Лабораторная работа № 4

Помехоустойчивое кодирование

1. Основные сведения о методах помехоустойчивого кодирования

Помехоустойчивые коды применяют для уменьшения влияния помех на сообщения. Построение помехоустойчивых кодов основано на добавлении к исходной комбинации из k символов r контрольных символов. Закодированная комбинация будет составлять n символов. Поэтому помехоустойчивые коды часто называют (n, k)-коды.

К простейшим помехоустойчивым кодам относят следующие коды для обнаружения ошибок:

1. код четности, который образуется путем добавления к передаваемой комбинации, состоящей из k информационных символов, одного контрольного символа (0 или 1), так, чтобы общее число единиц в передаваемой комбинации было четным;

2. код с постоянным весом, который содержит постоянное число единиц и нулей;

3. корреляционный код (код с удвоением), при построении которого 1 преобразуется в 10, а 0 – в 01;

4. инверсный код, получаемый при добавлении к исходной комбинации такой же комбинации по длине: если в исходной комбинации четное число единиц, то добавляемая комбинация повторяет исходную комбинацию, если нечетное – то добавляемая комбинация является инверсной относительно исходной;

5. код Грея, для построения которого используются следующие правила:

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

$$A_i = \begin{cases} a_n, i = n, \\ a_i, a_{i-1} \otimes a_{i-2} \otimes \ldots \otimes a_0 = 0, \\ \overline{a}_i, a_{i-1} \otimes a_{i-2} \otimes \ldots \otimes a_0 = 1; \end{cases}$$

где $A_nA_{n-1}\ldots A_0$ – исходная двоичная комбинация, а $a_na_{n-1}\ldots a_0$ – соответствующий ей код Грея.

Свойства помехоустойчивых кодов определяются кодовым расстоянием. Кодовое расстояние d – это минимальное число символов, в которых одна кодовая комбинация отличается от другой. Если $d = 1$, то код не обладает помехоустойчивыми свойствами, если $d = 2$, то код позволит обнаружить одиночные ошибки и т.д. Таким образом, увеличивая кодовое расстояние можно увеличить помехоустойчивость кода. В общем случае кодовое расстояние определяется по формуле

$$d = t + l + 1,$$

где t – число исправляемых ошибок, l – число обнаруживаемых ошибок (обычно $l > t$).

Коды, которые позволяют обнаруживать и исправлять ошибки, называют корректирующими кодами. Большинство корректирующих кодов являются линейными кодами. Линейные коды – это такие коды, у которых контрольные символы образуются путем линейной комбинации информационных символов. Кроме того, корректирующие коды являются групповыми кодами. Групповые коды G_n – это такие коды, которые имеют одну основную операцию. При этом должно соблюдаться условие замкнутости (т.е. 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.

Для построения кода способного обнаруживать и исправлять одиночную ошибку необходимое число контрольных разрядов будет составлять

$$n - k \geq \log(n + 1),$$

где k – число разрядов исходной кодовой комбинации, n – число разрядов после добавления контрольных символов. Если необходимо исправить две ошибки, то

$$n - k \geq \log_2\left(1 + C_n^1 + C_n^2\right).$$

В этом случае обнаруживаются однократные и двукратные ошибки. В общем случае, число контрольных символов определяется неравенством Хэмминга:

$$n - k \geq \log(1 + C_n^1 + C_n^2 + \ldots + C_n^t) = \log_2 \sum_{i=0}^{t} C_n^i .$$

Одними из наиболее широко применяемых корректирующих кодов являются циклические коды. Циклическими кодами называют специальную группу кодов, которые описываются полиномиально. Полиномиальное описание кодовых комбинаций заключается в следующем. Пусть, например, имеется кодовая комбинация 101101, тогда ее можно представить в виде полинома

$$A(X) = 1 \cdot x^5 + 0 \cdot x^4 + 1 \cdot x^3 + 1 \cdot x^2 + 0 \cdot x^1 + 1 = x^5 + x^3 + x^2 + 1.$$

Циклические коды относятся к систематическим (n, k)-кодам, в которых контрольные r и информационные k разряды расположены на строго определенных местах: $n = k + r$. При выполнении действий над циклическими кодами в многочленной форме операции умножения и вычитания выполняются как сложение по модулю 2.

Для получения циклического кода заданный многочлен $h(x)$ сначала умножается на одночлен x^{n-k}, затем делится на образующий многочлен $g(x)$. В результате получаем:

$$\frac{h(x)x^{n-k}}{g(x)} = Q(x) + \frac{R(x)}{g(x)}.$$

После этого к произведению $h(x)x^{n-k}$ добавляется остаток $R(x)$:

$$F(x) = Q(x) \cdot g(x) = x^{n-k}h(x) + R(X).$$

При декодировании, принятую кодовую комбинацию необходимо разделить на $g(x)$. Наличие остатка указывает на ошибку. Образующий

полином $g(x)$ является сомножителем при разложении двучлена x^n+1. Сомножителями разложения двучлена являются неприводимые полиномы из таблицы 2.

Образующий полином выбирают следующим образом. По заданной кодовой комбинации k определяют число контрольных символов из соотношения $r = \log_2(n + 1)$ или по эмпирической формуле

$$r = [\log_2\{(k + 1) + [\log_2(k + 1)]\}].$$

Соотношение значений n, k, r показано в таблице 1.

Таблица 1

Соотношение между n, k, r

n	3	5	6	7	9…15	17…31	33…63	65…127
k	1	2	3	4	5…11	12…26	27…57	28…120
r	2	3	3	3	4	5	6	7

Затем из таблицы 2 выбирают самый короткий неприводимый полином со степенью, равной числу контрольных символов.

Например, пусть требуется закодировать комбинацию вида 1101, что соответствует $h(x) = x^3 + x^2 + 1$.

1. Определяем число контрольных символов: $r = 3$.

2. Выбираем образующий полином: $g(x) = x^3 + x + 1$, т.е. 1011.

3. Умножаем $h(x)$ на x^r:

$$h(x)x^r = (x^3 + x^2 + 1)x^3 = x^6 + x^5 + x^3 ,$$

что соответствует 11010000.

4. Разделим полученное произведение на образующий полином $g(x)$:

$$\frac{h(x)x^r}{g(x)} = \frac{x^6 + x^5 + x^3}{x^3 + x + 1} = x^3 + x^2 + x + 1 + \frac{1}{x^3 + x + 1} = 1111 + \frac{0001}{1011}$$

5. Остаток суммируем с $h(x)x^r$:

$$F(x) = (x^3 + x^2 + 1)(x^3 + x + 1) = (x^3 + x^2 + 1)x^3 + 1, \text{ т. е. } 1101001.$$

Полученная кодовая комбинация $F(x)$ циклического кода содержит исходную комбинацию $h(x)$ = 1101 и контрольные символы $R(x)$ = 001. Очевидно, что закодированное сообщение делится на образующий полином без остатка.

Для рассмотренного примера исходное сообщение является одной из 16 возможных комбинаций 4-разрядного кода. Это значит, что если все сообщения необходимо преобразовать в циклический код, то каждое из них необходимо кодировать, выполняя такую же последовательность вычислений, что и для $h(x) = 1101$. Однако выполнять дополнительные 15 расчетов (в общем случае $2^{n-k} - 1$ расчетов) нет необходимости, если составить образующую (порождающую) матрицу $G_{k\times n}$, которая составляется на основе единичной матрицы I_k, к которой справа дописывается матрица остатков $R_{k\times(n-k)}$:

$$G_{k\times n} = \left(I_k R_{k\times(n-k)}\right).$$

Матрица $R_{k\times(n-k)}$ содержит остатки от последовательного деления единицы с нулями на образующий многочлен $g(x)$, например:

1000000 / 1011 = 1 (1–й остаток – 011)

011000 / 1011 = 0 (2–й остаток – 110)

11000 / 1011 = 1 (3–й остаток – 111)

1110 / 1011 = 1 (4–й остаток – 101).

При делении, начиная со второго шага, в качестве делимого используется остаток, найденный на предыдущем шаге. Так же отметим, что располагать остатки в матрице нужно, начиная с последнего. Таким образом, для рассматриваемого примера образующая матрица имеет вид

$$G_{4\times 7} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$

Матрица $G_{4\times 7}$ уже содержит 4 комбинации циклического кода, а остальные 12 ненулевых комбинаций находятся суммированием по модулю 2 всевозможных сочетаний строк образующей матрицы

Для обнаружения и исправления ошибок принятая комбинация делится на образующий многочлен $g(x)$. Если остаток $R(x)$ будет равен 0, то комбинация принята без ошибок. Наличие ненулевого остатка

свидетельствует о том, что комбинация принята искаженной. Значение остатка совпадет с одним из опознавателей транспонированной проверочной матрицы $H^T_{(n-k)\times n}$, который и укажет на местоположение. Проверочная матрица имеет вид:

$$H_{(n-k)\times n} = \left(R^T_{k\times(n-k)} I_{n-k}\right),$$

например, для циклического кода из примера проверочная матрица будет следующей

$$H_{3\times 7} = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

Тогда, если вместо 1101001 получена кодовая комбинация 1100001, то остаток от деления 1100001 на 1011 будет равен 011. Остаток совпадает с четвертой строкой матрицы $H^T_{(n-k)\times n}$:

$$H^T_{3\times 7} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Это означает, что ошибка содержится в 4-м разряде, исправив который получаем правильную комбинацию 1101001.

Таблица 2

Образующие полиномы

$g(x)$	Полином	$g(x)$	Полином
$g(x)$	$x+1$	$g_1(x^6)$	x^6+x+1
$g(x^2)$	x^2+x+1	$g_2(x^6)$	x^6+x^3+1
$g_1(x^3)$	x^3+x+1	$g_3(x^6)$	x^6+x^5+1
$g_2(x^3)$	x^3+x^2+1	……	……
$g_1(x^4)$	x^4+x+1	$g_1(x^7)$	x^7+x+1
$g_2(x^4)$	x^4+x^3+1	$g_2(x^7)$	x^7+x^3+1
$g_3(x^4)$	$x^4+x^3+x^2+x+1$	$g_3(x^7)$	$x^7+x^3+x^2+x+1$
$g_1(x^5)$	x^5+x^2+1	……	……
$g_2(x^5)$	x^5+x^3+1	$g_1(x^8)$	$x^8+x^4+x^3+x+1$
$g_3(x^5)$	$x^5+x^3+x^4+x+1$	$g_2(x^8)$	$x^8+x^4+x^3+x^2+1$
$g_4(x^5)$	$x^5+x^4+x^2+x+1$	……	……
$g_5(x^5)$	$x^5+x^4+x^3+x+1$	$g_1(x^9)$	x^9+x+1
$g_6(x^5)$	$x^5+x^4+x^3+x^2+1$	$g_2(x^9)$	x^9+x^4+1

2. Порядок выполнения лабораторной работы

1. Ознакомиться с основными сведениями по помехоустойчивому кодированию.

2. Получить задание на выполнение лабораторной работы.

3. Выполнить необходимые расчеты.

4. Сделать выводы по результатам выполнения лабораторной работы.

5. Оформить отчет о выполнении лабораторной работы.

6. Ответить на контрольные вопросы.

3. Контрольные вопросы

1. Назначение помехоустойчивых кодов?

2. Как строятся коды с проверкой на четность, с удвоением, с постоянным весом, инверсные?

3. Какие коды называются корректирующими?

4. Что определяет минимальное кодовое расстояние?

5. Как определяются линейные коды?

6. Как определяются циклические коды?

7. Как выбирается образующий многочлен?

8. Как построить образующую и проверочную матрицы циклического кода?

9. Как выполняются кодирование и декодирование циклического кода?

10. Как выявляется и исправляется ошибка в циклическом коде?

4. Задания на лабораторную работу

Дан алфавит латинских символов A = {a, …, z}.

1. Определить кодовые комбинации символов алфавита

2. Выбрать неприводимый полином из таблицы 2.

3. Выполнить кодирование сообщения из таблицы 3 с использованием циклического кода.

4. Выполнить декодирование закодированного сообщения.

5. Внести ошибку в закодированное сообщение.

6. Выполнить декодирование закодированного сообщения с ошибкой.

7. Выполнить кодирование сообщения с использованием кодов с проверкой на четность, с удвоением, с постоянным весом, инверсные для исходного сообщения.

Таблица 3

Сообщения дискретного источника

№	Сообщение
1	*abcaaaabacabbacbbaccbbaccbbddadadaa*
2	*ecaeecdabacbbacbbddcbbaccbbdbdadaae*
3	*dddbhjgffsdkkkdffgdhgdfhfhjfghkfgjk*
4	*abcazzgabbacbaacljabaccbbaccssddadd*
5	*aaddaddabacabbacbmmncbbaccjjidadada*
6	*cccaddabttdccaabcaassxacabbacbbacbb*
7	*dbdaadaxghabbacbbacnvbaccbbddadadac*
8	*bbbiiaabacabbacbbacllbacdbffdadadac*
9	*aacabaarrcdbbacbdewcbbacqobddadadbo*
10	*dppuadabbttrbacbbacyubaccbuuwdaddaa*
11	*abccppjbacabnscbbadsmdaccbmndadadcj*
12	*abcbbgqqacabqqwddacerbaccbuydadyydd*