

Декодирование

После исправления ошибки в S -м разряде (если она произошла) для декодирования достаточно оставить информационные разряды.

Пример 3.6.1. Пусть $t=4$. Определим наименьшее число l , удовлетворяющее

неравенству $2^l \geq 2l / (l - 1) : l=7, k=3$.

Информационными будут разряды с номерами 3, 5, 6, 7. Разряды с номерами 1, 2, 4, которые являются степенями 2, будут проверочными. Содержимое проверочных разрядов определяется следующими равенствами:

$$c_1 = c_3 \oplus c_5 \oplus c_7 \pmod{2},$$

$$c_2 = c_3 \oplus c_6 \oplus c_7 \pmod{2},$$

$$c_4 = c_5 \oplus c_6 \oplus c_7 \pmod{2}.$$

Код Хэмминга запишем в виде таблицы, где i -й столбец соответствует i -му разряду закодированного слова. Проверочные разряды помечены символом

*

1*	2*	3	4*	5	6	7
0	0	0	0	0	0	0
1	1	0	1	0	0	1
0	1	0	1	0	1	0
1	0	0	0	0	1	1
1	0	0	1	1	0	0
0	1	0	0	1	0	1

1	1	0	0	1	1	0
0	0	0	1	1	1	1
1	1	1	0	0	0	0
0	0	1	1	0	0	1
1	0	1	1	0	1	0
0	1	1	0	0	1	1
0	1	1	1	1	0	0
1	0	1	0	1	1	0
0	0	1	0	1	1	0
1	1	1	1	1	1	1

Пусть получено слово $\beta = 0110111$. Определим, произошла ли при передаче ошибка и в каком разряде. Вычислим S' .

$$S'_1 = \beta'_1 + \beta'_3 + \beta'_5 + \beta'_7 \pmod{2} = 0 + 1 + 1 + 1 = 1,$$

$$S'_2 = \beta'_2 + \beta'_3 + \beta'_6 + \beta'_7 \pmod{2} = 1 + 1 + 1 + 1,$$

$$S'_4 = \beta'_4 + \beta'_5 + \beta'_6 + \beta'_7 \pmod{2} = 0 + 1 + 1 + 1.$$

Получаем, что $S' = 101$ в двоичной системе счисления, или $S' = 5$. Значит, ошибка произошла в 5-м разряде. После исправления ошибки получим слово 0110011. Выделив в нем информационные разряды, получим исходное слово 1011.

Пусть исходное слово (кодовое слово без контрольных разрядов) - 0110101_2 .

Обозначим P_i - контрольный разряд № i ; а D_i - информационный разряд № i , где $i = 1,2,3$,

	1	2	3	4	5	6	7	8	9	10	11
№ разряда:	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011
Распределение контрольных и информационных разрядов	p_1	p_2	d_1	p_3	d_2	d_3	d_4	p_4	d_5	d_6	d_7
Информационное кодовое слово :			0		1	1	0		1	0	1
p_1	1		0		1		0		1		1
p_2		0	0			1	0			0	1
p_3				0	1	1	0				
p_4								0	1	0	1
Кодовое слово с контрольными разрядами:	1	0	0	0	1	1	0	0	1	0	1

Предположим теперь, для примера, что при передаче данного кодового слова 10001100101 произошла ошибка в 11-м разряде, так, что было принято новое кодовое слово 10001100100. Произведя в принятом коде проверку четности внутри контрольных групп, мы обнаружили бы, что количество единиц нечетно в 1-й, 2-й и 4-й контрольных группах, и четно в 3-й контрольной группе. Это указывает, во-первых, на наличие ошибки, во-вторых, означает, что номер ошибочно принятого разряда в двоичном представлении содержит единицы на первом, втором и четвертом местах и нуль - на третьем месте, т. к. ошибка только одна, и 3-я контрольная

	p_4	p_3	p_2	p_1
В двоичном представлении	1	0	1	1
В десятичном представлении	8		2	1 $\Sigma = 11$

Из таблицы следует, что ошибка произошла в 11-м разряде и её можно исправить. Построенный код не рассчитан на возможность одновременной ошибки в двух разрядах.

Пример 1. Построим по методу Хэмминга кодовое слово для сообщения $\tilde{\alpha} = (1011)$. Имеем $m = 4$, $n = \min\{l: 2^m \leq 2^l / (l + 1)\} = 7$, $k = n - m = 3$. Кодовое слово $\tilde{\beta}$ имеет вид $\beta_1\beta_2\beta_3\beta_4\beta_5\beta_6\beta_7 = p_0p_11p_2011$. Значения проверочных символов определяются из равенств

$$p_0 = \beta_3 \oplus \beta_5 \oplus \beta_7 = 1 \oplus 0 \oplus 1 = 0,$$

$$p_1 = \beta_3 \oplus \beta_6 \oplus \beta_7 = 1 \oplus 1 \oplus 1 = 1,$$

$$p_2 = \beta_5 \oplus \beta_6 \oplus \beta_7 = 0 \oplus 1 \oplus 1 = 0.$$

Таким образом, кодовым словом для $\tilde{\alpha}$ является вектор $\tilde{\beta} = (0110011)$.

Декодирование состоит в том, что по вектору $\tilde{\beta} = (\beta_1, \dots, \beta_n)$, полученному из некоторого кодового слова путем искажения не более чем в одном разряде, восстанавливается исходное сообщение $\tilde{\alpha} = (\alpha_1, \dots, \alpha_m)$. Декодирование осуществляется следующим образом.

Пусть $m = \lceil \log_2(2^n / (n + 1)) \rceil$, $k = n - m$. Вычислим по вектору $\tilde{\beta} = (\beta_1, \dots, \beta_n)$ k сумм вида

$$v_i = \beta_{2^i} \oplus \beta_{2^i+1} \oplus \dots, \quad i = 0, \dots, k - 1,$$

где в i -ю сумму включаются все координаты β_j ($2^i \leq j \leq n$), у которых двоичное разложение индекса j имеет коэффициент при 2^i , равный единице. В результате получаем двоичный вектор $\mathbf{v} = (v_0, \dots$

$\dots, v_{k-1})$ и число $\bar{\nu}(\mathbf{v}) = \sum_{0 \leq i < k} v_i \cdot 2^i$. Это число указывает номер

разряда, в котором произошла ошибка. Если $\bar{\nu}(\mathbf{v}) = 0$, то считаем, что ошибки при передаче не было. Если $\bar{\nu}(\mathbf{v}) > n$, то считаем, что передавалось слово, которое не является ни кодовым словом сообщения $\tilde{\alpha}$, ни кодовым словом, искаженным в одном разряде.

Пример 2. Декодировать вектор $\tilde{\beta} = (1001110)$. Имеем $n = 7$, $m = \lceil \log_2(2^7/8) \rceil = 4$, $k = n - m = 3$.

Вычислим вектор $\mathbf{v} = (v_0, v_1, v_2)$. Имеем

$$v_0 = \beta_1 \oplus \beta_3 \oplus \beta_5 \oplus \beta_7 = 1 \oplus 0 \oplus 1 \oplus 0 = 0,$$

$$v_1 = \beta_2 \oplus \beta_3 \oplus \beta_6 \oplus \beta_7 = 0 \oplus 0 \oplus 1 \oplus 0 = 1,$$

$$v_2 = \beta_4 \oplus \beta_5 \oplus \beta_6 \oplus \beta_7 = 1 \oplus 1 \oplus 1 \oplus 0 = 1.$$

Получаем, что $v(\mathbf{v}) = 1 \cdot 2^2 + 1 \cdot 2^1 + 0 \cdot 2^0 = 6$. Следовательно, ошибка произошла в шестом разряде. Неискаженный кодовый вектор имеет вид $\tilde{\beta}' = (1001100)$. Вычеркивая проверочные разряды с номерами 1, 2, 4, получаем исходное сообщение $\tilde{\alpha} = (0100)$.

Построение префиксного кода по набору длин элементарных кодов

Пусть задан набор натуральных чисел l_1, l_2, \dots, l_m , удовлетворяющих неравенству Мак-Миллана:

$$\sum_{i=1}^m 2^{-l_i} \leq 1$$

Алгоритм К.Шеннона построения префиксного кода по набору длин:

Будем полагать $l_1 \leq l_2 \leq \dots \leq l_m$. Построим последовательность чисел q_1, q_2, \dots, q_m , которая строится по следующим правилам:

$$q_1 = 0,$$

$$q_{i+1} = q_i + 2^{-l_i} \quad (i = 1, 2, \dots, m-1).$$

Очевидно, $0 \leq q_i < 1$ и q_i имеет единственное представление в виде двоичной

дроби с l_i знаками после запятой: $q_i = \sum_{j=1}^{l_i} c_j^{(i)} \cdot 2^{-j}$, где $c_j^{(i)} = 0$ или $c_j^{(i)} = 1$.

Рассмотрим код $V = (v_1, v_2, \dots, v_m)$, где $v_i = c_1^{(i)} c_2^{(i)} \dots c_{l_i}^{(i)}$.

Так как наборы длин упорядочены по неубыванию, при $h > i$ выполняются неравенства $l_h \geq l_i$ и $q_h \geq q_i + 2^{-l_i}$. Поэтому элементарный код v_h отличается от элементарного кода v_i в l_i первых разрядах. Следовательно, построенный код является префиксным.

Пример : Рассмотрим набор чисел $L=(2,3,3,3,4,4,4)$, удовлетворяющий неравенству Мак-Миллана

Построим последовательность чисел $q_1, q_2, q_3, q_4, q_5, q_6, q_7$, записывая их в двоичной системе счисления.

$$q_1 = 0,00;$$

$$q_2 = 0 + 2^{-2} = 0,010;$$

$$q_3 = 0,01 + 2^{-3} = 0,01 + 0,001 = 0,011;$$

$$q_4 = 0,011 + 2^{-3} = 0,011 + 0,001 = 0,100;$$

$$q_5 = 0,1 + 2^{-3} = 0,1 + 0,001 = 0,1010;$$

$$q_6 = 0,101 + 2^{-4} = 0,101 + 0,0001 = 0,1011;$$

$$q_7 = 0,1011 + 2^{-4} = 0,1011 + 0,0001 = 0,1100.$$

Далее строим схему f алфавитного кодирования, выбирая в качестве элементарного кода v_i последовательность из 0 и 1 длины l_i , образующую дробную часть числа q_i (при этом недостающие разряды в записи дробной части дополняем справа нулями):

$$f : \begin{cases} v_1 = 00 \\ v_2 = 010 \\ v_3 = 011 \\ v_4 = 100 \\ v_5 = 1010 \\ v_6 = 1011 \\ v_7 = 1100. \end{cases}$$