## Федеральное государственное бюджетное образовательное учреждение высшего образования «Сибирский государственный университет телекоммуникаций и информатики» (СибГУТИ)

Кафедра прикладной математики и кибернетики

Практическая работа № 4 Оптимальное побуквенное кодирование

> Выполнил: студент группы ИП-715 Комашко Т.М.

Работу проверила: доцент кафедры ПМ и К Мачикина Е. П. Результаты, полученные при работе программы:

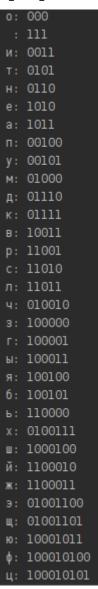


Рисунок 1. Символы с соответствующими им кодами

Средняя длина 4.4040784595144125

Рисунок 2. Средняя длина кода

Средняя энтропия русского языка примерно равна 2.

Метод кодирова ния	Название текста	Оценка избыточно сти кодирован ия	Оценка энтропии выходной послед-ти (частоты символов)	Оценка энтропии выходной послед-ти (частоты пар символов)	Оценка энтропии выходной послед-ти (частоты троек символов)
Метод Хаффмана	Эрнест Хеммингуэ й «По ком звонит колокол» (отрывок на русском)	2.4040784 595144125	0.9961512 06650011 6	0.9961133 26900221 3	0.9959896 80715705 3

## Вывод:

В качестве метода кодирования был выбран метод Хаффмана, который позволяет получить оптимальный префиксный код. Оптимальным данный код является по двум причинам:

- 1) Чем больше вероятность появления символа, тем меньше длина его колового слова:
- 2) Самые длинные коды различаются только в последнем порядке.

При сравнении полученных результатов из предыдущей лабораторных работ и данной, можно заметить, что оценки энтропии после кодирования находятся намного ближе к максимальному значению энтропии, чем в случае художественного текста.

Из-за того, что средняя длина кодового слова вычислялась по 1 источнику, а энтропия текста по другому источнику получилось очень высокое значение энтропии.