|  |  |
| --- | --- |
| Череповецкий государственный университет  Кафедра «Математического и программного обеспечения ЭВМ» | |
| ОТЧЕТ ПО ЛАБОРАТОРНОЙ РАБОТЕ  по дисциплине «Теория информации»  КОДЫ ФАНО И ХАФФМАНА | |
|  | Принял:  преподаватель Е.Н. Руденко    подпись, дата  Выполнил:  студент гр. 1ПИб-02-1оп-22  Остапенко Степан Ярославович  подпись, дата |
| Череповец, 2023 | |

**Реферат**

Цель работы – познакомиться с методами сжатия информации и применить их на практике.

В ходе работы исследуются три метода сжатия информации: RLE, Шеннона-Фано и Хаффмана.

**Ход работы**

Вариант 17

Задание 1.

Выполнить вручную кодирование сообщения методом RLE. В качестве исходной фразы взять текст из табл. 3.1. С помощью таблицы CP-1251 (см. Приложение 1) перевести символы заданной фразы в десятичные числа, а затем десятичные числа перевести в двоичные. Выполнить сжатие информации, вычислить контрольные суммы и коэффициент сжатия.

 Табл. 3.1.

|  |  |
| --- | --- |
| Вар | Текст |
| 17 | Ккккктттттттто тттттам? |

Решение:

|  |  |  |
| --- | --- | --- |
| Символ | Десятичные числа | Двоичный код |
| К | 202 | 11001010 |
| к | 234 | 11101010 |
| т | 242 | 11110010 |
| о | 238 | 11101110 |
| [пробел] | 32 | 00100000 |
| а | 224 | 11100000 |
| м | 236 | 11101100 |
| ? | 63 | 00111111 |

|  |  |  |
| --- | --- | --- |
| Символ | Десятичные числа | Двоичный код |
| 1 | 49 | 00110001 |
| 4 | 52 | 00110100 |
| 5 | 53 | 00110101 |
| 8 | 56 | 00111000 |

Переведем исходное сообщение в двоичный код:

11001010 11101010 11101010 11101010 11101010 11110010 11110010 11110010 11110010 11110010 11110010 11110010 11110010 11101110 00100000 11110010 11110010 11110010 11110010 11110010 11100000 11101100 00111111

Проведём сжатие на основе фиксированной длины кодового слова:

1К 4к 8т 1о 1 5т 1а 1м 1?

Переведем это сообщение в двоичный код:

00110001 11001010 00110100 11101010 00111000 11110010 00110001 11101110 00110001 00100000 00110101 11110010 00110001 11100000 00110001 11101100 00110001 00111111

До сжатия: 184 бит

После сжатия: 144 бит

Кс = 144/184 = 0,783

**Задание 2. Выполнить сжатие информации методом Шеннона-Фано**

Используя фразу из табл. 3.1, построить кодовое дерево и определить коэффициент сжатия методом Шеннона-Фано

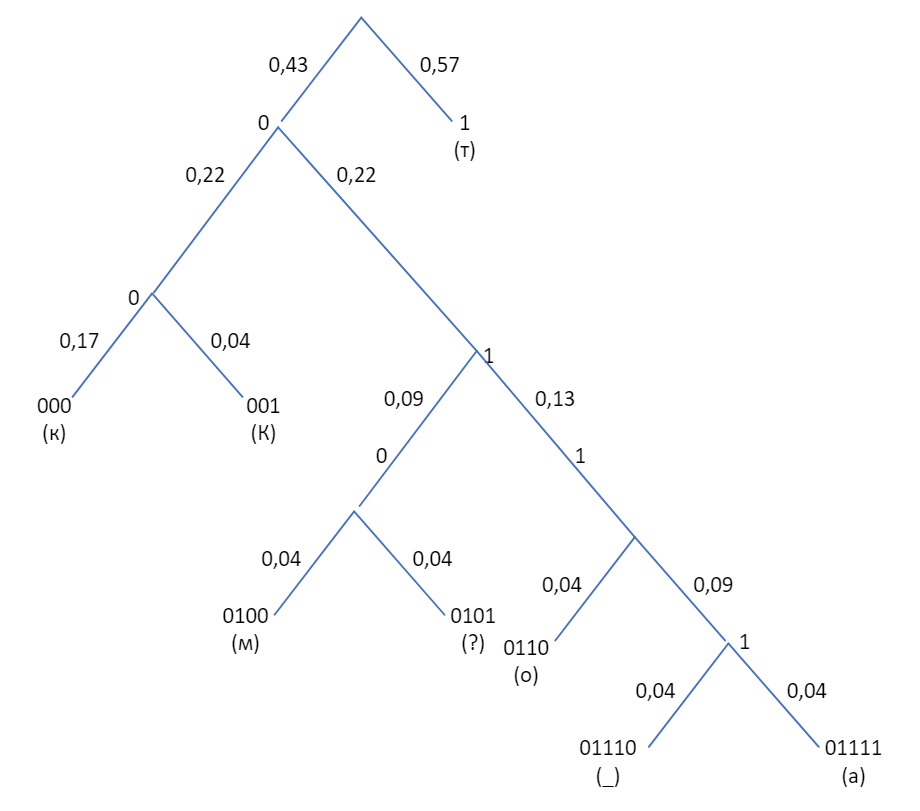
|  |  |
| --- | --- |
| Вар | Текст |
| 17 | Ккккктттттттто тттттам? |

Решение:

Найдём вероятности появления символа в сообщении:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Символ | Вероятность появления | | | | | |
| т | 0,57 | 0,57 | 0,57 | 0,57 | 0,57 | 1,00 |
| к | 0,17 | 0,22 | 0,22 | 0,22 | 0,43 |  |
| К | 0,04 |
| о | 0,04 | 0,04 | 0,13 | 0,22 |
| \_ | 0,04 | 0,09 |
| а | 0,04 |
| м | 0,04 | 0,09 | 0,09 |
| ? | 0,04 |

Построим кодовое дерево:



Из дерева следует, что

|  |  |
| --- | --- |
| Символ | Двоичный код |
| К | 001 |
| к | 000 |
| т | 1 |
| о | 0110 |
| \_ | 01110 |
| а | 01111 |
| м | 0100 |
| ? | 0101 |

Закодируем исходное сообщение в соответствии с деревом Фано. Закодированное сообщение выглядит следующим образом:

001 000 000 000 000 1 1 1 1 1 1 1 1 0110 01110 1 1 1 1 1 01111 0100 0101

До сжатия: 184 бит

После сжатия: 50 бит

Кс = 50/184 = 0,272

**Задание 3. Выполнить сжатие информации методом Хаффмана**

Используя фразу из табл. 3.1, построить кодовое дерево и определить коэффициент сжатия методом Хаффмана

|  |  |
| --- | --- |
| Вар | Текст |
| 17 | Ккккктттттттто тттттам? |

Решение:

Построим таблицу по методу Хаффмана:



В результате получаем:



Закодируем исходное сообщение в соответствии с методом Хаффмана. Закодированное сообщение выглядит следующим образом:

1010 100 100 100 100 0 0 0 0 0 0 0 0 1011 1110 0 0 0 0 0 1111 1100 1101

До сжатия: 184 бит

После сжатия: 49 бит

Кс = 49/184 = 0,266

**Задание 4. Архивирование файлов**

С помощью стандартного архиватора (WinZip, WinRar, 7-Zip и т.п.) выполнить сжатие различных документов, тип которых указан в таблице 3.4.1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Документ | Расширение | Объем файла до архивации, Кбайт | Объем файла после архивации, Кбайт | Коэффициент сжатия |
| Текст | .doc | 913,5 | 186,4 | 0,20405 |
| Фотография | .jpg | 270,2 | 270,4 | 1,00074 |
| Рисунок | .bmp | 1037,8 | 15,2 | 0,01465 |
| Видео | .avi | 745 | 736,9 | 0,98913 |
| Звук | .mp3 | 4675,2 | 4664,4 | 0,99769 |

В качестве текстового документа нужно взять файл, который не содержит рисунков. Число символов должно быть более 3000 знаков.

Фотографии нужно взять на сайте **samara.psati.ru** в соответствии с вариантом (таблица 3.4.2.). Следует взять одну фотографию из указанного раздела.

В качестве рисунка формата BMP нужно взять изображение флага государства в соответствии с таблицей 3.4.3.

Видеоклип желательно снять самостоятельно (с помощью видеокамеры, цифрового фотоаппарата, мобильного телефона, планшетника).

В качестве звукового файла используйте своё любимое музыкальное произведение

**Заключение**

Познакомился с такими методами сжатия информации, как метод RLE, дерево Шеннона-Фано и дерево Хаффмана. Также воспользовался стандартным архиватором WinRar.