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 №3

**Тема «Элементы теории информации. Информативность данных в различных кодировках»**

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

**Задачи:**

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

**Теоретические сведения**

Энтропия максимальна при равномерном появлении букв на любом месте сообщения. Для характеристики источника сообщений с различным алфавитом представляет интерес сравнение фактической энтропии источника с максимально возможной. В этом смысле введено понятие *избыточности источника сообщений*, или *избыточности алфавита*.

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

Вычисляется избыточность по формуле:



Уже созданное к тому времени и работающее программное обеспечение зачастую было приспособлено для семибитных кодировок, что приводило, например, к тому, что почтовый сервер при передаче письма обнулял старшие биты в каждом байте сообщения. 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).

*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.*

Если на входе доступно менее 24 битов, входная группа дополняется (справа) нулями до формирования целого числа 6-битных групп. Заполнение в конце данных осуществляется как раз с использованием символа «=». 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 «=».

Если, к примеру, входная строка имеет вид «qwe», то выходная – «cXdl», при расширении входной строки на один символ («qwer») на выходе получим «cXdlcg==».

**Использование операции XOR**

Таблица 1 – таблица истинности операции XOR

|  |  |  |
| --- | --- | --- |
| a | b | a XOR b |
| 0 | 0 | 0 |
| 0 | 1 | 1 |
| 1 | 0 | 1 |
| 1 | 1 | 0 |

Если а и b имеют длину более 1 бита, к примеру 1 байт, то рассматриваемая операция над ними выполняется побитово. Указанным байтам могут соответствовать символы в определенной кодировке. Положим, символу «М» (hex4d) соответствует 8-битный код 01001101, а символу «а» (hex61) соответствует код 01100001, тогда операция сложения по модулю 2 этих двух бинарных кодов дает 00101100 ((hex2с), или символ «,».

**Практические задания**

*Задание 1.*

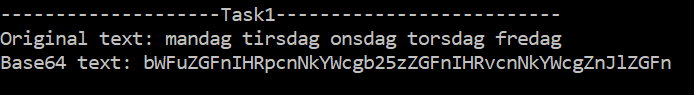


Рис. 1 – Результат выполнения программы для задания 1

*Задание 2.*

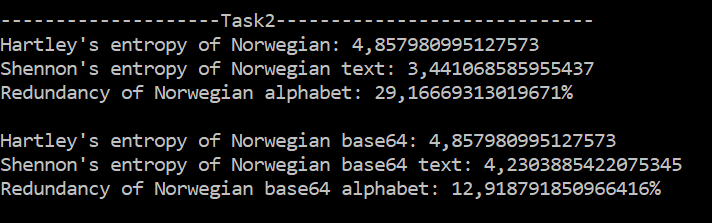


Рис. 2 – Результат выполнения программы для задания 2

Во втором задании вычислялась энтропия Хартли и Шеннона и избыточность алфавита.

*Задание 3.*

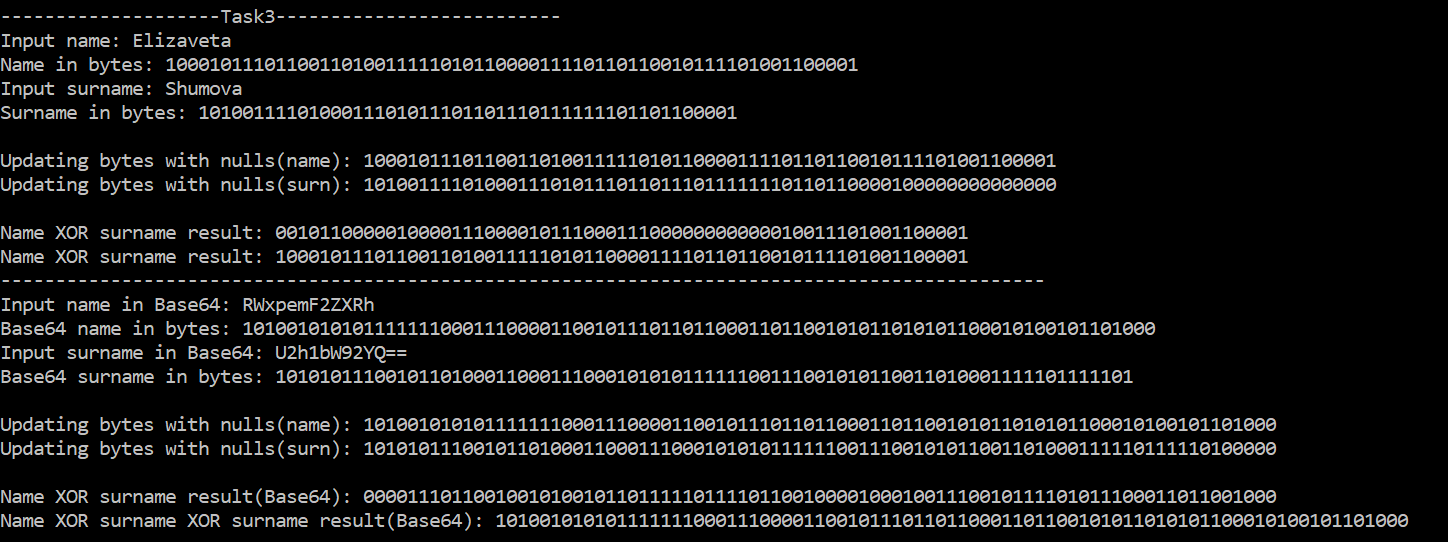


Рис. 3 – Результат выполнения программы для задания 3

В 3-ем задании необходимо было вычислить значения операций a XOR b и a XOR b XOR b для кодировок ASCII и Base64, где a – имя студента, b –фамилия.

**Вывод:** в ходе лабораторной работы были приобретены практические навыки трансформации данных и сопоставлены энтропийные свойства используемых при этом алфавитов.