Учреждение образования

«БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ»

Кафедра информационных систем и технологий

Отчет

по дисциплине «Защита информации и надежность информационных систем»

Студент: Шумова Е.И.

ФИТ 3 курс, 1 группа, 1 подгруппа

Преподаватель: Савельева М.Г.

Минск 2022

Лабораторная работа №2

**Тема «Элементы теории информации. Параметры и характеристики дискретных информационных систем»**

**Цель:** приобретение практических навыков расчета и анализа параметров и информативных характеристик дискретных ИС

**Задачи:**

* Закрепить теоретические знания по основам теории информации.
* Разработать приложение для расчета и анализа параметров и информативных характеристик дискретных ИС.
* Результаты выполнения лабораторной работы оформить в виде описания разработанного приложения, методики выполнения экспериментов с использованием приложения и результатов эксперимента.

**Теоретические сведения**

*Основные понятия из предметной области.*

Передача информации (данных) осуществляется между двумя абонентами, называемыми *источником сообщения* (ИcС) и *получателем сообщения* (ПС). Третьим элементом информационной системы является *канал (среда) передачи*, связывающий ИсС и ПС.

Отметим также, что и в системах с хранением информации всегда можно выделить ИcС и ПС. В данном случае каналом передачи здесь выступает устройство хранения информации (память). Например, при записи данных в ОЗУ (оперативное запоминающее устройство) компьютера в качестве ИcС и ПС может выступать процессор (соответственно при записи и чтении данных).

Таким образом, простейшая информационная система состоит из трех элементов: источника сообщения, канала передачи сообщения и получателя сообщения.

Отображение сообщения обеспечивается изменением какой-либо физической величины, характеризующей процесс (например, амплитуда, частота, фаза). Эта величина является *информационным параметром сигнала* (в общем случае – информационной системы).

Сигналы, как и сообщения, могут быть *непрерывными* и *дискретными*. Информационный параметр непрерывного сигнала с течением времени может принимать любые мгновенные значения в определенных пределах. Непрерывный сигнал часто называют *аналоговым*, а каналы и устройства, функционирующие на основе такого типа сигналов – *аналоговыми*.

*Дискретный сигнал* (устройство или канал передачи) характеризуется конечным числом значений информационного параметра.

*Дискретные сообщения* состоят из последовательности *дискретных знаков*. Часто этот параметр принимает всего два значения (0 или 1). Сообщение или канал его передачи на основе этих двух значений сигнала называют *двоичным* или *бинарным*.

Построение сигнала по определенным правилам, обеспечивающим соответствие между сообщением и сигналом, называют *кодированием*.

Кодирование в широком смысле – *преобразование сообщения в сигнал*.

Кодирование в узком смысле – *представление исходных знаков*, называемых символами, в другом алфавите с меньшим числом знаков. Оно осуществляется с целью повышения надежности и преобразования сигналов к виду, удобному для передачи по каналам связи. Последний тип кодирования относится к так называемой *прикладной теории кодирования информации*, занимающейся поиском и реализацией методов и средств обнаружения несоответствий (*ошибок*) между переданным Xk и принятым Yk сообщениями.

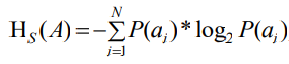
*Понятия и формулы, использованные в лабораторной работе.*

Алфавит, А– это общее число знаков или символов (N), используемых генерации или передачи сообщений. Символы алфавита будем обозначать через {аi}, где 1 ≤ i≤ N; N – мощность алфавита.

Информационной характеристикой алфавита (источника сообщений на основе этого алфавита) является *энтропия*.

Этот термин применительно к техническим системам был введен Шенноном и Хартли.

Энтропию алфавита А={ai} по К. Шеннону рассчитывают по следующей формуле:



С физической точки зрения *энтропия алфавита показывает, какое количество информации (бит) приходится в среднем на один символ алфавита*.

Частным случаем энтропии Шеннона является энтропия Хартли. Дополнительным условием при этом является то, что все вероятности одинаковы и постоянны для всех символов алфавита. С учетом этого формулу можно преобразовать к виду:



Сообщение Хk, которое состоит из k символов, должно характеризоваться определенным *количеством информации*, I(Хk):



Для ДСК: p + q = 1, где q – вероятность правильной передачи бита сообщения, а p – вероятность передачи бита с ошибкой.

He(A) = 1 – H(Y|X), где H(Y|X) – условная энтропия и

H(Y|X) = –p log2 p – q log2 q

**Практические задания**

*Задание 1.*

*Входной документ на латинице (норвежский):*

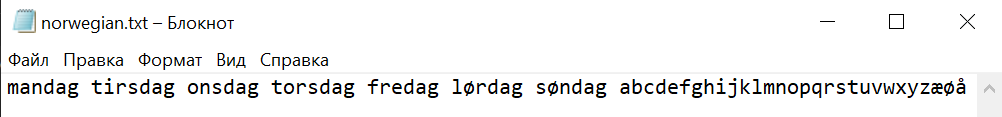


Рисунок 1 – Файл на норвежском языке

Энтропия данного норвежского текста: 4,483591603121156

Частота появления символов и их вероятность:

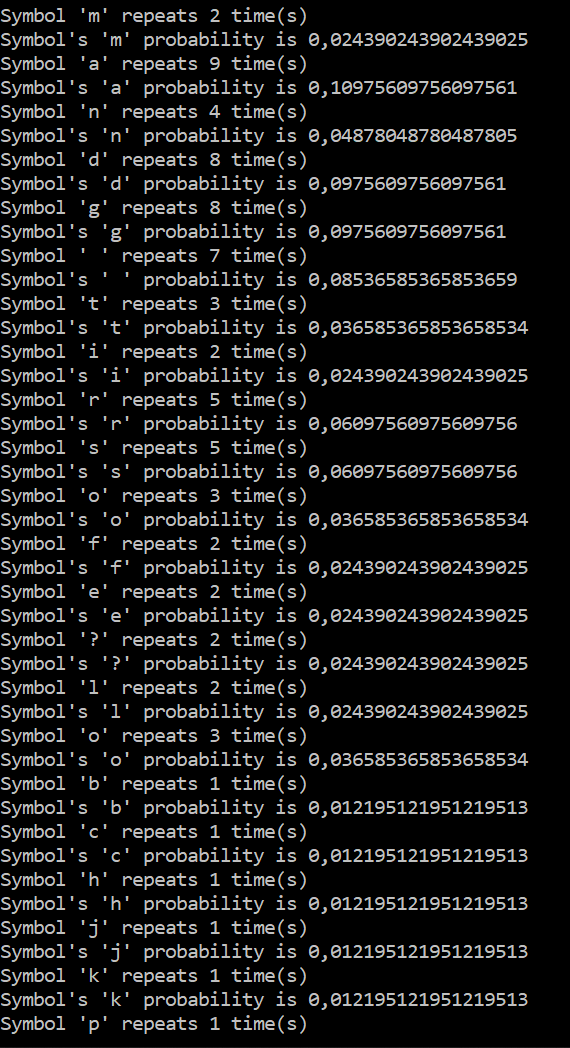


Рисунок 2 – Частота и вероятность символов

Гистограмма частоты появления символов:

Рисунок 3 – Гистограмма частоты символов

*Входной документ на кириллице (старославянский):*

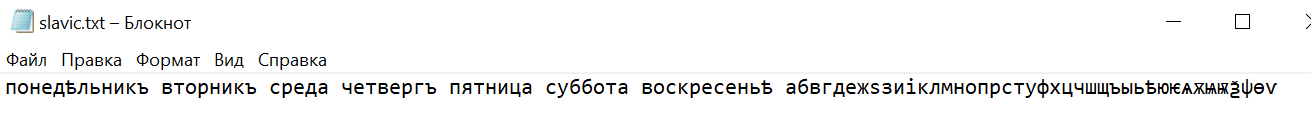


Рисунок 4 – Файл на старославянском языке

Энтропия данного славянского текста: 5,115760249184779

Частота появления символов и их вероятность:

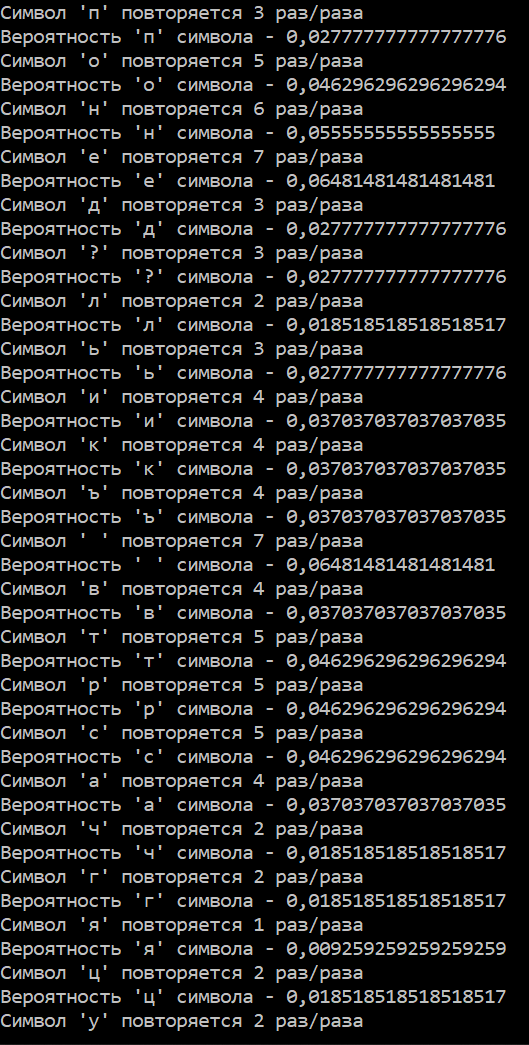


Рисунок 5 – Частота и вероятность символов

Рисунок 6 – Гистограмма частоты символов

*Задание 2.*

Энтропия бинарного алфавита: 1.

*Задание 3.*

ФИО на норвежском: Shumova Yelizaveta Igorevna.

ФИО на славянском: Шоумова Ѥлисавета Игоревна.

Количество информации в сообщении на основе норвежского алфавита: 121,05697328427122 бит.

Количество информации в сообщении на основе славянского алфавита: 133,00976647880427 бит.

Количество информации в сообщении на основе бинарного алфавита: 194,82074866345295 бит.

*Задание 4.*

ФИО на норвежском: Shumova Yelizaveta Igorevna.

ФИО на славянском: Шоумова Ѥлисавета Игоревна.

*С вероятностью ошибочной передачи 0.1:*

Количество информации в сообщении в кодах ASCII для норвежского алфавита: 18,026899061964045 бит.

Количество информации в сообщении в кодах ASCII для славянского алфавита: 17,359236133743156 бит.

Количество информации в сообщении в кодах ASCII для бинарного алфавита: 198,9382386196716 бит.

*С вероятностью ошибочной передачи 0.5:*

Количество информации в сообщении в кодах ASCII для норвежского алфавита: 0 бит.

Количество информации в сообщении в кодах ASCII для славянского алфавита: 0 бит.

Количество информации в сообщении в кодах ASCII для бинарного алфавита: 0 бит.

*С вероятностью ошибочной передачи 1:*

Количество информации в сообщении в кодах ASCII для норвежского алфавита: 26,99556676580855 бит.

Количество информации в сообщении в кодах ASCII для славянского алфавита: 25,995730959667494 бит.

Количество информации в сообщении в кодах ASCII для бинарного алфавита: 198,96732542207045 бит.

**Вывод:** в ходе лабораторной работы были приобретены практические навыки расчета и анализа параметров и информативных характеристик дискретных ИС. Также была разработана программа, вычисляющая энтропию латинского алфавита и кириллицы.