**Отчет по выполнению лаб. работы №1.**

**Жумигиной Евгении, 12 группа, 4 курс.**

### Задания

1. Реализовать программное средство, осуществляющее шифрование и дешифрование текстового файла, содержащего текст на заданном языке.
   * Реализовать программное средство, осуществляющее криптоанализ зашифрованного по методу Виженера текста. Для криптоанализа использовать тест Касиски.
   * Провести экспериментальное исследование зависимости вероятности успешного проведения атаки по методу Касиски от длины шифротекста.
   * Провести экспериментальное исследование зависимости вероятности успешного проведения атаки по методу Касиски от длины использованного при шифровании ключевого слова.

### Ход выполнения

1. Был подобран тестовый набор данных, являющийся текстом фиксированной длины, взятого из произведения «Pride and prejudice».
2. Исходные документы подверглись предварительной обработке - токенизации. Были исключены знаки препинания, буквы были приведены к единому регистру.
3. Далее произведено шифрование функцией encrypt, которая принимает в качестве параметров исходный текст и ключ. На выходе получаем зашифрованный текст.
4. Атака методом Касиски. Сперва с помощью теста Касиски определяется длина ключа. Определяются l-граммы для l. Далее находим НОД для всех расстояний между одинаковыми l-граммами и определяем длину ключа. Как только длина ключа известна, можно переходить к определению самого ключа. Так как символы зашифрованы циклично с помощью ограниченного последовательного набора сдвигов, равному длине ключа, можно рассматривать группы символов с одинаковым сдвигом. Производится поиск среди этих групп символов наиболее часто встречаемый и разность по модулю с английским символом 'e' (т.к. в английском он встречается чаще всего). Таким образом, последовательно получаем каждый из символов ключа.
5. Также программа предусматривает возможность дешифрования текста.
6. Анализ результатов экспериментов приведен на скриншотах. Представлены данные об использовании метода Касиски на разном размере исходного текста и при разной длине ключа. Линейная зависимость между длиной текста и вероятностью его взлома не прослеживается, что во многом обусловлено недостаточным количеством и малой репрезентативностью выборки исходных данных.



