

Лабораторная работа № 2

Тема занятия: Взлом моноалфавитного подстановочного шифра методом частотной атаки

Цель работы: Ознакомиться на практике с использованием частотной криптоатаки при взломе подстановочных шифров.

Теоретическая часть:

Моноалфавитный подстановочный шифр - шифр, в котором каждой букве исходного алфавита поставлена в соответствие одна буква шифра.

Например, возьмем слово «КУКУРУЗА». Пусть букве «К» текста соответствует буква «А» шифра, букве «У» текста соответствует буква «Б» шифра, букве «Р» текста соответствует буква «В» шифра, букве «З» текста соответствует буква «Г» шифра, букве «А» текста соответствует буква «Д» шифра. После подстановки букв шифра вместо букв исходного теста слово «КУКУРУЗА» в зашифрованном виде будет выглядеть как «АБАБВБГД».

Недостатком подобного шифрования является то, что, если какая-то буква встречается в исходном тексте

чаще всего (например, буква «О» в русском алфавите), то и соответствующая ей буква шифра в зашифрованном тексте также встречается чаще всего.

В нижеприведенной таблице приведены частоты встречаемости букв в английском тексте (в процентах):

Высокая	Средняя	Низкая
E 12,31	L 4,03	B 1,62
T 9,59	D 3,65	G 1,61
A 8,05	C 3,20	V 0,93
O 7,94	U 3,10	K 0,52
N 7,19	P 2,29	Q 0,20
I 7,18	F 2,28	X 0,20
S 6,59	H 2,25	J 0,10
R 6,03	W 2,03	Z 0,09
H 5,14	Y 1,88	

Зная частоты наиболее встречающихся букв и подсчитав, какие буквы чаще всего встречаются в шифровке, криptoаналитик может подобрать расшифровку для некоторых букв текста. Затем, анализируя короткие слова, найти еще буквы, истинные значения которых можно с высокой степенью уверенности предугадать. Например, если уже расшифрована буква «О» и в тексте есть слово «ОБ_ИО» (подчеркнуты уже расшифрованные буквы), то, скорее

всего, шифру «Ы» соответствует буква «Н» в исходном тексте («ОНО»). Чем дальше расшифровывается текст, тем легче идет процесс расшифровки.

Методические указания по выполнению:

1. Разработать программу моноалфавитного подстановочного шифрования (например, методом Цезаря). Программу можно реализовать в одном модуле с программой расшифровки. Исходный и зашифрованный текст должны загружаться и выгружаться в текстовые файлы.
2. Реализовать программу расшифровки текста методом частотной атаки. Зашифрованный текст должен загружаться из текстового файла. Должна быть возможность сохранения частично расшифрованного текста в файл.
3. Прототип пользовательского интерфейса окна программы расшифровки текста (рис. 1):

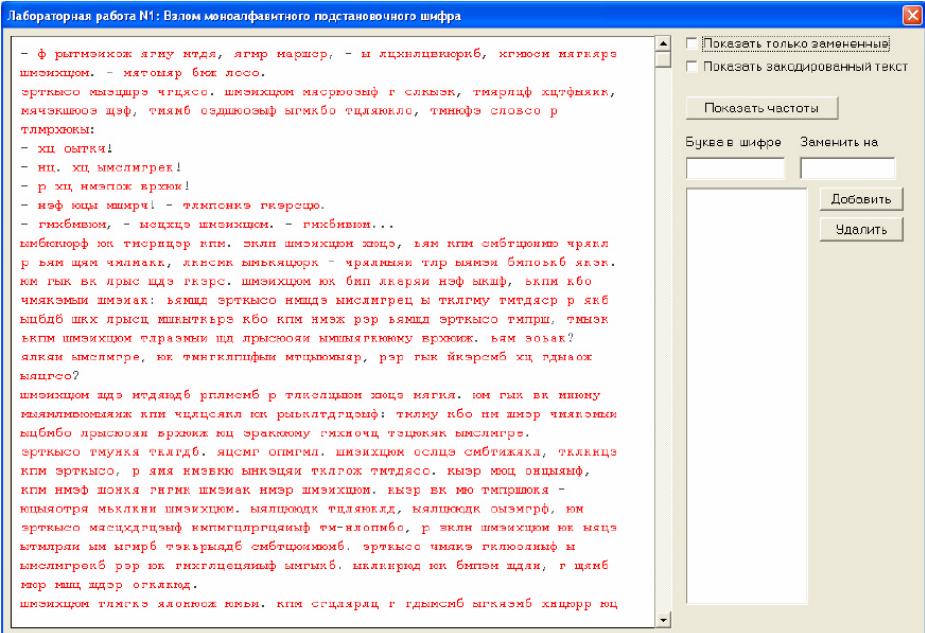


Рисунок 1. Окно программы расшифровки текста

В левой части окна должен находиться зашифрованный текст (буквы, выделенные красным цветом). В процессе расшифровки расшифрованные (правильно или неправильно) буквы текста должны менять цвет с красного на черный.

Необходимо реализовать ввод значений для таблицы замены букв. Например, чтобы указать для какой-либо буквы шифра ее истинное (расшифрованное) значение, нужно в поле «Буква в шифре» указать значение буквы, например, “б”, а в поле «Заменить на» -

ее истинное значение, например, “и”, а затем нажать кнопку “Добавить”. Результат такого действия приведен на рис. 2. На рис. 3. Приведен прототип окна работы программы после добавления расшифровок нескольких букв.

Чтобы отменить указанную расшифровку буквы, нужно в списке расшифровок мышкой указать соответствующую пару букв и нажать кнопку «Удалить» (рис. 4). Полоса вертикального скроллинга должна служить для навигации по расшифровываемому тексту.

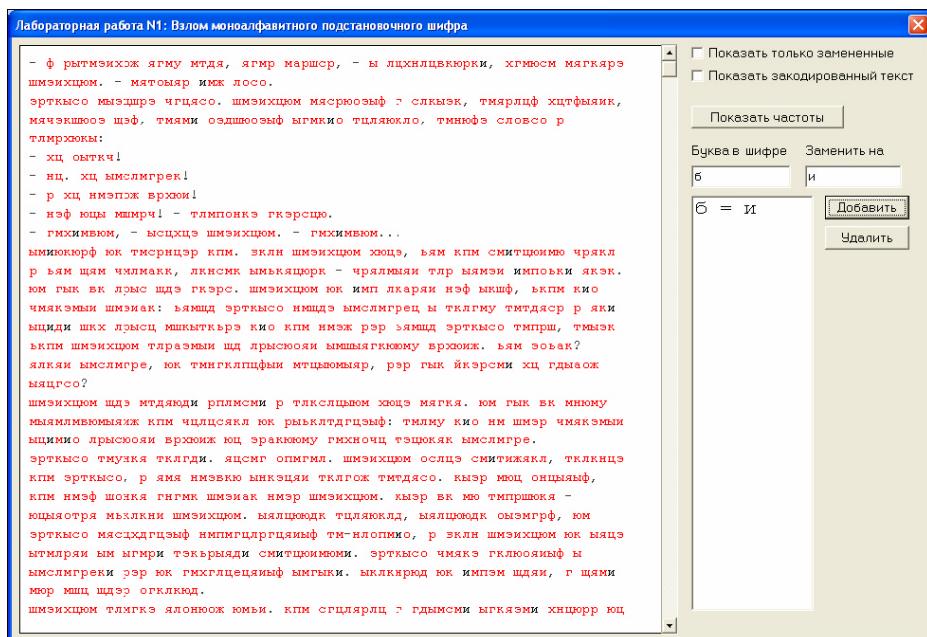


Рисунок 2. Изменения окна лабораторной работы
после расшифровки одной буквы

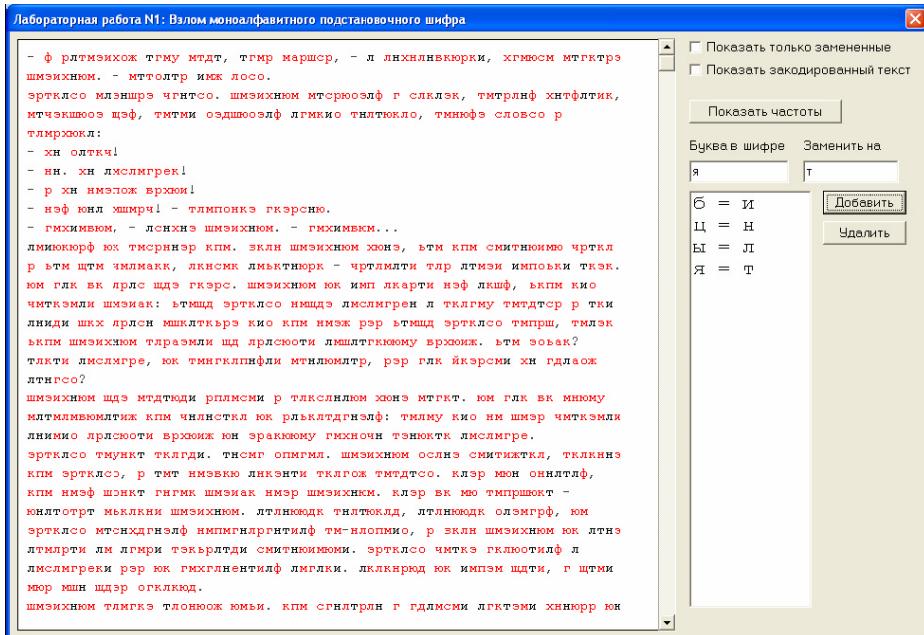


Рисунок 3. Окно лабораторной работы
после расшифровки нескольких букв

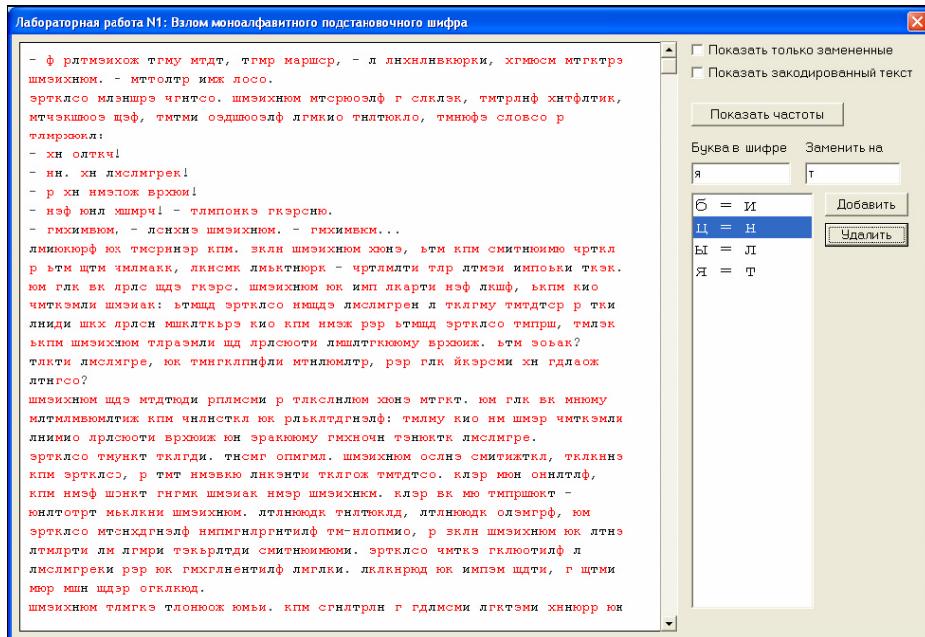


Рисунок 4. Процедура удаления ошибочно указанных расшифрововок

4. Начинается частотная атака с анализа частот встречаемости букв в шифровке. Для этих целей в окне выполнения лабораторной работы нужна кнопка «Показать частоты». При ее нажатии на экран должно выводиться перечень десяти наиболее часто встречающихся букв в шифре, а также перечень букв, наиболее часто встречающихся в русском языке (рис. 5).

Information



Первые 10 самых часто встречающихся букв:
(скорее всего одними из первых будут "о", "е", "а", "н", "и", "т", "л")

"м" = 503
"к" = 437
"р" = 315
"ц" ("ч") = 307
"ю" = 303
"я" ("т") = 267
"ы" ("л") = 262
"з" = 244
"л" = 226
"г" = 182

OK

Рисунок 5. Информация о частотах
встречаемости букв в шифре

Первым шагом в расшифровке текста может быть
указание расшифровки для самой часто встречаемой
буквы - буквы «о». Для случая, приведенного на рис. 5,
указывается «о» как расшифровка буквы «м» шифра.

Следует помнить, что для конкретного текста
частота встречаемости букв может быть несколько иной,
чем в среднем для русского языка. Если в русском языке,
например, буква «т» встречается чаще, чем буква «л», то
в каком-то конкретном тексте буква «л» вполне может
встречаться чаще буквы «т». Поэтому слепо опираться на
данные частотного анализа не следует.

5. В зашифрованном тексте осуществляется поиск коротких слов, зашифрованные буквы которых можно предсказать по уже расшифрованным буквам и частотной информации из рис. 5.

6. Когда так много букв уже известно, зашифрованные буквы могут мешать для понимания слов. Для облегчения дальнейшего анализа в программе должна быть предусмотрена возможность выставления флага «Показать только замененные», при выставлении которого все зашифрованные буквы выводятся на экран в виде символов решетки (рис. 6).

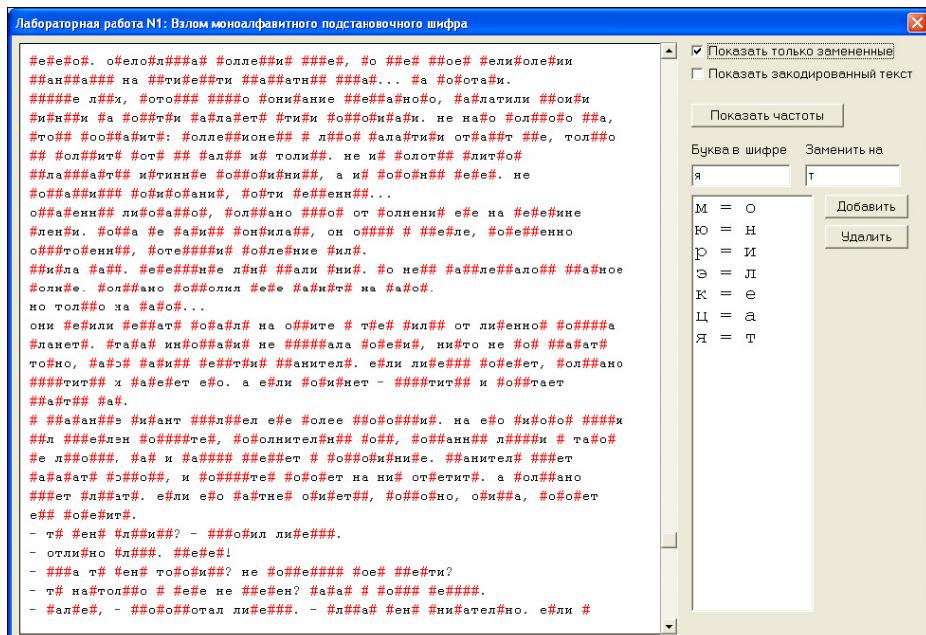


Рисунок 6. Использование флага «Показать только замененные буквы»

Когда же все буквы текста расшифрованы, на экран выводится информационное окно (рис. 7):

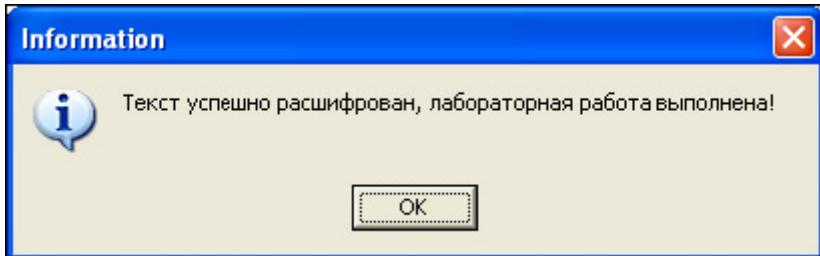


Рисунок 7. Информационное окно,
свидетельствующее об успешной расшифровке текста

Появление этого окна на экране должно свидетельствовать об успешном выполнении лабораторной работы.

Содержание работы:

1. Знакомство с теоретическими основами: принцип моноалфавитного подстановочного шифрования.
2. Реализовать программу моноалфавитного подстановочного шифрования.
3. Реализовать программу расшифровки текста методом частотной атаки.

4. Выполнить расшифровку предоставленного фрагмента.