МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ   
РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

**«БЕЛГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ**

**ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ им. В. Г. ШУХОВА»**

**(БГТУ им. В.Г. Шухова)**

Кафедра программного обеспечения вычислительной техники и автоматизированных систем

**Лабораторная работа №1**

по дисциплине: Теория информации

тема: «Исследование кодирования по методу Хаффмана. Оценка эффективности кода»

Выполнил: ст. группы ПВ-233

Ситников Алексей Павлович

Проверил: Твердохлеб Виталий

Викторович

Белгород 2025 г.

**Цель работы:** изучить метод Хаффмана и научиться его применять.

Задание 1. Построить кодовое представление сообщения, вероятности появления символов в пределах алфавита которого приведены в табл.1.

Таблица 1 – Вероятности появления символов в пределах алфавита исходного сообщения

Изображение выглядит как текст, Шрифт, линия, число

Автоматически созданное описание

Расположим вероятности по возрастанию:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Символ | s8 | s7 | s5 | s6 | s3 | s4 | s2 | s1 |
| Вероятность | 0.01 | 0.05 | 0.10 | 0.10 | 0.16 | 0.16 | 0.19 | 0.23 |

Изображение выглядит как диаграмма, круг

Автоматически созданное описание

s1 = 01

s2 = 00

s3 = 110

s4 = 111

s5 = 1011

s6 = 100

s7 = 10101

s8 = 10100

Пример предложения с данным алфавитом:

s1s2s2s4s4s5s1s1s7s8s8s4s3s1s2s2s3 кодируется в: 01000011111110110101101011010010100111110010000110

Найдём B: B = 17 \* 8 = 136

Найдём B`: B` = 4 \* 2 + 4 \* 2 + 2 \* 3 + 3 \* 3 + 4 + 5 + 2 \* 5 = 50

Коэффициент сжатия K comp= 136/50 = 2,72

Найдём = 2 \* 0,23 + 2 \* 0,19 + 3 \* 0,16 + 3 \* 0,16 + 3 \* 0,1 + 4 \* 0,1 + 5 \* 0,05 + 5 \* 0,01 = 2,4

Величина дисперсии:

= 0,23 \* (2 – 2,4) 2 + 0,19 \* (2 – 2,4) 2 + 0,16 \* (3 – 2,4) 2 + 0,16 \* (3 – 2,4) 2 + 0,1 \* (4 – 2,4) 2 + 0,1 \* (3 – 2,4) 2 + 0,05 \* (5 – 2,4) 2 + 0,01 \* (5 – 2,4) 2 = 0,57

Задание 2. Построить кодовое представление сообщения, вероятности появления символов в пределах алфавита которого приведены в табл.2.

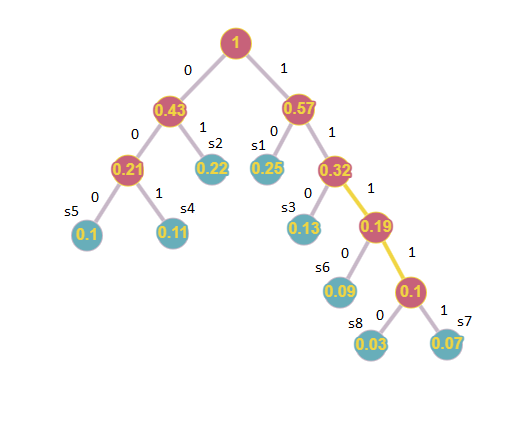
Таблица 2 – Вероятности появления символов в пределах алфавита исходного сообщения

Изображение выглядит как текст, Шрифт, линия, число

Автоматически созданное описание

Расположим вероятности по возрастанию:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Символ | s8 | s7 | s6 | s5 | s4 | s3 | s2 | s1 |
| Вероятность | 0.03 | 0.07 | 0.09 | 0.1 | 0.11 | 0.13 | 0.22 | 0.25 |



s1 = 10

s2 = 01

s3 = 110

s4 = 001

s5 = 000

s6 = 1110

s7 = 11111

s8 = 11110

Пример предложения с данным алфавитом:

s1s2s2s4s4s5s1s1s7s8s8s4s3s1s2s2s3 кодируется в: 1001010010010001010111111111011110001110100101110

Найдём B: B = 17 \* 8 = 136

Найдём B`: B` = 4 \* 2 + 4 \* 2 + 2 \* 3 + 3 \* 3 + 3 + 5 + 2 \* 5 = 49

Коэффициент сжатия K comp= 136/49 2,78

Найдём = 2 \* 0,25+ 2 \* 0,22 + 3 \* 0,13+ 3 \* 0,11 + 3 \* 0,1 + 4 \* 0,09 + 5 \* 0,07 + 5 \* 0,03 = 2,82

Величина дисперсии:

= 0,25 \* (2 – 2,82) 2 + 0,22\* (2 – 2,82) 2 + 0,13 \* (3 – 2,82) 2 + 0,11 \* (3 – 2,82) 2 + 0,1 \* (3 – 2,82) 2 + 0,09 \* (4 – 2,82) 2 + 0,07 \* (5 – 2,82) 2 + 0,03 \* (5 – 2,82) 2 = 0,93

Задание 3. Построить кодовое представление сообщения:

**оитомии о ими оооитми о о о ооиимтомиимотоим оои тоо и и м оио и омтоо тоимо т и**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Символ | о | и | т | м | \_ |
| Вероятность | 0.31 | 0.24 | 0.1 | 0.13 | 0.22 |

Изображение выглядит как круг, снимок экрана, диаграмма

Автоматически созданное описание

и = 00

о = 01

- = 10

т = 110

м = 111

010011001111000010011000111001000110010010101001101100100011001100110010100001111100111100001110111001001111001010010110010110001000101111001000110001001111110010110110010011101101101000

Задание 4. Для условий, приведенных в заданиях 1 и 2 и 3, выявить возможность построения альтернативных кодовых моделей сообщения. В случае обнаружения таковых выявить наиболее эффективные из них по критериям  и  .

Рассмотрим задание 1.

После второго склеивания получим три одинаковые вероятности, но при замене мест будет меняться только код символов, но не размер, перестановка не даст разницы.

Рассмотрим задание 2.

После первого склеивания получаем две одинаковые вероятности 0.1, выбор, какую вероятность разместить раньше, повлияет на итог.  
В задание выше первым взяли полученную на предыдущем этапе вероятность, но если взять из таблицы вероятность символа получится такое дерево:

Изображение выглядит как диаграмма, круг

Автоматически созданное описание

Получим такие коды:  
s1 = 10

s2 = 01

s3 = 110

s4 = 001

s5 = 1111

s6 = 1110

s7 = 0001

s8 = 0000

Высота дерева уменьшилась на 1.

Пример предложения с данным алфавитом:

s1s2s2s4s4s5s1s1s7s8s8s4s3s1s2s2s3 кодируется в: 10010100100111111010000100000000001110100101110

Найдём B: B = 17 \* 8 = 136

Найдём B`: B` = 4 \* 2 + 4 \* 2 + 2 \* 3 + 3 \* 3 + 4 + 4 + 2 \* 4 = 47

Коэффициент сжатия K comp= 136/47 2,89

Найдём = 2 \* 0,25 + 2 \* 0,22 + 3 \* 0,13 + 3 \* 0,11 + 4 \* 0,1 + 4 \* 0,09 + 4 \* 0,07 + 4 \* 0,03 = 2,82

Величина дисперсии:

= 0,25 \* (2 – 2,82) 2 + 0,22\* (2 – 2,82) 2 + 0,13 \* (3 – 2,82) 2 + 0,11 \* (3 – 2,82) 2 + 0,1 \* (4 – 2,82) 2 + 0,09 \* (4 – 2,82) 2 + 0,07 \* (4 – 2,82) 2 + 0,03 \* (4 – 2,82) 2 = 0,72

Коэффициент сжатия был 2,78; а стал 2,89. Дисперсия была 0,93; а стала 0,72. Чем больше коэффициент сжатия, тем лучше, чем меньше дисперсия, тем лучше, значит второй вариант по обоим параметрам обходит первый вариант.

В третьем задании дерево, построенное по методу Хаффмана, имеет один вариант.

**Приложение Хаффман:**

Изображение выглядит как текст, диаграмма, Шрифт, снимок экрана

Автоматически созданное описание

**Приложение greedy2.exe:**

Входные данные для задания №1, получил такое дерево:  
Изображение выглядит как снимок экрана, диаграмма, линия, круг

Автоматически созданное описание

Совпало с вычисленным вручную.

Входные данные для задания №1, получил такое дерево:

Изображение выглядит как снимок экрана, диаграмма, Цвет электрик, линия

Автоматически созданное описание

Совпало с вычисленным вручную.

**Вывод:** после проделанной работы я изучил метод Хаффмана и научился его применять, изучил критерии эффективности кодирования.