ЛАБОРАТОРНАЯ РАБОТА №1 «ДОМАШИННЫЕ МЕТОДЫ ШИФРОВАНИЯ»

СОДЕРЖАНИЕ

Теоретический минимум

Перестановочный шифр с ключевым словом

Буквы открытого текста записываются в клетки прямоугольной таблицы по ее строчкам. Буквы ключевого слова пишутся над столбцами и указывают порядок этих столбцов (по возрастанию номеров букв в алфавите). Чтобы получить зашифрованный текст, надо выписывать буквы по столбцам с учетом их нумерации:

Пример:

Открытый текст: `Прикладная математика` Ключ: `Ш и ф р`

```
4  1  3  2
П  р  и  к
л  а  д  н
а  я  м  а
т  е  м  а
т  и  к  а
```

Криптограмма: `Раяеикнаааидммкплатт`

Частотный анализ

Таблицы распределения букв

В русском языке					
Буква	Частота	Буква	Частота	Буква	Частота
а	0.062	л	0.035	ц	0.004
б	0.014	м	0.026	ч	0.012
в	0.038	н	0.053	ш	0.006
г	0.013	о	0.090	щ	0.003
д	0.025	п	0.023	ы	0.016
е	0.072	р	0.040	ъ, ь	0.014
ж	0.007	с	0.045	э	0.003
з	0.016	т	0.053	ю	0.006
и	0.062	у	0.021	я	0.018
й	0.010	ф	0.002	разделитель	0.174
к	0.028	х	0.009		

В английском языке					
Буква	Частота	Буква	Частота	Буква	Частота
a	0.0804	b	0.0154	c	0.0306
d	0.0399	e	0.1251	f	0.0230
g	0.0196	h	0.0549	i	0.0726
j	0.0016	k	0.0067	l	0.0414
m	0.0253	n	0.0709	o	0.0760
p	0.0200	q	0.0011	r	0.0612
s	0.0654	t	0.0925	u	0.0271
v	0.0099	w	0.0192	x	0.0019
y	0.0173	z	0.0009		

Хотя нет таблицы, которая может учесть все виды текстов, но есть вещи общие для всех таблиц, например, в английском языка буква E всегда возглавляет список частот, а T идет на второй позиции. A и O почти всегда третьи. Кроме того девять букв английского языка E, T, A, O, N, I, S, R, H всегда имеют частоту выше, чем любые другие. Эти девять букв заполняют примерно 70% английского текста.

Ниже приведены соответствующие таблицы для различных языков.

Русский	
Буква	Частота
о	0.1090
е	0.0872
а	0.0751
и	0.0751
н	0.0642
т	0.0642
с	0.0545
р	0.0484
в	0.0460
Всего	0.6235

Английский	
Буква	Частота
e	0.1251
t	0.0925
a	0.0804
o	0.0760
i	0.0726
n	0.0709
s	0.0654
r	0.0612
h	0.0549
Всего	0.6990

Немецкий	
Буква	Частота
e	0.1846
n	0.1142
i	0.0802
r	0.0714
s	0.0704
a	0.0538
t	0.0522
u	0.0501
d	0.0494
Всего	0.7263

Французский	
Буква	Частота
e	0.1587
a	0.0942
i	0.0841
s	0.0790
t	0.0726
n	0.0715
r	0.0646
u	0.0624
l	0.0534
Всего	0.7405

Итальянский	
Буква	Частота
e	0.1179
a	0.1174
i	0.1128
o	0.0983
n	0.0688
l	0.0651
r	0.0637
t	0.0562
s	0.0498
Всего	0.7500

Финский	
Буква	Частота
a	0.1206
i	0.1059
t	0.0976
n	0.0864
e	0.0811
s	0.0783
l	0.0586
o	0.0554
k	0.0520
Всего	0.7359

Заметим, что буквы I, N, S, E, A (И, Н, С, Е, А) появляются в высокочастотном классе каждого языка.

Таблица частот биграмм

Таблица частот биграмм отражает количество повторений в тексте определенных пар букв (биграмм). Например, из приведенной ниже таблицы (составленной для некоторого текста) видно, например, что сочетание “ЛА” встречается в нем 25 раз, а сочетание “ПО” – 46.

Часть1																
	А	**Б**	**В**	**Г**	**Д**	**Е**	**Ж**	**З**	**И**	**Й**	**К**	**Л**	**М**	**Н**	**О**	**П**
А	2	12	35	8	14	7	6	15	7	7	19	27	19	45	5	11
Б	5					9	1		6			6		2	21	
В	35	1	5	3	3	32		2	17		7	10	3	9	58	6
Г	7				3	3			5		1	5		1	50	
Д	25		3	1	1	29	1	1	13		1	5	1	13	22	3
Е	2	9	18	11	27	7	5	10	6	15	13	35	24	63	7	16
Ж	5	1			6	12			5					6		
З	35	1	7	1	5	3			4		2	1	2	9	9	1
И	4	6	22	5	10	21	2	23	19	11	19	21	20	32	8	13
Й	1	1	4	1	3		1	2	4		5	1	2	7	9	7
К	24	1	4	1		4	1	1	26		1	4	1	2	66	2
Л	25	1	1	1	1	33	2	1	36		1	2	1	8	30	2
М	18	2	4	1	1	21	1	2	23		3	1	3	7	19	5
Н	54	1	2	3	3	34			58		3		1	24	67	2
О	1	28	84	32	47	15	7	18	12	29	19	41	38	30	9	18
П	7					15			4			9		1	46	

Часть2															
	Р	**С**	**Т**	**У**	**Ф**	**Х**	**Ц**	**Ч**	**Ш**	**Щ**	**Ы**	**Ь**	**Э**	**Ю**	**Я**
А	26	31	27	3	1	10	6	7	10	1			2	6	9
Б	8	1		6						1	11				2
В	6	19	6	7		1	1	2	4	1	18	1	2		3
Г	7			2											
Д	6	8	1	10			1	1	1		5	1			1
Е	39	37	33	3	1	8	3	7	3	3			1	1	2
Ж		1													
З	3	1		2							4				4
И	11	29	29	3	1	17	3	11	1	1			1	3	17
Й	3	10	2				1	3	2						
К	10	3	7	10			1								
Л		3	1	6		4		1			3	20		4	9
М	2	5	3	9	1			2			5	1	1		3
Н	1	9	9	7	1		5	2			36	3			5
О	43	50	39	3	2	5	2	12	4	3			2	3	2
П	41	1		6							2				2

Часть3																
	А	**Б**	**В**	**Г**	**Д**	**Е**	**Ж**	**З**	**И**	**Й**	**К**	**Л**	**М**	**Н**	**О**	**П**
Р	55	1	4	4	3	37	3	1	24		3	1	3	7	56	2
С	8	1	7	1	2	25			6		40	13	3	9	27	11
Т	35	1	27	1	3	31		1	28		5	1	1	11	56	4
У	1	4	4	4	11	2	6	3	2		8	5	5	5	1	5
Ф	2					2			2						1	
Х	4	1	4	1	3	1		2	3		4	3	3	4	18	5
Ц	3					7			10		2				1	
Ч	12					23			13		2			6		
Ш	5					11			14		1	2		2	2	
Щ	3					8			6					1		
Ы		1	9	1	3	12		2	4	7	3	6	6	3	2	10
Ь		2	4	1	1	2		2	2		6		3	13	2	4
Э											1			1		
Ю		2	1	2	1			3	1		1		1	1	1	3
Я	1	3	9	1	3	3	1	5	3	2	3	3	4	6	3	6

Часть4															
	Р	**С**	**Т**	**У**	**Ф**	**Х**	**Ц**	**Ч**	**Ш**	**Щ**	**Ы**	**Ь**	**Э**	**Ю**	**Я**
Р	1	5	9	16		1	1	1	2		8	3			5
С	4	11	82	6		1	1	2	2		1	8			17
Т	26	18	2	10				1			11	21			4
У	7	14	7			1		8	3	2				9	1
Ф	1	1													
Х	3	4	2	2	1			1							
Ц				1							1				
Ч			7	1					1			1			
Ш				1								1			
Щ				1											
Ы	3	9	4	1		16		1	2						
Ь	1	11	3					1	4				1	3	1
Э		1	9												
Ю	1	1	7				1	1		4					
Я	3	6	10			2	1	4	1	1			1	1	1

Коды Азбуки Морзе

Морзе Самюэл Финли Бриз (1791-1872) — американский художник и изобретатель. В 1837 г. изобрел электро-механический телеграфный аппарат. В 1838 г. разработал телеграфный код (азбука Морзе) — неравномерный код, в котором каждая буква или знак представлены комбинацией коротких (точки) и длинных (тире) электрических сигналов.

Русский алфавит	Латинский алфавит	Код Азбуки Морзе
А	**A**	**.-**
Б	**B**	**-...**
В	**W**	**.--**
Г	**G**	**--.**
Д	**D**	**-..**
Е	**E**	**.**
Ж	**V**	**...-**
З	**Z**	**--..**
И	**I**	**..**
Й	**J**	**.---**
К	**K**	**-.-**
Л	**L**	**.-..**
М	**M**	**--**
Н	**N**	**-.**
О	**O**	**---**
П	**P**	**.--.**
Р	**R**	**.-.**
С	**S**	**...**
Т	**T**	**-**
У	**U**	**..-**
Ф	**F**	**..-.**
Х	**H**	**....**
Ц	**C**	**-.-.**
Ч	-	**---.**
Ш	-	**----**
Щ	**Q**	**--.-**
Ы	**Y**	**-.--**
Ь	**X**	**-..-**
Э	-	**..-..**
Ю	-	**..--**
Я	-	**.-.-**

Цифры и знаки препинания	Код Азбуки Морзе
1	**.----**
2	**..---**
3	**...--**
4	**....-**
5	**.....**
6	**-....**
7	**--...**
8	**---..**
9	**----.**
0	**-----**
,	**.-.-.-**
.	**.....**
;	**-.-.-**
:	**---...**
?	**..--..**
№	**-..-.**
"	**.-..-.**
'	**.----.**
( )	**-.--.-**
!	**--..--**
-	**-....-**

Задания

Вариант № 1.

Реализовать перестановочный шифр с ключевым словом. При этом ключевое слово должно задаваться пользователем. Допускается использование фиксированной длины слова. Предусмотреть возможность дешифрования ранее зашифрованного текста.

Вариант № 2.

Реализовать афинную криптографическую систему, выполняющую шифрование по формуле

$$A_{a,b}(j)=(a*j+b)(mod\ n)$$

и обратное преобразование по формуле

$$A^{-1}_{a,b}(j)=(j-b)*a^{-1}(mod\ n).$$

В качестве переменных должны использоваться *a, b* и алфавит. Допускается взаимную простоту *a* и *n* при вводе пользователем не проверять, условившись сделать это условие входным при вводе данных. Предусмотреть возможность дешифрования ранее зашифрованного текста.

Вариант № 3.

Реализовать криптосистему «Доска Полибея»:

	А	**Б**	**В**	**Г**	**Д**	**Е**
А	А	Б	В	Г	Д	Е
Б	Ж	З	И	Й	К	Л
В	М	Н	О	П	Р	С
Г	Т	У	Ф	Х	Ц	Ч
Д	Ш	Щ	Ъ	Ы	Ь	Э
Е	Ю	Я	.	,	-	

Входной переменной является исходный текст, составленный из символов алфавита. Предусмотреть возможность дешифрования ранее зашифрованного текста.

Вариант № 4.

Реализовать криптосистему «Шифр Цезаря с ключевым словом». Входными переменными являются исходный текст, алфавит, смещение и ключевое слово. Предусмотреть возможность дешифрования преобразованного текста.

Вариант № 5.

Реализовать шифр с автоключом. Входными переменными являются исходный текст, алфавит и ключ. Предусмотреть возможность дешифрования преобразованного текста.

Вариант № 6.

Реализовать частотный анализ вводимого текста. В качестве выхода должны выступать статистика по числу повторений символа и % от общего числа введенных символов. Предусмотреть возможность ввода произвольного алфавита.

Вариант № 7.

Применив частотный анализ, определить тип шифра (шифр перестановки или шифр замены).

РГШВГЕЕКГИШБТНВВЦОВБФЕОЫЛАНЖВЕЕВБОГШЕТГ3ЧЕЬНВЕНЕРБВБДНИЖГР
ЕУБАНЖН3ИГРБРЫГТГШГ3ИЕ3РГНЕ3ТЛЯЕТГВРЫРБЖШЕЕРЦФНТРГИ3ИБРСЛРВБУ
БТНЫГШБЛНОБТР3РГЭШНЖНРВЭЕ3ИНОДГЖГВГИИЛШБВНРЦНКЯБТГВАЦТЯНВБ
ИВБАНШВГЕШРГЖЮВСНСГИГЖБЮЛЬНЖТБРЖГШБОРИГРЖНЬЮСБСГВВБОГШЕТ
3ЮРГИЧНКЯНЬДГТНОГКЮЕ3ИРНВВЦНЛДЖБЯВНВЕЮ3СГЖГНЫГЛИНФЕТЕГВРЦ3
ИЖГЕТШГЪДГ3ГА3ИРНВВГЬЛДТБВЛКБРНТЛ3НАЮ3ЛСГВВЛЭМБАЖЕСЛЛИЖГЕТ
ШГОГШЦЕ3ИБТДГУЕИБИЧ3НАЮЛЬВНЕФЕЬУНТГРНСГЬРГРЗНЬГСГТГИСНРУНЬЕВ
НДЖНСГ3ТГРЕТЕНЬЛ3Г3НШЕДЖЕНКЯБРФЕНСВНЬЛЫГ3ИЕИЧ33РГЕЬЕ3НЬНЕ3ИРБ
ЬЕЕ3ГАБСБЬЕРАЛШВЕГВОГШЕТРДТЕ3ГРГЕСЛЖИСНДГДЖБКШВЕСБЬВБШНРБТ3Н
ЖИЛСЕК3ЛСВБШГЬБФВНЕЖБАГИЦ3БЬКБДЕ3ЦРБТЖБ3ОГШЕВЕУНЫГВНУЕИБТСЖ
ГЬН3НВБИ3СЕОРНШГЬГ3ИНЕ

Вариант № 8.

Составить таблицы частот биграмм для 3 текстов на русском языке длиной 1000 символов, взятых из художественной и технической литературы, а также разговорной речи и сравнить их на предмет наиболее часто встречающихся биграмм.

Вариант № 9.

Реализовать возможность кодирования открытого текста и декодирования шифрограммы по правилам азбуки Морзе. Предусмотреть поддержку русского и английского алфавитов.

Вариант № 10.

Реализовать возможность расшифровки криптограммы, полученной путем применения шифра Цезаря, с помощью метода полосок. Входной информацией должны являться алфавит и криптограмма. Количество полосок принять равное 8.

Вариант № 11.

Разработать и реализовать собственный алгоритм шифра перестановки.

Вариант № 12.

Разработать и реализовать собственный алгоритм блочного шифра замены.

Указания по выполнению, оформлению и отчету заданий

- Задания выполняются на любом языке программирования высокого уровня;
- Необходим минимальный пользовательский интерфейс, позволяющий ввести произвольные входные данные и просмотреть результат;
- В качестве отчета выступают:
    - Титульный лист с информацией об исполнителе (ФИО, группа) и лабораторной работе (№, вариант, дата выполнения)
    - Исходный код программы;
    - Экранные формы и/или пользовательские диалоги;
    - Комментарии к алгоритму (по необходимости);
- На отчете могут быть заданы вопросы по соответствующему лекционному материалу.