Коды и антиподы

Выполнил: Мельник Пётр

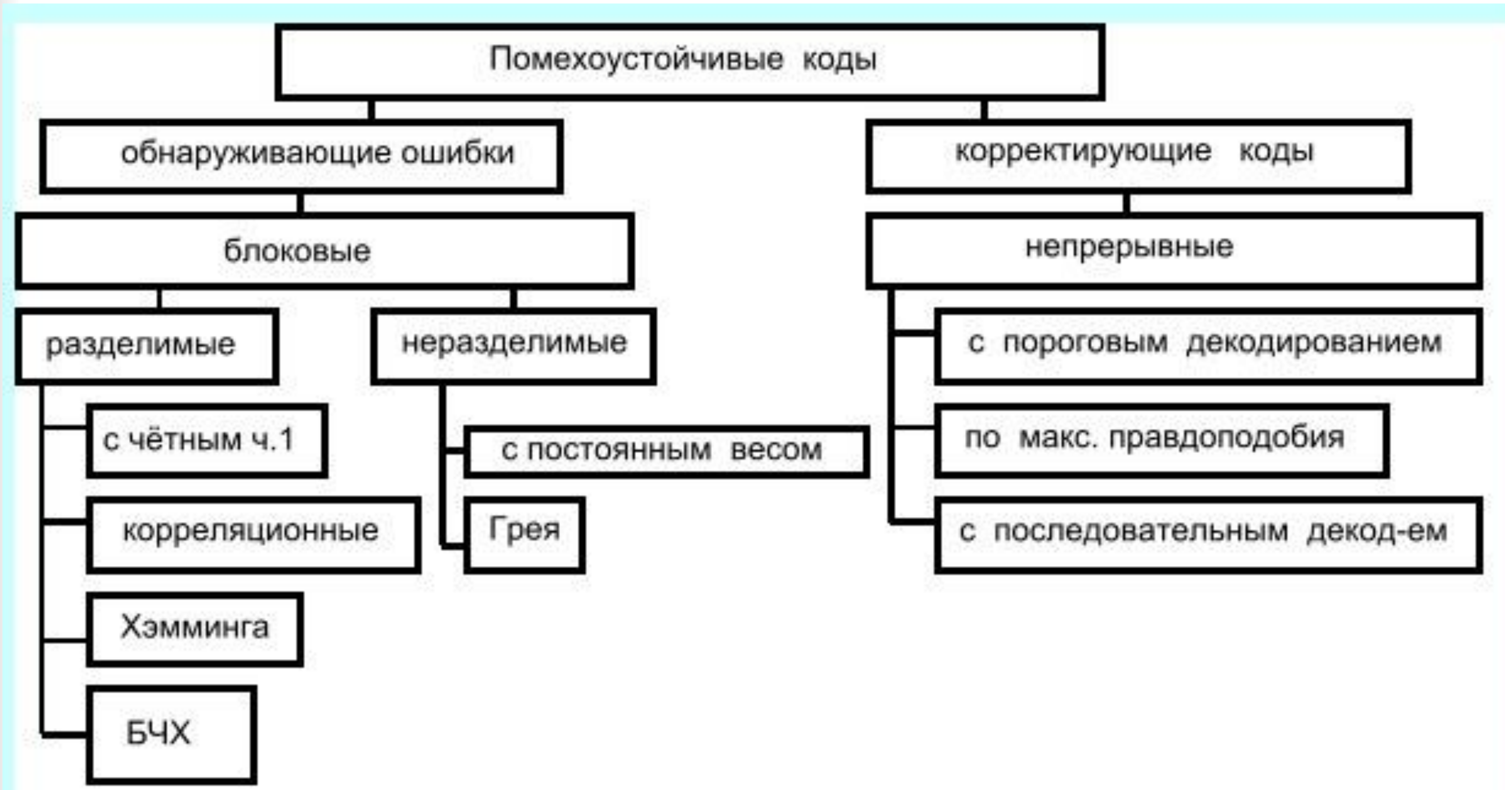
студент 5 курса математического факультета

Помехоустойчивое кодирование

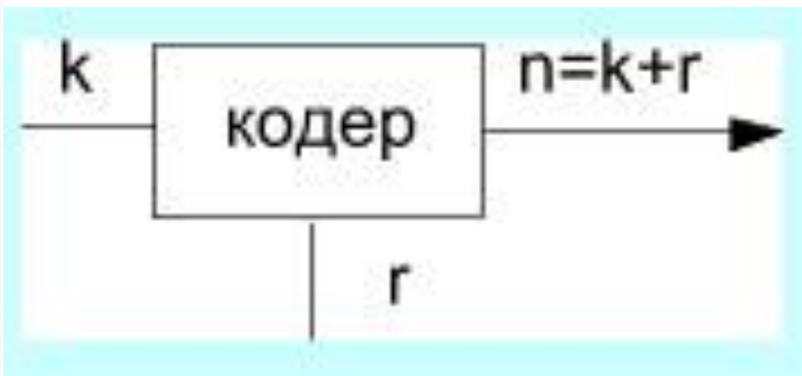
Под помехой понимается любое воздействие, накладывающееся на полезный сигнал и затрудняющее его приём. Ниже приведена классификация помех и их источников.

Внешние источники помех вызывают в основном импульсные помехи, а внутренние – флуктуационные

Приведём классификацию помехоустойчивых кодов:



Построение помехоустойчивых кодов в основном связано с добавлением к сходной комбинации (k-символов) контрольных (r-символов). Закодированная комбинация будет составлять n-символов. Эти коды часто называют (n,k) - коды



k – число символов в исходной комбинации

r – число контрольных символов

Рис. 1.Получение (n,k) кодов.

Коды с обнаружением ошибок

1. Код с проверкой на чётность.

Такой код образуется путём добавления к передаваемой комбинации, состоящей из k информационных символов, одного контрольного символа (0 или 1), так, чтобы общее число единиц в передаваемой комбинации было чётным.

Пример 1. Построим коды для проверки на чётность, где k – исходные комбинации, r контрольные символы.

k	r	n
11011	0	110110
11100	1	111001

Определим, каковы обнаруживающие свойства этого кода. Вероятность P_{00} обнаружения ошибок будет равна

$$P_{00} = C_n^1 p^1 (1-p)^{n-1} + C_n^3 p^3 (1-p)^{n-3} + C_n^5 p^5 (1-p)^{n-5} \mathrm{K}$$

Так как вероятность ошибок p<<1 является весьма малой величиной, то можно ограничится

$$P_{00} = C_n^1 P^1 (1-p)^{n-1}$$

Вероятность появления всевозможных ошибок, как обнаруживаемых так и не обнаруживаемых, равна $P_\Sigma = 1 - Q$, где $Q = (1-p)^n$ - вероятность отсутствия искажений в кодовой комбинации. Тогда $P_\Sigma = 1 - (1-p)^n$.

При передаче большого количества кодовых комбинаций N_k, число кодовых комбинаций, в которых ошибки обнаруживаются, равно:

$$N_{00} = N_k \cdot P_{00} = N_k C_n^1 p (1-p)^{n-1}$$

Общее количество комбинаций с обнаруживаемыми и не обнаруживаемыми ошибками равно

$$N_\Sigma = N_k \cdot P_\Sigma = N_k(1-(1-p)^n)$$

Тогда коэффициент $K_{обн}$ обнаружения для кода с чётной защитой будет равен

$$K_{обн} = \frac{N_{00}}{N_\Sigma} = \frac{P_{00}}{P_\Sigma}$$

Например, для кода с k=5 и вероятностью ошибки $p = 10^{-2}$ коэффициент обнаружения составит $K_{обн} = 0.9$.
То есть 90% ошибок обнаруживаем, при этом избыточность будет составлять $L = 1 - \frac{5}{6} = 0.17$
или 17%.

2. Код с постоянным весом.

Этот код содержит постоянное число единиц и нулей. Число кодовых комбинаций составит

$$N = C_n^k = \frac{n!}{k!(n-k)!}$$

Пример 2. Коды с двумя единицами из пяти и тремя единицами из семи.

$N = C_5^2 = 10$	$N = C_7^3 = 35$
11000	0000111
10010	1001001
00101	1010100

Этот код позволяет обнаруживать любые одиночные ошибки и часть многократных ошибок. Не обнаруживаются этим кодом только ошибки смещения, когда одновременно одна единица переходит в ноль и один ноль переходит в единицу, два ноля и две единицы меняются на обратные символы и т.д.

Рассмотрим код с тремя единицами из семи. Для этого кода возможны смещения трёх типов.

$$\left.\begin{matrix}0 \to 1\\ 1 \to 0\end{matrix}\right\} см1 \qquad \left.\begin{matrix}00 \to 11\\ 11 \to 00\end{matrix}\right\} см2 \qquad \left.\begin{matrix}000 \to 111\\ 111 \to 000\end{matrix}\right\} см3$$

Вероятность появления не обнаруживаемых ошибок смещения $P_{H0} = P_{a1} + P_{a2} + P_{a3}$, где

$$P_{a1} = C_3^1 * C_4^1 * p^2 * (1-p)^5$$
$$P_{a2} = C_3^2 * C_4^2 * p^4 * (1-p)^3$$
$$P_{a3} = C_3^3 * C_4^3 * p^6 * (p-1)^1$$

При p<<1 $P_{a3} \ll P_{a2} \ll P_{a1}$ тогда $P_{H0} = 12p^2(1-p)^5$

Вероятность появления всевозможных ошибок как обнаруживаемых, так и не обнаруживаемых будет составлять

$$P_{\Sigma} = 1 - Q = 1 - (1-p)^7$$

Вероятность обнаруживаемых ошибок $P_{00} = P_{\Sigma} - P_{НО}$. Тогда коэффициент обнаружения будет равен

$$K_{обн} = \frac{P_{00}}{P_{\Sigma}} = \frac{1-(1-p)^{n} - 12 * p^{2} * (1-p)^{5}}{1-(1-p)^{n}}$$

Например, код C_7^3 при $p = 10^{-2}$ коэффициент обнаружения составит $K_{обн} = 0.985$ избыточность L=27%.

3. Корреляционный код (Код с удвоением).

Элементы данного кода заменяются двумя символами, единица «1» преобразуется в 10, а ноль «0» в 01.

Вместо комбинации 1010011 передается 10011001011010. Ошибка обнаруживается в том случае, если в парных элементах будут одинаковые символы 00 или 11 (вместо 01 и 10).

Например, при k=5, n=10 и вероятности ошибки $p = 10^{-2}$, $K_{обн} = 0.995$. Но при этом избыточность будет составлять 50%

Back 3

4. Инверсный код.

К исходно комбинации добавляется такая же комбинация по длине. В линию посылается удвоенное число символов. Если в исходной комбинации чётное число единиц, то добавляемая комбинация повторяет исходную комбинацию, если нечётное, то добавляемая комбинация является инверсной по отношению к исходной.

k	r	n
11011	11011	1101111011
11100	00011	1110000011

Приём инверсного кода осуществляется в два этапа. На первом этапе суммируются единицы в первой основной группе символов. Если число единиц чётное, то контрольные символы принимаются без изменения, если нечётное, то контрольные символы инвертируются. На втором этапе контрольные символы суммируются с информационными символами по модулю два.

Нулевая сумма говорит об отсутствии ошибок. При ненулевой сумме, принятая комбинация бракуется. Покажем суммирование для принятых комбинаций без ошибок (1,3) и с ошибками (2,4).

1 11011	2 11111	3 11100	4 11000
11011	00100	11100	11100
00000	11011	00000	00100

Обнаруживающие способности данного кода достаточно велики. Данный код обнаруживает практически любые ошибки, кроме редких ошибок смещения, которые одновременно происходят как среди информационных символов, так и среди соответствующих контрольных. Например, k=5, n=10 и $p = 10^{-2}$. Коэффициент обнаружения будет составлять $K_{обн} = 1 - 10^{-5}$.

5. Код Грея.

Код Грея используется для преобразования угла поворота тела вращения в код. Принцип работы можно представить по рисунку 3.

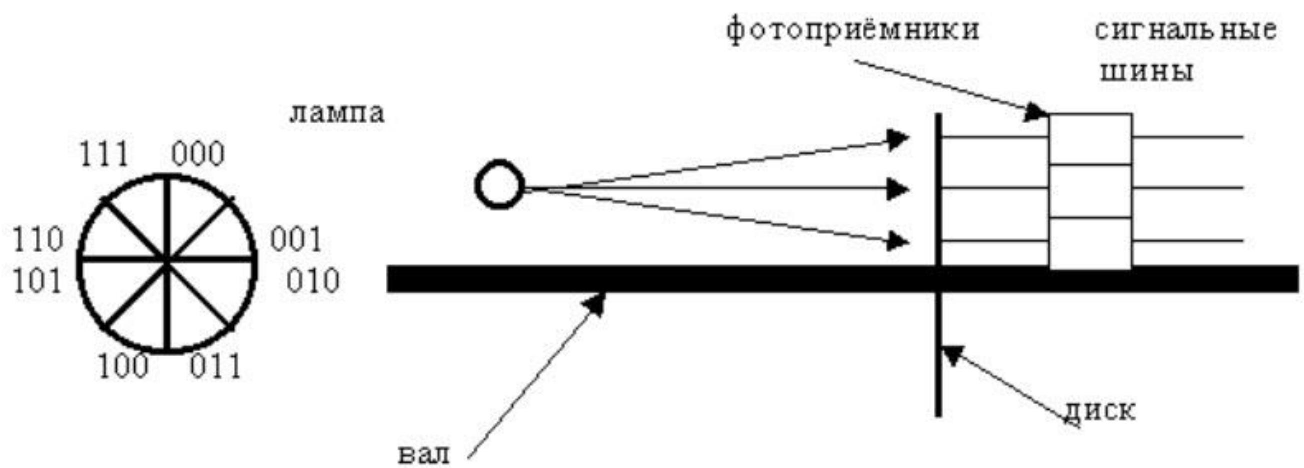


Рисунок 3

На пластинке, которая вращается на валу, сделаны отверстия, через которые может проходить свет . Причём, диск разбит на сектора, в которых и сделаны эти отверстия. При вращении, свет проходит через них, что приводит к срабатыванию фотоприёмников. При снятии информации в виде двоичных кодов может произойти существенная ошибка.

Например, возьмём две соседние цифры 7 и 8. Двоичные коды этих цифр отличаются во всех разрядах.

7	0111	->	1111
8	1000	->	0000

Если ошибка произойдёт в старшем разряде, то это приведёт к максимальной ошибке, на 360 градусов. А код Грея, это такой код в котором все соседние комбинации отличаются только одни символом, поэтому при переходе от изображения одного числа к изображению соседнего происходит изменение только на единицу младшего разряда. Ошибка будет минимальной.

Код Грея записывается следующим образом.

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

Разряды в коде Грея не имеют постоянного веса. Вес k-разряда определяется следующим образом

$$\pm\sum_{i=0}^{k} 2^{i} = \pm(2^{k+1} - 1)$$

При этом все нечётные единицы, считая слева направо, имеют положительный вес, а все чётные единицы отрицательный.

Например: $1110 = 1\cdot(2^{3+1} - 1) - 1\cdot(2^{2+1} - 1) + 1\cdot(2^{1+1} - 1) = 15 - 7 + 3 = 11_{10}$

Непостоянство весов разрядов затрудняет выполнение арифметических операций в коде Грея, поэтому необходимо уметь делать перевод кода Грея в обычный двоичный код и наоборот. Алгоритм перевода чисел можно представить следующим образом.

Пусть $A_n A_{n-1} \mathrm{K}\, A_0$ - двоичный код, $a_n a_{n-1} \mathrm{K}\, a_0$ - код Грея.

Тогда переход из двоичного кода в код Грея выполнится по следующему алгоритму:

$$a_i = \begin{cases} A_n, \text{ при } i = n \\ A_i, \text{ если } A_{i+1} = 0 \\ \overline{A}_i, \text{ если } A_{i+1} = 1 \end{cases}$$

Например: $1101_2 = 1011_Г$.

Обратный переход из кода Грея в двоичный код

$$A_i = \begin{cases} a_n, \text{ при } i = n \\ a_i, \text{ если } i - \text{символу предшествует четное число } 1 \\ \overline{a}_i, \text{ если } i - \text{символу предшествует нечетное число } 1 \end{cases}$$

Например: $1101_Г = 1011_2$.

Корректирующие коды

Корректирующими называются коды позволяющие обнаруживать и исправлять ошибки. Идею представления корректирующих кодов можно представить с помощью N-мерного куба.

Возьмём трёхмерный куб (рис 4), длина рёбер, в котором равна одной единице.

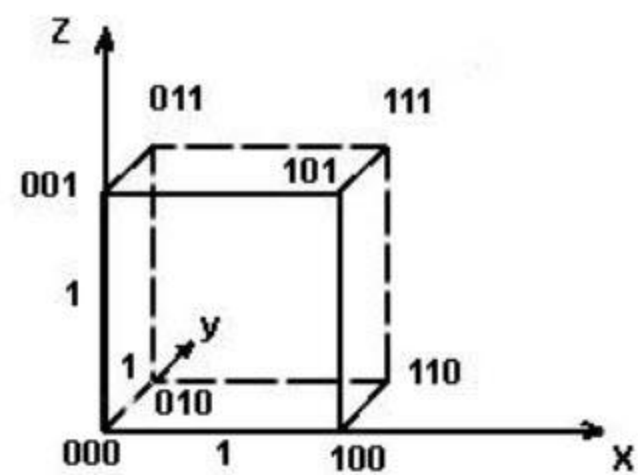


Рисунок 4. Представление двоичных кодов с помощью куба

Вершины такого куба отображают двоичные коды. Минимальное расстояние между вершинами определяется минимальным количеством рёбер, находящихся между вершинами. Это расстояние называется кодовым (или хэмминговым) и обозначается буквой d.

Иначе, кодовое расстояние – это то минимальное число элементов, в которых одна кодовая комбинация отличается от другой.

Для определения кодового расстояния достаточно сравнить две кодовые комбинации по модулю 2. Так, сложив две комбинации

$$\begin{array}{r} 10110101101 \\ \underline{11001010101} \\ 01111111000 \end{array}$$

Определим, что расстояние между ними d=7.

Для кода с N=3 восемь кодовых комбинаций размещаются на вершинах трёхмерного куба. Такой код имеет кодовое расстояние d=1, и для передачи используются все восемь кодовых комбинаций 000, 001, …., 111. Такой код является не помехоустойчивым, он не в состоянии обнаружить ошибку.

Если выберем комбинации с кодовым расстоянием d=2, например, 000, 110, 101, 011, то такой код позволит обнаруживать однократные ошибки. Назовём эти комбинации разрешенными, предназначенными для передачи информации. Все остальные 001, 010, 100,111 – запрещённые.

Любая одиночная ошибка приводит к тому, что разрешенная комбинация переходит в ближайшую, запрещённую комбинацию (рис.4). Получив запрещённую комбинацию, мы обнаружим ошибку.

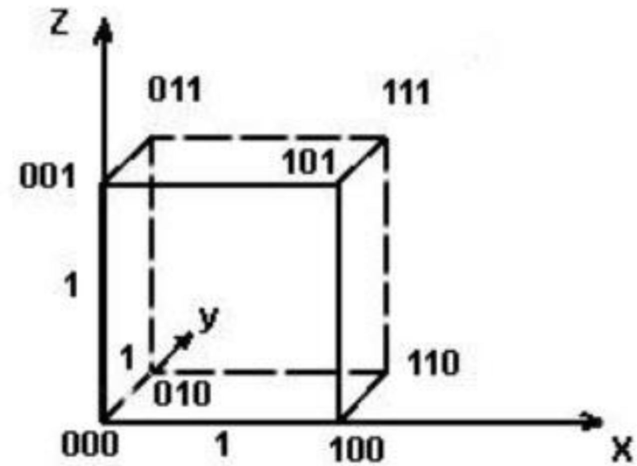
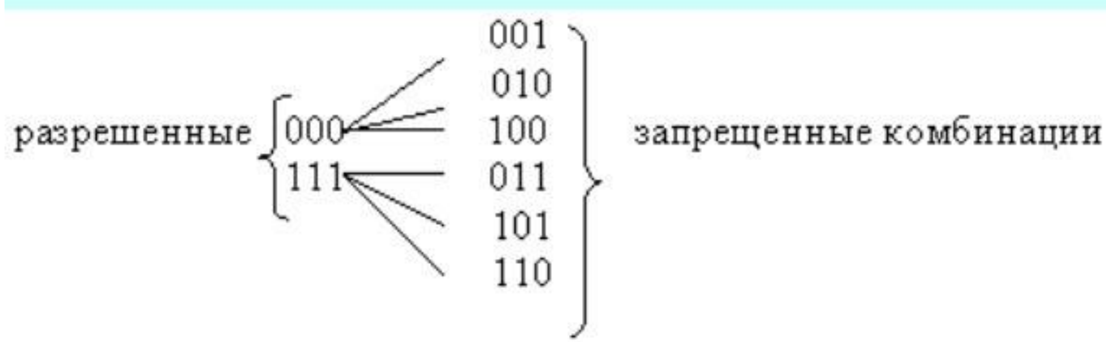


Рисунок 4. Представление двоичных кодов с помощью куба

Выберем далее вершины с кодовым расстоянием d=3.



Такой код может исправить одну одиночную ошибку или обнаружить две ошибки. Таким образом, увеличивая кодовое расстояние можно увеличить помехоустойчивость кода. В общем случае кодовое расстояние определяется по формуле d=t+l+l где t – число исправляемых ошибок, l- число обнаруживаемых ошибок. Обычно l>t.

Большинство корректирующих кодов являются линейными кодами.

Кроме того, корректирующие коды являются групповыми кодами.

Ниже приведены кодовые комбинации, являющиеся группой или нет.

1) 1101 1110 0111 1011 – не группа, так как нет нулевого элемента

2) 0000 1101 1110 0111 - не группа, так как не соблюдается условие замкнутости (1101 +1110 = 0011)

3) 000 001 010 011 100 101 110 111 - группа

Большинство корректирующих кодов образуются путём добавления к исходной k – комбинации r-контрольных символов. В итоге в линию передаются n=k+4 символов. При этом корректирующие коды называются (n,k) кодами. Как можно определить необходимое число контрольных символов?

Для построения кода способного обнаруживать и исправлять одиночную ошибку необходимое число контрольных разрядов будет составлять $n-k \geq \log(n+1)$

Это равносильно известной задачи о минимуме числа контрольных вопросов, на которые могут быть даны ответы вида «да» или «нет» , для однозначного определения одного из элементов конечного множества

Если необходимо исправить две ошибки, то число различных исходов будет составлять C_n^2 . Тогда

$n-k \geq \log\left(1+C_n^1+C_n^2\right)$, в этом случае обнаруживаются однократные и двукратные ошибки. В общем случае, число контрольных символов должно быть не меньше

$$n-k \geq \log\left(1+C_n^1+C_n^2+\ldots+C_n^t\right)=\log\sum_{i=0}^{t} C_n^i$$

$$n-k \geq \log\left(1+C_n^1+C_n^2+\ldots+C_n^t\right)=\log\sum_{i=0}^{t} C_n^i$$

Эта формула называется неравенством Хэмминга, или нижней границей Хэмминга для числа контрольных символов.

Упражнения

1. **Определить величину кодового расстояния между двумя комбинациями 1101101, 1001011.**

2. **Определить величину кодового расстояния d, обеспечивающего исправление s кратных ошибок при s=1,3,5.**

Упражнения

3. Перекодировать сообщение 100101010010 при помощи <u>корреляционного кода.</u>

Спасибо за внимание!!!

Линейные коды – это такие коды, у которых контрольные символы образуются путём линейной комбинации информационных символов.

Back 1

Групповые коды - это такие коды, которые имеют одну основную операцию. При этом, должно соблюдаться условие замкнутости (то есть, при сложении двух элементов группы получается элемент принадлежащий этой же группе). Число разрядов в группе не должно увеличиваться. Этому условию удовлетворяет операция поразрядного сложения по модулю 2. В группе, кроме того, должен быть нулевой элемент.

Back 2