

4. ЭФФЕКТИВНОЕ КОДИРОВАНИЕ

4.1. Объем сигнала и емкость канала связи, условия их согласования

В разделе 2 обсуждены вопросы согласования дифференциальных характеристик источника дискретной информации и предоставленного дискретного канала связи в терминах потока информации, выводимой из ИДИ, и пропускной способности КС. В данном разделе эта задача решается на уровне интегральных характеристик сигнала и канала в виде *объема сигнала и емкости канала связи*.

Сигнал как *транспортная модель* сообщения (информации) имеет «габаритные размеры», характеризующие *объем сигнала*, аналитически определяемый выражением

$$V_c = T_c F_c W_c = T_c F_c \log_2 \left(1 + \frac{P_c}{P_n} \right), \quad (4.1)$$

где T_c – временная длительность сигнала, F_c – эффективный спектр сигнала, определяемый эффективным спектром элементарного сигнала кода и типом модуляции. В (4.1) компонент $W_c = \log_2 \left(1 + \frac{P_c}{P_n} \right)$

именуется *логарифмическим превышением сигнала над помехой*, в котором P_c – мощность сигнала, P_n – мощность помехи, сопровождающей процесс формирования сигнала.

Аналогичным образом канал связи как *транспортная среда* характеризуется «габаритными размерами», определяющими *емкость канала связи*, аналитически задаваемую выражением

$$V_k = T_k F_k W_k = T_k F_k \log_2 \left(1 + \frac{P_c}{P_n} \right), \quad (4.2)$$

где T_k – длительность интервала времени, на который предоставлен канал связи, F_k – эффективная полоса пропускания канала связи, которая может быть определена аналитически или экспериментально по амплитудной частотной характеристике *четырёхполюсника*, который представляет собой канал связи. В (4.2) компонент $W_k = \log_2 \left(1 + \frac{P_c}{P_n} \right)$ именуется *логарифмическим превышением сигнала*

над помехой, в котором P_c – мощность сигнала, фиксируемая в канальной среде, P_n – мощность помехи в канальной среде.

Нетрудно понять, что передача сигнала по предоставленному каналу связи *возможна* только тогда, когда размеры *транспортируемого средства (сигнала)* не превышают размеров *транспортной среды (канала связи)*.

Таким образом *необходимым* условием *согласования* сигнала с предоставленным каналом связи является выполнение неравенства

$$V_c \leq V_k. \quad (4.3)$$

Достаточными условиями *согласования* сигнала с предоставленным каналом связи является выполнение неравенств

$$T_c \leq T_k, F_c \leq F_k, W_c \leq W_k. \quad (4.4)$$

И наконец, условием *эффективного* использования предоставленного канала связи является выполнение равенства

$$V_c = V_k. \quad (4.5)$$

Конструктивным инструментом *согласования* сигнала с предоставленным каналом связи путем уменьшения объема сигнала V_c за счет уменьшения компонента T_c является использование возможностей *эффективного кодирования*.

Теорема К.Шеннона об эффективном кодировании символов ИДИ (основная теорема К.Шеннона)

Пусть источник дискретной информации (ИДИ) генерирует алфавит $X = \{x_i : p(x_i); i = \overline{1, n}\}$, составленный из n дискретных *статистически независимых* символов, характеризующийся

энтропией $H(X) = -\sum_{i=1}^n p(x_i) \log_2 p(x_i)$, тогда:

1. существует такой способ кодирования символов, при котором среднее на символ число двоичных разрядов кода

$l_{cp}\{K(X)\} = \sum_{i=1}^n p(x_i) l\{K(x_i)\}$ будет сколь угодно близким к энтропии $H(X)$ ИДИ;

2. не существует такого способа кодирования, при котором среднее на

символ число двоичных разрядов кода $l_{cp}\{K(X)\} = \sum_{i=1}^n p(x_i) l\{K(x_i)\}$

будет меньше энтропии $H(X)$ ИДИ. \square

Доказательство. Доказательство теоремы использует связь пропускной способности канала связи C с энтропией ИДИ $H(X)$ и максимальной скорости вывода символов $v_{\max}(x)$ из ИДИ, а также определение пропускной способности КС. Ситуация вывода сообщений, составленных из символов, из ИДИ и передача их по двоичному КС контролируется на бесконечном интервале T . Тогда может быть записана цепочка равенств с использованием обозначений n_c – число выведенных из ИДИ символов за интервал T , $l\{K(x_i)\}$ – длина (число разрядов) двоичного кода символа $x_i (i = \overline{1, n})$ и $\{K_v(x_i)\}$ – множество методов двоичного кодирования символа $x_i (i = \overline{1, n})$

$$H(X) = \frac{C}{v_{\max}(x)} = \left\{ \lim_{T \rightarrow \infty} \frac{\log_2 N(T)}{T} \right\} \frac{T}{n_c} = \lim_{T, n_c \rightarrow \infty} \frac{\log_2 N(T)}{n_c} = \frac{\sum_{i=1}^n l\{K(x_i)\}}{n_c} = \quad \blacksquare (4.6)$$

$$= \min_{\{K_i(x_i)\}} l_{cp}\{K(X)\}$$

Наиболее распространенными алгоритмами *эффективного* кодирования, основанного на использовании основной теоремы К.Шеннона, являются:

1. Алгоритм К.Шеннона – Р.Фэно,
2. Алгоритм Д. Хаффмэна.

Следует сказать, что содержательный момент этих алгоритмов, несмотря на процедурные различия, позволяющий построить эффективный двоичный код, средняя на символ длина которого максимально приближена к энтропии источника, состоит в том, что символам с большей вероятностью появления на выходе ИДИ ставятся в соответствие коды меньшей длины, а символам с меньшей вероятностью – коды большей длины. Таким образом, эффективное кодирование реализуется в классе *неравномерных* кодов.

Эффективные коды принято характеризовать двумя показателями кода: *избыточность* и *эффективность*

Определение 4.1. *Избыточностью* эффективного кода называется показатель D_k , определяемый выражением

$$D_k = \frac{l_{cp}\{K(X)\} - l_{\min}\{K(X)\}}{l_{cp}\{K(X)\}} = \frac{\sum_{i=1}^n p(x_i) l\{K(x_i)\} - H(X)}{\sum_{i=1}^n p(x_i) l\{K(x_i)\}}. \quad \square (4.7)$$

Определение 4.2. *Эффективностью* эффективного кода называется показатель η , определяемый выражением

$$\eta = \frac{l_{\min}\{K(X)\}}{l_{cp}\{K(X)\}} = \frac{H(X)}{\sum_{i=1}^n p(x_i) l\{K(x_i)\}}. \quad \square (4.8)$$

Нетрудно видеть, что *избыточность* и *эффективность* эффективного кода связаны соотношением

$$D_k = 1 - \eta.$$

4.2. Алгоритм эффективного кодирования К.Шеннона – Р.Фэно

Алгоритм эффективного кодирования К.Шеннона – Р.Фэно в своей реализации опирается на табличную форму представления информации об источнике дискретной информации и процесса

формирования эффективных кодов каждого из символов, дополненный процедурой вычисления значения средней на символ длины полученных кодовых комбинаций с последующей оценкой близости полученного значения к энтропии источника. А Алгоритм эффективного кодирования К.Шеннона – Р.Фэно может быть представлен в следующем виде.

АЛГОРИТМ

эффективного кодирования К.Шеннона – Р.Фэно символов источника дискретной информации

1. Задать источник дискретной информации, описав его в стандартной форме $X = \{x_i : p(x_i); i = \overline{1, n}\}$ и охарактеризовав его значением энтропии $H(X)$, вычисляемой в силу определения

$$H(X) = -\sum_{i=1}^n p(x_i) \log_2 p(x_i);$$

2. Составить таблицу формирования эффективного кода символов ИДИ, имеющую $(n + 1)$ строку и не более $(n + 3)$ столбцов;

3. Выписать в порядке убывания значений вероятностей $p(x_i); i = \overline{1, n}$ символы $x_i; i = \overline{1, n}$, разместив их во втором столбце, а значения вероятностей их появления на выходе ИДИ – в третьем;

4. Провести *первый* шаг процедуры эффективного кодирования, для чего символы разделить на две группы с возможно *равными* суммами вероятностей символов групп, после чего *всем* символам верхней группы в *старший разряд* кода в четвертом столбце вписать единицу (1), а нижней группы – ноль (0);

5. Провести *второй* шаг процедуры эффективного кодирования, для чего символы каждой из двух групп, полученных в п.4 алгоритма, разделить на две подгруппы с возможно *равными* суммами вероятностей символов смежных подгрупп, после чего *всем* символам верхних подгрупп в *разряд* кода, следующий за старшим, в пятом столбце вписать единицу (1), а нижних подгрупп – ноль (0);

6. Провести *последующие* шаги процедуры эффективного кодирования по схеме, описанной в п.п.4 и 5 алгоритма, вписывая всякий раз в *последующий* разряд кода символов верхних подподгрупп единицу (1), а нижних подподгрупп – ноль (0), причем процедуру проводить до тех пор, пока не будут закодированы все символы алфавита, число итераций которой в зависимости от числа n кодируемых символов ИДИ и структуры распределения вероятностей $p(x_i)$ по символам может составлять от единицы до $(n - 1)$;

7. Выписать в *последний правый* столбец таблицы значения чисел *разрядов* сформированных эффективных кодов, представляющих собой

длины $l\{K(x_i)\}$ этих кодов с тем, чтобы вычислить среднее на символ число двоичных разрядов кода с помощью соотношения

$$l_{cp}\{K(X)\} = \sum_{i=1}^n p(x_i)l\{K(x_i)\}; \quad (4.9)$$

8. Оценить степень близости $l_{cp}\{K(X)\}$ к энтропии $H(X)$ ИДИ путем контроля выполнения неравенства

$$\frac{l_{cp}\{K(X)\} - H(X)}{H(X)} \leq \delta_l = (0.05 \div 0.1). \quad (4.10)$$

Если неравенство (4.10) выполняется, то осуществить переход к п.14 алгоритма, если неравенство (4.10) не выполняется, то осуществить переход к п.9 алгоритма;

9. Модифицировать источник путем введения агрегированных символов $\tilde{x}_l = x_i x_j, (i, j = \overline{1, n}; l = \overline{1, n^2})$, характеризующихся вероятностями $p(\tilde{x}_l)$ появления на выходе модифицированного ИДИ, равными $p(\tilde{x}_l) = p(x_i)p(x_j)$;

10. Осуществить эффективное кодирование агрегированных символов модифицированного ИДИ по схеме, представленной п.п.2 – 7 алгоритма;

11. Оценить с помощью соотношения (4.9) среднюю длину кода на агрегированный символ, с последующим пересчетом ее путем деления полученной длины на число элементов в агрегированном символе – блоке;

12. Выполнить проверку выполнения неравенства (4.10). Если неравенство (4.10) выполняется, то осуществить переход к п.14 алгоритма, если нет, то осуществить переход к п.13 алгоритма;

13. Осуществить очередную модификацию источника дискретной информации, увеличив число исходных символов в блоках на единицу, и выполнить п.п.2 – 7 и п.п.10 – 12 алгоритма;

14. Передать таблицы полученных эффективных кодов символов (в общем случае блоков символов) в алгоритмическую и аппаратную среду эффективного кодирования. \square

Проиллюстрируем алгоритм К.Шеннона – Р.Фэнно эффективного кодирования двумя примерами.

Пример 4.1. Осуществим эффективное кодирование символов ИДИ, генерирующего алфавит

$$X = \{x_i; i = 1 \div 8: p(x_i) = (2^{-i}); i = 1 \div 7; p(x_8) = p(x_7)\}. \quad \square$$

Решение 1. В соответствии с п.1 алгоритма К.Шеннона – Р.Фэнно (АШФ) произведем вычисление энтропии источника, что дает результат

$$H(X) = -\sum_{i=1}^8 p(x_i) \log_2(p(x_i)) = \sum_{i=1}^8 p(x_i) \log_2(p(x_i)^{-1}) = \frac{1}{2}1 + \frac{1}{4}2 + \frac{1}{8}3 + \frac{1}{16}4 +$$

$$+ \frac{1}{32}5 + \frac{1}{64}6 + \frac{1}{128}7 + \frac{1}{128}7 = \frac{127}{64} [\text{бит/символ}];$$

2. В соответствие с п.п.2÷7 АШФ сформируем и заполним таблицу 4.1 эффективными кодами и значениями их длин

Таблица 4.1

№ п/п	x_i	$p(x_i)$	$K(x_i)$							$l\{K(x_i)\}$
			1	2	3	4	5	6	7	
1.	x_1	1/2	1							1
2.	x_2	1/4	0	1						2
3.	x_3	1/8	0	0	1					3
4.	x_4	1/16	0	0	0	1				4
5.	x_5	1/32	0	0	0	0	1			5
6.	x_6	1/64	0	0	0	0	0	1		6
7.	x_7	1/128	0	0	0	0	0	0	1	7
8.	x_8	1/128	0	0	0	0	0	0	0	7

3. Оценим среднюю длину кода на символ с помощью (4.9)

$$l_{cp}\{K(X)\} = \sum_{i=1}^n p(x_i) l\{K(x_i)\} = \frac{1}{2}1 + \frac{1}{4}2 + \frac{1}{8}3 + \frac{1}{16}4 + \frac{1}{32}5 + \frac{1}{64}6 + \frac{1}{128}7 +$$

$$+ \frac{1}{128}7 = \frac{127}{64} [\text{бит/символ}];$$

4. Оценим степень близости $l_{cp}\{K(X)\}$ к энтропии $H(X)$ ИДИ путем контроля выполнения неравенства (4.10)

$$\frac{l_{cp}\{K(X)\} - H(X)}{H(X)} = \frac{127/64 - 127/64}{127/64} = 0 \leq \delta_l = (0.05 \div 0.1);$$

5. В связи с выполнением неравенства (4.10) с нулевой невязкой считаем, что задача получила идеальное решение, в связи с чем передаем таблицу 4.1 полученных эффективных кодов символов в алгоритмическую и аппаратную среду эффективного кодирования. ■

Примечание 4.1. Рассмотренный пример обнаруживает интересное системное свойство эффективных кодов, состоящее в том, что *наибольший эффект сжатия* кодового представления символов ИДИ имеет место в случае, когда вероятности $p(x_i)$ символов представляют собой отрицательные степени числа два. В этом случае *средняя на символ длина кода* будет всегда *совпадать с энтропией источника*. □

В качестве второго примера рассмотрим ИДИ, распределение вероятностей которого по символам источника приводит к необходимости модификации источника и блоковому кодированию.

Пример 4.2. Осуществим эффективное кодирование символов ИДИ, генерирующего алфавит $X = \{x_1, x_2 : p(x_1) = 0.9; p(x_2) = 0.1\}$. □

Решение

1. В соответствии с п.1 алгоритма К.Шеннона – Р.Фэнно (АШФ) произведем вычисление энтропии источника, что дает результат

$$H(X) = -\sum_{i=1}^2 p(x_i) \log_2(p(x_i)) = -0.9(\log_2 0.9) - 0.1(\log_2 0.1) = 0.9(0.14) + 0.1(0.32) = 0.47 \text{ [бит/символ]}$$

2. В соответствии с п.п.2÷7 АШФ сформируем и заполним таблицу 4.2.1 эффективными кодами и значениями их длин

Таблица 4.2.1

№ п/п	x_i	$p(x_i)$	$K(x_i)$							$l\{K(x_i)\}$
			1	2	3	4	5	6	7	
1.	x_1	0.9	1							1
2.	x_2	0.1	0							1

3. Оценим среднюю длину кода на символ с помощью (4.9)

$$l_{cp}\{K(X)\} = \sum_{i=1}^n p(x_i) l\{K(x_i)\} = 0.9 \cdot 1 + 0.1 \cdot 1 = 1 \text{ [бит/символ]}$$

4. Оценим степень близости $l_{cp}\{K(X)\}$ к энтропии $H(X)$ ИДИ путем контроля выполнения неравенства (4.10)

$$\frac{l_{cp}\{K(X)\} - H(X)}{H(X)} = \frac{1 - 0.47}{0.47} = 2.128 - 1 = 1.128 \gg \delta_l = (0.05 \div 0.1);$$

5. В связи с невыполнением неравенства (4.10) осуществим переход к п.9 алгоритма с целью модификации исходного ИДИ путем введения *мультипликативно* агрегированных символов (блоков) $\tilde{x}_l = x_i x_j, (i, j = \overline{1, n}; l = \overline{1, n^2})$, характеризующихся вероятностями $p(\tilde{x}_l)$ появления на выходе модифицированного ИДИ, равными $p(\tilde{x}_l) = p(x_i)p(x_j)$;

6. Выполним п.п.2 – 7 алгоритма, заполнив таблицу 4.2.2 эффективными кодами блоков символов и значениями их длин

Таблица 4.2.2

№ п/п	\tilde{x}_i	$p(\tilde{x}_i)$	$K(\tilde{x}_i)$							$l\{K(\tilde{x}_i)\}$
			1	2	3	4	5	6	7	
1.	$\tilde{x}_1 = x_1 x_1$	0.81	1							1
2.	$\tilde{x}_2 = x_1 x_2$	0.09	0	1						2
3.	$\tilde{x}_3 = x_2 x_1$	0.09	0	0	1					3
4.	$\tilde{x}_4 = x_2 x_2$	0.01	0	0	0					3

7. Оценим *среднюю длину кода на блок* символов с помощью (4.7)

$$l_{\text{срб}}\{K(\tilde{X})\} = \sum_{i=1}^n p(\tilde{x}_i) l\{K(\tilde{x}_i)\} = 0.81 \cdot 1 + 0.09 \cdot 2 + 0.09 \cdot 3 + 0.01 \cdot 3 = 1.29 [\text{бит}/\text{блок}];$$

и на исходный символ с помощью

$$l_{\text{ср}}\{K(X)\} = \frac{l_{\text{срб}}\{K(\tilde{X})\}}{n_{\sigma}} = \frac{1.29}{2} = 0.645 [\text{бит}/\text{символ}];$$

8. Оценим степень близости $l_{\text{ср}}\{K(X)\}$ к энтропии $H(X)$ ИДИ путем контроля выполнения неравенства (4.10)

$$\frac{l_{\text{ср}}\{K(X)\} - H(X)}{H(X)} = \frac{0.645 - 0.47}{0.47} = 1.37 - 1 = 0.37 > \delta_l = (0.05 \div 0.1);$$

9. В связи с невыполнением неравенства (4.10) осуществим очередную модификацию источника путем введения *мультипликативно* агрегированных символов $\tilde{\tilde{x}}_l = x_i x_j x_k, (i, j, k = \overline{1, n}; l = \overline{1, n^2})$, характеризующихся вероятностями $p(\tilde{\tilde{x}}_l)$ появления на выходе модифицированного ИДИ, равными $p(\tilde{\tilde{x}}_l) = p(x_i)p(x_j)p(x_k)$;

10. Выполним п.п.2 – 7 алгоритма, заполнив таблицу 4.2.3 эффективными кодами блоков символов и значениями их длин

Таблица 4.2.3

№ п/п	\tilde{x}_i	$p(\tilde{x}_i)$	$K(\tilde{x}_i)$							$l\{K(\tilde{x}_i)\}$
			1	2	3	4	5	6	7	
1.	$\tilde{x}_1 = x_1x_1x_1$	0.729	1							1
2.	$\tilde{x}_2 = x_1x_1x_2$	0.081	0	1	1					3
3.	$\tilde{x}_3 = x_1x_2x_1$	0.081	0	1	0					3
4.	$\tilde{x}_4 = x_2x_1x_1$	0.081	0	0	1					3
5.	$\tilde{x}_5 = x_1x_2x_2$	0.009	0	0	0	1	1			5
6.	$\tilde{x}_6 = x_2x_1x_2$	0.009	0	0	0	1	0			5
7.	$\tilde{x}_7 = x_2x_2x_1$	0.009	0	0	0	0	1			5
8.	$\tilde{x}_8 = x_2x_2x_2$	0.001	0	0	0	0	0			5

11. Оценим среднюю длину кода на блок символов с помощью (4.7) и на исходный символ с помощью соотношений

$$l_{срб} \{K(\tilde{X})\} = \sum_{i=1}^n p(\tilde{x}_i) l\{K(\tilde{x}_i)\} = 0.729 \cdot 1 + 3 \cdot 0.081 \cdot 3 + 3 \cdot 0.009 \cdot 5 + 0.001 \cdot 5 =$$

$$= 1.59 [\text{бит/блок}];$$

$$l_{ср} \{K(X)\} = \frac{l_{срб} \{K(\tilde{X})\}}{n_{\tilde{b}}} = \frac{1.29}{2} = 0.53 [\text{бит/символ}];$$

12. Оценим степень близости $l_{ср} \{K(X)\}$ к энтропии $H(X)$ ИДИ путем контроля выполнения неравенства (4.10)

$$\frac{l_{ср} \{K(X)\} - H(X)}{H(X)} = \frac{0.53 - 0.47}{0.47} = 1.127 - 1 = 0.127 \approx \delta_l = (0.05 \div 0.1);$$

13. В связи с выполнением неравенства (4.10) с некоторой *натяжкой* считаем, что задача получила *приемлемое* решение, в связи с чем передаем таблицу 4.2.3 полученных эффективных кодов символов в алгоритмическую и аппаратную среду эффективного кодирования. ■

Примечание 4.2. Строго говоря, для корректного удовлетворения неравенства (4.10) требуется еще одна итерация модификации ИДИ, но стала понятной тенденция приближения средней на символ длины кода к энтропии ИДИ с ростом размеров блоков при блоковом эффективном кодировании средствами АШФ. □

Примечание 4.3. Завершая рассмотрение процедуры (алгоритма) К.Шеннон – Р.Фано эффективного кодирования, следует заметить, что процедура АШФ характеризуется возможностью получения

неоднозначного решения, так как при разбиении символов на группы (подгруппы) можно сделать большей по суммарной вероятности как верхнюю группу (подгруппу), в результате чего код эффективный всякий раз получает свою конкретную реализацию. Это различие растет с ростом основания эффективного кода. □

4.3. Алгоритм эффективного кодирования Д.Хаффмэна

Алгоритм эффективного кодирования Д.Хаффмена (АХ) свободен от недостатка АШФ, изложенного в примечании 4.3. По технологии построения он является таблично – графическим и реализуется в форме трех глобальных итераций. *Первая итерация* является табличной, она сводится к построению за $(n - 1)$ шаг системы аддитивно агрегированных символов. *Вторая итерация* является графической, она сводится к построению графа эффективного кода на основании таблицы, построенной при реализации первой итерации. *Третья итерация* алгоритма состоит в формировании эффективных кодов первичных символов, путем их списывания с ветвей графов при перемещении вдоль графа от начальной вершины до соответствующей конечной.

В случае, если построенный эффективный код не удовлетворяет неравенству (4.10) осуществляется его мультипликативная модификация с целью перехода к построению эффективного кода блока.

АЛГОРИТМ

эффективного кодирования Д. Хаффмэна символов источника дискретной информации

1. Задать источник дискретной информации, описав его в стандартной форме $X = \{x_i : p(x_i); i = \overline{1, n}\}$ и охарактеризовав его значением энтропии $H(X)$, вычисляемой в силу определения

$$H(X) = -\sum_{i=1}^n p(x_i) \log_2 p(x_i);$$

2. Составить таблицу формирования эффективного кода символов ИДИ, имеющую $(n + 1)$ строк и $(n + 1)$ столбцов;

3. Выписать в порядке убывания значений вероятностей $p(x_i); i = \overline{1, n}$ символы $x_i; i = \overline{1, n}$, разместив их во втором столбце, а значения вероятностей их появления на выходе ИДИ – в третьем;

4. Провести *первый* шаг процедуры аддитивного агрегирования символов путем суммирования двух наименьших вероятностей, приняв полученное значение за вероятность появления на выходе ИДИ первого аддитивно агрегированного символа;

5. Выписать в порядке убывания значений вероятностей $p(x_i); i = \overline{1, n-1}$ символов $x_i; i = \overline{1, n-1}$, в состав которых включен первый агрегированный символ, разместив их в четвертом столбце;

6. Прodelать процедуру п.5 алгоритма еще $(n-2)$ – раза с тем, чтобы сформировался единый *глобальный аддитивно агрегированный* символ, характеризующийся вероятностью появления, равной единице, которая вписывается в последний $(n+1)$ – й столбец таблицы;

7. Сформировать граф кода по следующему правилу. Начальному узлу графа кода приписываются *глобальный аддитивно агрегированный* символ и значение его вероятности равно единице (1), а n финальным узлам приписываются исходные символы $x_i (i = \overline{1, n})$, наблюдаемые на выходе ИДИ, и значения их вероятностей $p(x_i); i = \overline{1, n}$; промежуточным узлам графа, соответствующим промежуточным *аддитивно агрегированным* символам, приписываются соответствующие *аддитивно агрегированные вероятности*. Ветвям графа, соответствующим «движению» от значения вероятности появления на выходе ИДИ *глобального аддитивно агрегированного* символа через вершины, соответствующие промежуточным *аддитивно агрегированным* символам, к финальным вершинам, в случае, если движение по графу осуществляется в направлении «сверху – вниз» приписываются *единичные значения* (1), если эта ветвь *левая*, и *нулевые* (0) – если эта ветвь *правая*.

8. Сформировать эффективные коды путем их списывания с ветвей графов при перемещении вдоль графа от начальной вершины до соответствующей конечной;

9. Оценить степень близости $l_{cp}\{K(X)\}$ к энтропии $H(X)$ ИДИ путем контроля выполнения неравенства (4.10). Если неравенство (4.10) выполняется, то осуществить переход к п.15 алгоритма, если неравенство (4.10) не выполняется, то осуществить переход к п.10 алгоритма;

10. Модифицировать источник путем введения мультипликативно агрегированных символов $\tilde{x}_l = x_i x_j, (i, j = \overline{1, n}; l = \overline{1, n^2})$, характеризующихся вероятностями $p(\tilde{x}_l)$ появления на выходе модифицированного ИДИ, равными $p(\tilde{x}_l) = p(x_i)p(x_j)$;

11. Осуществить эффективное кодирование агрегированных символов модифицированного ИДИ по схеме, представленной п.п.2 – 8 алгоритма;

12. Оценить с помощью соотношения (4.9) среднюю длину кода на агрегированный символ, с последующим пересчетом ее путем

деления полученной длины на число элементов в агрегированном символе – блоке;

13. Выполнить проверку выполнения неравенства (4.10). Если неравенство (4.10) выполняется, то осуществить переход к п.15 алгоритма, если нет, то осуществить переход к п.13 алгоритма;

14. Осуществить очередную модификацию источника дискретной информации, увеличив число исходных символов в блоках на единицу, и выполнить п.п.2 – 8 и п.п.11 – 13 алгоритма;

15. Передать таблицы полученных эффективных кодов символов (в общем случае блоков символов) в алгоритмическую и аппаратную среду эффективного кодирования. □

Проиллюстрируем алгоритм Д. Хаффмэна эффективного кодирования примером.

Пример 4.3. Осуществим эффективное кодирование символов ИДИ, генерирующего алфавит

$$X = \left\{ \begin{array}{l} x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8 : p(x_1) = 0.22; p(x_2) = 0.2; p(x_3) = 0.16; \\ p(x_4) = 0.16; p(x_5) = 0.10; p(x_6) = 0.10; p(x_7) = 0.04; p(x_8) = 0.02; \end{array} \right\}.$$

□

Решение

1. В соответствии с п.1 алгоритма Д. Хаффмэна (АХ) произведем вычисление энтропии источника, что дает результат

$$\begin{aligned} H(X) &= -\sum_{i=1}^8 p(x_i) \log_2(p(x_i)) = -0.22 \log_2 0.22 - 0.2 \log_2 0.2 - 0.16 \log_2 0.16 - \\ &- 0.16 \log_2 0.16 - 0.1 \log_2 0.1 - 0.1 \log_2 0.1 - 0.04 \log_2 0.04 - 0.02 \log_2 0.02 \\ &= 0.22 \cdot 2.1847 + 0.2 \cdot 2.3229 + 2 \cdot 0.16 \cdot 2.6442 + 2 \cdot 0.1 \cdot 3.3223 + \\ &+ 0.04 \cdot 4.6439 + 0.02 \cdot 5.6446 = 2.7545 \text{ (бит/символ)}; \end{aligned}$$

2. Составим таблицу формирования эффективного кода

Таблица 4.3.1

№ п/п	x_i	$p(x_i)$	Вероятности аддитивно модифицированного источника						
			1	2	3	4	5	6	7
1.	x_1	0.22	0.22	0.22	0.26	0.32	0.42	0.58	1.0
2.	x_2	0.2	0.2	0.2	0.22	0.26	0.32	0.42	
3.	x_3	0.16	0.16	0.16	0.2	0.22	0.26		
4.	x_4	0.16	0.16	0.16	0.16	0.2			
5.	x_5	0.1	0.1	0.16	0.16				
6.	x_6	0.1	0.1	0.1					
7.	x_7	0.04	0.06						
8.	x_8	0.02							

3. Сформируем граф эффективного кода на основании таблицы 4.3.1, следуя п.7 АХ, который приведен на рисунке 4.1;

4. Сформируем эффективные коды символов путем их списывания с ветвей графов при перемещении вдоль графа рисунок 4.1 от начальной вершины до соответствующей конечной, разместив результаты в таблице 4.3.2;

Таблица 4.3.2

№ п/п	x_i	$p(x_i)$	$K(x_i)$						$l\{K(x_i)\}$
			1	2	3	4	5	6	
1.	x_1	0.22	0	1					2
2.	x_2	0.2	0	0					2
3.	x_3	0.16	1	1	1				3
4.	x_4	0.16	1	1	0				3
5.	x_5	0.1	1	0	0				3
6.	x_6	0.1	1	0	1	1			4
7.	x_7	0.04	1	0	1	0	1		5
8.	x_8	0.02	1	0	1	0	0		5

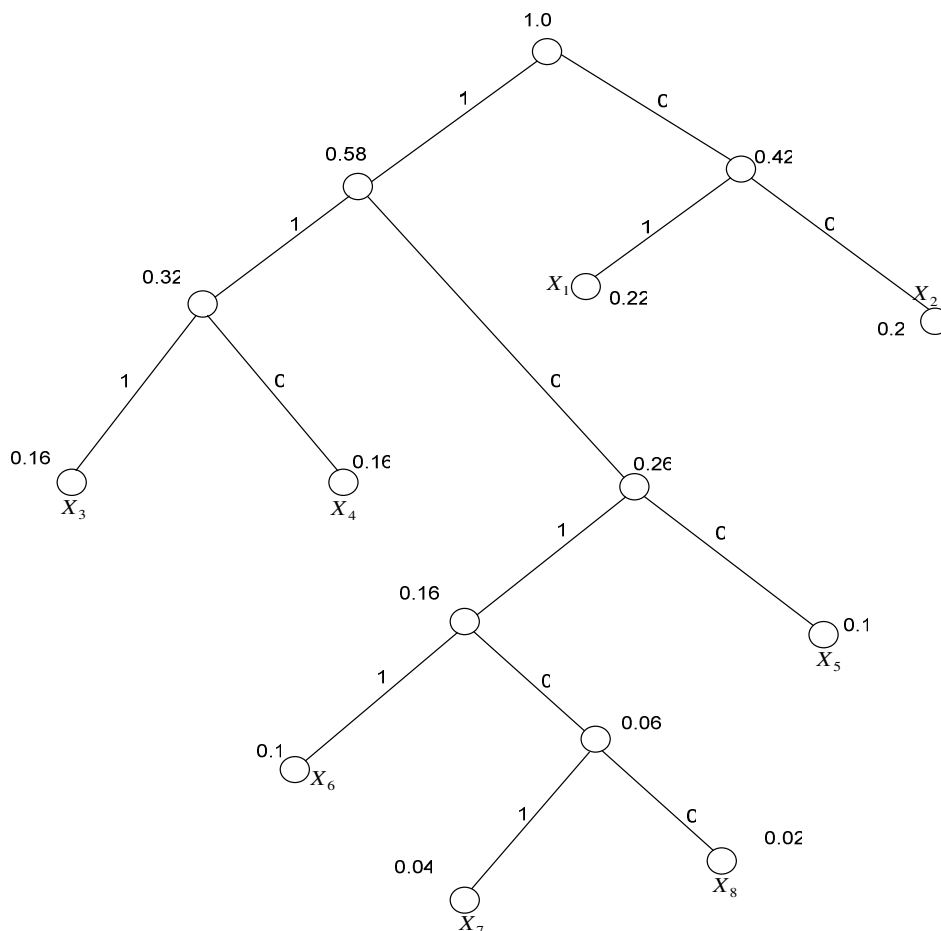


Рисунок 4.1

5. Оценим среднюю длину кода на символ с помощью соотношения (4.9)

$$l_{cp}\{K(X)\} = \sum_{i=1}^n p(x_i) l\{K(x_i)\} = 0.22 \cdot 2 + 0.2 \cdot 2 + 0.16 \cdot 3 + 0.16 \cdot 3 + 0.1 \cdot 3 + 0.1 \cdot 4 + \\ + 0.04 \cdot 5 + 0.02 \cdot 5 = 2.8 [\text{бит/символ}];$$

6. Выполним проверку выполнения неравенства (4.10) с целью оценки степени близости средней на символ длины эффективного кода к энтропии ИДИ

$$\frac{l_{cp}\{K(X)\} - H(X)}{H(X)} = \frac{2.8 - 2.7545}{2.8} = 0.0163 \leq \delta_l = (0.05 \div 0.1);$$

7. Так как неравенство (4.10) выполняется, то передаем таблицы полученных эффективных кодов символов в алгоритмическую и аппаратную среду эффективного кодирования. ■

Завершая рассмотрение проблемы построения эффективных кодов, необходимо отметить положительные и отрицательные свойства этих кодов.

1. Эффективный код является эффективным средством уменьшения объема передаваемого сигнала за счет сокращения средней длительности сигнала;

2. Несмотря на неравномерный характер эффективного кода, он гарантирует однозначность декодирования этих кодов, что обеспечивается свойством *префиксности* эффективных кодов, построенных с помощью алгоритмов К.Шеннона – Р.Фэнно и Д.Хаффмэна, при котором ни один более короткий код не является начальным фрагментом любого более длинного кода;

3. Эффективное использование предоставленной линии связи предполагает необходимость поступления закодированных символов равномерно, что нарушается неравномерной природой эффективного кода, что особенно сильно проявляется при эффективном кодировании блоков символов;

4. Эффективные коды относятся к классу помехонезащищенных, так как минимальное кодовое расстояние между эффективными кодами одной длины равно единице, в результате чего искажение в одном разряде трансформирует одну кодовую комбинацию;

5. Основной областью использования возможностей эффективных кодов является не *канал передачи*, а *канал хранения*, ярким примером использования возможностей эффективных кодов являются *архиваторы* в современных компьютерах.

Примеры и задачи

4.1. Осуществить с помощью АШФ эффективное кодирование символов ИДИ, генерирующего алфавит

$$X = \{x_i; i = 1 \div 8: p(x_i) = (2^{-i}); i = 1 \div 7; p(x_8) = p(x_7)\};$$

4.2. Осуществить с помощью АШФ эффективное кодирование символов ИДИ, генерирующего алфавит

$$X = \{x_1, x_2: p(x_1) = 0.9; p(x_2) = 0.1\};$$

4.3. Осуществить с помощью АХ эффективное кодирование символов ИДИ, генерирующего алфавит

$$X = \left\{ \begin{array}{l} x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8: p(x_1) = 0.22; p(x_2) = 0.2; p(x_3) = 0.16; \\ p(x_4) = 0.16; p(x_5) = 0.10; p(x_6) = 0.10; p(x_7) = 0.04; p(x_8) = 0.02; \end{array} \right\};$$

4.4. Осуществить с помощью АХ эффективное кодирование символов ИДИ, генерирующего алфавит

$$X = \{x_i; i = 1 \div 8: p(x_i) = (2^{-i}); i = 1 \div 7; p(x_8) = p(x_7)\};$$

4.5. Осуществить с помощью АХ эффективное кодирование символов ИДИ, генерирующего алфавит

$$X = \{x_1, x_2: p(x_1) = 0.9; p(x_2) = 0.1\};$$

4.6. Осуществить с помощью АШФ эффективное кодирование символов ИДИ, генерирующего алфавит

$$X = \left\{ \begin{array}{l} x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8: p(x_1) = 0.22; p(x_2) = 0.2; p(x_3) = 0.16; \\ p(x_4) = 0.16; p(x_5) = 0.10; p(x_6) = 0.10; p(x_7) = 0.04; p(x_8) = 0.02; \end{array} \right\};$$

4.7. Осуществить с помощью АШФ эффективное кодирование символов ИДИ, генерирующего алфавит

$$X = \{x_i; i = 1 \div 8: p(x_i) = (2^{-3})\};$$

4.8. Осуществить с помощью АШФ эффективное кодирование символов ИДИ, генерирующего алфавит

$$X = \{x_1, x_2: p(x_1) = 0.75; p(x_2) = 0.25\};$$

4.9. Решить задачу 4.8. с помощью АХ;

4.10. Осуществить с помощью АШФ эффективное кодирование символов ИДИ, генерирующего алфавит

$$X = \{x_1, x_2, x_3: p(x_1) = 0.6; p(x_2) = 0.25; p(x_3) = 0.15\};$$

4.11. Решить задачу 4.10. с помощью АХ;

4.12. Осуществить с помощью АШФ эффективное кодирование символов ИДИ, генерирующего алфавит

$$X = \left\{ \begin{array}{l} x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8: p(x_1) = 0.4; p(x_2) = 0.2; p(x_3) = 0.1; \\ p(x_4) = 0.06; p(x_5) = 0.06; p(x_6) = 0.06; p(x_7) = 0.06; p(x_8) = 0.06; \end{array} \right\};$$

4.13. Осуществить с помощью АХ эффективное кодирование символов ИДИ, генерирующего алфавит

$$X = \left\{ \begin{array}{l} x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8: p(x_1) = 0.4; p(x_2) = 0.2; p(x_3) = 0.1; \\ p(x_4) = 0.06; p(x_5) = 0.06; p(x_6) = 0.06; p(x_7) = 0.06; p(x_8) = 0.06; \end{array} \right\};$$

Решение вариантов задач

Задача 4.1. Осуществить с помощью АШФ эффективное кодирование символов ИДИ, генерирующего алфавит

$$X = \{x_i; i = 1 \div 8: p(x_i) = (2^{-i}); i = 1 \div 7; p(x_8) = p(x_7)\};$$

Решение (см. в тексте параграфа 4.2.).

Задача 4.2. Осуществить с помощью АШФ эффективное кодирование символов ИДИ, генерирующего алфавит

$$X = \{x_1, x_2: p(x_1) = 0.9; p(x_2) = 0.1\};$$

Решение (см. в тексте параграфа 4.2.).

Задача 4.3. Осуществить с помощью АХ эффективное кодирование символов ИДИ, генерирующего алфавит

$$X = \left\{ \begin{array}{l} x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8: p(x_1) = 0.22; p(x_2) = 0.2; p(x_3) = 0.16; \\ p(x_4) = 0.16; p(x_5) = 0.10; p(x_6) = 0.10; p(x_7) = 0.04; p(x_8) = 0.02; \end{array} \right\};$$

Решение (см. в тексте параграфа 4.3.).