**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ**

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ**

**ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ**

**«ДОНСКОЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ» (ДГТУ)**

Факультет «Информатика и вычислительная техника»

Кафедра «Кибербезопасность информационных систем»

**Лабораторная работа №** **1**

на тему «Изучение методов Шеннона – Фано и Хафмана по построению эффективных кодов»

|  |
| --- |
| Выполнил: студент группы ВКБ43 |
| Ковалев Данил Петрович |
| (Фамилия, имя, отчество) |
| Проверил: |
| Егорова Римма Викторовна |
| (Фамилия, имя, отчество) |

**Цель работы** - изучить возможности эффективного кодирования информации по методам Шеннона – Фано и Хаффмана.

Данная лабораторная работа предполагает выполнение следующих этапов:

1) Изучить методические указания к лабораторной работе.

2) Пройти собеседование с преподавателем и получить задание для выполнения работы.

3) На основе заданного первичного алфавита и вероятностей появления символов этого алфавита получить в форме таблицы двоичный код Шеннона-Фано.

4) Определить эффективность кода, полученного по методу Шеннона-Фано.

5) Выполнить пункты 3, 4 для метода Хаффмана.

6) Сравнить эффективности методов Шеннона-Фано и Хаффмана. Сделать выводы о полученных результатах.

7) Оформить и защитить отчет по выполнению лабораторной работы.

Вариант 11

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Вар. |  | | | | | | | | | | | |
| S |  |  |  |  |  |  |  |  |  |  |  |  |
| 11 | 0.09 | 0.11 | 0.05 | 0.03 | 0.10 | 0.12 | 0.15 | 0.02 | 0.08 | 0.14 | 0.04 | 0.07 |

## Ход работы:

### Задание 3

Для двоичного кода метод Шеннона-Фано сводится к следующему:

1) Буквы алфавита располагаются в порядке убывания вероятностей.

2) Алфавит букв разбивается на две группы таким образом, чтобы суммарные вероятности букв обеих групп были по возможности равны. Первой группе присваивается символ 1, второй символ – 0.

3) Каждую из образованных групп вновь делят на две части с приблизительно равными суммарными вероятностями и присваивают им 1 и 0. Таким образом, получают вторые цифры кода.

4) Процесс повторяется до тех пор, пока в каждой подгруппе не останется по одной букве.

Результат выполнения алгоритма на заданном первичном алфавите представлены в таблице 1.

Таблица 1 – Кодирование заданного алфавита методом Шеннона-Фано

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Символы |  | Уровни группировки | | | | | Кодовые слова |
| 1 | 2 | 3 | 4 | 5 |
|  | 0.15 | 0 – 0,52 | 0 – 0,29 | 0 – 0,15 |  |  | 000 |
|  | 0.14 | 1 – 0,14 | 001 |
|  | 0.12 | 1 – 0,23 | 0 – 0,12 | 010 |
|  | 0.11 | 1 – 0,11 | 011 |
|  | 0.10 | 1 – 0,48 | 0 – 0,27 | 0 – 0,19 | 0 – 0,10 | 1000 |
|  | 0.09 | 1 – 0,09 | 1001 |
|  | 0.08 | 1 – 0,08 |  | 101 |
|  | 0.07 | 1 – 0,21 | 0 – 0,12 | 0 – 0,07 | 1100 |
|  | 0.05 | 1 – 0,05 | 1101 |
|  | 0.04 | 1 – 0,09 | 0 – 0,04 | 1110 |
|  | 0.03 | 1 – 0,05 | 0 – 0,03 | 11110 |
|  | 0.02 | 1 – 0,02 | 11111 |

Построим дерево для всех символов и соответственно кодов. Результат представлен на рисунке 1.

Изображение выглядит как линия, диаграмма, График

Содержимое, созданное искусственным интеллектом, может быть неверным.

Рисунок 1 – дерево Шенона-Фано

### Задание 4

Определим эффективность кода по методу Шеннона-Фано. Согласно формуле:

=0.15⋅3+0.14⋅3+0.12⋅3+0.11⋅3+0.10⋅4+0.09⋅4+0.08⋅3+0.07⋅4+0.05⋅4+0.04⋅4+0.03⋅5+0.02⋅5 = 3.45  
 Для двоичного кода считаем, что

Теперь имея все данные рассчитаем эффективность данного кода:

### Задание 5

Выполним кодирование заданного алфавита методом Хаффмана.

Для двоичного кода метод Хаффмана сводится к следующему:

1) Буквы алфавита выписываются в столбец в порядке убывания вероятностей.

2) Две последние буквы объединяются в одну вспомогательную букву, которой приписывается суммарная вероятность.

3) Вероятности букв, участвующих, в объединении, и полученная суммарная вероятность вновь располагаются в порядке убывания вероятностей в дополнительном столбце, а две последние буквы объединяются.

4) Процесс продолжается до тех пор, пока не будет получена единственная вспомогательная буква с суммарной вероятностью, равной 1.

Для получения кодовой комбинации, соответствующей данной букве необходимо проследить путь перехода по строкам и столбцам таблицы.

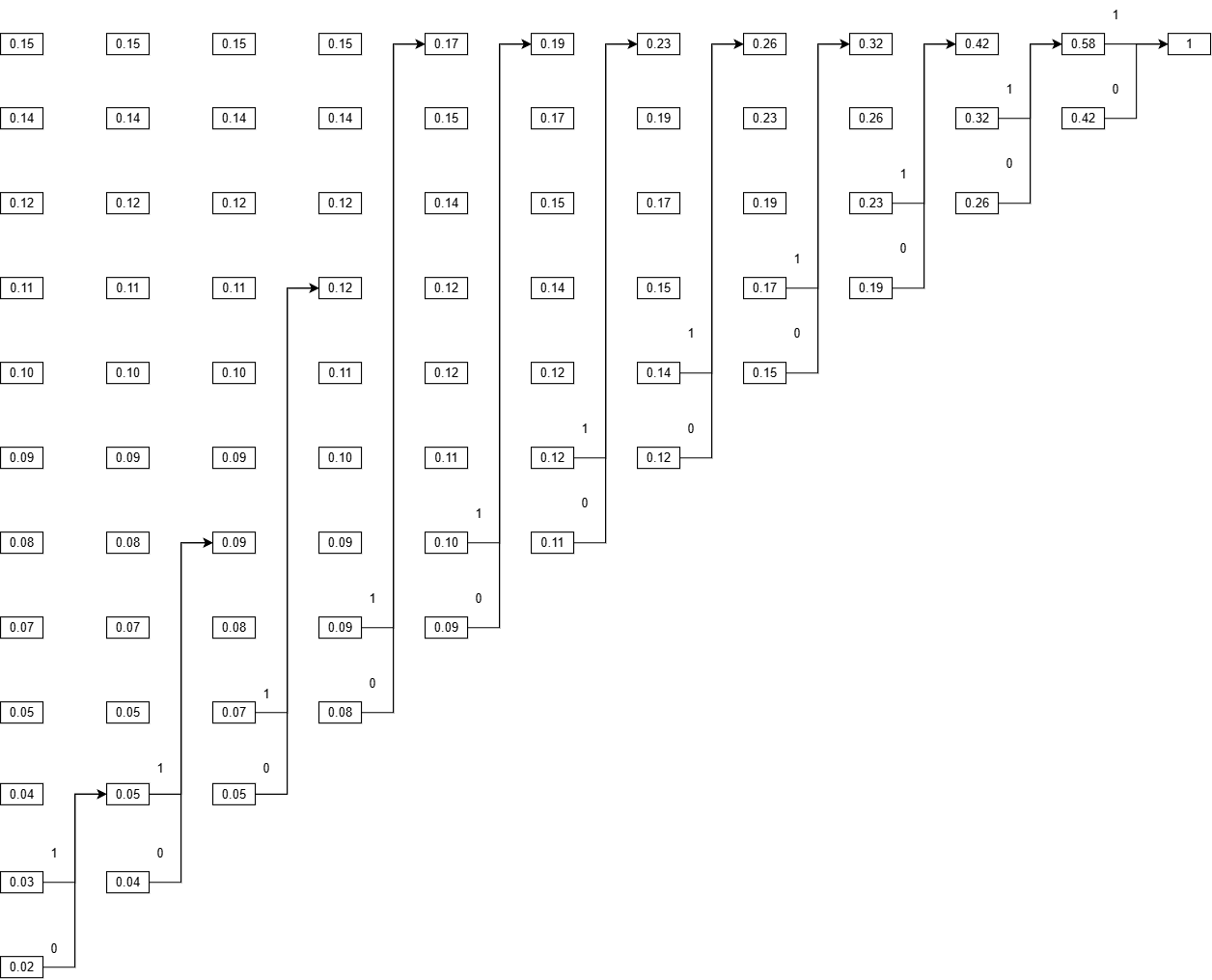


Схема 1 – Кодирование заданного алфавита методом Хаффмана

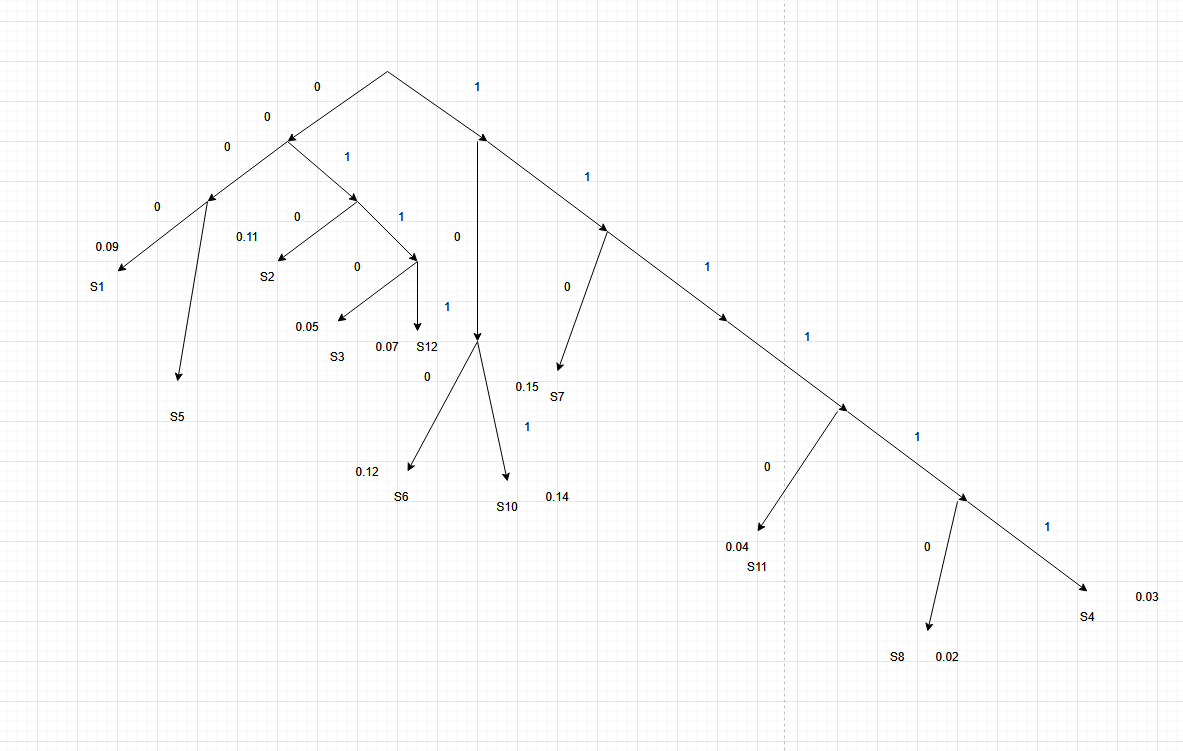


Рисунок 2 – Дерево для алгоритма Хаффмана

На основе схемы составим таблицу сопоставления символов исходного алфавита с их кодовым словом.

Таблица - сопоставление символам исходного алфавита кодовым словам по методу Хаффмана

|  |  |  |
| --- | --- | --- |
|  |  | Кодовое слово |
|  | 0,09 | 000 |
|  | 0,11 | 010 |
|  | 0,05 | 0110 |
|  | 0,03 | 111111 |
|  | 0,10 | 001 |
|  | 0,12 | 100 |
|  | 0,15 | 110 |
|  | 0,02 | 111110 |
|  | 0,08 | 1110 |
|  | 0,14 | 101 |
|  | 0,04 | 11110 |
|  | 0,07 | 0111 |

Определим эффективность кода, созданного на основе заданного алфавита методом Хаффмана:

.

Для двоичного кода считаем, что .

## Вывод:

В данной лабораторной работе были изучены методы кодирования данных Шеннона – Фано и Хафмана по построению эффективных кодов. Для определения более эффективного кода были рассчитаны эффективности для соответствующих кодов на примере варианта № 1. Было установлено, что код Хаффмана, в данном случае, оказался более предпочтительным. Эффективность кода Шеннона-Фано составила 0.579, а эффективность кода Хаффмана – 0.583.