УДК 621.394.147

А. И. Армоник

ПОСТРОЕНИЕ ПРОСТЫХ ПОМЕХОУСТОЙЧИВЫХ БЛОКОВЫХ КОДОВ

Рассмотрена проблема уменьшения ошибок при передаче сообщений по каналам с помехами с использованием блоковых кодов. Предложена методика построения помехоустойчивого кода с помощью таблицы кодовых расстояний и простых алгоритмов помехоустойчивой обработки информации. Методика применима к системам управления и передачи информации с небольшим словарем сообщений (10–20 команд).

Реальная система передачи информации (СПИ) всегда подвержена влиянию помех, как внутренних, так и внешних. Это связано и с особенностями конструкции системы, и с условиями ее эксплуатации, и с другими причинами. Для повышения достоверности и качества работы СПИ можно применить помехоустойчивое кодирование. Основой такого кодирования является введение избыточности кода, позволяющей так задать передаваемые последовательности символов, чтобы они удовлетворяли дополнительным условиям, проверка выполнения которых на приемной стороне дает возможность обнаружить и исправить ошибки. Обобщенная функциональная схема такой помехоустойчивой СПИ представлена на рис. 1. Будем рассматривать двоичное представление информации, как наиболее распространенное, и устройства, работающие с двоичными кодами.

Код называют равномерным блочным со словами (блоками) длиной n, если взаимосвязь между символами кодовой последовательности заканчивается через каждые n символов. Искажения информации в канале передачи помехоустойчивой СПИ можно описать с помощью следующих параметров: наибольшей кратности t ошибки; вероятности p появления искаженного символа сообщения.

Первый параметр задает максимальное число искаженных бит информации на блок данных длиной n, при появлении которых необходимо предотвратить пропуск ошибки в передаваемом сообщении. Пара-

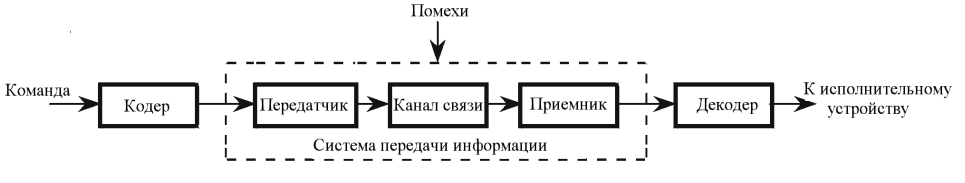

Рис. 1. Обобщенная функциональная схема помехоустойчивой СПИ

метр t характеризует необходимые потенциальные возможности кода к обнаружению ошибок. Реализация этих возможностей обеспечивается устройством декодирования.

Второй параметр определяет качество исполнения канала передачи информации и зависит от вида модуляции, мощности передатчика, уровня шумов, вида помех, чуствительности приемника и других факторов. Так, для когерентной двоичной фазовой манипуляции в условиях аддитивного белого шума со спектральной плотностью N_0 вероятность p появления ошибочного бита определяется формулой [1]

$$p = \frac{1}{2} - \Phi\left(\sqrt{\frac{2E_b}{N_0}}\right),$$

где E_b — энергия, приходящаяся на бит информации; $\Phi(x)$ — функция Лапласа. Если вероятности p_{01} перехода сигнала 0 в сигнал 1 и p_{10} перехода сигнала 1 в сигнал 0 не существенно отличаются друг от друга, то канал передачи является симметричным и принимается $p_{10} = p_{01} = p$.

Значения параметров t и p наиболее точно определяются статистическими данными, полученными в реальных условиях эксплуатации СПИ.

Основные принципы помехоустойчивого кодирования. Рассмотрим симметричный канал с независимыми ошибками. Пусть на вход кодирующего устройства поступает последовательность из k информационных символов, определяющих передаваемое слово. На выходе ей соответствует последовательность из n символов, причем $n > k$. Всего может быть N ($N \le 2^k$) различных входных и 2^n различных выходных последовательностей. Из общего числа 2^n выходных последовательностей только N последовательностей соответствуют входным; они называются разрешенными кодовыми комбинациями, или мощностью кода. Остальные $2^n - N$ возможных выходных последовательностей для передачи не используются и называются запрещенными комбинациями. Поскольку каждая из N разрешенных комбинаций в результате действия помех может трансформироваться в любую другую, то всего имеется $2^n N$ возможных случаев передачи. В это число входят:

— N случаев безошибочной передачи;

— $N(N-1)$ случаев перехода в другие разрешенные комбинации (необнаруживаемых ошибок);

— $N(2^n - N)$ случаев перехода в неразрешенные комбинации (ошибок, которые могут быть обнаружены).

Любой метод декодирования можно рассматривать как правило разбиения всего множества запрещенных кодовых комбинаций на N непересекающихся подмножеств M_i, $i = 1, 2, \ldots, N$, каждое из которых

ставится в соответствие одной из разрешенных комбинаций A_i. При получении запрещенной комбинации, принадлежащей подмножеству M_i, принимают решение, что передавалась разрешенная комбинация A_i. Ошибка будет исправлена в тех случаях, когда полученная комбинация действительно образовалась из комбинации A_i. Таким образом, при наличии избыточности в любом коде появляется возможность исправления ошибки. Способ разбиения на подмножества зависит от того, какие ошибки должны исправляться с помощью данного конкретного кода.

Для уменьшения вероятности $P_{\text{ош}}$ ошибочного декодирования в подмножество M_i следует включать те запрещенные кодовые комбинации B_k, которым с наибольшей вероятностью соответствует переданная комбинация A_i, т.е.

$$P(A_i)P(B_k|A_i) > P(A_j)P(B_k|A_j), \qquad (1)$$

где $i, j = 1, 2, \ldots, N$; $i \neq j$. Выражение (1) определяет критерий максимального правдоподобия.

Мерой различия кодовых комбинаций служит кодовое расстояние d, определяемое числом позиций, в которых одна кодовая комбинация отличается от другой:

$$d = d(A_i, A_j) = \sum_{k=1}^{n} (a_{ik} \oplus a_{jk}),$$

где a_{ik} — k-й бит i-го кодового слова, $\oplus$ — операция суммирования по модулю 2. Набор всех значений $d(A_i, A_j)$ определяет таблицу D кодовых расстояний.

Рассмотрим случай равновероятной передачи сообщений A_i, $i = 1, 2, \ldots, N$. Если передаваемая кодовая комбинация A_i трансформируется в комбинацию B_k, то имеется ошибка в $t = d(A_i, B_k)$ знаках кодового слова (вероятность этого события p^t); тогда остальные $n - t$ знаков должны быть правильными (вероятность этого события $(1 - p)^{n-t}$). В результате получим

$$P(B_k|A_i) = p^t(1 - p)^{n-t}.$$

Поскольку $p < 1 - p$, то вероятность $P(B_k|A_i)$ монотонно убывает с возрастанием t, принимая максимальное значение для кодовой комбинации A_i, которая отличается от принятой комбинации B_k меньшим числом символов. Таким образом, если декодирование производится так, что принятая кодовая комбинация отождествляется с той разрешенной, которая находится от нее на наименьшем кодовом расстоянии, то такое декодирование называется декодированием по методу максимального правдоподобия.

Число возможных комбинаций B_k, находящихся на расстоянии $d(A_i, B_k) = t$ от комбинации A_i, равно числу сочетаний C_n^t. Тогда вероятность появления ошибки кратности t составляет

$$P_t = C_n^t p^t (1-p)^{n-t}. \qquad (2)$$

Из выражения (2) видно, что наиболее вероятны ошибки меньшей кратности, их следует обнаруживать и исправлять в первую очередь.

Известно [2], что минимальное расстояние $d_{\min}$ между разрешенными кодовыми комбинациями должно удовлетворять следующим условиям:

— для обнаружения ошибок кратности не более t необходимо $d_{\min} \geq t+1$;

— для исправления всех ошибок кратности не более s необходимо $d_{\min} \geq 2s+1$;

— для исправления всех ошибок кратности s и одновременного обнаружения всех ошибок кратности t, $t \geq s$, необходимо $d_{\min} \geq t+s+1$.

Наибольшее возможное число N разрешенных комбинаций n-значного кода, находящихся друг от друга на расстоянии $d_i \geq d_{\min}$, можно найти из неравенства [2]

$$N \leqslant \frac{2^n}{\sum\limits_{i=0}^{s} C_n^i}.$$

Таким образом, число N зависит от значности кода n и кодового расстояния d. Поскольку $s = s(d) = (d-1)/2$, то для нечетных d получим

$$N(n,d) \leqslant \frac{2^n}{\sum\limits_{i=0}^{s(d)} C_n^i}.$$

Для четных d, когда $s(d)$ — не целое число, действует правило $N(n,d) = N(n-1, d-1)$, справедливость которого подтверждена экспериментально для $d \geq 5$ или $n \geq 8$. Например, $N(3,2) = 4$. Следует заметить, что при рассмотрении всего множества 2^n слов имеем $N(n,d) = 2^k$, т.е. количество кодовых слов принадлежит множеству $\{2^0, 2^1, 2^2, \ldots, 2^k\}$. Величина k/n называется относительной скоростью кода и показывает, во сколько раз уменьшится скорость передачи информации при помехоустойчивом кодировании.

Построение блокового кода. Для определения разрешенных кодовых комбинаций проанализируем таблицу D кодовых расстояний. Ис-

0	0	0	1-я комбинация
0	0	1	2-я комбинация
0	1	0	3-я комбинация
0	1	1	4-я комбинация
1	0	0	5-я комбинация
1	0	1	6-я комбинация
1	1	0	7-я комбинация
1	1	1	8-я комбинация

а

	1	2	3	4	5	6	7	8
1	0	1	1	2	1	2	2	3
2	1	0	2	1	2	1	3	2
3	1	2	0	1	2	3	1	2
4	2	1	1	0	3	2	2	1
5	1	2	2	3	0	1	1	2
6	2	1	3	2	1	0	2	1
7	2	3	1	2	1	2	0	1
8	3	2	2	1	2	1	1	0

б

Рис. 2. Возможные кодовые комбинации (*а*) и таблица D кодовых расстояний (*б*) для трехзначного кода

ходными данными для анализа являются значность кода n и необходимое минимальное кодовое расстояние $d_{\min}$. Предлагаемый метод заключается в следующем.

1. Выписываются все комбинации (слова) n-значного кода и нумеруются в порядке возрастания двоичных чисел от 1 до 2^n. Для $n = 3$ результат представлен на рис. 2, *а*.

2. Формируется таблица D кодовых расстояний между всеми комбинациями кода. Номера строк i и столбцов j элементов d_{ij} таблицы D задаются номерами соответствующих кодовых комбинаций (рис. 2, *б*).

3. Формируется вспомогательная таблица D', элементы которой удовлетворяют условию

$$d'_{ij} = \begin{cases} 0, & \text{если } d_{ij} < d_{\min}, \quad i \neq j; \\ 1 & \text{иначе.} \end{cases} \tag{3}$$

Для $d_{\min} = 2$ вспомогательная таблица D' представлена на рис. 3, *а*. Элементы таблицы $d'_{ij} = 1$ задают пары комбинаций, обеспечивающих необходимое кодовое расстояние. Сочетающиеся между собой C_N^2 таких пар представляют собой N искомых разрешенных комбинаций кода. Причем таблица D', состоящая только из строк и столбцов, соответствующих этим C_N^2 парам, будет заполнена единичными элементами.

4. Определяются возможные сочетания разрешенных кодовых комбинаций путем "склеивания" таблицы D', включающего в себя следующие шаги:

1) выбирается базовый столбец — первый по порядку, в котором есть хотя бы один нулевой элемент (см. рис. 3, *а*);

2) из таблицы исключаются все строки, в которых элементы базового столбца имеют значение 0, а также столбцы, номера которых соот-

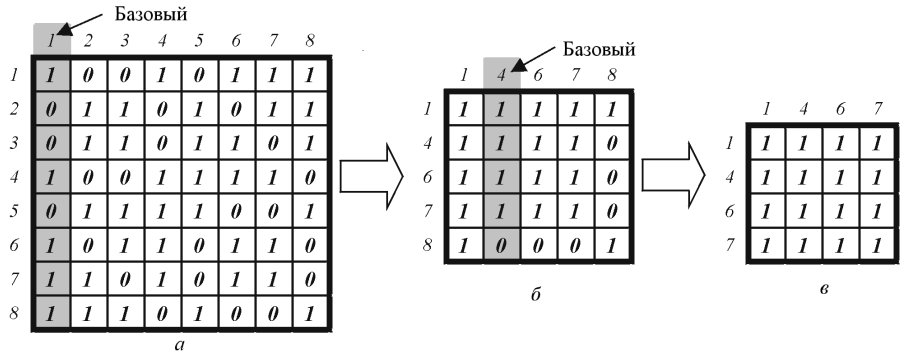

Рис. 3. "Склеивание" вспомогательной таблицы для трехзначного кода ($d_{\min} = 2$):
исходный вид (*а*), промежуточный вариант (*б*), "склеенная" таблица (*в*)

ветствуют номерам исключенных строк (рис. 3, *б*); номера оставшихся строк и столбцов сохраняются;

3) осуществляется переход к шагу 1), если в "склеенной" таблице остался хотя бы один нулевой элемент; в результате получается "склеенная" таблица D' без нулевых элементов, номера строк и столбцов которой указывают на одно из искомых сочетаний разрешенных комбинаций кода (рис. 3, *в*);

4) процесс поиска заканчивается, когда первым базовым столбцом будут выбраны по очереди все столбцы вспомогательной таблицы D'.

В рассматриваемом случае получается два кода с $N = 4$ разрешенными комбинациями: код, включающий кодовые слова, соответствующие 1-й, 4-й, 6-й и 7-й комбинациям, — код 1–4–6–7 и код, включающий кодовые слова, соответствующие 2-й, 3-й, 5-й и 8-й комбинациям, — код 2–3–5–8 (см. рис. 2). Заметим, что код 1–4–6–7 является групповым, т.е. обладает следующими свойствами:

— содержит нулевое кодовое слово;

— сумма любых двух кодовых слов по модулю 2 также является кодовым словом.

Подробные сведения о групповых кодах приведены в работах [2–10]. Таким образом, для построения с помощью предложенного метода всех групповых кодов, включающих 2^n слов, достаточно рассмотреть в качестве первой образующей нулевую комбинацию, т.е. комбинацию с номером 1.

Для нахождения эквидистантных кодов, т.е. кодов, в которых кодовое расстояние между всеми словами одинаковое, достаточно изменить условие (3) для значений элементов таблицы D':

$$d'_{ij} = \begin{cases} 0, \text{ если } d_{ij} \neq d_{\min}, \ i \neq j, \\ 1 \text{ иначе.} \end{cases}$$

Особый интерес могут представлять коды с определенным количеством единичных элементов, задающим вес w кодовых комбинаций. Значение веса может быть одним для всех кодовых комбинаций или же вес может принимать несколько значений. Чтобы искомый код обладал определенным набором весов W, при построении таблицы D должны учитываться только те кодовые комбинации из множества 2^n, для которых $w \in W$.

Методика определения разрешенных комбинаций с помощью анализа таблицы D достаточно легко реализуется при $n \leq 8$, например, с помощью программы MS Excel. Коды с бóльшим количеством символов можно построить простыми способами на основе уже построенных кодов с меньшим числом символов.

Рассмотрим метод замещения, когда каждый символ кодовой комбинации X_i заменяется кодом, состоящим из двух слов Y_0, Y_1, так, что каждый элемент исходного кода $x_j = 0$ переходит в комбинацию Y_0, а $x_j = 1$ — в Y_1. Например, кодовая комбинация 0101 преобразуется в комбинацию 01100110, если символ 0 заменить последовательностью 01, а символ 1 — последовательностью 10. Для полученного кода $d = d_X d_Y$, $n = n_X n_Y$. Заметим, что если комбинации Y_0 и Y_1 равновесные, т.е. содержащие одно и то же количество нулей и единиц, то комбинации преобразованного кода также будут обладать одним весом $w = n_X w_Y$.

Еще один способ — метод объединения комбинаций, принадлежащих различным кодам. Объединение символов исходных кодов должно происходить в строго определенной последовательности: i-му символу строящегося кода всегда соответствует j-й символ одного из исходных кодов. В простейшем случае комбинации кодов записывают последовательно:

$$x_1 x_2 \ldots x_{n_X} y_1 y_2 \ldots y_{n_Y} \ldots z_1 z_2 \ldots z_{n_Z}.$$

Для такого кода $d = \sum d_k$, $n = \sum n_k$, где d_k и n_k — кодовое расстояние и количество символов k-го кода, входящего в объединенный код.

Методика помехоустойчивого кодирования информации. Поиск кодовых слов с помощью анализа таблицы кодовых расстояний является универсальным, так как позволяет включить в рассмотрение все возможные блоковые коды и выбрать из них наиболее подходящие. Построенный код позволяет обнаруживать ошибки, но не обла-

дает корректирующей способностью. Кроме того, передаваемые команды должны быть связаны с кодом таблицей соответствия, которая хранится непосредственно в памяти кодирующего и декодирующего устройств. Таким образом, использование кода становится мало эффективным при числе команд $N > 10 \ldots 20$ из-за необходимости хранения большого количества данных в памяти кодирующего и декодирующего устройств. К тому же, время исполнения алгоритма помехоустойчивого кодирования информации будет тем продолжительнее, чем больше число N.

Составления таблицы соответствия можно избежать, если задать соответствие между командами и кодовыми словами в виде алгоритма. Например, команда задает адрес ячейки памяти, в которой хранится кодовое слово, или входит в состав кодового слова.

Для обеспечения возможности исправления ошибок применим метод λ-кратного повторения кодовой комбинации $x_1x_2 \ldots x_n$, который является частным случаем метода объединения:

$$x_1x_2 \ldots x_nx_1x_2 \ldots x_n \ldots x_1x_2 \ldots x_n;$$
$$x_1x_1 \ldots x_1x_2x_2 \ldots x_2 \ldots x_nx_n \ldots x_n.$$

(в первом случае исходная комбинация X_i повторяется λ раз подряд, во втором — сначала λ раз подряд записывается первый символ комбинации X_i, затем λ раз подряд — второй символ и т.д.).

При такой структуре кода в памяти кодирующего и декодирующего устройств хранится только исходная кодовая последовательность $x_1x_2 \ldots x_n$. Удобно, если λ — нечетное число, тогда суждение о правильном значении каждого элемента x_i, $i = 1, 2, \ldots, n$, выносится по наибольшему количеству одинаковых значений соответствующих элементов кода (“два из трех”, “три из пяти” и т.д.). Например, кодовая комбинация 01101 при последовательном трехкратном повторении была частично искажена помехами, и на выходе приемника появился код 10101 01110 01001. В результате проверки правильной считается комбинация 01101.

Два главных достоинства кодов с повторением — простота реализации и возможность исправления не только независимых ошибок, но и пакетов ошибок. Выбор числа копий кодовых комбинаций, характеристик исходных комбинаций, метода объединения зависит от конкретных условий применения этих кодов.

Расчет вероятностей ошибок. Вероятность появления ошибочной комбинации на выходе помехоустойчивой СПИ зависит от выбранного помехоустойчивого кода и алгоритма декодирования. Рассмотрим 2^n возможных комбинаций на входе декодирующего устройства, разделив их на три области: разрешенные кодовые комбинации (N кодовых

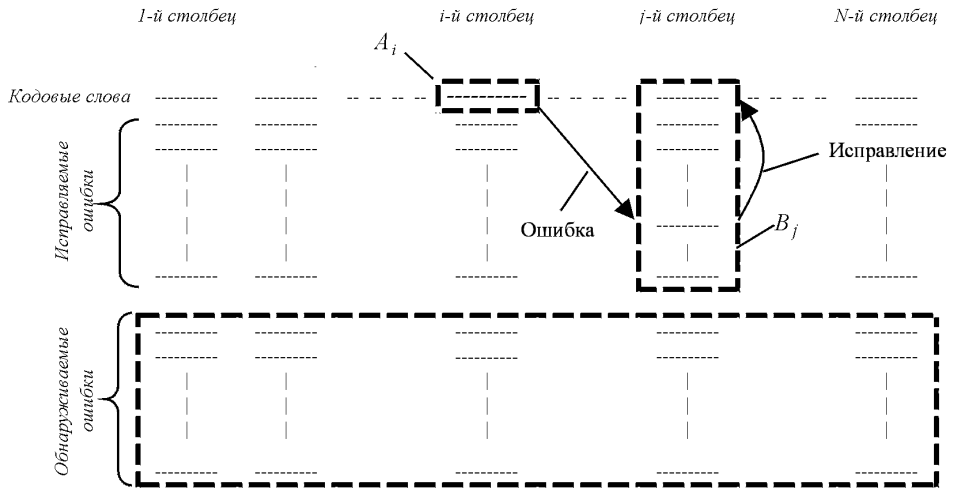

Рис. 4. Обобщенная таблица декодирования

слов), исправляемые ошибки, обнаруживаемые ошибки. Представим их в виде обобщенной таблицы декодирования (рис. 4).

В первой строке таблицы выписываются все кодовые слова, под каждым кодовым словом — все принятые наборы, которые можно при декодировании преобразовать в это кодовое слово. Эти наборы будут составлять область исправляемых ошибочных последовательностей. Ниже выписываются все кодовые комбинации, которые однозначно декодировать нельзя. Они составляют область ошибок, которые нельзя исправить, но можно обнаружить.

Будем рассматривать следующие события: A_i (передано i-е кодовое слово); B_i — принятое слово соответствует последовательности из i-го столбца, содержащей исправляемую ошибку; C — принятое слово соответствует последовательности, содержащей обнаруживаемую ошибку. Тогда вероятность правильного декодирования составляет

$$P = \sum_{i=1}^{2^k} P\left(A_i B_i\right) = \sum_{i=1}^{2^k} P(A_i) P(B_i | A_i),$$

вероятность появления ошибочной последовательности —

$$P_{\text{ош}} = 1 - P = 1 - \sum_{i=1}^{2^k} P(A_i) P(B_i | A_i),$$

$$P_{\text{ош}} = P_{\text{ош}}^{\text{об}} + P_{\text{ош}}^{\text{необ}},$$

где

$$P_{\text{ош}}^{\text{об}} = \sum_{i=1}^{2^k} P(A_i C) = \sum_{i=1}^{2^k} P(A_i) P(C | A_i)$$

— вероятность обнаруженной, но неисправленной ошибочной последовательности,

$$P_{\text{ош}}^{\text{необ}} = P_{\text{ош}} - P_{\text{ош}}^{\text{об}} \tag{4}$$

— вероятность появления необнаруживаемой ошибки.

Пусть A_0, B_0 — события передачи и приема нулевых кодовых слов соответственно. Для группового кода справедливо следующее: для любых значений i и j существует такое значение m, что $P(B_m|A_0) = P(B_j|A_i)$ и, более того, $P(B_0|A_0) = P(B_i|A_i)$. Тогда имеем

$$P_{\text{ош}} = 1 - P(B_0|A_0)\sum_{i=1}^{2^k} P(A_i)$$

или с учетом того, что

$$\sum_{i=1}^{2^k} P(A_i) = 1,$$

имеем

$$P_{\text{ош}} = 1 - P(B_0|A_0), \quad P_{\text{ош}}^{\text{об}} = P(C|A_0).$$

Рассмотрим декодирующее устройство, исправляющее все комбинации из не более чем b взаимно независимых ошибок. В этом случае

$$P = \sum_{i=0}^{b} C_n^i p^i (1-p)^{n-i},$$

$$P_{\text{ош}} = 1 - \sum_{i=0}^{b} C_n^i p^i (1-p)^{n-i}.$$

Пусть B'_j — событие, когда принятое слово совпадает с некоторой кодовой комбинацией, содержащей исправляемую ошибку и расположенной в таблице декодирования под некоторым кодовым словом, имеющим вес w, не являющимся тем словом, которое передавалось; N_j — количество кодовых слов, имеющих вес $w = j$. Поскольку декодирующее устройство исправляет b ошибок, то область значений весов кодовых слов, в которых возможно появление необнаруживаемой ошибки, — это $2b + 1 \le j \le n$. Необнаруживаемая ошибка появляется в ситуации, когда переданное кодовое слово искажается и при декодировании преобразуется в другое кодовое слово, т.е.

$$P_{\text{ош}}^{\text{необ}} = \sum_{j=2b+1}^{n} N_j P(B'_j|A_0).$$

Вероятность того, что переданное кодовое слово будет распознано как некоторое другое кодовое слово, не соответствующее передаваемому, составляет $P(B'_j|A_0) = P_1 + P_2 + P_3$, где

$$P_1 = \sum_{v=1}^{b} C_j^v p^{j-v} (1-p)^{n-(j-v)}, \quad P_2 = p^j (1-p)^{n-j},$$

$$P_3 = \begin{cases} \sum\limits_{v=1}^{b} C_{n-j}^v p^{j+v} (1-p)^{n-(j+v)} & \text{при} \quad n-j \geqslant b, \\ \sum\limits_{v=1}^{n-j} C_{n-j}^v p^{j+v} (1-p)^{n-(j+v)} & \text{при} \quad 0 < n-j < b, \\ 0 \quad \text{при} \quad n-j = 0. \end{cases}$$

Тогда вероятность $P_{\text{ош}}^{\text{об}}$ можно получить из формулы (4).

Рассмотрим системы, которые только обнаруживают ошибки и не исправляют их. Правильная передача кодового слова возможна в случае отсутствия ошибок, т.е.

$$P = \sum_{i=0}^{n} C_n^i p^i (1-p)^{n-i} = (1-p)^n, \quad P_{\text{ош}} = 1 - (1-p)^n.$$

Тогда, учитывая, что область исправляемых ошибок отсутствует, получим

$$P_{\text{ош}}^{\text{необ}} = \sum_{j=1}^{n} N_j^{\text{п}} p^j (1-p)^{n-j};$$

здесь $p^j(1-p)^{n-j}$ — вероятность события, когда под воздействием ошибки кратности $t = j$ передаваемое кодовое слово X_0 перешло в другое кодовое слово, имеющее вес $w = j$; $N_j^{\text{п}} \equiv N_j$ — возможное количество таких переходов для слов, имеющих вес $w = j$. Таким образом, вероятность обнаружения ошибочной последовательности имеет вид

$$P_{\text{ош}}^{\text{об}} = P_{\text{ош}} - P_{\text{ош}}^{\text{необ}} = 1 - \sum_{j=0}^{n} N_j p^j (1-p)^{n-j}.$$

Заметим, что в случае использования алгоритма только обнаружения ошибок вероятность $P_{\text{ош}}^{\text{необ}}$ всегда меньше, чем при частичной корректировке ошибок, когда исправляются все ошибки кратности не более чем t, и существенно меньше, чем в случае, когда исправляются все принятые последовательности. Кроме того, при любом коде имеет место неравенство $P_{\text{ош}}^1 \geq P_{\text{ош}}^2 \geq P_{\text{ош}}^{\text{необ}}$, называемое теоремой Финка,

причем оно переходит в равенство только для безызбыточных кодов [5]; здесь $P^1_{\text{ош}}$ — вероятность появления ошибки $P_{\text{ош}}$ в случае использования алгоритма обнаружения без исправления; $P^2_{\text{ош}}$ — вероятность появления ошибки $P_{\text{ош}}$ для случая исправления максимально возможного числа ошибок.

Особенности расчета ассиметричного канала и пакетов ошибок. Вероятности p_{01} перехода символа 0 в символ 1 и p_{10} перехода символа 1 в символ 0 могут существенно отличаться друг от друга, т.е. канал передачи может быть ассиметричным. Полученные результаты можно использовать и для этого случая, учитывая, что необходимо рассматривать ошибки двух видов: переход сигнала 1 в сигнал 0 и переход сигнала 0 в сигнал 1. Для кодового расстояния d получим [11]

$$
\begin{aligned}
&d \geqslant (s_{10} + t_{10} + 1) + (s_{01} + t_{01} + 1) - \Delta w_{\min},\\
&\qquad \text{если } \; 2s_{01} + r + 1 \leqslant \Delta w_{\min} \leqslant s_{10} + s_{01} + r + 1,\\
&d \geqslant s_{10} + t_{10} + 1, \;\; \text{если } \; \Delta w_{\min} \leqslant 2s_{01} + r + 1,\\
&s_{10} > s_{01};
\end{aligned}
$$

здесь s_{01}, s_{10} — кратности исправляемых ошибок, а $t_{01} = s_{01} + r$, $t_{10} = s_{10} + r$ — кратности обнаруживаемых ошибок переходов сигнала 0 в сигнал 1 и сигнала 1 в сигнал 0 соответственно; $\Delta w_{\min}$ — наименьшая разность между весами w_i кодовых комбинаций. Для случая $s_{01} > s_{10}$ выражение аналогично приведенному: индекс “01” заменяется индексом “10”, и наоборот. Таким образом, необходимое кодовое расстояние для кода в асимметричном канале может быть меньше кодового расстояния в симметричном канале, что определяется величиной параметра $\Delta w_{\min}$. Для кодов, для которых $\Delta w_{\min} < 2$, результаты совпадают. Завышение значения d позволяет увеличить устойчивость кода к ошибкам большей кратности, что повышает помехоустойчивость системы. Поэтому уменьшение кодового расстояния имеет значение только в случае его влияния на другие параметры кода, значения которых ограничены.

При расчете вероятностей используют тот же подход, что и в случае симметричного канала. Однако формулы, в которых для подсчета возможных комбинаций ошибок используется сочетание C^j_i, необходимо изменить, так как в них не учитываются различия в переходах сигнала 1 в сигнал 0 и сигнала 0 в сигнал 1. Так, для декодирующего устройства, исправляющего все комбинации из не более чем $s = s_{01} + s_{10}$ взаимно независимых ошибок, в случае равновесного кода имеем

$$
P = \sum_{i=0}^{s_{01}} C^i_{n-w} p^i_{01} \left(1 - p_{01}\right)^{n-w-i} + \sum_{i=0}^{s_{10}} C^i_w p^i_{10} \left(1 - p_{10}\right)^{w-i}.
$$

В случае группирующихся ошибок (пакетов ошибок) подбор помехоустойчивого кода и расчет его параметров можно вести по рассмотренной методике с учетом наибольшей кратности возникающих ошибок. В результате код будет обладать большей помехозащищенностью, так как будет устойчив не только к пакетам ошибок. Однако параметры кода будут не оптимальными.

Таким образом, результаты, полученные для симметричного канала с независимыми ошибками, показывают, что существует запас по помехоустойчивости кода и для асимметричного канала, и для пакетов ошибок. Получить точные результаты для этих случаев можно только с учетом статистики ошибок.

Границы вероятности появления ошибки. Определим границы для вероятности $P_{\text{ош}}$. Из теории вероятностей известно, что если некоторое событие может быть представлено в виде объединения нескольких подсобытий, то вероятность этого события не превосходит суммы вероятностей подсобытий:

$$P(A_1 \dots A_i) \le P(A_1) + \dots + P(A_i).$$

Это может быть применено к любому блоковому или сверточному коду. При вычислении вероятности появления ошибки для группового кода достаточно рассмотреть событие передачи нулевого кодового слова. Таким образом, вероятность появления ошибки ограничена сверху суммой вероятностей появления отдельных комбинаций ошибок. Пусть событие B''_j соответствует случаю, когда расстояние между принятым словом и некоторым кодовым словом, имеющим вес j, меньше расстояния между принятым словом и нулевым кодовым словом, т.е. произошла ошибка кратности $t > d/2$. Тогда

$$P_{\text{ош}} \leqslant \sum_{j=1}^{n} N_j P(B''_j | A_0), \tag{5}$$

где $P(B''_j|A_0)$ вычисляется суммированием по всем словам, которые отличаются от данного кодового слова, имеющего вес j, в не более чем $j/2$ ненулевых позициях этого кодового слова, т.е.

$$P(B''_j|A_0) = \begin{cases} \displaystyle\sum_{i=\frac{j+1}{2}} C_j^i p^i (1-p)^{j-i}, \ \text{если } j \text{ нечетно}, \\ \displaystyle\frac{1}{2} C_j^{\frac{j}{2}} p^{\frac{j}{2}} (1-p)^{\frac{j}{2}} + \sum_{i=\frac{j}{2}+1} C_j^i p^i (1-p)^{j-i}, \ \text{если } j \text{ четно}. \end{cases}$$

Полученное неравенство (5) называется аддитивной границей вероятности ошибки.

Получим также верхнюю границу случайного кодирования для $P_{\text{ош}}$. Для определения этой границы используется тот факт, что иногда легче вычислить среднюю вероятность ошибки для класса кодов, чем для каждого отдельного кода, и что в классе кодов должен существовать хотя бы один код со свойствами не хуже усредненных. Вероятность ошибочного декодирования для наилучшего группового (n, k)-кода ограничена сверху [6]:

$$P_{\text{ош}} \le \sum_{j=\frac{d^*}{2}}^{d^*-1} \sum_{i=d^*-j}^{j} \sum_{h=d^*}^{i+j} C_{n-j}^{\frac{h+i-j}{2}} C_j^{\frac{h+j-1}{2}} \frac{C_n^j p^i (1-p)^{n-j}}{2^{n-k-1}} + \sum_{i=d^*}^{n} C_n^j p^j (1-p)^{n-j},$$

где d^* — наибольшее целое число, удовлетворяющее условию

$$2^{n-k} \ge 2 \cdot \sum_{i=0}^{d^*-1} C_n^i.$$

Кроме того, справедливо следующее утверждение: минимальное кодовое расстояние бóльшей части (более половины) групповых кодов равно по меньшей мере d^*.

Для любого блокового (n, k)-кода справедливо неравенство, называемое границей сферической упаковки [6]:

$$P_{\text{ош}} \ge \left(C_n^{t+1} - \alpha_{t+1}\right) p^{t+1} (1-p)^{n-t-1} + \sum_{i=t+2}^{n} C_n^i p^i (1-p)^{n-i},$$

где t и α_{t+1} определяются из условия

$$\alpha_{t+1} = 2^{n-k} - \sum_{i=0}^{t} C_n^i \ge 0,$$

причем t — наибольшее целое число.

Заключение. Выбор конкретного кода из всех найденных зависит от значений P, $P_{\text{ош}}$, $P_{\text{ош}}^{\text{об}}$, $P_{\text{ош}}^{\text{необ}}$, определяющих его потенциальную помехоустойчивость. В таблице представлены параметры нескольких блоковых (n, k)-кодов, рассчитанные для симметричного канала с независимыми ошибками и разных способов декодирования. В этой таблице для всех кодов вероятность p появления искаженного бита кода в канале передачи выбрана равной 0,1. Расчет показывает, что для помехоустойчивых кодов значения параметров P, $P_{\text{ош}}$, $P_{\text{ош}}^{\text{об}}$, $P_{\text{ош}}^{\text{необ}}$ сильно зависят от значений величины p. Так, код БЧХ (31,6) (код Боуза–Чоудхури–Хоквингема с параметрами $n = 31$, $k = 6$) в режиме исправления ошибок для $p = 0{,}01$ обеспечивает значения $P \approx 1$,

$P_{\text{ош}} \approx P_{\text{ош}}^{\text{об}} = 1{,}2 \cdot 10^{-5}$, $P_{\text{ош}}^{\text{необ}} = 2{,}8 \cdot 10^{-7}$, но эти параметры существенно изменяются, если p увеличить на порядок (см. таблицу). Таким образом, при значении $p = 0{,}1$ для многих кодов происходит резкое снижение помехоустойчивости. Появление же одной–трех ошибок на 10 бит кода в СПИ небольшой стоимости — вполне ожидаемая величина для p.

Код	$d_{\min}$	Способ декодирования	P	$P_{\text{ош}}$	$P_{\text{ош}}^{\text{об}}$	$P_{\text{ош}}^{\text{необ}}$
Хемминга (7,4)	3	исправление	0,850	0,150	0	0,150
		обнаружение	0,478	0,522	0,517	$5{,}100 \cdot 10^{-3}$
Голея (23,12)	7	исправление	0,807	0,193	0	0,193
Голея (24,12)	8	исправление	0,807	0,193	0	0,193
БЧХ (15,5)	8	исправление	0,944	0,056	0,043	$1{,}272 \cdot 10^{-2}$
		обнаружение	0,206	0,794	0,794	$6{,}700 \cdot 10^{-8}$
БЧХ (31,6)	14	исправление	0,969	0,031	0,021	$9{,}587 \cdot 10^{-3}$
		обнаружение	0,038	0,962	0,962	$2{,}018 \cdot 10^{-14}$
Групповой код (3, 2)	2	обнаружение	0,729	0,271	0,244	$2{,}700 \cdot 10^{-2}$
Тройное повторение	6	проверка “2 из 3”	0,911	0,089	0,086	$2{,}817 \cdot 10^{-3}$
Дублирование тройного повторения	12	проверки “2 из 3” и совпадение результата в дублях	0,830	0,170	0,170	$2{,}645 \cdot 10^{-6}$

Заметим, что режим обнаружения обеспечивает хорошие показатели для необнаруживаемых ошибок, но получение переданного слова становится маловероятным событием. В режиме исправления, наоборот, подавляющее количество команд будет принято правильно, но при этом велика возможность перехода одного слова команды в другое ($P_{\text{ош}}^{\text{необ}}$).

Помимо нескольких широко известных кодов, в таблице приведены характеристики группового кода 1–4–6–7. Такой код можно использовать только в режиме обнаружения ($d = d_{\min} = 2$, $n = 3$). Однако метод трехкратного повторения кодовой комбинации с проверкой “два из трех” ($d = 6$, $n = 9$) позволяет исправить все одиночные, часть двойных и тройных ошибок, его помехоустойчивость приближается к помехоустойчивости одного из лучших блоковых кодов — кода БЧХ. Простое дублирование таких кодовых посылок с проверкой на совпадение результата в дублях ($d = 12$, $n = 18$) еще больше улучшает веро-

ятностные характеристики: вероятность $P_{ош}^{необ}$ уменьшается на три порядка при увеличении P приблизительно на 10 %. При этом сложность алгоритма кодированиия и декодирования изменяется незначительно. Таким образом, подбором количества копий исходного кода в кодовой посылке и повторением таких кодовых посылок можно изменять характеристики P, $P_{ош}$, $P_{ош}^{об}$, $P_{ош}^{необ}$.

Полученные результаты позволяют сделать вывод о возможности практического применения рассмотренной методики. Решение проблемы построения помехоустойчивых кодов в применении к конкретной помеховой обстановке возможно с помощью математической модели канала связи, которая будет учитывать различные виды помех и позволит получить сведения о характеристиках и оптимальной структуре кода для этих видов помех, а также возможных их сочетаний.

СПИСОК ЛИТЕРАТУРЫ

1. Скляр Б. Цифровая связь. Теоретические основы и практическое применение. – М.: Изд. дом "Вильямс", 2003.
2. Дмитриев В. И. Прикладная теория информации. – М.: Высшая школа, 1989.
3. Темников Ф. Е. и др. Теоретические основы информационной техники. – М.: Энергия, 1979.
4. Злотник Б. М. Помехоустойчивые коды в системах связи. – М.: Радио и связь, 1989.
5. Бородин Л. Ф. Введение в теорию помехоустойчивого кодирования. – М.: Сов. радио, 1968.
6. Питерсон У., Уэлдон Э. Коды, исправляющие ошибки. – М.: Мир, 1976.
7. Хемминг Р. В. Теория кодирования и теория информации. – М.: Радио и связь, 1983.
8. Блейхут Р. Теория и практика кодов, контролирующих ошибки. – М.: Мир, 1986.
9. Касами Т., Токура Н., Ивадари И., Инагаки Я. Теория кодирования. – М.: Мир, 1978.
10. Кларк Дж. мл., Кейн Дж. Кодирование с исправлением ошибок в системах цифровой связи / Под ред. Б.С. Цыбакова. – М.: Радио и связь, 1987. – (Статистическая теория связи. Вып. 28).
11. Харкевич А. А. Борьба с помехами. – М.: Наука, 1965.

Статья поступила в редакцию 24.05.2004

Андрей Игоревич Армоник родился в 1977 г., окончил в 2001 г. МГТУ им. Н.Э. Баумана. Аспирант кафедры "Автономные информационные и управляющие системы" МГТУ им. Н.Э. Баумана. Специализируется в области помехоустойчивого кодирования применительно к радиокомандным системам управления.

A. Armonik (b. 1977) graduated from the Bauman Moscow State Technical University in 2001. Post-graduate of "Autonomous Information and Control Systems" department of the Bauman Moscow State Technical University. Specializes in the field of noise immunity coding applied to radio command control systems.