Лабораторная работа №2

Частотный анализ. Основы криптоанализа

Требования к сдаче работы: в ходе лабораторной работы необходимо сохранить расшифрованные тексты в отдельные файлы, заполнить предлагаемую таблицу и ответов письменно на вопросы.

Основным средством обеспечения конфиденциальности являются различные криптографические алгоритмы. История развития криптографии насчитывает около четырёх тысяч лет. Одними из средств криптографии являются шифры.
Первые шифры были **моноалфавитными.** Например, шифр Атбаш, шифр квадрат Полибия, шифр Цезаря и т.д.
Моноалфавитные шифры подвержены частотному анализу, так как при шифровании выполняется взаимо-однозначное преобразование символов одного алфавита в символы другого (или этого же) алфавита. То есть, зная частотное распределение букв алфавита, можно дешифровать крипто-текст.

Задание 1. Используя самостоятельно разработанную программу для частотного анализа, **расшифруйте предлагаемые тексты**: tis_text1.txt, tis_text2.txt, tis_text3.txt.

Для выполнения частотного анализа необходимо наличие распределения частот появления букв, которое зависит от языка, стиля, автора и прочих свойств текста. Поэтому распределения частот букв для разных текстов могут отличаться. Чем больше длина текстов, тем более сходны распределения частот букв этих текстов. В общем случае криптоанализ на основе частот появления букв представляет подбор из нескольких вариантов соответствия букв закрытого и открытого текста. Для сокращения времени расшифровки предлагается воспользоваться распределениями частот букв, которые содержатся в файлах freq_text1.txt, freq_text2.txt и freq_text3.txt, соответственно для каждого текста.

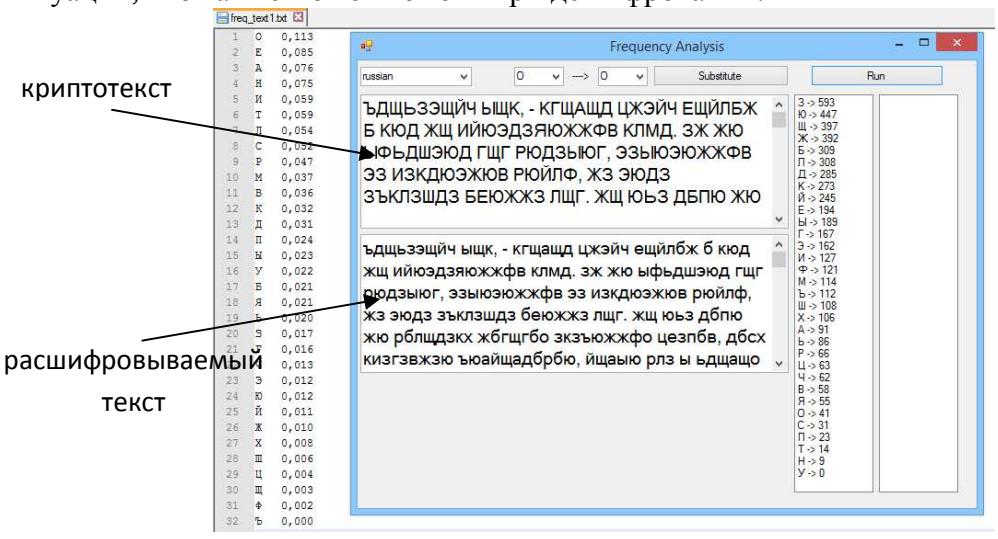
(!) Расшифрованные тексты сохраните для отчета.

Заполните следующую таблицу:

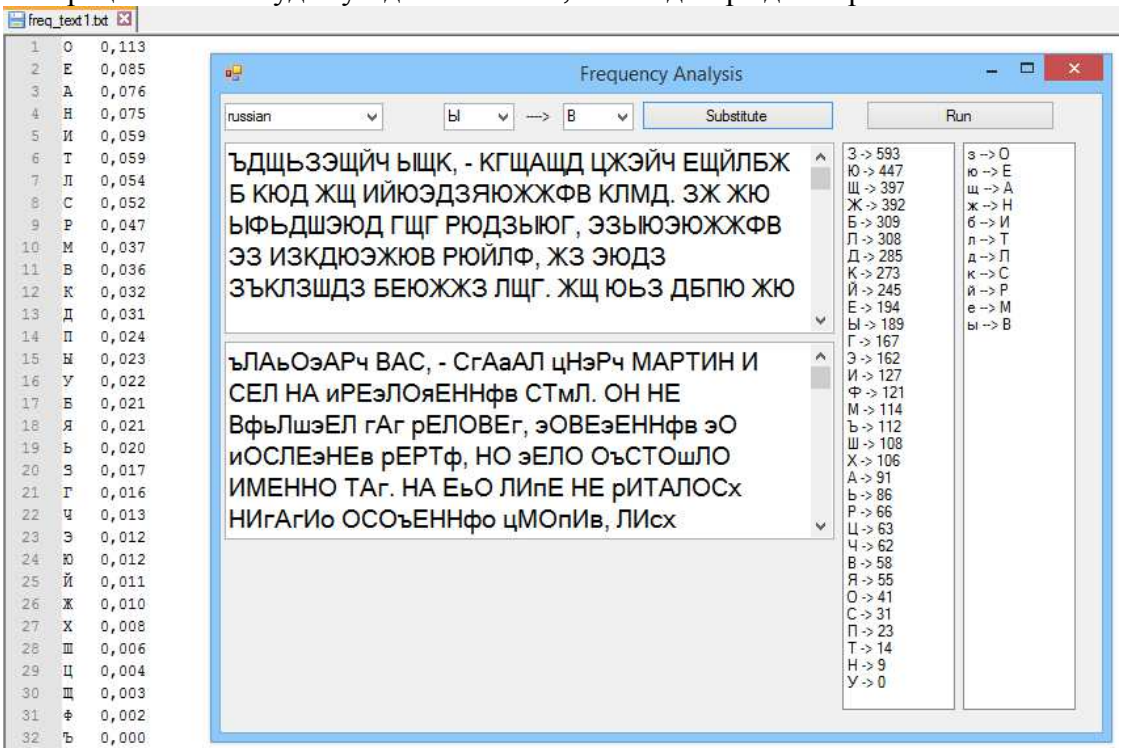
Документ	Файл частот букв	Источник (откуда текст)	Предполагаемый шифр
tis_text1.txt	freq_text1.txt		
tis_text2.txt	freq_text2.txt		
tis_text3.txt	freq_text3.txt		

Проверьте совпадение распределений частот букв расшифрованных текстов с данными для выполнения задания, используя самостоятельно разработанную программу подсчета частот появления букв.

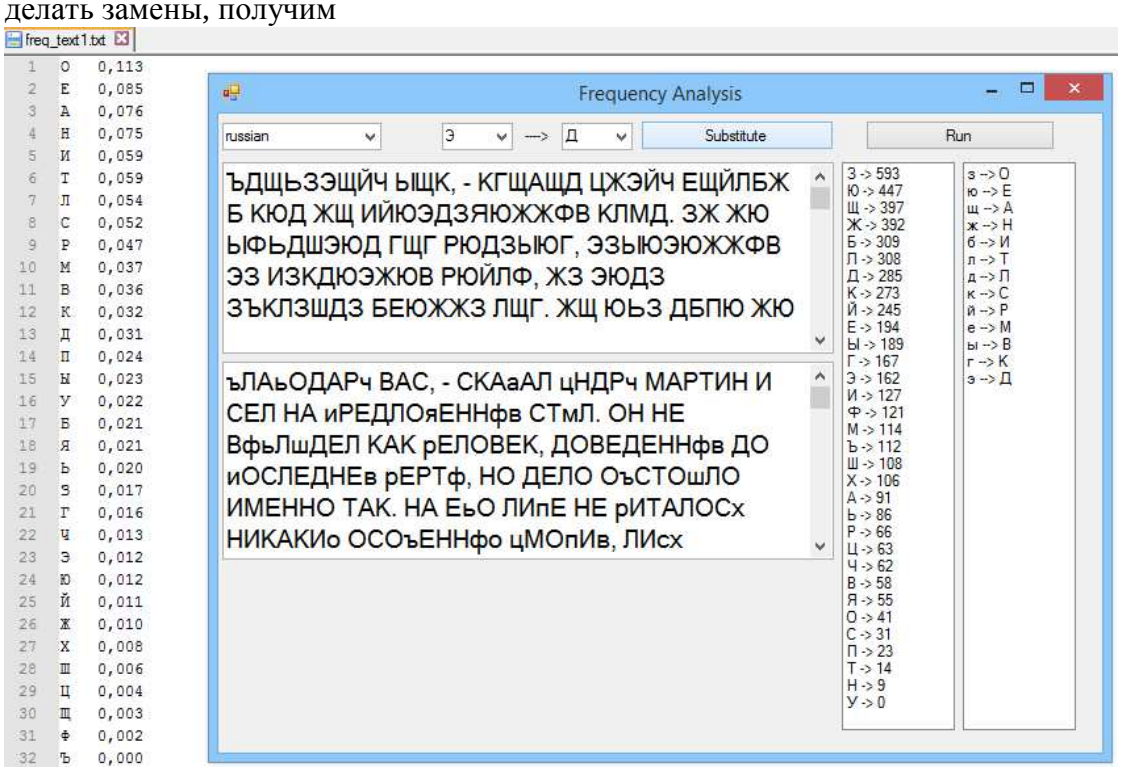
Пример демонстрации программы. Расшифруем текст из первого файла (tis_text1.txt), используя распределение частот из соответствующего файла freq_text1.txt. Заметим, что в криптотексте сохранены знаки пунктуации, что также может помочь при дешифровании.



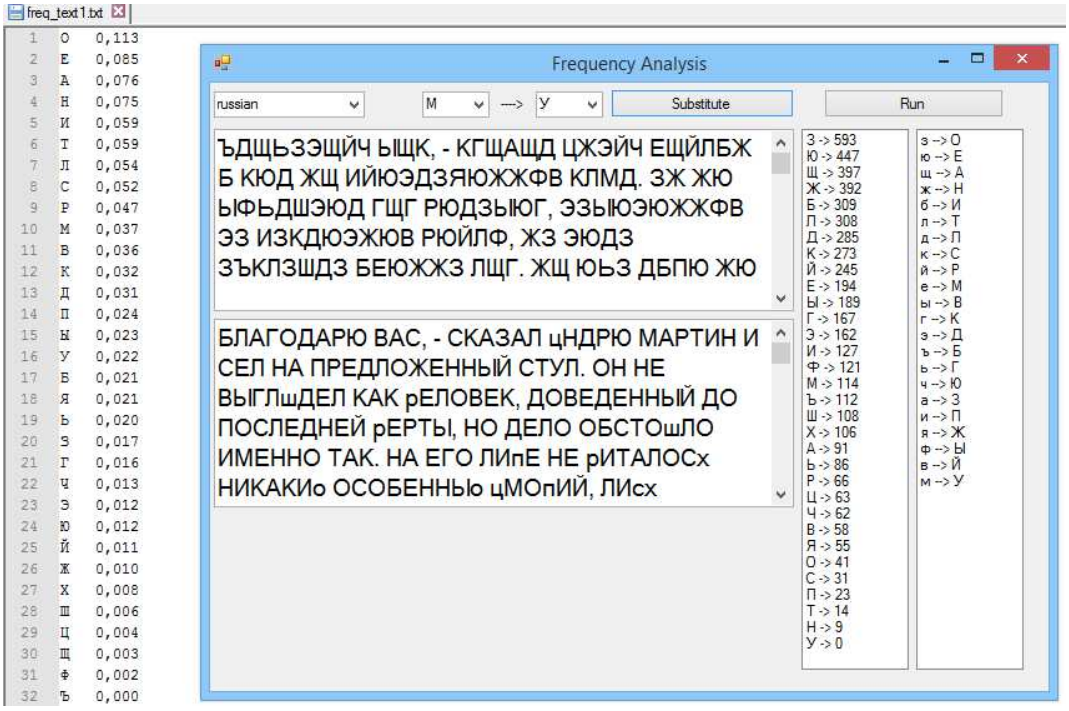
Наиболее часто встречающейся буквой по частотному распределению является буква О, а в криптотексте чаще всех встречается буква З. Поэтому можно сделать замену З->О.
И так далее…
С некоторой итерации можно будет угадывать слова, что подтверждает правильность замен.



В окне расшифрованного текста большие буквы относятся к открытому тексту, а маленькие к криптотексту. После ряда замен можно делать дополнительные предположения о замене букв. Например, третье слово предположительно «сказал», последнее слово в первом предложении – «стол» или «стул». Но слово «стол» не подходит, так как буква О уже была расшифрована.
Продолжая делать замены, получим



Предполагая первое слово «благодарю», можно сделать замены ъ->Б, ь->Г и ч->Ю. Скорее всего третье слово – «сказал», поэтому добавим замену а->З. После слов «мартин и сел на» скорее всего следует «предложенный стул». Поэтому можно добавить замену и->П, я->Ж, ф->Ы, в->Й и м->У. Получим



Таким образом, можно расшифровать текст полностью.

Задание 2. Ответьте **<u>письменно</u>** на следующие вопросы:

1. Приведите пример шифра, который не подвержен частотному анализу. Объясните, благодаря какому свойству достигается стойкость шифра к частотному анализу.

2. Какие части слова «угадываются» лучше при расшифровке? Как вы считайте с чем это связанно?