|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | |  |  |  | | МИНОБРНАУКИ РОССИИ | | | | Федеральное государственное бюджетное образовательное учреждение  высшего образования  **«МИРЭА – Российский технологический университет»**  **РТУ МИРЭА** | | | | |
|  | Институт информационных технологий (ИТ) |
|  | Кафедра инструментального и прикладного программного обеспечения (ИиППО) |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ОТЧЕТ ПО ПРАКТИЧЕСКОЙ РАБОТЕ** | | | | |
| **по дисциплине** | | | | |
| **«Структуры и алгоритмы обработки данных»** | | | | |
| Выполнил студент группы ИКБО-01-19 | | Солуянов М.А. | | |
|  | |  | | |
| Принял | | Грушицын А.С. | | |
| Практические работы выполнены | «\_\_\_»\_\_\_\_\_\_\_2020 г. | | (подпись студента) |
| «Зачтено» | «\_\_\_»\_\_\_\_\_\_\_2020 г. | | (подпись руководителя) |
|  |  | |  |

Москва 2020

# Практическая работа №14: «Алгоритмы сжатия и кодирования данных»

**Задача.**

Провести кодирование исходной строки символов «Солуянов Максим Алексеевич» с использованием алгоритма Хаффмана. Для этого:

1. Построить таблицу частот встречаемости символов
2. Отсортировать алфавит в порядке убывания частот появления символов
3. Построить дерево кодирования Хаффмана
4. Присвоить коды каждой букве
5. Провести кодирование исходной строки и рассчитать коэффициенты сжатия относительно кодировки ASCII и относительно равномерного кода.
6. Рассчитать среднюю длину полученного кода и его дисперсию
7. **Таблица частот**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Алфавит | с | о | л | у | я | н | в | < > |
| Кол. вх. | 3 | 2 | 2 | 1 | 1 | 1 | 2 | 2 |
| Вероятн. | 0.115 | 0.077 | 0.077 | 0.038 | 0.038 | 0.038 | 0.077 | 0.077 |

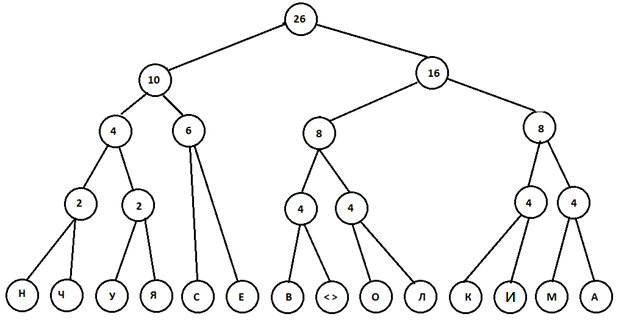
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Алфавит | м | а | к | и | е | ч |
| Кол. вх. | 2 | 2 | 2 | 2 | 3 | 1 |
| Вероятн. | 0.077 | 0.077 | 0.077 | 0.077 | 0.115 | 0.038 |

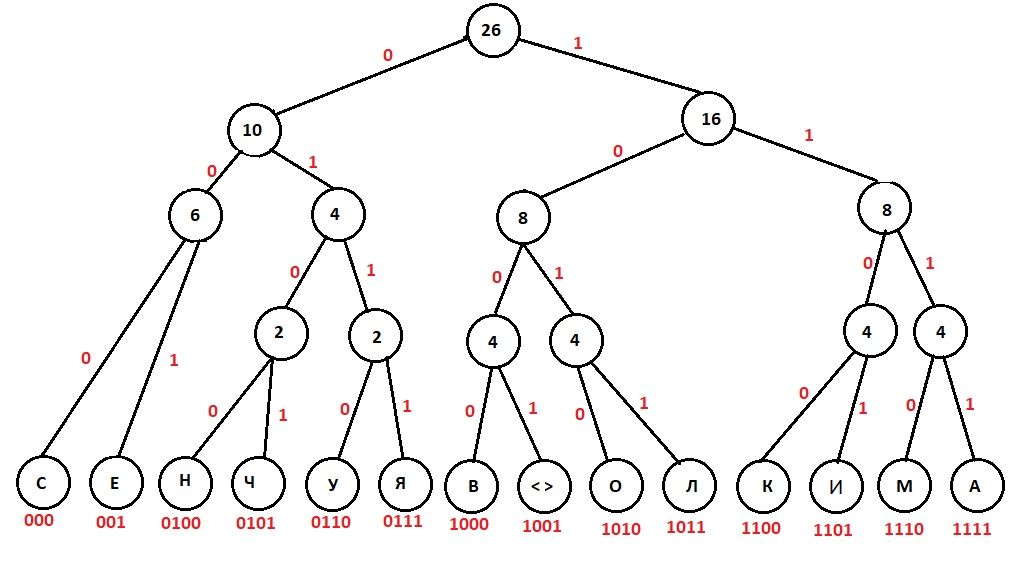
Всего – 26 символов

1. **Отсортированный алфавит**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Алфавит | с | е | о | л | в | < > | м | а | к | и |
| Кол. вх. | 3 | 3 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| Вероятн. | 0.115 | 0.115 | 0.077 | 0.077 | 0.077 | 0.077 | 0.077 | 0.077 | 0.077 | 0.077 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Алфавит | у | я | н | ч |
| Кол. вх. | 1 | 1 | 1 | 1 |
| Вероятн. | 0.038 | 0.038 | 0.038 | 0.038 |

1. **Дерево кодирования Хаффмана**
2. **Коды символов**



1. **Кодирование исходной строки и коэффициенты сжатия**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| с | о | л | у | я | н | о | в | < > | м | а | к | с | и | м |
| 000 | 1010 | 1011 | 0110 | 0111 | 0100 | 1010 | 1000 | 1001 | 1110 | 1111 | 1100 | 000 | 1101 | 1110 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| < > | а | л | е | к | с | е | е | в | и | ч |
| 1001 | 1111 | 1011 | 001 | 1100 | 000 | 001 | 001 | 1000 | 1101 | 0101 |

Рассчитаем коэффициент сжатия относительно использования кодировки ASCII (8 бит/символ):

Рассчитаем коэффициент сжатия относительно равномерного кода (4 бит/символ):

1. **Средняя длинна полученного кода и его дисперсия**  
     
   Рассчитаем среднюю длину полученного кода по формуле:  
     
   где s — множество символов алфавита; — вероятность появления символа; — количество бит в коде символа

Дисперсия рассчитывается по формуле:

**Вывод**

Я изучил алгоритм оптимального префиксного кодирования Хаффмана и применил знания на практике, закодировав своё ФИО