ПРАКТИЧЕСКОЕ ЗАНЯТИЕ

Тема «**Элементы теории кодирования. Коды Хемминга**»

Код Хемминга в РГР это $(4,7)$-код (общий случай разобран на лекции), т.е. он переводит (кодирует) исходные слова (сообщения) $a = a_1 a_2 a_3 a_4$ длины 4 в кодовые слова $b = b_1 b_2 b_3 b_4 b_5 b_6 b_7$ длины 7, кратко: $a \mapsto b$.

1. **Схема кодирования.** В кодовом слове b символы b_{2^i}, где $i = 0,1,2$, являются ***контрольными*** (т.е. добавочными, которые и позволят обнаруживать и исправлять возможные ошибки передачи сообщений), а остальные символы в естественном порядке – символы исходного сообщения, т.е.

$$a = a_1 a_2 a_3 a_4 \mapsto b = b_1 b_2 b_3 b_4 b_5 b_6 b_7 = b_1 b_2 a_1 b_4 a_2 a_3 a_4.$$

Составим систему уравнений для нахождения контрольных символов. Рассмотрим матрицу $M = M_{7\times 3}$ такую, что в i-й строке этой матрицы $(i = 1,\dots,7)$ находятся символы двоичного разложения числа i:

$$M = \begin{bmatrix} 0\ 0\ 1 \\ 0\ 1\ 0 \\ 0\ 1\ 1 \\ 1\ 0\ 0 \\ 1\ 0\ 1 \\ 1\ 1\ 0 \\ 1\ 1\ 1 \end{bmatrix}.$$

Запишем систему уравнений $bM = 0$:

$$\begin{cases} b_4 + b_5 + b_6 + b_7 = 0, \\ b_2 + b_3 + b_6 + b_7 = 0, \\ b_1 + b_3 + b_5 + b_7 = 0, \end{cases} \Leftrightarrow \begin{cases} b_4 = a_2 + a_3 + a_4, \\ b_2 = a_1 + a_3 + a_4, \\ b_1 = a_1 + a_2 + a_4, \end{cases}.$$

Пример 1. Найти кодовое слово $b = b_1 b_2 b_3 b_4 b_5 b_6 b_7$ для исходного сообщения $a = a_1 a_2 a_3 a_4 = 1011.$

Решение. $a = a_1 a_2 a_3 a_4 = 1011 \mapsto b = b_1 b_2 b_3 b_4 b_5 b_6 b_7 =$
$= b_1 b_2 a_1 b_4 a_2 a_3 a_4 = b_1 b_2 1 b_4 011,$

$$\begin{cases} b_4 = a_2 + a_3 + a_4 = 0+1+1 = 0, \\ b_2 = a_1 + a_3 + a_4 = 1+1+1 = 1, \\ b_1 = a_1 + a_2 + a_4 = 1+0+1 = 0, \end{cases}$$

откуда $b = b_1 b_2 1 b_4 011 = 0110011.$

2. **Схема декодирования.** Пусть принято слово $c = b + e,$ где e - ошибка. Тогда $bM = 0$, а следовательно,

$$cM = (b+e)M = bM + eM = eM.$$

Если $cM = 0,$ то считается, что ошибок не было. Это действительно так при $e = 0.$ Если вектор ошибок имеет только одну единицу в i-й позиции, то eM есть вектор, совпадающий с i-й строкой матрицы $M,$ являющейся двоичным разложением числа i. В этом случае следует изменить символ в i-й позиции слова $c = b + e$. В случае двух ошибок будет выполняться $cM \neq 0,$ что даст нам информацию о том, что принятое слово c содержит ошибки, однако их положение определить не можем.

Пример 2. Декодируем принятое слово $c = 1011100,$ полученное при передаче некоторого кодового слова $b,$ предполагая, что при передаче произошла ошибка не более, чем в одной позиции. Имеем:

$$cM = 1011100 \begin{bmatrix} 0\ 0\ 1 \\ 0\ 1\ 0 \\ 0\ 1\ 1 \\ 1\ 0\ 0 \\ 1\ 0\ 1 \\ 1\ 1\ 0 \\ 1\ 1\ 1 \end{bmatrix} = 011 (=3).$$

Таким образом, при передаче кодового слова b произошла ошибка в 3-й позиции, исправляя которую, получим, что было передано слово $b = 1001100,$ декодируя которое, получаем исходное сообщение $a = 0100.$

Замечание 1. Матрица M обладает следующим свойством: сумма элементов каждого столбца равна 0. Используя это свойство, нетрудно обосновать следующий метод **«быстрого» умножения** вектора $C = c_1c_2c_3c_4c_5c_6c_7$ на матрицу M. Определяем, каких элементов (0 или 1) в векторе C меньше. Например, в случае $c = 1011100$ (см. пример 2) меньше нулей (их 3, а единиц 4). Затем складываем (поэлементно, по модулю 2) строки матрицы M с номерами строк, соответствующими номерам позиций символа 0 в векторе C:

$$0\ 1\ 0 + 1\ 1\ 0 + 1\ 1\ 1 = 0\ 1\ 1 (=3).$$

Можно также сложить строки матрицы M с номерами этих строк, соответствующими номерам позиций символа 1 в векторе C:

$$0\ 0\ 1 + 0\ 1\ 1 + 1\ 0\ 0 + 1\ 0\ 1 = 0\ 1\ 1 (=3).$$

Ответы одинаковые, но в первом случае складывается меньшее число строк.

Замечание 2. Код Хемминга является матричным, т.е. можно указать матрицу $G = G_{4\times7}$, называемую *порождающей*, такую, что

$$b_1b_2b_3b_4b_5b_6b_7 = a_1a_2a_3a_4G.$$

Из системы уравнений

$$\begin{cases} b_1 = a_1 + a_2 + a_4, \\ b_2 = a_1 + a_3 + a_4, \\ b_3 = a_1, \\ b_4 = a_2 + a_3 + a_4, \\ b_5 = a_2, b_6 = a_3, b_7 = a_4, \end{cases}$$

следует, что порождающая матрица (4,7)-кода Хемминга имеет вид:

$$G = \begin{bmatrix} 1\ 1\ 1\ 0\ 0\ 0\ 0 \\ 1\ 0\ 0\ 1\ 1\ 0\ 0 \\ 0\ 1\ 0\ 1\ 0\ 1\ 0 \\ 1\ 1\ 0\ 1\ 0\ 0\ 1 \end{bmatrix}.$$

Для **«быстрого» умножения** вектора $a = a_1 a_2 a_3 a_4$ на G складываем (поэлементно, по модулю 2) строки матрицы G с номерами этих строк, соответствующими номерам позиций символа 1 в векторе a. Например,

$$1010G = 1\,1\,1\,0\,0\,0\,0 + 0\,1\,0\,1\,0\,1\,0 = 1\,0\,1\,1\,0\,1\,0.$$

Следует также иметь в виду тривиальные случаи:

$$0000G = 0\,0\,0\,0\,0\,0\,0,\ 1111G = 1\,1\,1\,1\,1\,1\,1.$$

Заметим, что сумма элементов в каждом столбце матрицы G равна 1. Из этого следует простой способ умножения a на G в случае наличия в a ровно одного элемента 0. В этом случае aG - вектор, двойственный к строке матрицы G, номер которой совпадает с номером нулевого элемента в a. Например (см. пример 1),

$$1011G = \neg(1\,0\,0\,1\,1\,0\,0) = 0\,1\,1\,0\,0\,1\,1.$$

При решении РГР кодирование сообщений произвести двумя способами: аналогично примеру 1 и с помощью порождающей матрицы (см. замечание 2).