|  |  |
| --- | --- |
| Череповецкий государственный университет  Кафедра «Математического и программного обеспечения ЭВМ» | |
| ОТЧЕТ ПО ЛАБОРАТОРНОЙ РАБОТЕ  по дисциплине «Теория информации»  КОДИРОВАНИЕ СООБЩЕНИЙ МЕТОДОМ RLE | |
|  | Принял:  преподаватель Е.Н. Руденко    подпись, дата  Выполнил:  студент гр. 1ПИб-02-3оп-22  Маркелов С.А.    подпись, дата |
| Череповец, 2023 | |

Реферат

Предметом исследование является кодирование сообщений методом кодирования повторов (RLE). Кодирование повторов – элементарный метод сжатия без потерь, при котором повторы символов кодируются числом повторов.

Цель: научиться кодировать и декодировать по методу RLE.

В ходе работы проводились теоретические исследования различных методов кодирования.

В результате аналитических исследований были найдены наиболее оптимальные методы кодирования. Численно были получены значения контрольных сумм и коэффициентов сжатия. Проведено кодирование различных сообщений методами RLE, Шеннона-Фано и Хаффмана. Произведено сжатие файлов различных форматов с помощью архиватора.

Введение

Сжатие данных – это процесс, обеспечивающий уменьшение объема данных. Целью сжатия информации является уменьшение ее объема для экономии ресурсов при хранении или передаче данных. Сжатие без потерь полностью обратимо. Оно действует за счет уменьшения избыточности. Сжатие с потерями необратимо. Оно действует за счет уменьшения избыточности и отбрасывания части информации.

К способам сжатия относятся:

* Изменение содержания данных (уменьшение избыточности данных)
* Изменение структуры данных (эффективное кодирование)
* Изменение содержания и структуры данных

Архивация – процесс, при котором файлы сжимаются без потери информации. При разархивации данные и программы восстанавливаются в исходном виде. Архиватор – программа, выполняющая сжатие (архивирование) файлов для более компактного хранения во внешней памяти и восстановление (разархивирование) сжатых файлов в первоначальном состоянии.

Коэффициент сжатия – это величина для обозначения эффективности метода сжатия, равная отношению количества информации до и после сжатия. Вычисляется по формуле , где Sc – объем информации после сжатия, S – объем информации до сжатия.

Алгоритм сжатия данных (RLE) основан на замене повторяющихся битов (в тексте может иметься последовательность одинаковых символов, в графическом файле – закрашенная одним цветом область и т.д.). При применении этого алгоритма вместо последовательности одинаковых по цвету пикселей в строке изображения записывается цвет и количество его повторений. Такой подход используется при хранении изображений в формате BMP.

Например, в тексте подряд идут 10 пробелов, которые кодируются 10-тью байтами. При архивации они заменяются 3-мя байтами (первый байт – кодирует заменяемый символ; второй байт – специальный байт "флажка" архивации, который указывает на необходимость развернуть первый байт в последовательность байтов; третий байт указывает количество повторяющихся байтов).

Ход работы

Вариант 14

**Задание 1.** Выполнить вручную кодирование сообщения методом RLE. В качестве исходной фразы взять текст: Цена 2598888666611 коп. С помощью таблицы CP-1251 перевести символы заданной фразы в десятичные числа, а затем десятичные числа перевести в двоичные. Выполнить сжатие информации, вычислить контрольные суммы и коэффициент сжатия.

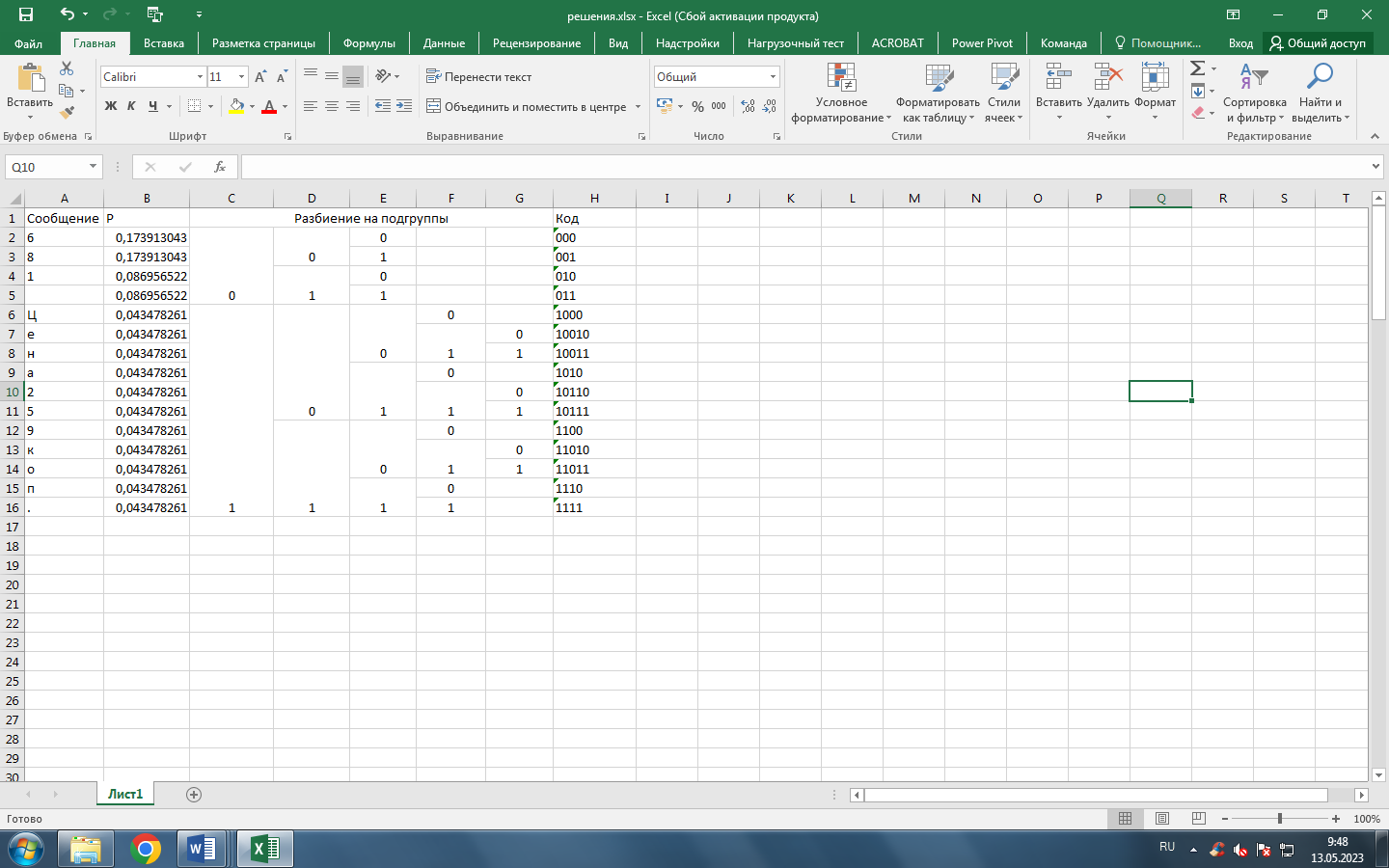
|  |  |  |
| --- | --- | --- |
| Символы | Десятичные числа | Двоичные числа |
| «Ц» | 1.214 | 1.11010110. |
| «е» | 1.229 | 1.11100101. |
| «н» | 1.237 | 1.11101101. |
| «а» | 1.224 | 1.11100000. |
| « » | 1.32 | 1.100000. |
| «2» | 1.50 | 1.110010. |
| «5» | 1.53 | 1.110101. |
| «9» | 1.57 | 1.111001. |
| 4 «8» | 4.56 | 100.111000. |
| 4 «6» | 4.54 | 100.110110. |
| 2 «1» | 2.49 | 10.110001. |
| « » | 1.32 | 1.100000. |
| «к» | 1.234 | 1.11101010. |
| «о» | 1.238 | 1.11101110. |
| «п» | 1.239 | 1.11101111. |
| «.» | 1.46 | 1.101110. |

1Ц 1е 1н 1а 1 12 15 19 48 46 21 1 1к 1о 1п 1.

1.11010110.1.11100101.1.11101101.1.11100000.1.100000.1.110010.1.110101.1.111001.100.111000.100.110110.10.110001.1.100000.1.11101010.1.11101110.1.11101111.1.101110.

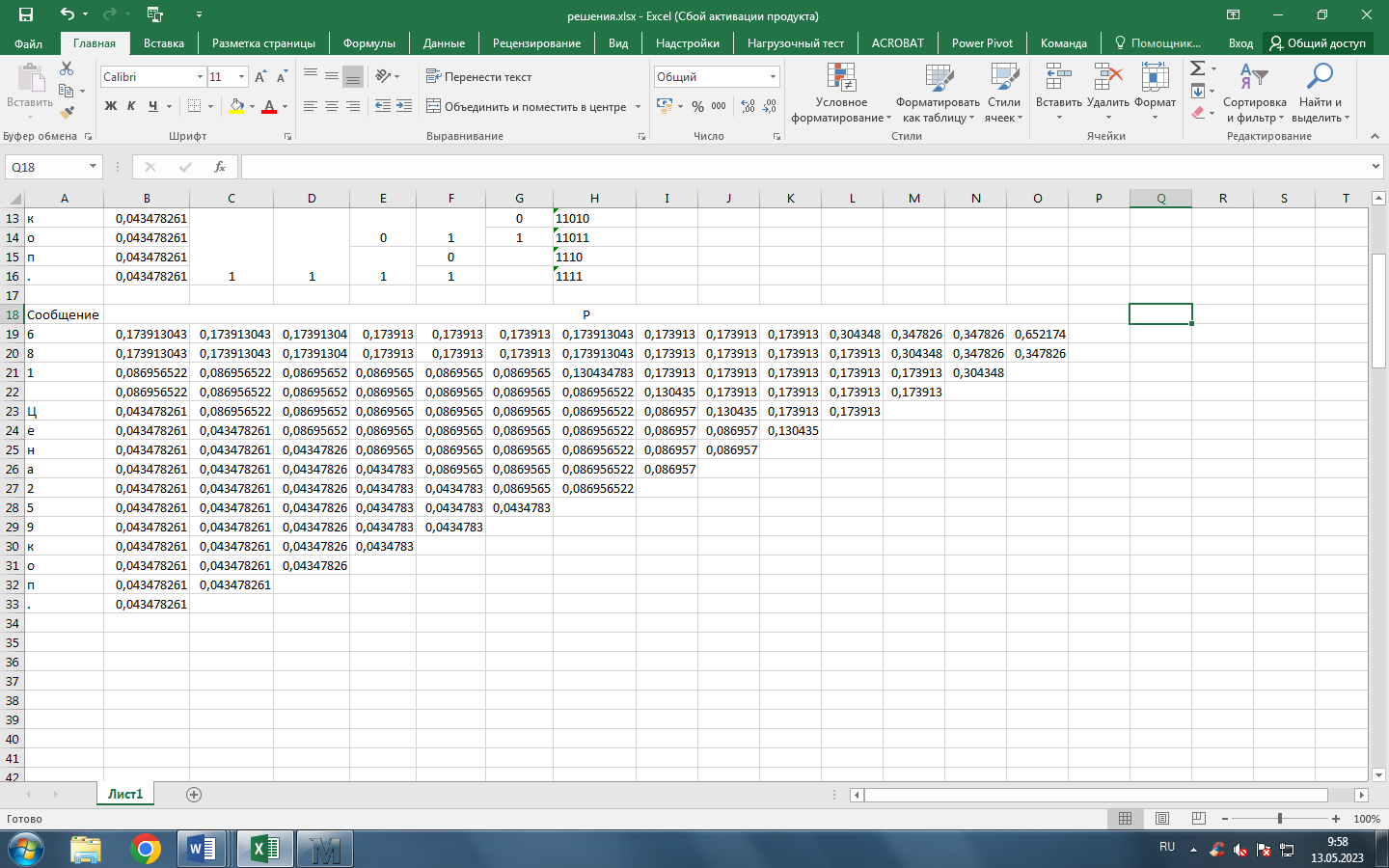
Коэффициент сжатия:

**Задание 2.** Используя фразу «Цена 2598888666611 коп.», построить кодовое дерево и определить коэффициент сжатия методом Шеннона-Фано.



Кс=0,17\*3+0,17\*3+0,08\*3+0,08\*3+0,04\*4+0,04\*5+0,04\*5+0,04\*4+0,04\*5+0,04\*5+0,04\*4+0,04\*5+0,04\*5+0,04\*4+0,04\*4=3,5 бит

**Задание 3.** Используя фразу «Цена 2598888666611 коп.», построить кодовое дерево и определить коэффициент сжатия методом Хаффмана



|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Ц | е | н | а |  | 2 | 5 | 9 | 8 | 6 |
| 00000 | 00001 | 00010 | 00011 | 0110 | 00101 | 00110 | 00111 | 00 | 01 |
| 1 | к | п | р | . |
| 0101 | 0111 | 01000 | 01001 | 00100 |

Кс=0,17\*2+0,17\*2+0,08\*4+0,08\*4+10\*(0,04\*5)+0,04\*4=3,48 бит

**Задание 4.** С помощью стандартного архиватора (WinZip, WinRar, 7-Zip и т.п.) выполнить сжатие различных документов, тип которых указан в таблице

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Документ | Расширение | Объем файла до архивации, Кб | Объем файла после архивации, Кб | Коэффициент сжатия |
| Текст | .doc | 83 | 18 | 4,61 |
| Фотография | .jpg | 9921 | 9916 | 1,0005 |
| Рисунок | .bmp | 6076 | 1688 | 3,6 |
| Видео | .avi | 154048 | 154031 | 1,0001 |
| Звук | .mp3 | 7791 | 7788 | 1,0003 |

Заключение

В ходе лабораторной работы был повторен теоретический материал по темам «сжатие информации», «архивация файлов». С помощью методов RLE, Шеннона-Фано и Хаффмана было выполнено кодирование текстового сообщения, вычислены контрольные суммы и коэффициент сжатия этого сообщения. С помощью архиватора было произведено сжатие файлов различных форматов (doc, jpg, bmp, avi, mp3)

Список использованных источников

1 Советов, Б. Я. Информационная технология [Текст] : Учеб. для студ. вузов по спец. «Автоматизир. системы обраб. информ. и управления» / Б. Я. Советов. – М. : Высш. шк., 1994. – 366 c.

2 Дмитриев, В. И. Прикладная теория информации [Текст] : Учеб. для студ. вузов по спец. «Автоматизированные системы обработки информации и управления» / В. И. Дмитриев. – М. : Высш. шк., 1989. – 320 с. : ил.

3 ГОСТ 7.32–2001. Отчёт о научно-исследовательской работе. Структура и правила оформления [Текст]. – Взамен ГОСТ 7.32–91 ; введ. 2001–07–01. – Минск : Межгос. совет по стандартизации, метрологии и сертификации ; М. : Изд-во стандартов, 2001. – 16 с. – (Система стандартов по информации, библиотечному и издательскому делу).

4 Теоретические основы информационных процессов [Текст] : Учеб. пособие для вузов по спец. «Автоматизация и механизация процессов обработки и выдачи информации» / Л. Ф. Куликовский, В. В. Мотов. – М. : Высш. шк., 1987. – 248 с.

Скрытый служебный текст:

рисунков 2,

всего таблиц 2,

источников 4.