I. ВВЕДЕНИЕ. ОСНОВНЫЕ ПОНЯТИЯ И ТЕРМИНЫ ТЕОРИИ КОДИРОВАНИЯ

В данном разделе книги приведено изложение некоторых основополагающих вопросов теории кодирования. В частности, здесь определён ряд новых понятий и рассматриваются вопросы:

понятие корректирующего кода;

потенциальные обнаруживающая и исправляющая способности корректирующего кода;

основные способы обнаружения и исправления ошибок, лежащие в основе создания корректирующих кодов;

комбинация и её вес, хэммингово расстояние между двумя комбинациями, кодовое расстояние;

вектор ошибки и назначение ошибок к исправлению;

корректирующая способность кода и кодовое расстояние;

общие соображения о сложности кодера и декодера, реализующих рассматриваемый алгоритм декодирования принятых комбинаций;

определён ряд основных корректирующих кодов:

блоковые коды:

коды с многократной передачей символов,

коды с одной проверкой на чётность,

коды с разбиением множества запрещённых комбинаций на попарно не пересекающиеся подмножества,

коды с декодированием по методу максимального правдоподобия,

коды со стиранием символов,

коды с проверкой на чётность – групповые (n, k)-коды;

не блоковые коды – свёрточные (непрерывные) корректирующие коды;

введены следующие сравнительные характеристики корректирующих кодов: вероятность ошибки на сообщение P_e, скорость передачи данных R, простота реализации кодера и декодера; для блоковых кодов введены также характеристики: длина кода n, кодовое расстояние $d_{\min}$, число кодовых комбинаций N и другие характеристики;

связь между расстоянием точки, изображающей сигнал, от начала координат в метрическом $2WT$-мерном пространстве и энергией сигнала;

приведены некоторые необходимые справочные данные.

Для оценки эффективности использования корректирующего кода необходимо знать статистические свойства шумов, действующих в канале, или, иначе говоря, знать модель используемого канала. Моделью, с которой мы чаще всего будем иметь дело, является СБПК и её подвид – модель ДСК [1.8].

1.1. НАЧАЛЬНЫЕ СВЕДЕНИЯ О БЛОКОВЫХ КОРРЕКТИРУЮЩИХ КОДАХ

Ранее мы ознакомились с некоторыми начальными сведениями об экономичных кодах и о блоковых корректирующих кодах, [1.8, с. 226, 434], а также рассмотрели постановку задачи о *блоковом* кодировании и декодировании бесконечной последовательности информационных двоичных символов при передаче данных по каналу с шумами, [1.8, с. 419 – 428].

Теперь мы продолжим изложение начальных сведений о блоковых корректирующих кодах.

Число символов алфавита A, используемого для образования кода, L ($L \geq 2$) называется *основанием* кода.

ПРИМЕР 1.1. Код с основанием 2 называется *двоичным* кодом, с основанием 3 – *троичным*, …, с основанием L называется *L-ичным* кодом.

END

В настоящем пособии мы будем преимущественно рассматривать двоичные коды.

Число элементов *последовательности символов* называется её *длиной*. Длина бесконечной последовательности равна бесконечности; длина конечной последовательности выражается натуральным числом.

Безызбыточные комбинации информационных символов одной и той же длины $n_0 = k$, появляющиеся на выходе кодера источника сообщений, иногда называют «сообщениями»; *безызбыточные комбинации* поступают на вход кодера канала, [1.8, с. 419 – 428]. *Число сообщений*

$$N = 2^k. \qquad (1\text{-}1)$$

Сопоставляемые взаимно однозначным образом кодером канала этим сообщениям *комбинации* длины $n > k$, поступающие с выхода кодера канала на вход канала с шумами, называются *кодовыми комбинациями* (а также *кодовыми последовательностями, кодовыми словами* или *кодовыми векторами*).

Избыточностью кодовых комбинаций или *избыточностью* блокового кода называют величину $r = n - k \geq 0$.

Число *всевозможных* (не обязательно кодовых) комбинаций символов алфавита A одной и той же длины n

$$N_0 = L^n. \qquad (1\text{-}1а)$$

Комбинация, набор или последовательность символов алфавита A (из числа всевозможных комбинаций заданной длины n) записываются обычно в виде:

$$\boldsymbol{a}_i = a_{i1}, a_{i2}, \ldots, a_{ij}, \ldots, a_{in}, \qquad (1\text{-}2)$$

или

$$\boldsymbol{a}_i = (a_{i1}, a_{i2}, \ldots, a_{ij}, \ldots, a_{in}), \qquad (1\text{-}2а)$$

или

$$\boldsymbol{a}_i = [a_{i1}\ a_{i2} \ldots a_{ij} \ldots a_{in}]; \qquad (1\text{-}2б)$$

где i – номер последовательности, $i = 1, 2, \ldots, L^n$, j – номер «позиции», «проекции», «координаты», «компоненты» или «разряда» комбинации, характеризующий положение рассматриваемого символа в комбинации; при этом a_{ij} может быть любым символом алфавита A. Заметим, что при написании $\boldsymbol{a}_i$ согласно (1-2) или (1-2а) с использованием в качестве a_{ij} элементов конечного числового поля (например, элементов из множества {0, 1}) запятые обычно опускаются; в (1-2б) запятые отсутствуют, т.к. $\boldsymbol{a}_i$ представляется здесь матрицей-строкой.

Иногда, там, где не возникает недоразумений, индекс комбинации опускают и пишут один из вариантов выражений (1-2), (1-2а) и (1-2б) без индекса i; например,

$$\boldsymbol{a} = (a_1, a_2, \ldots, a_j, \ldots, a_n). \qquad (1\text{-}2в)$$

Далее, если не оговорено или не подразумевается противное, мы будем рассматривать двоичные ($L = 2$) коды. Компонентами кодовых комбинаций двоичного кода являются элементы множества {0, 1}.

Пусть B^n множество всевозможных последовательностей символов двоичного алфавита $A = \{0, 1\}$ одной и той же длины n:

$$B^n = \{(a_1, a_2, \ldots, a_i, \ldots, a_n) |\ a_i \in A,\ \ 1 \leq i \leq n\}.$$

Код, все кодовые последовательности которого имеют одну и ту же длину n, называется *равномерным* или *блоковым* n-значным кодом и, в противном случае, код называется *не-*

равномерным; часто n называют *длиной блокового кода* или *длиной кода*.

Процесс взаимно однозначного преобразования на основе некоторого заданного алгоритма поступающих на вход кодера канала с шумом *безызбыточных комбинаций* информационных символов постоянной длины $n_0 = k < n$ в появляющиеся на выходе кодера *кодовые комбинации* одной и той же длины n называется *кодированием*.

Процесс кодирования подлежащих передаче сообщений является детерминированным процессом в том смысле, что каждое сообщение отображается взаимно однозначным образом в одну, вполне определённую кодовую комбинацию. *В процессе кодирования информация не теряется*.

Следует отметить, что сложность реализации процесса кодирования сообщений, подлежащих передаче, определяется в первую очередь видом применяемого кода и способом его описания. Простейшим по его описанию, но не оптимальным в смысле сложности реализации процесса кодирования, способом задания кода является:

-перечисление совокупности всех N сообщений, подлежащих передаче,

-перечисление совокупности N кодовых комбинаций

и

-задание одного конкретного взаимно однозначного отображения между элементами этих множеств из совокупности $N!$ возможных отображений.

Процесс преобразования (возможно с ошибками) принятых комбинаций (не обязательно кодовых) в породившие их переданные кодовые комбинации в соответствии с некоторым заранее установленным правилом декодирования называется *декодированием*.

Процесс декодирования принятых комбинаций в переданные кодовые комбинации (в соответствии с некоторым правилом декодирования) является, вообще говоря, *вероят-*

ностным процессом в том смысле, что полученная на выходе канала комбинация может произойти из кодовых комбинаций некоторого соответствующего ей подмножества кодовых комбинаций; причём полученной комбинации и этому подмножеству кодовых комбинаций соответствует некоторое условное распределение вероятности. Именно поэтому процесс декодирования может быть охарактеризован введённой нами ранее (см. [1.8, с. 420 – 431]) средней по множеству сообщений вероятностью ошибки декодирования на блок (или на символ блока символов) P_e , зависящей от свойств канала и заданного правила декодирования. Очевидно, что *в процессе декодирования информация* в общем случае *может теряться*.

Устройства, осуществляющие процесс кодирования и процесс декодирования, называются *кодером* и *декодером* соответственно.

Применение блоковых кодов подразумевает использование *блокового* кодирования и декодирования [1.1, 1.3].

Блоковое кодирование в случае двоичных кодов состоит в том, что поступающая с выхода кодера источника последовательность *равновероятных и статистически независимых двоичных символов* делится кодером канала на блоки информационных символов одной и той же длины $n_0 = k$; таким образом, все передаваемые сообщения имеют одну и ту же вероятность 2^{-k} [1.8]. Кодер в соответствии с конкретным алгоритмом кодирования добавляет к блоку информационных символов длины k r контрольных или проверочных или избыточных символов ($r = n - k > 0$) и передаёт образованную таким способом кодовую последовательность символов постоянной длины n на вход канала.

Описанный блоковый кодер кодирует каждый поступающий на его вход блок информационных символов *независимо* от других блоков информационных символов, так что каждое новое кодовое слово на его выходе *не зависит* от пре-

дыдущих кодовых слов; это характерное свойство блоковых кодов.

В отличие от сказанного в случае *свёрточных* кодов каждый блок символов на выходе кодера *зависит*, вообще говоря, от предыдущих блоков [1.7].

Как уже говорилось [1.8], код Бодо является безызбыточным *двоичным*, *равномерным*, *пятизначным кодом* ($n = k = 5$, $L = 2$), для которого выполняются соотношения

$$N = L^k = L^n = N_0 = 2^5 = 32, \qquad (1\text{-}2г)$$

где N – число сообщений исходного ансамбля (в данном случае букв русского алфавита) или число всевозможных кодовых комбинаций *безызбыточного* (или *простейшего* или *примитивного*) кода, k – число информационных символов, N_0 – число всевозможных комбинаций (каждая из которых является кодовой) длины $n = k = 5$; $r = n - k = 0$.

Очевидно, что если N – общее число *кодовых комбинаций* L-ичного n-значного кода, то в общем случае имеет место соотношение

$$L^k = N \le N_0 = L^n. \qquad (1\text{-}2д)$$

Например, в случае двоичного ($L = 2$) безызбыточного блокового 5-значного кода Бодо согласно (1-2г) $N = N_0 = 32$.

Число проверочных (контрольных или избыточных) символов *произвольного* блокового кода

$$r = n - k \ge 0. \qquad (1\text{-}2е)$$

Так, в случае кода Бодо $r = 0$.

Если в канале действует шум, а мощность сигналов ограничена, то передаваемые по каналу кодовые комбинации могут искажаться (т.е. в канале с шумами в передаваемых кодовых комбинациях *могут возникать ошибки*) и, следовательно, принимаемые на выходе канала комбинации могут отличаться от переданных кодовых комбинаций; при этом предполагается, что переданная кодовая комбинация длины n порождает принятую не обязательно кодовую последовательность той же длины n. Для обеспечения помехоустойчивой передачи данных по каналу используют *корректирующие коды*; если эти коды являются блоковыми, то $r = n - k > 0$.

Практические ограничения на создание и использование кодеров и декодеров накладывает не пропускная способность канала, а сложность и стоимость декодера (обычно в первую очередь) и кодера.

Таким образом, вследствие действия шума в канале при передаче по нему кодовой комбинации длины n в отдельных её разрядах (позициях) возможно возникновение элементарных ошибок, заключающихся в замене переданных (верных) символов другими – неверными символами. Если подобные элементарные ошибки возникают в нескольких разрядах (позициях) одной и той же кодовой комбинации, то можно говорить о *кратности ошибки* q.

ОПРЕДЕЛЕНИЕ 1.1. *Кратностью ошибки q называется суммарное число искажённых в пределах одной и той же кодовой комбинации символов.*

END

Очевидно, что q всегда заключено в пределах:

$$0 \le q \le n, \qquad (1\text{-}3)$$

где n – длина кодовой комбинации. Ниже приведен пример вычисления вероятности q-кратной ошибки для *двоичного симметричного канала* (ДСК) [1.8], когда ошибки, возникающие в пределах одной и той же кодовой комбинации, являются *статистически независимыми ошибками*.

ПРИМЕР 1.2. Предположим, что для передачи данных по ДСК используется двоичный блоковый код длины n. ДСК ха-

рактеризуется *одной и той же вероятностью* искажения двоичного символа (нуля или единицы) шумом в нём p_0, $0 \le p_0 < 0.5$; согласно определению ДСК при этом имеет место симметрия: вероятность замены единицы нулём равна вероятности замены нуля единицей. Очевидно, что вероятность возникновения *некоторой* конкретной q-кратной ошибки в определённой кодовой комбинации равна вероятности возникновения ошибки в *некоторых* q конкретных разрядах этой кодовой комбинации p_0^q, умноженной на вероятность не возникновения при этом ошибок в остальных $(n - q)$ разрядах этой же самой кодовой комбинации $(1 - p_0)^{n-q}$

$$p_0^q (1 - p_0)^{n-q}, \qquad (1\text{-}4)$$

а вероятность возникновения *произвольной* q-кратной ошибки в рассматриваемой кодовой комбинации

$$p_q = C_n^q p_0^q (1 - p_0)^{n-q}, \qquad (1\text{-}4а)$$

где

$$C_n^q = n!/(q!(n - q)!). \qquad (1\text{-}5)$$

Легко убедиться в том, что функция p_q при $p_0 < 0.5$ с ростом q спадает, причём спадает тем быстрее, чем меньше p_0; откуда следует важный вывод: в случае *статистической независимости* элементарных ошибок при $p_0 < 0.5$ *более вероятными* являются ошибки *низших кратностей q*.

Распределение вероятности случайной величины q, определяемое соотношением (1-4а), характеризуется двумя параметрами p_0 и n; при этом распределение вероятности $\{p_q | q=0, 1, 2, \ldots n\}$, определяемое соотношением (1-4а), называется *биномиальным распределением*.

Заметим, что

$$\sum_{q=0}^{n} p_q = \sum_{q=0}^{n} C_n^q p_0^q (1 - p_0)^{n-q} = (p_0 + (1 - p_0))^n = 1. \qquad (1\text{-}6)$$

Таким образом, согласно сказанному, мы рассмотрели передачу данных по ДСК с помощью двоичного (равномерного) блокового кода длины n и при этом имели место условия:

-передаваемые по каналу сообщения являются *равновероятными*;

-используемый канал – ДСК – канал без памяти и, следовательно, элементарные ошибки, возникающие в различных разрядах кодовых комбинаций одной и той же длины n, являются *статистически независимыми*;

-вероятности замены единицы нулём и наоборот *равны одной и той же величине* p_0 (свойство ДСК).

При этом вероятность возникновения произвольной q-кратной ошибки в ДСК выражается соотношением (1-4а).

END

Далее мы всюду будем предполагать, что передаваемые сообщения являются *равновероятными*.

Для описания процесса искажения двоичных кодовых комбинаций $\boldsymbol{a}_i$, $i = 1, 2, \ldots, N$, шумом в канале введём понятие *вектора ошибки* $\boldsymbol{e}_j$, $j = 1, 2, \ldots, N_0$, $N_0 = 2^n > 2^k = N$, который будем записывать в виде двоичной комбинации символов той же длины n, что и длина кодовых комбинаций $\boldsymbol{a}_i$, причём нули в определённых разрядах $\boldsymbol{e}_j$ означают, что ошибки в аналогичных разрядах $\boldsymbol{a}_i$ не происходят; единицы же стоят в тех разрядах $\boldsymbol{e}_j$, которые соответствуют искажаемым разрядам $\boldsymbol{a}_i$. Таким образом, суммарное число единиц в векторе $\boldsymbol{e}_j$ равно кратности q ошибки, характеризуемой этим вектором.

Принятая кодовая комбинация $\boldsymbol{y}$ получается в результате поразрядного сложения $\boldsymbol{a}_i$ и $\boldsymbol{e}_j$ по модулю 2 (см. приложение 3.18.):

$$\boldsymbol{y} = \boldsymbol{a}_i \oplus \boldsymbol{e}_j. \qquad (1\text{-}7)$$

При этом заметим, что в данной книге выражение

$$y = a_i + e_j \pmod 2 \qquad (1\text{-}7а)$$

равносильно выражению (1-7); более того, далее там, где это не вызывает недоразумений, вместо (1-7а) мы будем писать

$$y = a_i + e_j\,. \qquad (1\text{-}7б)$$

ПРИМЕР 1.3.

а)

$a_i = 10101$

$e_j = 00000$ — e_j – нулевой вектор;

$a_i \oplus e_j = 10101$ — ошибки во всех разрядах a_i отсутствуют.

б)

$a_i = 10101$

$e_j = 10101$ — $e_j = a_i$; поэтому $a_i \oplus e_j$ – нулевой вектор.

$a_i \oplus e_j = 00000$

в)

$a_i = 10101$

$e_j = 11111$ — $e_j = 11111$; поэтому во всех пяти разрядах a_i имеют место ошибки.

$a_i \oplus e_j = 01010$

END

1.2. КОРРЕКТИРУЮЩИЕ КОДЫ И ИХ ОСНОВНЫЕ ВИДЫ

В [1.8] отмечалось, что основными факторами, определяющими рассеяние передаваемой по каналу информации, в самом общем случае являются «дефицит алфавита» и шум.

Очевидно, что при проектировании системы передачи сообщений следует учитывать необходимость исключения «дефицита алфавита»; требование исключения «дефицита

алфавита» в системе передачи сообщений является, как правило, разрешимой технической задачей. Однако возможны случаи, когда это требование является трудно выполнимым. Так, например, при разработке аналоговых дистанционных измерительных систем возможен случай, когда диапазон изменения измеряемой аналоговой величины составляет несколько порядков (вплоть до десяти порядков), а измерительный преобразователь имеет диапазон линейного преобразования не более двух порядков; в этом случае может возникать проблема рангирования или логарифмирования при удовлетворительно малых искажениях сигнала измерительной информации.

В дальнейшем мы будем считать, что в канале действует только один рассеивающий информацию фактор – шум.

Эффективным средством борьбы с ошибками, возникающими при передаче данных по каналу под воздействием шума, действующего в канале, является применение *корректирующих кодов*, позволяющих *обнаруживать* и *исправлять* ошибки (т.е. *корректировать* принятые комбинации, содержащие ошибки). Аналогичной цели служит применение корректирующих кодов для хранения информации в различных запоминающих средах.

Мы уже касались понятия «избыточность кода» или «избыточность источника» и говорили о том, что именно «… избыточность языка позволяет передавать без потерь информацию при наличии различного вида ошибок и сокращений в письменной и устной речи» [1.8, с. 317 – 319].

ОПРЕДЕЛЕНИЕ 1.2. *Избыточные коды, позволяющие* ***обнаруживать и исправлять*** *ошибки, возникающие при передаче кодовых комбинаций из-за наличия шума в канале, называются* ***корректирующими***.

END

Основная концепция *обнаружения* ошибок с помощью избыточных кодов состоит в том, что множество всевозмож-

ных комбинаций разбивается на два подмножества: подмножество *разрешённых* (передаваемых или кодовых) комбинаций и подмножество *запрещённых* (не передаваемых или не кодовых) комбинаций и при получении на выходе канала запрещённой комбинации принимается решение о наличии ошибки, превратившей переданную кодовую комбинацию в полученную запрещённую.

Собственно процедура обнаружения ошибки сводится к следующему:

некоторым способом, зависящим от специфических свойств конкретного избыточного кода, определяется, к какому из подмножеств – подмножеству разрешённых или подмножеству запрещённых комбинаций относится принятая комбинация;

если принятая комбинация является запрещённой, то принимается решение *о наличии ошибки* в принятой комбинации; в противном случае, принимается (возможно, ложное) решение *об отсутствии ошибки* в принятой комбинации; очевидно, что, если принятая комбинация является кодовой и не совпадающей с переданной кодовой комбинацией, то это последнее решение является *ошибочным*.

Известны, например, следующие способы обнаружения ошибки в принятой комбинации (т.е. способы отнесения принятой комбинации к подмножеству запрещённых комбинаций):

последовательное сравнение (сличение) принятой комбинации с элементами множества всех кодовых комбинаций или с элементами множества всех запрещённых комбинаций;

сравнение веса (вес – суммарное число единиц в комбинации; см. ниже) принятой комбинации с весами кодовых комбинаций (в случае кода с многократной передачей символов);

общая проверка на чётность;

проверка на постоянство веса (в случае равновесного кода);

и др…

ПРИМЕР 1.4. Обнаружение ошибок путём последовательного сличения принятой комбинации с элементами множества всех кодовых комбинаций.

Предположим, что имеется блоковый L-ичный n-значный код и, следовательно, имеются:

всевозможные комбинации *длины* n $\{\mathbf{s}_l\}$, $l = 1, 2, \ldots, N_0$; $N_0 = L^n$;

кодовые комбинации $\{\boldsymbol{a}_i\}$, $i = 1, 2, \ldots, N$; $N = L^k$, где k – *число информационных символов*, $k < n$;

запрещённые комбинации $\{\boldsymbol{b}_j\}$, $j = 1, 2, \ldots, N_0 - N$; $(N_0 - N) > 0$;

при этом *избыточность* $r = n - k > 0$.

Сделаем естественное предположение о том, что выполняется соотношение: $N << N_0$. При получении некоторой комбинации $\mathbf{s}$ производится её последовательное покомпонентное сличение с кодовыми комбинациями из множества $\{\mathbf{a}_i\}$, $i = 1, 2, \ldots, N$; если для какого-то номера $i=i^+$ имеет место покомпонентное равенство сличаемых комбинаций, то процесс сличения прекращается и принимается решение (возможно неправильное) об отсутствии ошибок; если же принятая комбинация не совпадает ни с одной кодовой комбинацией, то принимается решение о наличии в ней ошибки.

END

ПРИМЕР 1.5. Рассмотрим пример обнаружения ошибок путём сравнения веса принятой комбинации с весами кодовых комбинаций.

Если используется двоичный ($L = 2$) избыточный равномерный блоковый код с многократной передачей символов и n – длина комбинаций или длина кода, то при получении некоторой комбинации $\boldsymbol{s}$, в которой число единиц $n_{[1]}$ удовлетворяет неравенствам $0 < n_{[1]} < n$, принимается решение о на-

личии ошибки в комбинации ***s***. Как следует из ниже изложенного, вес комбинации ***s***

$$w(\boldsymbol{s}) = n_{[1]}. \qquad (1\text{-}8)$$

При этом имеют место следующие соотношения:

кодовые комбинации: $\{\boldsymbol{a}_i\}$, $i = 1, 2$;
запрещённые комбинации: $\{\boldsymbol{b}_j\}$, $j = 1, 2, \ldots, (N_0 - N)$;
$(N_0 - N) = (2^n - 2^k) = (2^n - 2^1) > 0$;
$k = 1$;
избыточность: $r = n - k = n - 1 > 0$.

END

ПРИМЕР 1.6. Рассмотрим код с общей или, как его ещё иногда называют, с одной проверкой на чётность.

Предположим, что избыточный блоковый код с одной проверкой на чётность характеризуется параметрами: $L = 2$, число информационных символов в кодовой комбинации $k = (n - 1)$, где n – длина кодовой комбинации, причём последний – *избыточный* символ a_n выбирается таким образом, чтобы он определялся всеми k *информационными* символами:

$$a_1 \oplus a_2 \oplus a_3 \oplus \ldots \oplus a_{n-1} \oplus a_n = 0, \qquad (1\text{-}9)$$

т.е. все кодовые комбинации имеют чётное число единиц; избыточность $r = n - k = 1$. Этот код позволяет обнаруживать одну элементарную ошибку (или нечётное число элементарных ошибок); однако чётное число элементарных ошибок не обнаруживается. При приёме комбинации с нечётным числом единиц принимается решение о наличии ошибки и, если это возможно, то делается запрос на повторную передачу.

END

ПРИМЕР 1.7. Обнаружение ошибок путём проверки комбинаций на постоянство веса.

Код с постоянным весом кодовых комбинаций – *равновесный* код – избыточный блоковый (равномерный) код, все кодовые комбинации которого имеют одну и ту же длину n и одинаковое число единиц $n_{[1]}$ (т.е. имеют одинаковый вес). Эти коды позволяют лишь обнаруживать ошибки любой кратности, при которых для некоторой принятой комбинации **s** имеет место соотношение

$$w(\boldsymbol{s}) \neq n_{[1]}. \qquad (1\text{-}10)$$

Однако ошибки сдвига, когда из-за действия шума в канале одна из единиц преобразуется в нуль, а один из нулей в единицу (или другие аналогичные ошибки, не изменяющие числа $n_{[1]}$) не обнаруживаются.

END

ПРИМЕР 1.8. Дан код 1000, 0100, 0010, 0001. Этот код имеет особенность – хеммингово расстояние (см. ниже) между любой парой его комбинаций равно двум. Коды с подобными свойствами называются *эквидистантными*. Этот же код является равновесным: вес каждой из его комбинаций равен единице.

END

Важным подвидом помехоустойчивой передачи данных является *передача с переспросом*, которая подразумевает использование *избыточного кода с обнаружением ошибок* и дополнительного *двустороннего канала*; при обнаружении ошибки в принятой комбинации или при наличии в ней искажённых (сомнительных) символов делается запрос на повторную передачу кодовой комбинации, породившей принятую комбинацию, содержащую обнаруженную ошибку [1.3, 1.4]. Достоинством передачи с переспросом является высокая скорость передачи, а при наличии в принятой комбинации сомнительных символов – возможность избежать опасных

ситуаций, если они возможны при последующем использовании этой принятой комбинации.

Исправление ошибок, как правило, является *более сложной задачей*, нежели их *обнаружение*, так как исправление ошибок, вообще говоря, предполагает их предварительное обнаружение.

Ниже с различной степенью подробности будут изложены начальные сведения о некоторых корректирующих кодах; читателю предлагается затем самостоятельно рассмотреть эти коды более детально с помощью цитируемых в пособии литературных источников. К кодам, о которых идёт речь, в первую очередь могут быть отнесены избыточные блоковые коды:

коды с многократной передачей символов;

коды с одной проверкой на чётность;

коды с разбиением множества запрещённых комбинаций на попарно не пересекающиеся подмножества; при этом предполагается существование взаимно однозначного отображения множества этих подмножеств в множество кодовых комбинаций;

коды с декодированием по методу максимального правдоподобия, когда принятая запрещённая комбинация декодируется в ближайшую (в смысле хэммингова расстояния) кодовую комбинацию;

блоковые корректирующие коды со стиранием символов;

групповые или линейные коды;

циклические коды.

Имеются также не блоковые коды – непрерывные коды (называемые также свёрточными, рекуррентными, конволюционными или цепными кодами [1.9, с. 244]).

1.3. БЛОКОВЫЕ КОРРЕКТИРУЮЩИЕ КОДЫ С МНОГОКРАТНОЙ ПЕРЕДАЧЕЙ СИМВОЛОВ

Предположим, что по каналу с шумом передаётся последовательность из нулей и единиц [1.8].

Простейшими избыточными блоковыми двоичными кодами являются *коды с повторением*: число информационных символов $k = 1$, длина кодовых комбинаций $n > k$, избыточность $r = n - k = n - 1$; единица кодируется последовательностью единиц длины n (единичная кодовая последовательность), нуль – последовательностью нулей длины n (нулевая кодовая последовательность).

Правило декодирования: в принимаемой последовательности подсчитывается число нулей $n_{[0]}$ и число единиц $n_{[1]}$ $(n_{[0]} + n_{[1]} = n)$;

если число нулей больше числа единиц, то принимается решение, что было передано кодовое слово, состоящее из нулей;

если число единиц больше числа нулей, то принимается решение, что была передана кодовая комбинация, состоящая из единиц;

если, при чётном n, число нулей равно числу единиц, то решение не принимается – т.е. имеет место «отказ от декодирования» или, иначе говоря, декодирование является *неполным*;

при нечётном n, отказ от декодирования невозможен и декодирование является *полным*.

При отказе от декодирования, если имеется возможность, организуется переспрос.

Полное декодирование приведёт к правильному результату в том случае, если шумом будет искажено менее половины символов кодовой комбинации и даст неверный результат – в противном случае. Если будет искажена половина символов кодовой комбинации (при чётном n), то будет иметь

место отказ от декодирования. Очевидно, что при редких ошибках и достаточно больших n вероятность ошибки при декодировании и вероятность отказа от декодирования будут малыми.

Очевидно, что рассматриваемому коду при больших n соответствует низкая скорость передачи данных

$$R = k/n = 1/n \text{ [бит/симв]}. \qquad (1\text{-}11)$$

Важно иметь в виду, что исполнение ошибочной команды может в некоторых случаях иметь катастрофические последствия и «отказ от декодирования с последующим переспросом» позволяет избежать их; поэтому алгоритм неполного декодирования в ряде случаев является более предпочтительным по сравнению с алгоритмом полного декодирования.

Код с многократной передачей символов является линейным или групповым (n, k)-кодом с кодовым расстоянием (см. ниже) $d_{\min} = n$.

ПРИМЕР 1.9. Пусть код с многократной передачей символов имеет параметры: $n = 5$; тогда $r = n - 1 = 4$. Так как n является нечётным числом, то будет использоваться *алгоритм полного декодирования*.

Если $0 \le n_{[0]} \le 2$ (и $3 \le n_{[1]} \le 5$), то принятая комбинация декодируется в переданную кодовую комбинацию 11111, в противном случае ($3 \le n_{[0]} \le 5$ и $0 \le n_{[1]} \le 2$) она декодируется в переданную кодовую комбинацию 00000.

Если речь идёт об использовании ДСК с вероятностью искажения нуля или единицы $p_0 \ll 1$, то вероятность ошибки декодирования

$$P_{ош} = \sum_{j=3}^{5} C_5^j p_0^j (1 - p_0)^{5-j}. \qquad (1\text{-}12)$$

Однако и для рассматриваемого кода можно модернизировать приведенный алгоритм *полного декодирования* до *алгоритма неполного декодирования*; вот этот алгоритм:

если $0 \le n_{[0]} \le 1$, то принятая комбинация декодируется в кодовую комбинацию 11111;

если $4 \le n_{[0]} \le 5$, то принятая комбинация декодируется в кодовую комбинацию 00000;

если $2 \le n_{[0]} \le 3$, то решение не принимается.

END

1.4. БЛОКОВЫЕ КОРРЕКТИРУЮЩИЕ КОДЫ С ОДНОЙ ПРОВЕРКОЙ НА ЧЁТНОСТЬ

В примере 1.10 описано образование кода с одной проверкой на чётность. Очевидно, что при больших n рассматриваемому коду соответствует высокая скорость передачи:

$$R = k/n = (n - 1)/n = (1 - 1/n) \text{ [бит/симв]}. \qquad (1\text{-}13)$$

Рассматриваемый код является линейным групповым (n, k)-кодом с кодовым расстоянием $d_{\min} = 2$ (как доказать это строго?).

ПРИМЕР 1.10. Предположим, что код с одной проверкой на чётность имеет длину кодовых комбинаций $n = 5$; следовательно, $k = n - 1 = 4$. Код используется для передачи сообщений по ДСК с параметром $p_0 = 0.01$ (вероятность искажения одиночного символа шумом в канале).

Характеристики рассматриваемого кода:

число кодовых комбинаций $N = 2^k = 16$;

число всевозможных комбинаций $N_0 = 2^n = 32$;

скорость передачи информации $R = 0.8$ [бит/симв];

кодовое расстояние $d_{\min} = 2$;

согласно (1-4) вероятность обнаружения однократной ошибки

$$p_1 = C_5^1 \cdot p_0 \cdot (1 - p_0)^4 = 0.048. \qquad (1\text{-}14)$$

END

ЗАМЕЧАНИЕ 1.1. В табл. 1.1. приведены характеристики кода с многократной передачей символов и кода с одной проверкой на чётность. Оба кода имеют одну и ту же длину n.

Таблица 1.1

Характеристики корректирующих кодов	Блоковый код с многократной передачей символов	Блоковый код с одной проверкой на чётность
Вид кода	Групповой $(n, 1)$-код	Групповой $(n, n-1)$-код
Длина кода	n	n
Число информационных символов k	1	$(n-1)$
Число всевозможных комбинаций N_0	2^n	2^n
Число кодовых комбинаций $N=2^k$	2	2^{n-1}
N/N_0	$1/2^{n-1}$	1/2
Скорость $R = k/n$	$1/n$	$(1 - 1/n)$
Кодовое расстояние d_{min}	n	2
Порождающая матрица $\boldsymbol{G}_{(n, k)}$	Матрица-строка размерности $1\times n$ [1 1 1 … 1]	Прямоугольная матрица размерности $(n-1)\times n$; например: $\begin{bmatrix} 1&0&0&\dots&0&1 \\ 0&1&0&\dots&0&1 \\ 0&0&1&\dots&0&1 \\ \multicolumn{6}{c}{\dots\dots\dots} \\ 0&0&0&\dots&1&1 \end{bmatrix}$
Проверочная матрица $\boldsymbol{H}_{(n, k)}$	Прямоугольная матрица размерности $(n-1)\times n$; например: $\begin{bmatrix} 1&1&0&0&\dots&0 \\ 1&0&1&0&\dots&0 \\ 1&0&0&1&\dots&0 \\ \multicolumn{6}{c}{\dots\dots\dots} \\ 1&0&0&0&\dots&1 \end{bmatrix}$	Матрица-строка размерности $1\times n$ [1 1 1 … 1]

Следует подчеркнуть, что код с многократной передачей символов и код с одной проверкой на чётность *при одной и той же заданной длине кодовых комбинаций n* имеют ряд

практически противоположных свойств (табл. 1.1). При первом чтении книги две последних строки табл. 1.1 можно не рассматривать.

END

1.5. БЛОКОВЫЕ КОРРЕКТИРУЮЩИЕ КОДЫ С РАЗБИЕНИЕМ МНОЖЕСТВА ЗАПРЕЩЁННЫХ КОМБИНАЦИЙ НА ПОПАРНО НЕ ПЕРЕСЕКАЮЩИЕСЯ ПОДМНОЖЕСТВА

Если под действием шума каждая поступающая на вход канала кодовая комбинация $\boldsymbol{a}_i$, $i = 1, 2, \dots, N$, преобразуется *только в комбинации из соответствующего ей подмножества запрещённых комбинаций* M_i и если при этом $\{M_i\}$, $i= 1, 2, \dots, N$, является разбиением множества всевозможных запрещённых комбинаций $\{\boldsymbol{b}_j\}$, то потери информации в таком канале отсутствуют [1.8]. В связи с этим возникает вопрос: всегда ли для некоторого конкретного канала существует разбиение $\{M_i\}$, $i= 1, 2, \dots, N$, обеспечивающее отсутствие потерь информации при передаче сообщений по каналу?

Очевидно, что если $\{M_i\}$, $i= 1, 2, \dots, N$, не является разбиением множества всевозможных запрещённых комбинаций $\{\boldsymbol{b}_j\}$, в котором, например, M_i и M_j – пересекающиеся подмножества, то при декодировании принятой на выходе канала запрещённой комбинации, принадлежащей одновременно M_i и M_j, возникнет *неопределённость*.

Ниже мы рассмотрим потенциальные обнаруживающую и исправляющую способности корректирующего кода допуская, что существует некоторое разбиение $\{M_i\}$, $i= 1, 2, \dots, N$, множества $\{\boldsymbol{b}_j\}$; при этом для некоторого канала найденные результаты реализуются, если и на самом деле передаваемая по этому каналу кодовая комбинация $\boldsymbol{a}_i$, $i = 1, 2, \dots, N$, преобразуется под действием шума только в комбинации из соответствующего ей подмножества запрещённых комби-

наций M_i; если же это не так, то найденные потенциальные обнаруживающая и исправляющая способности корректирующего кода, вообще говоря, не реализуются.

Итак, предположим, что X^n и Y^n – множества всевозможных допустимых входных и выходных последовательностей символов длины n на входе и выходе канала соответственно, а $\boldsymbol{x}$ и $\boldsymbol{y}$ – некоторые произвольные последовательности символов из этих множеств соответственно.

Пусть $\{\boldsymbol{a}_i\}$, i= 1, 2, …, N, – множество разрешённых (кодовых) комбинаций, $\{\boldsymbol{b}_j\}$, j= 1, 2, …, N_0 - N, – множество запрещённых комбинаций; $\{M_i\}$, i= 1, 2, …, N, – разбиение $\{\boldsymbol{b}_j\}$ на N (попарно непересекающихся) подмножеств, объединение которых исчерпывает всё множество запрещённых комбинаций.

Возможность обнаружения ошибок при приёме комбинаций на выходе канала обеспечивается тем, что в качестве кодовых комбинаций используются не все $N_0 = L^n$ различные всевозможные комбинации длины n, а лишь их часть $(N/N_0) < 1$.

Как уже говорилось, фактически используемые для кодирования передаваемых сообщений N комбинаций из числа N_0 всевозможных комбинаций длины n называются *разрешёнными* или *кодовыми комбинациями*, а остальные $(N_0 - N)$ – *запрещёнными*.

Естественно, что возникает вопрос о том, какие из q-кратных ошибок, q = 1, 2, …, n, являются *обнаруживаемыми* и какие не являются.

Если при передаче данных по каналу с помощью блочного L-ичного n-значного кода некоторая разрешённая комбинация превратится из-за некоторой возникшей ошибки в какую-то из запрещённых комбинаций, то, очевидно, что ошибка, обусловившая такое превращение, является *обнаруживаемой ошибкой* (но не обязательно исправляемой).

Поэтому любой код, удовлетворяющий единственному условию: $N < N_0$, позволяет обнаруживать ошибки в $N(N_0 - N)$ случаях переходов N разрешённых комбинаций в $(N_0 - N)$ запрещённых комбинаций при общем числе NN_0 возможных переходов.

Следовательно, при

$$N/N_0 \longrightarrow 0 \qquad (1\text{-}15)$$

число переходов, обусловленных обнаруживаемыми ошибками, составляют от общего числа всевозможных переходов часть, выражаемую соотношением

$$N(N_0 - N)/NN_0 = (1 - N/N_0) \xrightarrow[N/N_0 \longrightarrow 0]{} 1. \qquad (1\text{-}16)$$

Так как N обычно фиксировано (как правило, задан ансамбль сообщений, подлежащих передаче), то условие $N/N_0 \to 0$ реализуется за счёт реализации условия $N_0 \to \infty$ (т.е. за счёт реализации условия $n \to \infty$), что, вообще говоря, сопровождается уменьшением скорости передачи информации.

Заметим, однако, что если из-за наличия ошибки некоторая разрешённая (кодовая) комбинация перейдёт в другую, также разрешённую комбинацию, то ошибка, обусловившая такой переход, является уже *не обнаруживаемой*.

Исправление ошибок сводится к следующему. Разобьём множество запрещённых комбинаций $\{\boldsymbol{b}_j\}$, j=1, 2, …, $(N_0 - N)$, (см. рис. 1.1) на N попарно не пересекающихся подмножеств $\{M_i\}$, i=1, 2, …, N:

$$\{\boldsymbol{b}_j\} = \bigcup_{j=1}^{N_0-N} \boldsymbol{b}_j = \bigcup_{i=1}^{N} M_i\,, \qquad (1\text{-}17)$$

$$M_j \bigcap_{i,j=1}^{N} M_i = \begin{cases} M_i, & i = j, \\ \varnothing, & i \neq j, \end{cases} \qquad (1\text{-}18)$$

где Ø – пустое множество.

Установим взаимно однозначное соответствие

$$M_i \Leftrightarrow \boldsymbol{a}_i, \quad i = 1, 2, \ldots, N, \qquad (1\text{-}19)$$

где $\boldsymbol{a}_i$ – разрешённая (кодовая) комбинация.

Условимся, что в случае приёма любой запрещённой комбинации

$$\boldsymbol{b}_j \in M_i \qquad (1\text{-}20)$$

принимается решение:

«была передана кодовая комбинация $\boldsymbol{a}_i$». (1-21)

При таком способе приёма ошибка в передаваемых по каналу кодовых комбинациях действительно исправляется, но только в том случае, если и на самом деле при выполнении (1-20) была передана кодовая комбинация $\boldsymbol{a}_i$; если же при выполнении (1-20) кодовая комбинация $\boldsymbol{a}_i$ не была передана, то ошибка не исправляется, а только определяется.

Очевидно, что число различных переходов входных последовательностей канала в выходные, вызванных исправляемыми ошибками, равно $(N_0 - N)$ – т.е. суммарному числу запрещённых кодовых комбинаций, которое равно суммарному числу элементов $\bigcup_i M_i$.

Число всевозможных переходов, вызванных исправляемыми ошибками, $(N_0 - N)$ составляет от числа всевозможных переходов входных последовательностей канала в выходные, вызванных обнаруживаемыми ошибками, $N(N_0 - N)$ часть

$$Q_p = \begin{cases} (N_0 - N)/(N(N_0 - N)) = 1/N < 1, & \text{при } N < N_0; \\ \text{«0/0» - величина не определена,} & \text{при } N = N_0. \end{cases} \qquad (1\text{-}22)$$

Величина Q_p может быть названа *потенциальной исправляющей способностью корректирующего кода* [1.9]. Эффективность использования потенциальной исправляющей способности кода Q_p зависит от конкретного способа разбиения $\{\boldsymbol{b}_j\}$ на подмножества M_i, $i = 1, 2, \ldots, N$, и от статистики ошибок, имеющих место в канале.

ЗАМЕЧАНИЕ 1.2. Очевидно, что при $N = N_0$ любая ошибка является *не обнаруживаемой* и *не исправляемой*; и, следовательно, такой *код не является корректирующим*.

Поэтому условие

$$N < N_0 \qquad (1\text{-}23)$$

является *необходимым условием* того, чтобы код являлся *корректирующим*.

END

Число всевозможных переходов, вызванных исправляемыми ошибками, составляет от числа всевозможных переходов входных последовательностей канала в выходные, часть

$$(N_0 - N)/(NN_0) = 1/N - 1/N_0 \xrightarrow[n \to \infty]{} 1/N = Q_p, \quad \text{при } N \neq N_0. \qquad (1\text{-}24)$$

Очевидно, что все исправляемые ошибки являются обнаруживаемыми, так как процесс исправления ошибок начинается с установления факта наличия ошибок – т.е. с процесса обнаружения; однако не все обнаруживаемые ошибки являются исправляемыми.

ТАБЛИЦА ДЕКОДИРОВАНИЯ

Алгоритм исправления ошибок при приёме комбинации $\boldsymbol{y}$ на выходе канала

$$\boldsymbol{x} = \begin{cases} \boldsymbol{a}_i, & \text{если } \boldsymbol{y} = \boldsymbol{b}_j \text{ и } \boldsymbol{b}_j \in M_i, \\ \boldsymbol{a}_i, & \text{если } \boldsymbol{y} = \boldsymbol{a}_i. \end{cases} \qquad (1\text{-}25)$$

При этом должно быть обеспечено выполнение следующих соотношений