Лабораторная работа 3.
Взлом моноалфавитного подстановочного шифра методом частотной атаки

6.4.2 Криптоанализ шифра простой замены

Источник https://intuit.ru/studies/courses/13837/1234/lecture/31196?page=4

Криптоанализ шифра *простой замены* основан на использовании статистических закономерностей языка. Приведем таблицы частот букв русского и английского языков (найти самостостельно).

Таблица 6.3. Частоты f(l) букв l русского языка в 32-буквенном алфавите со знаком пробела

l	$f(l)$	l	$f(l)$	l	$f(l)$	l	$f(l)$
-	0,175	О	0,09	Е, Ё	0,72	А	0,062
И	0,062	Т	0,053	Н	0,053	С	0,045
Р	0,040	В	0,038	Л	0,035	К	0,028
М	0,026	Д	0,025	П	0,023	У	0,021
Я	0,018	Ы	0,016	З	0,016	Ь,Ъ	0,014
Б	0,014	Г	0,013	Ч	0,012	Й	0,010
Х	0,009	Ж	0,007	Ю	0,006	Ш	0,006
Ц	0,004	Щ	0,003	Э	0,003	Ф	0,002

Следующая задача посвящена криптоанализу текста на русском языке, зашифрованного двумя способами.

Пример 6.7 *Для этой задачи образцом послужила* [2, задача 6.3] *шифробозначения и текст мы заменили). Первый шифртекст получен из исходного текста перестановкой букв. Второй шифртекст получен из того же исходного текста заменой каждой буквы на другую букву так, что разные буквы заменены разными, а одинаковые - одинаковыми. Восстановите исходный текст.*

Первый шифртекст:

```
И Т Ш И Ь О К Т С О Г М А О Ф О К Е Т А П С С Е О Н С С Ы А
В М Ь Ю З Т Ы Т А Ф О Ь В В Б А С О Ж Е З Т С И Н Й А Я Р Р Р
Т О С Н М Я П Н Н О А Т Ш А О В О
```

Второй шифртекст:

Ф Я Р Ф Р У Ч Р Ф Ц Ы С А Б О В Я О Р Ц А Г Р Ф Ц Р Э Ц Ы Г
Ф И Г Р Х Н Р Ш Ч Д Н В Ц В Т В Н Ч В Ч И Ж Р Ч В Х Д Г В И Ц
Ф Э Ф Ц Р Л Т Р Ф Ц Ы М С А Б О В

Решение. Подсчитаем частоты букв первого шифртекста. Эти частоты будут частотами букв исходного текста, так как первый шифртекст получен перестановкой букв из исходного текста. Одновременно подсчитаем частоты шифробозначений второго текста.

Первый шифртекст	Второй шифртекст
О - 11	Р - 11
Т, А, С - 8	Ц, Ф, В - 8
Н - 5	Ч - 5
В - 4	Г - 4
И, Ь, М, Е, Р - 3	Ы, А, О, И, Н - 3
Ш, К, Ф, П, Ы, З, Я - 2	Я, С, Б, Э, Х, Д, Т - 2
Й, Г, Ю, Б, Ж - 1	У, Ш, Ж, Л, М - 1

Сразу можно сделать вывод: шифробозначению Р соответствует буква О открытого текста, Ч - соответствует Н, а Г - В.

Будем постепенно "проявлять" текст. Запишем второй шифртекст (в нем порядок следования букв не менялся), заменяя уже известные нам шифробозначения их значениями (прописные буквы - это буквы открытого текста, строчные - шифробозначения):

фяОфОуНОфцысабов ...

Далее, шифробозначение Ф скрывает одну из букв: Т, А, С. Шифробозначение Я скрывает либо букву Я, либо согласную. Предположим, что Я - Я, пробуем читать начало:

Ф ⟶ Т: ТЯОТО...
Ф ⟶ А: АЯОАО...
Ф ⟶ С: СЯОСО...

Не читается. Вывод: Я скрывает согласную, причем одну из следующих: Ш, К, Ф, П, З (вариант Я ⟶ Ы мы отбросили сразу). Но если Ф ⟶ Т, то слово не читается. Проверим вариант Ф ⟶ С. Пробуем читать: СяОСОуНОСцы..., и из всех возможных замен для Я подходит только П, тогда начало текста: СПОСОБНОСТЬ, и мы знаем теперь, что Я ⟶ П, У ⟶ Б, Ц ⟶ Т, Ы ⟶ Ь.

Обратим внимание на фрагмент ...яорцагрфцрэцы... С учетом наших знаний это: ПоОТаВОСТОэТЬ, и мы находим еще соответствия: О ⟶ Р, А ⟶ И, Э ⟶ Я. Кроме того, так как Ф ⟶ С, Ц ⟶ Т, то из таблицы: В ⟶ А.

Теперь исследуем последний фрагмент текста: ...фцрлтрфцымсабов. Заменим обозначения их значениями: ...СТОлтОСТЬм... Обратимся к таблице. Обозначению Л соответствует одна из букв Й, Г, Ю, Б, Ж. Пробуем читать, получаем: Л ⟶ Й, и из таблицы: Т ⟶ К. Но тогда М ⟶ Ю, и слово получилось СТОЙКОСТЬЮ.

Теперь начало второй строки: ...АТАКАнНА..., ясно, что Н ⟶ М. Далее, записываем вторую строку:

...АТАКАМНАНижОНАхдВАиТСЯСТОЙКОСТЬЮ...

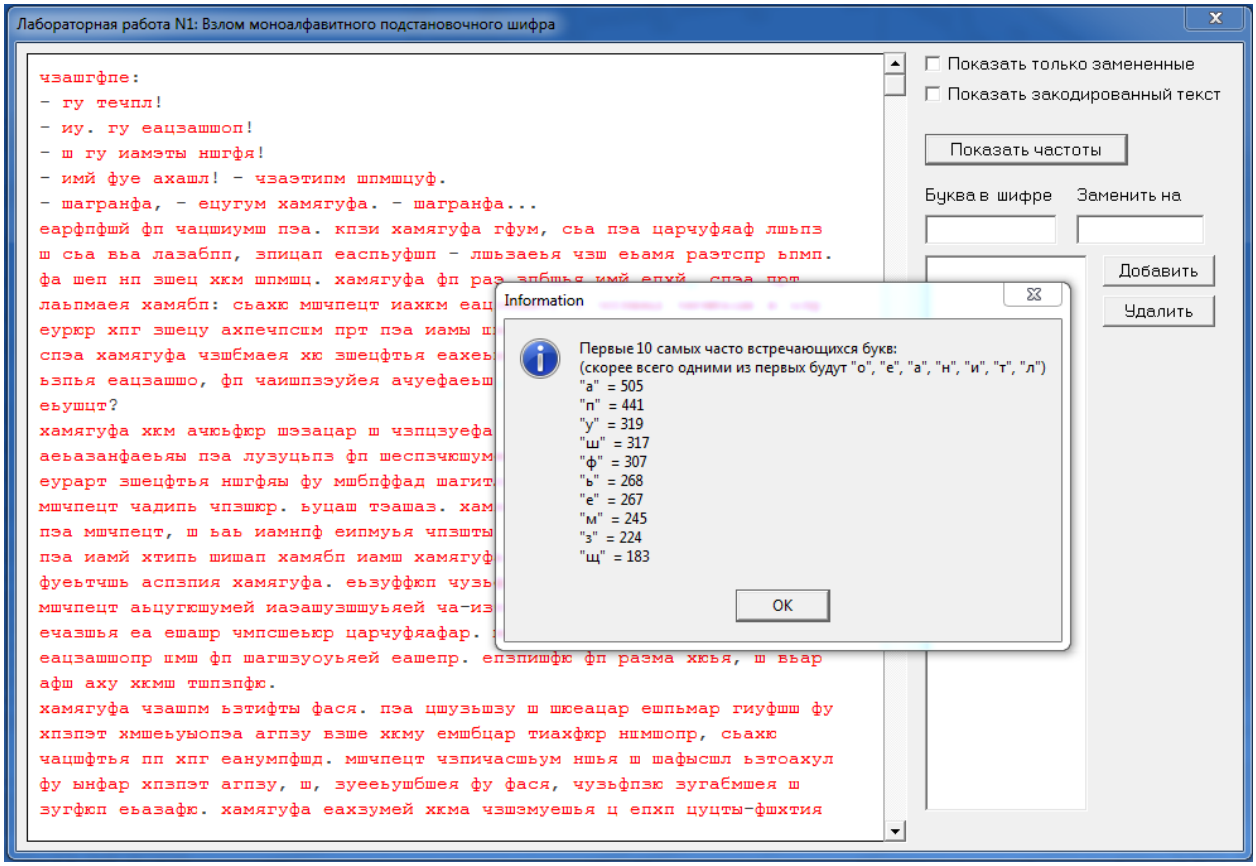
И мы знаем: И ⟶ Е, Ж ⟶ Г, Х ⟶ З, Д ⟶ Ы.

Текст: "Способность шифра противостоять всевозможным атакам на него называется стойкостью шифра".

Содержание отчета по работе (пример)

начало работы 12.58

Начальное окно и частоты букв



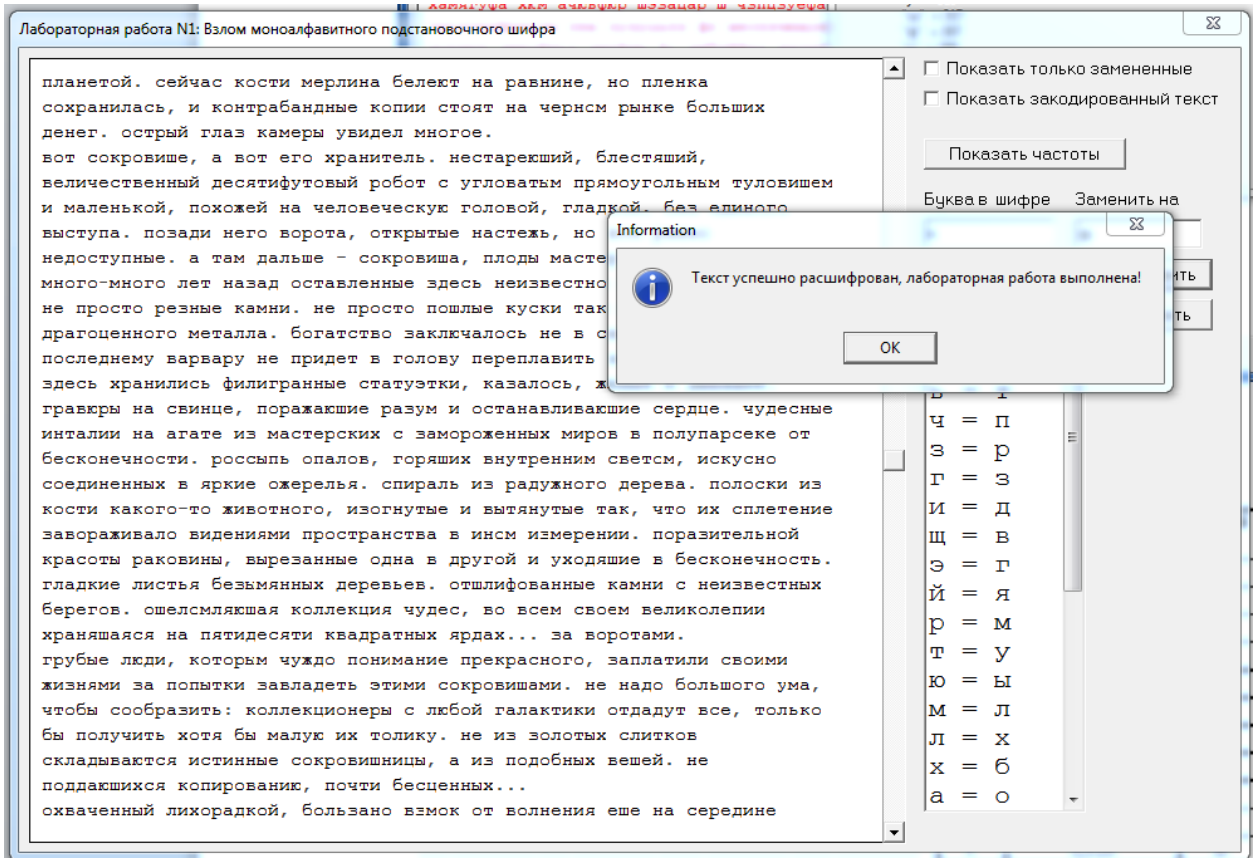
Используем относительные частоты русских букв

*(в таблице 6.3 опечатка – правильная частота для буквы **О = 0,90**)*

Таблица 6.3. Частоты f(l) букв l русского языка в 32-буквенном алфавите со знаком пробела

l	$f(l)$	l	$f(l)$	l	$f(l)$	l	$f(l)$
-	0,175	О	0,09	Е, Ё	0,72	А	0,062
И	0,062	Т	0,053	Н	0,053	С	0,045
Р	0,040	В	0,038	Л	0,035	К	0,028
М	0,026	Д	0,025	П	0,023	У	0,021
Я	0,018	Ы	0,016	З	0,016	Ь,Ъ	0,014
Б	0,014	Г	0,013	Ч	0,012	Й	0,010
Х	0,009	Ж	0,007	Ю	0,006	Ш	0,006
Ц	0,004	Щ	0,003	Э	0,003	Ф	0,002

Буква	Относительная частота	Буква	Относительная частота
а	0,062	р	0,040
б	0,014	с	0,045
в	0,038	т	0,053
г	0,013	у	0,021
д	0.025	ф	0,002
е, ё	0,072	х	0,009
ж	0,007	ц	0,004
з	0,016	ч	0,012
и	0,062	ш	0,006
й	0,010	щ	0,003
к	0,028	ы	0,016
л	0,035	ь, ъ	0,014
м	0,026	э	0,003
н	0,053	ю	0,006
о	0,090	я	0,018
п	0,023		



окончание работы 13.17

таблица замен букв шифра