Тема 3. Кодирование в системах связи

1. Основные понятия теории кодирования

Кодом называется совокупность знаков, а также система правил, позволяющая представлять информацию в виде набора таких знаков.

Кодовым словом называют любой ряд допустимых знаков. Например, двоичное число 1100 можно считать двоичным 4-разрядным кодовым словом

Кодирование – это преобразование сообщений в последовательность элементарных символов. Пример – код Морзе.

По цели различают три вида кодирования:

1. ***Криптостойкое кодирование*** **-** применяется для защиты передаваемой информации от посторонних, т. е. обеспечивает секретность передаваемой информации по каналам связи.

2. ***Экономное кодирование (сжатие, компрессия)*** **-** применяется для уменьшения избыточности информации. Используется в каналах без помех.

Сжатие информации представляет собой процесс преобразования исходного сообщения из одной кодовой системы в другую, в результате которого уменьшается размер сообщения.

Алгоритмы, предназначенные для сжатия информации, можно разделить на две группы:

реализующие *сжатие без потерь* (обратимое сжатие);

реализующие *сжатие с потерями* (необратимое сжатие).

Обратимое сжатие позволяет абсолютно точно восстановить данные после декодирования и может применяться для сжатия любой информации.

Оно всегда приводит к снижению объема выходного потока информации без изменения его информативности, т.е. без потери информационной структуры. Более того, из выходного потока, при помощи восстанавливающего или декомпрессирующего алгоритма, можно получить входной поток.

Процесс восстановления называется **декомпрессией** или распаковкой и только после процесса распаковки данные пригодны для обработки в соответствии с их внутренним форматом.

Сжатие без потерь применяется для текстов, исполняемых файлов, высококачественного звука и графики.

Необратимое сжатие имеет гораздо более высокую степень сжатия, но допускает некоторые отклонения декодированных данных от исходных.

На практике существует широкий круг задач, в которых соблюдение требования точного восстановления исходной информации после декомпрессии не является обязательным.

Это, в частности, относится к сжатию мультимедийной информации: звука, фото- или видеоизображений.

Так, например, широко применяются форматы мультимедийной информации *JPEG* и *MPEG*, в которых используется необратимое сжатие.

Необратимое сжатие обычно не используется совместно с криптографическим шифрованием, так как основным требованием к криптосистеме является идентичность расшифрованных данных исходным.

3. *Помехоустойчивое кодирование*

Необходимость помехоустойчивого кодирования:

Если в канале есть помехи, то при приеме кодовых символов могут произойти ошибки, тогда кодовые комбинации (полученные при эффективном кодировании) будут декодированы неправильно!

Задача: повышение верности передачи
Один из путей ее решения – *помехоустойчивое (канальное)* кодирование.

Помехоустойчивыми (корректирующими) кодами называются коды, обеспечивающие автоматическое обнаружение и/или исправление ошибок в кодовых комбинациях.

Такая возможность обеспечивается целенаправленным *введением избыточности* в передаваемые сообщения.

Наиболее простой способ:

например, вместо слова стол можно передавать слово *сссттоооллл*

Помехоустойчивое кодирование - предназначено для обнаружения и по возможности исправления ошибок, возникших вследствие действия помех при передаче сигналов по каналам связи.

Помехоустойчивое кодирование предполагает введение избыточности в кодированный дискретный сигнал. При этом канал связи должен позволять пропускать кодированный сигнал с избыточностью.

Общая идея помехоустойчивого кодирования состоит в том, что из всех возможных кодовых слов считаются допустимыми не все, а лишь некоторые из них. Например, в коде с контролем по четности считаются допустимыми лишь слова с четным числом единиц.

Ошибка превращает *допустимое слово* в *недопустимое* и поэтому обнаруживается.

Если канал связи не позволяет вводить избыточность, то сначала проводят экономное кодирование с последующим введением избыточности.

Общим для всех трех видов кодирования является то, что информация каким-либо образом меняет форму представления, но не смысл.

Отличия разных видов кодирования связаны с целью проводимых преобразований.

Таким образом, криптографическое шифрование, помехоустойчивое кодирование и сжатие отчасти дополняют друг друга и их комплексное использование помогает эффективно использовать каналы связи для надежной защиты передаваемой информации.

2-я Теорема Шеннона (Основная теорема кодирования для каналов с помехами (шумами)

Если производительность источника $H'(\mathrm{A})$ ***меньше пропускной способности канала*** C ***то*** *существует* ***по крайней мере одна процедура кодирования/декодирования, при которой вероятность ошибочного декодирования и ненадежность*** $H(\mathrm{A}|\mathrm{B})$ ***могут быть сколь угодно малы. Если*** $H'(\mathrm{A}) > C$ ***то такой процедуры не существует.***

В вышеприведённом примере ясно, что при фиксированном уровне помех для стремления вероятности ошибки к нулю количество повторений должно стремиться к бесконечности.

Скорость передачи информации при этом стремится к нулю.

2. Параметры кодов

Число *m* используемых для кодирования элементарных символов, определяет *основание кода*.

Коды с основанием ***m* =2** называются *двоичными*.

В двоичном коде элементарные символы «1» и «0».

Коды с основанием ***m* >2** называются *многопозиционными*.

Длина кодовой комбинации – ***n*** (*разрядность кода*).

Для кодов, имеющих одинаковую длину кодовой комбинации существует понятие **кодового расстояния**.

Расстоянием по Хэммингу между двумя кодовыми словами называется число разрядов, в которых они различны.

При этом в качестве **минимального кодового расстояния** выбирается наименьшее из всех расстояний по Хэммингу для любых пар различных кодовых слов, образующих код.

Чему равно кодовое расстояние для кодовых слов

001100100110101

100100100111000

*_*_____**_*

Чтобы получить кодовое расстояние между двумя комбинациями двоичного кода, достаточно подсчитать число единиц в сумме этих комбинаций по модулю 2.

$$1 \oplus 1 = 0 \qquad 1 \oplus 0 = 1 \qquad 0 \oplus 1 = 1 \qquad 0 \oplus 0 = 0$$

Кодовая комбинация составляется из элементарных символов, например, при n=5 может быть представлена в виде 10101.

Число возможных комбинаций M при данных m и n:

$$\boldsymbol{M = m^n}.$$

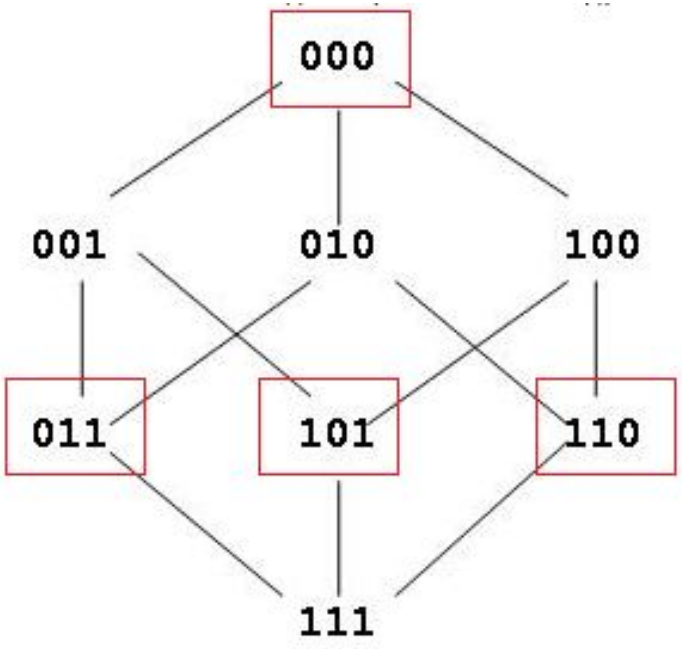
Например, для кода Бодо (телеграфный код) при $\boldsymbol{m=2}$, $\boldsymbol{n=5}$

$$\boldsymbol{M = 2^5 = 32}.$$

Пример: рассмотрим обычный (не помехоустойчивый) трехразрядный двоичный код.

Для него возможное количество кодовых комбинаций $M=2^3=8$ и все комбинации являются допустимыми

Минимальное кодовое расстояние между словами, как и для всякого обычного, не помехоустойчивого кода, равно единице



Если применить для определения разрешенных кодовых комбинаций правило контроля четности, то допустимыми будут только 4 комбинации.

Минимальное кодовое расстояние с контролем по четности равно 2.

Платой за помехоустойчивость является необходимость увеличения длины слов по сравнению с обычным кодом.

В данном примере только два разряда являются информационными. Это они образуют четыре разных информационных слова.

Третий разряд является *контрольным* и служит только для увеличения расстояния между допустимыми словами.

В передаче информации контрольный разряд не участвует, так как является *линейно зависимым* от информационных.

Код с контролем по четности позволяет обнаружить нечетное количество ошибок в блоках при передаче данных, причем не указывает, сколько таких ошибок.

Четное количество ошибок (двукратные, четырехкратные и т.д.) он не обнаруживает.

Таким образом, для того чтобы код мог обнаруживать и устранять ошибки, необходимо отказаться от его безызбыточности.

Для этого и разделяют всё множество возможных комбинаций двоичных символов на два подмножества: допустимых кодовых слов и недопустимых.

Разбиение осуществляется таким образом, чтобы увеличить минимальное кодовое расстояние между допустимыми словами.

В этом случае ошибка превращает допустимое кодовое слово в недопустимое, что позволяет её обнаружить.

Введение дополнительных контрольных разрядов увеличивает затраты на хранение или передачу кодированной информации. При этом фактический объем полезной информации остается неизменным.

В этом случае можно говорить об **избыточности** помехоустойчивого кода, которую формально можно определить как отношение числа контрольных разрядов k к общему числу n разрядов кодового слова

$$Q=k/n$$

Избыточность является важной характеристикой кода, причем чрезмерное увеличение избыточности нежелательно.

Важной задачей теории информации является синтез кодов с минимальной избыточностью, обеспечивающих заданную обнаруживающую и корректирующую способность.

1.Чтобы код обладал свойствами обнаруживать одиночные ошибки, необходимо ввести избыточность, которая обеспечивала бы минимальное расстояние между любыми двумя разрешенными комбинациями не менее двух.

В общем случае при необходимости обнаруживать ошибки кратности t_0 минимальное хэммингово расстояние должно быть, по крайней мере, на единицу больше , т.е.

$$d_{min} \geq t_0 + 1$$

2. Для исправления одиночной ошибки каждой разрешенной кодовой комбинации необходимо сопоставить подмножество запрещенных кодовых комбинаций. Чтобы эти подмножества не пересекались, хэммингово расстояние должно быть не менее трех.

В общем случае исправляемые ошибки кратности t_u связаны с кодовым расстоянием соотношением

$$d_{min}=2t_u+1$$

Важно! Каждый конкретный корректирующий код не гарантирует исправления любой комбинации ошибок. Коды предназначены для исправления комбинаций ошибок, наиболее вероятных для заданного канала связи.

Для ориентировочного определения необходимой избыточности кода при заданном кодовом расстоянии ***d*** можно воспользоваться верхней граничной оценкой для ***r = n - k***, называемой оценкой Хэмминга

$$r = n - k \log_2 \left(1 + \sum_{t=1}^{t_и} C_n^t \right),$$

Если, например, n=7, t_u=1,

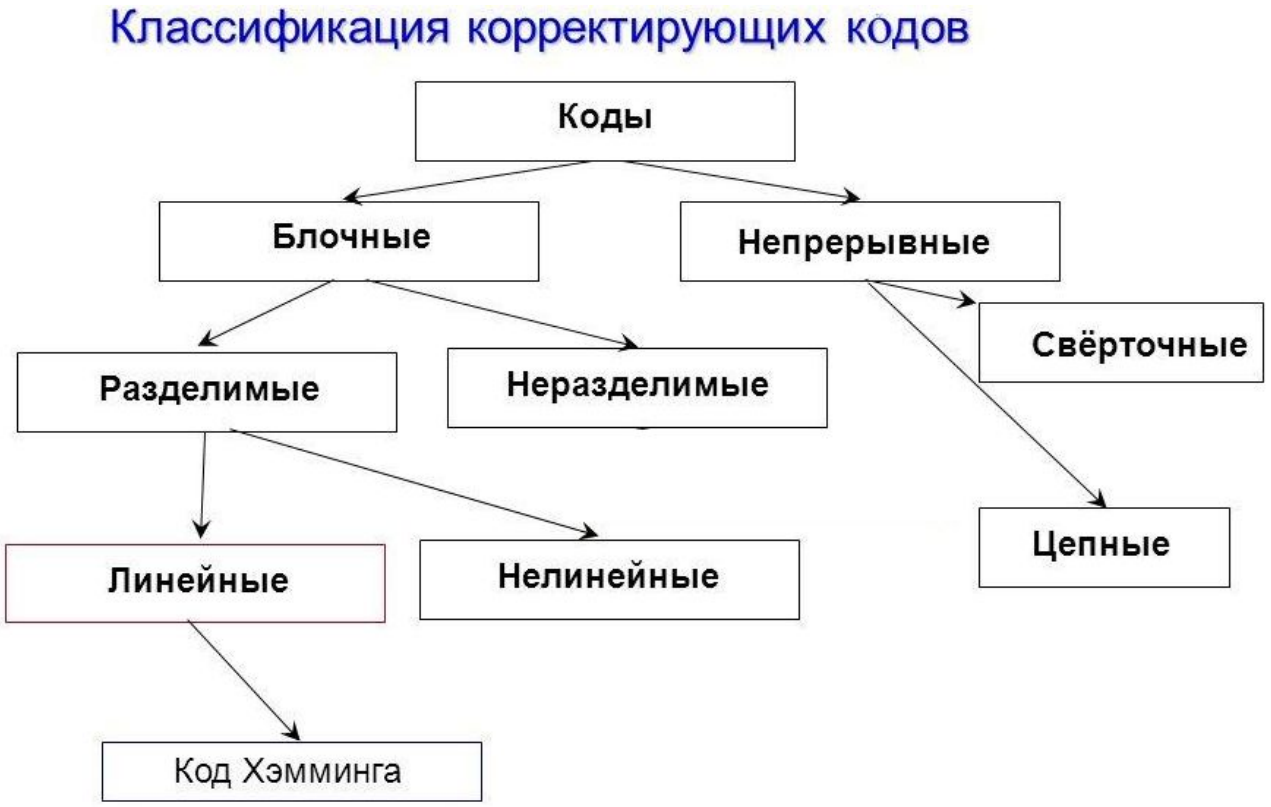
$$d_{\min} = 3,\ n - k \log_2 (1+7) = 3 .$$

3. Классификация кодов

Различают коды избыточные (корректирующие) и безызбыточные.

Весом называется число единиц, содержащихся в кодовых комбинациях.

Если число единиц во всех комбинациях кода будет постоянным, то такой код будет кодом с постоянным весом.



Равномерные и неравномерные коды.

У равномерных кодов ***n=const,*** например, пятиразрядный код Бодо.

У неравномерных кодов *n* является переменной величиной, например, в коде Морзе кодовые комбинации имеют различную длину.

Особенность неравномерного кода состоит в том, что более длинные кодовые комбинации присваиваются знакам передаваемой информации, которые встречаются редко и наоборот, более часто встречающимся знакам присваиваются наиболее короткие кодовые комбинации.

Это обеспечивает повышение пропускной способности канала связи.

Код Морзе

Код Морзе - это способ знакового кодирования, представление букв алфавита, цифр, знаков препинания и других символов последовательностью сигналов: длинных ("тире") и коротких ("точек").

За единицу времени принимается длительность одной точки. Длительность тире равна трём точкам. Пауза между элементами одного знака — одна точка, между знаками в слове — 3 точки, между словами — 7 точек.

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

—•• •• ••• •——• • — |———• • •—•

Код Морзе статистически согласован с английским языком.

Так буква *E, которая* в английском языке имеет наибольшую вероятность $p(E)$=0,11, имеет самую короткую комбинацию, состоящую из одного элемента (точки).

С алфавитом русского языка код Морзе менее согласован.

Например, буква *О*, которая имеет наибольшую вероятность в русском языке $p(O)$=0,1, передаётся комбинацией из 5 элементов – – – (три тире и два разделительных интервала).

В 2004 году Международный союз электросвязи (МСЭ) ввёл в азбуку Морзе новый код для символа @ (· — — · — ·), для удобства передачи адресов электронной почты.

Неравномерность является основной особенностью кода Морзе, которая позволяет учитывать статистику сообщения.

Однако код Морзе менее экономичный, чем равномерный код Бодо. Если принять одинаковыми длительности кодовых элементов в кодах Бодо и Морзе, то средняя длина комбинации в коде Морзе вдвое больше средней длины комбинации равномерного пятизначного кода Бодо.

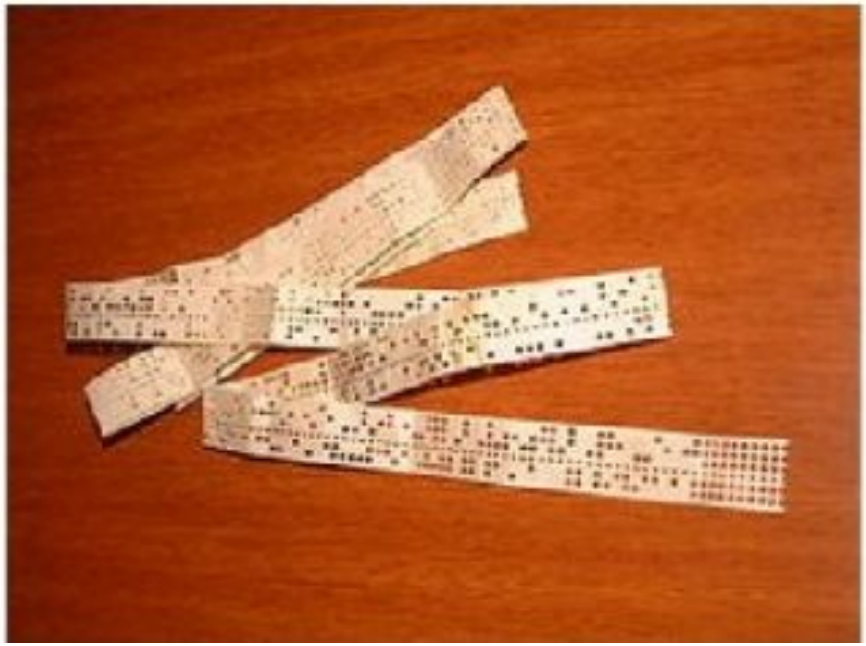
Это объясняется тем, что в коде Морзе обязательным является наличие разделительных знаков, как между кодовыми комбинациями, так и между точками и тире внутри кодовой комбинации.

Неравномерность кода Морзе не позволяет осуществить слитную передачу кодовых комбинаций, а следовательно, и осуществить кодом Морзе автоматизированную систему связи.

Достоинством кода Морзе является его простота, облегчающая приём на слух (каждая кодовая комбинация имеет свою «мелодию»).

Телеграф и радиотелеграф первоначально использовали азбуку Морзе; позже стали применяться код Бодо и ASCII, которые более удобны для автоматизации

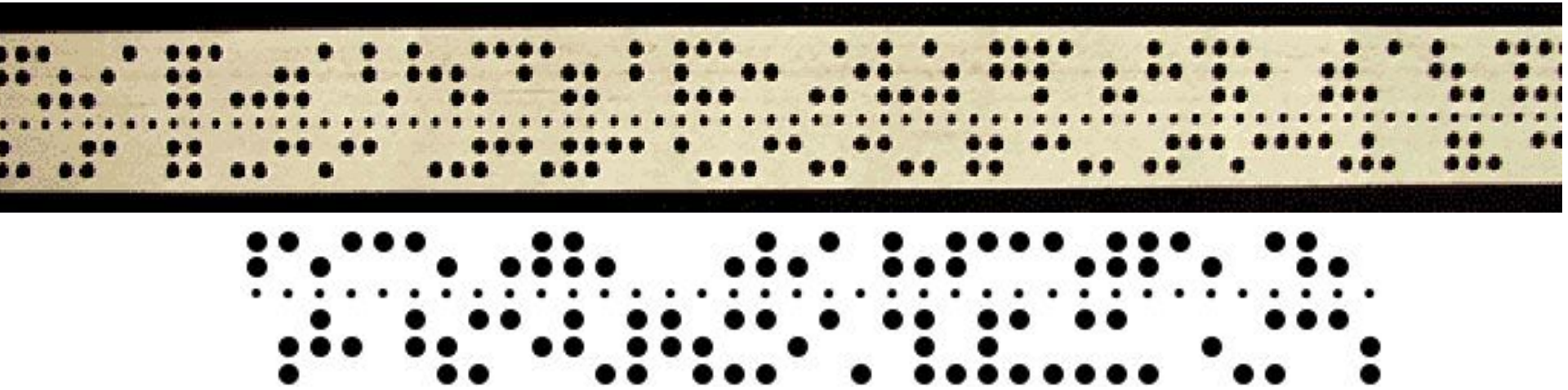
Код Бодо



Управляющие символы	
o. ...	пробел, перейти к таблице букв
.o ...	пробел, перейти к таблице цифр
oo ...	удалить последний знак

таблица букв							таблица цифр				
.. o..	A		oo o..	K			.. o..	1		o. o..	.
.. oo.	É		oo oo.	L			.. .o.	2		o. .o.	9/
.. .o.	E		oo .o.	M			.. ..o	3		o. ..o	7/
.. .oo	I		oo .oo	N			.. o.o	4		o. o.o	2/
.. ooo	O		oo ooo	P			.. ooo	5		o. ooo	'
.. o.o	U		oo o.o	Q			.. oo.	1/		o. oo.	:
.. ..o	Y		oo ..o	R			.. .oo	3/		o. .oo	?
.o ..o	B		o. ..o	S			.o o..	6		oo o..	(
.o o.o	C		o. o.o	T			.o .o.	7		oo .o.	)
.o ooo	D		o. ooo	V			.o ..o	8		oo ..o	-
.o .oo	F		o. .oo	W			.o o.o	9		oo o.o	/
.o .o.	G		o. .o.	X			.o ooo	0		oo ooo	+
.o oo.	H		o. oo.	Z			.o oo.	4/		oo oo.	=
.o o..	J		o. o..	—			.o .oo	5/		oo .oo	£

Русский шрифт		**Е**	Перевод строки ≡	Пробел	Возврат каретки ∨	**Т**	**А**	**И**	**Н**	**О**	**С**	**Р**	**Х**	**Д**	**Л**	**З**	**У**	**Ц**	**М**	**Ф**	**Й**	**Г**	**П**	**Ы**	**Б**	**В**		**К**	**Ж**	**Ь**	**Я**						
Цифры		**3**				**5**	**-**	**8**	**,**	**9**	**'**	**Ч**	**Щ**	кто там?	**)**	**+**	**7**	**:**	**.**	**Э**	**Ю** (зв)	**Ш**	**0**	**5**	**?**	**2**	Цифры	**(**	**=**	**/**	**1**	Буквы лат.	Буквы рус.				
Латинский шрифт		**Е**				**Т**	**А**	**I**	**N**	**О**	**S**	**R**	**H**	**D**	**L**	**Z**	**U**	**C**	**M**	**F**	**J**	**G**	**P**	**Y**	**B**	**W**		**K**	**V**	**X**	**Q**						
Ведущие отверстия	1	●					●				●			●		●	●			●	●			●	●	●	●	●		●	●	●				
	2		●				●	●				●			●		●	●			●	●	●			●	●	●	●		●	●				
		○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○			
	3			●				●	●		●		●				●	●	●	●			●	●				●	●	●	●	●				
	4				●				●	●		●		●				●	●	●	●	●			●		●	●	●	●		●				
	5					●				●			●		●	●			●			●	●	●	●	●	●		●	●	●	●	●				



Равномерные корректирующие коды подразделяются на блочные и непрерывные.

В блочных двоичных кодах последовательность элементарных сообщений источника разбивается на отрезки, каждый из которых независимо преобразуется в определённую последовательность (блок) кодовых символов равной длины *M*.

Непрерывные коды представляют непрерывную последовательность кодовых символов, её разделение на отдельные кодовые комбинации не производится.

Среди блочных кодов наиболее распространены линейные систематические коды, особенностью которых является то, что они строятся путём добавления к комбинации из «*k*» информационных символов «*r*» проверочных символов, число которых равно $r = n-k$.

Проверочные символы получаются путём некоторых линейных комбинаций (суммирование по модулю два) над информационными символами. Среди них наиболее изучены циклические коды.

Среди помехоустойчивых кодов выделяют разделимые коды и неразделимые коды.

В разделимых кодах разряды могут быть принципиально разделены на проверочные и информационные. При этом место проверочных и информационных разрядов в кодовой комбинации четко определено.

В неразделимых кодах (например, коды с постоянным весом) деление на информационные и проверочные разряды отсутствуют.

Разделимые коды подразделяются на систематические и несистематические.

Систематическими называют такие коды, у которых сумма по модулю два двух разрешенных комбинаций кода дает также разрешенную комбинацию этого же кода.

Кроме того, в систематических кодах проверочные символы могут образовываться путем различных линейных комбинаций информационных символов.

Для систематического кода применяется обозначение ***(n,m)***–код,

где

n – число всех разрядов в кодовой комбинации,

m – число информационных разрядов.

Декодирование систематических кодов основано на проверке линейных соотношений между символами, стоящими на определенных проверочных позициях.

К систематическим относятся циклические коды .

Линейные блочные коды

Блочный равномерный код – множество кодовых слов (комбинаций) одинаковой длины n.

Элементы кодовых слов выбираются из некоторого алфавита (канальных символов) объемом q.

Если q =2, код называется двоичным.

Поскольку все кодовые слова имеют одинаковую длину, удобно считать их векторами, принадлежащими линейному пространству размерности n.

00110100101101

.....

01001110100100

Для линейных ко́дов справедливо утверждение:
линейная комбинация кодовых слов является кодовым словом.

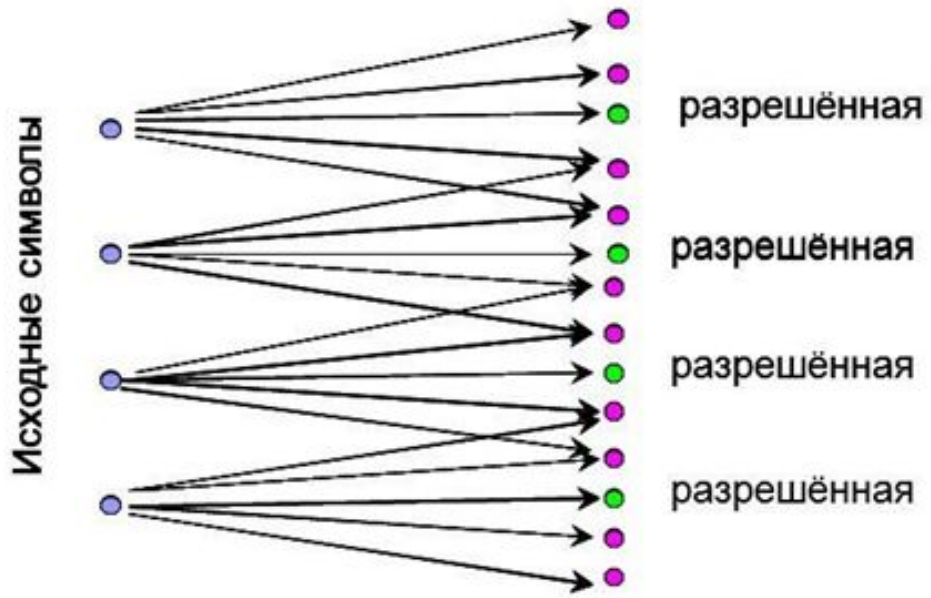
Всего 2^n *n*-мерных векторов с двоичными компонентами (кодовых комбинаций или слов).

Из них только $M = 2^k$ комбинаций являются *разрешёнными* и составляют код, который называется (n,k)-*кодом* (отношение $k/n = R$ называется относительной *скоростью кода*).

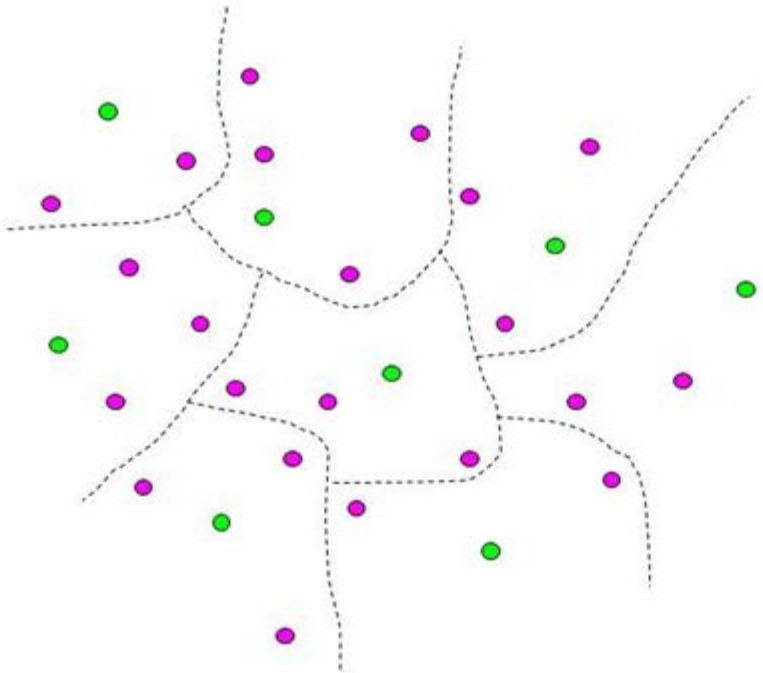
Остальные комбинации являются *запрещёнными*, образуются из разрешённых в канале под воздействием помех.

Разрешённые комбинации – векторы линейного пространства.

Чем больше расстояние между разрешёнными комбинациями, тем меньше вероятность преобразования их друг в друга под действием помех, тем выше способность кода к *обнаружению и исправлению* ошибок.



Работа декодера сводится к разбиению всего пространства на области, каждая из которых содержит одну разрешённую комбинацию



Для кодирования и декодирования линейных блочных кодов применяются действия, описываемые операциями над векторами в линейном пространстве над конечным полем целых чисел

Сложение и умножение в конечном поле понимаются как сложение и умножение по модулю *q*

Простейшее из таких полей, называемых *полями Галуа́* – поле по модулю 2, обозначаемое *GF*(2)

ЭВАРИСТ ГАЛУА
Évariste Galois
(1811-1832)
выдающийся французский математик, основатель современной алгебры.

Линейные коды являются разделимыми, то есть k символов – информационные, остальные $(n-k)$ – проверочные.

Информационные символы зависят от передаваемого сообщения и могут быть какими угодно.

Проверочные символы однозначно определяются информационными т.к. *формируются из информационных символов* кодером, работающим по определённому алгоритму)

Отсюда следует, что каждая кодовая комбинация, будучи вектором n-мерного пространства, принадлежит его k-мерному подпространству

Обозначим

$\mathbf{X} = (x_1, \ldots, x_k)$ *информационный вектор*

$\mathbf{C} = (c_1, \ldots, c_n)$ кодовый вектор

Кодирование описывается линейным преобразованием (оператором), отображающим векторы, соответствующие подпространству S_k^0, в векторы из S_n:

$$\mathbf{C} = \mathbf{XG}$$

Матрица кодирования

$$\mathbf{G} = \begin{pmatrix} g_{11} & g_{12} & \cdot & \cdot & \cdot & g_{1n} \\ g_{21} & g_{22} & \cdot & \cdot & \cdot & g_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ g_{k1} & g_{k2} & \cdot & \cdot & \cdot & g_{kn} \end{pmatrix}$$

Подробнее:

$$(x_1,\ldots,x_k)\begin{pmatrix} g_{11} & g_{12} & \cdot & \cdot & \cdot & g_{1n} \\ g_{21} & g_{22} & \cdot & \cdot & \cdot & g_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ g_{k1} & g_{k2} & \cdot & \cdot & \cdot & g_{kn} \end{pmatrix} =$$

$$= \left(\sum_{i=1}^{k} x_i g_{i1}, \ldots \sum_{i=1}^{k} x_i g_{in} \right)$$

Или $c_j = x_1 g_{1j} + x_2 g_{2j} + \ldots + x_k g_{kj},\ j = 1,\ldots,n$

Путём линейных операций над строками и перестановки столбцов любую такую матрицу можно привести к *систематическому* виду

$$\mathbf{G} = (\mathbf{I}_k \ \vdots \ \mathbf{P}) = \begin{pmatrix} 1 & 0 & 0 & . & . & 0 & p_{11} & p_{12} & \cdot & p_{1(n-k)} \\ 0 & 1 & 0 & . & . & 0 & p_{21} & p_{22} & \cdot & p_{2(n-k)} \\ . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & 1 & p_{k1} & p_{k2} & \cdot & p_{k(n-k)} \end{pmatrix}.$$

k

k первых символов повторяют символы информационного вектора, а остальные (*n*-*k*) символов формируются из информационных и являются проверочными (паритетными).

В этом случае код называют *систематическим*.

Пример. Систематический код (7,4) порождается матрицей

$$\mathbf{G} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$

Кодовые слова имеют структуру

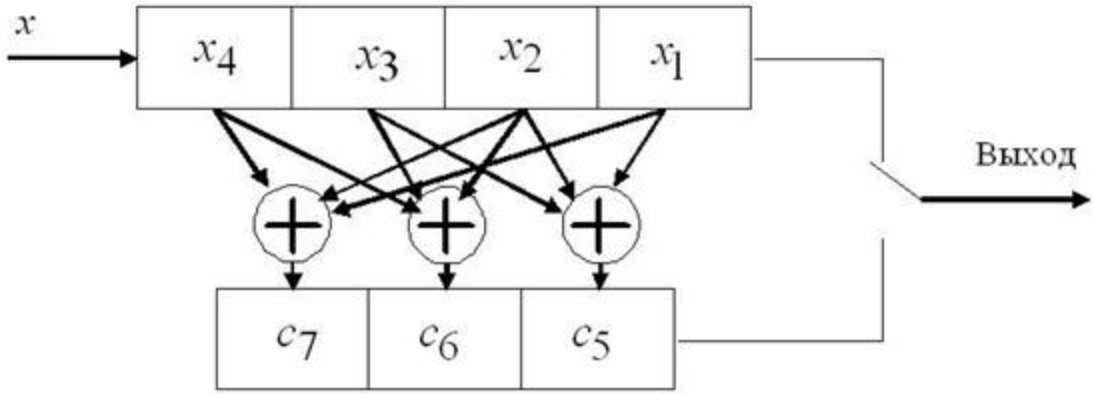
$$\mathbf{C} = (x_1, x_2, x_3, x_4, c_5, c_6, c_7)$$

где $c_5 = x_1 + x_2 + x_3$

$$c_6 = x_2 + x_3 + x_4$$

$$c_7 = x_1 + x_2 + x_4$$

Структурная схема кодера



Применение любого кода предполагает реализацию не только кодирования, но и декодирования. Декодирование систематического линейного блочного кода могло бы заключаться в простом отбрасывании проверочных символов, но это не обеспечивало бы обнаружения и исправления ошибок

4. Циклические коды

Циклические коды относятся к блоковым кодам.

Последовательность кодовых комбинаций в циклическом коде разбивается на отдельные блоки, состоящие из информационных и проверочных разрядов и в пределах этих блоков производится исправление ошибок.

Блок состоит из:

m – информационных элементов (разрядов);

k – проверочных элементов (разрядов);

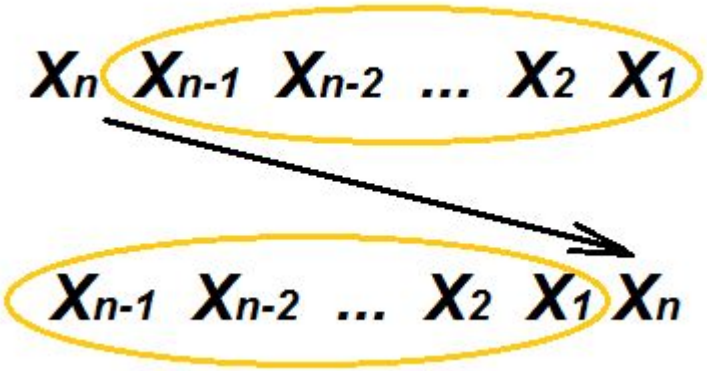
n=m+k - общее число разрядов циклического кода.

В кодовой комбинации циклического кода, как правило, в начале идут информационные элементы, а потом проверочные.

$$x_m\ x_{m-1} \ldots x_2\ x_1\ \ r_k \ldots r_2\ r_1$$

Циклические коды имеют высокую эффективность обнаружения ошибок и сравнительно просты в реализации кодирующих и декодирующих устройств.

Основным свойством циклических кодов является то, что циклический сдвиг одной кодовой комбинации приводит к другой тоже разрешённой кодовой комбинации..



Разрешенная кодовая комбинация образуется также при сложении по модулю 2 двух других разрешенных кодовых комбинаций

При описании свойств циклических кодов пользуются представлением кодовых комбинаций в виде многочленов (полиномов) от некоторой переменной «***X***» с коэффициентами «**1**» или «**0**».

Цифры двоичного кода можно рассматривать как коэффициенты многочлена переменной ***x***. Например, для кодовой комбинации **1011011** (*n=7*) полином

$$A(x)=1*x^6+0*x^5+1*x^4+1*x^3+0*x^2+1*x^1+1*x^0= x^6+x^4+x^3+x+1.$$

При таком представлении кодов математические операции с полученными многочленами производятся в соответствии с законами обычной алгебры, за исключением того, что сложение осуществляется по модулю 2:

$$x^a + 0 = x^a, \quad 0 + 0 = 0, \quad x^a + x^a = 0$$

Если число разрядов кодовой комбинации ***n***, то многочлен имеет степень (***n–1***).

Принцип обнаружения ошибок при помощи циклического кода заключается в том, что в качестве разрешенных кодовых комбинаций принимаются такие, которые делятся без остатка на некоторый заранее выбранный исходный (образующий) полином ***G(x)***.

Если принятая комбинация искажена, то это условие на приемной стороне не будет выполнено, в результате чего формируется сигнал, указывающий на наличие ошибки.

Построение комбинаций циклического кода возможно путем умножения исходной комбинации $A(x)$ на образующий полином $G(x)$ с приведением подобных членов по модулю 2:

если старшая степень произведения не превышает $(n-1)$, то полученный полином будет представлять кодовую комбинацию циклического кода;

если старшая степень произведения больше или равна n, то полином произведения делится на заранее выбранный полином степени n и результатом умножения считается полученный остаток от деления.

Таким образом, все полиномы, отображающие комбинации циклического кода, будут иметь степень ниже n.

В процессе кодирования сообщения:

1) Многочлен ***A(x)***, отображающий двоичный код исходного передаваемого сообщения, умножается на x^k.

При этом длина кодовой комбинации увеличивается на ***k*** разрядов, которые предназначены для проверочных разрядов. Эти разряды заполняются «0».

2) Произведение $A(x)x^k$ делят на образующий полином ***G(x)***, и остаток от этого деления ***R(x)*** суммируют с произведением $A(x)x^k$.

Полученная кодовая комбинация, описываемая кодовым многочленом $F(x)= A(x)x^k+R(x)$, делится без остатка на образующий полином ***G(x).***

При таком методе построения коэффициенты при высших степенях ***x*** являются обозначениями информационных разрядов, а коэффициенты при степенях порядка ***k-1*** и ниже – проверочными.

Пример.

Дано: $n=7$, $m=4$, $k=3$ и $G(x)=x^3+x^2+1$.

Требуется закодировать сообщение **1 0 1** 1.

Сообщению 1**011** соответствует многочлен $A(x)=x^3+x+1$.

Выполним умножение $A(x)$ на x^k, где k=3. Получим $A(x)x^3 = x^6+x^4+x^3$.

Разделим $A(x)x^3 = x^6+x^4+x^3$ на образующий полином $G(x)$:

$$
\begin{array}{l|l}
x^6+x^4+x^3 & x^3+x^2+1 \\
\hline
x^6+x^5+x^3 & x^3+x^2 \\
\hline
x^5+x^4 & \\
x^5+x^4+x^2 & \\
\hline
R(x)=x^2 &
\end{array}
$$

В итоге этой операции получим остаток $R(x)=x^2$.

Суммируя произведение $A(x)x^3$ с полученным остатком, получим кодовый многочлен

$$F(x)=A(x)x^3+R(x)=x^6+x^4+x^3+x^2.$$

В двоичном коде этому многочлену соответствует кодовая комбинация **1 0 1 1 1 0 0**, в которой проверочные разряды занимают три последние позиции.

При применении циклического кода в качестве кода с исправлением ошибок места искаженных разрядов определяются путем анализа остатка, получившегося после деления принятой кодовой комбинации на образующий полином.

Алгоритм построения кода:

По заданному объему информационного кода однозначно определяется число информационных разрядов ***k***. Далее необходимо найти наименьшее ***n***, обеспечивающее обнаружение или исправление ошибок заданной кратности.

Для циклического кода эта проблема сводится к нахождению образующего полинома ***G(x).***

Образующий полином следует выбирать как можно более коротким: наибольшая степень его должна быть равна числу контрольных разрядов, а число ненулевых членов должно быть не меньше минимального кодового расстояния.

Достоинство циклического кода в том, что он используется для обнаружения как одиночных ошибок, так и групповых.

5. Код Хемминга

Ричард Хемминг разработал код, который обеспечивает обнаружение и исправление одиночных ошибок при минимально возможном числе дополнительных проверочных бит.

Для каждого числа проверочных символов используется специальная маркировка вида ***(n, m),***

где

n — количество символов в сообщении,

m — количество информационных символов в сообщении.

Например, существуют коды **(7, 4), (15, 11), (31, 26).**

Каждый проверочный символ в коде Хэмминга представляет сумму по модулю 2 некоторой последовательности данных.

Рассмотрим пример, когда количество информационных бит ***m*** в блоке равно 4. Это код **(7,4),** количество проверочных символов равно **3**.

Классически проверочные символы располагаются на позициях, равных степеням двойки в порядке возрастания

первый проверочный бит на позиции $\mathbf{2^0 = 1}$;

второй проверочный бит на позиции $\mathbf{2^1 = 2}$;

третий проверочный бит на позиции $\mathbf{2^2 = 4}$.

Теперь рассчитаем значения проверочных символы по алгоритму:

$k1 = m1 \oplus m2 \oplus m4$

$k2 = m1 \oplus m3 \oplus m4$

$k3 = m2 \oplus m3 \oplus m4$

Итак, в закодированном сообщении у нас получится следующее:

k1 k2 m1 k3 m2 m3 m4

Проверочные символы можно разместить и в конце передаваемого блока данных, но тогда алгоритм для их расчета будет другим.

Построение корректирующего кода Хэмминга производится исходя из требуемого объема информационных сообщений и статистических данных о наиболее вероятных векторах ошибок в используемом канале связи.

Вектором ошибки будем называть кодовую комбинацию, имеющую единицы в разрядах, подвергшихся искажению, и нули во всех остальных разрядах.

Любую искаженную кодовую комбинацию можно рассматривать как сумму по модулю 2 разрешенной кодовой комбинации и вектора ошибки.

В коде Хэмминга необходимое число проверочных разрядов ***k*** определяется из известного соотношения

$$2^{n-k} - 1 \geq n$$

Значения символов в проверочных разрядах устанавливаются в результате суммирования по модулю 2 значений символов в определенных информационных разрядах.

В принципе, место расположения контрольных разрядов в коде Хэмминга безразлично, но определенные удобства создает такое размещение, при котором контрольные разряды входили бы в возможно меньшее число сумм, получаемых при проверке кода.

Это будет, если контрольные разряды размещать в позициях, номера которых равны целой степени числа 2, т.е. в разрядах:

1, 2, 4, 8, 16, 32 и т.д.

Вычисление контрольных бит.

Значение каждого контрольного бита зависит от значений информационных бит, но не от всех, а только от тех, которые этот контрольный бит контролирует.

Важно: контрольный бит с номером ***N*** контролирует все последующие ***N*** бит через каждые ***N*** бит, начиная с позиции ***N***.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	1	1	1	0	1
1	X		X		X		X		X		X		X		X		X		X		X
2		X	X			X	X			X	X			X	X			X	X		
4				X	X	X	X					X	X	X	X					X	X
8								X	X	X	X	X	X	X	X						
16																X	X	X	X	X	X

Знаком «X» обозначены те биты, которые контролирует контрольный бит. Так, бит номер 12 контролируется битами с номерами 4 и 8.

Значение контрольного бита 0 – если сумма контролируемых им бит четная, 1 – если нечетная.

Проверка на приемной стороне принятой кодовой комбинации осуществляется следующим образом: создаются контрольные суммы S1, S2, S3, S4 и т.д.

$$S_1 = a_1 \oplus a_3 \oplus a_5 \oplus a_7 \ldots$$

$$S_2 = a_2 \oplus a_3 \oplus a_6 \oplus a_7 \ldots$$

$$S_3 = a_4 \oplus a_5 \oplus a_6 \oplus a_7 \ldots$$

$$S_4 = a_8 \oplus a_9 \oplus a_{10} \oplus a_{11} \ldots$$

Правило построения контрольных сумм:

S1 – суммируются все нечетные разряды

S2 – суммируются начиная со 2-го разряда по два разряда подряд через 2 разряда

S3 –суммируются начиная с 4-го разряда по 4 разряда через 4 разряда

S4 – суммируются начиная с 8-го разряда по 8 разрядов через 8 разрядов.

Если все суммы равны нулю, то в принятой кодовой комбинации нет ошибки. В случае, когда одна или несколько контрольных сумм равны единице, то эти суммы располагаются слева направо в порядке возрастания индексов и полученная запись в двоичном коде указывает на номер разряда, где произошла ошибка.

Сущность обнаружения позиции ошибки, а значит и ее исправления кодом Хэмминга состоит в том, что производятся многократные проверки на четность различных вариантов сумм разрядов принятого кода, в результате которых получается двоичный код номера искаженного разряда.

Для кода (7,4), исправляющего одиночные ошибки, соотношения для нахождения ошибки имеют вид

$$a_1 = a_3 \oplus a_5 \oplus a_7$$

$$a_2 = a_3 \oplus a_6 \oplus a_7$$

$$a_4 = a_5 \oplus a_6 \oplus a_7$$

Пример. Построить код Хэмминга с исправлением одиночной ошибки при 11 информационных разрядах, т.е. ***m=11***.

Определим число контрольных разрядов (***k=n – m = n-11***)

$$2^{n-m} - 1 = n$$

$$2^{n-11} - 1 \geq 11+k$$

$$n=15$$

Число контрольных разрядов ***k= 4***.

Пусть необходимо закодировать сообщение:

10110100111

Представим это информационное сообщение в виде кода Хэмминга, установив контрольные разряды на 1, 2, 4, 8 позициях.

15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
1	0	1	1	0	1	0	a_8	0	1	1	a_4	1	a_2	a_1

Определим значение контрольных разрядов, запишем их в соответствующих местах, и в окончательном виде код Хэмминга без ошибок будет выглядеть так:

15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
1	0	1	1	0	1	0	**0**	0	1	1	**1**	1	**0**	**0**

Если при передаче данного сообщения произошло искажение в каком-либо информационном разряде, например, в третьем, то принятая комбинация

Найдем контрольные суммы:

$$S1 = a_3 \oplus a_5 \oplus a_7 \oplus a_9 \oplus a_{11} \oplus a_{13} \oplus a_{15}$$
$$S2 = a_3 \oplus a_6 \oplus a_7 \oplus a_{10} \oplus a_{11} \oplus a_{14} \oplus a_{15}$$
$$S3 = a_5 \oplus a_6 \oplus a_7 \oplus a_{12} \oplus a_{13} \oplus a_{14} \oplus a_{15}$$
$$S4 = a_9 \oplus a_{10} \oplus a_{11} \oplus a_{12} \oplus a_{13} \oplus a_{14} \oplus a_{15}$$

S1=1
S2=1
S3=0
S4=0

По полученному коду

$S_4S_3S_2S_1$
0 0 1 1

видно, что искажение произошло в третьем разряде.

6. Коды Рида-Соломона

Коды Рида-Соломона были предложены в 1960 сотрудниками Линкольнской лаборатории МТИ Ирвином Ридом и Густавом Соломоном.

Коды Рида-Соломона базируются на блочном принципе коррекции ошибок и используются в огромном числе приложений в сфере цифровых телекоммуникаций и при построении запоминающих устройств.

Коды Рида-Соломона применяются для исправления ошибок во многих системах, включая:

- устройства памяти (включая магнитные ленты, CD, DVD, штриховые коды, и т.д.);
- беспроводные или мобильные коммуникации (включая сотовые телефоны, микроволновые каналы и т.д.);
- спутниковые коммуникации;
- цифровое телевидение / DVB (digital video broadcast);
- скоростные модемы, такие как ADSL, xDSL и т.д.

Коды Рида-Соломона это *недвоичные циклические* коды, символы которых представляют собой s-битовые последовательности.

Код Рида-Соломона обозначается как ***RS(n,k)*** ***s***-битных символов.

Это означает, что кодировщик воспринимает ***k*** информационных символов по ***s*** бит каждый и добавляет символы четности для формирования ***n*** символьного кодового слова.

Имеется ***n-k*** символов четности по ***s*** бит каждый.

Декодер Рида-Соломона может корректировать до ***t*** символов, которые содержат ошибки в кодовом слове, где ***2t = n-k***. Т.е. для исправления ***t*** ошибок код должен иметь ***2t*** символов.

В отличие от кодов Хемминга, коды Рида-Соломона могут исправлять любое разумное количество ошибок при вполне приемлемом уровне избыточности.

Это достигается тем, что если в кодах Хемминга контрольные биты контролируют лишь те информационные биты, что находятся только по одну сторону от них, то в кодах же Рида-Соломона контрольные биты распространяют свое влияние на все информационные биты. Поэтому с увеличением количества контрольных бит увеличивается и количество распознаваемых/устраняемых ошибок.

Именно благодаря последнему обстоятельству вызвана огромная популярность корректирующих кодов Рида-Соломона.