

ФЕДЕРАЛЬНОЕ АГЕНТСТВО СВЯЗИ

**Федеральное государственное образовательное бюджетное
учреждение высшего профессионального образования
«САНКТ-ПЕТЕРБУРГСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ТЕЛЕКОММУНИКАЦИЙ
им. проф. М. А. БОНЧ-БРУЕВИЧА»**

С. С. Владимиров

**МАТЕМАТИЧЕСКИЕ ОСНОВЫ
ТЕОРИИ ПОМЕХОУСТОЙЧИВОГО
КОДИРОВАНИЯ**

Курс лекций

СПб ГУТ)))

**Санкт-Петербург
2014**

1. Помехоустойчивое кодирование

Помехоустойчивое кодирование (англ. Error Correcting Coding, ECC) — процесс преобразования информации, предоставляющий возможность обнаружить и исправить ошибки, возникающие при передаче информации по каналам передачи данных.

Процесс помехоустойчивого кодирования заключается во введении *избыточности*, т. е. для передачи информации используется код, у которого используются не все возможные комбинации, а только некоторые из них. Такие коды называют избыточными или корректирующими.

Соответственно, процесс введения избыточности (преобразование информационных символов в кодовое слово) называется *кодированием*, а обратный процесс восстановления информации из кодового слова, возможно содержащего ошибки — *декодированием*.

Еще одним важным свойством помехоустойчивого кодирования является эффект *усреднения шума*, который достигается за счет того, что избыточные символы зависят от нескольких информационных символов.

В рамках цифровой системы передачи данных задачи кодирования и декодирования возложены на *кодер* и *декодер* соответственно. Структура цифровой системы передачи данных показана на рис. 1.1 [1].

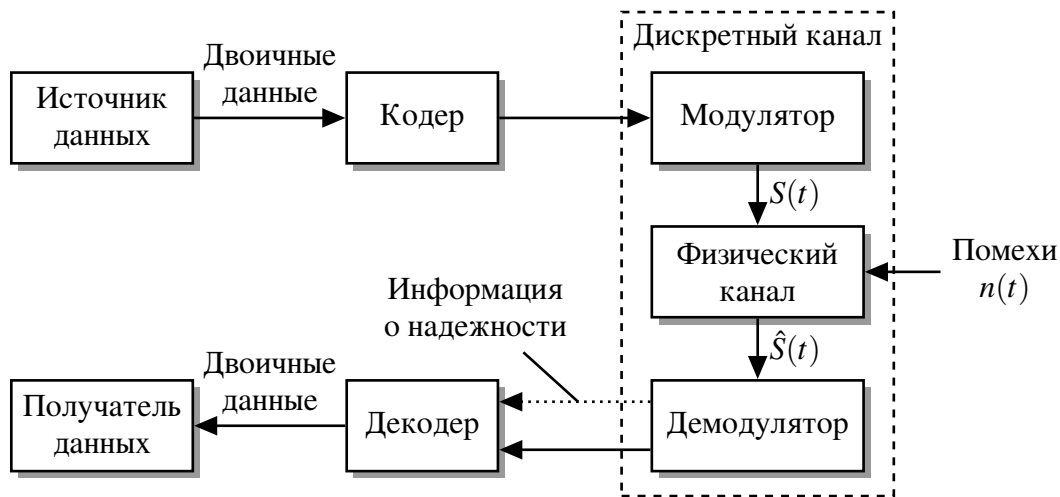


Рис. 1.1. Структура цифровой системы передачи данных

На рис. 1.1 также показано, что часто декодеру доступна информация, указывающая на надежность решений, принимаемых о различных символах кодового слова. Такая информация часто может быть использована для упрощения процесса декодирования, либо для улучшения его характеристик [1].

В целом, способность помехоустойчивых кодов определять и исправлять ошибки — их корректирующие свойства — зависят от правил построения этих кодов и параметров кода (числа разрядов, избыточности и др.), а также от используемых алгоритмов декодирования.

1.1. Основные параметры помехоустойчивых кодов

Основными параметрами, характеризующими корректирующие свойства кодов являются

1. Избыточность кода.
2. Кодовое расстояние.
3. Кратность гарантированно обнаруживаемых ошибок.
4. Кратность гарантированно исправляемых ошибок.

1.1.1. Избыточность корректирующего кода

Избыточность корректирующего кода может быть абсолютной и относительной. Под *абсолютной избыточностью* понимают число вводимых дополнительных разрядов

$$r = n - k,$$

где n — число кодовых символов на выходе кодера, соответствующих k информационных символов на его входе.

Относительной избыточностью корректирующего кода называют величину

$$R_{\text{отн}} = \frac{r}{n} = \frac{n - k}{n} = 1 - \frac{k}{n}.$$

С ней связана так называемая *относительная скорость передачи информации* или *скорость кода*, которая показывает, какую часть общего числа символов кодовой комбинации составляют информационные символы.

$$\frac{k}{n} = 1 - R_{\text{отн}}.$$

Если производительность источника равна H символов в секунду, то скорость передачи после кодирования этой информации будет равна

$$R = H \cdot \frac{k}{n}.$$

1.1.2. Кодовое расстояние

Кодовое расстояние d или *расстояние Хемминга* характеризует степень различия любых двух кодовых комбинаций. Оно выражается числом символов, которыми комбинации отличаются одна от другой.

Чтобы получить кодовое расстояние между двумя комбинациями двоичного кода, достаточно подсчитать число единиц в сумме этих комбинаций по модулю 2.

$$10011 \oplus 11001 = 01010 \quad \Rightarrow \quad d = 2.$$

Кодовое расстояние может быть различным. Так, в первичном натуральном безызбыточном коде это расстояние для различных комбинаций может различаться от единицы до n , где n — значность/длина кода.

Для помехоустойчивого кода наиболее важным является *минимальное кодовое расстояние* d_{\min} — наименьшее кодовое расстояние из всех между всеми парами кодовых комбинаций.

В безызбыточном коде все комбинации являются разрешенными, $d_{\min} = 1$. Достаточно только исказиться одному символу, и будет ошибка в сообщении.

1.1.3. Кратности гарантированно обнаруживаемой и гарантированно исправляемой ошибки

Эти параметры напрямую зависят от минимального кодового расстояния. Под *кратностью* понимается количество поражённых ошибкой символов кодовой комбинации.

В общем случае при необходимости обнаруживать ошибки кратности $t_{\text{обн}}$ минимальное кодовое расстояние должно быть, по крайней мере, на единицу больше $t_{\text{обн}}$, то есть

$$d_{\min} \geq t_{\text{обн}} + 1.$$

Соответственно, *кратность гарантированно обнаруживаемой кодом ошибки* равна

$$t_{\text{обн}} \leq d_{\min} - 1.$$

Кратность гарантированно исправляемой кодом ошибки вычисляется по формуле

$$t \leq \frac{d_{\min} - 1}{2}.$$

Таким образом, код, имеющий минимальное кодовое расстояние $d_{\min} = 3$, позволяет гарантированно обнаружить $t_{\text{обн}} = 2$ и менее ошибок и гарантированно исправить $t = 1$ ошибку.

1.2. Классификация помехоустойчивых кодов

Помехоустойчивые коды классифицируются по различным признакам. Одной из основных классификаций является деление кодов на *блочные* и *непрерывные*.

Блочный (блоковый) код является *кодом без памяти*. Кодер блочного кода отображает подающийся на вход блок информационных символов длиной k в кодовую последовательность из n выходных символов. Термин «без памяти» указывает, что каждый блок из n символов зависит только от соответствующего блока из k символов и не зависит от других блоков [1].

Основными параметрами блочных кодов являются длина информационного блока k , длина кодового слова n , скорость кода $\frac{k}{n}$ и минимальное кодовое расстояние d_{\min} .

Непрерывные или *древовидные* коды — это коды, исправляющие ошибки, которые используют непрерывную, или последовательную, обработку информации короткими фрагментами (блоками). Чаще всего используются линейные древовидные коды, называемые *сверточными*. Кодер древовидного кода является устройством *с памятью*. На вход поступают наборы из k входных информационных символов, а на выходе появляются наборы из n кодовых символов. Каждый набор n кодовых символов зависит от текущего входного набора и от v предыдущих входных символов. Следовательно кодер должен содержать устройство памяти на $m = k + v$ входных символов. Параметр m часто называют *длиной кодового ограничения* кода [1].

Также *непрерывные* коды характеризуются скоростью кода $\frac{k}{n}$ и свободным расстоянием $d_{\text{св}}$ [1].

Особое место в такой классификации занимают каскадные коды и турбо коды, представляющие из себя комбинации блочных и/или непрерывных кодов [2].

Другой подход к классификации делит коды на *линейные* и *нелинейные*. Линейные коды образуют векторное пространство. Два кодовых слова линейного кода при сложении по определенному правилу дают в результате третье кодовое слово [1].

Практически все применяемые на практике схемы кодирования основаны на использовании линейных кодов. Двоичные линейные блочные коды часто называют *групповыми* кодами, так как их кодовые слова образуют математическую структуру, называемую *группа* [1].

По способу кодирования коды делятся на *систематические* и *несистематические*. В первом случае информационные символы передаются на выход декодера без изменения и к ним добавляются проверочные символы. В случае несистематического кодирования информационные символы в явном виде в кодовом слове отсутствуют.

Ещё одним вариантом деления помехоустойчивых кодов является разделение их на *коды, исправляющие случайные ошибки*, и *коды, исправляющие пакеты (пачки) ошибок*. Хотя для исправления пачек ошибок было разработано большое количество кодов с хорошими характеристиками, часто оказывается выгодным использовать коды, исправляющие случайные ошибки, совместно с устройствами перемежения/деперемежения [1]. Также стоит отметить, что существуют алгоритмы декодирования, позволяющие использовать коды, рассчитанные для исправления случайных ошибок, для исправления пачек ошибок без использования перемежителей.