

УДК 519.725 (045)

А. Я. Белецкий, д-р техн. наук, проф.,
П. Л. Ткаченко

НЕФОРМАЛЬНЫЙ СИНТЕЗ КОДОВ ХЭММИНГА

Институт электроники и систем управления, e-mail: abelnau@ukr.net

Описан неформальный способ построения порождающих и проверочных матриц несистематических кодов Хэмминга по алгоритму, подобному алгоритму построения матриц, соответствующих систематическим кодам.

Ключевые слова: коды Хэмминга, порождающие и проверочные матрицы.

Введение и постановка задачи. Коды Хэмминга относятся к множеству линейных блочных кодов. В этом классе кодов различают систематические и несистематические коды. Систематическим двоичным линейным кодом называется код длины n , каждое кодовое слово которого начинается с k информационных символов (битов), а оставшиеся $n - k$ бит являются проверочными символами [1]. Такие коды принято обозначать (n, k) -кодами.

Наиболее удобной формой описания линейных блочных кодов является представление их в виде двоичной $(n \times k)$ -матрицы $\mathbf{G}$, состоящей из k строк и n столбцов, которая называется порождающей матрицей кода. Пространство строк матрицы $\mathbf{G}$ – это линейный код L , в котором любое кодовое слово $\mathbf{c}$, представляющее собой двоичный вектор длины n – линейная комбинация строк из $\mathbf{G}$. Для придания строкам матрицы $\mathbf{G}$ линейной независимости принято в качестве левой части $\mathbf{G}$ использовать единичную $\mathbf{E}$ матрицу k -го порядка. Пусть $\mathbf{P}$ есть $(k, n - k)$ -матрица, в которой размещены проверочные биты для каждой строки матрицы $\mathbf{E}$. В таком случае порождающую матрицу $\mathbf{G}$ можно записать в виде [1]

$$\mathbf{G} = [\mathbf{E} \quad \mathbf{P}]. \quad (1)$$

Матрица (1) является порождающей матрицей в систематическом виде. Наиболее естественным способом кодирования является отображение

$$\mathbf{c} = \mathbf{iG},$$

где $\mathbf{i}$ – информационное слово, представляющее собой k -последовательность кодируемых информационных символов, а $\mathbf{c}$ – образующая кодовое слово n -последовательность.

В качестве примера рассмотрим алгоритм составления порождающей матрицы для циклического $(7, 4)$ -кода над неприводимым полиномом (НП) третьей степени $\phi_3 = 1011$. Для выбранного НП строки матрицы $\mathbf{P}$ в (1) представляют собой двоичные векторы длины 3 и образуются как остаток от деления соответствующей строки единичной матрицы $\mathbf{E}$, дополненной справа тремя нулями, на полином ϕ_3 , т. е.

$$\mathbf{P} = \mathbf{E}000 \bmod \phi_3. \quad (2)$$

На основании соотношений (1) и (2) получим образующую матрицу циклического $(7, 4)$ -кода над НП $\phi_3 = 1011$

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}. \quad (3)$$

В этом случае, например, информационное слово

$$\mathbf{i} = [1 \ 0 \ 1 \ 1] \quad (4)$$

будет преобразовано в кодовое слово

$$\mathbf{c} = \left([1 \ 0 \ 1 \ 1] \cdot \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} \right)_2 = [1 \ 0 \ 1 \ 1 \ 0 \ 0 \ 0]. \quad (5)$$

В соотношении (5) запись $(*)_2$ означает, что матричные преобразования выполняются над кольцом вычетов по $\text{mod } 2$. В дальнейшем в подобного рода преобразованиях для простоты круглые скобки и нижний индекс будем опускать.

В дополнении к порождающей матрице $\mathbf{G}$ вводится матрица $\mathbf{H}$, состоящая из n строк и $n - k$ столбцов, такая, что

$$\mathbf{cH} = \mathbf{0}. \quad (6)$$

Условие (6) означает, что n -последовательность $\mathbf{c}$ является кодовым словом в том и только в том случае, когда она ортогональна каждому вектор-столбцу матрицы $\mathbf{H}$. Это равенство позволяет проверить, является ли данное слово кодовым. Матрица $\mathbf{H}$ называется *проверочной матрицей* кода. Поскольку равенство (6) выполняется при подстановке вместо $\mathbf{c}$ любой строки матрицы $\mathbf{G}$, то допустимым является также соотношение

$$\mathbf{GH} = \mathbf{0}. \quad (7)$$

Предположим, что $\mathbf{G} = [\mathbf{E} \ : \ \mathbf{P}]$. Тогда в соответствии с соотношением (7) естественным определением проверочной матрицы в систематическом виде, очевидно, является равенство

$$\mathbf{H} = \begin{bmatrix} \mathbf{P} \\ \dots \\ \mathbf{E} \end{bmatrix}, \quad (8)$$

так как для двоичной модулярной арифметики

$$\mathbf{GH} = [\mathbf{E} \ : \ \mathbf{P}] \cdot \begin{bmatrix} \mathbf{P} \\ \dots \\ \mathbf{E} \end{bmatrix} = \mathbf{P} + \mathbf{P} = 2\mathbf{P} = \mathbf{0}.$$

На основании выражений (1), (3) и (8) для рассматриваемого (7, 4)-систематического кода приходим к следующему значению проверочной матрицы:

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (9)$$

Естественно, что равенство (7) соблюдается не только для всех строк матрицы (3), но также и для любой линейной комбинации этих строк.

А теперь предположим, что в кодовое слово, заданное правой частью выражения (5), внесена одиночная ошибка такая, что левый нуль замещен единицей, т. е. принята кодовая комбинация

$$\mathbf{c}^* = [1 \ 1 \ 1 \ 1 \ 0 \ 0 \ 0]. \quad (10)$$

Умножив вектор $\mathbf{c}^*$ на проверочную матрицу (9), получим трехразрядный двоичный вектор $\mathbf{s}$, называемый *синдромом*. Если синдром не равен нулю, то это означает, что в принятую кодовую комбинацию вкралась ошибка. В рассматриваемом примере

$$\mathbf{s} = \mathbf{c}^* \mathbf{H} = [1 \ 1 \ 1].$$

Для обнаружения и исправления одиночного разряда в циклических кодах производят следующие операции [2].

1. Принятую комбинацию делят на образующий полином. В рассматриваемом примере циклического кода, устраняющего одиночную ошибку, образующий полином совпадает с выбранным неприводимым полиномом третьей степени $\phi_3 = 1011$.

2. Подсчитывают количество единиц в остатке (вес остатка ω). Если $\omega = 1$, принятую комбинацию складывают по mod 2 с полученным остатком.

3. Если $\omega > 1$, то производят циклический сдвиг принятой комбинации влево на один разряд. Комбинацию, полученную в результате циклического сдвига, делят на образующий полином. Если в результате повторного деления $\omega = 1$, то делимое суммируют с остатком.

4. Производят циклический сдвиг вправо на один разряд комбинации, полученной в результате суммирования последнего делимого с последним остатком. Полученная комбинация уже не содержит ошибок.

5. Если после первого циклического сдвига и последующего деления остаток получается таким, что его вес $\omega > 1$, то повторяют операцию п. 3 до тех пор, пока не будет достигнуто равенство $\omega = 1$. В этом случае комбинацию, полученную в результате последнего циклического сдвига, суммируют с остатком от деления этой комбинации на образующий многочлен.

6. Производят циклический сдвиг вправо ровно на столько разрядов, насколько сдвинута суммируемая с последним остатком комбинация относительно принятой комбинации. В результате получим исправленный код.

Легко проверить, что искаженный код (10) может быть исправлен за два циклических сдвига влево.

Отличительная особенность кодов Хэмминга состоит в том, что его проверочные разряды не плотно расположены в правых разрядах кодовых слов, как это имеет место в рассмотренных выше циклических кодах, а «вмонтированы» в разряды, номера которых заданы двоично-рациональными числами 1, 2, 4, ... [3]. Отмеченные свойства кодов Хэмминга приводят к некоторым проблемам при вычислении синдрома, связанными с необходимостью «выковыривания» проверочных разрядов.

Классические коды Хэмминга, устраняющие одиночные ошибки, принадлежат подмножеству *несистематических блочных кодов*.

Основная задача, решению которой посвящена эта работа, состоит в разработке метода построения образующих и проверочных матриц для кодов Хэмминга, подобных соответствующим матрицам циклических кодов.

Синтез образующих и проверочных матриц кодов Хэмминга. Обратимся к анализу семиразрядного кода Хэмминга, в котором четыре разряда отведены для размещения информационных символов, а оставшиеся три – для размещения проверочных символов, которые на рис.1 отмечены звездочками.

7	6	5	4	3	2	1
			*		*	*

Рис. 1. Разметка проверочных разрядов

Впишем в пустые клетки однострочной таблицы на рис. 1 информационное слово (4).

Алгоритм вычисления проверочных разрядов (ПР), отвечающих информационному слову (4), а также приведенному на рис. 2, чрезвычайно простой [4]. С этой целью достаточно выписать в столбик двоичные номера разрядов N , в которых находятся единицы, и просуммировать числа в столбцах по mod 2, как это показано в табл. 1.

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

Рис. 2. Размещение информационных элементов

Размещая полученные значения проверочных разрядов в соответствующих ячейках на рис. 2, приходим к такому выражению для кодового слова

$$c = [1 \ 0 \ 1 \ 0 \ 1 \ 0 \ 1]. \quad (11)$$

Если снова просуммировать двоичные номера разрядов слова (11), то приходим к нулевому значению синдрома. Это означает, что кодовое слово (11) составлено без искажения.

Алгоритм вычисления проверочных разрядов (ПР), отвечающих информационному слову (4), а также приведенному на рис. 2, чрезвычайно простой [4]. С этой целью достаточно выписать в столбик двоичные номера разрядов N , в которых находятся единицы, и просуммировать числа в столбцах по mod 2, как это показано в табл. 1.

Размещая полученные значения проверочных разрядов в соответствующих ячейках на рис. 2, приходим к такому выражению для кодового слова:

$$c = [1 \ 0 \ 1 \ 0 \ 1 \ 0 \ 1]. \quad (11)$$

Если снова просуммировать двоичные номера разрядов слова (11), то приходим к нулевому значению синдрома. Это означает, что кодовое слово (11) составлено без искажения.

А теперь внесем сбой, как в предыдущем разделе, в шестой разряд слова (11), заменив в нем 0 на 1. Получим

$$c^* = [1 \ 1 \ 1 \ 0 \ 1 \ 0 \ 1]. \quad (12)$$

Повторим вычисление синдрома по схеме (табл. 2), подобной той, что приведена в табл. 1.

Предложенное Хэммингом размещение проверочных символов в ячейках (разрядах) кодового слова, номера которых образуют двоично-рациональную последовательность чисел, имеет глубокий смысл. Во-первых, при этом вес ω_p номера ячейки каждого проверочного символа будет

Таблица 1

К вычислению ПР

N	Двоичный эквивалент N		
7	1	1	1
5	1	0	1
3	0	1	1
ПР	0	0	1

Таблица 2

К вычислению синдрома

N	Двоичный эквивалент N		
7	1	1	1
6	1	1	0
5	1	0	1
3	0	1	1
1	0	0	1
s	0	0	1

равен единице, т. е. $\omega_p = 1$. И, как следствие, – вес каждого информационного слова ω_i становится не меньше двух, т. е. $\omega_i \geq 2$. А это означает, во-вторых, что вес ω_c ненулевых кодовых слов оказывается не меньше трех, т. е. $\omega_c \geq 3$. Полный ансамбль допустимых кодовых слов (7, 4)-кода Хэмминга приведен в табл. 3.

Таблица 3

Полное множество кодовых слов (7, 4) – кода Хэмминга

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

Из анализа табл. 3 следует, что минимальное расстояние по Хэммингу между любыми парами кодовых слов равно трем, что необходимо и достаточно для обнаружения и устранения одиночной ошибки в (7, 4)-кодах Хэмминга.

Несмотря на то, что рассмотренный выше способ формирования кодов Хэмминга достаточно простой, но вместе с тем он и утомительный. Попробуем перейти к образующим и проверочным матрицам (7, 4)-кодов Хэмминга, подобных соответствующим матрицам циклических кодов. Положим в основу формирования данных матриц выражения (1) и (8). С этой целью поменяем местами четвертый и третий разряды в (7, 4)-кодах Хэмминга. Это делается для того, чтобы формально проверочные разряды размещались справа от информационных разрядов, как и в образующей матрице (3) для циклических кодов. Получим

$$\begin{array}{ccccccc}
 & 7 & 6 & 5 & 3 & 4 & 2 & 1 \\
 \mathbf{G} = & \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}
 \end{array} \quad (13)$$

Над матрицей (13) указаны номера столбцов исходных (без перестановки разрядов) кодов Хэмминга.

На основании соотношений (8) и (13) составим проверочную матрицу

$$H = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (14)$$

Легко убедиться в том, что матрицы (13) и (14) удовлетворяют условию (7). А это означает, что как отдельные кодовые слова, представленные строками матрицы (13), так и их любые линейные комбинации, умноженные справа на матрицу (14), образуют нулевой синдром. И это правильно, поскольку кодовые слова не искажены помехой.

Вычислим кодовое слово $\mathbf{c}$, отвечающее информационному слову (4). Имеем

$$\mathbf{c} = [1 \ 0 \ 1 \ 1] \cdot \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} = [1 \ 0 \ 1 \ 1 \ 0 \ 0 \ 1]. \quad (15)$$

Условимся нумеровать разряды сформированного кодового слова (15) справа налево натуральным рядом чисел. Легко проверить, что если произошел сбой в одном из следующих разрядов слова $\mathbf{c}$ $\{1, 2, 5, 6, 7\}$, то синдром $\mathbf{s}$ в точности будет равен двоичному номеру разряда, который подвергся искажению. Если же помеха наложилась на третий или четвертый разряды, то синдром указывает на наличие сбоя в смежном (соответственно, в четвертом или третьем) разрядах, что является следствием перестановки столбцов в матрице (13).

А теперь составим порождающую матрицу (15, 11) – кода Хэмминга, группируя, как и в матрице (13), проверочные разряды в правой части матрицы.

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \end{bmatrix}. \quad (16)$$

Проверочная матрица $\mathbf{H}$ для (15, 11)-кода Хэмминга представляет собой (15, 4)-матрицу, в верхней части которой размещены четыре правых столбца матрицы (16), а в нижней – единичная матрица четвертого порядка. Таким образом,

$$\mathbf{H}^T = \begin{matrix} & 15 & 14 & 13 & 12 & 11 & 10 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \end{matrix}. \quad (17)$$

Выберем, для примера, 11-разрядное информационное слово

$$\mathbf{i} = [1 \ 1 \ 0 \ 1 \ 0 \ 0 \ 1 \ 1 \ 0 \ 1 \ 1],$$

умножив которое справа на матрицу (16), преобразуем его в 15-разрядное кодовое слово

$$\mathbf{c} = \begin{matrix} & 15 & 14 & 13 & 12 & 11 & 10 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ [1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1] \end{matrix}. \quad (18)$$

Номера разрядов вектора (18) и столбцов матрицы (17) обозначены верхней строчкой цифр. Ниже вектора (18) выписана перестановка столбцов порождающей матрицы кода Хэмминга (и, соответственно, разрядов кодового слова), посредством которой матрица (16) приведена к форме порождающей матрицы систематического кода. Такую перестановку удобнее представить в табличной форме в виде подстановки (табл. 4).

Таблица 4

Подстановка для (15, 11)-кода Хэмминга

<i>x</i>	3	4	5	6	7	8
<i>s</i>	4	8	3	5	6	7

Предположим, что в процессе передачи информационного слова (18) по зашумленному каналу связи в одном из его разрядов возможно произошел сбой (замена 0 на 1, или наоборот) и на вход приемника поступает слово $\mathbf{c}^*$. Для определения номера пораженного разряда необходимо вычислить синдром

$$\mathbf{s} = \mathbf{c}^* \cdot \mathbf{H},$$

где $\mathbf{H}$ – проверочная матрица, заданная соотношением (17).

Нулевое значение синдрома указывает на то, что кодовое слово принято без искажения. При значении $\mathbf{s} > 0$ возможны два решения относительно номера пораженного разряда. Если номер разряда, в котором произошел сбой, не входит в таблицу подстановок (табл. 4), то десятичное значение синдрома точно соответствует номеру пораженного разряда. В том случае, когда сбой произошел в разряде, номер которого принадлежит подмножеству номеров из табл. 4, то синдром вырабатывает число, содержащееся в строке *s* таблицы подстановок. Фактически при этом номер пораженного разряда следует искать в соответствующей ячейке строки *x* табл. 4. Например, пусть $\mathbf{s} = 1000$, что соответствует числу 8 в строке *s* табл. 4. Это означает, что сбой произошел в четвертом разряде принятого информационного слова.

Приведем таблицу подстановок для (31, 26)-кода Хэмминга (табл. 5).

Таблица 5

Подстановка для (31, 26)-кода Хэмминга

<i>x</i>	3	4	5	6	7	8	9	10	11	12	13	14	15	16
<i>s</i>	4	8	16	3	5	6	7	9	10	11	12	13	14	15

Следуя изложенной выше методике, легко составить как образующие и проверочные матрицы, так и таблицы подстановок для кодов Хэмминга с произвольными n, k параметрами. Синдром $\mathbf{s}$, образуемый произведением принятой кодовой комбинации $\mathbf{c}^*$ и проверочной матрицы $\mathbf{H}$, численно равен номеру разряда $\mathbf{c}^*$, в котором произошел сбой, но при условии, что значение $\mathbf{s}$ не входит в подмножество чисел, содержащихся в подстановке. В противном случае номер пораженного разряда определяется из соответствующей ячейки строки x таблицы подстановки.

Выводы. Таким образом, на основе неформального подхода разработан алгоритм синтеза образующих и проверочных матриц для (n, k) – кодов Хэмминга, а также решена задача локализации и устранения одиночной ошибки в любом разряде кода. Разработанные алгоритмы допускают достаточно простую как программную, так и аппаратную реализацию.

Список литературы

1. Блейхут Р. Теория и практика кодов, контролирующих ошибки: пер. с англ. – М.: Мир, 1986. – 576 с.
2. Березюк Н. Т. Кодирование информации: справ. – Х.: Изд-во «Вища школа», 1978. – 250 с.
3. Хэмминг Р. В. Теория кодирования и теория информации. – М.: Радио и связь, 1983. – 176 с.
4. Семенов Ю. А. Коррекция ошибок http://book.itep.ru/2/28/corec_28.htm.

А. Я. Білецький, П. Л. Ткаченко

Неформальний синтез кодів Хеммінга

Описано неформальний спосіб побудови породжувальних та перевірних матриць несистематичних кодів Хеммінга за алгоритмом, подібним алгоритму побудови матриць, що відповідають систематичним кодам.

A. J. Beletsky, P. L. Tkachenko

Informal synthesis of Hamming codes

We describe an informal method of constructing the generators and the check matrix RIP nonsystematic Hamming algorithm, such an algorithm for constructing matrices corresponding systematic codes.