Криптографические средства

Лекция 2

Технологии защиты информации

Содержание

- История криптографии
- Роль криптографии
- Криптоанализ
- Дешифрование шифра простой замены

Введение

С незапамятных времен существовали три основных способа защиты информации

Первый способ	Второй способ	Третий способ защиты информации
предполагал чисто силовые методы охраны документа (носителя информации) физическими лицами, его передача специальным курьером и т. д.	получил название «стеганография» и заключался в сокрытии самого факта наличия секретной информации	заключался в преобразовании смыслового текста в некий хаотический набор знаков (букв алфавита). Получатель донесения имел возможность преобразовать его в исходное осмысленное сообщение, если обладал «ключом» к его построению. Этот способ защиты информации называется криптографическим.

ШИФР Гая Юлия Цезаря

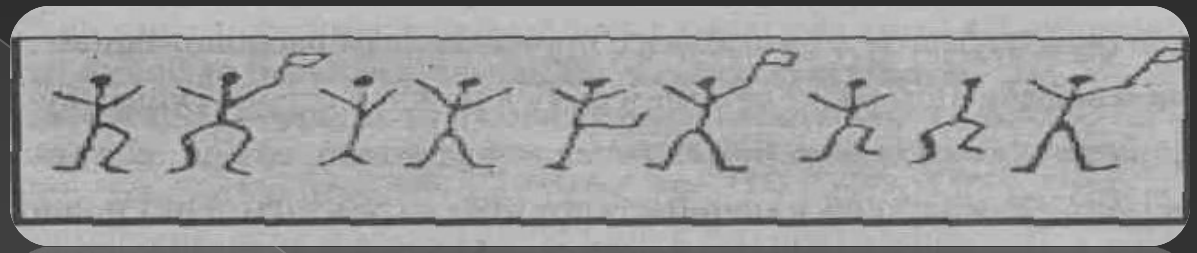
- Выписывался алфавит:
- А, Б, В, Г, Д, Е,...; затем под ним выписывался тот же алфавит, но с циклическим сдвигом на 3 буквы влево:

АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЫЪЬЭЮЯ

ГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЫЪЬЭЮЯАБВ

- При зашифровании буква А заменялась буквой Г, Б заменялась на Д, В - Е и так далее.
- Так, например, слово «РИМ» превращалось в слово «УЛП».
- Ключом в шифре Цезаря является величина сдвига 2-й нижней строки алфавита.

«Пляшущие человечки»



- «I'm here Abe Slaney» («Я здесь Аб Слени»).
- Использован шифр простой замены букв на фигурки людей; флажок в руках означает конец слова.

- Ключом такого шифра являлись позы человечков, заменяющих буквы

ШИФР перестановки

- Выберем целое положительное число, скажем, 5; расположим числа от 1 до 5 в двухстрочной записи, в которой вторая строка - произвольная перестановка чисел верхней строки:

1	2	3	4	5
3	2	5	1	4

- Эта конструкция носит название подстановки, а число 5 называется ее степенью

ШИФР перестановки

1	2	3	4	5
3	2	5	1	4

Зашифруем фразу «СВЯЩЕННАЯ РИМСКАЯ ИМПЕРИЯ».

В этой фразе 23 буквы. Дополним её двумя произвольными буквами (например, Ь, Э) до ближайшего числа, кратного 5, то есть 25. Выпишем эту дополненную фразу без пропусков, одновременно разбив её на пятизначные группы:

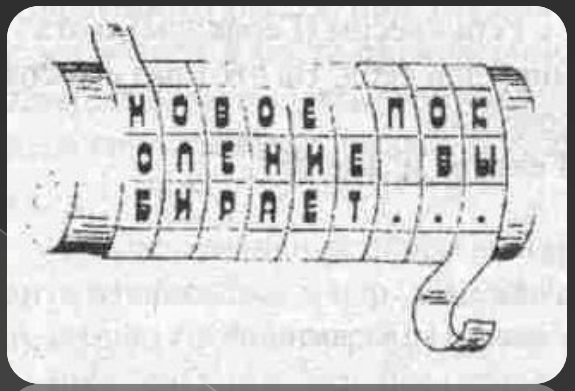
СВЯЩЕ ННАЯР ИМСКА ЯИМПЕ РИЯЬЭ

Буквы каждой группы переставим в соответствии с указанной двухстрочной записью. Полученный текст выписывается без пропусков:

ЩВСЕЯЯННРАКМИАСПИЯЕМЬИРЭЯ

При расшифровании текст разбивается на группы по 5 букв и буквы переставляются в обратном порядке

Прибор Сцитала



- Один из первых физических приборов, реализующих шифр перестановки
- В этом примере ключом шифра являлся диаметр цилиндра и его длина

Прибор Сцитала

Открытый текст выписывается в прямоугольную таблицу из **n** строк и **m** столбцов.

Предполагается, что длина текста **t<n-m** (в противном случае оставшийся участок текста шифруется отдельно по тому же шифру). Если **t** строго меньше **n-m**, то оставшиеся пустые клетки заполняются произвольным набором букв алфавита.

Шифртекст выписывается по этой таблице по заранее оговоренному «маршруту» - пути, проходящему по одному разу через все клетки таблицы.

Ключом шифра являются числа **n**, **m** и указанный маршрут.

Прибор Сцитала

В такой трактовке шифр «Сцитала» приобретает следующий вид.

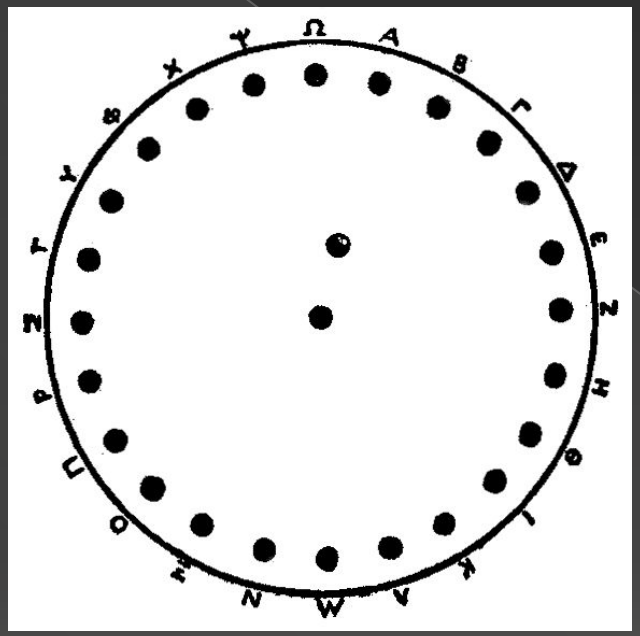
Пусть **m** - количество витков ремня на цилиндре, **n** - количество букв, расположенных на одном витке.

Тогда открытый текст, выписанный построчно в указанную таблицу, шифруется путем последовательного считывания букв по столбцам.

Поскольку маршрут известен и не меняется, то ключом шифра являются числа **m** и **n**, определяемые диаметром цилиндра и длиной ремешка.

При перехвате сообщения (ремешка) единственным секретным ключом является диаметр.

Диск и линейка Энея



+

Нить

Диск Энея

- Шифр, реализуемый линейкой Энея, является одним из примеров шифра замены: в нем буквы заменяются на расстояния между узелками на нитке. Ключом шифра являлся порядок расположения букв по отверстиям в линейке. Посторонний, получивший нить (даже имея линейку, но без нанесённых букв), не сможет прочитать передаваемое сообщение

УЗЕЛКОВОЕ письмо

- Свои сообщения индейцев Центральной Америки передавали в виде нитки, на которой завязывались разноцветные узелки, определявшие содержание сообщения.

КНИЖНЫЙ шифр 1

- Эней предложил прокалывать малозаметные дырки в книге или в другом документе над буквами секретного сообщения.

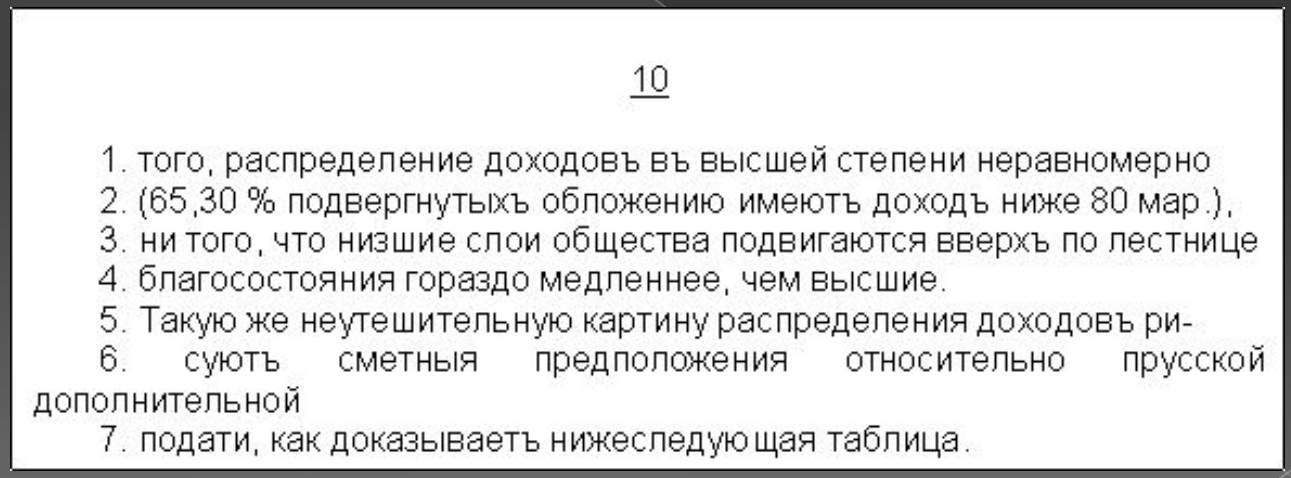
КНИЖНЫЙ шифр 2

- Буквы заменяются на номер строки и номер этой буквы в строке в заранее оговоренной странице некоторой книги. Ключом такого шифра является книга и используемая страница в ней.

КНИЖНЫЙ шифр

- «Наш адрес… Уфа, Уфимское 1/6 5/3 3/5 3/2 3/12 3/1 6/2 7/1 1/8 4/3 1/27 4/2 1/13 3/1 3/14 3/15 3/1 3/2 5/3 1/2 4/2 4/3 6/3 4/4 4/8 1/27 1/8 3/2 4/2 7/2 1/27 7/6 3/7 5/5 1/27 4/3 4/4 3/2 3/1 6/2» (325).

- При шифровке криптограммы были задействованы следующие семь строк 10-й ключевой страницы «Геркнера»:



10

1. того, распределение доходовъ въ высшей степени неравномерно
2. (65,30 % подвергнутыхъ обложению имеютъ доходъ ниже 80 мар.),
3. ни того, что низшие слои общества подвигаются вверхъ по лестнице
4. благосостояния гораздо медленнее, чем высшие.
5. Такую же неутешительную картину распределения доходовъ ри-
6. суютъ сметныя предположения относительно прусской дополнительной
7. подати, как доказываетъ нижеследующая таблица.

Согласно приведенной ключевой книжной таблице криптограмма без труда разбирается: «Акцизное управление. Николаю Гавриловичу Вагину».

КВАДРАТ Полибия

	A	B	C	D	E
A	A	B	C	D	E
B	F	G	H	I	K
C	L	M	N	O	P
D	Q	R	S	T	U
E	V	W	X	Y	Z

T	H	E	A	B
L	C	D	F	G
I	K	M	N	0
P	Q	R	S	и
V	W	X	Y	Z

Пароль: «THE TABLE»

Шифруемая буква заменялась на координаты квадрата, в котором она записана. Так, B заменялась на AB, F на BA, R на DB и т. д.

Секретом в данном случае является сам способ замены букв. Ключ в этой системе отсутствует.

Усложненный вариант шифра Полибия заключается в записи букв в квадрат в произвольном (неалфавитном) порядке. Этот произвольный порядок является ключом.

ТЮРЕМНЫЙ шифр

	1	2	3	4	5
1	A	B	C	D	E
2	F	G	H	I	K
3	L	M	N	O	P
4	Q	R	S	T	U
5	V	W	X	Y	Z

- Стороны квадрата обозначаются не буквами (ABCDE), а числами (12345). Число 3, например, передаётся путём тройного стука. При передаче буквы сначала «отстукивается» число, соответствующее строке, в которой находится буква, а затем номер соответствующего столбца.

Криптоанализ

Основная целью криптографии - хранение открытого текста в тайне от шпионов (взломщиков, перехватчиков, оппонентов, врагов).

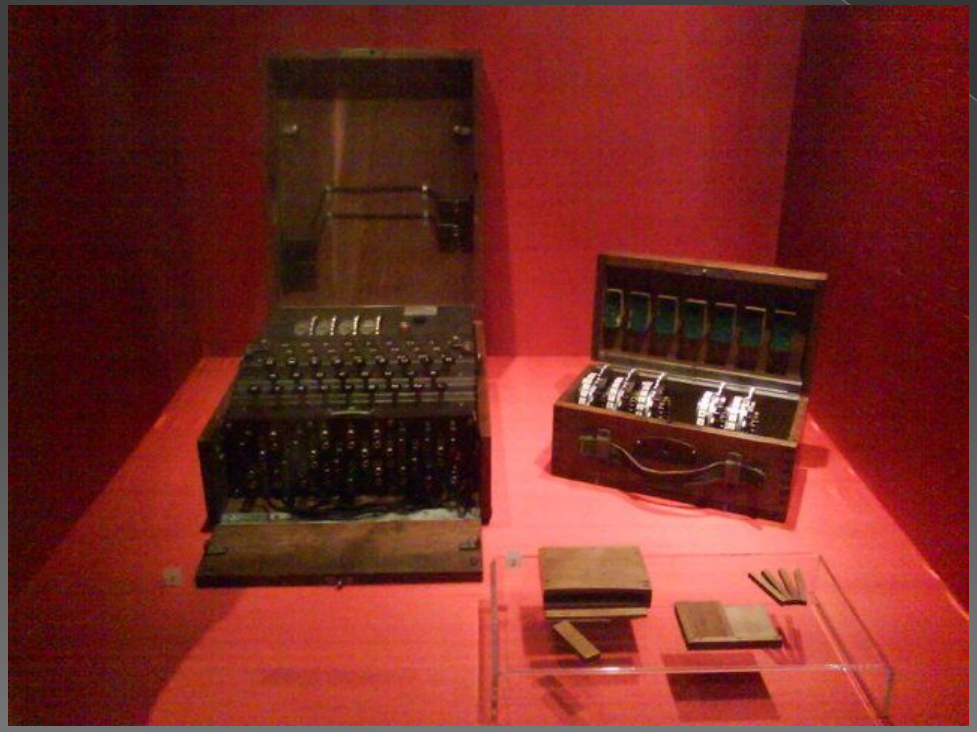
Шпионы предположительно имеют полный доступ к коммуникациям между отправителем и получателем.

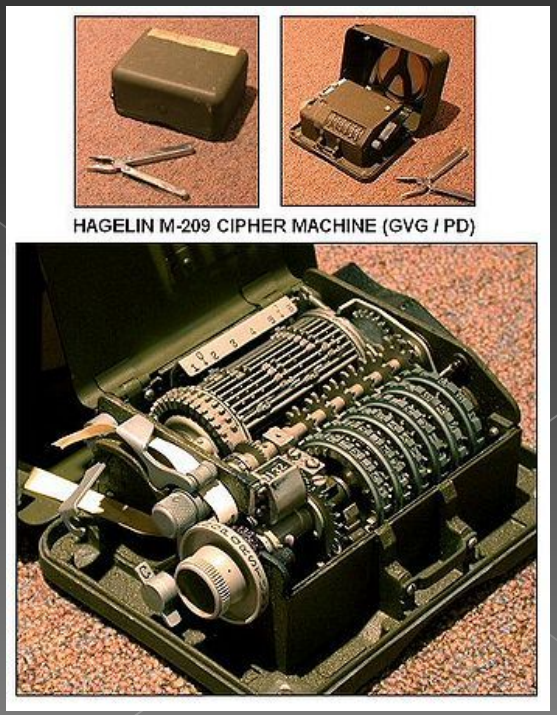
Криптоанализ - наука восстановления открытого текста сообщения без доступа к шифру.

Применение криптоанализа называется атакой.

Криптоанализ

- Классический криптоанализ
  - Частотный анализ
  - метод Касиски
  - дешифровальные машины





Криптоанализ

- Современный криптоанализ
  - В 1998 было обнаружена уязвимость к атакам на основе шифротекста у блочного шифра MADRYGA,
  - Уничтожен блочный шифр FEAL, предложенный как замена DES в качестве стандартного алгоритма шифрования
  - Было установлено, что поточные шифры A5/1, A5/2, блочный шифр CMEA, и стандарт шифрования DECT могут быть взломаны за считанные часы или минуты, а порою и в режиме реального времени.

Основные типы атак криптоаналитика

- *Атака зашифрованного текста*
- *Атака с известным открытым текстом*
- *Атака выборочного открытого текста*
- *Метод «резиновой дубинки»*

Дешифрование шифра простои замены

К наиболее устойчивым закономерностям открытого сообщения относятся следующие:

- 1) В осмысленных текстах любого естественного языка различные буквы встречаются с разной частотой. То же самое можно сказать и о частотах пар, троек букв открытого текста;
- 2) Любой естественный язык обладает так называемой избыточностью, что позволяет с большой вероятностью «угадывать» смысл сообщения, даже если часть букв в сообщении не известна.

Относительные частоты букв алфавита русского языка

1	а - 0,062	12	л - 0,035	23	ц - 0,004
2	б - 0,014	13	м - 0,026	24	ч - 0,012
3	в - 0,038	14	н - 0,053	25	ш - 0,006
4	г - 0,013	15	о - 0,090	26	щ - 0,003
5	д - 0,025	16	п - 0,023	27	ы - 0,016
6	е, ё - 0,072	17	р - 0,040	28	ъ, ь - 0,014
7	ж - 0,077	18	с - 0,045	29	э - 0,003
8	з - 0,016	19	т - 0,Q53	30	ю - 0,006
9	и - 0,062	20	у - 0,021	31	я - 0,018
10	й - 0,010	21	ф - 0,002	32	- 0,175
11	к - 0,28	22	х - 0,009		

Если упорядочить буквы по убыванию вероятностей, то мы получим вариационный ряд

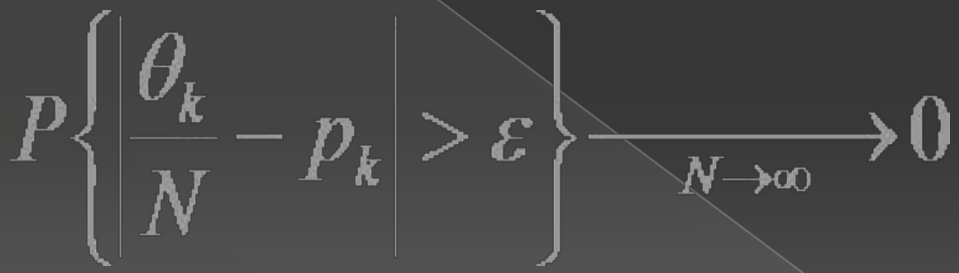
О, Е, А, И, Н, Т, С, Р, В, Л, К, М, Д, П, У, Я, З, Ы, Б, Ь, Г, Ч, Й, Х, Ж, Ю, Ш, Ц, Щ, Э, Ф.

СЕНОВАЛИТР

- Частотная диаграмма зависит от языка.
- Относительные частоты наиболее употребляемых букв некоторых языков.

Английский язык	e - 12,75	t - 9,25	r -8,50	i - 7,75	h - 7,75	o - 7,50
Французский язык	e- 17,75	a - 8,25	s - 8,25	i - 7,25	n - 7,25	r - 7,25
Немецкий язык	e- 18,50	n - 11,50	i - 8,00	r - 7,50	s - 7,00	a - 5,00
Арабский язык	a -17,75					
Греческий язык	a -14,25					
Японский язык	p- 15,75					
Латинский язык	a -11,00					
Малайский язык	a -20,25					
Санскрит	a -31,25					

- Частотная диаграмма является устойчивой характеристикой текста.
- Из теории вероятностей следует, что при достаточно слабых ограничениях на вероятностные свойства случайного процесса справедлив закон больших чисел, т. е. относительные частоты у знаков сходятся по вероятности к значениям их вероятностей P_k:



- Шифры перестановки и простой замены не полностью разрушают вероятностно-статистические свойства, имеющиеся в открытом сообщении.

Дешифрование шифра простой замены

При дешифровании текста, зашифрованного шифром простой замены, используют частотные характеристики открытого текста.

Если подсчитать частоты встречаемости знаков в шифрованном тексте

↓

упорядочить их по убыванию

↓

сравнить с вариационным рядом вероятностей открытого текста

↓

то эти две последовательности будут близки

↓

Скорее всего на первом месте окажется пробел, далее будут следовать буквы О, Е, А, И.

Дешифрование шифра простой замены

Пары соседних букв (биграммы) открытого текста

Наиболее частая биграмма русского открытого текста – СТ.

Для получения устойчивой картины длина последовательности должна быть большой.

Более устойчивой характеристикой биграмм является отсутствие в осмысленном тексте некоторых биграмм, как говорят, наличие запретных биграмм, имеющих вероятность, равную практически 0.

ЪЬ, «гласная»Ь, «пробел»Ь

Знание и использование указанных особенностей открытого текста значительно облегчает дешифрование шифра перестановки и замены.

Пример дешифрования шифра простой замены

- Пусть имеется следующий шифртекст

ДОЧАЛЬ ИЬЦИО ЛИОЙО ВНЫИЮШ ХЕМВЛНХЕИ ДОСОЛЬ ЧСО ИА ТЬЖАТСР ЬАС АКЕИОЙО ДОКЩОКЗЖАЙО КПЗ РТАЩ ТПЬЧНАР ТДО-ТОУН ХЕМВОРНИЕЗ ЕИМОВЛНЯЕЕ РЮУОВ БВЕД СОЙВНМЕЧАТБОГ ТЕТСАЛЮ ЫНРЕТЕС ОС ОТОУ АИИОТСАГ ЕИМОВЛНЯЕЕ АА ЯАИИ-ОТСЕ Е РОЫЛОЦИОТСАГ РПНКАПШЯАР ДО ЫНЖУСА ТРОАГ ЕИМОВЛНЯЕЕ ДВАЦКА РТАЙО ДОКЧАВБЙАЛ УОПШХОА ВНЫИООУ ВНЫЕА РЕКОР ЫНЖЕЖНАЛОГ ЕИМОВЛНЯАА КОБЬЛАИСНПШИНЗ САПАМОИИНЗ САПАРЕЫЕОИИНЗ БОЛДШЭСАВИНЗ БНЦКЮГ РЕК ЕИМОВЛНЯЕЕ УЛААС ТРОЕ ТДАЯЕМЕЧАТБЕА ОТОУАИИОТСЕ Е ФСБ ОТОЧАИИОТСЕ ТЕПШИО РПЕЗЭС ИН РЮУОВ ЛАСОКОР ХЕМВОРНИЕЗ ЕИМОВЛНЯЕЕ УОПШХОА ЫИНЧАИЕА ЕЛАЭС ОУЪАЛЮ Е СВАУЬАЛНЗ ТБОВОТСШ ДАВАКНЧЕ ХЕМВОРНИИОГ

Пример дешифрования шифра простой замены

Подсчет частот символов в шифрованном тексте.

Упорядочим символы по убыванию частот.

Под ними подпишем вариационный ряд вероятностей знаков в открытом тексте.

О Е А И НТСРВЛКМДПУЯЗЫБЬГЧЙХЖЮШЦЩЭФ

При достаточно большой длине шифртекста, для того чтобы из шифрованного текста получить открытый, достаточно заменить b1 на О, b2 на Е, b3 на А и т. д.

Пример дешифрования шифра простой замены

Угадываем замену с учетом статистических особенностей открытого текста.

В шифртексте через пробел, скорее всего, обозначается пробел, через букву О скорее всего обозначена О или А, через Е - О, Е, А, через А - Е, А или И и т.д.

В тексте есть слово АА из двух часто встречающихся букв. В русском языке нет слов ОО, ИИ, НН и т.д. -> ЕЕ, значит Е была заменена на А.

Очень часто в шифртексте слова кончаются биграммами ЕЕ. В русском языке типичными окончаниями являются сочетания ЕЕ, ИИ. Учитывая, что замену для буквы Е мы уже угадали, приходим к выводу, что в шифртексте буква И заменена на Е.

Автоматизация дешифрования шифра простой замены

Поместить в память компьютера словарь русского языка.

просматривая шифртекст, программа осуществляет пробные обратные замены, предполагая, что на данном фиксированном месте в открытом тексте находилось проверяемое слово.

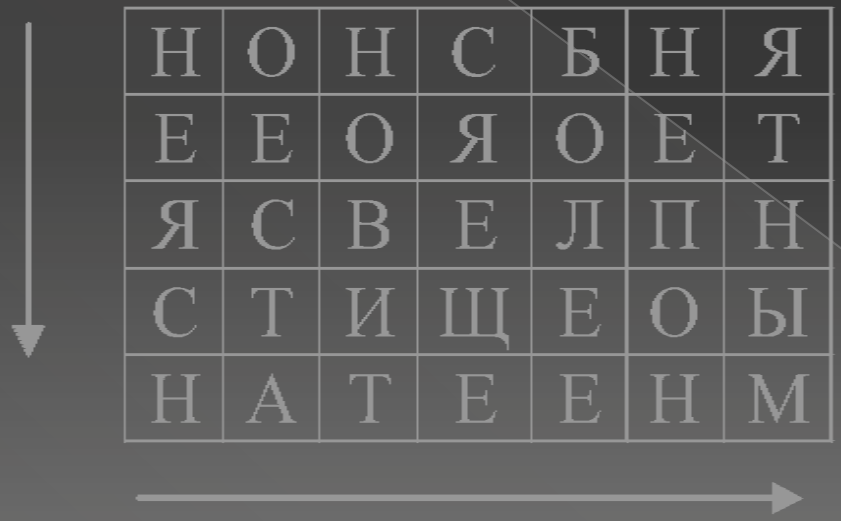
После каждой замены программа частично восстанавливает ключевую подстановку и отсеивает вариант подстановки, если в каком-то месте восстановленного текста оказывается буквосочетание, которое не может быть в открытом тексте.

После восстановления некоторого числа замен в тексте появляются участки, в которых определено значительное число букв.

Оставшиеся буквы подбираются путем перебора словаря и подстановки в текст слов, которые не противоречат восстановленным ранее заменам.

Вскрытие шифров перестановок

Сообщение
«**НЕЯСНОЕ СТАНОВИТСЯ ЕЩЕ БОЛЕЕ НЕПОНЯТНЫМ**»
записывается в таблицу по столбцам.



Метод вскрытия шифров двойной перестановки

Метод вскрытия шифров двойной перестановки основан на определении маловероятных сочетаний букв и нахождении на их основе истинной последовательности столбцов в шифровальной таблице

$$P_{\text{столбца}} = P_{\text{строки 1}} \cdot P_{\text{строки 2}} \cdot \ldots \cdot P_{\text{строки i}} \cdot \ldots \cdot P_{\text{строки n}}$$

или

$$\log P_{\text{столбца}} = \log P_{\text{строки 1}} + \log P_{\text{строки 2}} + \ldots + \log P_{\text{строки n}}$$

- Была перехвачена радиограмма противника: АЗЮЖЕ СШГТООИПЕР
- Сообщение укладывается в таблицу 4х4

	1	2	3	4
1	А	З	Ю	Ж
2	Е		С	Ш
3	Г	Т	О	О
4	И	П	Е	Р

- При следовании за первым столбцом второго, получим:

F(1→2) =F(А3) + F(Е_) + F(ГТ) + F(ИП)=7+9+0+5=21.

- При следовании за первым столбцом третьего получим:

F(1→3) =F(АЮ) + F(ЕС) + F(ГО) + F(ИЕ) =6+8+8+8=30.

- При следовании за первым столбцом четвертого получим:

F(1→4) =F(АЖ) + F(ЕШ) + F(ГО) + F(ИР) =7+5+8+7=27.

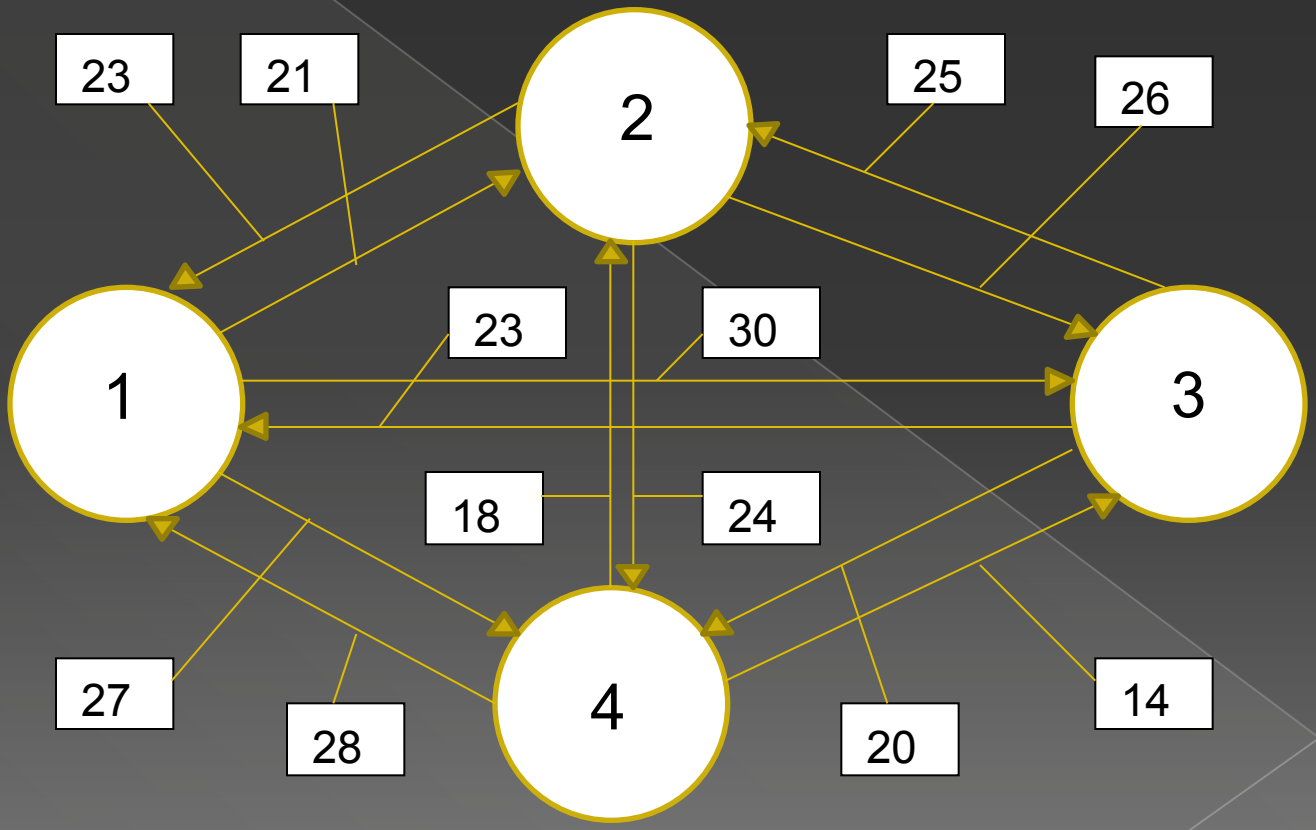
- Рассматривая все возможные варианты следования столбцов получим:

F(2→1) = F(3А) + F(_Е) + F(ТГ) + F(ПИ) =8+7+1+7=23;
F(2→3) = F(3Ю) + F(_С) + F(ТО) + F(ПЕ) =0+9+9+8=26;
F(2→4) = F(3Ж) + F(_Ш) + F(ТО) + F(ПР) =1+5+9+9=24;

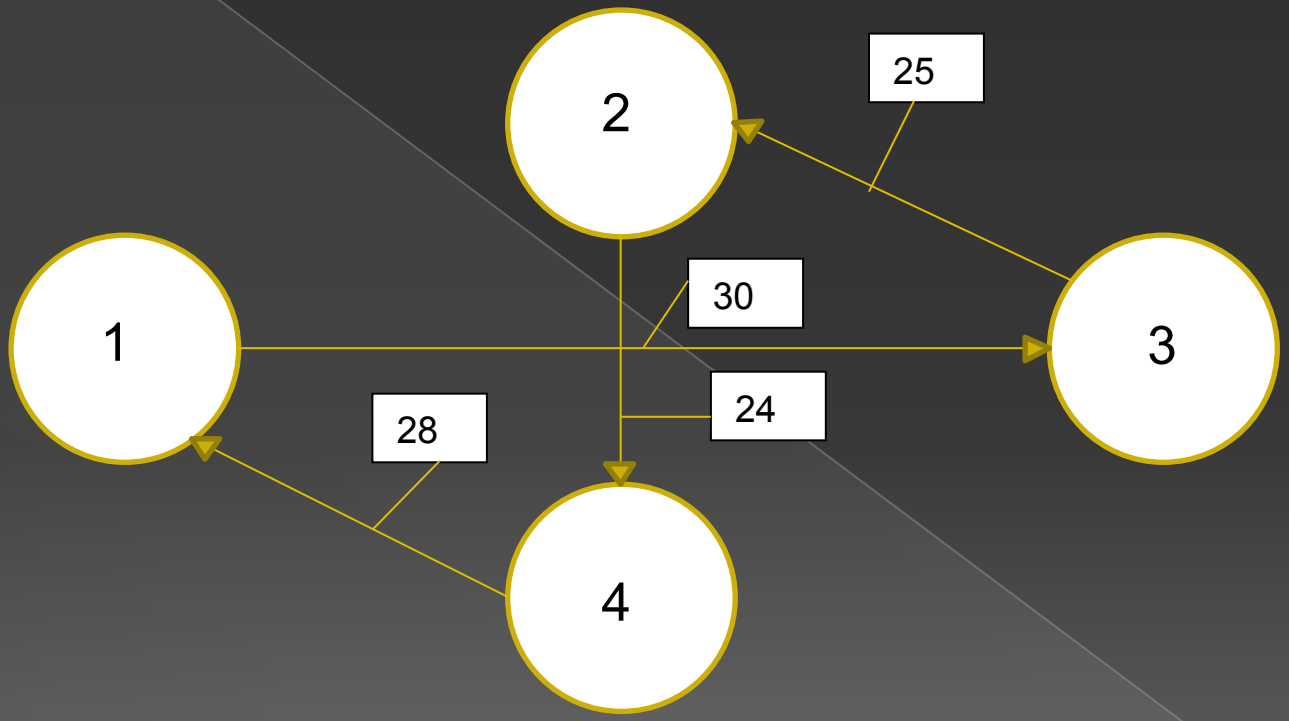
F(3→1) = F(ЮА) + F(СЕ) + F(ОГ) + F(ЕИ) =0+7+8+8=23;
F(3→2) = F(Ю3) + F(С_) + F(ОТ) + F(ЕП) =2+7+8+8=25;
F(3→4) = F(ЮЖ) + F(СШ) + F(ОО) + F(ЕР) =1+5+6+8=20;

F(4→1) = F(ЖА) + F(ШЕ) + F(ОГ) + F(РИ) =6+6+8+8=28;
F(4→2) = F(Ж3) + F(Ш_) + F(ОТ) + F(РП) =1+5+8+4=18;
F(4→3) = F(ЖЮ) + F(ШС) + F(ОО) + F(РЕ) =0+0+6+8=14.

- необходимо найти такой порядок следования столбцов, чтобы **получить максимальную сумму логарифмов** (наиболее вероятное следование столбцов в исходном сообщении)



Подграф максимального пути



Для полученного графа составляем возможные варианты таблиц

	1	3	2	4
1	А	Ю	З	Ж
2	Е	С		Ш
3	Г	О	Т	О
4	И	Е	П	Р

	4	1	3	2
1	Ж	А	Ю	З
2	Ш	Е	С	
3	О	Г	О	Т
4	Р	И	Е	П

	2	4	1	3
1	З	Ж	А	Ю
2		Ш	Е	С
3	Т	О	Г	О
4	П	Р	И	Е

	3	2	4	1
1	Ю	З	Ж	А
2	С		Ш	Е
3	О	Т	О	Г
4	Е	П	Р	И

	2	4	1	3
4	П	Р	И	Е
1	З	Ж	А	Ю
2		Ш	Е	С
3	Т	О	Г	О

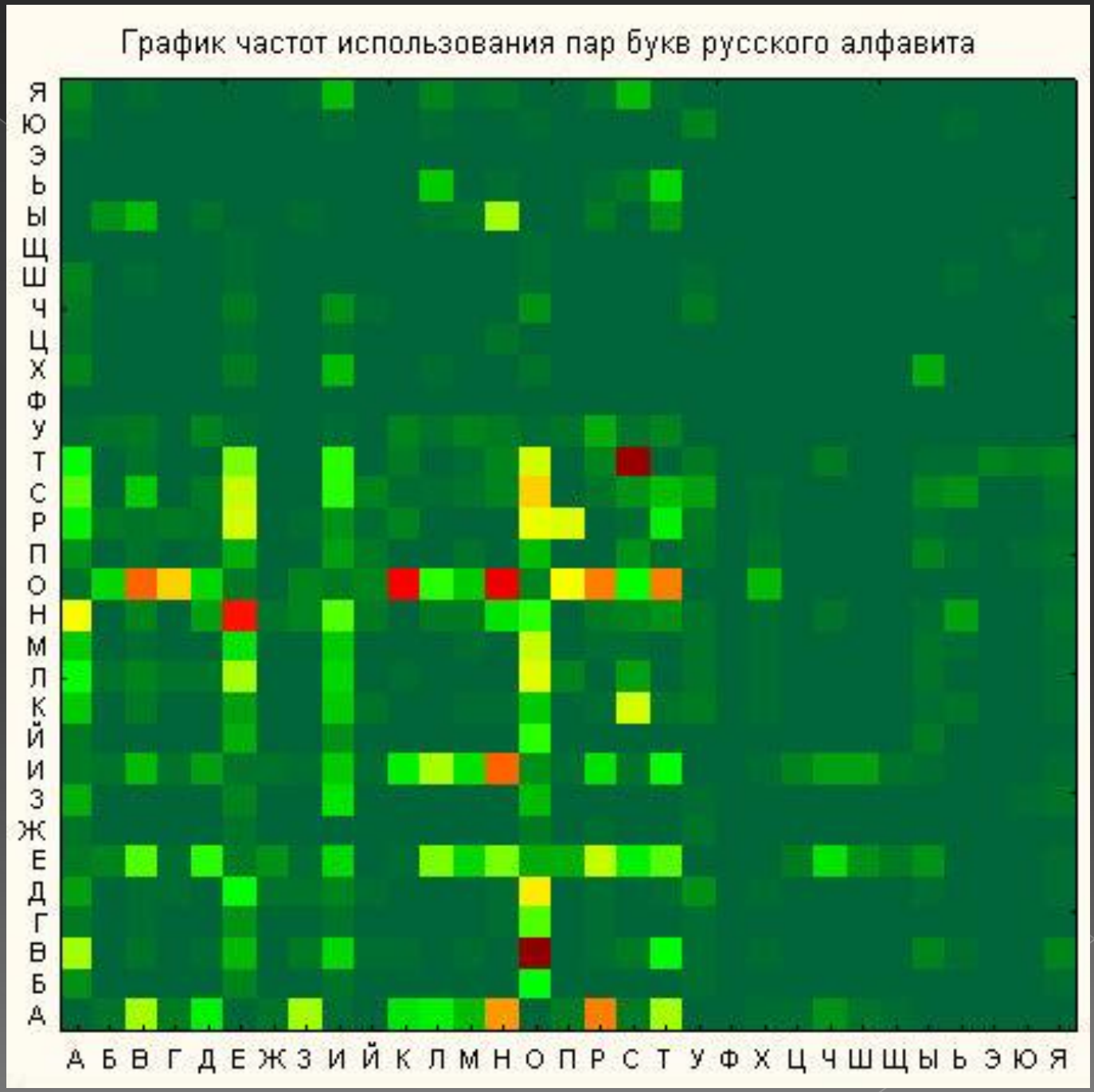


График частот использования пар букв русского алфавита
Я Ю Э Ь Ы Щ Ш Ч Ц Х Ф У Т С Р П О Н М Л К Й И З Ж Е Д Г В Б А
А Б В Г Д Е Ж З И Й К Л М Н О П Р С Т У Ф Х Ц Ч Ш Щ Ы Ь Э Ю Я