В. А. АРАКЕЛОВ, Г. М. ТЕНЕНГОЛЬЦ

НЕКОТОРЫЕ КЛАССЫ КОРРЕКТИРУЮЩИХ КОДОВ

В теории кодирования используются сравнения вида

$$W_{a,\beta} = \sum_{i=1}^{n} \beta_i \gamma(\alpha_i) \equiv a \pmod{m}, \qquad (1)$$

где β_i - члены заданной последовательности, $a = 0, 1, \cdots, m-1$; $\alpha_i = 0, 1, 2, \cdots, q-1$, $q \geqslant 2$ — целое, $\gamma(\alpha_i)$ — коэффициенты, зависящие от величины α_i.

Использование (1) при различных начальных данных позволяет решить ряд технических задач. В частности, в [1] было показано, что совокупность всевозможных бинарных решений выражения (1) при $\beta_i = i$, $m = n+1$; $\gamma(\alpha_i) = \alpha_i$ образует код длины n, исправляющий одиночные несимметрические ошибки (т. е. ошибки типа $0 \to 1$ или $1 \to 0$). Этот же код [2] исправляет одиночные вставки или выпадения символов в сигнале. При $m = 2n$, $\beta_i = i$, $\gamma(\alpha_i) = \alpha_i$ множество двоичных решений сравнения (1) образует код длины n, исправляющий одиночные вставки или выпадения, или симметрические искажения символов в сигнале. В случае, когда $m = n^2 + n + 1$, $n = p^k$; p — простое; $k > 0$ — произвольное целое, β_i являются членами последовательности Зингера*) [3], совокупность всевозможных двоичных решений сравнения (1) образует код длины n, исправляющий двойные и одиночные несимметрические ошибки [4].

В настоящей работе сравнение (1) используется для построения кодов с основанием $q > 2$, устойчивых к одиночным сбоям типа вставок или выпадений символов, а также с исправлением одиночных несимметрических ошибок. Исследуется частный случай сравнения (1), позволяющий строить двоичный код, исправляющий одиночные симметрические ошибки. Указанные коды в бинарном случае рассматриваются в систематическом виде. Приводятся схемы кодирования и декодирования этих кодов.

---

*) Последовательность Зингера обладает тем свойством, что любое целое число $a \not\equiv 0 \pmod{n}$ единственным образом представимо в виде $b_k - b_l \equiv a \pmod{n}$, $k \neq l$.

§ 1. Класс кодов, исправляющих одиночные выпадения или вставки символов в сигналах

В системах связи возможна такая ситуация, когда полученное сообщение имеет длину, отличную от длины посланного сигнала. Такого рода искажения могут иметь место, например, при нарушении синхронизации работы приемника и передатчика системы связи.

Поэтому представляет интерес исследование каналов, в которых допускаются сбои вида $\alpha \to \Lambda$, называемые выпадениями, и сбои вида $\Lambda \to \alpha$, называемые вставками (здесь Λ — пустое слово; $\alpha = 0, 1, \cdots$ $\cdots, q-1$, $q \geqslant 2$ — целое). При синтезе кодов, устойчивых к сбоям указанного типа, может быть использовано сравнение (1) в следующем виде:

$$W = \sum_{i=1}^{n} i\,\beta(\alpha_i) \equiv a \pmod{m}, \tag{2}$$

где $\beta(\alpha_i)$ — коэффициенты, зависящие от величины α_i,

$$\alpha_i = 0, 1, \cdots, q-1;\ q \geqslant 2,$$

a — произвольное целое, $m = \sum_{v=1}^{q-1} n^v$,

$$\beta(\alpha_i) = \sum_{t=0}^{\alpha_i - 1} n^t \quad \left(\text{по определению } \sum_{t=0}^{-1} n^t = 0\right).$$

Имеет место

Теорема 1. Множество $K_a(q)$ последовательностей вида $\alpha_1, \alpha_2, \cdots, \alpha_n$, где $\alpha_1, \alpha_2, \cdots, \alpha_n$ удовлетворяют сравнению (2), является кодом, исправляющим одиночные вставки или выпадения символов в сигнале.

Доказательство. Пусть на передающем конце сигнал имеет вид $x = \alpha_1, \alpha_2, \cdots, \alpha_n$. При выпадении одного символа он воспринимается на приемном конце в виде $y = \alpha'_1, \alpha'_2, \cdots, \alpha'_{n-1}$. Для восстановления слова x по слову y надо знать: 1) какой из q возможных символов выпал; 2) какую позицию занимал выпавший символ.

Зная слово y, можно определить наименьшее неотрицательное решение сравнения

$$t \equiv -\sum_{i=1}^{n-1} i\,\beta(\alpha'_i) + a \left(\bmod \sum_{v=0}^{q-1} n^v \right)$$

и совокупность чисел $|y|^{(1)}, |y|^{(2)}, \cdots, |y|^{(q-1)}$, где $|y|^{(l)}$ $(l = 1, 2, \cdots$ $\cdots, q-1)$ — число символов „l“ в слове y. Обозначим через n_l $(l = 0, 1, 2, \cdots, q-1)$ — число символов „l“ левее выпавшего символа. Нетрудно видеть, что число t единственным образом представимо в виде:

$$t = \sum_{l=1}^{q-1} a_l \sum_{u=0}^{l-1} n^u, \tag{3}$$

где a_l ($l = 1, 2, \cdots, q-1$) — некоторые целые неотрицательные числа, не превосходящие n.

Если выпал символ „0", то коэффициенты a_l удовлетворяют соотношению

$$a_l = |y|^{(l)} - n_l. \tag{4}$$

Если же выпал символ „j", отличный от нуля, то

$$\begin{cases} a_l = |y|^{(l)} - n_l, & l \neq j \\ a_j = |y|^{(j)} + \sum\limits_{\substack{\nu=0 \\ \nu \neq j}}^{q-1} n_\nu + 1. \end{cases} \tag{5}$$

Из сказанного следует, что при выпадении символа „0" $a_l \leqslant |y|^{(l)}$, а при выпадении символа „j" $a_j > |y|^{(j)}$. При этом место выпадения в первом случае определится решениями системы (4), а во втором — системы (5). Следовательно, по слову y однозначно восстанавливается слово x, а рассматриваемое множество является кодом, устойчивым к одиночным выпадениям символов в сигнале. Так как код, исправляющий одиночные ошибки типа выпадения, является также кодом с исправлением одиночных вставок, то теорема доказана.

Следует отметить, что процедура декодирования в случае одиночных вставок такая же, как и в случае выпадений, с той лишь разницей, что в качестве t берется наименьшее неотрицательное решение сравнения $t \equiv \sum\limits_{i=1}^{n-1} i\,\beta(x_i') - a \pmod{\sum\limits_{\nu=0}^{q-1} n^\nu}$, а в системе (5) a_j примет вид:

$$a_j = |y|^{(j)} + \sum_{\substack{\nu=0 \\ \nu \neq j}}^{q-1} n_\nu.$$

Если $q = 2$, то предлагаемый код совпадает с кодом, рассматриваемым в работе [2]. В этом случае при условии одиночного выпадения

$$t \equiv -\sum_{i=1}^{n-1} i x_i' + a \pmod{n+1} \tag{6}$$

и имеют место следующие равенства:

$$t = |y| - n_1 = n^1, \tag{7}$$

$$t = |y| + n_0 + 1, \tag{8}$$

где $|y|$ — норма кодового слова y, n^1 — число единиц правее выпавшего символа „0". Если $t < |y|$, то выпал символ „0", и $t > |y|$ при выпадении символа „1", а номер выпавшей позиции определится согласно (7) или (8) соответственно.

В случае одиночных вставок

$$t \equiv \sum_{i=1}^{n+1} i z_i - a \pmod{n+1}, \tag{6a}$$

а (7) и (8) примут соответственно вид:

$$t = |y| - n_1 = n^1 \tag{7a}$$

$$t = |y| + n_0. \tag{8a}$$

Число сигналов в коде с исправлением одиночных выпадений или вставок

Для мощности M наилучшего (в смысле количества элементов) кода в предлагаемом классе кодов $K_a(q)$ справедлива следующая оценка снизу:

$$M \geqslant \frac{q^n}{\sum_{v=0}^{q-1} n^v}.$$

В случае $q = 2$ в качестве наилучшего согласно [5] можно взять код $K_0(q)$, причем, как показано в [2], этот код является асимптотически оптимальным.

Система кодирования, устойчивая к одиночным сбоям типа вставок или выпадений символов

Рассматриваемый выше код является несистематическим. Для удобства реализации видоизменим его так, чтобы получить систематический код. Будем рассматривать наиболее важный для практики двоичный код. Систематический двоичный код представляет собой множество последовательностей вида $\alpha_1, \alpha_2, \cdots, \alpha_k, \bar{\alpha}_k, \bar{\alpha}_k, \alpha_{k+1}, \cdots, \alpha_n, 1, 0$, где значения α взяты из поля $GF(2)$. Информационную часть составляют символы $\alpha_1, \alpha_2, \cdots, \alpha_k$; $\alpha_{k+1}, \cdots, \alpha_n$ — проверочные символы, представляющие собой двоичную запись наименьшего неотрицательного вычета t обобщенного веса $W = \sum_{i=1}^{k} i\alpha_i$ по mod $k+1$ в виде последовательности длины $n - k =]\log_2 k[$*), т. е.

$$t = \sum_{i=k+1}^{k} \alpha_i \cdot 2^{i-k-1}.$$

Разделительные символы $\bar{\alpha}_k \bar{\alpha}_k$, являющиеся инверсией α_k, служат для отделения информационной части от проверочной. Комбинация „1,0“ играет роль запятой между передаваемыми сообщениями.

---

*) Функция $]x[$ означает наименьшее целое число, не меньше x.

Рассмотрим работу кодирующего устройства, приведенного на рис. 1.

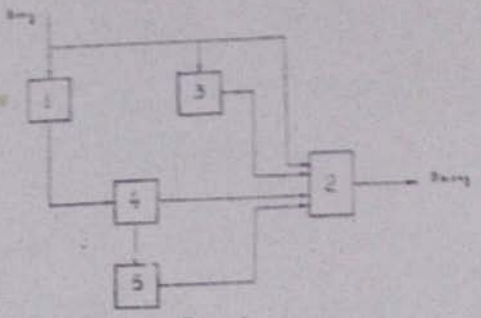

Рис. 1.

Двоичная информационная последовательность $a_1, a_2, \cdots, a_k$ поступает на вход блока 1, определяющего наименьший неотрицательный вычет t обобщенного веса $W = \sum_{i=1}^{k} i a_i$ этой последовательности по mod $k+1$. Одновременно информационные символы через блок задержек 2 поступают в канал связи. Селектор 3, выделяя k-й символ из последовательности $a_1, a_2, \cdots, a_k$, выдает на выходе символы $\bar{a}_k$ $\bar{a}_k$, которые поступают в блок 2. Полученный в блоке 1 в виде некоторой двоичной последовательности вычет t поступает на нормирующий блок 4, который пропускает через себя на блок 2 лишь $n-k$ двоичных символов. После этого включается разделительный блок 5 с целью выдачи комбинации „1,0". Поступающие на блок задержек 2 последовательности формируются в нем в кодовое слово $a_1, a_2, \cdots, a_k, \bar{a}_k, \bar{a}_k, a_{k+1}, \cdots, a_n, 1, 0$, которое передается в канал связи.

Процесс декодирования может быть осуществлен с помощью устройства, блок-схема которого приведена на рис. 2.

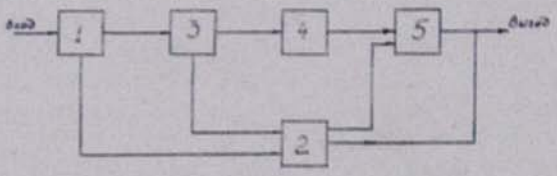

Рис. 2.

Рассмотрим сначала канал связи, в котором возможны одиночные сбои типа выпадения символов. Принимаемая последовательность символов*) поступает на блок 1 декодирующего устройства, который анализирует $n+3$-й символ. Если он окажется равным „1", то ошибки не произошло, и первые $n+2$ символа передаются в блок 2, отсчитывающий первые k символов полученного сообщения, которые выдаются на выход устройства. Процесс декодирования на этом закан-

*) Начальный момент передачи фиксирован.

чивается. Если же $n+3$-й символ равен „0“, то в сообщении произошло выпадение, и первые $n+1$ символов передаются в блок 3, который проводит анализ k-го и $k+1$-го символов. Если k-й и $k+1$-й символы противоположны, то ошибки в информационной части не произошло, а поэтому слово передается в блок 2, выделяющий первые k символов принятого сообщения с выдачей на выход декодирующего устройства. Если же k-й и $k+1$-й символы совпадают, то ошибка произошла в информационной части, и сообщение передается в блок 4, который подсчитывает наименьший неотрицательный вычет t выражения

$$-\sum_{i=1}^{k-1} i\alpha_i' + \sum_{i=k+2}^{n+1} \alpha_i' - 2^{i-k-2} \quad \text{по} \quad \mathrm{mod}\,(k+1),$$

и одновременно в блок 2, выделяющий из слова искаженную информационную часть $\alpha_1', \cdots, \alpha_{k-1}'$, поступающую в блок коррекции ошибки 5. В последний поступает также из блока 4 величина t. К поступившей в блок 5 последовательности $\alpha_1', \cdots, \alpha_{k-1}'$ слева добавляется „0“. Если $t=0$, это означает, что выпал символ „0“, расположенный в k-й позиции, и на выходе выдается информационная последовательность в виде $\alpha_1', \cdots, \alpha_{k-1}', 0$. В случае же $t \neq 0$, добавленный слева „0“ учитывается только при подсчете и в последовательности $\alpha_1', \cdots, \alpha_{k-1}', 0$ отсчитываются сначала единицы справа налево, а затем нули слева направо до тех пор, пока в сумме не наберется число t. После той позиции, на которой остановились, по ходу счета вставляется символ, отличный от предыдущего, и на выходе блока 5, таким образом, получаем исправленную информационную последовательность.

Пример. Пусть передается слово 101010011010 ($k=5$; $n=8$) и пусть в результате одиночного выпадения на вход декодирующего устройства оно поступает в виде 11010011010.

Анализ $n+3$-го символа в блоке 1 показывает, что в слове произошла ошибка, и поэтому первые $n+1$ символов 110100110 передаются в блок 3, в котором в результате сравнения k-го и $k+1$-го символов устанавливается, что ошибка произошла в информационной части.

Блок 4 определят $t \equiv -(1\cdot 1+1\cdot 2+1\cdot 4)+1\cdot 2^0+1\cdot 2^1 \equiv$
$\equiv -4 \pmod 6 \cdot t=2$.

На блок 5 поступает выделенная блоком 2 последовательность 1101 и одновременно величина $t=2$, после чего ведется счет единиц и нулей в слове 01101 так, чтобы их суммарное количество равнялось двум. Останавливаемся на единице, стоящей в третьей позиции. Следовательно, слева от нее вставляется „0“ и на выход блока 5 поступит исправленная информационная последовательность, имеющая вид 10101.

В случае канала связи, в котором возможны одиночные сбои типа вставок символов, блок-схема декодирующего устройства оста-

ется такой же, с той лишь разницей, что блок 1 анализирует $n+4$-й символ. Если $n+4$-й символ равен „0", то ошибки в сообщении не произошло, и блок 2 выделяет первые k символов сообщения, поступающие на выход декодирующего устройства. Если же $n+4$-й символ равен „1", то произошла ошибка, и первые $n+3$ символа поступают на блок 3, который сравнивает $k+1$-й и $k+2$-й символы. Если они равны, то ошибка произошла в проверочной или разделительной части, и блок 2 отделяет неискаженные информационные символы. Если же $k+1$-й и $k+2$-й символы последовательности противоположны, то ошибка произошла в информационной части, и блок 4 вычисляет наименьший неотрицательный вычет t выражения $\sum_{i=1}^{s+1} i z_i' - \sum_{i=s+4}^{n+3} z_i' 2^{i-s-4}$ по mod $k+1$, поступающий на блок 5. Одновременно на последний поступает выделенная в блоке 2 искаженная информационная последовательность $z_1', \ldots, z_{k+1}'$.

Если $t=0$, то вычеркивается $k+1$-й символ. Если же $t \neq 0$, то отсчитываются сначала только единицы, а затем только нули (счет единиц, как и в случае выпадений, ведется справа налево, а нулей — слева направо) так, чтобы в сумме набрать число, равное t. Если останавливаемся на первой позиции слова, то символ, соответствующий ей, опускается. В противном случае опускается символ после той позиции, на которой остановились по ходу счета.

Пример. Пусть передается слово 101010011010 ($k=5$, $n=8$) и пусть в результате одиночной вставки оно поступает на вход декодирующего устройства в виде 11010100011010. Анализ $n+4$-го символа в блоке 1 показывает, что в слове произошла ошибка, и поэтому первые $n+3$ символа 11010101000110 передаются в блок 3, в котором в результате сравнения $k+1$-го и $k+2$-го символов устанавливается, что произошла ошибка в информационной части.

Блок 4 определяет $t \equiv 1 \cdot 1 + 1 \cdot 2 + 1 \cdot 4 + 1 \cdot 6 - (1 \cdot 2^0 + 1 \cdot 2^1) \equiv 10 \pmod 6$. $t=4$.

На блок 5 поступает выделенная блоком 2 последовательность 110101 и одновременно величина $t=4$, после чего ведется счет единиц и нулей так, чтобы их суммарное количество равнялось 4. Остановка происходит на символе, стоящем в первой позиции, который и вычеркивается, и на выход блока 5 поступит исправленная информационная последовательность, имеющая вид 10101.

§ 2. Класс несимметрических кодов, устойчивых к одиночным искажениям

В технике связи получают распространение устройства хранения и передачи информации с несимметрической характеристикой повреждений, поэтому представляет интерес исследование несимметриче-

ских систем кодирования. В случае произвольного основания кода $q \geqslant 2$ может быть два типа несимметрических ошибок: 1) малые искажения типа $+1$ (или -1), когда каждый символ (за исключением символа $q-1$, который не искажается) может быть подвергнут искажению вида $i \to i+1$ (или $i \to i-1$); 2) большие искажения типа $(+)$ (или $(-)$), когда все символы (за исключением символа $q-1$) подвержены искажению вида

$$i \to i+k \text{ (или } i \to i-k);\ k = 1, 2, \cdots, q-1-i;\ i = 0, 1, \cdots, q-2.$$

Заметим, что множество $K_a(q)$ всевозможных последовательностей вида $\alpha_1, \alpha_2, \cdots, \alpha_n$, где $\alpha_1, \alpha_2, \cdots, \alpha_n$ — решения сравнения

$$W = \sum_{i=1}^{n} i\alpha_i \equiv a \pmod{(n+1)}, \tag{9}$$

$\alpha_i = 0, 1, 2, \cdots, q-1$, q — основание кода, a — произвольное целое, является кодом, исправляющим одиночные несимметрические ошибки типа $+1$ (или -1). Причем, согласно [5], при $a = 0$ получается код, имеющий наибольшую мощность в данном классе кодов.

Рассмотрим код, исправляющий одиночные „большие“ искажения типа $(+)$ или $(-)$.

Воспользуемся сравнением

$$\sum_{i=1}^{n} p_i \alpha_i \equiv a \pmod{(q-1)\, p_n + 1}, \tag{10}$$

где $0 \leqslant \alpha_i \leqslant q-1$; p_i — члены последовательности, обладающей тем свойством, что для любого i и j, $i \neq j$,

$$p_i \alpha_i \neq p_j \alpha_j, \text{ где } \alpha_i, \alpha_j = 1, 2, \cdots, q-1;\ i, j = 1, 2, \cdots, n$$

справедлива

Теорема 2. Совокупность $K_a(q)$ всевозможных последовательностей вида $\alpha_1, \alpha_2, \cdots, \alpha_n$, где $\alpha_1, \alpha_2, \cdots, \alpha_n$ — решения сравнения (10), является кодом, исправляющим одиночные „большие“ искажения типа $(+)$ или $(-)$.

Доказательство. Пусть по каналу передается сигнал $\alpha_1, \alpha_2, \cdots, \alpha_n$. В результате одиночной несимметрической ошибки он исказился и на приемном конце воспринимается в виде $\alpha_1', \alpha_2', \cdots, \alpha_n'$ (для определенности считаем, что имеют место „большие“ ошибки типа $(+)$).

Подсчитывается наименьшее неотрицательное решение сравнения.

$$t \equiv -a + \sum_{i=1}^{n} p_i \alpha_i' \pmod{(q-1)\, p_n + 1}.$$

В случае $t = 0$ ошибок нет, при $t \neq 0$ номер k искаженной позиции определяется из условия

$$t = p_k \cdot z_k,$$

а искажение в символе $l = \frac{t}{p_k}$ *).

При этом k и l определяются однозначно в силу выбора последовательности p_n. Тем самым теорема доказана.

Замечание 1. Для оценки мощности кодов, получаемых с помощью сравнения (10), необходимо исследовать поведение последовательности p_n, используемой в (10).

В случае основания кода $q = 3$ имеет место следующее соотношение между числом членов последовательности n и n-ым членом последовательности p_n:

$$n = \sum_{v=0}^{[\log_4 p_n]} \left[\frac{p_n}{4^v} + 1 \right] \quad **),$$

откуда имеем $p_n = \frac{3}{2} n + 0(\ln n)$.

Замечание 2. При $q = 2$ код, рассмотренный выше, является известным кодом, исправляющим одиночные несимметричные ошибки. В этом случае номер искаженной позиции t определяется как наименьшее неотрицательное решение сравнения

$$t = \pm W' \pmod{n+1} \quad ***) \qquad (11)$$

($a = 0$; $W' = \sum_{i=1}^{n} i\alpha_i'$, где $\alpha_1', \alpha_2', \cdots, \alpha_n'$ — принятый сигнал).

Указанный двоичный код с исправлением несимметрических ошибок можно видоизменить так, чтобы сделать его систематическим. Тогда кодовые последовательности будут иметь вид:
$\alpha_1, \alpha_2, \cdots, \alpha_k, \alpha_{k+1}, \alpha_{k+2}, \cdots, \alpha_n$, где $\alpha_1, \alpha_2, \cdots, \alpha_k$ — информационные символы, $\alpha_{k+1} = \sum_{i=1}^{k} \alpha_i$, $\alpha_{k+2}, \cdots, \alpha_n$ — проверочные символы, представляющие собой двоичную запись в виде последовательности длины $]\log_2 k[$ наименьшего неотрицательного вычета t выражения $\sum_{i=1}^{k} i\alpha_i$ по $\bmod k+1$, т. е. $t = \sum_{i=k+2}^{n} \alpha_i \cdot 2^{i-k-2}$. Символ α_{k+1} служит для локализации ошибки. По нему можно определить, находится ли ошибка в информационной или в проверочной части.

---

*) В случае „больших" искажений типа $(-)$ перед суммой в выражении для t и перед дробью в формуле для l берется знак $(-)$.

**) Функция $[x]$ — наибольшее целое число, не превосходящее x.

***) В случае ошибок вида $0 \to 1$ левая часть (11) берется со знаком $(+)$, а в случае ошибок вида $1 \to 0$ — со знаком $(-)$.

Опишем алгоритм декодирования. Определяется сумма $\sum_{i=1}^{k} \alpha_i'$ принятого сообщения $\alpha_1', \alpha_2', \cdots, \alpha_n'$. Если $\alpha_{k+1}' = \sum_{i=1}^{k} \alpha_i'$, то ошибки либо нет, либо она в проверочной части. Если же $\alpha_{k+1}' \neq \sum_{i=1}^{k} \alpha_i'$, то ошибка находится или в информационной части, или искажен символ α_{k+1}'. Тогда определяется

$$t \equiv \sum_{i=1}^{k} i\alpha_i' - \sum_{i=k+2}^{n} \alpha_i' \, 2^{i-k-2} \text{ *)} \pmod{k+1}.$$

Если $t = 0$, то исказился символ α_{k+1}', и первые k символов выдаются на выходе декодирующего устройства. Если же $t \neq 0$, то символ α_t' надо заменить на противоположный.

§ 3. Исследование одной числовой функции; класс двоичных кодов, исправляющих одиночные симметрические ошибки

Из предыдущего изложения видно, как важно знать, при каком a сравнение (1) имеет максимальное число решений. Задача о выборе оптимального a для случая, когда $\beta_i = i$; $\gamma(\alpha_i) = \alpha_i = 0, 1, \cdots, q-1$ полностью была решена Р. Р. Варшамовым [5], а Б. Р. Гинзбург [6] нашел точную формулу числа решений сравнения (1) указанного частного случая в точке максимума.

В данном параграфе исследуется (определяется максимальное число решений) частный случай сравнения (1), а также рассматривается связанный с ним двоичный код, исправляющий одиночные симметрические ошибки.

Рассмотрим сравнение вида

$$\sum_{i=1}^{n} i\alpha_i \equiv a \pmod{qn}, \tag{12}$$

где $\alpha_i = 0, 1, 2, \cdots, q-1$; q — произвольное целое.

Имеет место

Теорема 3. Сравнение (12) имеет максимальное число решений при значениях $a \equiv 0 \pmod{n}$, причем максимальное число решений

$$M_{n,\,qn}^{0,\,q} = \frac{1}{n} \sum_{\substack{(u,\,q)=1 \\ u \mid n}} q^{\frac{n}{u}-1} \cdot \varphi(u),$$

где $\varphi(u)$ — функция Эйлера

*) В случае ошибок типа $0 \to 1$ выражение для t берется со знаком $(+)$, в противном случае $(1 \to 0)$ — со знаком $(-)$.

Доказательство. Обозначим через $M_{n,\,qn}^{a,\,q}$ — число решений сравнения (12), $S(k,\,j\,n)$ — число способов представления числа k в виде суммы целых неотрицательных слагаемых вида iz_i ($i = 1,\,2,\,\cdots,\,n-1$; $z_i = 0,\,1,\,2,\,\cdots,\,q-1$) и $j\cdot n$. Имеем

$$M_{n,\,qn}^{a,\,q} = \sum_{l=0}^{\left[\frac{(n+1)(q-1)}{2}\right]} \sum_{j=0}^{q-1} S(a + iqn,\ jn). \tag{13}$$

Используя равенство (13), легко показать, что

$$M_{n,\,qn}^{a,\,q} = M_{n,\,qn}^{b,\,q}, \quad \text{где} \quad b \equiv a \pmod{n},$$

а потом можно полагать $a < n$.

Принимая во внимание (13) и тождество $S(a + i\,q\,n,\,j\,n) = S(a + i\,(q \pm l)\,n,\,(j \pm il)\cdot n)$, нетрудно получить следующее равенство:

$$q \cdot M_{n,\,qn}^{a,\,q} = M_{n,\,n}^{a,\,q},$$

где $M_{n,\,n}^{a,\,q}$ — число решений сравнения

$$\sum_{i=1}^{n} iz_i \equiv a \pmod{n} \qquad (z_i = 0,\ 1,\ 2,\ \cdots,\ q-1).$$

Заметим, что

$$M_{n,\,n}^{a,\,q} = q \cdot M_{n-1,\,n}^{a,\,q}, \tag{14}$$

где $M_{n-1,\,n}^{a,\,q}$ — число решений сравнения вида:

$$\sum_{i=1}^{n-1} iz_i \equiv a \pmod{n} \qquad (z_i = 0,\ 1,\ 2,\ \cdots,\ q-1).$$

Согласно [6]: $M_{n-1,\,n}^{a,\,q}$ — имеет максимум при $a = 0$, причем

$$M_{n-1,\,n}^{0,\,q} = \frac{1}{n} \sum_{\substack{(u,\,q)=1 \\ u/n}} \varphi(u) \cdot q^{\frac{n}{u}-1}, \tag{15}$$

где $\varphi(u)$ — функция Эйлера.

Из (14) и (15) следует утверждение теоремы.

Теорема 3. Множество решений сравнения

$$\sum_{i=1}^{n} iz_i \equiv a \pmod{2n}, \tag{16}$$

где $z_i = 0,\ 1$, образует код, исправляющий одиночные симметрические ошибки. Причем максимальное число решений и, следовательно, оптимальный в данном классе кодов код будет при значениях $a = 0$ и $a = n$. Число элементов кода при $a = 0$ дается формулой

$$M_{n,\,2n}^{0,\,2} = \frac{1}{n} \sum_{\substack{(u,\,2)=1 \\ u/n}} \varphi(u) \cdot 2^{\frac{n}{u}-1}.$$

Доказательство. Действительно, множество решений сравнения (16) образует код, исправляющий одиночные симметрические ошибки. Номер t позиции, в которой произошла ошибка, определяется как $t = \min(t_1, t_2)$, где t_1 — наименьший неотрицательный вычет выражения $W' = \sum_{i=1}^{n} i z_i'$ — соответственно выражения — W' по $\bmod 2n$.

Последние утверждения теоремы 3 автоматически следуют из теоремы 1 при $q = 2$.

Рассматриваемый выше код может быть записан в систематическом виде, так же как и код, исправляющий одиночные несимметрические ошибки. Отличие заключается лишь в том, что вычет t берется по $\bmod 2k$ и проверочная часть имеет длину $]\log_2 2k[$. При декодировании t находится как $\min(t_1, t_2)$, где t_1, t_2 — наименьшие неотрицательные вычеты соответствующих выражений по $\bmod 2k$.

§ 4. Системы кодирования с исправлением одиночных несимметрических (симметрических) ошибок или ошибок типа вставок (выпадений)

Хотя систематические коды с исправлением одиночных симметрических (несимметрических) ошибок имеют несколько большую избыточность, чем коды Хэмминга*), использование их представляет практический интерес. Так, например, их можно успешно применить в устройствах, корректирующих одиночные ошибки типа вставок (выпадений) или симметрических (несимметрических) ошибок. Это связано с тем, что большая часть оборудования для систем кодирования с исправлением одиночных вставок (выпадений) и симметрических (несимметрических) ошибок — общая.

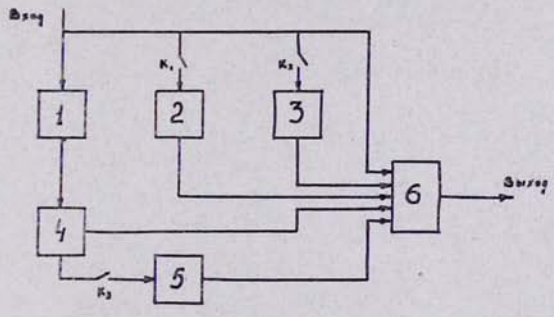

Рис. 3.

На рис. 3 приведена блок-схема кодирующего устройства системы с исправлением одиночных вставок (выпадений) или симметрических (несимметрических) ошибок, а на рис. 4 — блок-схема декодирующего устройства этой системы.

Процесс кодирования осуществляется следующим образом. Информационная последовательность $\alpha_1, \alpha_2, \cdots, \alpha_k$ поступает на блок опре-

*) Число проверочных символов в коде с исправлением одиночных симметрических (соответственно несимметрических) ошибок на два (соответственно на один) больше, чем в коде Хемминга.

деления вычета 1, определяющий наименьший неотрицательный вычет t выражения $\sum_{i=1}^{k} i z_i$ по $\bmod m$, где $m = k+1$ в случае сбоев типа вставок (выпадений) или несимметрических ошибок и $m = 2k$ в случае симметрических ошибок. Одновременно информационные символы поступают на блок 2 или 3 в зависимости от типа ошибок и на блок задержек 6. Вычет t поступает на нормирующий блок 4, который пропускает через себя $\lceil \log_2 m \rceil$ символов, и одновременно в случае ошибок типа вставок (выпадений) через замкнутый ключ K_3 запускает разделительный блок 5, вырабатывающий комбинацию „10". При сбоях типа вставок (выпадений), кроме ключа K_3, в замкнутом положении находится ключ K_1, а ключ K_2 разомкнут, и блок 2 вырабатывает символы $\bar{z}_k$, $\bar{z}_k$, поступающие на блок 6. В случае симметрических (несимметрических) ошибок наоборот, ключи K_1 и K_3 разомкнуты, а ключ K_2 — замкнут, и блок 3 вырабатывает символ проверки на четность $z_{k+1} = \sum_{i=1}^{k} z_i$, поступающий на блок 6.

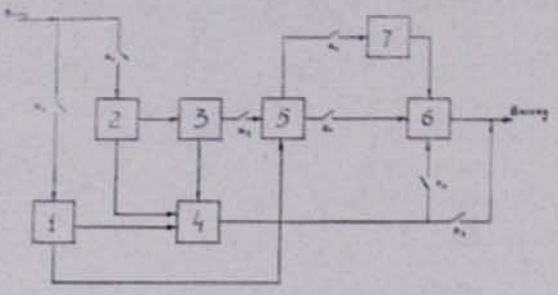

Рис. 4.

Декодирование производится следующим образом.

В случае сбоев типа вставок (выпадений) ключ K_2 — замкнут, а ключ K_1 — разомкнут, и информация из канала связи поступает на блок 2, анализирующий $n+3$-й символ в случае сбоев типа выпадений и $n+4$-й символ в случае вставок. Блок 2, в зависимости от результатов анализа, включает или блок 4, выделяющий неискаженные информационные символы с выдачей на выход, или блок 3, сравнивающий k-й и $k+1$-й символы, если имеет место выпадение, или $k+1$-й и $k+2$-й символы в случае вставки. Если результат анализа показывает, что ошибок в информационной части сообщения нет, то ключ K_3 разомкнут, и информационные символы с блока 4 через замкнутый ключ K_4 поступают на выход. Если в информационной части есть ошибка, то ключ K_3 замкнут, а K_4 — разомкнут, и блок 5 определяет соответствующий вычет t и выдает его через замкнутый ключ K_6 на блок 6, куда также через ключ K_5 с блока 4 поступают искаженные информационные символы. Блок 6 корректирует

ошибку и выдает на выход исправленную информационную последовательность.

В случае симметрических (несимметрических) ошибок информационные символы через замкнутый ключ K_7 поступают на блок 1, который подсчитывает $\sum_{i=1}^{k} x_i$ и сравнивает результат с x_{k+1}. В зависимости от результата сравнения последовательность поступает либо на блок 4, отсчитывающий информационные символы с последующей выдачей их на выход, либо на блок 5, определяющий вычет t по соответствующему модулю. Если $t = 0$, то блок 4 через замкнутый ключ K_4 выдает на выход информационную последовательность. Если $t \neq 0$, то замыкаются ключи K_5 и K_6 или K_7 (в случае несимметрической ошибки замкнут ключ K_6, а в случае симметрической — K_7). Заметим, что при симметрической ошибке определяются два вычета — t_1 и t_2, и блок 7 находит $\min(t_1, t_2)$. Блок 6 корректирует ошибку и выдает на выход исправленную последовательность информационных символов

Վ. Ա. ԱՌԱՔԵԼՈՎ, Գ. Մ. ՏԵՆԵՆՀՈԼՑ

ՈՒՂՂԻՉ ԿՈԴԵՐԻ ՄԻ ՔԱՆԻ ԴԱՍԵՐ

Ա մ փ ո փ ո ւ մ

Դիտարկվում են կամայական հիմունքով այնպիսի կոդերի դասեր, որոնք ուղղում են սիմվոլների ներդրման կամ վայրանկման տիպի առանձին խանգարումները։ Երկակի դեպքում կառուցվում են սիստեմատիկ կոդեր։ Բերվում է կոդավորման սիստեմ՝ ինչպես առանձին անհամաչափ կամ համաչափ սխալների, այնպես էլ ներդրման կամ վայրանկման տիպի սխալների ուղղումով։

ЛИТЕРАТУРА

1. *Р. Р. Варшамов, Г. М. Тененгольц.* Код, исправляющий одиночные несимметрические ошибки. Автоматика и телемеханика, т. XXVI, № 2, 1965, стр. 288—292.
2. *В. И. Левенштейн.* Двоичные коды с исправлением выпадений, вставок и замещений. ДАН СССР, т. 163, № 4, 1965, стр. 845—848.
3. *J. Singer.* A theorem in finite projective geometry and some applications to number theory. Trans. of the American math. society. vol. 43, 1938, pp. 377—385.
4. *Г. М. Тененгольц.* Об одном классе кодов для несимметрического бинарного канала. Сб. „Кибернетика“ (Мир глазами молодого ученого). Изд-во „Наука“, М., 1967, стр. 120—129.
5. *Р. Р. Варшамов.* Об одной арифметической функции, имеющей приложение в теории кодирования. ДАН СССР, т. 161, № 3, 1965, стр. 540—543.
6. *Б. Р. Гинзбург.* Определение числа элементов в коде Варшамова—Тененгольца, исправляющем одну асимметричную ошибку. Труды второй Всесоюзной конференции по теории кодирования и ее приложениям, М., 1965, секция 1, часть 1; стр. 25—27.