилами, определенными для алгебраического поля с заданными элементами. При m=2 поле имеет два элемента 0 и 1. Алгебраические поля с конечным числом элементов называются конечными полями Галуа GF(m). Для каждого m определяются правила умножения и сложения.

Для m=2:

+	0	1		*	0	1
0	0	1		0	0	0
1	1	0		1	0	1

Алгоритм кодирования

Рассмотрим алгоритм кодирования для двоичного блочного кода (7,3), у которого каждое слово имеет 7 символов, из которых 3 – информационные и 4 – проверочные.

Алгоритм формирования кодовых комбинаций следующий :

1.Присваиваем каждому символу кода номер : $a_1, a_2, a_3, a_4; a_5, a_6, a_7$.

Первые три символа (a_1, a_2, a_3) являются информационными. Последние четыре символа - корректирующие (проверочные): a_4, a_5, a_6, a_7.

2. Составляем порождающую матрицу G. Эта матрица должна иметь n столбцов и k строк. Левая часть матрицы – это единичная матрица размером k*k. Правая часть G – это матрица-дополнение P размером (n-k)*k :

$$G = \underbrace{\left|\begin{matrix}1&0&0\\0&1&0\\0&0&1\end{matrix}\right.}_{\text{единичная матрица}}\underbrace{\left|\begin{matrix}0&1&1&1\\1&0&1&1\\1&1&0&1\end{matrix}\right|}_{\text{матрица - дополнение}} \tag{11}$$

Матрица-дополнение P имеет вид :

$$P = \begin{vmatrix}0&1&1&1\\1&0&1&1\\1&1&0&1\end{vmatrix} \tag{12}$$

3. Формируем кодовые комбинации. Для этого сначала записываем все возможные информационные комбинации из трех символов (всего восемь комбинаций): **000,001,010,011,100,101, 110,111**.

4. К информационным символам приписываем четыре проверочных символа, получающихся в результате умножения информационного вектора-строки ($a_1a_2a_3$) на матрицу-дополнение P. Произведение есть вектор-строка ($a_4a_5a_6a_7$):

$(a_1a_2a_3) * P = (a_4, a_5, a_6, a_7)$.

Очевидно, для заданной матрицы P: $a_4 = a_2 \oplus a_3$; $a_5 = a_1 \oplus a_3$;

$a_6 = a_1 \oplus a_2$; $a_7 = a_1 \oplus a_2 \oplus a_3$.

Знак ⊕ означает суммирование по модулю 2, т.е. $0 \oplus 0 = 0$; $1 \oplus 0 = 1$; $0 \oplus 1 = 1$; $1 \oplus 1 = 0$.

Составим кодовую таблицу разрешенных кодовых комбинаций:

Структурная схема декодера

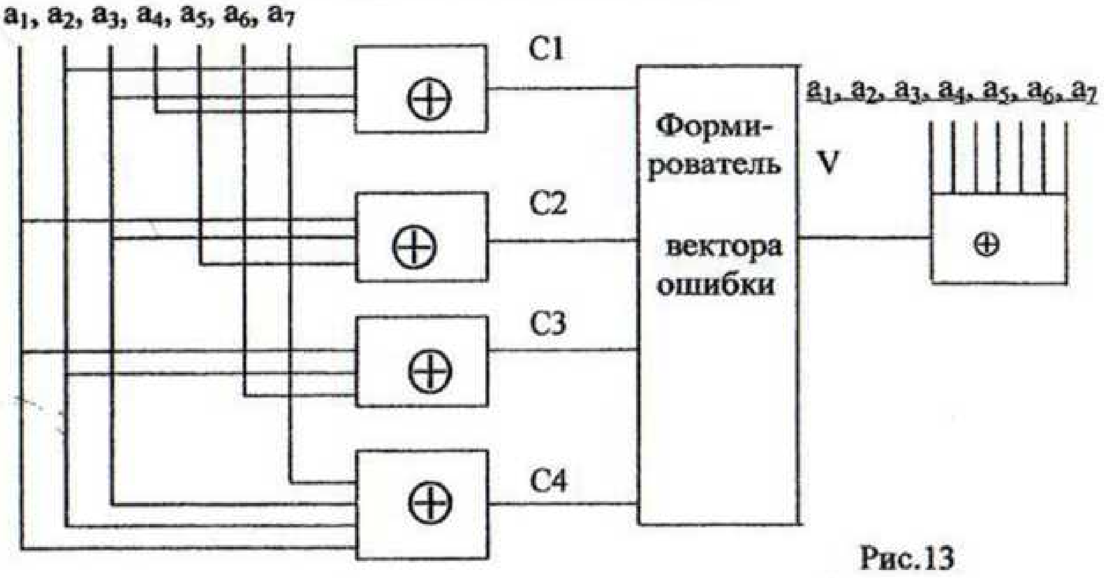


Рис.13

ПРИМЕРЫ БЛОЧНЫХ КОДОВ

В аппаратуре связи используются различные блочные коды:

- коды Хемминга, для которых $n = 2^m-1$, $k = 2^m-1-m$; $d_{min}=3$;
- коды Адамара, для которых кодовые слова есть столбцы матрицы Адамара A_n ($n=2^m$, $k=m+1$, $d_{min}=2^{m-1}$);
- код Голея (23,12), $d_{min}=7$.

11.Циклический код

Характерной особенностью этих кодов является то, что циклическая перестановка символов одной комбинации, например, 1001011 дает новую комбинацию того же кода 1100101.

Теория циклических кодов базируется на теории двоичных полиномов. Каждая комбинация записывается в виде двоичного полинома степени (n-1) с коэффициентами $a_к = 0$ или 1:

$$A(Z) = a_{n-1}Z^{n-1} + a_{n-2}Z^{n-2} + \ldots + a_1Z + a_0 \;.$$

Например:

$$\begin{array}{ccc} 1 & 0 & 1 \\ a_2 & a_1 & a_0 \end{array} \Rightarrow A(Z) = \underset{1}{\underset{|}{a_2}}Z^2 + \underset{0}{\underset{|}{a_1}}Z + \underset{1}{\underset{|}{a_0}} = Z^2 + 1 \;.$$

Алгоритм формирования циклического кода на примере кода (7, 4).

Комбинации данного циклического кода состоят из 7 символов, из которых 4 символа информационные и 3 – проверочные.

1) Записываем возможные информационные комбинации из 4-х символов: 0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111.

2) Каждую комбинацию записываем в виде полинома. Например:

$$\underset{a_3}{1}\ \underset{a_2}{0}\ \underset{a_1}{0}\ \underset{a_0}{1} \Longrightarrow A(Z) = Z^3 + 1$$

3) Из таблицы выбираем образующий полином P(Z) , степень которого соответствует количеству проверочных символов.

В данном случае количество проверочных символов n-k=3.

Полином: $P(Z) = Z^3 + Z^2 + 1 \Longrightarrow 1\ 1\ 0\ 1$

4) Полином, соответствующий каждой информационной комбинации, умножается на P(Z):

$$(Z^3 + 1)(Z^3 + Z^2 + 1) = Z^6 + Z^5 + Z^2 + 1 \Longrightarrow 1100101$$

В результате получим 16 комбинаций циклического кода (7,4): 0000000, 0001101, 0011010 и т.д.....1001011.

Минимальное кодовое расстояние равно 3, т.е. данный код исправляет все одиночные ошибки.

Алгоритм декодирования циклического кода на примере кода(7,4).

1) Принятая кодовая комбинация делится на образующий полином. Остаток от деления есть синдром, который указывает на позицию, где произошла ошибка. Т.к. синдром не зависит от передаваемой комбинации, а зависит только от позиции, в которой произошла ошибка, то синдромы можно вычислить заранее. Например, передавали комбинацию 0000000, под действием помехи она превратилась в 0100000, т.е. ошибка в 6-ом символе справа. Разделим 0100000 на P(Z)=1101:

```
0100000 | 1101
0000      0111
 1000
 1101
  1010
  1101
   1110
   1101
    011
```

Остаток 011 и есть синдром, указывающий, что ошибка произошла в 6-ом символе справа.

2) В соответствии с синдромом формируется вектор ошибки, т.е. кодовая комбинация, которая содержит 1 в той позиции, где произошла ошибка. Для данного примера вектор ошибки V=0100000.

3)Вектор ошибки суммируется по модулю 2 с принятой комбинацией:

$0100000 \oplus 0100000 = 0000000$

Ошибка исправлена.

СТРУКТУРНАЯ СХЕМА КОДЕРА ЦИКЛИЧЕСКОГО КОДА (7,4)

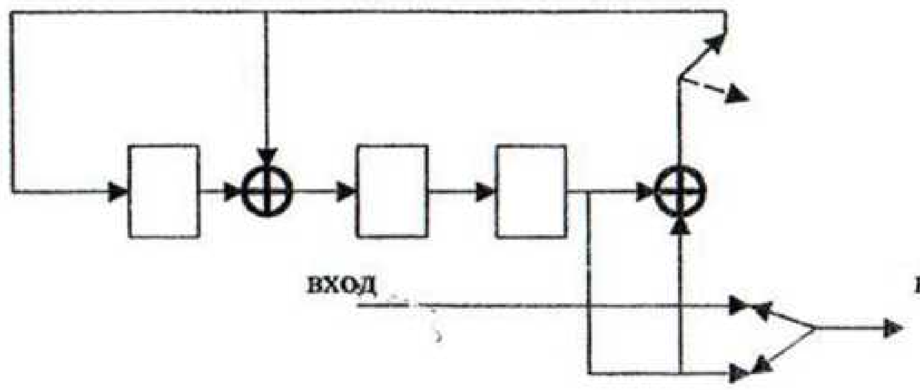


Рис.14

Кодер содержит два сумматора по модулю 2($\oplus$), три элемента памяти и два переключателя. Четыре информационных бита поступают на вход схемы и одновременно на выход (переключатели находятся в верхнем положении). После этого переключатели переходят в нижнее положение и на выход идут три проверочных бита.

СТРУКТУРНАЯ СХЕМА ДЕКОДЕРА ЦИКЛИЧЕСКОГО КОДА (7,4)

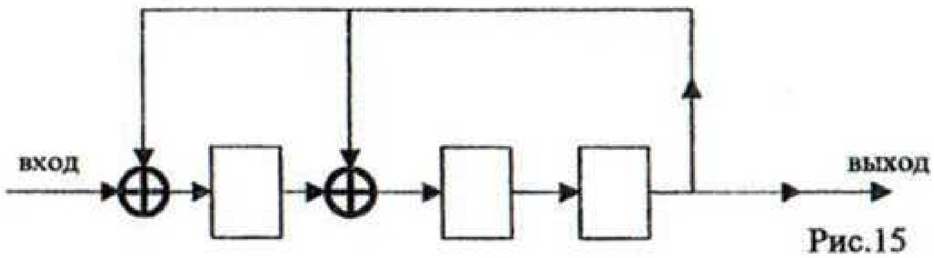


Рис.15

Семь бит (четыре информационных бита и три проверочных) поступают на вход декодера (ключ – в верхнем положении). После этого ключ переходит в нижнее положение и синдром поступает на выход схемы. Синдром поступает на блок формирования вектора ошибки и сумматор по модулю 2 вектора ошибки и принятой комбинации (см. Рис. 13).

ПРИМЕРЫ ЦИКЛИЧЕСКИХ КОДОВ

В аппаратуре связи используются различные типы циклических кодов:

- циклические коды Хемминга (2^m-1, 2^m-1-m);
- циклический код Голея (23,12);
- коды сдвигового регистра максимальной длины (см. [2]);

- БЧХ коды (2^m-1, 2^m-1-mt), d_{min}=2t+1, где m>2, t- целые положительные числа.

12. Сверточный код (решетчатый)

Свёрточный код относится к классу непрерывных кодов, у которых выходные символы – это функции входных символов и конструкции кодера. Сверточный кодер представляет собой линейный сдвиговый регистр с конечным числом состояний. Кодер содержит ячейки памяти и сумматоры по модулю 2 (Рис.16). Под действием тактовых импульсов информационные импульсы продвигаются к следующей ячейке. С выходов регистра импульсы поступают в блок выходных сумматоров. За один такт коммутатор последовательно проходит все n выходов и дает n импульсов. Кодовая скорость равна R=k/n.

Длина регистра K, т.е. в простейшем случае количество ячеек, называется длиной кодового ограничения. На рис.16 показан сверточный кодер с R=1/2, длина кодового ограничения K=3. Параметр K характеризует, на какое количество выходных символов влияет данный входной.

Выходная битовая последовательность может быть определена путем перемножения информационного полинома на порождающие полиномы кодера. Т.к. кодер имеет два выхода, то следует составить два порождающих полинома. Если ячейка подключена к выходу, то в порождающем полиноме соответствующий коэффициент равен 1, если не подключена, то равен 0.

Порождающие полиномы для данного кодера равны: $g_1(D)=1+D^2$; $g_2(D)=1+D+D^2$.

Эти полиномы записывают сокращенно как одну восьмиричную цифру: код (5,7). Чтобы найти выходную последовательность, надо или проследить по схеме, или перемножить входной полином на g_1 и g_2 и совместить их.

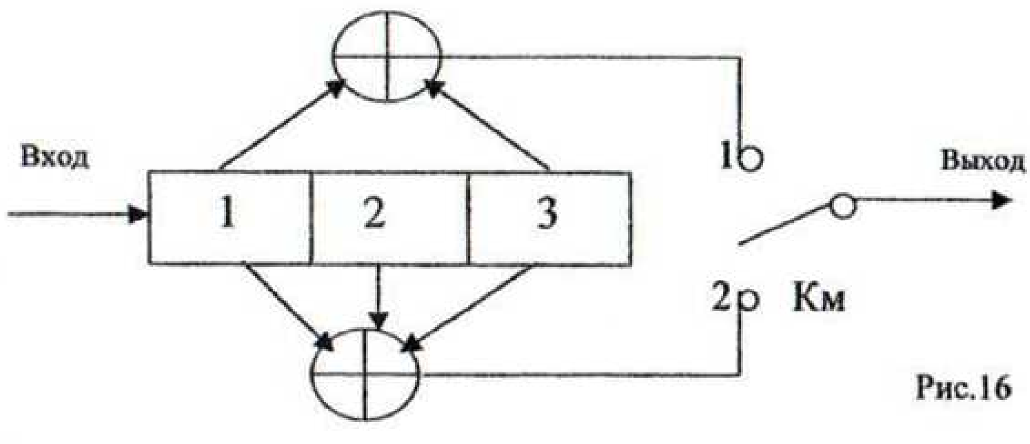


Рис.16

Сверточный код – частный случай решетчатых кодов. Решетчатая диаграмма на рис.17 для сверточного кода, формируемого кодером рис.16 - это один из способов задания сверточного кода. Каждое сечение соответствует внутреннему состоянию кодера; ребро соответствует передаваемому символу (верхнее – 0; нижнее - 1) на входе. Около ребра написана комбинация из двух бит, которая появится на выходе.

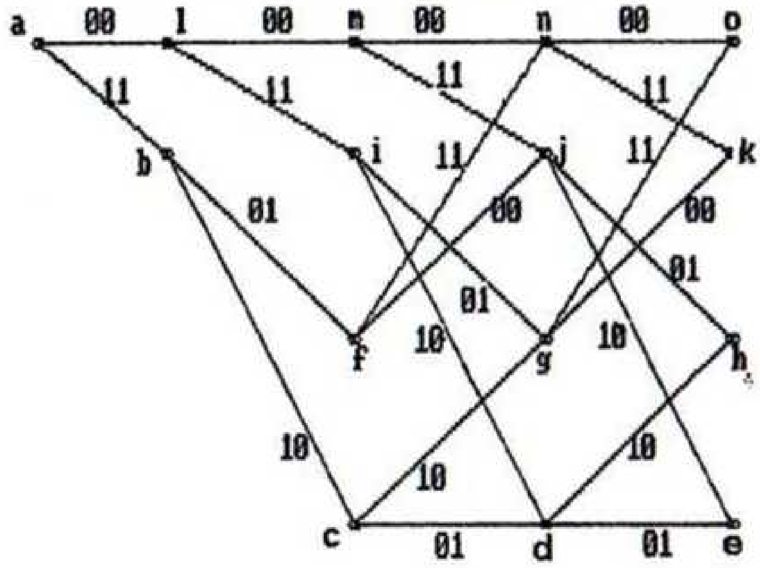


Рис.17

Пусть на вход кодера рис.16 поступила комбинация 101. Ей соответствует полином $X(D)=1+D^2$. На первом выходе кодера получим $Y_1(D)=X(D)*g_1(D)=(1+D^2)*(1+D^2)=1+D^4$. Этот полином соответствует комбинации 10001.

На втором выходе кодера получим $Y_2(D)=X(D)*g_2(D)=(1+D^2)*\ )*(1+D+D^2)=1+D+D^3+D^4$. Это соответствует комбинации 11011.

Следует учесть, что сложение коэффициентов осуществляется по модулю 2, т.е. $D+D=(1\oplus1)D=0$.

Совместим эти комбинации: сначала передается первый символ с первого выхода, потом первый символ со второго выхода, потом второй символ с первого выхода, потом второй символ со второго выхода и т.д. Получим 110100……

Аналогичный результат получим с помощью решетчатой диаграммы: при передаче 1 идем по нижнему ребру и в канал связи передается 11, далее передается 0, идем по верхнему ребру и в канал передается 01, далее передается 1, а в канал идет 00 и т.д.

Структурная схема декодера сверточного кода

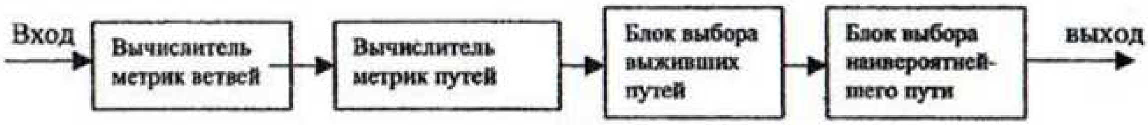


Рис.18

Декодирование сверточного кода осуществляется в соответствии со схемой рис.18. Пусть передавали 110100…., приняли 100100…Ошибка во 2-ом символе. Покажем, как декодер исправляет эту ошибку.

Декодирование осуществляется поэтапно путем анализа n бит. Для рассматриваемого кодера n=2. Приняли первые два бита 10. Кодовое расстояние между 10 и путем al, которому соответствует 00 по решетчатой диаграмме, равно d=1(метрика пути al). Кодовое расстояние между 10 и путем ab, которому соответствует 11, равно d=1 (метрика пути ab). Сохраняем оба пути.

Приняли следующие два бита 01. Метрики путей:

alm – 2; ali – 2; abf – 1; abc - 3. Сохраняем пути (выжившие пути): alm, ali, abf.

Приняли еще два бита 00. Метрики путей:

almn – 2; almj – 4; alig – 3; alid – 3; abfn – 3; abfj – 1.

Выбираем наивероятнейший путь abfj, метрика которого наименьшая – 1. Т.е. считаем, что передавалась комбинация 110100 – ошибка исправлена.

13. Перемежение

Перемежение - эффективный способ борьбы с пакетами ошибок.

В реальных каналах связи принимаемый сигнал флуктуирует по амплитуде.

Когда уровень сигнала падает практически до нуля, принимаемые кодовые комбинации содержат очень много ошибок, т.е. мы принимаем «пакет ошибок». Использование мощных кодов – неэффективно. На передаче осуществляется перемежение, т.е. сначала передаем первый символ первой комбинации a_{11}, потом первый символ второй комбинации a_{21} и т.д. первый символ r –ой комбинации a_{r1}. Далее передаются вторые символы и т.д.

Пусть пакет ошибок поразил группу символов a_{11}, a_{21}…. a_{r1}.

На приеме осуществляется операция «деперемежение». Символы возвращаются на свои места в кодовых комбинациях. Следовательно, в каждой комбинации будет по одной ошибке в первом символе. Одиночная ошибка исправляется достаточно простым кодом.

14.Итеративные и каскадные коды

Теория кодирования позволяет построить мощные коды большой длины n, позволяющие исправлять ошибки большой кратности. Основная трудность заключается в нахождении кодов, имеющих приемлемую сложность технической реализации и разумное время задержки решения.

Хорошие результаты могут быть получены с помощью итеративных кодов.

Простейший пример показан в таблице.

	Информационные символы 1-й ступени				Проверочные символы 1-й ступени		
Информа-	0	0	0	1	1	0	1
ционные	0	0	1	0	1	1	1
символы	0	1	1	0	1	0	0
2-ой ступени	1	1	1	1	1	1	1
Проверочные	1	0	0	0	1	1	0
символы	0	1	0	1	1	1	0
2-ой ступени	1	0	1	1	1	0	0

Первые четыре строки в таблице – это кодовые комбинации кода первой ступени (7,4). Первые четыре символа - информационные, последние три – проверочные.

Первый столбец – это кодовая комбинация кода (7,4) второй ступени. Верхние четыре символа – информационные символы, нижние три символа – проверочные символы кода второй ступени.

Полученный итеративный код (n_1n_2 , k_1k_2), т.е. в данном случае это код (49,16).

Каскадный код содержит внешний код: k бит рассматриваются как укрупненный символ кода с основанием $m=2^k$. Информационная часть внешнего кода содержит k_2 таких m-ичных символов. К ним добавляется ($n_2 - k_2$) символов m-ичного кода.

Внутренний код формируется из k бит внешнего кода, к которым добавляются (n-k) проверочных бит. Полученный каскадный код (nn_2, kk_2) позволяет исправлять пачки ошибок при достаточно простой реализации.

ЗАДАНИЕ НА КУРСОВУЮ РАБОТУ (ЧАСТЬ 2.1. « Разработка кодека»)

Курсовой проект посвящен разработке основных блоков цифровой системы связи и состоит из двух взаимосвязанных частей:

- часть 2.1 «Разработка кодека»;
- часть 2.2 «Разработка модема».

В разделе 2.1 необходимо разработать структурные схемы и определить основные параметры кодека.

1. Нарисовать структурную схему цифровой системы связи и указать назначение основных блоков (см. гл. 1 данных МУ, гл.1 в [1]).

2. Записать свои фамилию, имя, отчество и выбрать первые 10 букв. Каждая буква - это импульс-отсчет некоторого процесса. Амплитуда отсчета равна порядковому номеру буквы (см. задачу №1 в данных МУ). Закодировать эти отсчеты двоичным кодом (m = 2, n = 5), нарисовать эти отсчеты и соответствующий им сигнал ИКМ.

3. Рассчитать дисперсию шума квантования, если U_{max} в вольтах равна количеству букв в Вашей фамилии (см. формулу (3) в данных МУ).

4. Определить вероятность дибитов 00, 01, 10, 11 в двоичной последовательности сигнала ИКМ, полученной в пункте 2. Рассчитать энтропию источника с полученной вероятностью дибитов. Закодировать дибиты двоичным кодом с префиксными свойствами и определить его энтропию, избыточность и среднюю длину кодовой комбинации (см. задачу №2 в данных МУ).

5. Осуществить помехоустойчивое кодирование двоичных информационных комбинаций, используя для этого код, указанный в таблице.

Последняя цифра номера студенческого билета	Способ кодирования
0, 1, 2, 3	Циклический код (см. гл. 11)
4, 5, 6	Сверточный код (см. гл. 12)
7, 8, 9	Блочный двоичный код (см. гл. 10)

Необходимо описать в соответствии с указанной главой алгоритм кодирования и декодирования; записать разрешенные комбинации на выходе кодера для всех возможных информационных комбинаций на входе; зарисовать структурные схемы кодера и декодера.

СПИСОК ЛИТЕРАТУРЫ

1. Андреев А.Н., Краснов Р.П., Чепелев М.Ю. Теория электрической связи. Курс лекций. Учебное пособие для вузов. Москва: Горячая линия - Телеком, 2014.

2. Белов В.М., Новиков С.Н., Солонская О.И. Теория информации. Учебное пособие. Москва: Горячая линия - Телеком, 2012

3. Данилов В.А., Львов В.Л., Бородин А.В. Методическое пособие для проведения практических занятий по дисциплине «Общая теория связи» Ростов-на-Дону: СКФ МТУСИ, 2016.