Министерство образования Республики Беларусь

Учреждение образования
«Гомельский государственный технический
университет имени П. О. Сухого»

Кафедра «Промышленная электроника»

В. В. Щуплов

ЭФФЕКТИВНЫЕ КОДЫ

ЛАБОРАТОРНЫЙ ПРАКТИКУМ
по курсу «Теория электросвязи»
для студентов специальности 1-36 04 02
«Промышленная электроника»
дневной формы обучения
В двух частях
Часть 2

Гомель 2011

УДК 621.394.14(075.8)
ББК 32.811.4я73
Щ97

Рекомендовано научно-методическим советом
факультета автоматизированных и информационных систем
ГГТУ им. П. О. Сухого
(*протокол № 10 от 28.06.2010 г.*)

Рецензент: канд. техн. наук, доц. каф. «Электроснабжение»
ГГТУ им. П. О. Сухого *О. Г. Широков*

Щуплов, В. В.

Щ97 Эффективные коды : лаборатор. практикум по курсу «Теория электросвязи» для студентов специальности 1-36 04 02 «Промышленная электроника» днев. формы обучения : в 2 ч. Ч. 2 / В. В. Щуплов. – Гомель : ГГТУ им. П. О. Сухого, 2011. – 27 с. – Систем. требования: PC не ниже Intel Celeron 300 МГц ; 32 Mb RAM ; свободное место на HDD 16 Mb ; Windows 98 и выше ; Adobe Acrobat Reader. – Режим доступа: http://lib.gstu.local. – Загл. с титул. экрана.

Даны необходимые сведения для освоения теоретического материала и практического закрепления знаний по курсу «Теория электросвязи». Рассмотрены основные понятия теории информации и методы эффективного кодирования на примере эффективного кода Хафмена.

Для студентов специальности 1-36 04 02 «Промышленная электроника» дневной формы обучения.

УДК 621.394.14(075.8)
ББК 32.811я73

Лабораторная работа 3

Э Ф Ф Е К Т И В Н Ы Е   К О Д Ы

ЦЕЛЬ РАБОТЫ. Закрепление знаний по теории информации и методам эффективного статистического кодирования на примере эффективного кода Хафмена.

1. ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ ИНФОРМАЦИИ

1.1. Информация, сообщение, кодирование, сигнал

В общем случае под информацией понимают совокупность сведений о каких-либо сообщениях, явлениях или предметах, которые получает потребитель. В теории электросвязи информация - это сведения, являющиеся объектом хранения, передачи и преобразования. Информацию, представленную в форме, которая позволяет осуществлять ее преобразование с целью передачи, обработки и практического использования, называют сообщением.

Всякое сообщение является совокупностью сведений о состоянии какой-либо материальной системы, которые передаются человеком (или устройством), наблюдающим эту систему, другому человеку, (или устройству), не имеющему возможностей получить эти сведения из непосредственных наблюдений. Наблюдаемая материальная система вместе с наблюдателем представляет собой источник информации (корреспондент), а человек или устройство, которому предназначаются результаты наблюдения - получатель информации (абонент).

Источник информации может вырабатывать непрерывное или дискретное сообщения. Передача сообщений на расстояние осуществляется с помощью какого-либо материального носителя (бумага, магнитная лента и т.п.) или физического процесса (звуковых или электромагнитных волн, тока и.т.п.).

Физический процесс, отображающий (несущий) передаваемое сообщение, называется сигналом. Сигналы формируются путём изменения тех или иных параметров физического носителя по закону передаваемых сообщений. Этот процесс изменения параметров сигнала в соответствии с передаваемым сообщением называется модуляцией и осуществляется в системах связи с помощью модуля-тора. Любое преобразование сообщения в определённый сигнал путём установления между ними однозначного соответствия называют в широком смысле кодированием. Кодирование может включать в себя процессы преобразования и дискретизации непрерывных сообщений (аналого - цифровое преобразование), мо-дуляцию (манипуляцию в цифровых системах связи) и непосредственно кодирование в узком смысле слова. Обратная операция называется декодированием.

Рассмотрим передачу дискретных сообщений. Последовательный процесс преобразования сообщения в кодирующем устройстве (кодере) может быть представлен структурной схемой, показанной на рис.1

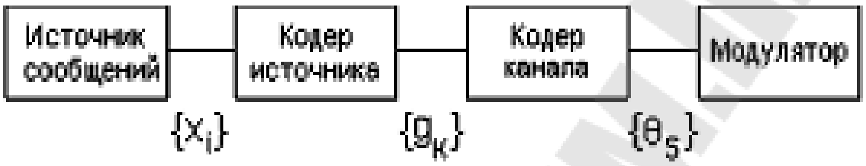

Рис. 1

Источник дискретных сообщений в общем случае характеризуется ансамблем

$X = \{x_1,...,x_i,...,x_{mi}\}$ сообщений, представляющих собой конечное число символов x_i. Совокупность символов $x_1,...,x_i,...,x_{mi}$ называется алфавитом источника сообщений, а число различных символов mi - является объёмом алфавита источника сообщений. Для полного описания источника сообщений необходимо задать вероятности появления символов $p(x_1),..., p(x_i),..., p(x_{mi})$, причём их сумма равна 1. В частном случае символами алфавита источника могут быть буквы. Наиболее часто повторяющаяся буква русского языка - "О". Её вероятность появления $p(x) = 0.09$ означает, что на 1000 букв русского текста приходится в среднем 90 букв "О". Наиболее редко встречающаяся в тексте - буква "Ф", для которой $p(x) = 0.002$. Вероятность появления промежутка между словами в тексте (пробел) наибольшая и равна 0.175. Отметим, что по этой причине на клавиатуре пишущих машинок клавиша "пробел", которой наиболее часто пользуются, выполнена не в виде кнопки, как для других букв, а в виде протяжённой планки, размещаемой ниже всей клавиатуры.

Кодер источника, который иногда может и отсутствовать, служит для представления сообщений в более удобной для передачи и компактной форме без по-тери информации. Кодер источника имеет свой алфавит $G = \{g_k\}$, k = 1,2,...,m_k , где m_k - объём алфавита кодера источника. Например, русские буквы алфавита источника могут в кодере источника перекодироваться в двоичный код с алфавитом 0 и 1 и, следовательно, с объёмом алфавита m = 2 (например первичный код МТК - 2). С помощью кодера источника возможно устранение избыточности источника сообщений путём применения эф-фективного статистического кодирования.

Кодер канала может иметь свой алфавит $\theta = \{\theta_S\}$, s = 1, 2, . . ., m_S. С помощью кодера канала может вводиться избыточность при применении корректирующих кодов в целях повышения помехозащищённости системы связи. Таким образом, в процессе преобразования сообщения в сигнал операция кодирования позволяет в итоге уменьшить влияние различных помех и искажений на передачу сообщений.

В процессе кодирования сообщений могут выполняться следующие операции:

- преобразование сообщений из одной формы в другую, например, непрерывных в дискретные (натуральное, первичное кодирование);
- устранение естественной избыточности источника сообщений (эффективное или статистическое кодирование);
- введение специально рассчитанной искусственной избыточности в сообщение (помехоустойчивое кодирование).

Соответствующие кодеры (рис. 1) можно построить либо для каждой из указанных выше операций отдельно, либо объединить их в единое устройство.

1.2. Информационные характеристики системы передачи сообщений

Такие понятия теории информации, как количество информации, передаваемое по каналу связи, энтропия, избыточность, пропускная способность канала являются интегральными оценками эффективности системы связи. Теория указывает потенциальные возможности системы связи, которые надо стремиться реализовать на практике.

Мера количества информации

Мера количества информации должна отражать сущность работы систем передачи сообщений и служит основой для сравнения их между собой. Для систем связи в большинстве случаев не имеет значение конкретное содержание сообщений, их ценность, важность, истинность или ложность. Системы передачи информации можно уподобить почте, для которой важен только сам факт отправления письма, а содержание пересылаемых писем никак не учитывается. Поэтому понятие количества информации, применяемое для характеристики технических систем, значительно беднее, чем используемое нами в повседневной жизни.

Тем не менее мера количества информации должна согласовываться с интуитивными представлениями о существенных сторонах сообщений. При этом разумно руководствоваться следующими соображениями:

- чем длиннее сообщение, тем большее количество информации оно должно содержать;
- количество информации в сообщении тем больше, чем больше число возможных сообщений;
- количество информации должно обладать свойством аддитивности, т.е. количество информации, содержащееся, например, в двух независимых сообщениях, должно равняться сумме количества информации, переносимой каждым сообщением;
- большее количество информации несут маловероятные сообщения (сенсации, которые трудно угадать).

Понятие количества информации прошло следующие этапы эволюции:

1). Сообщение состоит не из одного, а из многих символов (букв, знаков, элементов). Число возможных элементов определяется объёмом m соответствующего алфавита (m_i, m_k , m_S - i - источник, k - кодер, s - сигнал), а число элементов в сообщении - n. При выборе первого элемента сообщения производится выбор из m возможных элементов. При выборе второго делается выбор из того же числа m элементов, но число возможных комбинаций выбора двух элементов составляет уже m^2(при m = 2, например, "0" и "1", возможных комбинаций - 4: "00", "01", "10" и "11"). Если же сообщение содержит n элементов, то число различных сочетаний этих элементов равно:

$$N = m^n \qquad (1)$$

Значение N определяет число возможных сообщений. Оно и может служить мерой количества информации. Однако мера N не обладает свойством аддитивности. Действительно, количество информации в сообщении из n символов не равно сумме количеств информации из n_1 и n_2 символов, так как $m^n = m^{n1} + m^{n2}$, если $n_1 + n_2 = n$.

2). Для удовлетворения условию аддитивности можно выбрать в качестве меры количества информации не само число N, а некоторую его функцию $J = f(N)$. Р. Хартли в 1928 году предложил логарифмическую меру количества информации:

$$J = \log(N) = n \times \log(m) \qquad (2)$$

Эта мера обладает свойством аддитивности, а именно:

$$n \times \log(m) = n1 \times \log(m) + n2 \times \log(m), \quad \textit{если } n1 + n2 = n.$$

Основание логарифма в (2) не имеет существенного значения. Широко пользуются логарифмом по основанию 2 (причём обозначение "2" опускается). В этом случае количество информации измеряется в двоичных единицах (дв.ед.) или битах. Однако мера (2) не удовлетворяет четвёртому интуитивному требованию, так как не учитывается зависимость количества информации, содержащейся в сообщении, от вероятности появления сообщения. В то же время эта вероятность характеризует неожиданность данного сообщения для получателя.

3. К. Шеннон учёл требуемую зависимость и предложил определять количество информации, содержащееся в сообщении x_i (i = 1, 2, . . . , m_i) и относящееся к выбору данной буквы x алфавита источника, в виде:

$$J(x_j) = \log[\frac{1}{p(x_j)}] = -\log[p(x_j)] \qquad (3)$$

где $p(x_j)$ - вероятность появления сообщения x_j, причём сумма всех $p(x_j)=1$.

Как следует из (3), количество информации, содержащееся в сообщении, тем больше, чем меньше вероятность этого сообщения. Такая зависимость соответствует интуитивным представлениям об информации. Действительно, сообщения, ожидаемые с большей вероятностью, легко угадываются получателем, а достоверные сообщения, вероятность которых равна 1, вообще не содержат информации,так как всегда могут быть предсказаны точно (очевидно,если $p(x_j)=1$, то $J(x_j)=0$. Наоборот, сообщения, являющиеся сенсациями, имеют малую вероятность появления и их трудно предсказать, поэтому они содержат больше информации.

Количество информации, определяемое (3), является случайной величиной, принимающей значение $J(x_i)$ с вероятностью $p(x_i)$ в зависимости от появления буквы x_i в сообщении источника. Однако при передаче больших массивов сообщений важно не количество информации в одном конкретном символе $J(x_i)$, а количество информации, усреднённое по всем возможным сообщениям, содержащим n символов. Такой мерой количества информации является математическое ожидание (среднее значение) случайной величины $J(x_i)$, содержащей n символов (букв), усреднённое по всему ансамблю *X*:

$$J(X)=n\cdot\sum_{i=1}^{m_i} p(x_i)\cdot J(x_i)=-n\cdot\sum p(x_i)\cdot\log p(x_i) \qquad (4)$$

Это соотношение носит название формулы Шеннона. Для равновероятных сообщений $p(x_i)=1/m$ меры информации по Хартли (2) и по Шеннону (4) совпадают:

$$J(X)=-n\cdot\sum_{i=1}^{m} 1/m\cdot\log(1/m)=n\cdot\log m$$

Поэтому меру Шеннона (4) можно рассматривать как обобщение меры Хартли на ансамбль сообщений с распределением вероятностей, отличающимся от равномерного.

Энтропия источника дискретных сообщений

Для характеристики источника сообщений более удобной величиной является средняя величина (математическое ожидание) количества информации, содержащегося в одном символе (букве) сообщения. Эта величина называется энтропией источника сообщений. В случае отсутствия статистической связи между символами, энтропия источника равна:

$$H(X) = J(X)/n = -\sum_{i=1}^{m} p(x) \cdot \log p(x). \qquad (5)$$

Понятие энтропии (от греческого "эн-тропе" - обращение) распространилось на ряд областей знания. Энтропия характеризует неопределённость каждой ситуации. Энтропия в термодинамике определяет вероятность теплового состояния вещества (закон Больцмана), в математике - степень неопределённости ситуации или задачи, в теории информации она характеризует способность источника "отдавать" информацию. Приобретение информации сопровождается уменьшением неопределённости, поэтому количество информации можно измерять количеством исчезнувшей неопределённости, т.е. энтропией. Энтропию называют также информационной содержательностью сообщения.

Анализируя выражение (5), можно отметить некоторые свойства энтропии дискретной случайной величины.

1).Энтропия источника является величиной вещественной и положительной - $H(x) \geq 0$. Энтропия равна 0 в случае, когда отсутствует возможность выбора, т.е. когда величина X может принимать только одно значение с вероятностью $p(x) = 1$. В передаче такого сообщения нет смысла, поскольку результат заранее известен получателю. Источники с малой энтропией не являются информативными. Они выдают знаки, которые с большой вероятностью известны получателю. В этом смысле энтропия источника характеризует его информационную ёмкость.

2).Энтропия случайной величины, имеющей всего два значения x_1 и x_2, не превышает 1. При объёме алфавита источника $m_i = 2$ и одинаковой вероятности сообщений $p(x_1) = p(x_2) = 0.5$ энтропия достигает максимального значения $H_{\max}(x) = 1$ дв.ед. Следовательно, в качестве единицы измерения информации (дв.ед., бит) взята информация, содержащаяся в одном из двух равновероятных сообщений.

3).Максимальная энтропия источника $H_{\max}(x)$достигается лишь в случае равных вероятностей выбора букв алфавита, т.е. когда $p(x_i) = 1/m$, (i= 1,2,...,m), тогда:

$$H_{\max}(x) = -n \cdot \sum_{i=1}^{m} 1/m \cdot \log(1/m) = \log m \qquad (6)$$

Такой источник называют идеальным (оптимальным), так как каждый его символ несет максимальное количество информации.

В теории информации доказывается, что энтропия источника зависимых сообщений всегда меньше энтропии источника независимых сообщений при том же объёме алфавита и тех же безусловных вероятностях сообщений.

Если источник выдаёт последовательность букв из алфавита объёмом m =32 и буквы выбираются равновероятно и независимо друг от друга, то энтропия источника (6) $H_{\max}(x) = \log m = 5$ бит. Однако таким источником могла бы быть обезьяна, нажимающая в хаотическом порядке клавиши пишущей машинки (идеальный источник!).

Если буквы передаются не хаотически, а составляют связный, например, русский текст, то появление их неравновероятно (см. выше - вероятность появления буквы "О" в 45 раз больше, чем буквы "Ф"), и, главное, буквы в тексте зависимы. Так, после гласных не может появиться "Ь", мала вероятность сочетания более трёх согласных подряд, вероятность последовательности, не образующей осмысленных слов (идеальный источник), практически равна нулю. Расчёты показывают, что для текстов русской художественной прозы энтропия оказывается менее 1.5 бит на букву. Еще меньше, около 1 бита на букву, энтропия поэтических произведений, так как в них имеются дополнительные вероятностные связи, обусловленные ритмом и рифмами. Слово, рифмуемое с окончанием предыдущей стихотворной строки, легко угадывается без произнесения или чтения его, и поэтому информации не несет ($H(x) = 0$). Энтропия телеграмм обычно не превышает 0.8 бит на букву, поскольку их тексты довольно однообразны (особенно поздравительных).

Количественно эта характеристика источника оценивается его избыточностью.

Избыточность источника сообщений.

Абсолютная избыточность источника определяется формулой

$$\chi_a = H_{\max}(x) - H(x). \tag{7}$$

Чаще используется понятие относительной избыточности, которую и называют избыточностью источника:

$$\chi = \frac{H_{\max}(x) - H(x)}{H_{\max}(x)} = 1 - \frac{H(x)}{H_{\max}(x)} = 1 - \mu, \tag{8}$$

где $\mu = H(x) / H_{\max}(x)$ - относительная энтропия.

Избыточность $0 \le \chi \le 1$ и учитывает как взаимосвязь (корреляцию) символов в передаваемой последовательности, так и неопределённость каждого символа. Она является важной характеристикой источника, так как указывает, насколько можно сократить число символов и довести его до минимального $n_{\min}$ в последовательности данного источника, если то же количество информации будет передаваться последовательностью, составленной из равновероятных и независимых символов, т.е. при $H(x) = H_{\max}(x)$. Действительно, для данного (реального) источника количество информации, содержащееся в последовательности из n символов, равно (5) $J = n \cdot H(x)$, а для идеального $J = n_{\min} \cdot H_{\max}$ Приравнивая количества информации этих источников, получим

$$H(x) / H_{\max} = \frac{n_{\min}}{n}$$

или избыточность кода источника

$$\chi = \frac{n - n_{\min}}{n} = 1 - \frac{n_{\min}}{n} = 1 - \mu \tag{9}$$

где отношение $\mu = \frac{n_{\min}}{n}$ получило название коэффициента сжатия, равного относительной энтропии.

Таким образом, источник с избыточностью $\chi = 0$ формирует по-

следовательности сообщений, число n символов в которых больше минимально необходимого $n_{\min}$ для передачи данного количества информации.
Установлено, что избыточность текстов на русском и английском языках $\chi = 0.7$, т.е. объём книги и другой печатной продукции примерно в 3.3 раза больше, чем это необходимо для отображения содержащейся в ней информации (при $\chi = 0.7$ значение $\mu = 0.3 = 1/3.3$).

Однако это не даёт повод утверждать, что такая избыточность бесполезна. Избыточность текста обеспечивает высокую достоверность передачи информации, позволяет легко находить опечатки и исправлять ошибки. В частности, получатель телеграммы догадывается о её подлинном содержании даже при нескольких ошибочно переданных буквах. Отметим, что именно необходимость разговаривать при воздействии акустических помех явилась причиной того, что все национальные языки в процессе своего возникновения и развития оказались избыточными, и значение избыточности для всех языков близко к $\chi = 0.7 - 0.9$.

В технических приложениях естественную избыточность источников трудно использовать для повышения помехоустойчивости систем связи. Лишние символы в последовательности сообщений часто нежелательны, так как увеличивают время передачи информации, а при её хранении требуют дополнительной памяти в запоминающих устройствах. Вопросам устранения избыточности сообщений уделяется большое внимание и с этой целью осуществляют статистическое (эффективное) кодирование дискретных сообщений, в частности, применяют коды Шеннона - Фано и Хафмена.

Отметим, что для повышения помехозащищенности канала связи целесообразно вводить избыточность снова, что делается при помехоустойчивом кодировании.

Производительность источника

Производительность источника *H'(x)* есть среднее количество информации, создаваемое источником в единицу времени:

$$H'(X) = \lim_{T \to \infty} \frac{H(X_T)}{T}, \qquad (10)$$

где $H(X_T)$ - энтропия случайной последовательности, заданной на интервале *T*.

При наличии кодера источника выражение (10) преобразуется к виду:

$$H'(X) = \frac{H_1(X)}{\overline{\tau_X}} = V_X H_1(X), \qquad (11)$$

где $\overline{\tau_X}$ - средняя длительность одного символа (разряда) кодового слова, $V_X = 1/\overline{\tau_X}$ – скорость формирования символов кодером источника.

Из (11) следует весьма важный вывод о том, что источник может генерировать сообщения с большой скоростью, но, тем не менее, его производительность с информационной точки зрения будет чрезвычайно низкой, если $H_1(X) \ll 1$. Причиной этого является избыточность источника.

Различие понятий производительности и скорости формирования символов объясняется тем, что количество информации характеризует не сам факт появления сообщения, а определённое его свойство - степень его неожиданности, нетривиальность выбора этого сообщения из множества других.

Призводительностью источника можно управлять, изменяя длительность символов $\overline{\tau_X}$. Поэтому различают неуправляемые и управляемые источники. Для неуправляемых источников производительность - постоянная величина. Так, телеметрические датчики обычно выдают информацию с постоянной скоростью и могут служить примером неуправляемых источников с фиксированной скоростью создания сообщений. Для управляемых источников формирование символов сообщений происходит по внешним командам и, следовательно, длительность символа может изменяться. Например, чтение чисел из запоминающего устройства осуществляется импульсами, интервал между которыми определяется возможностями и быстродействием периферийных устройств. Очевидно, что производительность управляемого источника может меняться в широких пределах.

Производительность источника является основной характеристикой при решении задач согласования источника с каналом связи.

1.3. Эффективное кодирование дискретных сообщений

Основной целью эффективного (статистического) кодирования является преобразование сообщения в сигнал с меньшей, чем у сообщения, избыточностью (в пределе - без избыточности). Сигналы без избыточности имеют максимальную удельную энтропию. Для передачи информации с помощью таких сигналов требуется минимальное количество символов. Поэтому такие коды называют эффективными, а также экономными и оптимальными. В результате эффективного кодирования скорость передачи информации по дискретному каналу связи может быть приближена к его пропускной способности. В этом случае говорят о согласовании источника с каналом.

Вопросы статистического кодирования сообщений рассмотрим в предположении, что передача ведётся по каналу связи в отсутствии помех. В этом случае принятый за время T сигнал Y_T совпадает с переданным S_T, равны их энтропии $H(S_T) = H(Y_T)$. Скорость передачи информации

$$R = \lim_{T \to \infty} \frac{H(S_T)}{T} = \frac{H_1(S)}{\tau} \qquad (12)$$

где $H_1(S)$ удельная энтропия сигнала; τ - длительность символов сигналов неуправляемого источника (τ = const).

При однозначном преобразовании сообщения в сигнал $H(X_T) = H(S_T)$ и скорость передачи может быть выражена через удельную энтропию $H_1(S)$)сообщения:

$$R = \lim_{T \to \infty} \frac{H(S_T)}{T} = \lim_{T \to \infty} \frac{H(X_T)}{T} = \frac{H_1(X)}{\tau} = H'(X), \qquad (13)$$

т.е. скорость равняется производительности источника - (11).

Пропускная способность канала C, характеризующая его потенциальные возможности, определяется как верхняя граница (или максимум) скорости передачи информации R. Для дискретного канала без помех максимальное значение скорости достигается при равновероятных и независимых символах сигнала (6):

$$C = \max_{p(x)} R = \max \frac{H_1}{\tau} = \frac{H_1(S) X_{\max}}{\tau} = \frac{\log m_S}{\tau}, \qquad (14)$$

где m_S - алфавит кодера канала (сигнала).

В частности, для двоичного канала $m_S = 2$ и $C = 1/\tau$ численно совпадает со скоростью манипуляции символов в канале. Полное согласование источника с каналом достигается при $R / C = 1$, а качество согласования определяется отношением:

$$\frac{R}{C} = \frac{H_1(S)}{\log m_S} = 1 - \chi_S \qquad (15)$$

Где $\chi_S = 1 - H_1(S) / \log m_S$ - избыточность сигнала по аналогии с (8).

При заданном канале отношение *R/C* полностью определяется избыточностью сигнала и его удельной энтропией $H_1(S)$.

Чтобы $R / C \to 1$, следует выбирать такой способ кодирования сообщений источника, при котором $H_1(S) \longrightarrow \log m_S$, т.е. в результате кодирования должна получаться последовательность, составленная из равновероятных и независимых символов.

Основная теорема кодирования

Вопрос о возможности передачи информации со скоростью, равной пропускной способности канала без помех, решается положительно при применении безизбыточного кодирования (см. (15) - при $\chi_S = 0$). Это утверждение доказывается одной из основных теорем теории информации, которая называется теоремой кодирования для источника. Поскольку при этом предполагается, что последовательность кодовых символов принимается без ошибок, то эту теорему называют также теоремой кодирования для канала без помех.

Одна из возможных формулировок этой теоремы следующая. Если производительность источника $H'(X) = C - \varepsilon$, где C - пропускная способность канала связи, а $\varepsilon > 0$ - сколь угодно малая величина, то существует способ кодирования, обеспечивающий передачу всех сообщений, вырабатываемых источником, со скоростью $R = H'(X) = C - \varepsilon$. Если $H'(X) > C$, то длительная передача всех сообщений невозможна.

Избыточность источника возникает за счёт:

–неравной вероятности набора символов, составляющих алфавит источника;

–зависимости выбора последующего символа от предыдущего (так, в связном русском тексте после гласных не может появиться "ъ", мала вероятность сочетания более трёх согласных подряд и т.п.).

Устранение избыточности достигается следующим образом.

1-й этап - применяется укрупнение алфавита источника для устранения статистической связи между соседними символами (кодируются не отдельные буквы, а целые слова текста), при этом уменьшается неравновероятность букв укрупнённого алфавита.

2-й этап - при последующем кодировании используются неравномерные коды; при этом наиболее вероятные буквы ранее укрупнённого алфавита источника передаются меньшим количеством символов.

Теорема кодирования является теоремой существования, т.е. она доказывает, что оптимальные (эффективные) коды существуют, но не даёт указаний о том, как построить такие коды.

В настоящее время разработано большое количество эвристических приёмов, позволяющих осуществить статистическое кодирование и найти код, близкий к оптимальному. Однако основные свойства и особенности, которыми должны обладать такие эффективные коды, следуют из теоремы кодирования.

1).Для обеспечения минимальной средней длины кодового слова избыточность должна быть сведена к минимуму (желательно к нулю). Для этого эффективный код должен состоять из кодовых слов, в которых все символы равновероятны и статистически независимы. Это позволяет уравнять скорость передачи с пропускной способностью канала связи, что и является целью безизбыточного кодирования.

2).Ни одна из кодовых комбинаций не должна получаться из другой, более короткой, путём добавления новых символов. Эффективные коды не требуют разделительных символов (маркеров) и при этом должно выполняться их однозначное декодирование. Коды, удовлетворяющие этому условию, называются <u>префиксными кодами</u>, так как ни одно кодовое слово не является передней частью ("префиксом" - приставкой) другого кодового слова.

3).Эффективные коды являются неравномерными, т.е. для передачи разных символов сообщения m_i используются кодовые комбинации разной длины. Наиболее вероятные сообщения кодируются самыми короткими кодовыми словами, вследствие чего средняя длина кодового слова в сообщении уменьшается, что и позволяет решить задачу равенства скорости передачи и пропускной способности канала.

При неравномерном эффективном кодировании средняя длина кодового слова n_{cp} определяется выражением:

$$n_{cp} = \sum_{k=1}^{m_i} n_k p(x_k), \qquad (16)$$

где $p(\chi_k)$ - вероятность появления сообщения (кодового слова), причём их сумма равна 1; $n_л$ - длина кодовых слов x_k $(k = 1, 2, \ldots, m_i)$

По аналогии с выражением (9), где предполагалось применение равномерных кодов с постоянной длиной кодовых слов (*n* = *const*), при использовании эффективных неравномерных кодов (*n* = *var*) избыточность кода источника

$$\chi_k = \frac{n_{cp} - n_{\min}}{n_{cp}} = 1 - \frac{n_{\min}}{n_{cp}} \qquad (17)$$

Очевидно, когда $n = n_{\min}$, что эквивалентно равенству $H_{\max}(X) = H(X)$, избыточность кода χ_k равна нулю и при применении эффективных кодов обеспечивается полное согласование источника сообщений с каналом (15). При этом энтропия источника *H(X)* является оценкой среднего числа n_{cp} двоичных символов, требуемых для кодирования сообщений.

Процедуру построения эффективного кода, близкого к оптимальному, предложили практически одновременно Шеннон и Фано (код Шеннона – Фано.

В данной лабораторной работе студенты знакомятся с процедурой построения эффективных кодов, предложенной Хафменом. При малом алфавите источника и неравновероятных символах x_i выгодно кодировать не отдельные символы, а целые блоки из нескольких символов (букв). В этом студенты убеждаются, исследуя в лабораторной работе метод кодирования Хафмена.

Метод кодирования Хафмена

Д. А. Хафменом был предложен систематический метод кодирования, который всегда приводит к получению оптимального множества кодовых слов для кодирования данного множества сообщений.

Для дискретных систем с двоичным алфавитом кодера (m=2) методика построения кода Хафмена сводится к следующей процедуре.

1. Все $m_i = M$ сообщений (буквы алфавита источника) выписываются в порядке убывания вероятностей $p(x_i)$ (см. таблицу).

2. Две последние буквы алфавита, имеющие наименьшие вероятности $p(x_{M-1})$ и $p(x_M)$, группируются вместе и объединяются в одну вспомогательную букву, которой приписывается суммарная вероятность

$$p_{\Sigma} = p(x_{M-1}) + p(x_M)$$

3. Вероятности букв, не участвовавших в объединении, и полученная суммарная вероятность снова располагаются в порядке убывания вероятностей (в следующем столбце таблицы). Объём нового алфавита таким образом уменьшается на единицу: *M-1*.

4. Производят второе укрупнение алфавита, состоящего уже из *M-1* символов, путём объединения двух символов с наименьшими вероятностями и вычисляют их общую вероятность. Получают новый алфавит объёмом *M-2*.

5. Упорядочивают по вероятности символы этого нового алфавита.

6. Образуют последовательность укрупнённых алфавитов путём последовательного повторения операций пунктов 4 и 5, пока в ансамбле не останется единственное сообщение с вероятностью, равной 1 (шаговая процедура, записываемая в столбцах таблицы).

7. Проведя линии, соединяющие символы при последовательном укрупнении алфавита, получают так называемое кодовое дерево, концы ветвей которого являются символами исходного алфавита источника сообщений. Приписывая ветвям дерева, исходящим из каждого промежуточного узла, различные символы алфавита кодера (0 или 1), получают кодовые слова, соответствующие кодируемым сообщениям источника.

Методика поясняется примером, представленным на рисунке 2, где для алфавита источника с объёмом $M = 8$ приняты произвольные значения вероятностей $p(x_i)$, но

$$\sum_{i=1}^{M} p(x_i) = 1$$

x_i	Вероят-ности	Шаговая процедура (кодовое дерево)							Кодовые слова
	$p(x_i)$	1	2	3	4	5	6	7	
x_1	0,729	0,729	0,729	0,729	0,729	0,729	0,729	1	1
x_2	0,081	0,081	0,081	0,081	0,109	0,162	0,271		011
x_3	0,081	0,081	0,081	0,081	0,081	0,109			010
x_4	0,081	0,081	0,081	0,081	0,081				001
x_5	0,009	0,010	0,018	0,028					00011
x_6	0,009	0,009	0,010						00010
x_7	0,009	0,009							00001
x_8	0,001								00000

Рис. 2

Для составления кодовой комбинации, соответствующей данному сообщению x_i , необходимо проследить путь перехода сообщения по строкам и столбцам таблицы. Для наглядности кодовое дерево построено в поле таблицы. Целесообразно строить кодовое дерево, начиная с первого столбца таблицы, располагая ветви против группируемых попарно вероятностей $p(x_i)$ и соединяя их со зна-чением суммарной вероятности p_Σ, располагаемой в следующем столбце. Ветвям с большей вероятностью присваивается символ 1, а с меньшей – 0 (или наоборот). Такое последовательное ветвление продолжается до тех пор, пока ветвь не закончится узлом с вероятностью $p(x_i)$ каждой буквы алфавита источника. Отдельно кодовое дерево для алфавита источника, рассматриваемого в примере, приведено на рис. 3

Перемещаясь по кодовому дереву сверху вниз, можно записать для каждой буквы алфавита x_i соответствующую ей кодовую комбинацию

x1	x2	x3	x4	x5	x6	x7	x8
1	011	010	001	00011	00010	00001	00000

Код Хафмена при любом распределении вероятностей $p(x_i)$ даёт однозначный ансамбль набора кодовых слов, в то время как при коде Шеннона – Фано на выходной ансамбль кодовых слов влияет субъективный выбор границ последовательного разбиения алфавита на две группы.

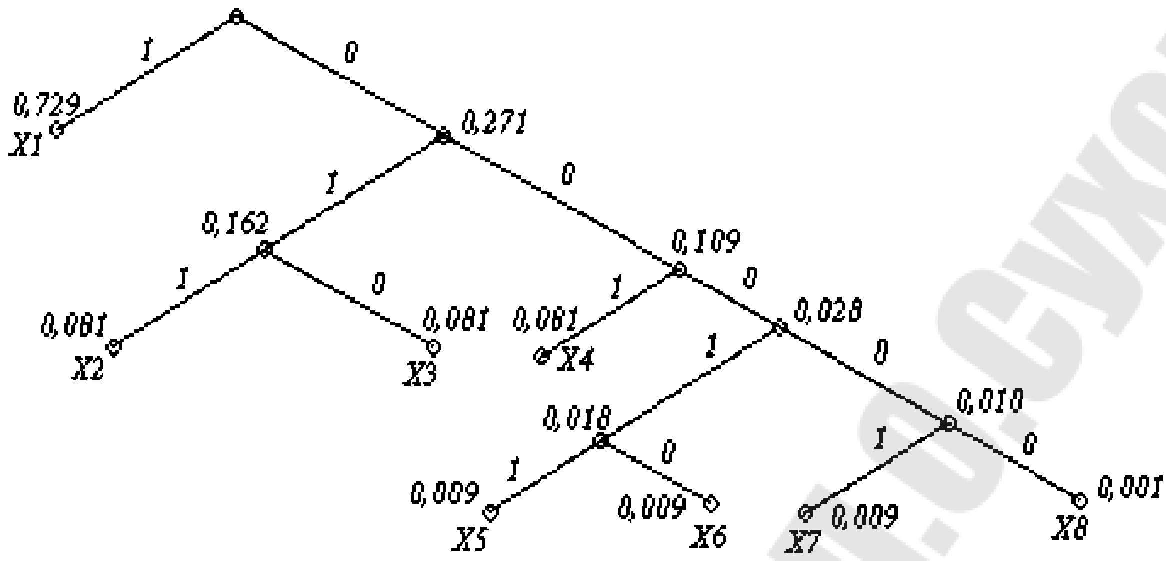

Рис. 3 Кодовое дерево по данным таблицы

Существенное преимущество кода Хафмена по сравнению с кодом Шеннона - Фано проявляется при применении кодов с основанием большим 2 ($M > 2$) и заключается в том, что методика Хафмена гарантирует однозначное построение кода с наименьшим для данного распределения вероятностей средним количеством символов на букву.

Достоинства и недостатки эффективных кодов

Кратко сформулируем перечисленные выше достоинства оптимальных эффективных кодов.

1. При эффективном кодировании, учитывающим вероятности появления букв алфавита источника сообщений, удаётся построить коды с максимальной удельной энтропией на символ.

2. Обеспечивается преобразование сообщения в сигнал с меньшей, чем у сообщения избыточностью (в пределе - без избыточности).

3. На передачу сообщения затрачивается минимальное количество символов.

4. Решается задача согласования источника сообщений с каналом связи, в результате чего скорость передачи информации может быть

приближена к пропускной способности канала.

5. Не требуется введения специальных разделительных символов (маркеров), как, например, в коде Морзе, для отделения одной кодовой комбинации от другой, так как ни одна комбинация эффективного кода не совпадает с началом другой, более длинной. Такое свойство кода называется "неприводимостью", и коды называются префиксными или кодами без запятой.

К недостаткам эффективных кодов можно отнести следующее.

1. Эффективные коды являются неравномерными, т.е. кодовые комбинации имеют различное количество символов. Если линия связи работает с постоянной скоростью передачи, то на выходе кодера необходимо буферное запоминающее устройство ("упругая задержка") для записи в него "пульсирующих" по длительности кодовых групп и последующего считывания в канал символов с постоянной скоростью. Аналогичная "упругая задержка" должна быть и на стороне приёма.

2. Наибольший эффект оптимальные коды дают при кодировании исходного сообщения длинными блоками, поскольку при этом достигается равновероятность и статистическая независимость блоков. Однако блочное кодирование вызывает необходимость накапливать слова алфавита источника, прежде чем поставить им в соответствие определённую кодовую группу эффективного кода. Это приводит к большим задержкам при передаче и приёме сообщений, что затрудняет (или исключает) применение эффективных кодов в системах, работающих в реальном масштабе времени. Эффективное кодирование (кодом Хафмена) применяется при записи информации на магнитные носители (системы архивации) и в системах факсимильной связи.

3. Существенным недостатком эффективных кодов является то, что они непомехозащищённые. Любая одиночная ошибка при приёме переводит передаваемую комбинацию в другую, не равную ей по длительности, что влечёт за собой неправильное декодирование целого ряда последующих кодовых групп. Такое специфическое влияние помех называется "треком ошибок" или пакетом ошибок. В чистом виде эффективное кодирование можно применять только для каналов без помех.

Таким образом, непосредственная передача сообщений при применении эффективных кодов по каналу связи с шумами приводит к недопустимо большим искажениям (потере информации). Однако, эффективное кодирование, устраняющее статистическую избыточность в передаваемом сообщении, наилучшим образом подготавливает непрерывную кодовую последовательность, полученную после первичного кодирования сообщений источника, к последующему помехоустойчивому кодированию с помощью корректирующих кодов в кодере канала (рис. 1). Целенаправленное введение избыточности при помехоустойчивом кодировании путём добавления дополнительных проверочных символов в кодовые информативные группы позволяет при декодировании обнаруживать и исправлять ошибки, вызванные помехами.

2. ПОРЯДОК ВЫПОЛНЕНИЯ РАБОТЫ

Для выполнения лабораторной работы студенты должны предварительно изучить раздел 1 настоящего методического руководства и получить у преподавателя допуск к работе.

Лабораторная работа выполняется с помощью математического пакета Mathcad.

Вероятности элементов x_1 и x_2 алфавита источника задаются преподавателем.

Перед выполнением лабораторной работы целесообразно подготовить таблицы 1-5.

После введения заданных студентами вероятностей $p(x_1))$ и $p(x_2)$ появления 2-х сообщений x_1 и x_2 осуществляется последовательное группирование символов алфавита источника сообщений в блоки с длиной ℓ =2, 3 и 4 символа. Каждой новой комбинации источника сообщений присваивается обозначение, соответственно, *Yi* , *Zi* ,*Qi* и рассчитывается вероятность появления этой комбинации. Далее выполняется операция упорядочения вероятностей по величине от большего значения к меньшему.

Студентам необходимо построить дерево и получить кодовые слова кода Хафмена для каждой из букв алфавита источника данного ансамбля. Результаты работы и расчеты характеристик кодов сводятся в приведенные ниже таблицы.

Таблица 1

Алфавит источника	Обозначения кодовых слов	Вероятность $P(S_i)$	Код Хафмена	Число символов
x_1 x_2	*X1* *X2*			

Таблица 2

Группировка по 2

Алфавит источника	Обозначения кодовых слов	Вероятность $P(S_i)$	Код Хафмена	Число символов
$x_1\ x_1$ $x_1\ x_2$ $x_2\ x_1$ $x_2\ x_2$	*Y1* *Y2* *Y3* *Y4*			

Таблица 3

Группировка по 3

Алфавит источника	Обозначения кодовых слов	Вероятность $P(S_i)$	Код Хафмена	Число символов
$x_1\ x_1\ x_1$ $x_1\ x_1 x_2$ $x_1 x_2\ x_1$ $x_2\ x_2\ x_2$	*Z1* *Z2* *Z3* *Z8*			

Таблица 4

Группировка по 4

Алфавит источника	Обозначения кодовых слов	Вероятность $P(S_i)$	Код Хафмена	Число символов
$x_1\ x_1\ x_1\ x_1$ $x_1 x_1\ x_1 x_2$ $x_1 x_1 x_2\ x_1$ ………… $x_2\ x_2\ x_2\ x_2$	Q1 *Q2* *Q3* ... Q16			

Таблица 5

ℓ	$p(x_k)$	*H(S)*	$H_1(S)$	$n_{\underline{min.c}}$	χ_u	n_{cp}	$n_{cp.c}$	$\frac{n_{\min.c}}{n_{cp.c}}$	χ_k	*R/C*
1 2 3 4										

ℓ – длина группирования символов

n_{cp} – средняя длина кодового слова

$n_{cp.c}$ – средняя длина на символ

$H(S)$ – энтропия источника

$H_1(S)$ – удельная энтропия

χ_u – избыточность источника

χ_k – избыточность кода

R/C – качество согласования источника с каналом

3. КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Дайте определения понятий "информация", "сигнал", модуляция", "кодирование", "объём алфавита".
2. Сравните три вида определения меры количества информации. Достоинства определения меры по Шеннону.
3. Поясните определение "энтропия". Перечислите свойства энтропии.
4. Дайте определение избыточности источника сообщений, коэффициента сжатия и коэффициента избыточности.
5. Поясните различие понятий производительности и скорости формирования символов источника сообщений.
6. Объясните, зачем нужно производить эффективное (статистическое) кодирование и в чём его суть.
7. Поясните связь между скоростью передачи информации и пропускной способностью дискретного канала.
8. Дайте определение теоремы кодирования для канала без помех. Какие свойства оптимальных кодов являются следствием этой теоремы?
9. Почему оптимальные коды называются "префиксными"? Какой факт подчёркивается этим названием?
10. Поясните преимущества блочного кодирования и его особенности.
11. Поясните процедуру кодирования по методу Хафмена. Назовите достоинства процедуры Хафмена.
12. Поясните процедуру кодирования по методу Шеннона-Фано. Назовите недостатки этого метода.
13. Перечислите достоинства эффективных кодов и возможности их применения.
14. Что такое "неприводимость" кода?
15. Что понимается под термином "упругая задержка"?
16. Как влияют помехи на декодирование сообщений при зффективном кодировании?
17. Определите избыточность алфавита двоичного источника, выдающего независимые сообщения "0" и "1", на выходе которого вероятность появления символа "0" равна p(0) = 0.2.

Литература

1.Клюев Л.Л. Теория электрической связи. Мн.:Техноперспектива, 2008.

2.Передача дискретных сообшений. Учебник для вузов /В.П.Шувалов, Н.В.Захарченко и др.; Под ред. В.П.Шувалова. М.: Радио и связь, 1990.

3. Клюев Л.Л. Теория электрической связи. Мн.: Дизайн ПРО, 1998.

4. Теория передачи сигналов. Учебник для вузов/ А.Г.Зюко и др. М.: Радио и связь, 1986.

5. Помехоустойчивость и эффективность систем передачи информации /А.Г.Зюко, А.И.Фалько, И.П.Панфилов и др.; Под ред. А.Г.Зюко. — М.: Радио и связь, 1985.

Содержание

Лабораторная работа 3. Эффективные коды……………………………3
1. Основные понятия теории информации……………………………...3
2. Порядок выполнения работы…………………………………………...22
3. Контрольные вопросы…………………………………………………...25
Литература……………………………………………………………….26

Щуплов Вячеслав Валентинович

ЭФФЕКТИВНЫЕ КОДЫ

ЛАБОРАТОРНЫЙ ПРАКТИКУМ
по курсу «Теория электросвязи»
для студентов специальности 1-36 04 02
«Промышленная электроника»
дневной формы обучения
В двух частях
Часть 2

Подписано в печать 27.01.11.
Формат 60x84/$_{16.}$ Бумага офсетная. Гарнитура «Таймс».
Ризография. Усл. печ. л. 1,63. Уч.-изд. л. 1,4.
Изд. № 78.
E-mail: ic@gstu.by
http://www.gstu.by

Отпечатано на цифровом дупликаторе
с макета оригинала авторского для внутреннего использования.
Учреждение образования «Гомельский государственный
технический университет имени П. О. Сухого».
246746, г. Гомель, пр. Октября, 48.