Учреждение образования

«Белорусский государственный технологический университет»

**Отчет по лабораторной работе №2**

по дисциплине: «Информационная безопасность»

Студента 3 курса 6 группы ФИТ

Гвоздовского Кирилла Владимировича

**Лабораторная работа №2**

**ЭЛЕМЕНТЫ ТЕОРИИ ИНФОРМАЦИИ. ИНФОРМАТИВНОСТЬ ДАННЫХ В РАЗЛИЧНЫХ КОДИРОВКАХ**

**Цель:** приобретение практических навыков трансформации данных и сопоставление энтропийных свойств используемых при этом алфавитов.

**Задачи:**

1. Закрепить теоретические знания по взаимной конвертации данных, представленных в кодах ASCII и base64.
2. Разработать приложение для конвертации произвольного документа в формат base64 и обратно.
3. Исследовать энтропийные характеристики используемых в конвертерах алфавитов.
4. Изучить особенности практической реализации операции XOR над данными, представленными в разных форматах.
5. Результаты выполнения лабораторной работы оформить в виде описания разработанного приложения, методики выполнения экспериментов с использованием приложения и результатов эксперимента.

**Задания к выполнению:**

1. Создать собственное приложение или воспользоваться *Base64-онлайн-кодировщиком*, с помощью которого конвертировать произвольный документ (а) на латинице (можно использовать документ из лабораторной работы № 2) в документ (б) формата base64. В качестве входных данных можно использовать указанный преподавателем вариант из списка:

• входные параметры;

• текстовый файл (\*.txt);

• документ Word (\*.doc);

• документ Word (\*.docx);

• документ PowerPoint (\*.ppt, \*.pptx);

• архив (\*.zip);

• текстовая строка;

• случайное число (от 999999);

• PDF-файл;

• архив (\*.rar);

• архив (\*.7z).

2. С помощью приложения, созданного в лабораторной работе №2, получить распределение частотных свойств алфавитов по документам (а) и (б). Вычислить энтропию Хартли и Шеннона, а также избыточность алфавитов. Объяснить полученный результат.

3. Написать функцию, которая принимает в качестве аргументов два буфера (*а* и *b*) одинакового размера и возвращает XOR (собственная фамилия (*а*) и имя (*b*); при разной длине меньшую дополнить нулями). Входные аргументы представлять: 1) в кодах ASCII; 2) в кодах base64. Что будет результатом операции *а*XOR*b*XOR*b*?

**Теоретические сведения**

**Избыточностью алфавита** называют уменьшение информационной нагрузки на один символ вследствие разной вероятности и взаимозависимости появления его символов в сообщениях.

В наиболее общем виде избыточность алфавита R можно оценить отношением энтропии по Хартли и по Шеннону:



Дополнительная избыточность обусловлена переносом сообщения из одной среды в другую или, иначе говоря, кодированием символов исходного алфавита.

*Проблемы с восьмибитной кодировкой (ASCII):*

* «обрезание восьмого бита» в системе электронной почты;
* 256 различных значений позволили уместить в одной кодовой таблице и общепринятые символы и символы кириллицы;
* большинство ПО приспособлены для семибитных кодировок => почтовый сервер при передаче письма обнулял старшие биты в каждом байте сообщения.

*Решение:*

Одним из решений проблемы стала кодировка (а точнее – алгоритм) **base64**. В PGP алгоритм base64 используется для кодирования бинарных данных.

**Кодирование base64 разработано для** представления произвольных последовательностей октетов в форме, позволяющей использовать строчные и прописные буквы.

Используется 65-символьное подмножество набора символов US-ASCII, обеспечивающее представление одним печатным символом 6 битов данных (дополнительный 65-й символ используется для обозначения функции специальной обработки).

Еще раз обратимся к процессу кодировки. Как было выше установлено, каждые 6 битов буфера, начиная с самых старших, используются как индексы строки «ABCDEFGHIJKLMNOPQRSTUVWXYZabcdefghijklmnopqrstuvwxyz0123456789+/», и ее символы, на которые указывают индексы, помещаются в выходную строку.

*Если кодируются только один или два байта, в результате получаются только первые два или три символа строки, а выходная строка дополняется двумя или одним символами «=». Это предотвращает добавление дополнительных битов к восстановленным данным.*

Поскольку входная информация base64 всегда включает целое число октетов, возможны лишь перечисленные ниже случаи:

• размер финального блока кодирования на входе кратен 24 битам, кодированный результат будет содержать целое число 4-символьных групп без заполнения символами «=»;

• размер финального блока кодирования на входе составляет 8 битов, выходной блок будет представлять 2 символа, дополненные последовательностью из двух символов заполнения «==»;

• размер финального блока кодирования на входе составляет 16 битов, выходной блок будет представлять 3 символа, дополненные символом заполнения «=».

**Ход работы**

**Задание 1.** Для конвертирования текста в base64 было разработано приложение, представленное на Рисунке 1

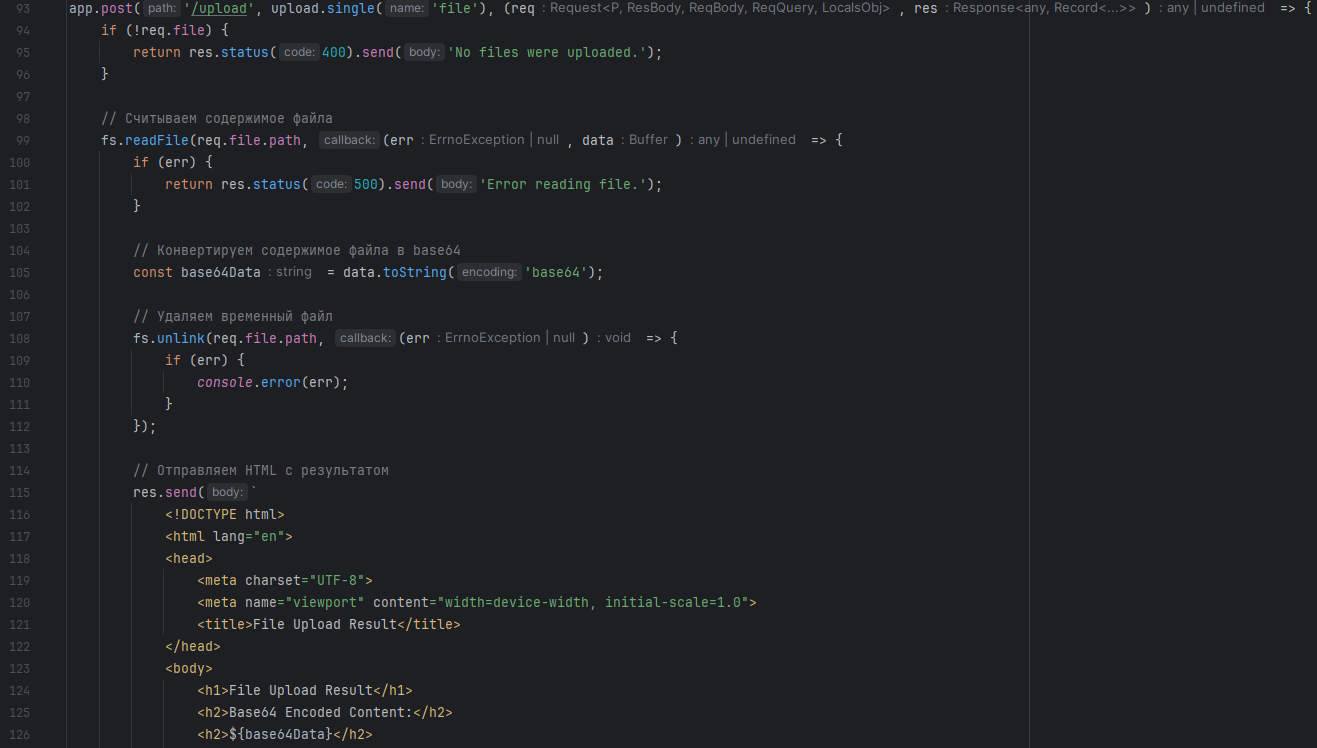


Рисунок 1 – Разработка приложения конвертирования в base64

Входными данными являются два файла с расширением .txt. Результат перевода в Base64 представлен на рисунке 2.

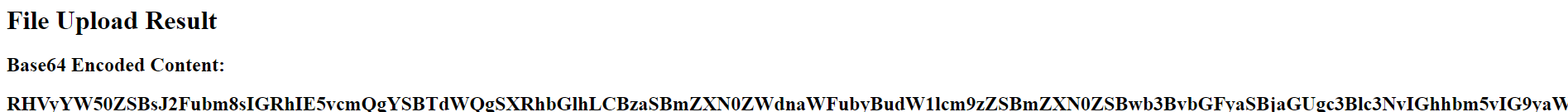


Рисунок 2 – Результат конвертирования в base64

**Задание 2.** Используя функции подчёта энтропии по Шеннону и Хартли, рассчитываем соответствующие энтропии, а также избыточность алфавитов, которая рассчитывается по формуле 1.



Результат выполнения программы приведен на рисунке 3.

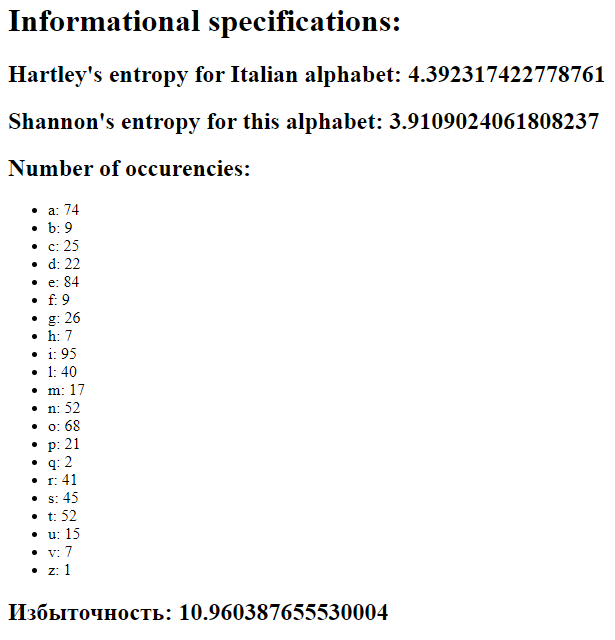


Рисунок 3 – Результат перевода в base64

**Задание 3.** Далее необходимо написать функцию, которая принимает два аргумента (а и b, где а – фамилия, а b – имя) и возвращает XOR. Пример разработанной функции приведен на рисунке 4, результат выполнения на рисунке 5. Бинарный конвертер и конвертер в base64 приведены на рисунке 4.



Рисунок 4 – Реализация программы с выполнением оператора XOR на JavaScript

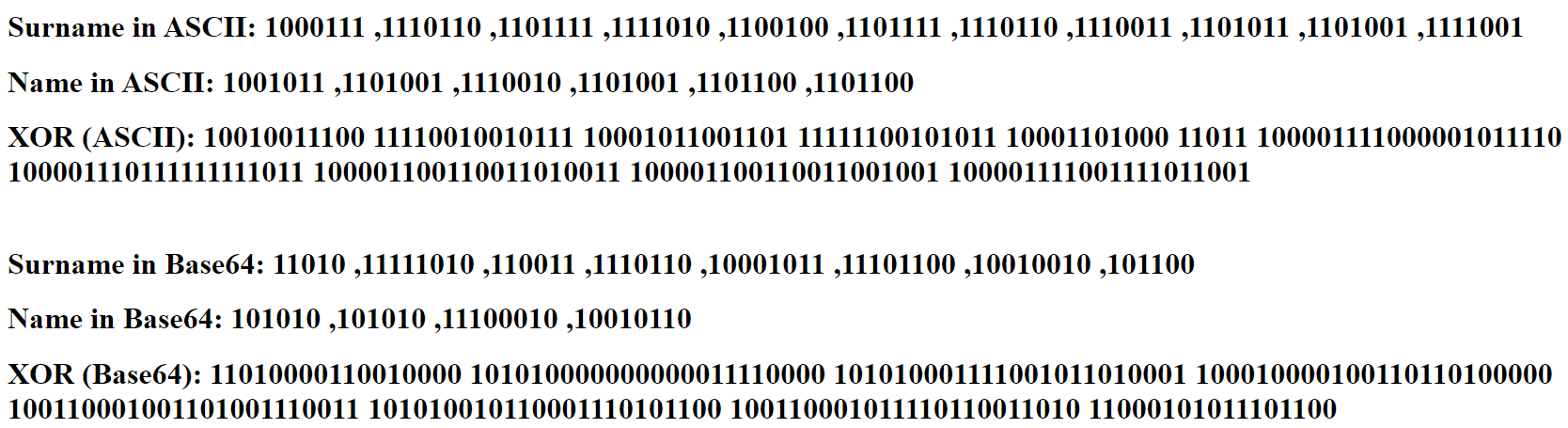


Рисунок 5 – Результат выполнения функции XOR

**Вывод:** в ходе лабораторной работы был изучен алгоритм и особенности кодирования base64. Base64 — стандарт кодирования двоичных данных при помощи только 64 символов ASCII. А также приобретены навыки по переводу текстов на основе естественного языка в base64 и использование оператора XOR.