**Введение**

Для решения задач к данной лекции используйте программу [Classic](https://www.intuit.ru/EDI/18_04_17_2/1492467688-23204/tutorial/1287/objects/6/files/Software.zip).

Традиционной задачей криптографии является обеспечение конфиденциальности информации при передаче сообщений *по* контролируемому противником каналу связи. Владелец информации, называемый обычно *отправителем*, осуществляет преобразование исходной *(открытой)* информации (сам процесс преобразования называется *шифрованием*}в форму передаваемых *получателю* *по* открытому каналу связи *шифрованных* сообщений с целью ее защиты от *противника*.

Под *противником*понимается любой субъект, не имеющий *права* ознакомления с содержанием передаваемой информации. В качестве противника может выступать *криптоаналитик*, владеющий методами раскрытия шифров. Законный получатель информации осуществляет *расшифровывание* полученных сообщений. Противник пытается овладеть защищаемой информацией (его действия обычно называют *атаками*). При этом он может совершать как пассивные, так и активные действия.

*Пассивные* атаки связаны с прослушиванием, анализом трафика, перехватом, записью передаваемых шифрованных сообщений, *дешифрованием*, т. е. попытками "взломать" защиту с целью овладения информацией.

При проведении *активных* атак противник может прерывать процесс передачи сообщений, создавать поддельные (сфабрикованные) или модифицировать передаваемые шифрованные сообщения. Эти активные действия называют попытками *имитации* и *подмены* соответственно.

Изложим кратко основные понятия, следуя [1].

Под *шифром* обычно понимается семейство обратимых преобразований, каждое из которых определяется некоторым параметром, называемым ключом, а также порядком применения данного преобразования, называемым *режимом шифрования*.

*Ключ* - это важнейший *компонент* шифра, отвечающий за выбор преобразования, применяемого для зашифровывания конкретного сообщения. Обычно *ключ* представляет собой некоторую буквенную или числовую последовательность. Эта последовательность как бы "настраивает" *алгоритм* шифрования.

Каждое преобразование однозначно определяется ключом и описывается некоторым *криптографическим алгоритмом*. Один и тот же криптографический *алгоритм* может применяться для шифрования в различных режимах. Тем самым реализуются различные способы шифрования (простая замена, гаммирование и т. п.). Каждый режим шифрования имеет как свои преимущества, так и недостатки. Поэтому выбор режима зависит от конкретной ситуации. При расшифровывании используется криптографический *алгоритм*, который в общем случае может отличаться от алгоритма, применяемого для зашифровывания сообщения. Соответственно могут различаться ключи зашифровывания и расшифровывания. Пару алгоритмов зашифровывания и расшифровывания обычно называют *криптосистемой (шифрсистемой)*, а реализующие их устройства - *шифртехникой*.

Если обозначить через M открытое, а через C шифрованное сообщения, то процессы зашифрования и расшифрования можно записать в виде равенств

E_{k_1}(M) = C,\qquad D_{k_2}(C)=M,

в которых алгоритмы зашифрования E и расшифрования D должны удовлетворять равенству

D_{k_2}(E_{k_1}(M))=M.

Дадим формальное *определение* шифра.

Пусть X, K, Y - *конечные* *множества* открытых текстов, ключей и шифрованных текстов, соответственно; {E}_{k}:X\rightarrow Y - правило зашифрования на ключе k{\in}K. Множество \left\{{E}_{k}\right|k{\in}K\} обозначим через E, а множество \left\{{E}_{k}\left(x\right)\right|x{\in}X\} - через {E}_{k}\left(X\right). Пусть {D}_{k}:{E}_{k}\left(X\right)\rightarrow X - правило расшифрования на ключе k{\in}K, и D - множество \left\{{D}_{k}\left|k{\in}K\right.\right\}.

Здесь и далее будем полагать, что, если *ключ* k{\in}K представляется в виде k=\left({k}_{\text{\textcyrillic{з}}},{k}_{\text{\textcyrillic{р}}}\right), где {k}_{\text{\textcyrillic{з}}} - *ключ* зашифрования, {k}_{\text{\textcyrillic{р}}} - *ключ* расшифрования, причем {k}_{\text{\textcyrillic{з}}}\neq{k}_{\text{\textcyrillic{р}}}, то {E}_{k} понимается, как *функция* {E}_{{k}_{\text{\textcyrillic{з}}}}, а {D}_{k} - как *функция* {D}_{{k}_{\text{\textcyrillic{р}}}}.

**Определение 6.1** *Шифром (шифросистемой) назовем совокупность {\Sigma }_{A}=(X,K,Y,E,D) введенных множеств, для которых выполняются следующие свойства:*

1. {D}_{k}\left({E}_{k}\left(x\right)\right)=x  {\forall}x{\in}X  {\forall}k{\in}K,
2. Y= \bigcup\limits_{k\in K} {E}_{k}(X).

Неформально, *шифр* - это совокупность множеств открытых текстов(то, что шифруется), возможных ключей(то, с помощью чего шифруется), возможных шифртекстов(то, во что шифруется), правил зашифрования и правил расшифрования.

Ниже следуют теоретические сведения и примеры решения задач, после чего представлены задачи для самостоятельного решения.

Рассмотрим вводные, подготовительные задачи.

**Пример 6.1** [3] *Перед вами два зашифрованных текста. Один шифрованный текст соответствует английскому открытому тексту, другой - русскому. Открытые тексты состоят только из букв. Шифрование заключалось в замене каждой буквы открытого текста двузначным числом. Разным буквам соответствуют разные числа, различия между строчной и прописной буквами не делалось, знаки препинания опущены. Определите, какой шифрованный текст соответствует русскому открытому тексту.*

*Текст 1*

11 25 28 33 35 42 47 53 69 72 83 11 19 22 17 31 47 39 45 17 11 57

83 91 19 11 59 62 53 45 11 51 11 65 25 25 28 11 83 91 51

45 11 19 25 33 62 35 69 72 83 19 25 65 51 83 91 11

*Текст 2*

97 11 63 19 44 58 17 93 28 25 31 33 22 35 39 83 11 17 19 35 11 39

59 47 51 17 91 69 33 28 42 45 17 53 57 62 33 65 22 72 97 11 63 69

11 47 17 19 93 17 25 31 53 22 33 39 11 19 11 63 35

**Решение.** Подсчитаем количество шифрообозначений в текстах. В первом тексте их 25, во втором - 28. Английский *алфавит* содержит 26 букв, следовательно, второй *шифртекст* не может соответствовать английскому открытому тексту. Значит, второй текст - это зашифрованный русский текст.

*Ответ:* первый текст - это зашифрованный английский текст, а второй - русский.

**6.1 Математические модели открытого текста**

([1], §2.5])

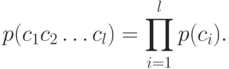
Потребность в математических моделях открытого текста продиктована, прежде всего, следующими соображениями. Во-первых, даже при отсутствии ограничений на временные и *материальные затраты* *по* выявлению закономерностей, имеющих *место* в открытых текстах, нельзя гарантировать того, что такие свойства указаны с достаточной полнотой. Например, хорошо известно, что частотные свойства текстов в значительной степени зависят от их характера. Поэтому при математических исследованиях свойств шифров прибегают к упрощающему моделированию, в частности, реальный *открытый текст* заменяется его моделью, отражающей наиболее важные его свойства. Во-вторых, при автоматизации методов криптоанализа, связанных с перебором ключей, требуется "научить" ЭВМ отличать *открытый текст* от случайной последовательности знаков. Ясно, что соответствующий критерий может выявить лишь *адекватность* последовательности знаков некоторой модели открытого текста.

Один из естественных подходов к моделированию открытых текстов связан с учетом их частотных характеристик, приближения для которых можно вычислить с нужной точностью, исследуя тексты достаточной длины. Основанием для такого подхода является *устойчивость* частот k-грамм или целых словоформ реальных языков человеческого общения (то есть отдельных букв, слогов, слов и некоторых словосочетаний). Основанием для построения модели может служить также и теоретико-информационный подход, развитый в работах К. Шеннона.

Учет частот k-грамм приводит к следующей модели открытого текста. Пусть P^{(k)}(A) представляет собой *массив*, состоящий из приближений для вероятностей p(b_1,b_2,\ldots,b_k) появления k-грамм b_1b_2\ldots b_k в открытом тексте, k\in \mathbb{N}, A=\{a_1,a_2,\ldots,a_n\} - *алфавит* открытого текста, b_i\in A$, $i=1,2,\ldots,k.

Тогда в последовательности c_1,c_2,\ldots,c_k,c_{k+1},\ldots знаков алфавита A, генерируемой источником "открытого текста", k-грамма c_1c_2\ldots c_k появляется с вероятностью p(c_1,c_2,\ldots,c_k)\in P^{(k)}(A), k-грамма c_1,c_2\ldots,c_k,c_{k+1} появляется с *вероятность* p(c_2,c_3,\ldots,c_{k+1})\in P^{(k)}(A). и т. д. Назовем построенную модель открытого текста *вероятностной моделью* k-го приближения.

Таким образом, простейшая модель открытого текста - *вероятностная модель первого приближения* - представляет собой последовательность знаков c_1,c_2,\ldots, в которой каждый знак c_i, i=1,2,\ldots появляется с вероятностью p(c_i)\in P^{(1)}(A), независимо от других знаков. Будем называть также эту модель *позначной моделью открытого текста*. В такой модели *открытый текст* c_1c_2\ldots c_l имеет *вероятность*

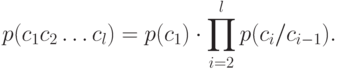


В вероятностной модели второго приближения первый знак c_1 имеет *вероятность* p(c_1)\in P^{(1)}(A), а каждый следующий знак c_i зависит от предыдущего и появляется с вероятностью

p(c_i/c_{i-1})=\frac{p(c_{i-1} c_i)}{p(c_{i-1})},

где p(c_{i-1}c_i)\in P^{(2)}(A), p(c_{i-1})\in P^{(1)}(A), i=2, 3,\ldots.

Другими словами, модель открытого текста второго приближения представляет собой *простую однородную цепь Маркова*. В такой модели *открытый текст* c_1c_2\ldots c_l имеет *вероятность*



Модели открытого текста более высоких приближений учитывают зависимость каждого знака от большего числа предыдущих знаков. Ясно, что чем выше степень приближения, тем более "читаемыми" являются соответствующие модели.

Отметим, что с более общих позиций *открытый текст* может рассматриваться как реализация *стационарного эргодического случайного процесса с дискретным временем и конечным числом состояний*.

**6.1.1 Критерии распознавания открытого текста**

Заменив реальный открытый текст его моделью, мы можем теперь построить критерий распознавания открытого текста. При этом можно воспользоваться либо стандартными методами различения статистических гипотез, либо наличием в открытых текстах некоторых запретов, таких, например, как биграмма ЪЪ в русском тексте. Проиллюстрируем первый подход при распознавании позначной модели открытого текста.

Итак, согласно нашей договоренности, открытый текст представляет собой реализацию независимых испытаний случайной величины, значениями которой являются буквы алфавита A=\{a_1,\ldots, a_n\}, появляющиеся в соответствии с распределением вероятностей P^{(1)}(A)=(p(a_1),\ldots,p(a_n)). Требуется определить, является ли случайная последовательность c_1c_2\ldots c_l букв алфавита A открытым текстом, или нет.

Пусть H_0 - гипотеза, состоящая в том, что данная последовательность - открытый текст, H_1 - альтернативная гипотеза. В простейшем случае последовательность c_1c_2\ldots c_l можно рассматривать при гипотезе H_1 как случайную и равновероятную. Эта альтернатива отвечает субъективному представлению о том, что при расшифровании криптограммы с помощью ложного ключа получается "бессмысленная" последовательность знаков. В более общем случае можно считать, что при гипотезе H_1 последовательность c_1c_2\ldots c_l представляет собой реализацию независимых испытаний некоторой случайной величины, значениями которой являются буквы алфавита A, появляющиеся в соответствии с распределением вероятностей Q^{(1)}(A)=(q(a_1),\ldots,q(a_n)). При таких договоренностях можно применить, например, *наиболее мощный критерий* различения двух простых гипотез, который дает *лемма Неймана-Пирсона*.

В силу своего вероятностного характера такой критерий может совершать ошибки двух родов. Критерий может принять открытый текст за случайный набор знаков. Такая ошибка обычно называется *ошибкой первого рода*, ее вероятность равна \alpha=p(H_1/H_0). Аналогично вводится *ошибка второго рода* и ее вероятность \beta=p(H_0/H_1). Эти ошибки определяют качество работы критерия. В криптографических исследованиях естественно минимизировать вероятность ошибки первого рода, чтобы не "пропустить" открытый текст. Лемма Неймана-Пирсона при заданной вероятности первого рода минимизирует также вероятность ошибки второго рода.

Критерии на открытый текст, использующие запретные сочетания знаков, например k-граммы подряд идущих букв, будем называть *критериями запретных k-грамм*. Они устроены чрезвычайно просто. Отбирается некоторое число s редких k-грамм, которые объявляются запретными. Теперь, просматривая последовательно k-грамму за k-граммой анализируемой последовательности c_1c_2\ldots c_l, мы объявляем ее случайной, как только в ней встретится одна из запретных k-грамм, и открытым текстом в противном случае. Такие критерии также могут совершать ошибки в принятии решения. В простейших случаях их можно рассчитать. Несмотря на свою простоту, критерии запретных k-грамм являются весьма эффективными.

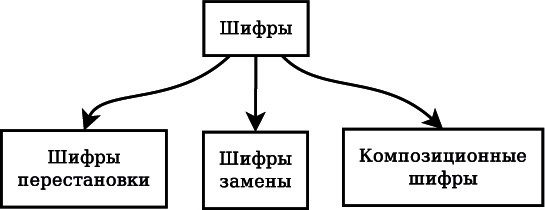
**6.2 Классификация шифров**

(см. также [1, Глава 3])

Большое влияние на развитие криптографии оказали появившиеся в середине XX века работы американского *математика* Клода Шеннона. В этих работах были заложены основы теории информации, а также был разработан математический аппарат для исследований во многих областях науки, связанных с информацией. Более того, принято считать, что *теория информации* как наука родилась в 1948 году после публикации работы К. Шеннона "Математическая теория связи".

В своей работе "Теория связи в секретных системах" Клод Шеннон обобщил накопленный до него *опыт* разработки шифров. Оказалось, что даже в очень сложных шифрах в качестве типичных компонентов можно выделить такие простые шифры как *шифры замены, шифры перестановки* или их *сочетания*.

В качестве первичного признака, *по* которому проводится классификация шифров, используется тип преобразования, осуществляемого с открытым текстом при шифровании. Если фрагменты открытого текста (отдельные буквы или группы букв) заменяются некоторыми их эквивалентами в шифртексте, то соответствующий *шифр* относится к классу *шифров замены*. Если буквы открытого текста при шифровании лишь меняются местами друг с другом, то мы имеем дело с *шифром перестановки*. С целью повышения надежности шифрования шифрованный текст, полученный применением некоторого шифра, может быть еще раз зашифрован с помощью другого шифра.



**Рис. 6.1.**Классификация шифров

Всевозможные такие композиции различных шифров приводят к третьему классу шифров, которые обычно называют *композиционными шифрами*. Заметим, что композиционный *шифр* может не входить ни в *класс* шифров замены, ни в *класс* шифров перестановки ([рис. 6.1](https://www.intuit.ru/studies/courses/13837/1234/lecture/31196?page=3#image.6.1) ).

**6.3 Шифры перестановки**

*Шифр* перестановки, как видно из названия, осуществляет преобразование перестановки букв в открытом тексте. Типичным примером шифра перестановки является *шифр* "Сцитала". Обычно *открытый текст* разбивается на отрезки равной длины и каждый *отрезок* шифруется независимо. Пусть, например, *длина* отрезков равна n и \sigma  - взаимнооднозначное *отображение* *множества* \{1,2,\ldots,n\} в себя. Тогда *шифр* перестановки действует так: *отрезок* открытого текста x_1x_2\ldots x_n преобразуется в *отрезок* шифрованного текста.

Классическим примером такого шифра является система, использующая карточку с отверстиями - *решетку*, которая при наложении на *лист* бумаги оставляет открытыми лишь некоторые его части. При зашифровке буквы сообщения вписываются в эти отверстия. При расшифровке сообщение вписывается в диаграмму нужных размеров, затем накладывается решетка, после чего на виду оказываются только буквы открытого текста.

Также возможны и другие варианты шифра перестановки, например, шифры столбцовой и двойной перестановки.

**6.3.1 Шифр столбцовой перестановки**

При расшифровывании буквы шифртекста записываются по столбцам в соответствии с последовательностью чисел ключа, после чего исходный текст считывается по строкам. Для удобства запоминания ключа применяют перестановку столбцов таблицы по ключевому слову или фразе, всем символам которых ставятся в соответствие номера, определяемые порядком соответствующих букв в алфавите.

При решении заданий на криптоанализ шифров перестановки необходимо восстановить начальный порядок следования букв текста. Для этого используется анализ совместимости символов, в чем может помочь таблица сочетаемости (см. [1]).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Таблица 6.1. Сочетаемость букв русского языка [1] | | | | | | |
| **Г** | **С** | **Слева** |  | **Справа** | **Г** | **С** |
| 3 | 97 | л, д, к, т, в, р, н | А | л, н, с, т, р, в, к, м | 12 | 88 |
| 80 | 20 | я, е, у, и, а, о | Б | о, ы, е, а, р, у | 81 | 19 |
| 68 | 32 | я, т, а, е, и, о | В | о, а, и, ы, с, н, л, р | 60 | 40 |
| 78 | 22 | р, у, а, и, е, о | Г | о, а, р, л, и, в | 69 | 31 |
| 72 | 28 | р, я, у, а, и, е, о | Д | е, а, и, о, н, у, р, в | 68 | 32 |
| 19 | 81 | м, и, л, д, т, р, н | Е | н, т, р, с, л, в, м, и | 12 | 88 |
| 83 | 17 | р, е, и, а, у, о | Ж | е, и, д, а, н | 71 | 29 |
| 89 | 11 | о, е, а, и | 3 | а, н, в, о, м, д | 51 | 49 |
| 27 | 73 | р, т, м, и, о, л, н | И | с, н, в, и, е, м, к, з | 25 | 75 |
| 55 | 45 | ь, в, е, о, а, и, с | К | о, а, и, р, у, т, л, е | 73 | 27 |
| 77 | 23 | г, в, ы, и, е, о, а | Л | и, е, о, а, ь, я, ю, у | 75 | 25 |
| 80 | 20 | я, ы, а, и, е, о | М | и, е, о, у, а, н, п, ы | 73 | 27 |
| 55 | 45 | д, ь, н, о, а, и, е | Н | о, а, и, е, ы, н, у | 80 | 20 |
| 11 | 89 | р, п, к, в, т, н | О | в, с, т, р, и, д, н, м | 15 | 85 |
| 65 | 35 | в, с, у, а, и, е, о | П | о, р, е, а, у, и, л | 68 | 32 |
| 55 | 45 | и, к, т, а, п, о, е | Р | а, е, о, и, у, я ,ы, н | 80 | 20 |
| 69 | 31 | с, т, в, а, е, и, о | С | т, к, о, я, е, ь, с, н | 32 | 68 |
| 57 | 43 | ч, у, и, а, е, о, с | Т | о, а, е, и, ь, в, р, с | 63 | 37 |
| 15 | 85 | п, т, к, д, н, м, р | У | т, п, с, д, н, ю, ж | 16 | 84 |
| 70 | 30 | н, а, е, о, и | Ф | и, е, о, а, е, о, а | 81 | 19 |
| 90 | 10 | у, е, о, а, ы, и | X | о, и, с, н, в, п, р | 43 | 57 |
| 69 | 31 | е, ю, н, а, и | Ц | и, е, а, ы | 93 | 7 |
| 82 | 18 | е, а, у, и, о | Ч | е, и, т, н | 66 | 34 |
| 67 | 33 | ь, у, ы, е, о, а, и, в | Ш | е, и, н, а, о, л | 68 | 32 |
| 84 | 16 | е, б, а, я, ю | Щ | е, и, а | 97 | 3 |
| 0 | 100 | м, р, т, с, б, в, н | Ы | Л, х, е, м, и, в, с, н | 56 | 44 |
| 0 | 100 | н, с, т, л | Ь | н, к, в, п, с, е, о, и | 24 | 76 |
| 14 | 86 | с, ы, м, л, д, т, р, н | Э | н, т, р, с, к | 0 | 100 |
| 58 | 42 | ь, о, а, и, л, у | Ю | д, т, щ, ц, н, п | 11 | 89 |
| 43 | 57 | о, н, р, л, а, и, с | Я | в, с, т, п, д, к, м, л | 16 | 84 |

При анализе сочетаемости букв друг с другом следует иметь в виду зависимость появления букв в открытом тексте от значительного числа предшествующих букв. Для анализа этих закономерностей используют понятие условной вероятности.

Систематически вопрос о зависимости букв алфавита в открытом тексте от предыдущих букв исследовался известным русским математиком А.А. Марковым (1856-1922). Он доказал, что появления букв в открытом тексте нельзя считать независимыми друг от друга. В связи с этим А.А. Марковым отмечена еще одна устойчивая закономерность открытых текстов, связанная с чередованием гласных и согласных букв. Им были подсчитаны частоты встречаемости биграмм вида гласная-гласная (*г, г*), гласная-согласная (*г, с*), согласная-гласная (*с, г*), согласная-согласная (*с, с*) в русском тексте длиной в 10^5 знаков. Результаты подсчета отражены в следующей таблице:

|  |  |  |  |
| --- | --- | --- | --- |
| Таблица 6.2. Чередование гласных и согласных | | | |
|  | Г | С | Всего |
| Г | 6588 | 38310 | 44898 |
| С | 38296 | 16806 | 55102 |

**Пример 6.2** [3] *Открытый текст, сохраняя пробелы между словами, записали в таблицу. Начало было в первой строке, текст записывали слева направо, переходя со строки на следующую, шифрование заключалось в перестановке столбцов. Найдите открытый текст.*

*Шифрованный текст:*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Д | В |  |  | Ы | Т |
|  | Г | О | Е | Р | О |
| У | Ь | Д |  | У | Б |
| М | М | Я | Ы | Р | П |

**Решение.** Присвоим столбцам номера в порядке их следования. Наша задача - найти такой порядок столбцов, при котором текст будет осмысленным.

Составим таблицу:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 | 6 |
| 1 | Х |  |  |  |  |  |
| 2 |  | Х |  |  |  |  |
| 3 |  |  | Х |  |  |  |
| 4 |  |  |  | Х |  |  |
| 5 |  |  |  |  | Х |  |
| 6 |  |  |  |  |  | Х |

Клетка (k, m) в этой таблице означает, что столбец с номером m следует за столбцом с номером k. Знаком "Х" отметим невозможные случаи.

Сочетания столбцов 1, 2 и 5, 2 невозможны, так как гласная не может находиться перед мягким знаком. Невозможны и следования столбцов 2, 1 и 2, 5. Теперь из третьей строки следует, что 1, 5 и 5, 1 невозможны, так как УУ - нехарактерная для русского языка биграмма. Далее, два пробела подряд не могут быть в тексте, значит, ставим "Х" в клетках 3, 4 и 4, 3. Снова обратимся к третьей строке. Если бы столбец 2 следовал за столбцом 4, то слово начиналось бы с мягкого знака. Ставим "Х" в клетке 4, 2. Из первой строки: невозможна комбинация 4, 5, невозможна и 3, 5. Итог наших рассуждений представлен в таблице:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 | 6 |
| 1 | Х | Х |  |  | Х |  |
| 2 | Х | Х |  |  | Х |  |
| 3 |  |  | Х | Х | Х |  |
| 4 |  | Х | Х | Х | Х |  |
| 5 | Х | Х |  |  | Х |  |
| 6 |  |  |  |  |  | Х |

Итак, после столбца 6 обязательно следует столбец 5. Но тогда поставим "Х" в клетке 6, 2 и получим: столбец 2 следует за столбцом 3. Далее, мы вычеркнули 5, 1 и 2, 1, следовательно, надо проверить варианты: ...6532... и ...65432... . Но (4, 3) вычеркнуто ранее. Итак, остались варианты расположения столбцов:

* 1, 6, 5, 3, 2, 4
* 6, 5, 3, 2, 4, 1
* 4, 1, 6, 5, 3, 2
* 1, 4, 6, 5, 3, 2

Запишем 6, 5, 3, 2 столбцы подряд:

|  |  |  |  |
| --- | --- | --- | --- |
| 6 | 5 | 3 | 2 |
| т | ы | - | в |
| о | р | о | г |
| б | у | д | ь |
| п | р | я | м |

Попытка поставить столбец 1 перед столбцом 6 приведет к биграмме МП в последней строке и сочетанию ДТЫ в первой. Остались варианты: 653241, 146532.

*Ответ:* 653241 - ключ, открытый текст: ты\\_в\\_дороге\\_будь\\_упрямым (строка из популярной в 1970-е годы песни).

Приведем еще один пример криптоанализа шифра столбцовой перестановки.

**Пример 6.3** [1] *Расшифровать: СВПООЗЛУЙЬСТЬ\\_ЕДПСОКОКАЙЗО*

**Решение.** Текст содержит 25 символов, что позволяет записать его в квадратную матрицу 5х5. Известно, что шифрование производилось по столбцам, следовательно, расшифровывание следует проводить, меняя порядок столбцов.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| С | В | П | О | О |
| З | Л | У | Й | Ь |
| С | Т | Ь | \_ | Е |
| Д | П | С | О | К |
| К | А | Й | З | О |

Необходимо произвести анализ совместимости символов (Таблица сочетаемости букв русского и английского алфавита, а также таблицы частот биграмм представлена выше). В первом и третьем столбце сочетание СП является крайне маловероятным для русского языка. Рассмотрим другие запрещенные и маловероятные сочетания букв: ВП (2,3 столбцы), ПС (3,1 столбцы), ПВ (3,2 столбцы). Перебрав их все, получаем наиболее вероятные сочетания биграмм по столбцам:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| В | О | С | П | О |
| Л | Ь | З | У | Й |
| Т | Е | С | Ь | \_ |
| П | О | Д | С | К |
| А | З | К | О | Й |

Получаем осмысленный текст: ВОСПОЛЬЗУЙТЕСЬ\_ПОДСКАЗКОЙ

**6.3.2 Шифр двойной перестановки**

**Пример 6.4** [1] *Дан шифртекст: ЫОЕЧТТОУ\_СНСОРЧТРНАИДЬН\_Е. Известно, что шифрование производилось сначала по столбцам, а затем по строкам.*

**Решение.** Запишем текст в квадратную матрицу 5х5. Расшифровывание следует проводить тем же способом, что и зашифровывание.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Ы | О | Е | Ч | Т |
| Т | О | У | \_ | С |
| Н | С | О | Р | Ч |
| Т | Р | Н | А | И |
| Д | Ь | Н | \_ | Е |

Производим анализ совместимости символов. Если в примере столбцовой перестановки можно было легко подобрать нужную комбинацию путем перебора, то здесь лучше воспользоваться таблицей частот букв русского языка (см. приложение). Для оптимизации скорости выполнения задания можно проверить все комбинации букв только в первой строке. Получаем ОЕ-15, ОЧ-12, ЕТ-33, ТЕ-31, ЧО-х, ЕО-7, ЧЫ-х, ОЫ-х, ТЫ-11, ТЧ-1, ЧЕ-23 (где х - запрещенная комбинация).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| О | Ч | Е | Т | Ы |
| О | \_ | У | С | Т |
| С | Р | О | Ч | Н |
| Р | А | Н | И | Т |
| Ь | \_ | Н | Е | Д |

Теперь необходимо переставить строки в нужном порядке: 3, 2, 4, 5, 1:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| С | Р | О | Ч | Н |
| О | \_ | У | С | Т |
| Р | А | Н | И | Т |
| Ь | \_ | Н | Е | Д |
| О | Ч | Е | Т | Ы |

Получаем осмысленный текст: СРОЧНО\_УСТРАНИТЬ\_НЕДОЧЕТЫ

**6.4 Шифры замены**

*Шифр* замены является простейшим, наиболее популярным шифром. Типичными примерами являются *шифр* Цезаря, "цифирная азбука" Петра Великого и "пляшущие человечки" А. Конан Дойла. Как видно из самого названия, *шифр* замены осуществляет преобразование замены букв или других "частей" открытого текста на аналогичные "части" шифрованного текста. Дадим математическое описание шифра замены. Пусть X и Y - два алфавита (открытого и шифрованного текстов соответственно), состоящие из одинакового числа символов. Пусть также g: X \rightarrow Y - взаимнооднозначное *отображение* X в Y. Тогда *шифр* замены действует так: *открытый текст* x_1x_2\ldots x_n преобразуется в шифрованный текст g(x_1)g(x_2)\ldots g(x_n).

**6.4.1 Математическая модель шифра замены**

([1])

Определим модель \Sigma_A=(X,K,Y,E,D) произвольного шифра замены. Будем считать, что открытые и шифрованные тексты являются словами в алфавитах A и B соответственно: X\subset A^{\ast}, Y\subset B^{\ast}, |A|=n, |B|=m. Здесь и далее C^{\ast} обозначает множество слов конечной длины в алфавите C.

Перед зашифрованием открытый текст предварительно представляется в виде последовательности подслов, называемых *шифрвеличинами*. При зашифровании шифрвеличины заменяются некоторыми их эквивалентами в шифртексте, которые назовем *шифробозначениями*. Как шифрвеличины, так и шифробозначения представляют собой слова из A^\ast и B^\ast соответственно.

Пусть U=\{u_1,u_2,\ldots,u_N\} - множество возможных шифрвеличин, V=\{v_1,v_2,\ldots,v_M\} - множество возможных шифробозначений. Эти множества должны быть такими, чтобы любые тексты x\in X, y\in Y можно было представить словами из U^\ast, V^\ast соответственно. Требование однозначности расшифрования влечет неравенства N\geq n, M\geq m, M\geq N.

**6.4.2 Криптоанализ шифра простой замены**

Криптоанализ шифра *простой замены* основан на использовании статистических закономерностей языка. Приведем таблицы частот букв русского и английского языков, следуя [1].

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 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 | 3 | 0,016 | Ь,Ъ | 0,014 |
| Б | 0,014 | Г | 0,013 | Ч | 0,012 | Й | 0,010 |
| X | 0,009 | Ж | 0,007 | Ю | 0,006 | Ш | 0,006 |
| Ц | 0,004 | Щ | 0,003 | Э | 0,003 | Ф | 0,002 |

Полезной является диаграмма [рис. 6.2](https://www.intuit.ru/studies/courses/13837/1234/lecture/31196?page=4#image.6.2)



**Рис. 6.2.**Диаграмма частот букв русского языка

Для получения более точных сведений об открытых текстах можно строить и анализировать таблицы k-грамм при k>2, однако для учебных целей вполне достаточно ограничиться биграммами. Неравновероятность k-грамм (и даже слов) тесно связана с характерной особенностью открытого текста - наличием в нем большого числа повторений отдельных фрагментов текста: корней, окончаний, суффиксов, слов и фраз. Так, для русского языка такими привычными фрагментами являются наиболее частые биграммы и триграммы:

СТ, НО, ЕН, ТО, НА, ОВ, НИ, РА, ВО, КО,

СТО, ЕНО, НОВ, ТОВ, ОВО, ОВА

Полезной является информация о сочетаемости букв, т. е. о предпочтительных связях букв друг с другом, которую легко извлечь из таблиц частот биграмм.

Имеется в виду таблица, в которой слева и справа от каждой буквы расположены наиболее предпочтительные "соседи" (в порядке убывания частоты соответствующих биграмм). В таких таблицах обычно указывается также доля гласных и согласных букв (в процентах), предшествующих (или следующих за) данной букве.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Таблица 6.4. Частоты биграмм русского языка, часть 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  | **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 | | 3 | | 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 | |  | |  |
| **3** | 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 |
| **И** | 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 | |  |
| **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 | |  |
| **X** | 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 |
| Таблица 6.5. Частоты биграмм русского языка, часть 2 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|  | | **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 | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |
| **3** | | 3 | | 1 | |  | | 2 | |  | |  | |  | |  | |  | |  | | 4 | |  | |  | |  | | 4 | |
| **И** | | 11 | | 29 | | 29 | | 3 | | 1 | | 17 | | 3 | | 11 | | 1 | | 1 | |  | |  | | 1 | | 3 | | 17 | |
| **И** | | 3 | | 10 | | 2 | |  | |  | |  | | 1 | | 3 | | 2 | |  | |  | |  | |  | |  | |  | |
| **К** | | 10 | | 3 | | 7 | | 10 | |  | |  | | 1 | |  | |  | |  | |  | |  | |  | |  | |  | |
| **Л** | |  | | 3 | | 1 | | 6 | |  | | 4 | |  | | 1 | |  | |  | | 2 | | 30 | |  | | 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 | |
| **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 | |  | |  | | И | | 21 | |  | |  | | 4 | |
| **У** | | 7 | | 14 | | 7 | |  | |  | | 1 | |  | | 8 | | 3 | | 2 | |  | |  | |  | | 9 | | 1 | |
| **Ф** | | 1 | | 1 | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |  | |
| **X** | | 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 | |

**Пример 6.5** [3] *Известно, что зашифровано стихотворение Р. Киплинга в переводе С.Я. Маршака. Шифрование заключалось в замене каждой буквы на двузначное число. Отдельные слова разделены несколькими пробелами, знаки препинания сохранены. Таблица частот букв русского языка приведена выше.*

29 15 10 17 29 22 25 31 15 33 35 41 43 45 35 57 45 25 17 59 15 10 25 41

25 69, 59 78 29 82 25 78 25 17 15 10 88 90 78 25 62 25 22 10 57 73 79 35 67

78 90 88 29 45 35 29, 54 57 90 31 90 73 22 88 15 88 29 15 17 69 41 25 15, 70

17 90 57 43 59 15 78 15 62 22 25 17 57 25 69 88 15 82 17 25 88 29 45 35. . .

**Решение.** Подсчитаем частоты шифробозначений:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Обозначение** | 29 | 15 | 10 | 17 | 22 | 25 | 31 | 33 | 35 | 41 | 43 | 45 | 57 |
| **Количество** | 7 | 10 | 4 | 7 | 4 | 12 | 2 | 1 | 5 | 3 | 2 | 4 | 5 |
| **Обозначение** | 59 | 69 | 78 | 82 | 88 | 90 | 62 | 73 | 79 | 67 | 54 | 70 |  |
| **Количество** | 3 | 3 | 4 | 2 | 6 | 5 | 1 | 2 | 1 | 1 | 1 | 1 |  |

Из таблицы частот букв русского языка видно, что чаще всего встречается буква О, на втором месте Е. В нашем шифртексте чаще всего встречается обозначение 25 (12 раз), на втором месте идет обозначение 15 (10 раз), остальные обозначения им существенно уступают. Поэтому можем выдвинуть гипотезу: 25=О, 15=Е. Однако текст у нас не очень большой, поэтому закономерности русского языка проявляются в нем не обязательно в строгом соответствии с таблицей частот букв русского языка. Поэтому возможен и вариант: 25=Е, 15=О. Но тогда последнее слово в третьей строке имеет окончание ЕО, что возможно, но все же более вероятный вариант ОЕ. Итак, будем работать с текстом, считая, что 25=О, 15=Е.

Теперь нам поможет знак препинания: "29, {\dots}". Крайне маловероятно, чтобы запятая стояла после согласной. Итак, 29 - гласная, причем вероятнее всего 29=И или 29=А, т. к. гласные Я, Ю, Э, У встречаются в осмысленных текстах на русском языке намного реже, чем И и А, что не противоречит таблице частот шифртекста.

В последней строке: 88 15, но 15=Е, следовательно, 88 - согласная, причем наиболее вероятные значения - это Н и Т. Итак, 25=О, 15=Е, 29=\binom{\text{А}}{\text{И}}, 88=\binom{\text{Н}}{\text{Т}}. Теперь третье слово в третьей строке имеет 4 варианта:

* 29=И, 88=Н: 22 Н Е Н И Е
* 29=И, 88=Т:22 Т Е Т И Е
* 29=А, 88=Н:22 Н Е Н А Е
* 29=А, 88=Т: 22 Т Е Т А Е

Из рассмотренных вариантов лишь один является осмысленным, и он позволяет найти значение 22. Имеем: 22=М и третье слово в третье строке - М Н Е Н И Е.

Теперь рассмотрим второе слово в первой строке. Е 10 17 И, причем 10 и 17 - согласные, и это не М и не Н. Наиболее вероятное слово Е С Л И, т. е. 10=С, 17=Л. Конечно, если мы, продолжая работать с текстом, вдруг получим "нечитаемое" слово, то придется вернуться к этому этапу и рассмотреть другие варианты. Однако это маловероятно, поскольку вряд ли в стихотворении были слова наподобие Е Р Т И, Е В Л И и т. п.

Далее, первое слово второй строки: 59 78 И, причем 59 и 78 - согласные, и это не С, не Л, не М и не Н. Так что это слово П Р И, т. е. 59=П, 78=Р. Тогда шестое слово первой строки 45 О Л П Е, что дает значение 45=Т, и тогда при 57=В получаем фрагмент "... В Т О Л П Е ...". Также второе слово последней строки П Е Р Е 62 дает нам значение 62=Д.

Далее рассмотрим начало второй строки: "П Р И 82 О Р О Л Е С Н 90 Р О Д О М ...". Из него следует, что 82=К и 90=А. Зная, что 82=К, посмотрим на самое последнее слово К Л О Н И Т 35, откуда станет ясно, что 35=Ь.

Перед последней атакой выпишем текст, заменяя известные обозначения буквами.

И Е С Л И М О 31 Е 33 Ь 41 43 Т Ь В Т О Л П Е С О 41 О 69,

П Р И К О Р О Л Е С Н А Р О Д О М С В 73 79 Ь 67 Р А Н И Т Ь

И, 54 В А 31 А 73 М Н Е Н И Е Л 69 41 О Е,

70 Л А В 43 П Е Р Е Д М О Л В О 69 Н Е К Л О Н И Т Ь ...

Из последней строки: 69=Ю, тогда слова Л Ю 41 О Е и С О 41 О Ю определяют 41: 41=Б. Теперь из четвертого слова первой строки Б 43 Т Ь получаем, что 43=Ы. А первое слово из последней строки 70 Л А В Ы - это Г Л А В Ы. Слово в первой строке М О 31 Е 33 Ь угадывается из контекста: М О Ж Е Ш Ь, т. е. 31=Ж, 33=Ш. Теперь второе слово в третьей строке запишется как 54 В А Ж А 73, откуда с учетом контекста: 54=У, 73=Я. После этого окончание второй строки имеет вид "... С В Я 79 Ь 67 Р А Н И Т Ь". Легко определяются буквы 79=З, 67=Х.

**Ответ:**

И Е С Л И М О Ж Е Ш Ь Б Ы Т Ь В Т О Л П Е С О Б О Ю,

П Р И К О Р О Л Е С Н А Р О Д О М С В Я З Ь Х Р А Н И Т Ь

И, У В А Ж А Я М Н Е Н И Е Л Ю Б О Е,

Г Л А В Ы П Е Р Е Д М О Л В О Ю Н Е К Л О Н И Т Ь...

**Пример 6.6** *То же самое задание:*

56 67 92 18 58 39 99 27 87 67 56 25 56 80 67 10 17 92 39 62 25 56 27 24

95 56 31 95 46 27 73 56 31 17 58 39 58 67 95 58 92 56 95 40 24 40 17 92 39

62 69 39 40 17 56 67 58 - 56 18 99 92 46 67 56 87, 69 56 69 39 36 80 17 92

67 27 39 40 87 56 17 58 73 40. 25 56 39 73 56 10 17 92, 56 43 92 80 40 10, 95

56 23 80 40 23 17 40 24 40 25 46 92 69 14 95 67 27 73 95 73 58 87 67 56 73

58. 69 39 58 69 56 95 46 27 23 25 46 92 67 10 17 56 38 58 73 95 92 58 56 38

58 46 73 40 67 92 10. 25 46 92 18 56 46 56 69 92 25 27 17 62 73 56 69 24 80

58 39 62 18 14 17 56 25 46 58 69 58 17 92 95 56 58 87 67 56 43 58 39 73 69

56, 23 17 40 24 40 46 40 24 18 58 23 40 17 92 39 62. 56 80 67 40 95 56 18 17

40 23 56 80 40 46 10 73 58 87 43 58 80 69 27 87 67 58 80 58 17 10 87 73 46

58 67 92 46 56 69 56 95 67 40 87 40 95 58 73 58 92 73 14 39 10 38 58 95 46

40 73 67 56 25 56 69 73 56 46 58 67 67 14 87 67 40 39 73 40 69 17 58 67 92

10 87 92 67 39 73 46 27 95 73 56 46 40 56 67 92 39 56 69 58 46 99 58 67 67

156 73 56 38 67 56 24 67 40 17 92, 24 40 38 58 87 25 46 92 99 17 92. 25 56 67

10 73 92 10 67 58 92 87 58 17 92, 80 17 10 38 58 23 56 95 56 67 95 46 58 73

67 56 25 46 58 80 67 40 24 67 40 38 58 67 14 69 39 58 71 73 92 99 73 27 95

92 - 67 56 56 73 67 92 82 71 73 56 23 56 92 67 58 73 46 58 18 56 69 40 17 56

39 62.

67 58 25 46 56 99 17 56 92 87 92 67 27 73 14, 95 40 95 56 67 27 69 92 80

58 17 51 58 17 62 92 82 67 58 17 58 23 95 56 23 56 92 71 95 24 56 73 92 38

58 39 95 56 23 56 25 27 73 58 99 58 39 73 69 92 10 - 73 46 92 25 27 17 62 73

40 25 56 25 46 40 69 56 87 27 18 56 46 73 27, 27 39 14 25 40 67 67 14 58 38

58 46 73 56 69 56 31 27 31 87 56 31 73 27 87 18 17 58 46 56 69, 17 40 87 25

56 38 58 95, 25 58 46 58 95 17 36 38 40 73 58 17 58 31 92 95 67 56 25 56 95.

73 46 92 69 14 25 27 95 17 14 82 71 95 46 40 67 40 69 69 92 80 58 69 58 46

73 92 95 40 17 62 67 14 82 25 46 10 87 56 27 23 56 17 62 67 92 95 56 69 - 56

67 92 39 40 87 14 58, 67 92 95 40 95 56 31 56 99 92 18 95 92. . .

18 56 80 46 56 39 73 92 46 40 80 92, 56 67 25 56 69 73 56 46 92 17 25 46

56 39 58 18 10 25 56 17 36 18 92 69 99 27 36 39 10 51 92 73 40 73 27: ѕ38

73 56 56 80 92 67 38 58 17 56 69 58 95 25 56 39 73 46 56 92 17, 80 46 27 23

56 31 24 40 69 39 58 23 80 40 46 40 24 17 56 87 40 73 62 39 87 56 43 58 73ї.

92, 25 56 82 17 56 25 40 69 25 56 25 17 58 38 27 39 73 46 40 99 92 17 27 69

24 67 40 95 73 56 23 56, 38 73 56 67 40 25 40 46 67 92 95 80 56 17 43 58 67

18 80 92 73 58 17 62 67 56 39 73 56 10 73 62 67 40 99 27 82 58 46 58, 80 56

39 73 40 17 95 92 67 43 40 17 92 24 25 46 92 99 92 73 14 82 67 40 80 95 56

17 58 67 56 87 67 56 43 58 67.

Приведем ответ без подробного решения.

Ключ:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Буква** | **Число** | **Буква** | **Число** | **Буква** | **Число** | **Буква** | **Число** | **Буква** | **Число** |
| **а** | 40 | **ж** | 43 | **м** | 87 | **т** | 73 | **ы** | 14 |
| **б** | 18 | **з** | 24 | **н** | 67 | **у** | 27 | **ь** | 92 |
| **в** | 69 | **и** | 92 | **о** | 56 | **х** | 82 | **э** | 71 |
| **г** | 23 | **й** | 31 | **п** | 25 | **ц** | 51 | **ю** | 36 |
| **д** | 80 | **к** | 95 | **р** | 46 | **ч** | 38 | **я** | 10 |
| **е** | 58 | **л** | 17 | **с** | 39 | **ш** | 99 |  |  |

**Ответ:**

О Н И Б Е С Ш У М Н О П О Д Н Я Л И С Ь П О У З К О Й К Р У

Т О Й Л Е С Е Н К Е И О К А З А Л И С Ь В С А Л О Н Е - О Б Ш И Р

Н О М, В О В С Ю Д Л И Н У С А М О Л Е Т А. П О С Т О Я Л И, О

Ж И Д А Я, К О Г Д А Г Л А З А П Р И В Ы К Н У Т К Т Е М Н О Т

Е. В С Е В О К Р У Г П Р И Н Я Л О Ч Е Т К И Е О Ч Е Р Т А Н И Я.

П Р И Б О Р О В И П У Л Ь Т О В З Д Е С Ь Б Ы Л О П Р Е В Е Л И

К О Е М Н О Ж Е С Т В О, Г Л А З А Р А З Б Е Г А Л И С Ь. О Д Н А

К О Б Л А Г О Д А Р Я Т Е М Ж Е Д В У М Н Е Д Е Л Я М Т Р Е Н И

Р О В О К Н А М А К Е Т Е И Т Ы С Я Ч Е К Р А Т Н О П О В Т О Р

Е Н Н Ы М Н А С Т А В Л Е Н И Я М И Н С Т Р У К Т О Р А О Н И С

О В Е Р Ш Е Н Н О Т О Ч Н О З Н А Л И, З А Ч Е М П Р И Ш Л И. П

О Н Я Т И Я Н Е И М Е Л И, Д Л Я Ч Е Г О К О Н К Р Е Т Н О П Р Е

Д Н А З Н А Ч Е Н Ы В С Е Э Т И Ш Т У К И - Н О О Т Н И Х Э Т О

Г О И Н Е Т Р Е Б О В А Л О С Ь.

Н Е П Р О Ш Л О И М И Н У Т Ы, К А К О Н У В И Д Е Л Ц Е Л Ь

И Х Н Е Л Е Г К О Г О И Э К З О Т И Ч Е С К О Г О П У Т Е Ш Е С Т

В И Я - Т Р И П У Л Ь Т А П О П Р А В О М У Б О Р Т У, У С Ы П А Н

Н Ы Е Ч Е Р Т О В О Й У Й М О Й Т У М Б Л Е Р О В, Л А М П О Ч Е

К, П Е Р Е К Л Ю Ч А Т Е Л Е Й И К Н О П О К. Т Р И В Ы П У К Л

Ы Х Э К Р А Н А В В И Д Е В Е Р Т И К А Л Ь Н Ы Х П Р Я М О У Г

О Л Ь Н И К О В - О Н И С А М Ы Е, Н И К А К О Й О Ш И Б К И. . .

Б О Д Р О С Т И Р А Д И, О Н П О В Т О Р И Л П Р О С Е Б Я П О

Л Ю Б И В Ш У Ю С Я Ц И Т А Т У: "Ч Т О О Д И Н Ч Е Л О В Е К

П О С Т Р О И Л, Д Р У Г О Й З А В С Е Г Д А Р А З Л О М А Т Ь С

М О Ж Е Т". И, П О Х Л О П А В П О П Л Е Ч У С Т Р А Ш И Л У В

З Н А К Т О Г О, Ч Т О Н А П А Р Н И К Д О Л Ж Е Н Б Д И Т Е Л Ь

Н О С Т О Я Т Ь Н А Ш У Х Е Р Е, Д О С Т А Л К И Н Ж А Л И З П

Р И Ш И Т Ы Х Н А Д К О Л Е Н О М Н О Ж Е Н.

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\rightarrow Я, пробуем читать начало:

|  |
| --- |
| Ф\rightarrow Т: ТЯОТО... |
| Ф\rightarrow А: АЯОАО... |
| Ф\rightarrow С: СЯОСО... |

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

Обратим внимание на фрагмент ...яорцагрфцрэцы... С учетом наших знаний это: ПоОТаВОСТОэТЬ, и мы находим еще соответствия: О\rightarrow Р, А\rightarrow И, Э\rightarrow Я. Кроме того, так как Ф\rightarrow С, Ц\rightarrow Т, то из таблицы: В\rightarrow 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\rightarrow Й, и из таблицы: Т\rightarrow К. Но тогда М\rightarrow Ю, и слово получилось СТОЙКОСТЬЮ.

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

...АТАКАМНАНижОНАхдВАиТСЯСТОЙКОСТЬЮ...

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

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

**6.4.3 Советы по выполнению частотного анализа английских текстов**

Начните с подсчета частоты появления каждой из букв шифртекста. Примерно пять букв должны появляться с частотой менее 1 они, вероятно, представляют собой **j, k, q, х** и **z**. Одна из букв должна появляться с частотой более 10 она, по-видимому, представляет собой **е**. Если шифртекст не подчиняется этому распределению частот, то, возможно, исходное сообщение написано не на английском языке. Вы можете определить, какой это язык, если проанализируете частотное распределение букв в шифртексте. К примеру, в итальянском языке ecть три буквы с частотностью более 10 букв с частотностью менее 1 немецком языке буква **е** имеет чрезвычайно высокую частотность - 19 любой шифртекст, в котором одни из букв встречаются столь же часто, является, вполне возможно, немецким. После того, как вы определили язык, для выполнения частотного анализа вам следует воспользоваться соответствующей таблицей частотности букв для данного языка. Если у вас есть нужная таблица частотности букв, то нередко удается дешифровать даже шифртексты на неизвестном языке.

Если установлена взаимосвязь с английским языком, то, как часто и происходит, сразу же открытый текст не появляется, тогда обратите внимание на пары повторяющихся букв. В английском языке чаще всего повторяющимися буквами будут **ss, ее, tt, ff, 11, mm** и **оо**. Если в шифртексте имеются какие-либо повторяющиеся символы, то вы можете считать, что они представляют собой одну из этих пар.

Если в шифртексте имеются пробелы между словами, то постарайтесь определить слова, состоящие из одной, двух или трех букв. Единственными словами в английском языке, состоящими из одной буквы, являются **а** и **I**. Чаще всего встречающимися двухбуквенными словами будут **of**, **to**, **in**, **it**, **is**, **be**, **as**, **at**, **so**, **we**, **he**, **by**, **or**, **on**, **do**, **if**, **me**, **my**, **up**, **an**, **go**, **no**, **us**, **am**. Наиболее часто появляющиеся трехбуквенные слова - **the** и **and**.

Если удастся, подготовьте таблицу частотности букв для сообщения, которое вы стараетесь дешифровать. Например, в военных донесениях стремятся опускать местоимения и артикли, и отсутствие таких слов, как **I**, **he**, **а** и **the**, будет снижать частотность некоторых из чаще всего встречающихся букв.

Если вы знаете, что работаете с военным донесением, вам следует использовать таблицу частотности букв, созданную на основе других военных донесений.

Одно из самых полезных для криптоаналитика умений - это способность благодаря собственному опыту или чисто интуитивно - распознавать слова или даже целые фразы. Аль-Халил, один из первых арабских криптоаналитиков, продемонстрировал свои способности, когда взломал греческий шифртекст. Он предположил, что шифртекст начинается с приветствия "Во имя Бога". Установив, что эти буквы соответствуют определенному фрагменту шифртекста, он смог использовать их в качестве лома и раскрыть остальной шифртекст. Это получило название криб.

В некоторых случаях наиболее часто встречающейся буквой в шифртексте может быть **Е**, следующей по частоте появления - **Т** и т. д. Другими словами, частотность букв в шифртексте уже совпадает с частотностью букв в таблице. По-видимому, буква **Е** в шифртексте является действительно **е**, и то же самое, похоже, справедливо и для других букв, и все же шифртекст выглядит тарабарщиной. В этом случае вы столкнулись не с шифром замены, а с шифром перестановки. Все буквы остались теми же самыми, но находятся они не на своих местах.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Таблица 6.9. Сочетаемость букв английского языка | | | | | | | | | | | | | | | | | | |
| **Г** | **С** | | **Слева** | | | | |  | | **Справа** | | | | | **Г** | | **С** | |
| 19 | 81 | | l,c,d,m,n,s,w,t,r,e,h | | | | | **A** | | n,t,s,r,l,d,c,m | | | | | 6 | | 94 | |
| 55 | 45 | | y,b,n,t,u,d,o,s,a,e | | | | | **В** | | e,l,u,o,a,y,b,r | | | | | 70 | | 30 | |
| 61 | 39 | | u,o,s,n,a,i,l,e | | | | | **С** | | h,o,e,a,i,t,r,l,k | | | | | 59 | | 41 | |
| 52 | 48 | | r,i,l,a,n,e | | | | | **D** | | e,i,t,a,o,u | | | | | 54 | | 46 | |
| 8 | 92 | | c,b,e,m,v,d,s,l,n,t,r,h | | | | | **E** | | r,d,s,n,a,t,m,e,c,o | | | | | 21 | | 79 | |
| 69 | 31 | | s,n,f,d,a,i,e,o | | | | | **F** | | t,o,e,i,a,r,f,u | | | | | 52 | | 48 | |
| 36 | 64 | | o,d,u,r,i,e,a,n | | | | | **G** | | e.h.o.r.a.t.f.w.i.s | | | | | 42 | | 58 | |
| 7 | 93 | | g,e,w,s,c,t | | | | | **H** | | e,a,i,o | | | | | 90 | | 10 | |
| 13 | 87 | | f,m,w,e,n,l,d,s,r,h,t | | | | | **I** | | n,t,s,o,c,r,e,m,a,l | | | | | 17 | | 83 | |
| 28 | 72 | | y,w,t,s,n,e,c,b,a,c | | | | | **J** | | u,o,a,e,m,w | | | | | 88 | | 12 | |
| 53 | 47 | | y,u,i,n,a,r,o,c | | | | | **К** | | e,i,n,a,t,s | | | | | 68 | | 32 | |
| 52 | 48 | | m,p,t,i,b,u,o,e,l,a | | | | | **L** | | e,i,y,o,a,d,u | | | | | 65 | | 35 | |
| 69 | 31 | | s,d,m,r,i,a,o,e | | | | | **M** | | e,a,o,i,p,m | | | | | 71 | | 29 | |
| 89 | 11 | | u,e,o,a,i | | | | | **N** | | d,t,g,e,a,s,o,i,c | | | | | 32 | | 68 | |
| 21 | 79 | | o,d,l,p,h,n,e,c,f,s,i,r,t | | | | | **о** | | n,f,r,u,t,m,l,s,w,o | | | | | 18 | | 82 | |
| 47 | 53 | | r,l,t,n,i,p,m,a,o,u,e,s | | | | | **P** | | o,e,a,r,l,u,p,t,i,s | | | | | 59 | | 41 | |
| 20 | 80 | | o,n,l,e,d,r,s | | | | | **Q** | | u | | | | | 100 | | 0 | |
| 70 | 30 | | p,i,u,t,a,o,e | | | | | **R** | | e,o,a,t,i,s,y | | | | | 61 | | 39 | |
| 48 | 52 | | d,t,o,u,r,n,s,i,a,e | | | | | **S** | | t,e,o,i,s,a,h,p,u | | | | | 41 | | 59 | |
| 43 | 57 | | u,o,d,t,f,e,i,n,s.a | | | | | **Т** | | h,i,o,e,a,t,r | | | | | 38 | | 62 | |
| 35 | 65 | | p,f,t,l,b,d,s,o | | | | | **и** | | n,s,t,r,l,p,b,c | | | | | 8 | | 92 | |
| 88 | 12 | | r,u,o,a,i,e | | | | | **V** | | e,i,o,a | | | | | 99 | | 1 | |
| 48 | 52 | | g,d,y,n,s,t,o,e | | | | | **W** | | a,h,i,e,o,n | | | | | 80 | | 20 | |
| 95 | 5 | | u,n,i,e | | | | | **X** | | p,t,i,a,u,c,k,o | | | | | 38 | | 62 | |
| 24 | 76 | | b,n,a,t,e,r,l | | | | | **Y** | | a,o,s,t,w,h,i,e,d,m | | | | | 38 | | 62 | |
| 88 | 12 | | o,n,a,i | | | | | **Z** | | e,i,w | | | | | 86 | | 14 | |
| Таблица 6.10. Частоты биграмм английского языка, часть 1 | | | | | | | | | | | | | | | | | | | | |
|  | | **А** | | **В** | **С** | **D** | **Е** | | **F** | | **G** | **Н** | **I** | **J** | | **К** | | **L** | | **М** |
| **А** | | 4 | | 20 | 28 | 52 | 2 | | 11 | | 28 | 4 | 32 | 4 | | 6 | | 62 | | 23 |
| **В** | | 13 | | 0 | 0 | 0 | 55 | | 0 | | 0 | 0 | 8 | 2 | | 0 | | 22 | | 0 |
| **С** | | 32 | | 0 | 7 | 1 | 69 | | 0 | | 0 | 33 | 17 | 0 | | 10 | | 9 | | 1 |
| **D** | | 40 | | 16 | 9 | 5 | 65 | | 18 | | 3 | 9 | 56 | 0 | | 1 | | 4 | | 15 |
| **Е** | | 84 | | 20 | 55 | 125 | 51 | | 40 | | 19 | 16 | 50 | 1 | | 4 | | 55 | | 54 |
| **F** | | 19 | | 3 | 5 | 1 | 19 | | 21 | | 1 | 3 | 30 | 2 | | 0 | | 11 | | 1 |
| **G** | | 20 | | 4 | 3 | 2 | 35 | | 1 | | 3 | 15 | 18 | 0 | | 0 | | 5 | | 1 |
| **Н** | | 101 | | 1 | 3 | 0 | 270 | | 5 | | 1 | 6 | 57 | 0 | | 0 | | 0 | | 3 |
| **I** | | 40 | | 7 | 51 | 23 | 25 | | 9 | | 11 | 3 | 0 | 0 | | 2 | | 38 | | 25 |
| **J** | | 3 | | 0 | 0 | 0 | 5 | | 0 | | 0 | 0 | 1 | 0 | | 0 | | 0 | | 0 |
| **К** | | 1 | | 0 | 0 | 0 | 11 | | 0 | | 0 | 0 | 13 | 0 | | 0 | | 0 | | 0 |
| **L** | | 44 | | 2 | 5 | 12 | 62 | | 7 | | 5 | 2 | 42 | 1 | | 1 | | 53 | | 2 |
| **М** | | 52 | | 14 | 1 | 0 | 64 | | 0 | | 0 | 3 | 37 | 0 | | 0 | | 0 | | 7 |
| **N** | | 42 | | 10 | 47 | 122 | 63 | | 19 | | 106 | 12 | 30 | 1 | | 6 | | 6 | | 9 |
| **O** | | 7 | | 12 | 14 | 17 | 5 | | 95 | | 3 | 5 | 14 | 0 | | 0 | | 19 | | 41 |
| **Р** | | 19 | | 1 | 0 | 0 | 37 | | 0 | | 0 | 4 | 8 | 0 | | 0 | | 15 | | 1 |
| **Q** | | 0 | | 0 | 0 | 0 | 0 | | 0 | | 0 | 0 | 0 | 0 | | 0 | | 0 | | 0 |
| **R** | | 83 | | 8 | 16 | 23 | 169 | | 4 | | 8 | 8 | 77 | 1 | | 10 | | 5 | | 26 |
| **S** | | 65 | | 9 | 17 | 9 | 73 | | 13 | | 1 | 47 | 75 | 3 | | 0 | | 7 | | 11 |
| **Т** | | 57 | | 22 | 7 | 1 | 76 | | 5 | | 2 | 330 | 126 | 1 | | 0 | | 14 | | 10 |
| **U** | | 11 | | 5 | 9 | 6 | 9 | | 1 | | 6 | 0 | 9 | 0 | | 1 | | 19 | | 5 |
| **V** | | 7 | | 0 | 0 | 0 | 72 | | 0 | | 0 | 0 | 28 | 0 | | 0 | | 0 | | 0 |
| **W** | | 36 | | 1 | 1 | 0 | 38 | | 0 | | 0 | 33 | 36 | 0 | | 0 | | 4 | | 1 |
| **X** | | 1 | | 0 | 2 | 0 | 0 | | 1 | | 0 | 0 | 3 | 0 | | 0 | | 0 | | 0 |
| **Y** | | 14 | | 5 | 4 | 2 | 7 | | 12 | | 2 | 6 | 10 | 0 | | 0 | | 3 | | 7 |
| **Z** | | 1 | | 0 | 0 | 0 | 4 | | 0 | | 0 | 0 | 0 | 0 | | 0 | | 0 | | 0 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Таблица 6.11. Частоты биграмм английского языка, часть 2 | | | | | | | | | | | | | |
|  | **N** | **О** | **Р** | **Q** | **R** | **S** | **Т** | **U** | **V** | **W** | **X** | **Y** | **Z** |
| **А** | 167 | 2 | 14 | 0 | 83 | 76 | 127 | 7 | 25 | 8 | 1 | 9 | 1 |
| **В** | 0 | 11 | 0 | 0 | 15 | 4 | 2 | 13 | 0 | 0 | 0 | 15 | 0 |
| **С** | 0 | 50 | 3 | 0 | 10 | 0 | 28 | 11 | 0 | 0 | 0 | 3 | 0 |
| **D** | 6 | 16 | 4 | 0 | 21 | 18 | 53 | 19 | 5 | 15 | 0 | 3 | 0 |
| **Е** | 146 | 35 | 37 | 6 | 191 | 149 | 65 | 9 | 26 | 31 | 12 | 5 | 0 |
| **F** | 0 | 51 | 0 | 0 | 26 | 8 | 47 | 6 | 3 | 3 | 0 | 2 | 0 |
| **G** | 4 | 21 | 1 | 1 | 20 | 9 | 21 | 9 | 0 | 5 | 0 | 1 | 0 |
| **Н** | 2 | 44 | 1 | 0 | 3 | 10 | 18 | 6 | 0 | 5 | 0 | 3 | 0 |
| **I** | 202 | 56 | 12 | 1 | 46 | 79 | 117 | 1 | 22 | 0 | 4 | 0 | 3 |
| **J** | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 |
| **К** | 2 | 2 | 0 | 0 | 0 | 6 | 2 | 1 | 0 | 2 | 0 | 1 | 0 |
| **L** | 2 | 25 | 1 | 1 | 2 | 16 | 23 | 9 | 0 | 1 | 0 | 33 | 0 |
| **М** | 1 | 17 | 18 | 1 | 2 | 12 | 3 | 8 | 0 | 1 | 0 | 2 | 0 |
| **N** | 7 | 54 | 7 | 1 | 7 | 44 | 124 | 6 | 1 | 15 | 0 | 12 | 0 |
| **О** | 134 | 13 | 23 | 0 | 91 | 23 | 2 | 55 | 16 | 28 | 0 | 4 | 1 |
| **Р** | 0 | 27 | 9 | 0 | 33 | 14 | 7 | 6 | 0 | 0 | 0 | 0 | 0 |
| **Q** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 17 | 0 | 0 | 0 | 0 | 0 |
| **R** | 16 | 60 | 4 | 0 | 24 | 37 | 55 | 6 | 11 | 4 | 0 | 28 | 0 |
| **S** | 12 | 56 | 17 | 6 | 9 | 48 | 116 | 35 | 1 | 28 | 0 | 4 | 0 |
| **Т** | 6 | 79 | 7 | 0 | 49 | 50 | 56 | 21 | 2 | 27 | 0 | 24 | 0 |
| **U** | 31 | 1 | 15 | 0 | 47 | 39 | 31 | 0 | 3 | 0 | 0 | 0 | 0 |
| **V** | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 |
| **W** | 8 | 15 | 0 | 0 | 0 | 4 | 2 | 0 | 0 | 1 | 0 | 0 | 0 |
| **X** | 0 | 1 | 5 | 0 | 0 | 0 | 3 | 0 | 0 | 1 | 0 | 0 | 0 |
| **Y** | 5 | 17 | 3 | 0 | 4 | 16 | 30 | 0 | 0 | 5 | 0 | 0 | 0 |
| **Z** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

**6.5 Шифр Виженера**

Наиболее известными являются шифры замены или подстановки, особенностью которых является замена символов (или слов, или других частей сообщения) открытого текста соответствующими символами, принадлежащими алфавиту шифртекста. Различают *одноалфавитную* и *многоалфавитную* замену. Вскрытие одноалфавитных шифров основано на учете частоты появления отдельных букв или их сочетаний (биграмм, триграмм и т. п.) в данном языке.

Примером многоалфавитного шифра замены является так называемая система Виженера. *Шифрование* осуществляется *по* таблице, представляющей собой квадратную матрицу размерностью n\times n, где n-число символов используемого алфавита. В [табл. 6.12](https://www.intuit.ru/studies/courses/13837/1234/lecture/31196?page=6#table.6.12) показана *таблица* Виженера для русского языка (*алфавит* Z_{33} - 32 буквы и *пробел*). Первая строка содержит все символы алфавита. Каждая следующая строка получается из предыдущей циклическим сдвигом последней на символ влево.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Таблица 6.12. Таблица Виженера для алфавита Z_{33} | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| A | 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 | Ъ | Ы | Ь | Э | Ю | Я |  |
| 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 | Ъ | Ы | Ь | Э | Ю | Я |  | А |
| В | Г | Д | Е | Ж | З | И | Й | К | Л | М | Н | О | П | Р | С | Т | У | Ф | Х | Ц | Ч | Ш | Щ | Ъ | Ы | Ь | Э | Ю | Я |  | А | Б |
| Г | Д | Е | Ж | З | И | Й | К | Л | М | Н | О | П | Р | С | Т | У | Ф | Х | Ц | Ч | Ш | Щ | Ъ | Ы | Ь | Э | Ю | Я |  | А | Б | В |
| Д | Е | Ж | З | И | Й | К | Л | М | Н | О | П | Р | С | Т | У | Ф | Х | Ц | Ч | Ш | Щ | Ъ | Ы | Ь | Э | Ю | Я |  | А | Б | В | Г |
| Е | Ж | З | И | Й | К | Л | М | Н | О | П | Р | С | Т | У | Ф | Х | Ц | Ч | Ш | Щ | Ъ | Ы | Ь | Э | Ю | Я |  | А | Б | В | Г | Д |
| Ж | З | И | Й | К | Л | М | Н | О | П | Р | С | Т | У | Ф | Х | Ц | Ч | Ш | Щ | Ъ | Ы | Ь | Э | Ю | Я |  | А | Б | В | Г | Д | Е |
| З | И | Й | К | Л | М | Н | О | П | Р | С | Т | У | Ф | Х | Ц | Ч | Ш | Щ | Ъ | Ы | Ь | Э | Ю | Я |  | А | Б | В | Г | Д | Е | Ж |
| И | Й | К | Л | М | Н | О | П | Р | С | Т | У | Ф | Х | Ц | Ч | Ш | Щ | Ъ | Ы | Ь | Э | Ю | Я |  | А | Б | В | Г | Д | Е | Ж | З |
| Й | К | Л | М | Н | О | П | Р | С | Т | У | Ф | Х | Ц | Ч | Ш | Щ | Ъ | Ы | Ь | Э | Ю | Я |  | А | Б | В | Г | Д | Е | Ж | З | И |
| К | Л | М | Н | О | П | Р | С | Т | У | Ф | Х | Ц | Ч | Ш | Щ | Ъ | Ы | Ь | Э | Ю | Я |  | А | Б | В | Г | Д | Е | Ж | З | И | Й |
| Л | М | Н | О | П | Р | С | Т | У | Ф | Х | Ц | Ч | Ш | Щ | Ъ | Ы | Ь | Э | Ю | Я |  | А | Б | В | Г | Д | Е | Ж | З | И | Й | К |
| М | Н | О | П | Р | С | Т | У | Ф | Х | Ц | Ч | Ш | Щ | Ъ | Ы | Ь | Э | Ю | Я |  | А | Б | В | Г | Д | Е | Ж | З | И | Й | К | Л |
| Н | О | П | Р | С | Т | У | Ф | Х | Ц | Ч | Ш | Щ | Э | Ы | Ь | Э | Ю | Я |  | А | Б | В | Г | Д | Е | Ж | З | И | Й | К | Л | М |
| О | П | Р | С | Т | У | Ф | Х | Ц | Ч | Ш | Щ | Ъ | Ы | Ь | Э | Ю | Я |  | А | Б | В | Г | Д | Е | Ж | З | И | Й | К | Л | М | Н |
| П | Р | С | Т | У | Ф | Х | Ц | Ч | Ш | Щ | Ъ | Ы | Ь | Э | Ю | Я |  | А | Б | В | Г | Д | Е | Ж | З | И | Й | К | Л | М | Н | О |
| Р | С | Т | У | Ф | Х | Ц | Ч | Ш | Щ | Ъ | Ы | Ь | Э | Ю | Я |  | А | Б | В | Г | Д | Е | Ж | З | И | Й | К | Л | М | Н | О | П |
| С | Т | У | Ф | Х | Ц | Ч | Ш | Щ | Ъ | Ы | Ь | Э | Ю | Я |  | А | Б | В | Г | Д | Е | Ж | З | И | Й | К | Л | М | Н | О | П | Р |
| Т | У | Ф | Х | Ц | Ч | Ш | Щ | Ъ | Ы | Ь | Э | Ю | Я |  | А | Б | В | Г | Д | Е | Ж | З | И | Й | К | Л | М | Н | О | П | Р | С |
| У | Ф | Х | Ц | Ч | Ш | Щ | Ъ | Ы | Ь | Э | Ю | Я |  | А | Б | В | Г | Д | Е | Ж | З | И | Й | К | Л | М | Н | О | П | Р | С | Т |
| Ф | Х | Ц | Ч | Ш | Щ | Ъ | Ы | Ь | Э | Ю | Я |  | А | Б | В | Г | Д | Е | Ж | З | И | Й | К | Л | М | Н | О | П | Р | С | Т | У |
| Х | Ц | Ч | Ш | Щ | Ъ | Ы | Ь | Э | Ю | Я |  | А | Б | В | Г | Д | Е | Ж | З | И | Й | К | Л | М | Н | О | П | Р | С | Т | У | Ф |
| Ц | Ч | Ш | Ш | Ъ | Ы | Ь | Э | Ю | Я |  | А | Б | В | Г | Д | Е | Ж | З | И | Й | К | Л | М | Н | О | П | Р | С | Т | У | Ф | Х |
| Ч | Ш | Щ | Ъ | Ы | Ь | Э | Ю | Я |  | А | Б | В | Г | Д | Е | Ж | З | И | Й | К | Л | М | Н | О | П | Р | С | Т | У | Ф | Х | Ц |
| Ш | Щ | Ъ | Ы | Ь | Э | Ю | Я |  | А | Б | В | Г | Д | Е | Ж | З | И | Й | К | Л | М | Н | О | П | Р | С | Т | У | Ф | Х | Ц | Ч |
| Щ | Ъ | Ы | Ь | Э | Ю | Я |  | 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 |
| Ъ | Ы | Ь | Э | Ю | Я |  | 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 |  | 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 | Ъ |
| Ь | Э | Ю | Я |  | 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 | Ъ | Ы |
| Э | Ю | Я |  | 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 | Ъ | Ы | Ь |
| Ю | Я |  | 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 | Ъ | Ы | Ь | Э |
| Я |  | 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 | Ъ | Ы | Ь | Э | Ю |
|  | 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 | Ъ | Ы | Ь | Э | Ю | Я |

Выбирается *ключ* или ключевая фраза. После чего процесс зашифровывания осуществляется следующим образом. Под каждой буквой исходного сообщения последовательно записываются буквы ключа; если *ключ* оказался короче сообщения, его используют несколько раз. Каждая буква шифртекста находится на пересечении столбца таблицы, определяемого буквой открытого текста, и строки, определяемой буквой ключа.

Расшифровывание осуществляется следующим образом. Под буквами шифртекста последовательно записываются буквы ключа; в строке таблицы, соответствующей очередной букве ключа, происходит *поиск* соответствующей буквы шифртекста. Находящаяся над ней в первой строке таблицы буква является соответствующей буквой исходного текста.

Для увеличения надежности шифра можно рекомендовать его использование после предварительной псевдослучайной перестановки букв в каждой строке таблицы. Возможны и другие модификации метода.

**6.5.1 Теория криптоанализа шифра Виженера**

(см. [1, §6.5])

Рассмотрим *шифр* модульного гаммирования с уравнением

b_i=(a_i+y_i) ~(\mod  n),

для которого гамма является периодической последовательностью знаков алфавита. Такая гамма обычно получается периодическим повторением некоторого ключевого слова. Например, *ключевое слово* KEY дает гамму KEYKEYKEY... . Рассмотрим задачу вскрытия такого шифра *по* тексту одной криптограммы достаточной длины.

Пусть \mu - *длина* ключевого слова. Обычно *криптоанализ* шифра Виженера проводится в два этапа. На первом этапе определяется число \mu, на втором этапе - само *ключевое слово*.

Для определения числа \mu применяется так называемый тест Казиски, названный в честь Ф. Казиски, применившего его в 1863 г. Тест основан на простом наблюдении того, что два одинаковых отрезка открытого текста, отстоящих друг от друга на расстоянии, кратном \mu, будут одинаково зашифрованы. В силу этого в шифртексте ищутся повторения длины, не меньшей трех и расстояния между ними. Обратим внимание на то, что случайно такие одинаковые отрезки могут появиться в тексте с достаточно малой вероятностью.

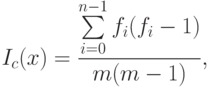
Пусть d_1, d_2,... - найденные расстояния между повторениями и d -наибольший общий делитель этих чисел. Тогда \mu должно делить d. Чем больше повторений имеет текст, тем более вероятно, что \mu совпадает с d. Для уточнения значения \mu, можно использовать так называемый *индекс* совпадения, введенный в практику У. Фридманом в 1920 г.

Для строки x=x_1x_2\ldots x_m длины m, составленной из букв алфавита A, индексом совпадения в x, обозначаемым I_c(x) будем называть *вероятность* того, что две случайно выбранные буквы из x совпадают.

Пусть A=\{a_1,\ldots,a_n\}. Будем отождествлять буквы алфавита с числами, так что

a_1\equiv 0, \ldots, a_{n-1}\equiv n-2,  a_{n}\equiv n-1.

**Теорема 6.1** *Индекс совпадения в x вычисляется по формуле:*



*где f_i - число вхождений буквы n_i в x, i\in Z_n.*

Пусть x - строка осмысленного текста (например, английского). Допустим, как и ранее, что буквы в *x* появляются на любом месте текста с соответствующими вероятностями p_0,..., p_{n-1} независимо друг от друга, где p_i - *вероятность* появления буквы i в осмысленном тексте, i\in Z_n. В такой модели открытого текста *вероятность* того, что две случайно выбранные буквы из x совпадают с i\in Z_n равна p_i^2, и, следовательно,

|  |  |
| --- | --- |
| I_c(x)\approx \sum\limits_{i=0}^{n-1} p_i^2. | ( 6.1) |

Взяв за основу значения вероятностей p_i для открытых текстов на английском языке, получаем приближение \sum_{i=0}^{25} p_i^2\approx 0,066. Тем самым, для английских текстов x можно пользоваться следующим приближением для индекса совпадения:

I_c(x)\approx 0,066.

Аналогичные приближения можно получить и для других языков. Так, для русского языка получаем приближение:

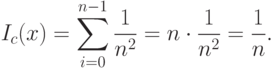
I_c(x)\approx 0,053.

Приведем значения индексов совпадения для ряда европейских языков.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Таблица 6.13. Индексы совпадения европейских языков | | | | | | |
| Язык | Русский | Алгл. | Франц. | Нем. | Итал. | Испан. |
| I_c(x)\approx | 0,0529 | 0,0662 | 0,0778 | 0,0762 | 0,0738 | 0,0775 |

Рассуждения, использованные при выводе формулы (4), остаются, очевидно, справедливыми и в случае, когда х - результат зашифровывания некоторого открытого текста простой заменой. В этом случае вероятности p_i переставляются местами, но сумма \sum_{i=0}^{n-1}p_i^2 остается неизменной.

Предположим, что х - реализация независимых испытаний случайной величины, имеющей равномерное распределение на Z_n. Тогда *индекс* совпадения вычисляется *по* формуле



Вернемся к вопросу об определении числа \mu.

Пусть y= y_1 y_2\dots y_n - данный *шифртекст*. Выпишем его с периодом \mu:

|  |  |  |  |
| --- | --- | --- | --- |
| Y_1^\downarrow | Y_2^\downarrow | \cdots | Y_\mu^\downarrow |
| y_1 | y_2 | \cdots | y_\mu |
| y_{\mu+1} | y_{\mu+2} | \cdots | y_{2\mu} |
| y_{2\mu+1} | y_{2\mu+2} | \cdots | y_{3\mu} |
| \cdots | \cdots | \cdots | \cdots |

и обозначим столбцы получившейся таблицы через Y^\downarrow_1,Y^\downarrow_2,\ldots,Y^\downarrow_\mu. Если \mu - это истинная *длина* ключевого слова, то каждый столбец Y^\downarrow_i, i=1,\ldots,\mu, представляет собой участок открытого текста, зашифрованный простой заменой, определяемой подстановкой

|  |  |
| --- | --- |
| \left(\begin{array}{ccccccc}0&1&2&\ldots&n-s&\ldots&n\\ s&s+1&s+2&\ldots&0&\ldots&s-1\end{array}\right) | ( 6.2) |

для некоторого s\in\{0,1,\ldots,n-1\} (числа берутся *по* модулю n).

В силу сказанного выше, (для английского языка) I_c (Y^\downarrow_i)\approx 0,066 при любом i. С другой стороны, если \mu отлично от длины ключевого слова, то столбцы Y^\downarrow_i будут более "случайными", поскольку они являются результатом зашифровывания фрагментов открытого текста некоторым многоалфавитным шифром. Тогда I_c (Y^\downarrow_i будет ближе (для английского языка) к числу 1/28 \approx 0,038.

Заметная разница значений I_c(x) для осмысленных открытых текстов и случайных последовательностей букв (для английского языка - 0,066 и 0,038, для русского языка - 0,053 и 0,030) позволяет в большинстве случаев установить точное *значение* \mu.

Предположим, что на первом этапе мы нашли длину ключевого слова \mu. Рассмотрим теперь вопрос о нахождении самого ключевого слова. Для его нахождения можно использовать так называемый *взаимный индекс совпадения*.

Пусть x=(x_1x_2\ldots x_m), y=(y_1y_2\ldots y_{m'}) - две строки букв алфавита A. Взаимным индексом совпадения x и y, обозначаемым MI_c(x,y), называется *вероятность* того, что случайно выбранная буква из x совпадает со случайно выбранной буквой из y.

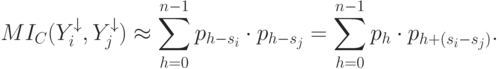
Пусть f_0, f_1,\ldots, f_n$ и $f^1_0, f^1_1,\ldots, f^1_{n-1} - числа вхождений букв алфавита в x и y соответственно.

*Взаимный индекс совпадения в x и y вычисляется по формуле*

|  |  |
| --- | --- |
| MI_c(x,y)=\frac{\sum\limits_{i=0}^{n-1} f_i\cdot f^1_i}{m\cdot m'}. | ( 6.3) |

Пусть k=(k_1,\ldots, k_\mu) - истинное *ключевое слово*. Попытаемся оценить индексы MI_c(Y^{\downarrow}_i, Y^\downarrow_j).

Для этого напомним, что Y^\downarrow_s является результатом зашифровывания фрагмента открытого текста простой заменой, определяемой подстановкой (5) при некотором s. *Вероятность* того, что в $Y^\downarrow_i и Y^\downarrow_j произвольная пара букв равна 0, имеет вид p_{n-s_i}\cdot p_{n-s_j} (где p_\alpha - *вероятность* появления буквы \alpha в открытом тексте); *вероятность* того, что обе буквы есть 1, равна p_{n-s_i+1}\cdot p_{n-s_j+1}, и т. д. На основании этого получаем:



Заметим, что сумма в правой части последнего равенства зависит только от разности (s_i-s_j) ~(\mod  n), которую назовем *относительным сдвигом* Y^\downarrow_i и Y^\downarrow_j. Заметим также, что

|  |  |
| --- | --- |
| \sum\limits_{j=0}^{n-1} p_j \cdot p_{j+s ~(\mod  n)} = \sum\limits_{j=0}^{n-1} p_j \cdot p_{(j-s) ~(\mod  n)}, | ( 6.4) |

поэтому Y^\downarrow_i и Y^\downarrow_j с относительными сдвигами s и n-s имеют одинаковые взаимные индексы совпадения. Приведем таблицу значений сумм (6.4) для английского языка:

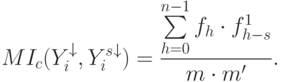
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Таблица 6.14. Взаимный индекс совпадения при сдвиге s | | | | | | | |
| Сдвиг s | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| MI_c(x,y)\approx | 0,066 | 0,039 | 0,032 | 0,034 | 0,044 | 0,033 | 0,036 |
| Сдвиг s | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| MI_c(x,y)\approx | 0,039 | 0,034 | 0,034 | 0,038 | 0,045 | 0,039 | 0,043 |

Обратим внимание на то, что ненулевые "сдвиги" дают взаимные индексы совпадения, изменяющиеся в пределах от 0,032 до 0,045, в то время как при нулевом сдвиге *индекс* MI_c(x,y) близок к 0,066. Это наблюдение позволяет определить величины относительных сдвигов s_i-s_j столбцов Y^\downarrow_i и Y^\downarrow_j. Для этого заметим, что при некотором значении s(i,j)\in \{0,1,\ldots, n-1\} столбец {Y^{s(i,j)_j}}^\downarrow , полученный из Y^\downarrow_j прибавлением к каждому его элементу числа s(i,j) (*по* модулю n), имеет нулевой относительный сдвиг с Y^\downarrow_i.

Пусть {Y^0_j}^\downarrow,{Y^1_j}^\downarrow,\ldots,{Y^{n-1}_j}^\downarrow - результаты зашифровывания Y^\downarrow_j каждой из простых замен (6.2). Несложно вычислить взаимные индексы

MI_c(Y^\downarrow_i, {Y^s_i}^\downarrow),\quad 0\leq s\leq n-1,\quad 1\leq i<j\leq \mu

(всего, таким образом, имеется C_\mu^2\cdot n значений). Для этого воспользуемся формулой, полученной из (6.3):



Если s равно s_i-s_j (относительному сдвигу Y^\downarrow_i и Y^\downarrow_j), то взаимный *индекс* впадения должен быть (для английского языка) близок к 0,066, так как относительный сдвиг Y^\downarrow_i и {Y^s_j}^\downarrow равен нулю. Если же s не равно s_i - s_j, то взаимный *индекс* совпадения должен колебаться в пределах 0,032-0,045.

Используя изложенный метод, мы сможем связать системой уравнений относительные сдвиги различных пар столбцов Y^\downarrow_i и Y^\downarrow_j. В результате останется 26 (для английского языка) вариантов для ключевого слова, из которых можно выбрать наиболее предпочтительный вариант (если *ключевое слово* является осмысленным).

Следует отметить, что предложенный метод будет эффективным для не слишком больших значений \mu. Это объясняется тем, что для хороших сближений индексов совпадения требуются тексты достаточно большой длины.

*Дан текст, зашифрованный шифром Виженера, требуется определить ключевое слово и прочитать открытый текст.*

влцдутжбюцхъяррмшбрхцэооэцгбрьцмйфктъъюьмшэсяцпунуящэйтаьэд

кцибрьцгбрпачкъуцпъбьсэгкцъгуущарцјэвърюуоюэкааэбрняфукабъарпяъа

фкъиьжяффнйояфывбнэнфуюгбрьсшьжэтбэјчюъюръегофкбьчябашвјэуъъ

юаднчжчужцјэвлрнчулбюпцуруньъшсэюъзкцхъяррнрювяспэмасчкпэужьж

ыатуфуярюравртубурьпэщлафоуфбюацмнубсюкйтаьэдйюнооэгюожбгкбрън

цэпотчмјодзцвбцшщвщепчдчдръюьскасэгъппэгюкдойрсрэвоопчщшоказръб

бнэугнялјкьсрбјуыэбдэулбюасшоуэтъшкрсдугэфлбубуъчнчтртпэгюкиугюэм

эгюккъъпэгяапуфуэзьрадзьжчюрмфцхраююанчјчюъ ыхьъцомэфъцпоирьк

нщпэтэузуябащущбаыэйчдфрпэцъьрьцъцпоилуфэдцойэдятррачкубуфнйтаь

эдкцкрннцюабугюуубурьпйюэъжтгюркующоъуфъэгясуоичщщчдцсфырэдщ

эъуяфшјчцюйрщвяхвмкршрпгюопэуцчйтаьэдкцибрьцыяжтюрбуэтэбдуящэ

убъибрювъежагибрбагбрымпуноцшяжцечкфодщоъчжшйуъцхчщвуэбдлдъэ

гясуахзцэбдэулькнъщбжяцэьрјдъьвювлрнуяфуоухфекьгцчччгэъжтанопчы

нажпачкъуъмэнкйрэфщэъьбудэндадъярьеюэлэтчоубъцэфэвлнјэгфдсэвэјкб

счоукгаутэыпуббцчкпэгючсаъбэнэфъркацхјваетуфяепьрювържадфјжбьфу

тощоявьъгупчршуитеачйчирамчюфчоуяюонкяжыкгсцбрясшчйотъъжрсщчл

**Решение.** Для определения числа букв в данном ключевом слове применяется так называемый тест Казиски. Тест основан на простом наблюдении того, что два одинаковых отрезка открытого текста, отстоящих друг от друга на расстоянии, кратном \mu (количество букв в слове), будут одинаково зашифрованы. В силу этого в шифртексте ищутся повторения длины не меньшей трех, и расстояния между ними. Необходимо обратить внимание на то, что случайно такие одинаковые отрезки могут появиться в тексте с достаточно малой вероятностью.

В данном тексте обнаружено четырехкратное повторение буквосочетания "брь". Выясним *расстояние* между ними и найдем наибольший общий делитель этих расстояний.

В результате получаем: 35, 85, 510, их НОД = 5.

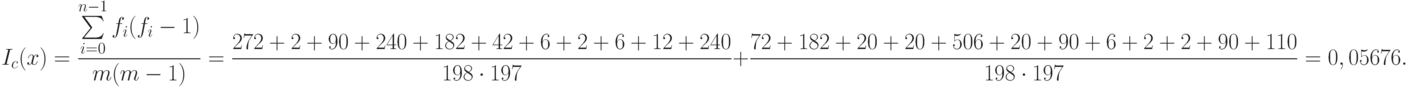
Следовательно, с определенной долей вероятности можно заключить, что *длина* кодового слова равна 5.

Для подтверждения гипотезы воспользуемся математической статистикой для определения длины ключевого слова. Для этого запишем *шифр*-текст в таблицу с 5 столбцами, предполагая, что *длина* ключевого слова равна 5.

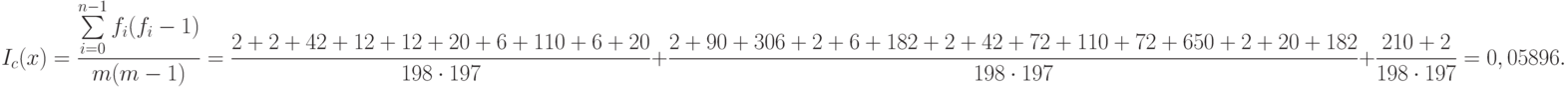
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Y_1^\downarrow | Y_2^\downarrow | Y_3^\downarrow | Y_4^\downarrow | Y_5^\downarrow |
| в | л | ц | д | у |
| т | ж | б | ю | ц |
| х | ъ | я | р | р |
| \ldots | \ldots | \ldots | \ldots | \ldots |
| т | ъ | ъ | ж | р |
| с | щ | ч | л |  |

Вычислим взаимные индексы совпадения I_c(x) букв в каждом из столбцов таблицы, для достоверного установления длины ключевого слова. Для этого посчитаем частоту повторения букв в каждом столбце.

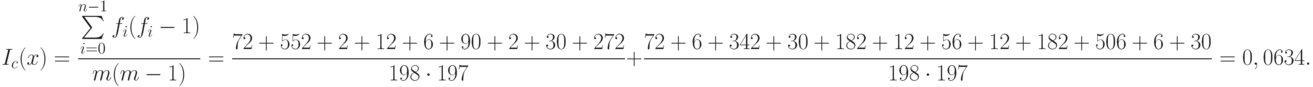
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Таблица 6.15. 1 столбец (общее количество букв m=198) | | | | | | | | | | | | | |
| Обозначение | а | б | в | г | д | е | ё | ж | з | и | й | к | л |
| Количество | 17 | 2 | 10 | 16 | 14 | 7 | 0 | 1 | 1 | 3 | 2 | 1 | 0 |
| Обозначение | м | н | о | п | р | с | т | у | ф | х | ц | ч | ш |
| Количество | 3 | 4 | 1 | 0 | 1 | 16 | 9 | 14 | 5 | 5 | 23 | 0 | 0 |
| Обозначение | щ | ъ | ы | ь | э | ю | я |  |  |  |  |  |  |
| Количество | 5 | 10 | 3 | 2 | 2 | 10 | 11 |  |  |  |  |  |  |



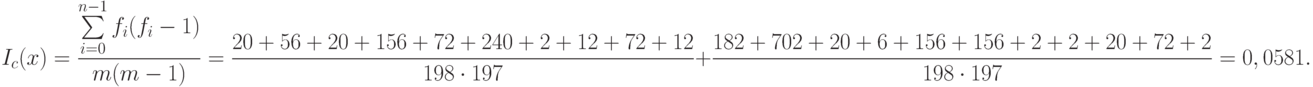
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Таблица 6.16. 2 столбец (общее количество букв m=198) | | | | | | | | | | | | | |
| Обозначение | а | б | в | г | д | е | ё | ж | з | и | й | к | л |
| Количество | 2 | 2 | 0 | 7 | 1 | 0 | 0 | 4 | 4 | 5 | 0 | 3 | 11 |
| Обозначение | м | н | о | п | р | с | т | у | ф | х | ц | ч | ш |
| Количество | 3 | 5 | 2 | 10 | 18 | 0 | 2 | 3 | 14 | 2 | 7 | 9 | 11 |
| Обозначение | щ | ъ | ы | ь | э | ю | я |  |  |  |  |  |  |
| Количество | 9 | 26 | 2 | 5 | 14 | 15 | 2 |  |  |  |  |  |  |



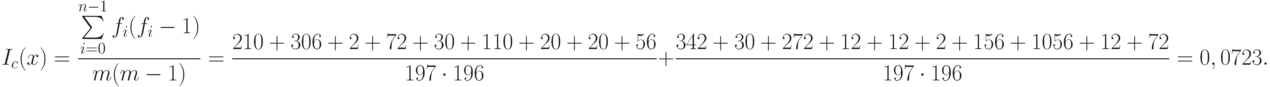
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Таблица 6.17. 3 столбец (общее количество букв m=198) | | | | | | | | | | | | | |
| Обозначение | А | б | в | г | д | е | ё | ж | з | и | й | к | л |
| Количество | 9 | 24 | 1 | 1 | 1 | 2 | 4 | 0 | 1 | 0 | 3 | 10 | 0 |
| Обозначение | М | н | о | п | р | с | т | у | ф | х | ц | ч | ш |
| Количество | 2 | 6 | 17 | 1 | 9 | 1 | 3 | 19 | 0 | 1 | 6 | 14 | 4 |
| Обозначение | Щ | ъ | ы | ь | э | ю | я |  |  |  |  |  |  |
| Количество | 1 | 8 | 4 | 14 | 23 | 3 | 6 |  |  |  |  |  |  |



|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Таблица 6.18. 4 столбец (общее количество букв m=198) | | | | | | | | | | | | | |
| Обозначение | А | б | в | г | д | е | ё | ж | з | и | й | к | л |
| Количество | 0 | 5 | 8 | 5 | 13 | 0 | 9 | 16 | 0 | 3 | 9 | 15 | 2 |
| Обозначение | М | н | о | п | р | с | т | у | ф | х | ц | ч | ш |
| Количество | 4 | 9 | 4 | 14 | 27 | 5 | 3 | 13 | 13 | 2 | 0 | 1 | 0 |
| Обозначение | Щ | ъ | ы | ь | э | ю | я |  |  |  |  |  |  |
| Количество | 2 | 5 | 0 | 0 | 0 | 9 | 2 |  |  |  |  |  |  |

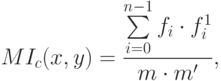


|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Таблица 6.19. 5 столбец (общее количество букв m=197) | | | | | | | | | | | | | |
| Обозначение | А | б | в | г | д | е | ё | ж | з | и | й | к | л |
| Количество | 15 | 18 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 2 | 9 | 1 |
| Обозначение | М | н | о | п | р | с | т | у | ф | х | ц | ч | ш |
| Количество | 1 | 6 | 11 | 5 | 5 | 0 | 8 | 19 | 0 | 1 | 6 | 17 | 0 |
| Обозначение | Щ | ъ | ы | ь | э | ю | я |  |  |  |  |  |  |
| Количество | 4 | 4 | 2 | 13 | 33 | 4 | 9 |  |  |  |  |  |  |



*По* полученным индексам совпадения можно сказать, что *длина* ключевого слова выбрана верно и равна 5.

После того как мы нашли длину ключевого слова, произведем *поиск* его истинного значения. Для его нахождения можно использовать так называемый взаимный *индекс* совпадения



где f_i, f_i^1 - частота буквы i в столбцах Y_i^\downarrow и (Y_i^1)^\downarrow соответственно; m, m' - число букв в этих столбцах.

Так как каждый из столбцов таблицы является результатом зашифровывания фрагмента открытого текста простой заменой, определяемой подстановкой, то попытаемся оценить взаимные индексы совпадения.

Взаимный *индекс* совпадения значения ключевого слова для русского языка должен находиться в приделах 0,053 - 0,07. И для его вычисления предварительно необходимо определить относительный сдвиг всех столбцов относительно первого.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Таблица 6.20. Сдвиг 2-го столбца на 6 позиций | | | | | | | | | | | | | |
| Обозначение | А | б | в | г | д | е | ё | ж | з | и | й | к | л |
| Количество | 26 | 2 | 5 | 14 | 15 | 2 | 2 | 2 | 0 | 7 | 1 | 0 | 0 |
| Обозначение | М | н | о | п | р | с | т | у | ф | х | ц | ч | ш |
| Количество | 4 | 4 | 5 | 0 | 3 | 11 | 3 | 5 | 2 | 10 | 18 | 0 | 2 |
| Обозначение | Щ | ъ | ы | ь | э | ю | я |  |  |  |  |  |  |
| Количество | 3 | 14 | 2 | 7 | 9 | 11 | 9 |  |  |  |  |  |  |

MI_c(Y_1^\downarrow,{Y_2^6}^\downarrow)= 0,05494.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Таблица 6.21. Сдвиг 3-го столбца на 3 позиции | | | | | | | | | | | | | |
| Обозначение | а | б | в | г | д | е | ё | ж | з | и | й | к | л |
| Количество | 23 | 3 | 6 | 9 | 24 | 1 | 1 | 1 | 2 | 4 | 0 | 1 | 0 |
| Обозначение | м | н | о | п | р | с | т | у | ф | х | ц | ч | ш |
| Количество | 3 | 10 | 0 | 2 | 6 | 17 | 1 | 9 | 1 | 3 | 19 | 0 | 1 |
| Обозначение | щ | ъ | ы | ь | э | ю | я |  |  |  |  |  |  |
| Количество | 6 | 14 | 4 | 1 | 8 | 4 | 14 |  |  |  |  |  |  |

MI_c(Y_1^\downarrow,{Y_3^3}^\downarrow)= 0,5798.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Таблица 6.22. Сдвиг 4-го столбца на 16 позиций | | | | | | | | | | | | | |
| Обозначение | а | б | в | г | д | е | ё | ж | з | и | й | к | л |
| Количество | 27 | 5 | 3 | 13 | 13 | 2 | 0 | 1 | 0 | 2 | 5 | 0 | 0 |
| Обозначение | м | н | о | п | р | с | т | у | ф | х | ц | ч | ш |
| Количество | 0 | 9 | 2 | 0 | 5 | 8 | 5 | 13 | 0 | 9 | 16 | 0 | 3 |
| Обозначение | щ | ъ | ы | ь | э | ю | я |  |  |  |  |  |  |
| Количество | 9 | 15 | 2 | 4 | 9 | 4 | 14 |  |  |  |  |  |  |

MI_c(Y_1^\downarrow,{Y_4^{16}}^\downarrow)= 0,06068.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Таблица 6.23. Сдвиг 5-го столбца на 3 позиции | | | | | | | | | | | | | |
| Обозначение | а | б | в | г | д | е | ё | ж | з | и | й | к | л |
| Количество | 33 | 4 | 9 | 15 | 18 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| Обозначение | м | н | о | п | р | с | т | у | ф | х | ц | ч | ш |
| Количество | 2 | 9 | 1 | 1 | 6 | 11 | 5 | 5 | 0 | 8 | 19 | 0 | 1 |
| Обозначение | щ | ъ | ы | ь | э | ю | я |  |  |  |  |  |  |
| Количество | 6 | 17 | 0 | 4 | 4 | 2 | 13 |  |  |  |  |  |  |

MI_c(Y_1^\downarrow,{Y_5^3}^\downarrow)= 0,106045.

*По* взаимным индексам совпадения можно судить, что сдвиги между столбцами выбраны верно. Составим уравнения для определения ключевого слова:

\begin{array}{lll}
    g[1]-g[2]=6 & g[1]=g[2] + 6 & g[2]=g[1] - 6 \\
    g[1]-g[3]=3 & g[1]=g[3] + 3 & g[3]=g[1] - 3 \\
    g[1]-g[4]=16 & g[1]=g[4] + 16 & g[4]=g[1] - 16\\
    g[1]-g[5]=3 & g[1]=g[5] + 3 & g[5]=g[1] - 3
    \end{array}

Теперь только необходимо вычислить *значение* g[1]:

\begin{array}{lll}
    g[l]=1: быюсю & g[l]=2: вьятя & g[l]=3: гэауа\\
    g[l]=4: дюбфб & g[l]=5: еявхв & g[l]=6: ёагцг\\
    g[l]=7: жбдчд & g[l]=8: звеше & g[l]=9: игёщё\\
    g[l]=10: йджъж & g[l]=ll: кезыз & g[l]=12: лёиьи\\
    g[1]=13: мжйэй & g[l]=14: нзкюк & g[l]=15: оилял\\
    g[l]=16: пймам & g[l]=17: pкнбн & g[l]=18: "\textbf{слово}"\\
    g[l]=19: тмпгп & g[l]=20: унрдр & g[l]=21: фосес \\
    g[l]=22: xптёт & g[l]=23: цружу & g[l]=24: чсфзф \\
    g[l]=25: штхих & g[l]=26: щуцйц & g[l]=27: ъфчкч \\
    g[l]=28: ыхшлш & g[l]=29: ьцщмщ & g[l]=30: эчънъ \\
    g[l]=31: юшыоы & g[l]=32: ящьпь &
    \end{array}

Найдено одно *ключевое слово* "*СЛОВО*".

Расшифруем зашифрованный текст:

развебытьздоровымтожесамоечтонебытьбольнымопределенноздоровьеэт

онечтобольшеедлянасфизическоездоровьеэтоисостояниеиспособностьиэнерг

иязаниматьсятемчтонамнеобходимополучатьприэтомудовольствиеивыздора

вливатьбезвсякойпомощиздоровьепарадоксальновынеможетенепосредствен

нозаставитьсебястатьздоровымвамостаетсятольконаблюдатьзатемкакудиви

тельнаяспособностьвашегоорганизмаисцелятьсебяначинаетдействоватьсама

собойивашебогатствоилибедностьжестокостьилидобродетельностьнеимеютз

десьповидимомуникакогозначенияздоровьеэтонечтопозитивноеононеознача

етотказотудовольствияздоровьеявляетсяестественнымследствиемнашегообр

азажизнивзаимоотношенийдиетыокружающейобстановкиздоровьеэтонепред

метсобственностиэтопроцессэтоточтомыделаемрезультатнашихмыслейичувс

твэтообразсуществованияинтересночтонаправлениемедицинскихисследован

ийвсебольшеибольшеотклоняетсявсторонутойобластикотораядосихпорсчита

ласьсферойдеятельностипсихологовисейчасужетруднопровестичеткиеразгр

аничениямеждуфизическимииментальнымифакторамизаболеваний

*Ключевое слово* верное, текст читается.

*Замечание.* *Криптоанализ* шифра Виженера является трудоемкой задачей для решения без использования программ. Лабораторный практикум, подготовленный авторами, содержит задачи, аналогичные представленной, для успешного решения этих задач предлагается программный *инструментарий*. Использование этого инструментария позволяет решать такие задачи за сравнительно короткое время.

**6.6 Преобразование сдвига**

См. [4])

Для ускорения зашифрования и расшифрования удобно иметь относительно простое правило реализации такой перестановки. Один из способов-рассматривать множество целых чисел \{0, 1, 2, \ldots, N-1\} и использовать на нем *операции* сложения и умножения *по* модулю N.

Итак, пусть используется N-буквенный *алфавит* с числовыми эквивалентами 0, 1, 2, {\ldots},N-1. Пусть b - фиксированное целое число. Преобразованием *сдвига* будем называть функцию f, определенную правилом C=f(P)=P+b ~(\mod  N). Случай N=26, b=3 - это шифрсистема, изобретенная Юлием Цезарем. Для расшифрования элемента шифртекста C\in\{0, 1, 2, {\ldots},N-1\} надо будет просто вычислить P=f^{-1}(C)\equiv C-b ~(\mod  N). *Параметр* b является ключом зашифрования.

**Пример 6.8** Предположим, что "FQOCUDEM", которое , как нам известно, зашифровано с использованием преобразования сдвига на буквах 26-буквенного алфавита. Нам надо найти b для успешного "взлома" шифра. Сделать это можно при помощи частотного анализа. Он работает следующим образом. Предположим, что нами перехвачен длинный *отрезок* шифртекста, скажем, несколько сотен букв. Мы знаем, что "E" - наиболее часто встречающаяся буква английского языка. Поэтому разумно предположить, что наиболее часто встречающаяся в шифртексте буква является результатом зашифрования буквы "E". Пусть чаще других в шифртексте встречается буква "U". Это значит, что сдвиг преобразует "E"=4 в "U"=20, то есть 20\equiv 4+b ~(\mod  26), так что b=16. Осталось вычесть 16 (*по* модулю 26) из числовых эквивалентов "FQOCUDEM":

"FQOCUDEM"=5, 16, 14, 2, 20, 3, 4, 12 \rightarrow 15, 0, 24, 12, 4, 13, 14, 22="PAYMENOW".

Если *шифрование* применяется к буквам 32-буквенного алфавита, то для b имеется всего 32 возможности, можно просто попробовать их все. И, скорее всего, лишь одному из значений b будет соответствовать осмысленное сообщение, такое b и есть *ключ* зашифрования.

Значит, этот тип криптосистем слишком прост, чтобы быть хорошим. Вскрыть его очень легко. Его можно усовершенствовать, используя более широкий *класс* преобразований, называемых аффинными отображениями: C=a\cdot P + b ~(\mod  N), где a и b-фиксированные целые числа(вместе они образуют *ключ* зашифрования).

Чтобы расшифровать сообщение, зашифрованное применением аффинного отображения, нужно просто выразить P через C: P\equiv s\cdot C+t ~(\mod  N), где s=a^{-1} ~(\mod  N) есть обратное к a *по* модулю N число, а t равно -a^{-1}b. Конечно, это возможно лишь при условии \GCD (a,N)=1. В противном случае нельзя выразить P через C: если \GCD (a,N)>1, то одной букве шифртекста соответствует несколько букв открытого текста, поэтому нельзя однозначно восстановить *открытый текст* *по* зашифрованному. *По* определению, такое преобразование не является шифрующим, так как последнее обязано быть взаимно однозначным, то есть *открытый текст* должен определяться шифртекстом однозначно.

Итак, аффинная *криптосистема* над N-буквенным алфавитом с параметрами a и b задается правилами

C=a\cdot P+b ~(\mod  N), \qquad  P = s\cdot C + t ~(\mod  N),

где s=a^{-1}, t=- a^{-1}b.

Как частный случай аффинной криптосистемы при a=1 получаем *преобразование сдвига*. В другом частном случае при b=0 получаем P \equiv a\cdot C ~(\mod  N), C=a^{-1}P ~(\mod  N). Такое преобразование называется *линейным*. Оно отображает сумму в сумму, т. е. если {P}_{1} при шифровании переходит в {C}_{1}, а {P}_{2} - в {C}_{2}, то {P}_{1}+ {P}_{2} переходит в {C}_{1}+ {C}_{2} (*сложение*, конечно, производится *по* модулю N).

Пусть теперь нам известно, что перехваченное сообщение зашифровано применением аффинного отображения букв - буквенного алфавита. Мы хотим определить *ключ* a, b, чтобы прочитать это сообщение. Для этого нужно знать, как зашифровываются какие-нибудь две буквы.

*Используется русский алфавит из 31 буквы: буквы Е и Ё не различаются, буквы И и Й не различаются. Каждой букве такого алфавита поставлено в соответствие число: А соответствует 0, Б-1, В-2,…., Ю-29, Я-30. Известен алгоритм зашифрования: номер буквы открытого текста умножается на число а, к результату прибавляется число b, находится остаток от деления полученного числа на 31. Этот остаток определяет номер буквы в шифрованном тексте. Пробелы и знаки препинания сохранены. Ключом является пара чисел (a, b). Итак, алгоритм известен, ключ Вам неизвестен. При зашифровывании противник не учел особенности русского языка, которые позволят Вам и без знания ключа расшифровать текст. Шифрованный текст:*

**ВЯТЖХШИАВБ-АНЯЯЖЬГМ, НН ЯИЕЖ ЫИЦВЦИГМ.**

**Решение.** Слову из двух одинаковых букв НН может соответствовать в открытом тексте только *слово* ЕЁ (напомним, Е и Ё при шифровании не различались). Итак, при расшифровывании нужно Н заменить на Е (Ё).

И у нас есть *информация* о ключе зашифрования: 5a+b=12 (\mod 31).

Обратим внимание на третью и четвертую буквы второго слова. Заметим, что наиболее вероятной буквой после Е является Н, биграмма ЕН встречается в русских осмысленных текстах примерно в два раза чаще, чем следующие *по* частоте биграммы с буквой Е: ЕР, ЕC, ЕЛ . С учетом удвоения буквы тем более вероятно, что символу Я шифрованного текста соответствует буква Н открытого текста.

Конечно, исследуемый текст может оказаться особенным, исключительным. И наша *гипотеза* может оказаться неверной. В таком случае мы придем к противоречию, к неосмысленному тексту, и нам придется вернуться к началу задачи и проверять другие гипотезы. Однако мы начнем с проверки наиболее вероятной гипотезы: Н\rightarrowЯ. На языке сравнений наша *гипотеза* записывается так: 12a+b = 30(\mod 31).

Вычитая из второго сравнения первое, получаем: 7a=18(\mod 31). Из этого сравнения найдем a. Воспользуемся расширенным алгоритмом Евклида. Имеем:

31=4\cdot 7+3,\qquad 7=2\cdot 3+1.

Из последнего равенства: 1=7-2\cdot 3, из первого равенства 3=31-4\cdot 7, и 1=7-2\cdot (31-4\cdot 7)=-2\cdot 31+9\cdot 7. Следовательно, 7^{-1}=9 (\mod 31), и a = 9\cdot 18 = 7 (\mod 31). Теперь b=12-5a=12-5\cdot 7=-23=8(\mod 31).

Мы нашли *ключ* шифрования. Чтобы расшифровать сообщение, зашифрованное применением отображения C=aP+b (\mod 31), нужно выразить P через C: P=sC+t(\mod 31), где s=a^{-1}(\mod 31) есть обратное к a *по* модулю 31 число, а t=-a^{-1}b. Так как число 31 простое, то s и t находятся также с помощью расширенного алгоритма Евклида: s=9, t=21.

Теперь мы можем проводить *расшифрование*. Букве И соответствует номер 8, и 9\cdot 8+21=93=0 (\mod 31), так что букве И в шифртексте соответствует буква А в открытом тексте. Подведем итог предыдущим рассуждениям. Символы шифртекста запишем строчными буквами, а уже найденные буквы открытого текста – прописными. Мы имеем:

**вНтжхшАавб-аЕННжьгм, ЕЕ НАеж ыАцвцАгм**.

Продолжим вычисления. Буква е имеет номер 5, и 9\cdot 5+21=66=4(\mod 31), а 4 - это номер буквы Д. Буква ж имеет номер 6, 9\cdot 6+21=75=13(\mod 31), а это номер буквы О. И мы знаем третье *слово* полностью: НАДО. Вычисляем букву открытого текста, зашифрованную символом ц: это Щ. Запишем результаты:

**вНтОхшАавб-аЕННОьгм, ЕЕ НАДО ыАЩвЩАгм**.

Тактически правильно будет найти буквы открытого текста, зашифрованные символами г и м. Вычисления дают: г\rightarrowТ,м\rightarrowЬ. И мы можем "проявить" еще часть текста:

**вНтОхшАавб-аЕННОьТЬ, ЕЕ НАДО ыАЩвЩАТЬ**.

Теперь уже последнее *слово* легко и угадать. Но все же вычислим еще две буквы. Действительно, ы\rightarrowЗ,в\rightarrowИ, последнее *слово* ЗАЩИЩАТЬ. Кроме того, мы узнали еще часть первого слова. Итог этого этапа:

**ИНтОхшАаИб-аЕННОьТЬ, ЕЕ НАДО ЗАЩИЩАТЬ**.

Далее, а\rightarrowЦ,ь\rightarrowС, что позволяет записать второе *слово*: ЦЕННОСТЬ. Легко вычислить и остальные буквы.

Ответ: **ИНФОРМАЦИЯ-ЦЕННОСТЬ, ЕЕ НАДО ЗАЩИЩАТЬ**.

**6.7 Преобразования биграмм**

Предположим теперь, что элементами открытого и шифрованного текстов являются двухбуквенные блоки, называемые *биграммами*. Это значит, что *открытый текст* разбивается на двухбуквенные *сегменты*. Если *открытый текст* состоит из нечётного числа букв, то, чтобы получить *целое число* биграмм, добавим к концу текста ещё одну букву, выбрав её так, чтобы не исказить смысл, например, добавим *пробел*, если он содержится в нашем алфавите.

Каждой биграмме приписывается далее её числовой эквивалент. Простейший способ - взять его в виде xN+y, где x-числовой эквивалент первой буквы биграммы, y-числовой эквивалент второй буквы биграммы, а N-число букв в алфавите. Таким образом, мы рассматриваем биграмму, как *запись* двузначного числа в системе счисления с основанием N. Это даёт взаимно однозначное соответствие между множеством всех биграмм в $N$-буквенном алфавите и множеством всех неотрицательных целых, меньших {N}^{2}.

Следующий шаг - выбор шифрующего преобразования, т. е. перестановки целых чисел \left\{0,1,2,{\dots},{N}^{2}-1\right\}. Примером простейших шифрующих преобразований служат *аффинные* преобразования: при шифровании P переходит в неотрицательное *целое число*, меньшее {N}^{2} и удовлетворяющее сравнению C\equiv aP+b(mod N^2). Здесь, как и раньше, число a должно не иметь общих множителей с N (что означает отсутствие общих множителей и с N^2) для того, чтобы существовало *обратное преобразование*.

**Пример 6.9** Пусть мы имеем дело с 33-буквенным алфавитом и используем биграммное шифрующее преобразование C\equiv 157 P + 580 (\mod  1089). Тогда биграмма "НО" имеет числовой эквивалент 14\cdot 33+15=477, и ей соответствует биграмма шифртекста 157\cdot 477+580=75469(\mod  1089) = 328= 9\cdot 33+31, то есть "ИЮ". Биграмма "ОН" имеет числовой эквивалент 15\cdot 33 + 14 = 509, и ей соответствует биграмма шифртекста 157\cdot 509 + 580 = 80493 = 996 = (\mod 1089)=30\cdot 33 + 6, то есть "ЭЁ". Заметим, что биграмма преобразуется как единое целое, и между зашифрованными биграммами, имеющими общую букву или даже составленными из одних и тех же букв, но в разном порядке, нет явной связи.

Для вскрытия биграммной системы, использующей аффинное преобразование C\equiv aP + b ~(\mod  N^2), надо знать элементы шифртекста, соответствующие двум разным элементам открытого текста. Так как элементами текста являются биграммы, то частотный *анализ* означает выделение в длинном отрезке шифртекста двухбуквенных блоков, встречающихся чаще других (разумеется, надо считать только элементы текста и игнорировать пары, образованные буквами соседних элементов), и сравнение с известными частотами биграмм в английских текстах, записанных в том же алфавите. Например, если используется 26-буквенный *алфавит*, то *статистический анализ*, скорее всего, покажет, что наиболее частыми биграммами являются "T H" и "H E" (в указанном порядке). Информации о двух парах отвечающих друг другу биграмм открытого и шифрованного текстов зачастую (но не всегда) бывает достаточно для определения a и b.

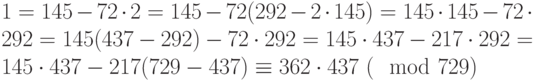
**Пример 6.10** *Вы знаете, что ваш противник использует криптосистему с 27-буквенным алфавитом, в котором буквы A-Z имеют числовые эквиваленты 0-25, и пробел=26. Каждой биграмме отвечает числовой эквивалент - целое число между 0 и 728= {27}^{2}-1, определяемое по правилу 27x+y - числовые эквиваленты букв биграммы. Пусть анализ длинного шифртекста показал, что чаще всего в нём встречаются биграммы "Z A", "I A" и "I W" (в указанном порядке). Предположим, что самыми частыми биграммами в английском языке (в текстах в нашем 27-буквенном алфавите) являются биграммы "E\_" (т. е. E и пробел), "S\_" и "\_T". Мы знаем, что криптосистема использует аффинное шифрующее преобразование по модулю 729. Найти ключ расшифрования и прочитать сообщение "NDXBHO". Найти также ключ шифрования.*

**Решение.** Мы знаем, что *открытый текст* шифруется *по* правилу C\equiv aP+b ~(\mod  729), и что *шифртекст* может быть расшифрован *по* правилу P\equiv sC+t ~(\mod  729), где a, b образуют *ключ* шифрования, а s, t - *ключ* расшифрования. Сначала мы хотим найти s и t. Нам известно, как расшифровываются три биграммы. Заменив эти биграммы их числовыми эквивалентами, получим три сравнения

|  |
| --- |
| 675s+t\equiv 134 ~(\mod  729), |
| 216s+t\equiv 512 ~(\mod  729), |
| 238s+t\equiv 721 ~(\mod  729). |

Если исключить t, взяв *разность* первых двух сравнений, то получим 459s \equiv 351 ~(\mod  729), что не даёт единственного решения s *по* модулю 729 (имеется 27 решений). Будет лучше, если вычесть из первого сравнения третье, что даст 437s\equiv 142 ~(\mod  729). Для решения последнего сравнения надо найти число, обратное к 437 *по* модулю 729. Воспользуемся алгоритмом Евклида:

729=437+292,\quad 437=292+145,\quad 292=2\cdot 145+2,\quad 145=72\cdot 2+1,

и затем .

Итак, s\equiv 362\cdot 142\equiv 374 ~(\mod  729), и, следовательно, t\equiv 134 - 675\cdot 374\equiv 647 ~(\mod  729). Теперь применим преобразование дешифрования к биграммам "N D", "X B" и "H O" нашего сообщения (им соответствуют целые числа 354, 622 и 203 соответственно) и получим числа 365, 724 и 24. Записав 365=13\cdot 27+14, 724=26\cdot 27+22, 24=0\cdot 27+24, мы соединим биграммы открытого текста в сообщение "NO WAY". Наконец, для нахождения ключа шифрования мы вычисляем a\equiv  {s}^{-1}\equiv {374}^{-1}{\equiv} 614 ~(\mod  729) (снова используя *алгоритм* Евклида) и b\equiv -{s}^{-1} t \equiv  -614\cdot 647\equiv 47 ~(\mod  729).

**6.7.1 Шифрующие матрицы**

Пусть имеется N-буквенный алфавит, и блоками сообщения являются биграммы. Ранее мы каждой биграмме сопоставляли целое число по модулю N^2, т.е. элемент кольца Z_{N^2}. Другой подход заключается в том, чтобы сопоставлять биграмме вектор \left(\begin{array}{l}x\\y\end{array}\right) с целыми числами x и y, рассматриваемыми по модулю N. Например, если используется русский алфавит из 33 букв, то биграмме ОН соответствует вектор \left(\begin{array}{l}15\\14\end{array}\right)(нумерация начинается с нуля, так что буква А имеет номер 0, буква Я имеет номер 32).

Для коммутативного кольца K обозначим через M_2(K) множество всех квадратных матриц порядка 2 с операциями сложения и умножения матриц. Как нетрудно видеть, мы получаем кольцо матриц над K.

**Теорема 6.3** *Пусть A\in M_2(Z_N), D - определитель матрицы A. Следующие условия равносильны:*

* *\GCD(D,N)=1,*
* *существует матрица, обратная к A,*
* *если хотя бы один из элементов x,y\in Z_N отличен от нуля, то A\cdot \left(\begin{array}{l}x\\y\end{array}\right)\neq \left(\begin{array}{l}0\\0\end{array}\right),*
* *матрица A задаёт взаимно однозначное отображение множества Z_N^2 на себя.*

На основании сформулированного утверждения мы можем выполнять шифрующее преобразование, используя матрицы из M_2(Z_N), у которых определители взаимно просты с N.

Элемент открытого текста P=\left(\begin{array}{l}x\\y\end{array}\right) преобразуется в элемент шифртекста C=\left(\begin{array}{l}u\\v\end{array}\right) по правилу C=A\cdot P, т.е.

\left(\begin{array}{l}u\\v\end{array}\right)=\left(\begin{array}{ll} a & b\\c & d \end{array}\right) \cdot\left(\begin{array}{l}x\\y\end{array}\right).

Расшифрование производится с помощью обратной матрицы: P=A^{-1} C.

**Пример 6.11** *Зашифруем открытый текст ШИФРОВАНИЕ с помощью матрицы A=\left(\begin{array}{ll}2&3\\4&7\end{array}\right).*

Биграмме ШИ соответствует вектор \left(\begin{array}{l}25\\9\end{array}\right), биграмме ФР - вектор \left(\begin{array}{l}21\\17\end{array}\right), и так далее. Слову ШИФРОВАНИЕ соответствует последовательность векторов \left(\begin{array}{lllll}25&21&15&0&9\\9&17&2&14&5\end{array}\right). Умножаем их на шифрующую матрицу, выполняя вычисления по модулю 33:

\left(\begin{array}{ll}2&3\\4&7\end{array}\right)\left(\begin{array}{lllll}25&21&15&0&9\\9&17&2&14&5\end{array}\right) = \left(\begin{array}{lllll}11&27&3&9&0\\31&5&8&32&5\end{array}\right).

Теперь записываем биграммы, соответствующие полученным векторам: \left(\begin{array}{l}11\\31\end{array}\right) соответствует КЮ, \left(\begin{array}{l}27\\5\end{array}\right) - ЪЕ, и так далее. Получаем шифртекст: КЮЪЕГЗИЯАЕ.

Шифрование триграмм проводится аналогично, триграмме сопоставляется вектор из трех компонент, а шифрующая матрица имеет порядок 3.

**6.8 Комбинированный (композиционный) шифр**

Наряду с шифрами перестановки и шифрами замены применяются и шифры, являющиеся комбинацией этих преобразований. Мы представим два варианта комбинированного шифра: поточный *шифр* и *шифр* с переменной длиной блока.

**Первый алгоритм**

Открытый текст, сохраняя пробелы между словами, записывается в таблицу. Начало - в первой строке, текст записывается слева направо, переходя со строки на следующую строку, после чего осуществляется перестановка столбцов. Теперь к каждому столбцу применяется простая замена: каждая буква первого столбца сдвигается циклически вправо на n_1 позиций в алфавите, буква второго столбца - на n_2 позиций, третьего - на n_3позиций, и так дале. Такая замена одну и ту же букву, стоящую в разных столбцах, переводит в разные буквы.

Описанный алгоритм является комбинацией перестановки и замены.

**Пример 6.12** *Открытый текст записан в таблицу, после чего столбцы переставили в соответствии с ключом зашифрования (5,1,3,2,4). На втором этапе буквы первого столбца сдвинули на 1 позицию в алфавите (а\rightarrow б, б\rightarrow в, ...), второго - на 2 позиции, третьего - на 3 позиции, четвертого - на 4 позиции, пятого - на 5 позиций. Расшифровать полученный таким образом шифртекст:*

**\_впфнубжкйцбфёйебххшгжхсу**

**Решение.**

Запишем шифртекст в таблицу:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| \_ | в | п | ф | н |
| у | б | ж | к | й |
| ц | б | ф | ё | й |
| е | б | х | х | ш |
| г | ж | х | с | у |

Зная величины сдвигов на втором этапе зашифрования, для расшифровки выполняем сдвиг i-го влево по алфавиту: 1-й столбец - на 1 позицию (а\rightarrow\_, б\rightarrowа, и т.д.), 2-й столбец - на 2 позиции, ... . Получаем:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| я | а | м | р | и |
| т | \_ | д | ж | е |
| х | \_ | с | в | е |
| д | \_ | т | с | у |
| в | е | т | н | о |

Зная ключ первого этапа зашифрования, осуществляем окончательное расшифрование, применяя обратную подстановку столбцов: (5,1,3,2,4)^{-1} = (2,4,3,5,1). Это даёт нам открытый текст:

**армия\\_ждет\\_всех\\_студентов**.

**6.8.1 Второй алгоритм: комбинированный шифр с переменным размером блока**

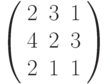
Блок текста обладает своей структурой, в большей или меньшей степени наследует особенности текстов на естественном языке. Поэтому для рассеяния таковых особенностей по шифртексту и для повышения качества итогового криптопреобразования приходится либо усложнять раундовые операции с блоками либо увеличивать количество раундов.

Можно, однако, для получения хороших показателей диффузии применить и несколько иную идею: работать с блоками переменной длины, менять размер блока от раунда к раунду. Эта идея реализована, например, в [5]. Кратко опишем схему шифрования.

Выбирается достаточно большое составное число N. В каждом раунде выполняются преобразования блоков длины k_i, числа k_i -делители N. Чередование уменьшения и увеличения длины блока в комбинации с перестановками позволяет добиться сокрытия статистических закономерностей текста за небольшое количество раундов. Здесь мы проиллюстрируем общую схему учебным алгоритмом шифрования с переменной длиной блока, применив замены и перестановку.

* Первое преобразование. Разбиваем текст на биграммы, шифруем каждую биграмму аффинным преобразованием.
* Второе преобразование. Перестановка букв.
* Третье преобразование. Разбиваем текст на триграммы, шифруя каждую с помощью матрицы, проводя вычисления по модулю 33.

**Пример 6.13** Первое преобразование: C=17\cdot P + 19\ (\mod 33^2), где P - числовой эквивалент шифруемой биграммы. Биграмма АТ имеет эквивалент 0x33+19=19, биграмма ЛА имеет эквивалент 12x33+0=396, биграмма НТ - 14x33+19=481, биграмма ТА - 627. Вычисления дают: Вычисления дают: АТ->17x19+19=342=10x33+12->ЙЛ, ЛА->17x396+19(mod 332) =6751 (mod 332)=217=6x33+19->ЁТ, НТ->17x481+19=8196=573=17x33+12->РЛ, ТА->17x627+19=10678=877=26x33+19->ЩТ и после первого преобразования получили тексты: ЙЛЁТРЛ и ЩТЁТРЛ. Второе преобразование: перестановка (462513), получаем: ТЛЛРЙЁ и ТЛТРЩЁ.

Третье преобразование: разбиваем текст на триграммы и шифруем с помощью матрицы . Триграмме ТЛЛ соответствует вектор , результат зашифрования:  -> УДЬ, триграмме РЙЁ соответствует вектор , результат зашифрования:  -> ДЖР и результатом зашифрования слова АТЛАНТ является шифртекст УДЬДЖР. Проведя вычисления аналогично, получим результат зашифрования слова ТАЛАНТ, это шифртекст ЪШГТЁА.

Итак, АТЛАНТ->УДЬДЖР, ТАЛАНТ->ЪШГТЁА и мы видим, что между результатами зашифрования первого и второго слова нет связи, несмотря на совпадение букв с 3 по 6 в исходных словах. Применение разных типов преобразований к блокам разной длины (в первом преобразовании длина блока 2, во втором 1, в третьем 3) дало хороший результат.

**6.9 Энтропия на знак, избыточность и расстояние единственности**

Мерой среднего количества информации, приходящейся на одну букву открытого текста языка F, который будем рассматривать как источник случайных текстов, служит величина H_F, называемая *энтропией языка*. Интерес представляет *энтропия* вероятностной схемы на r-граммах, деленная на r. Известно, что существует конечный *предел*, который и принимается за *определение* энтропии H_F языка F:

H_F=\underset{r\rightarrow {\infty}}{\lim }{\frac{Hr}{r}}.

При этом формула

R_F=1-\frac{{H}_{\Lambda }}{{\log }_{2}n}

определяет *избыточность* языка F. Как известно, если *длина* криптограммы не большая, то результат расшифрования может дать несколько осмысленных текстов.

Найдём оценку для числа ложных ключей. Для этого рассмотрим *связь* между энтропиями вероятностных распределений P\left(X\right), P\left(K\right), P(Y), заданных на компонентах X, K, Y произвольного шифра {\Sigma }_{B}.

Назовем условную энтропию H(K/Y) неопределенностью шифра \Sigma_B *по* *ключу*. Она измеряет среднее *количество информации* о ключе, которую дает *шифртекст*. Аналогично вводится неопределенность шифра *по* открытому тексту H(X/Y). Эти величины являются мерой теоретической стойкости шифра.

Рассмотрим произвольный поточный *шифр* замены {\Sigma }_{B}, для которого множество X открытых текстов представляет собой множество возможных осмысленных текстов в данном алфавите A (например, русском, английском или некотором другом), состоящим из n букв. Зафиксируем некоторое число L{\in}N и будем интересоваться числом ложных ключей, отвечающих данной криптограмме y{\in}{A}^{L}. Предполагается, что Aслужит также алфавитом шифрованного текста. Введем обозначение:

K\left(y\right)=\left\{k{\in}K\right|{\exists}x{\in}X,{E}_{k}\left(x\right)=y\}

K(y) есть множество ключей, для каждого из которых y является результатом шифрования некоторого осмысленного текста длины L. Если мы располагаем криптограммой y, то число ложных ключей равно \left|K\left(y\right)\right|=1, так как лишь один из допустимых ключей является истинным. Определим среднее число ложных ключей {\kappa }_{L} (относительно всех возможных шифртекстов длины L) формулой:

{\kappa }_{L}=\sum _{y{\in}Y}{p\left(y\right){\cdot}\left(\left|K\left(y\right)\right|-1\right),}

которая легко приводится к виду

{\kappa }_{L}=\sum _{y{\in}Y}{p\left(y\right)\left|K\left(y\right)\right|-1.}

**Теорема 6.4** *Для любого рассматриваемого шифра {\Sigma }_{B} с равновероятными ключами при достаточно больших значениях L имеет место неравенство*

{\kappa }_{L}{\geq}\frac{\left|K\right|}{{n}^{L{\cdot}{R}_{\Lambda }}}-1,

*где {R}_{\Lambda } - избыточность данного языка.*

Назовем расстоянием единственности для шифра {\Sigma }_{B} *натуральное число* (обозначим его {L}_{0}), для которого ожидаемое число ложных ключей \kappa _{\Lambda } равно нулю. *По* сути, *расстояние* единственности есть средняя *длина* шифртекста, необходимая для однозначного восстановления истинного ключа (без каких-либо ограничений на время его нахождения).

{L}_{0}=\left\lceil\frac{\log_2 |K|}{R_\Lambda\cdot \log_2 n}\right\rceil

Большинство криптосистем слишком сложны для точного определения расстояния единственности, однако в некоторых случаях оно может быть аппроксимировано соотношением L_0 = H(k)/R_F.

Оценим *расстояние* единственности шифра Виженера со случайным ключевым словом длиной 4 символа для сообщений на английском языке.

*Ключ* шифра - это набор из четырех чисел, каждое из которых является случайным сдвигом в диапазоне от 0 до 25. *Энтропия* ключа H(k) = log_2\ 26^4 \approx 18,8. Для языка из 26 символов R_0 = log_2 \ 26 \approx 4.7. Для длинных сообщений на английском языке оценка H_r/r составляет от 1 до 1,5. Взяв *значение* H_F = 1,5, можно вычислить L_0\approx 18,8 / 3,2 \approx 5,9. Таким образом, однозначное *дешифрование* шифра Виженера с 4-х символьным ключом возможно при длине шифртекста не менее 6 символов.

**Пример 6.14** ([6]) *Пусть источник порождает буквы из алфавита F=a,b,c с вероятностями P(a)=0,8, P(b)=0,15, P(c)=0,05. Пусть у нас источник без памяти. Шифрование состоит в замене буквы в исходном сообщении, используя перестановку символов в соответствии с ключом: k=1 : (abc),   k=2 : (acb),  k=3 : (bac),  k=4 : (bca),  k=5 : (cab), k=6 : (cba), то есть ключ принимает значения от 1 до 6 и если, например, k=5, то производится следующая замена символов текста: a\rightarrow c$, $b\rightarrow a$, $c\rightarrow b. Пусть противник перехватил шифрованное сообщение cccbc и пытается определить значение ключа. Оценим вероятности использования всех возможных ключей.*

Используем формулу Байеса:

P\left({K}_{i}\left|E\right.\right)=\frac{P\left({K}_{i}\right)P(E| {K}_{i})}{\sum _{j=1}^{t}{P\left({K}_{j}\right)P(E| {K}_{j})}},

где E, {K}_{1},{\dots}, {K}_{t} - некоторые события, {K}_{i} попарно несовместны и E{\subset}\bigcup\limits_{i=1}^{t}{{K}_{i}}. В нашем случае событие E - это получение шифрованного сообщения cccbc, t=6, а K_i означает, что выбран *ключ* k=i.

Мы предполагаем, что все ключи равновероятны, т.е.

P\left({K}_{1}\right)=P\left({K}_{2}\right)=P\left({K}_{3}\right)=P\left({K}_{4}\right)=P\left({K}_{5}\right)=P\left({K}_{6}\right)=1/6.

Тогда

|  |
| --- |
| P\left(E\left|{K}_{1}\right.\right)=P\left(\overline{{x}}=\mathit{cccbc}\right)={0.05}^{4} {\cdot}0.15{\approx}0.000001, |
| P\left(E\left|{K}_{2}\right.\right)=P\left(\overline{{x}}=\mathit{bbbcb}\right)={0.15}^{4}{\cdot}0.05{\approx}0.000025, |
| P\left(E\left|{K}_{3}\right.\right)=P\left(\overline{{x}}=\mathit{cccac}\right)=0.8{\cdot}{0.05}^{4}{\approx}0.000005, |
| P\left(E\left|{K}_{4}\right.\right)=P\left(\overline{{x}}=\mathit{bbbab}\right)=0.8{\cdot}{0.15}^{4}{\approx}0.000405, |
| P\left(E\left|{K}_{5}\right.\right)=P\left(\overline{{x}}=\mathit{aaaca}\right)={0.8}^{4}{\cdot}0.05=0.020480, |
| P\left(E\left|{K}_{6}\right.\right)=P\left(\overline{{x}}=\mathit{aaaba}\right)={0.8}^{4}{\cdot}0.15=0.061440. |

Отсюда легко находим



и получаем *по* формуле Байеса апостериорную *вероятность* того, что был использован *ключ* k = 1, при условии, что получено сообщение \overline{{y}}=\mathit{cccbc}:

P\left({K}_{1}\left|E\right.\right)=P\left(\overline{{x}}=\mathit{cccbc}\left|\overline{{y}}=\mathit{cccbc}\right.\right){\approx}\frac{(1/6){\cdot}0.000001}{0.13726}{\approx}0.000011.

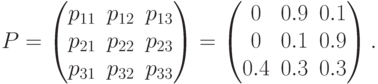
Продолжая аналогично, находим наиболее вероятные ключи для k = 5 и k = 6:

|  |
| --- |
| P\left({K}_{5}\left|E\right.\right)=P\left(\overline{{x}}=\mathit{aaaca}\left|\overline{{y}}=\mathit{cccbc}\right.\right){\approx}0.25, |
| P\left({K}_{6}\left|E\right.\right)=P\left(\overline{{x}}=\mathit{aaaba}\left|\overline{{y}}=\mathit{cccbc}\right.\right){\approx}0.75, |

а вероятности всех остальных ключей меньше 0.01.

Мы видим, что, перехватив всего пять букв, противник может определить *ключ* почти однозначно. Если *избыточность* сообщения равна нулю, то *ключ* никогда не будет определен. Уменьшение избыточности может быть достигнуто за счет сжатия данных. Это объясняется тем, что при сжатии *энтропия* "сжатого" текста сохраняется, а *длина* уменьшается. Следовательно, *энтропия* на букву в сжатом тексте больше, чем в исходном, а *избыточность* меньше. Таким образом, после сжимающего кодирования *расстояние* единственности шифра увеличивается. Поясним примером , как взаимная зависимость символов увеличивает *избыточность* и тем самым уменьшает *расстояние* единственности.

**Пример 6.15** *Пусть имеется марковский источник сообщений, вероятность p_ij появления j-го символа после i-го задаётся матрицей:*



*Заданы начальные вероятности: P(a)=0.19, P(b)=0.34, P(c)=0.47. Возможные ключи: {k}_{1}=\left(\begin{matrix}a&b&c\\a&b&c\end{matrix}\right), {k}_{2}=\left(\begin{matrix}a&b&c\\a&c&b\end{matrix}\right), {k}_{3}=\left(\begin{matrix}a&b&c\\b&a&c\end{matrix}\right), {k}_{4}=\left(\begin{matrix}a&b&c\\c&a&b\end{matrix}\right), {k}_{5}=\left(\begin{matrix}a&b&c\\b&c&a\end{matrix}\right), {k}_{6}=\left(\begin{matrix}a&b&c\\c&b&a\end{matrix}\right).*

*Пусть перехвачен шифртекст: bbacbac. Какой ключ использовался при шифровании?*

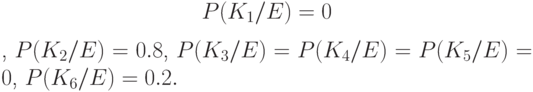
Зная матрицу переходов, мы можем сделать *вывод*: сочетание aa невозможно (после буквы a *вероятность* появления снова буквы a равна нулю), *вероятность* появления b после b равна 0,1 , то есть сочетание bb маловероятно. Поэтому наиболее вероятная первая пара букв открытого текста cc, у нас есть *гипотеза*: при зашифровании была проведена замена: c\rightarrow b. Тогда другие замены возможные: b\rightarrow a и a\rightarrow c либо a\rightarrow a и b\rightarrow c. В первом варианте паре букв шифртекста ac соответствует пара букв открытого текста ba. Но сочетание ba невозможно, как видно из матрицы переходов, а сочетание ab возможно. Следовательно,наиболее вероятным ключом является *подстановка*: {k}_{2}=\left(\begin{matrix}a&b&c\\a&c&b\end{matrix}\right). Для проверки нашей гипотезы вычислим вероятности использования различных ключей. Заметим, что *вероятность*сообщения равна произведению вероятности начальной буквы и вероятностей переходов от одной буквы к другой.

|  |
| --- |
| P(E/K_{1})=P(m=bbncbnc)=0.34\cdot0.1\cdot0=0, |
| P(E/K_{2})=P(m=ccnbcnb)=0.47\cdot0.3\cdot0.4\cdot0.9\cdot0.9\cdot0.4\cdot0.9=0.016446, |
| P(E/K_{3})=P(m=nnbcnbc)=0.19\cdot0=0, |
| P(E/K_{4})=P(m=nncbncb)=0.19\cdot0=0, |
| P(E/K_{5})=P(m=ccbncbn)=0.47\cdot0.3\cdot0.3\cdot0=0, |
| P(E/K_{6})=P(m=bbcnbcn)=0.34\cdot0.1\cdot0.9\cdot0.4\cdot0.9\cdot0.9\cdot0.4=0.003966. |

Применяем формулу полной вероятности:



Тепрерь *по* формуле Байеса апостериорные вероятности использования ключей при условии получения шифртекста bbacbac следующие:



Вычисления подтверждают нашу гипотезу.

**6.10 Разные задачи**

Как отмечалось ранее, при решении задач криптоанализа часто бывает необходимо учитывать особенности языка, на котором составлен *открытый текст*. Полезным является и *знание* способа перевода буквенного открытого текста в цифровую форму перед шифрованием. Следующие задачи, на наш взгляд, помогут развитию криптоаналитической интуиции.

**6.10.1 Задачи, которые условно можно назвать лингвистическими**

**Пример 6.16** *(Составлена по образцу задач 11.3, 12.2, 16.3 из* [2) *Используется русский алфавит, Е и Ё не различаются. Буквы алфавита кодируются следующим образом: каждой букве ставится в соответствие двоичная запись её номера, начиная с нуля. Таким образом, А соответствует 00000, Б соответствует 00001, В соответствует 00010, \dots, Ю \rightarrow 11110, Я \rightarrow 11111. Для передачи используется пять проводов. По каждому передается соответствующий разряд пятизначного двоичного числа. При монтаже 5 проводов были перепутаны, и теперь мы не знаем, по какому проводу приходит цифра старшего разряда и т. д. Перед вами полученный текст. Известно, что исходный текст был осмысленным. Восстановите его. Полученный текст: ЫАДАФЭС*

**Решение.** Перепутывание проводов не могло изменить количество единиц в двоичной записи буквы. Поэтому вторая и четвертая буквы исходного текста находятся сразу, это буквы А, так как А кодируется как 00000. Первая буква полученного текста Ы. Её код 11011, количество единиц четыре. Запишем буквы с четырьмя единицами в двоичной записи: П, Ч, Э, Ю. Варианты начала слова: ЫА, ПА, ЧА, ЭА, ЮА. Третья буква полученного текста - Д. Её двоичная запись 00100, найдём буквы, имеющие, как и Д, одну единицу в двоичной записи. Это буквы: Б, В, И, Р. Вариант начала ЫА отбросим как невозможный, имеем следующие варианты для первых четырех букв:


    \text{ПА}\left(\begin{array}{c}\text{Б}\\\text{В}\\\text{И}\\\text{Р}\end{array}\right)\text{A}, \quad
    \text{ЧА}\left(\begin{array}{c}\text{Б}\\\text{В}\\\text{И}\\\text{Р}\end{array}\right)\text{A}, \quad
    \text{ЭА}\left(\begin{array}{c}\text{Б}\\\text{В}\\\text{И}\\\text{Р}\end{array}\right)\text{A}, \quad
    \text{ЮА}\left(\begin{array}{c}\text{Б}\\\text{В}\\\text{И}\\\text{Р}\end{array}\right)\text{A}.
    

Шестая буква полученного текста Э, её код 11101, и есть следующие буквы, коды которых содержат те же четыре единицы: П, Ч, Ы, Ю. Учёт осмысленности исходного текста вариантов для 5 и 7 букв позволяют прочитать: ПАРАШЮТ.

**Пример 6.17** *Ниже записаны в зашифрованном виде: слева английские слова, справа - их переводы (иногда несколько переводов, но не обязательно все) на русский язык. Английские слова зашифрованы таким образом, что одной и той же букве латинского алфавита соответствует одно и то же число, записанное арабскими цифрами, разным буквам - разные числа. Аналогичным образом, но совершенно независимо, русские слова зашифрованы при помощи последовательности чисел, записанных римскими цифрами.*

*Буквы одного и того же слова отделены друг от друга запятыми, слова в заданиях - точками с запятой, фразы в заданиях - точками.*

*Надстрочные знаки и апострофы не учитываются, но* **и** и **й** *считаются разными буквами, а* **е** *и* **ё** - *не различаются.*

|  |
| --- |
| 1 - I |
| 2, 3 - II, III, IV или II, III, I или II, III, V |
| 4, 5, 6 - V, V |
| 7, 8, 9 - VI или VII |
| 8,10 - VIII, V, IX |
| 3, 10, 11 - IX, X или XI, X |
| 7, 12 - XII |
| 2, 10, 12, 4, 5, 6 - II, VII, IX, XIII |
| 13, 6, 1, 5, 8, 9, 14 - XIV, XV, XII, XVI, XIII, I |

*Переведите следующие зашифрованные русские слова и фразы в английские:*

1. *XI, III, IX; XIV, XI, V, XV, XIII; XIV, III, II, VII;*
2. *VI, XVI; II, III, XV, I;*
3. *XIV, V, XV, V, XI, XIII, I;*
4. *II, V, VIII, I; XVI, III, XI, XII, IX; IX, III, II.*

**Решение.** Единственное английское слово, состоящее из одной буквы, которое переводится русским словом тоже из одной буквы, - это "я" (1=I; I=я). Единственное русское слово из двух идентичных букв - это ее (её) (4=h, 5=e, 6=r; V=е). Единственное английское слово из трёх букв, переводящееся одним из двух русских слов, каждое из которых состоит из одной буквы, - это and (7=a, 8=n, 9=d; VI=(а или и), VII=(и или а)). Последнее английское слово: 13, r, i, e, n, d, 14, переведённое как XIV, XV, XII, XVI, XIII, я, должно быть friends "друзья" (13=f, 14=s; XII=у, XIII=ь, XIV=д, XV=р, XVI=з). Слово a, 12, переведённое как у, должно быть тогда at (12=t). Слово 2, 10, t, h, e, r, переведённое как II, и/а, IX, ь, - это мать (2=m, 10=o; II=м, VI=и, VII=а, IX=т).

Слова м, III, IV; м, III, я; м, III, е являются переводами одного и того же слова m, 3, которое, очевидно, является словом my (3=y; III=о, IV=й). И, наконец, 8, 10=no, а 3, 10, 11=you (11=u; VIII=н, X=ы, XI=в). Русские слова и фразы, которые нужно перевести, - это:

1. Вот дверь дома
2. из моря
3. деревья
4. Меня зовут Том.

**6.10.2 Разные задачи**

**Пример 6.18** ([6]) *Алфавит состоит из n букв. Какова максимальная длина слова, если выполнены два условия: а) в нем две рядом стоящие буквы всегда различны; б) из него нельзя получить вычеркиванием букв слова вида abab, где a и b - различные буквы?*

**Решение.** Если буква встречается только один раз, назовем ее буквой первого рода. В противном случае - буквой второго рода. Буквы, стоящие рядом с буквой второго рода, различны, это следует из свойства "б".

Если слово содержит хотя бы два вида букв, то оно содержит хотя бы одну букву первого рода. В противном случае мы можем получить из него слово вида abab, что противоречит условию. Вычеркнув все буквы второго рода, совпадающие с некоторой, получим слово из n - 1 различных букв. Вычеркивая буквы второго рода и далее, придем к слову, состоящему только из букв первого рода.

Докажем методом математической индукции, что слово имеет длину не более 2n - 1 буквы. Для n = 1 утверждение верно. Предположим теперь, что оно верно для n = k букв и докажем его для n = k + 1.

Пусть слово содержит k + 1 различную букву, \alpha - буква первого рода, а \beta - соседняя с ней буква. Если \beta - первого рода, то при вычеркивании \beta возникает слово из k различных букв, максимальная длина которого не превышает 2k-1 букв (по предположению индукции), а само исходное слово имеет длину не более 2k. Если \beta - второго рода, то обе соседние буквы пары \alpha, \beta или различны, или имеется только одна соседняя буква, т. е. пара \alpha, \beta стоит с краю. Поэтому пару \alpha, \beta можно вычеркнуть, оставшееся слово удовлетворяет условию "а" задачи. Длина оставшегося слова 2k-1 букв, а первоначального - 2k-1+2 = 2(k + 1) - 1.

Утверждение доказано. Итак, слово имеет длину не более 2n - 1 букв.

Приведем пример слова, имеющего длину в точности 2n - 1 букв. Из алфавита a_1,a_2,\ldots, a_n можно составить слово a_1a_2\ldots a_na_{n-1}a_{n-2}\ldots a_1.

**Пример 6.19** ([7]) *Даны пять правил подстановки:*

1. *D \rightarrow AB;*
2. *A \rightarrow CD;*
3. *B \rightarrow b;*
4. *C \rightarrow c;*
5. *D \rightarrow d*

*Правила означают следующее: если в какой-либо последовательности букв можно отыскать левую часть какого-либо из пяти правил, то её разрешается заменить правой частью этого правила и переписать последовательность в новом виде, не меняя остальных букв. После этого можно опять постараться применить какое-либо правило и т. д.*

|  |
| --- |
| *а) Какие из следующих последовательностей можно получить из буквы D применением какого-либо числа раз правил 1-5: bb, bbc, bbcd, bbcdd, bd, cbd, bdc, b, cdb, cb, cc, ccd, ccdbb, ccccd?* |
| *б) Охарактеризуйте все последовательности строчных букв, которые могут получиться из буквы D путём применения правил 1-5 и к которым уже нельзя применить ни одного из этих правил.* |

**Решение.** Если применить каждое из пяти правил по одному разу, то получится последовательность *cdb*. Но вместо того, чтобы использовать правило 5, мы всегда можем снова использовать правило 1, и тогда вместо *d* появится опять *cdb*. Это означает, что мы можем получать только следующие циклические последовательности: *cdb, ccdbb, cccdbbb* и т. д.; каждая такая последовательность описывается следующим образом: c может повторяться сколько угодно раз, потом идёт одно *d* и столько же раз *b*, сколько было *c*.

**Пример 6.20** *Будем рассматривать последовательности, состоящие только из букв A и B (например, AABABB, AA, B и т. п.). Разрешается преобразовывать каждую из последовательностей следующим образом:*

* *если в последовательности есть группа BA (подряд и именно в этом порядке), то её можно заменить на ABBB;*
* *ABBB можно заменить (при тех же условиях) на BA;*
* *можно вычеркнуть подряд идущую группу AA или BBBB;*
* *между любыми двумя стоящими рядом буквами последовательности, или левее всех букв, или правее всех букв можно написать группу AA или BBBB.*

*К каждой последовательности можно применить любое из этих преобразований, к полученной последовательности - снова любое из этих преобразований и т. д.*

|  |
| --- |
| *а) Какую цепочку преобразований нужно применить к последовательности AB, чтобы получить из неё последовательность BBBA (известно, что это можно сделать)?* |
| *б) Какую цепочку преобразований нужно применить к последовательностиAB, чтобы получить последовательностьABABAB?* |
| *в)\textbf{ }Можно ли путём указанных преобразований из последовательности AB получить последовательность ABAB?* |
| *г)\textbf{ }Докажите, что нельзя путём разрешённых преобразований получить из последовательности AB последовательность BA.* |

**Решение.**

|  |
| --- |
| а) AB \rightarrow BBBBAB \rightarrow BBBABBBB \rightarrow BBBA (сначала слева приписана последовательность BBBB, затем образовавшийся фрагмент BA заменён на ABBB, затем фрагмент ВВВВ в конце образовавшейся последовательности вычеркнут); |
| б) AB \rightarrow ABAA \rightarrow ABAABBBB \rightarrow ABABAB (справа приписана последовательность AA, затем справа приписана последовательность BBBB, затем фрагмент ABBB заменён на BA); |
| в) Нельзя, так как преобразования сохраняют чётность количества A (а также В) в группе букв; |
| г) Рассмотрим пару произвольных чисел (x; y), причём x \neq 0 и y \neq  0. |

Придадим буквам A и B смысл операций с парами чисел: A - "заменить знак первого числа на противоположный, второе число оставить без изменений", т. е. (x; y) A \rightarrow (-x; y); B - "поменять числа местами, после чего у числа, оказавшегося на втором месте, поменять знак на противоположный", т. е. (x; y) B \rightarrow  (y;-x).

Последовательность из букв A и B будем рассматривать как последовательность выполнения заданных этими буквами операций (в том порядке, как эти буквы записаны).

Очевидно, что последовательности операций AA и BBBB не меняют исходную пару чисел.

|  |
| --- |
| (x; y) A \rightarrow  (-x; y) A \rightarrow  (x; y) |
| (x; y) B \rightarrow (y; -x) B \rightarrow (-x;-y) B \rightarrow (-y; x) B \rightarrow  (x; y). |

Поэтому их можно добавлять или вычёркивать в любом месте последовательности операций, результат от этого не поменяется. Также легко убедиться, что BA и ABBB всегда приводят к одному и тому же результату, т. е. являются взаимозаменяемыми.

|  |
| --- |
| (x; y) B \rightarrow (y; -x) A \rightarrow  (-y; -x) |
| (x; y) A \rightarrow (-x; y) B \rightarrow  (y; x) B \rightarrow (x; -y) B \rightarrow (-y; -x). |

То есть, выполняя только разрешённые в условии задачи замены в любой последовательности букв A и B, мы всегда будем получать последовательности операций, в итоге приводящих к одному и тому же результату.

А последовательности операций AB и BA приводят к различным результатам.

|  |
| --- |
| (x; y) A \rightarrow (-x; y) B \rightarrow (y; x) |
| (x; y) B \rightarrow (y;-x) A \rightarrow (-y;-x). |

Поэтому с помощью разрешённых в условии замен получить из последовательности AB последовательность BA невозможно.

**Дополнение.** Пару чисел (x; y) можно рассматривать как координаты некоторой точки в декартовой прямоугольной системе координат. В этом случае операция A - это осевая симметрия относительно координатной оси Oy, а операция B - это поворот относительно начала координат на 90^\circ по часовой стрелке.

Тогда A, B порождают группу из 8 элементов: 4 поворотов на углы, кратные прямому, и 4 отражения относительно осей Ox, Oy и биссектрис четвертей, а условия 1)-4) задачи называются порождающими соотношениями этой группы. Каждое из этих преобразований можно ассоциировать с классом эквивалентных слов в алфавите \{A, B\}, где эквивалентными называются слова, получаемые одно из другого применением несколько раз правил 1)-4). В каждом классе можно выбрать представитель. Например, в качестве представителей можно взять: пустое слово (нейтральный элемент группы), B, BB, BBB, AB, ABB, ABBB.

Вопрос 4) можно переформулировать на языке теории групп следующим образом: является ли группа \langle A, B\rangle коммутативной? Или, по-другому, лежит ли BA в классе AB? Видно, что BA лежит в классе ABBB.

**Пример 6.21** (составлен по аналогии с задачей 6.4 из [2]) *На каждой из трех осей установлено по одной вращающейся шестеренке и неподвижной стрелке. Шестеренки соединены последовательно. На первой шестеренке 33 зубца, на второй - 10, на третьей - 7. На каждом зубце первой шестеренки по часовой стрелке написано по одной букве русского языка в алфавитном порядке:*

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

*На зубцах второй и третьей шестеренки в порядке возрастания по часовой стрелке написаны цифры от 0 до 9 и от 0 до 6 соответственно. Когда стрелка первой оси указывает на букву, стрелки двух других осей указывают на цифры. Буквы сообщения шифруются последовательно.*

*Зашифровывание производится вращением первой шестеренки против часовой стрелки до первого попадания шифруемой буквы под стрелку. В этот момент последовательно выписываются цифры, на которые указывают вторая и третья стрелки. В начале шифрования стрелка 1-го колеса указывала на букву А, а стрелки 2-го и 3-го колес - на цифру 0.*

*Расшифруйте сообщение* **948593312255334486**.

**Решение.** Очевидно, что если начать вращение первой шестерёнки ПРОТИВ часовой стрелки, то вторая начнёт вращаться ПО часовой стрелке, а третья ПРОТИВ часовой стрелки. Шифровалась первая буква. При этом цифра 0 на второй шестерёнке заменена на 9. Замена 0 на 9 могла произойти при повороте второй шестерёнки на 1 позицию или на 11 или на 21, 31. На третьей шестерёнке цифра 0 оказалась заменена на 4. Это могло произойти при повороте на 4 позиции или на 11 или 18 или 25 позиций. Из возможных вариантов для сдвига шестерёнок лишь число 11 является общим. Итак, при шифровании первой буквы произошёл сдвиг на 11 позиций на каждой шестерёнке, в том числе на первой. Таким образом, место под стрелкой заняла буква, отстоящая от А на 11 позиций. Сдвиг на одну позицию заменяет А на Б, сдвиг на две позиции заменит А на В, и т. д. Сдвиг на 11 позиций заменит А на К. Следовательно, первая буква открытого текста - это буква К.

Далее, пара цифр 94 при шифровании второй буквы перешла в 85. Цифра 9 на второй шестерёнке переходит в 8 при повороте на 1 позицию, а также на 11, 21, 31. А цифра 4 на третьей шестерёнке переходит в 5 при повороте на 1, 8, 15, 22, 29. Мы видим, что общим для вариантов сдвига на двух шестерёнках является число 1, и только оно. Вывод: при шифровании второй буквы первая шестерёнка вращалась так, что буква К была заменена на отстоящую от неё на ОДНУ ПОЗИЦИЮ, а это буква Л. И мы знаем первые две буквы открытого текста: КЛ, продолжая рассуждать описанным способом, находим (вычисляем!) все буквы.

**Ответ**: КЛЮЧШИФРА.

**6.11 Аутентификация**

Наряду с конфиденциальностью не менее важной задачей является *аутентификация*.

*Аутентификация* - проверка принадлежности субъекту доступа предъявленного им идентификатора; подтверждение подлинности. Иными словами, это процедура определения того, действительно ли *пользователь* или иной субъект передачи информации (приемопередающее оборудование, узел локальной сети и т. д.) *по* каналу связи является тем, за кого он себя выдает.

На сегодняшний день *механизмы* аутентификации широко применяются практически во всех компьютерных системах хранения, передачи и обработки информации, прежде всего, для обеспечения работы систем разграничения доступа и авторизации пользователей системы.

Все *механизмы* аутентификации основаны на предъявлении пользователем системе специальной информации, на основании которой *модуль* аутентификации выносит решение о том, считать ли данного пользователя легитимным или признать, что произошла попытка (возможно, неумышленная) подмены пользователя.

*Базовая* классификация типов систем аутентификации основывается на различии видов информации, используемой в качестве аутентификационной:

1. Последовательность информационных символов, которую пользователь предъявляет системе для успешного прохождения аутентификации. Простейший пример - парольная аутентификация на ПЭВМ, для которой пользователю достаточно ввести в систему свой идентификатор и пароль.
2. Уникальное содержимое или характеристики предмета. В этом качестве могут выступать, как правило, любые внешние носители информации: смарт-карты, электронные таблетки iButton, USB-токены и т. д. Реже для удаленной аутентификации узла сети могут использоваться уникальные данные о его аппаратном обеспечении.
3. Биометрическая информация человека, которая неотъемлема от него и уникальна для каждого индивида. Это может быть отпечаток пальца, рисунок радужной оболочки глаза, форма лица, параметры голоса и т. д.

Используются также комбинированные схемы, основанные на нескольких видах информации, передаваемой системе аутентификации. Типичный пример: аутентификационная *информация* пользователя для входа в программную систему хранится на смарт-карте, для доступа к которой нужно ввести *пароль* (PIN-код). Такая *аутентификация* называется двухфакторной. Намного реже встречаются системы с трехфакторной аутентификацией.

В ряде случаев требуется и взаимная *аутентификация* - когда оба участника информационного обмена проверяют подлинность идентификаторов друг друга. Например, перед передачей удаленному серверу каких-либо данных *пользователь* должен убедиться, что это именно тот *сервер*, который ему необходим.

Помимо этого, существует естественное *деление* механизмов аутентификации на локальную (в рамках локализованной информационной системы) и удаленную (при передаче информации *по* каналу связи).

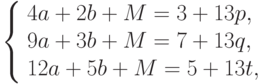
В случае удаленной аутентификации существует проблема передачи аутентификационной информации *по* недоверенным(ненадежным) каналам связи, так как возникает *угроза* перехвата аутентификационной информации злоумышленником(противником) с целью дальнейшего использования для атаки на систему защиты.

Рассмотрим несколько задач, моделирующих в упрощенном виде реальную работу системы парольной аутентификации в средствах защиты информации от несанкционированного доступа.

**Пример 6.22** (см. [8]) **Разделение секрета.** *Сейф с документами открывается, если набрано секретное число M. Для разделения секрета используется многочлен F(x)=ax^2+bx+M, где a, b - целые числа. Каждый из трех членов комиссии получил свою долю секрета. Именно: первый знает, что F(2) делится на 13 с остатком 3, второй знает: F(3) делится на 13 с остатком 7, третий знает: F(5) делится на 13 с остатком 5.*

*Докажите, что объединив свои знания, эти трое смогут открыть сейф (и найдите число M), а любые двое из них не смогут (числа a и b члены комиссии не знают).*

**Решение.** Нужно найти M из условий:



где p, q, t - некоторые целые числа.

Складывая первое *равенство* со вторым, и вычитая из их суммы третье, получим: a+M=5+13p', где p' - некоторое целое. Вычитая первое из второго, получим 5a+b=4+13q'.

Теперь подставим a=5-M+13p', b=4-5a+13q' = 4-5\cdot(5-M+13p') в первое уравнение. Получим M=11+13u, u - *целое число*. Так как каждое число заменяется его остатком от деления на 13, то M=11.

Поскольку a, b, M - неизвестны, то любые двое, объединившись, получат систему из двух уравнений относительно трех неизвестных.

**Пример 6.23 (Журнал "Квант")** *Семизначный код, состоящий из семи различных цифр, назовём хорошим. Известно, что сейф откроется, если введён хороший код и на каком-нибудь месте цифра кода совпала с соответствующей цифрой пароля. За какое наименьшее количество попыток можно с гарантией открыть сейф?*

Приведём пример, в котором сейф с гарантией открывается за 6 попыток:

|  |
| --- |
| 1234560, |
| 2345610, |
| 3456120, |
| 4561230, |
| 5612340, |
| 6123450. |

Среди первых 6 цифр пароля есть цифра от 1 до 6. Поскольку мы каждую из 6 цифр *по* разу набрали на каждом из первых 6 мест, значит, в одном из случаев она совпадает с соответствующей цифрой пароля.

Теперь докажем, что пяти попыток недостаточно. Пусть сделано пять попыток. Обозначим через M_1 множество цифр, которые не набирались в качестве первой цифры в этих попытках. Аналогично определим *множества* M_2,M_3,\ldots,M_7. Так как всего было 5 попыток, то |M_1|\geq 10-5=5 (здесь через |X| обозначаем количество элементов во множестве X). Так как каждая из цифр 0, 1, {\dots}, 9набирались не более одного раза в каждой попытке, то она содержится не менее чем в двух из семи множеств M_i. Покажем, что найдётся *пароль* a_1a_2\ldots a_7, где a_i - цифра из M_i; это и будет означать, что в случае такого пароля сейф не откроется.

*1 случай.* Пусть в двух из множеств M_i не менее семи цифр, скажем |M_6|\geq 7, |M_7|\geq 7. Тогда можно выбрать в M_1 любую цифру a_1, в M_2 - цифру a_2, отличную от a_1, и т. д., в M_7 - цифру a_7, отличную от a_1,\ldots, a_6.

*2 случай.* Пусть хотя бы в шести из множеств M_i не более 6 цифр. Тогда найдётся цифра x, которая не принадлежит двум из этих подмножеств, скажем x не принадлежит M_4 и M_5. Также выберем ещё одну цифру y\neq x, не принадлежащую M_5. Цифра y принадлежит хотя бы одному подмножеству M_i, i \neq 4, скажем y\in M_6. Цифра x принадлежит хотя бы одному подмножеству M_i, i\neq6, скажем x\in M_7. Положим a_6=y, a_7=x. Вычеркнув цифры x и y из подмножеств M_1,\ldots,M_5, мы приходим к подмножествам M'_1,\ldots, M'_5, для которых |M'_1|\geq 3$, $|M'_2|\geq 3, |M'_3|\geq 3, |M'_4|\geq 4, |M'_5|\geq 5 (M'_5 = M_5). Далее рассуждаем аналогично случаю 1: можно выбрать в M'_1 любую цифру a_1, в M'_2 - цифру a_2, отличную от а_1, и т. д., в M'_5 - цифру а_5, отличную от a_1, а}_2, а}_3, а}_4.

**Замечание.** Можно рассмотреть общую задачу про n-значный код в алфавите из m\leq n символов (в нашем случае n=7, m=10) и доказать, что минимальное число попыток при n\leq m/2 будет равно m+1-n, а при n<m/2 оно равно \left[\frac{m}{2}\right]+1.

**Пример 6.24 (фольклор)** *А теперь Вы по заданию Центра должны проникнуть на сервер противника. Для этого Вам предстоит подобрать пароль к серверу методом "грубой силы" (последовательного перебора). Известно, что пароль назначается компьютером, регулярно меняется, а предыдущие пароли выглядели следующим образом (все буквы - латинские):*

|  |
| --- |
| *abStwdRd* |
| *pVtrKRLp* |
| *iryzhToz* |
| *URbhhbEH* |
| *OJEXHZmJ* |
| *pzTDXJrZ* |

*Какие выводы надо сделать, чтобы ускорить процесс подбора пароля? (Время имеет значение!).*

**Решение.** Нетрудно заметить, что каждый из паролей является строкой из 8 латинских символов в верхнем и нижнем регистрах. Таким образом, *пространство* перебора имеет вид \{a,\ldots,z, A,\ldots,Z\}^8. *Мощность* парольного пространства равна 52^8\approx 5,3\cdot 10^{13}.

Для ускорения процесса подбора пароля необходимо обнаружить какие-либо закономерности в последовательности известных паролей, чтобы сократить *пространство* перебора.

Из приведенных в условии задачи паролей видно, что в каждом из них есть хотя бы одна буква в нижнем регистре и хотя бы одна - в верхнем. Можно предположить, что *генератор* парольных последовательностей исключает возможность генерации пароля, состоящего только из строчных или только из заглавных букв. Это предположение естественно, так как это является одним из классических критериев стойкости пароля к подбору.

Таким образом, из подбора следует исключить строки, состоящие только из строчных или только из заглавных букв. Это позволит уменьшить *пространство* перебора на 2\cdot 26^8 \approx 4,1\cdot 10^{11} попыток подбора.

**Пример 6.25** ([3]) *Специалист по защите информации А разработал собственную систему авторизации на компьютере. Пользователь вводит пароль - трехзначное натуральное число. Компьютер делит это число на n_1=31, полученный при этом остаток M умножает на 2 и получает число K. После этого число K делит на n_2=23 и полученный остаток A сохраняет на жестком диске. Например, если N=123, то M=30, соответственно, K=60 и A=14. Если пользователь ввел пароль P, и после указанных вычислений получилось число, совпадающее с числом, хранящимся в памяти компьютера, то он получает доступ.*

*Пользователь Б решил использовать на своем компьютере такую же систему. Но чтобы А не подал на него в суд за кражу интеллектуальной собственности, решил поменять местами числа 23 и 31. То есть сначала стал делить на n_2=23, а потом на n_1=31.*

*Известно, что в компьютере А и в компьютере Б хранится число x=5. Злоумышленник не знает паролей А и Б и поэтому перебирает их все подряд в случайном порядке. Чей компьютер он взломает быстрее?*

**Решение.** Всего различных паролей 900 (100 \leq N \leq 999). Вычислим, сколько из них дают в итоге число 5 *по* схеме пользователя А. *Остаток* 5 от деления на 23 дают натуральные числа 5, 28, 51. Больших чисел быть не может, так как K<62=31\cdot 2. Так как 5 и 51 - нечетные числа, то K=2M=28, значит, M=14. Числа N, дающие *остаток* 14 от деления на 31, имеют вид N=14 + 31s. Так как 100 \leq N \leq 999, то s=3, 4, 5, \ldots, 31. В итоге получаем 31-2=29 допустимых паролей (из 900 возможных).

Если те же вычисления провести *по* схеме Б, то получим следующих претендентов для числа K: 5, 36. Следовательно, M=18 и N=18+23s, где s= 4, 5, 6, \ldots, 42. Итого 39 допустимых паролей.

Итак, *вероятность* взлома системы Б выше в 39/29 \approx 1,34 раза, чем у А.

**Пример 6.26** ([3]) *Для подтверждения своих полномочий в компьютерной системе пользователь должен ввести свое имя и пароль, состоящий из 10 букв русского алфавита с исключенными Ё и Ь. Файл с паролями пользователей хранится на сервере в зашифрованном виде. Для их зашифровывания использовался следующий способ. Буквам алфавита поставлены в соответствие числа от 0 до 30: А - 0, Б - 1, ..., Я - 30. При зашифровывании пароля каждую его букву заменяют остатком от деления на 31 значения выражения (6a^3 + 5a^2 + 6a + 20), где a - число, соответствующее заменяемой букве. В начале каждого сеанса работы введенный пользователем пароль зашифровывается и сравнивается с соответствующей записью в файле. При совпадении сеанс продолжается, а при расхождении пароль запрашивается снова. Злоумышленник хочет войти в систему под чужим именем, а соответствующий этому имени пароль не знает. Он написал программу, которая в случайном порядке перебирает пароли. Какой из двух паролей - "ШИФРОВАНИЕ" или "КРИПТОГРАФ" - устойчивее к действию этой программы?*

**Решение.** В данной задаче необходимо вычислить все значения многочлена (6a^3 + 5a^2 + 6a + 20) *по* модулю 31 для всех чисел от 0 до 30. Результаты вычисления представлены в следующей таблице:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Таблица 6.26. Результат c шифрования буквы l с номером n | | | | | | | | | | |
| 0 | А | 20 |  | 11 | Л | 28 |  | 22 | Ц | 27 |
| 1 | Б | 6 |  | 12 | М | 20 |  | 23 | Ч | 10 |
| 2 | В | 7 |  | 13 | Н | 20 |  | 24 | Ш | 25 |
| 3 | Г | 28 |  | 14 | О | 2 |  | 25 | Щ | 15 |
| 4 | Д | 12 |  | 15 | П | 2 |  | 26 | Ъ | 16 |
| 5 | Е | 26 |  | 16 | Р | 25 |  | 27 | Ы | 2 |
| 6 | Ж | 13 |  | 17 | С | 14 |  | 28 | Э | 9 |
| 7 | З | 9 |  | 18 | Т | 5 |  | 29 | Ю | 11 |
| 8 | И | 19 |  | 19 | У | 3 |  | 30 | Я | 13 |
| 9 | Й | 17 |  | 20 | Ф | 13 |  |  |  |  |
| 10 | К | 8 |  | 21 | Х | 9 |  |  |  |  |

Таким образом, *слово* "*ШИФРОВАНИЕ*" зашифруется строкой чисел "25, 19, 13, 25, 2, 7, 20, 20, 19, 26". Перечисленные числа в качестве шифрсимволов встречаются соответственно 2, 1, 3, 2, 3, 1, 3, 3, 1 и 1 раза в таблице. Следовательно, число различных паролей, в зашифрованном виде имеющих вид "25, 19, 13, 25, 2, 7, 20, 20, 19, 26" будет равно 3^4 \cdot 2^2=324.

Для слова "КРИПТОГРАФ" мы после шифрования получаем последовательность "8, 25, 19, 2, 5, 2, 28, 25, 20, 13". Аналогичный подсчет дает число различных паролей, зашифровывающихся такой же последовательностью, равное 3^4\cdot 2^3=648.

*Пароль* "*ШИФРОВАНИЕ*" в 2 раза надежнее, чем "КРИПТОГРАФ".

**6.12 Краткий исторический очерк развития криптографии**

Наукой не установлен точный исторический период, когда появилась *криптография*, каковы были ее первоначальные формы и кто был ее создателем. Однако считается, что *криптография* старше египетских пирамид. Уже в исторических документах древних цивилизаций - Индии, Египта, Месопотамии-имеются сведения о системах и способах составления шифрованного письма. Сохранились достоверные сведения о системах шифров, применявшихся в Древней Греции. Свой след в истории криптографии оставили многие хорошо известные исторические личности.

В истории криптографии условно можно выделить четыре этапа:

1. *наивный*,
2. *формальный*,
3. *научный*,
4. *компьютерный*.

Рассмотрим кратко каждый из этапов.

1) Для *наивной криптографии* (до начала XVI в) характерно использование любых, обычно примитивных, способов запутывания противника относительно содержания шифруемых текстов. На начальном этапе для защиты информации использовались методы кодирования и стеганографии, которые родственны, но не тождественны криптографии.

Большинство из используемых шифров сводились к перестановке или моноалфавитной подстановке.

Первые сведения об использовании шифров в военном деле связаны с именем спартанского полководца Лисандра, разработавшего *шифр* "Сцитала". Этот *шифр* известен со времен войны Спарты против Афин в V веке до н.э. Для его реализации использовалась сцитала - жезл, имеющий форму цилиндра. На сциталу виток к витку наматывалась узкая папирусная лента (без просветов и нахлестов), а затем на этой ленте вдоль оси сциталы записывался текст. Лента разматывалась и получалось (для непосвященных), что поперек ленты в беспорядке написаны какие-то буквы. Затем лента отправлялась адресату. Адресат брал такую же сциталу, таким же образом наматывал на неё полученную ленту и читал сообщение вдоль оси сциталы. В этом шифре преобразование открытого текста в шифрованный заключается в определенной перестановке букв открытого текста. Поэтому *класс* шифров, к которым относится и *шифр* "Сцитала", называется шифрами перестановки.

Считается, что автором способа взлома шифра сциталы является Аристотель, который наматывал ленту на конусообразную палку до тех пор, пока не появлялись читаемые куски текста.

С именем Энея Тактика, полководца IV века до н.э., связывают несколько техник шифрования и тайнописи.

*Диск* Энея представлял собой *диск* диаметром 10-15см с отверстиями *по* числу букв алфавита. Для записи сообщения нитка протягивалась через отверстия в диске, соответствующим буквам сообщения. При чтении получатель вытягивал нитку, и получал буквы, правда, в обратном порядке. Хотя недоброжелатель мог прочитать сообщение, если перехватит *диск*, Эней предусмотрел способ быстрого уничтожения сообщения - для этого было достаточно выдернуть нить, закреплённую на катушке в центре диска.

Первым действительно криптографическим инструментом можно назвать линейку Энея, реализующей *шифр* замены. Вместо диска использовалась линейка с отверстиями *по* числу букв алфавита, катушкой и прорезью. Для шифрования нить протягивалась через прорезь и отверстие, после чего на нити завязывался очередной узел. Для дешифрования необходимо было иметь саму нить и линейку с аналогичным расположением отверстий. Таким образом, даже зная *алгоритм* шифрования, но не имея ключа (линейки), прочитать сообщение было невозможно.

В своём сочинении "О перенесении осады" Эней описывает ещё одну технику тайнописи, позже названную "книжный *шифр*". Он предложил делать малозаметные дырки рядом с буквами в книге или другом документе.

Эта система применялась даже в середине 20-го века, во время Второй мировой войны.

Во II веке до н.э. в Древней Греции был изобретён Квадрат Полибия. В нём буквы алфавита записывались в квадрат 5 на 5 (при использовании греческого алфавита одна *ячейка* оставалась пустой), после чего с помощью оптического телеграфа передавались номер строки и столбца, соответствующие символу исходного текста (на каждую букву приходилось два сигнала: число факелов обозначало разряд буквы *по* горизонтали и вертикали).

Некоторые исследователи полагают, что это можно рассматривать как первую систему, уменьшавшую (сжимавшую) исходный *алфавит*, и, в некотором смысле, как прообраз современной системы двоичной передачи данных.

Постепенно, *по* мере распространения техники частотного криптоанализа, шифры усложняются, что приводит к появлению шифров омофонической замены, а затем и полиалфавитных шифров.

Первые попытки систематизации и обобщения накопленного за века опыта были предприняты в арабских странах. Арабского происхождения и само *слово* "*шифр*".

Арабский филолог Халиль аль-Фарахиди первым обратил внимание на возможность использования стандартных фраз открытого текста для дешифрования. Он предположил, что первыми словами в письме на греческом языке византийскому императору будут "Во имя Аллаха", что позволило ему прочитать оставшуюся часть сообщения. Позже он написал книгу с описанием данного метода - "Китаб аль-Маумма" ("Книга тайного языка").

Первая книга, специально посвященная описанию и сравнению разных систем шифрования, появилась в 855 году. Это "Книга о большом стремлении человека разгадать загадки древней письменности" арабского учёного Абу Бакр Ахмед бен-Али бен-Вахшия ан-Набати, содержала описания нескольких шифров, в том числе с применением нескольких алфавитов. Также к IX веку относится первое известное упоминание о частотном криптоанализе - в книге Ал-Кинди "Манускрипт о дешифровке криптографических сообщений".

В книге X века "Адаб аль-Куттаб" ("Руководство для секретарей") ал-Сули есть инструкции *по* шифрованию записей о налогах, что подтверждает распространение криптографии в обычной, гражданской жизни.

В 1412 году выходит 14-томная энциклопедия Шехаба ал-Кашканди "Шауба ал-Аша", один из разделов которой "Относительно сокрытия в буквах тайных сообщений" содержал описание семи шифров замены и перестановки, частотного метода криптоанализа, а также таблицы частотности букв в арабском языке.

2) Этап *формальной криптографии* (конец XV-начало XX вв) связан с появлением формализованных и относительно стойких к ручному криптоанализу шифров. В европейских странах это произошло в эпоху Возрождения, когда развитие науки и торговли вызвало спрос на надежные способы защиты информации. Важная роль на этом этапе принадлежит Леону Батисте Альберти, итальянскому архитектору, который одним из первых предложил многоалфавитную подстановку. Данный *шифр*, получивший имя дипломата XVI в. Блеза Вижинера, состоял в последовательном "сложении" букв исходного текста с ключом (процедуру можно облегчить с помощью специальной таблицы). Его работа "Трактат о шифре" считается первой научной работой *по* криптологии. Одной из первых печатных *работ*, в которой обобщены и сформулированы известные на тот момент *алгоритмы шифрования*, является труд "*Полиграфия*" немецкого аббата Иоганна Трисемуса. Ему принадлежат два небольших, но важных открытия: способ заполнения полибианского квадрата (первые позиции заполняются с помощью легко запоминаемого ключевого слова, остальные - оставшимися буквами алфавита) и *шифрование* пар букв (биграмм). Простым, но стойким способом многоалфавитной замены (подстановки биграмм) является *шифр*Плейфера, который был открыт в начале XIXв. Чарльзом Уитстоном. Уитстону принадлежит и важное усовершенствование - *шифрование* "двойным квадратом". Шифры Плейфера и Уитстона использовались вплоть до первой мировой войны, так как с трудом поддавались ручному криптоанализу. В XIX в. голландец Керкхофф сформулировал главное требование к криптографическим системам, которое остается актуальным и поныне: секретность шифров должна быть основана на секретности ключа, но не алгоритма.

Большое внимание уделяли созданию и применению шифров такие известные исторические деятели, как Оливер Кромвель, Николо Макиавелли, кардинал Ришелье...

Традиции русского тайнописания уходят своими корнями в допетровскую эпоху. Известно, что крупные политические и военные акции Ивана Грозного оказали влияние и на развитие тайнописного дела. И традиции были продолжены. Так, в наказе царя Федора Иоанновича(1557-1589) - сына Ивана Грозного, - данном в 1589г. послу Николаю Воркачу, ему поручалось "писать письма мудрою азбукою, чтоб оприч Царского величества никто не разумел". С конца 16в. русские посланники за рубежом получают шифры в виде таблиц. И все же первым из российских государей, который предельно ясно осознавал важность развития шифровального дела для обеспечения безопасности государства, был Петр Великий. Государственные шифры того времени были шифрами простой замены. В российские шифры с начала 18в., как правило, вводятся "пустышки" - шифробозначения, которым не соответствует никакого знака открытого текста. Хотя обычно для этого использовалось всего пять-восемь шифрвеличин в качестве пустышек, очевидно, что введение их в *шифртекст* отражает стремление создателей шифров осмыслить процесс дешифрования. Эти пустышки разбивают структурные лингвистические связи открытого текста и изменяют статистические характеристики, то есть именно те особенности текста, которые используют в первую *очередь* при атаке на *шифр* простой замены. Кроме того, они изменяют длину передаваемого сообщения. Поэтому, видимо, не случайно, первый такой русский *шифр* был раскрыт англичанами лишь в 1725г.}

Наконец, последним словом в донаучной криптографии, которое обеспечило еще более высокую криптостойкость, а также позволило автоматизировать процесс шифрования стали роторные криптосистемы.

Одной из первых подобных систем стала изобретенная в 1790 г. Томасом Джефферсоном механическая машина. Многоалфавитная *подстановка* с помощью роторной машины реализуется вариацией взаимного положения вращающихся роторов, каждый из которых осуществляет "прошитую" в нем подстановку.

Практическое распространение роторные машины получили только в начале XX в. Одной из первых практически используемых машин, стала немецкая Enigma, разработанная в 1917 г. Эдвардом Хеберном и усовершенствованная Артуром Кирхом. Роторные машины активно использовались во время второй мировой войны. Помимо немецкой машины Enigma использовались также устройства Sigaba (США), Турех (Великобритания), *Red*, Orange и Purple (Япония). Роторные системы - *вершина* формальной криптографии, так как относительно просто реализовывали очень стойкие шифры. Успешные криптоатаки на роторные системы стали возможны только с появлением ЭВМ в начале 40-х гг.

И первая, и вторая мировые войны дали много примеров применения математики к анализу шифров. Криптографы воюющих стран напряженно работали, совершенствуя свои *алгоритмы шифрования* и "взламывая" шифры противника. И во второй половине 20в. ведущие державы мира уделяли большое внимание развитию криптографии.

3) Главная отличительная черта *научной криптографии* (1930 - 60-е гг.) - появление криптосистем со строгим математическим обоснованием криптостойкости. К началу 30-х гг. окончательно сформировались *разделы* математики, являющиеся научной основой криптологии: теория вероятностей и математическая *статистика*, общая *алгебра*, теория чисел, начали активно развиваться теория алгоритмов, *теория информации*, кибернетика. Своеобразным водоразделом стала работа Клода Шеннона "Теория связи в секретных системах", которая подвела научную базу под криптографию и *криптоанализ*. С этого времени стали говорить о криптологии (от греческого kryptos - тайный и logos - сообщение) - науке о преобразовании информации для обеспечения ее секретности. Этап развития криптографии и криптоанализа до 1949 г. стали называть донаучной криптологией.

Шеннон ввел понятия "рассеивание" и "перемешивание", обосновал возможность создания сколь угодно стойких криптосистем. В 1960-х гг. ведущие криптографические школы подошли к созданию блочных шифров, еще более стойких *по* сравнению с роторными криптосистемами, однако допускающих практическую реализацию только в виде цифровых электронных устройств

4) *Компьютерная криптография* (с 1970-х гг.) обязана своим появлением вычислительным средствам с производительностью, достаточной для реализации криптосистем, обеспечивающих при большой скорости шифрования на несколько порядков более высокую криптостойкость, чем "ручные" и "механические" шифры.

Первым классом криптосистем, практическое применение которых стало возможно с появлением мощных и компактных вычислительных средств, стали блочные шифры. В 70-е гг. был разработан американский стандарт шифрования *DES*. Один из его авторов, Хорст Фейстель описал модель блочных шифров, на основе которой были построены другие, более стойкие симметричные криптосистемы, в том числе отечественный стандарт шифрования ГОСТ 28147-89.

С появлением *DES* обогатился и *криптоанализ*, для атак на американский *алгоритм* был создано несколько новых видов криптоанализа (линейный, дифференциальный и т.д.), практическая реализация которых опять же была возможна только с появлением мощных вычислительных систем.

В середине 70-х гг. ХХ столетия произошел настоящий прорыв в современной криптографии - появление асимметричных криптосистем, которые не требовали передачи секретного ключа между сторонами. Здесь отправной точкой принято считать работу, опубликованную Уитфилдом Диффи и Мартином Хеллманом в 1976 г. под названием "Новые направления в современной криптографии". В ней впервые сформулированы принципы обмена шифрованной информацией без обмена секретным ключом. Независимо к идее асимметричных криптосистем подошел Ральф Меркли. Несколькими годами позже Рон Ривест, Ади Шамир и Леонард Адлеман открыли систему *RSA*, первую практическую асимметричную криптосистему, стойкость которой была основана на проблеме факторизации больших простых чисел. Асимметричная *криптография* открыла сразу несколько новых прикладных направлений, в частности системы электронной цифровой подписи (*ЭЦП*) и электронных денег.

В 1980-90-е гг. появились совершенно новые направления криптографии: вероятностное *шифрование*, квантовая *криптография* и другие. Актуальной остается и задача совершенствования симметричных криптосистем. В этот же период были разработаны нефейстелевские шифры (SAFER, *RC6* и др.), а в 2000 г. после открытого международного конкурса был принят новый национальный стандарт шифрования США - *AES*.

**6.13 Теоретические вопросы и упражнения**

1. Хотя аффинные криптосистемы с биграммами (т. е. по модулю N^2) лучше аналогичных однобуквенных систем (т. е. по модулю N), они также имеют недостаток. Укажите этот недостаток.
2. Каково максимальное число простых замен, из которых может состоять многоалфавитный шифр?
3. Алгоритмы сжатия данных хорошо подходят для совместного использования с криптографическими алгоритмами. Какой из двух вариантов предпочтительнее: сжать файл до того, как он будет зашифрован или сначала зашифровать, потом сжать?
4. В алгоритме DES таблицы замен одинаковы для всех пользователей и открыто опубликованы, секретным является только ключ. В российском алгоритме ГОСТ 28147-89 таблицы (а не только ключ) секретны и выбираются пользователем. Перечислите достоинства и недостатки этих двух подходов.
5. Перечислите причины, по которым криптостойкость алгоритма шифрования со временем изменяется.
6. Попробуйте назвать как можно большее количество причин ненадежности криптосистем.
7. Какие тексты на русском языке имеют бoльшую избыточность: литературные или технические?
8. Разделение секрета. Для доступа к серверу с информацией надо ввести три целых числа a, b, c, являющихся коэффициентами функции f(x) = ax^2 + bx + c. Представителям четырех делегаций были переданы следующие значения функции: "красным" f(21), "белым" f(24), "зеленым" f(25), "синим" - f(28). Когда представители партий встретились, чтобы совместно найти a, b, c и получить доступ к важной информации, один из представителей, чтобы сорвать мероприятие, предъявил неверное значение. Выясните, кто это был, если известно, что "красные" предъявили число 273, "белые" - 357, "зеленые" - 391, "синие" - 497.
9. Нами перехвачены два шифртекста. Алгоритм шифрования известен: это шифр Виженера, ключ неизвестен. Но наша агентура смогла добыть открытый текст, соответствующий первому шифртексту. Кроме того, стало известно, что противник не менял ключ. Теперь можно вычислить открытый текст, соответствующий второму шифртексту, сделайте это. Первый шифртекст: имидфщ, первый открытый текст: яблоко, второй шифртекст: фыееьщюгкцею.
10. (Повышенной сложности).Оценить расстояние единственности для шифра простой замены для : а) русского, б) английского языка.

**Список литературы**

1. Жданов О.Н., Куденкова И.А. Криптоанализ классических шифров. -- Красноярск: СибГАУ, 2007.
2. Жданов О.Н., Золотарев В.В. Методы и средства криптографической защиты информации: Учеб. пособие. -- Красноярск: Сиб. гос. аэрокосмич. ун-т, 2007. -- 253 с.
3. Сборник задач краевых олимпиад 2007--2011 гг. с решениями и комментариями / Под ред. О.Н. Жданова -- Красноярск:СибГАУ, 2011. -- 98с.
4. Коблиц Н. Курс теории чисел и криптографии -- М.: ТВП, 2003.
5. О.Н.Жданов, А.В.Соколов. Алгоритм шифрования с переменной фрагментацией блока, "Проблемы и достижения в науке и технике", Выпуск II. Сборник научных трудов по итогам международной научно-практической конференции (7 мая 2015г.) г. Омск, с.153-159.
6. Рябко Б.А., Филонов А.Н. Криптографические методы защиты информации: Учебное пособие для ВУЗов. Новосиб.:СибГУТИ, 2008. -- 229 с.
7. Задачи лингвистических олимпиад. 1965--1975 / Ред.-состав. В.И. Беликов, Е.В. Муравенко, М.Е. Алексеев -- М.:МЦНМО, 2006. --570 с.
8. Шнайер Б. Прикладная криптография. Протоколы, алгоритмы, исходные тексты на языке Си -- М. : Триумф, 2003. -- 816 с.