СОДЕРЖАНИЕ

1 Задание на практическую работу

Целью работы является приобретение навыков программной реализации и криптоанализа применительно к шифрам гаммирования.

В рамках практической работы необходимо выполнить следующее:

1) написать программную реализацию следующих шифра Виженера с тремя способами выработки гаммы на основе секретного ключа шифрования:
   - повторение короткого лозунга;
   - самоключ Виженера по открытому тексту;
   - самоключ Виженера по шифртексту;
2) изучить методы криптоанализа шифров гаммирования с использованием дополнительных источников;
3) провести криптоанализ данных шифров;
4) подготовить отчет о выполнении работы.

2 Краткая теоретическая часть

2.1 Описание шифров

Гаммирование заключается в наложении на открытый текст некоторой последовательности (гаммы), генерируемой на основе ключа шифрования. Под наложением гаммы на открытый текст обычно подразумевается сложение символов открытого текста с символами гаммы по модулю соответствующего алфавита. Однако в классических шифрах наложение гаммы может означать вычисление значений символов шифртекста на основе значений соответствующих символов открытого текста и гаммы по некоторому правилу.

Классическим представителем шифров гаммирования является шифр Виженера, устроенный следующим образом.

Символы алфавита A мощностью m представляются элементами множества $\mathbb{Z}_m$. Открытый текст и шифртекст обозначим соответственно $x = (x_1, \dots, x_l)$ и $y = (y_1, \dots, y_l)$, где $x_i, y_i \in Z_m$, $i = \overline{1,l}$.

Ключ шифрования представляет собой некоторую последовательность символов алфавита $k = (k_1, \dots, k_r)$, $k_j \in Z_m$, $j = \overline{1,r}$, $r \leq l$, которая служит для формирования гаммы $\gamma = (\gamma_1, \dots, \gamma_r)$, $\gamma_i \in Z_m$, $i = \overline{1,r}$.

Зашифрование заключается в сложении символов открытого текста с символами гаммы по модулю m:

$$y_i = (x_i + \gamma_i) \bmod m.$$

Расшифрование заключается в вычитании символов гаммы из символов шифртекста по модулю m.

В шифре Виженера в качестве ключа шифрования обычно использовалась короткая фраза, называемая лозунгом (паролем), которая циклически повторялась, формируя гамму.

Существует другой подход к формированию псевдослучайной ключевой последовательности — самоключ Виженера. Здесь в качестве начального ключа мы выбираем только один символ, к нему добавляем все символы открытого текста, за исключением последнего, и таким образом формируем гамму. Либо мы можем формировать гамму, добавляя к начальному символу поочередно символы шифртекста

2.2 Методы криптоанализа шифров

Криптоанализ классического шифра Виженера состоит из двух частей:

1) Нахождение длинны ключа;
2) Нахождение самого ключа.

Теперь о каждом пункте подробнее. Нахождение длины ключа основывается на понятии индекс совпадений. Любой осмысленный текст, написанный на естественном языке, сохраняет статистическое значение близкое к среднему. Само значение индекса совпадения вычисляется по формуле (2.2.1) [1]:

$$C_m = \sum \frac{f_i * (f_i - 1)}{l_i * (l_i - 1)} \tag{2.2.1}$$

где C_m – индекс совпадения для текста m;

f_i – частота i-ой буквы в выбранном тексте;

l_i – длина текста.

К примеру, для текстов, написанных на английском языке C_m приблизительно равен 0.066, а на русском 0.055. При криптоанализе шифра Виженера, если не рассматривать случай, где длина ключа равна единице, так как при этом индекс совпадения для такого текста будет примерно равен индексу осмысленного текста, данная величина используется по следующему принципу:

1) Необходимо вычислить индексы совпадения для текстов составленных из каждого второго, третьего, четвертого и т.д до длины ключа. символов исходного текста (2.2.2):

$$\begin{aligned} C_2 &= d_2 \\ C_3 &= d_3 \\ C_4 &= d_4 \\ &\dots \\ C_t &= d_t \\ &\dots \\ C_i &= d_i \end{aligned} \tag{2.2.2}$$

где C_i – индекс совпадения для текста составленного каждым i-ым символом исходного текста;

d_i – величина которой равен C_i.

2) Из полученных в предыдущем пункте индексов необходимо отобрать наиболее близкий по значению к индексу естественного языка, в выражении (2.2.2) он обозначен, как C_t, где t – длинна ключа, использованного при шифровании.

После установления длины ключа, использованного при зашифровании, исходный текст можно представить в виде t текстов зашифрование которых осуществлялось простым сдвигом по алфавиту, а значит, к ним можно применить частотный криптоанализ. При этом вскрытие состоит лишь в сопоставлении частоты появления символов в используемом алфавите с частотой появления символов в закрытом тексте.

Криптоанализ шифра Виженера с самоключом. При использовании данного типа шифра, как уже говорилось ранее гамма формируется следующим образом, к одному выбранному нами символу алфавита приписываются весь текст, кроме последнего символа. Полученная последовательность и используется, как ключ. Наилучшим способом криптоанализа данного шифра является простой перебор начальных символов, т.к. в данном случае его сложность равна мощности используемого алфавита, при этом последующие символы для каждой гаммы формируются по формуле (2.2.3):

$$\gamma_i = (y_{i-1} - \gamma_{i-1}) \bmod 26 \qquad (2.2.3)$$

где γ_i – i-ый символ гаммы;

y_i – i-ый символ шифротекста.

Криптоанализ шифра Виженера с самоключом по шифротексту. При криптоанализе данной вариации шифра, как и в предыдущем случае стоит использовать перебор начального символа гаммы. При этом получение каждого последующего символа гаммы происходит по следующей формуле (2.2.4):

$$\gamma_i = y_{i-1} \qquad (2.2.4)$$

где γ_i – i-ый символ гаммы;

y_i – i-ый символ шифротекста.

3 Примеры шифрования

Классический шифр Виженера. Для начала условимся во всех обозначениях (3.1), используемых далее, пусть:

$$\begin{gathered} A = \{a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z\}; \\ |A| = m = 26; \\ X = \{x = (x_1, x_2, x_3, \dots, x_r),\ \ x_i \in A\}; \\ Y = \{y = (y_1, y_2, y_3, \dots, y_r),\ \ y_i \in A\}; \\ k = \text{key} \\ n = 3 \\ \gamma \end{gathered} \tag{3.1}$$

где A – алфавит;

m – мощность алфавита;

X – открытый текст;

x_i – элемент открытого текста;

Y – закрытый текст;

y_i – элемент закрытого текста;

k – слоган;

n – длина слогана;

γ – гамма ключ.

Возьмем для примера X = ilyadolgov. Теперь необходимо представить каждый символ открытого текста в виде численного эквивалента равного порядковому номеру буквы в алфавите, при нумерации с нуля, после разбить текст на блоки размера длины ключа. В случае если текст не разбивается ровно на блоки такого размера, то последний оставить, не дополняя его символами (3.2).

$$\begin{gathered} X = (\text{ilyadolgov}) \\ i - 8 \\ l - 11 \\ y - 24 \end{gathered} \tag{3.2}$$

$$
\begin{gathered}
a - 0 \\
d - 3 \\
o - 14 \\
l - 11 \\
g - 6 \\
o - 14 \\
v - 21 \\
X = \{(8; 11; 24)(0; 3; 14)(11; 6; 14)(21)\}
\end{gathered}
\tag{3.2}
$$

После чего необходимо каждому блоку открытого текста сопоставить ключ, при этом каждый символ ключа так же необходимо заменить численным эквивалентом по тому же принципу, что и символы открытого текста (3.3).

$$
\begin{gathered}
X = \{(8; 11; 24)(0; 3; 14)(11; 6; 14)(21)\} \\
\gamma = \{(10; 4; 24)(10; 4; 24)(10; 4; 24)(10)\}
\end{gathered}
\tag{3.3}
$$

Теперь необходимо применяя формулу (3.4) необходимо получить закрытый текст.

$$
\begin{gathered}
y_i = (x_i + \gamma_i) \bmod 26 \\
y_1 = (x_1 + \gamma_1) \bmod 26 = 8 + 10 = 18 \\
y_2 = (x_2 + \gamma_2) \bmod 26 = 11 + 4 = 15 \\
y_3 = (x_3 + \gamma_3) \bmod 26 = 24 + 24 = 22 \\
y_4 = (x_4 + \gamma_4) \bmod 26 = 0 + 10 = 10 \\
y_5 = (x_5 + \gamma_5) \bmod 26 = 3 + 4 = 7 \\
y_6 = (x_6 + \gamma_6) \bmod 26 = 14 + 24 = 12 \\
y_7 = (x_7 + \gamma_7) \bmod 26 = 11 + 10 = 21 \\
y_8 = (x_8 + \gamma_8) \bmod 26 = 6 + 4 = 10 \\
y_9 = (x_9 + \gamma_9) \bmod 26 = 14 + 24 = 12 \\
y_{10} = (x_{10} + \gamma_{10}) \bmod 26 = 21 + 10 = 5
\end{gathered}
\tag{3.4}
$$

Тогда $Y = (18;\ 15;\ 22;\ 10;\ 7;\ 12;\ 21;\ 10;\ 12;\ 5)$ = spwkhmvkmf.

Теперь проведем обратную операцию, то есть расшифруем полученный закрытый текст. Для этого необходимо разбить шифротекст на блоки размера длины ключа, аналогично данному разбиению при шифровании, сопоставить каждому блоку символов

соответствующий ему блок символов ключа, после чего из каждого символа шифротекста вычесть соответствующий ему символ гаммы по модулю мощности алфавита, данные операции приведена в формуле (3.5).

$$
\begin{gathered}
Y = \{(18;\ 15;\ 22)(10;\ 7;\ 12)(21;\ 10;\ 12)(5)\} \\
\gamma\ = \{(10; 4; 24)(10; 4; 24)(10; 4; 24)(10)\} \\
x_i\ =\ (y_i - \gamma_i) \bmod 26 \\
x_1 = (y_1 - \gamma_1) \bmod 26 = 18 - 10 = 8 \\
x_2 = (y_2 - \gamma_2) \bmod 26 = 15 - 4 = 11 \\
x_3 = (y_3 - \gamma_3) \bmod 26 = 22 - 24 = 24 \\
x_4 = (y_4 - \gamma_4) \bmod 26 = 10 - 10 = 0 \\
x_5 = (y_5 - \gamma_5) \bmod 26 = 7 - 4 = 3 \\
x_6 = (y_6 - \gamma_6) \bmod 26 = 12 - 24 = 14 \\
x_7 = (y_7 - \gamma_7) \bmod 26 = 21 - 10 = 11 \\
x_8 = (y_8 - \gamma_8) \bmod 26 = 10 - 4 = 6 \\
x_9 = (y_9 - \gamma_9) \bmod 26 = 12 - 24 = 14 \\
x_{10} = (y_{10} - \gamma_{10}) \bmod 26 = 5 - 10 = 21
\end{gathered}
\tag{3.5}
$$

Тогда $X = (8; 11; 24; 0; 3; 14; 11; 6; 14; 21) = \text{ilyadolgov}$.

Шифр Виженера с самоключом. Для начала так же условимся с используемыми далее обозначениями (3.6):

$$
\begin{gathered}
A = \{a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z\}; \\
|A| = m = 26; \\
X = \{x = (x_1, x_2, x_3, \dots, x_r),\ \ x_i \in A\}; \\
Y = \{y = (y_1, y_2, y_3, \dots, y_r),\ \ y_i \in A\}; \\
k = a \\
\gamma = a \dots
\end{gathered}
\tag{3.6}
$$

где A – алфавит;

m – мощность алфавита;

X – открытый текст;

x_i – элемент открытого текста;

Y – закрытый текст;

y_i – элемент закрытого текста;

k – ключ;

γ – гамма ключ.

Пусть X = ilya. Необходимо представить каждый символ открытого текста в виде численного эквивалента являющимся порядковым номером символа в алфавите при нумерации с нуля. (3.7).

$$\begin{gathered} X = (ilya) \\ i - 8 \\ l - 11 \\ y - 24 \\ a - 0 \end{gathered} \tag{3.7}$$

После чего необходимо сформировать гамму, которая при данном типе шифра формируется, добавлением к начальному символу k описанному в (3.6) всех символов открытого текста, за исключением последнего (3.8).

$$\gamma = (0; 8; 11; 24) \tag{3.8}$$

Теперь остается применить формулу (3.9) и получить закрытый текст.

$$\begin{gathered} y_i = (x_i + \gamma_i) \bmod 26 \\ y_1 = (x_1 + \gamma_1) \bmod 26 = 8 + 0 = 8 \\ y_2 = (x_2 + \gamma_2) \bmod 26 = 11 + 8 = 19 \\ y_3 = (x_3 + \gamma_3) \bmod 26 = 24 + 11 = 9 \\ y_4 = (x_4 + \gamma_4) \bmod 26 = 0 + 24 = 24 \end{gathered} \tag{3.9}$$

Тогда $Y = (8; 19; 9; 24) = itjy$

Теперь проведем обратную операцию, то есть расшифруем полученный закрытый текст. Для этого проделаем с ним те же операции, что и описанные ранее с открытым текстом, то есть представим каждый символ в численном эквиваленте, а также сформируем гамму, для этого к начальному символу k, описанному в (3.6) будем поочередно добавлять элементы согласно формуле (3.10).

$$\gamma_i = (y_{i-1} - \gamma_{i-1}) \bmod 26, \text{при } i > 1$$
$$\gamma_1 = 0$$
$$\gamma_2 = (y_1 - \gamma_1) \bmod 26 = 8 \quad (3.10)$$
$$\gamma_3 = (y_2 - \gamma_2) \bmod 26 = 11$$
$$\gamma_4 = (y_3 - \gamma_3) \bmod 26 = 24$$

И того $\gamma = (0; 8; 11; 24)$. Теперь необходимо из символов закрытого текста вычитать соответствующие им символы гаммы по модулю мощности алфавита (3.11).

$$Y = (8; 19; 9; 24)$$
$$\gamma = (0; 8; 11; 24).$$
$$x_i = (y_i - \gamma_i) \bmod 26$$
$$x_1 = (y_1 - \gamma_1) \bmod 26 = 8 - 0 = 8 \quad (3.11)$$
$$x_2 = (y_2 - \gamma_2) \bmod 26 = 19 - 8 = 11$$
$$x_3 = (y_3 - \gamma_3) \bmod 26 = 9 - 11 = 24$$
$$x_4 = (y_4 - \gamma_4) \bmod 26 = 24 - 24 = 0$$

Тогда $X = (8; 11; 24; 0) = ilya$. Что и является использованным ранее открытым текстом.

Шифр Виженера с самоключом по шифротексту. Данный тип отличается от предыдущего тем, что гамма дополняется не символами открытого текста, а символами шифротекста. Условимся с обозначениями, используемыми далее (3.12).

$$A = \{a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z\};$$
$$|A| = m = 26;$$
$$X = \{x = (x_1, x_2, x_3, \dots, x_r),\ x_i \in A\};$$
$$Y = \{y = (y_1, y_2, y_3, \dots, y_r),\ y_i \in A\}; \quad (3.12)$$
$$k = a$$
$$\gamma = a \dots$$

где A – алфавит;

m – мощность алфавита;

X – открытый текст;

x_i – элемент открытого текста;

Y – закрытый текст;

y_i – элемент закрытого текста;

k – ключ;

γ – гамма ключ.

Пусть X = ilya. Необходимо представить каждый символ открытого текста в виде численного эквивалента являющимся порядковым номером символа в алфавите при нумерации с нуля. (3.13).

$$\begin{gathered} X = (ilya) \\ i - 8 \\ l - 11 \\ y - 24 \\ a - 0 \end{gathered} \tag{3.13}$$

После чего необходимо сформировать гамму, которая при данном типе шифра формируется, добавлением к начальному символу k описанному в (3.6) всех символов шифротекста кроме последнего (3.14).

$$\begin{gathered} \gamma_i = (x_{i-1} + \gamma_{i-1}) \bmod 26, \text{при } i > 1 \\ \gamma_1 = 0 \\ \gamma_2 = (x_1 + \gamma_1) \bmod 26 = 8 + 0 = 8 \\ \gamma_3 = (x_2 + \gamma_2) \bmod 26 = 11 + 8 = 19 \\ \gamma_4 = (x_3 + \gamma_3) \bmod 26 = 24 + 19 = 17 \end{gathered} \tag{3.14}$$

Тогда $\gamma = (0; 8; 19; 17)$. После формирования гаммы необходимо зашифровать открытый текст, для этого необходимо сложить символы открытого текста, представленных аналогично предыдущим представлениями открытого текста и сопоставленные им символ гаммы по модулю мощности алфавита (3.15).

$$\begin{gathered} y_i = (x_i + \gamma_i) \bmod 26 \\ y_1 = (x_1 + \gamma_1) \bmod 26 = 8 + 0 = 8 \\ y_2 = (x_2 + \gamma_2) \bmod 26 = 11 + 8 = 19 \\ y_3 = (x_3 + \gamma_3) \bmod 26 = 24 + 19 = 17 \\ y_4 = (x_4 + \gamma_4) \bmod 26 = 0 + 17 = 17 \end{gathered} \tag{3.15}$$

Тогда $Y = (8; 19; 17; 17) = \text{itrr}$.

Теперь выполним обратную операцию, то есть расшифруем полученный закрытый текст. Для этого представим каждый символ закрытого текста числовым эквивалентом, аналогично описанным ранее методу. После чего необходимо сформировать гамму. Гамма в данном случае формируется по следующей формуле (3.16).

$$\begin{gathered} \gamma_i = y_{i-1}, \text{при } i > 1 \\ \gamma_1 = 0 \\ \gamma_2 = y_1 = 8 \\ \gamma_3 = y_2 = 19 \\ \gamma_4 = y_3 = 17 \end{gathered} \tag{3.16}$$

После получения гаммы необходимо вычесть из символов шифротекста соответствующие им символы гаммы по модулю 26 (3.17).

$$\begin{gathered} Y = (8; 19; 17; 17) \\ \gamma = (0; 8; 19; 17). \\ x_i = (y_i - \gamma_i) \bmod 26 \\ x_1 = (y_1 - \gamma_1) \bmod 26 = 8 - 0 = 8 \\ x_2 = (y_2 - \gamma_2) \bmod 26 = 19 - 8 = 11 \\ x_3 = (y_3 - \gamma_3) \bmod 26 = 17 - 19 = 24 \\ x_4 = (y_4 - \gamma_4) \bmod 26 = 17 - 17 = 0 \end{gathered} \tag{3.11}$$

Тогда $X = (8; 11; 24; 0) = \text{ilya}$, что и является использованным открытым текстом.

4 Программная реализация шифров

Программная реализация была выполнена на языке python, с полным кодом можно ознакомиться по ссылке: https://github.com/il3241/crypto_pr3_github.

Примеры выполнения программного кода для последовательностей, показанных в разделе ручного зашифрования и расшифрования:

Классический шифр Виженера.

Зашифрование.

Пример входных данных

- Текст для зашифрования: ilyadolgov
- Ключ: key

Результат работы программы представлен на рисунке 1:

```
Введите текст на английском языке:
ilyadolgov
Введите тип шифра:
 1 - классический;
 2 - самоключ;
 3 - самоключ по шифротексту.
1
Введите необходимое действие de/en:
en
Введите ключ:
key
Полученный закрытый текст:
spwkhmvkmf
```

Рисунок 1.

Расшифрование.

Пример входных данных

- Текст для зашифрования: spwkhmvkmf
- Ключ: key

Результат работы программы представлен на рисунке 2:

```
Введите текст на английском языке:
spwkhmvkmf
Введите тип шифра:
 1 - классический;
 2 - самоключ;
 3 - самоключ по шифротексту.
1
Введите необходимое действие de/en:
de
Введите ключ:
key
Полученный открытый текст:
ilyadolgov
```

Рисунок 2.

Шифр Виженера с самоключом.

Зашифрование.

Пример входных данных

- Текст для зашифрования: ilya
- Ключ: a

Результат работы программы представлен на рисунке 3:

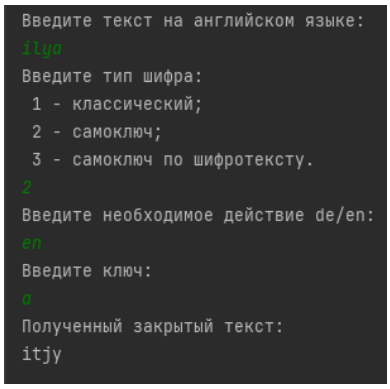

Рисунок 3.

Расшифрование.

Пример входных данных

- Текст для зашифрования: itjy
- Ключ: a

Результат работы программы представлен на рисунке 4:

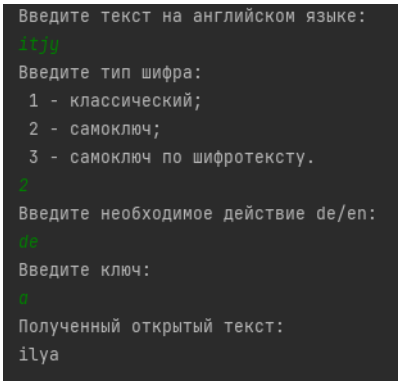

Рисунок 4.

Шифр Виженера с самоключом по шифротексту.

Зашифрование.

Пример входных данных

- Текст для зашифрования: ilya
- Ключ: a

Результат работы программы представлен на рисунке 5:

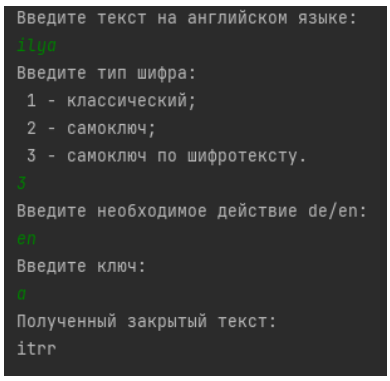

Рисунок 5.

Расшифрование.

Пример входных данных

- Текст для зашифрования: itrr
- Ключ: a

Результат работы программы представлен на рисунке 6:

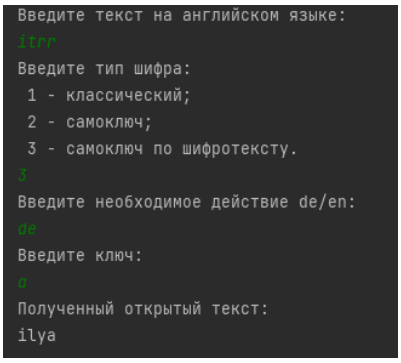

Рисунок 6.

5 Примеры криптоанализа

Начнем с классического шифра Виженера, в котором шифрование происходит через короткий повторяющийся лозунг. Зашифруем открытый текст X, для демонстрации был взят текст Оскара Уайльда “Мальчик-звезда” [2], с помощью ключа k = key написанным мной кодом: рисунок 7.

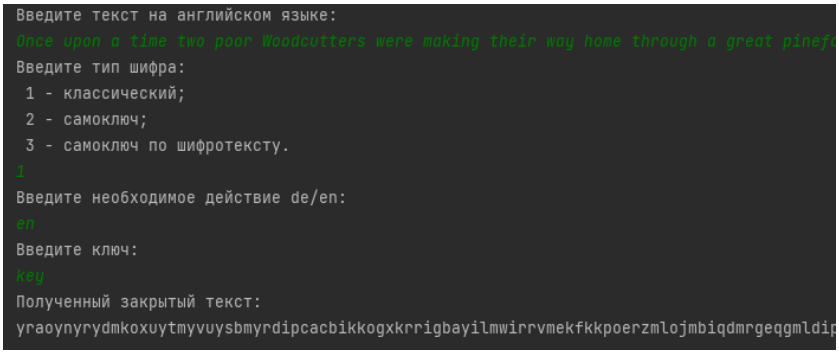

Рисунок 7.

Программа вывела закрытый текст Y, с ним можно ознакомиться по данной ссылке. Теперь согласно алгоритму описанному в 2.2 необходимо анализировать через индекс совпадений тексты из каждого второго, третьего, четвертого и т.д. символов шифротекста, для нахождения размера блока. Для этого мной была написана программа, считающая индекс совпадения, с полным кодом программы можно ознакомиться по данной ссылке: https://github.com/il3241/crypto_pr3_github. При выполнении программы выводится размер блока и индекс, совпадений соответствующий анализируемому тексту, этот результат показан на рисунке 8.

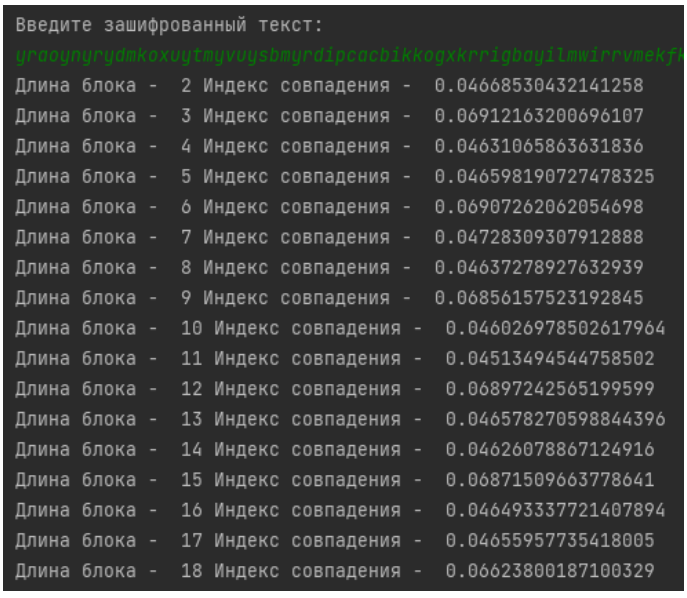

Рисунок 8.

При этом подсчет индекса совпадений происходил по формуле (5.1).

$$C_m = \sum \frac{f_i * (f_i - 1)}{l_i * (l_i - 1)} \quad (5.1)$$

где C_m – индекс совпадения для текста m;

f_i – частота i-ой буквы в выбранном тексте;

l_i – длина текста.

Как можно заметить наиболее близким по значению к естественному языку являются тексты, созданные каждым третьим символом алфавита, от сюда можно сделать вывод, что символы этого текста зашифрованы с помощью одного и того же символа ключа, а значит длина ключа равна трем. Так же стоит заметить, что индексом близким к реальному обладают все тексты, составленные из символов номер которых в тексте кратен трем, это происходит, потому что в таких случаях ключ с длинной три просто повторяется и символы открытого текста опять же кодируются с помощью одного и того же символа ключа. После того как была установлена длина ключа, необходимо разбить исходный текст на три, составленных из каждого первого, второго и третьего символа, если считать по блоку ключа, в более общем случае, при установлении длины ключа равной n текст необходимо разбить на n тексов составленных из каждого первого, второго и т.д. до n символа. Заметим, что каждый из текстов, на которые мы разбили зашифрован с помощью одной и той же буквы ключа, а значит этот шифротекст является результатом шифрования обычным шифром Цезаря, а следовательно, подвержен частотному криптоанализу, как и любой шифр, не размывающий статистику. Данный тип криптоанализа основывается на факте, что порядок букв в словах и фразах естественного языка подчиняется определенным статистическим закономерностям. Частотный анализ также учитывает частоту появления различных буквосочетаний: например, пара стоящих рядом букв «ся» в русском языке более вероятна, чем «цы», а «оь» не встречается никогда. Для большинства естественных языков такая статистика документирована [3]. В таблице 1 приведена статистика встречаемости букв английского алфавита.

Таблица 1 – Относительная частота появления букв английского алфавита

Буква	**Частота**	**Буква**	**Частота**	**Буква**	**Частота**	**Буква**	**Частота**
E	12.7	R	5.99	G	2.02	Q	0.1
T	9.06	D	4.25	Y	1.97	Z	0.05
A	8.17	L	4.03	P	1.93		
O	7.51	C	2.78	B	1.49		
I	6.97	U	2.76	V	0.98		

Продолжение таблицы 1

N	6.75	M	2.41	K	0.77		
S	6.33	W	2.36	X	0.15		
H	6.09	F	2.23	J	0.15		

Разбиение текста на несколько происходит с помощью программы расположенной по ссылке: https://github.com/il3241/crypto_pr3_github, ее функционал ограничен составлением текстов из каждой i-ой буквы исходного, составленные тексты находятся по данной ссылке, результат работы программы и часть текста приведена на рисунке 9.

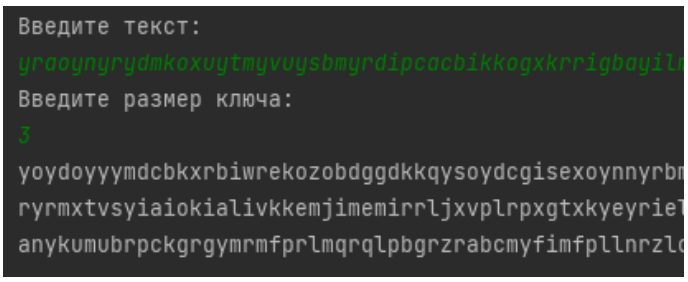

Рисунок 9.

Теперь к полученным текстам необходимо применить частотный анализ, для этого воспользуемся сервисом сбора статистики символов https://www.dcode.fr/frequency-analysis. Статистика для данного текста находится в таблице 2.

Таблица 2.

Буква	**Частота**	**Буква**	**Частота**	**Буква**	**Частота**	**Буква**	**Частота**
O	13.1%	C	5.52%	Q	1.88%	J	0.01%
D	9.57%	B	5.52%	I	1.67%		
R	8.76%	V	3.64%	Z	1.23%		
K	8.54%	G	3.13%	L	1.08%		
Y	7.7%	W	2.55%	F	0.74%		
S	6.34%	E	2.22%	U	0.71%		
X	5.87%	P	2.22%	T	0.07%		
N	5.76%	M	2.14%	A	0.05%		

После сбора статистических данных необходимо перейти к установлению ключа, поскольку зашифрование происходит по формуле (5.2):

$$y_i = (x_i + k) \bmod 26 \tag{5.2}$$

где y_i – i-ый символ шифротекста;

x_i – i-ый символ открытого текста;

k – ключ.

Нам необходимо составить уравнения для получения ключа, для точности стоит составить несколько уравнений, то есть систему, при этом пары символов открытый – закрытый текст будут подбираться по наиболее точному статистическому совпадению. К примеру, символ шифротекста O по статистике близок к E в естественном языке, D к T и т.д. В данном случае возьмем их для составления системы (5.3), при этом буквы представлены в их численном эквиваленте:

$$y_i = (x_i + k) \bmod 26 \Rightarrow k = (y_i - x_i) \bmod 26$$
$$k = (14 - 4) \bmod 26 = 10 = k \quad (5.3)$$
$$k = (3 - 19) \bmod 26 = 10 = k$$

Исходя из (5.3) можем установить, что первый символ ключа это k. Аналогичные действия необходимо проделать для двух оставшихся текстов, для установления оставшихся символов ключа. Необходимо собрать статистические показатели для второго текста, воспользуемся тем же сервисом, что и для первого, собранные данные представлен в таблице 3.

Таблица 3.

Буква	**Частота**	**Буква**	**Частота**	**Буква**	**Частота**	**Буква**	**Частота**
I	12.93%	V	5.59%	K	1.72%	B	0.01%
X	9.72%	W	5.38%	C	1.72%		
E	8.76%	P	3.71%	F	1.29%		
L	8.13%	A	2.72%	T	1.13%		
S	7%	J	2.45%	Z	0.91%		
R	6.84%	Y	2.4%	O	0.75%		
M	6.64%	Q	2.29%	U	0.09%		
H	5.72%	G	2.05%	N	0.02%		

Как можно заметить статистика I похожа на E их естественного языка, X на T. Составим систему уравнений (5.4) с помощью полученных ранее закономерностей.

$$y_i = (x_i + k) \bmod 26 \Rightarrow k = (y_i - x_i) \bmod 26$$
$$k = (8 - 4) \bmod 26 = 4 = e \quad (5.4)$$
$$k = (23 - 19) \bmod 26 = 4 = e$$

С помощью (5.4) можно установить, что второй символ ключа это e. Теперь остается собрать статистику для третьего текста и провести с его символами схожие операции. Статистические показатели приведены в таблице 4.

Таблица 4.

Буква	**Частота**	**Буква**	**Частота**	**Буква**	**Частота**	**Буква**	**Частота**
C	12.81%	P	5.86%	E	1.86%	X	0.02%
R	9.79%	Q	5.43%	W	1.58%		
F	8.61%	J	3.68%	N	1.25%		
M	7.57%	U	3%	Z	1.19%		
Y	7.45%	K	2.84%	T	0.83%		
L	6.53%	D	2.24%	I	0.66%		
B	6.49%	S	1.96%	H	0.05%		
G	6.33%	A	1.92%	O	0.04%		

Заметим, что C похожа на E по частотности, а R на T, составим систему (5.5) с этими данными.

$$y_i = (x_i + k) \bmod 26 \Rightarrow k = (y_i - x_i) \bmod 26$$
$$k = (2 - 4) \bmod 26 = 24 = y \quad (5.4)$$
$$k = (17 - 19) \bmod 26 = 24 = y$$

Отсюда делаем вывод, что третий символ ключа это y, а следовательно весь ключ это key, что и является ключом, который был использован при шифровании.

Криптоанализ шифра Виженера с самоключом, для этого зашифруем тот же текст произведения Оскара Уайльда “Мальчик-звезда” [2], с помощью ключа k = h написанным мной кодом: рисунок 10.

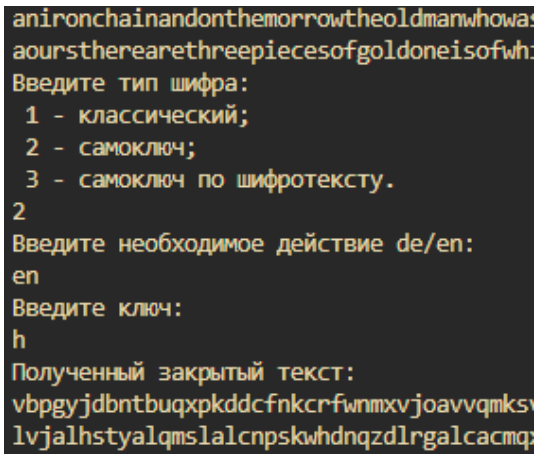

Рисунок 10.

Вариация шифра Виженера с самоключом не имеет стойкости к простому перебору ключа, поэтому наилучшем вариантом при криптоанализе будет простой перебор m возможных ключей, где m это мощность используемого алфавита. Для отображения получаемых текстов во время подбора ключа можно воспользоваться моей программой, результат работы которой представлен на рисунке 11.

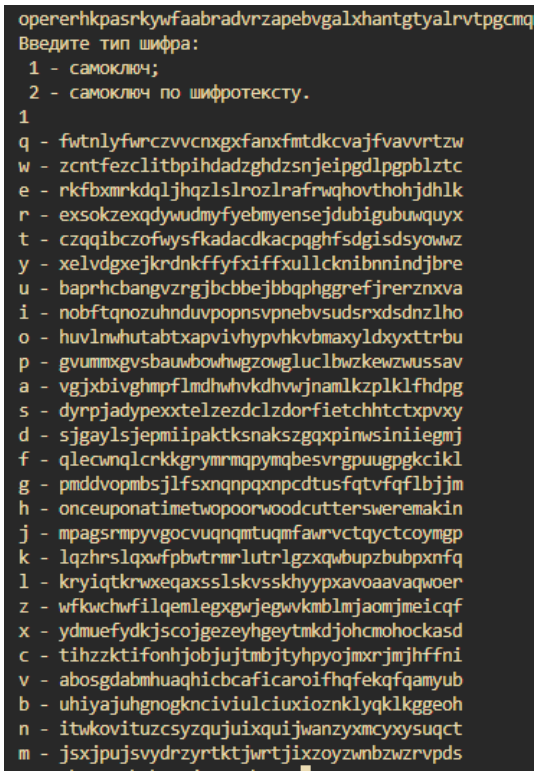

Рисунок 11.

Как видно, перебрав до m вариаций мы получаем открытый текст, в данном случае моей программой был установлен ключ h.

Криптоанализ шифра Виженера с самоключом по шифротексту так же заключается в переборе до m ключей, для нахождения использованного при зашифровании. Для примера зашифруем все тот же текст с ключом g. Результат шифрования приведен на рисунке 12.

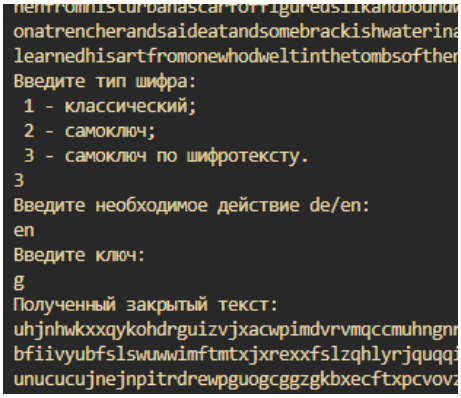

Рисунок 12.

К полученному результату опять требуется применить перебор, который осуществлён с помощью программы, описанной ранее, результат ее работы приведен на рисунке 13.

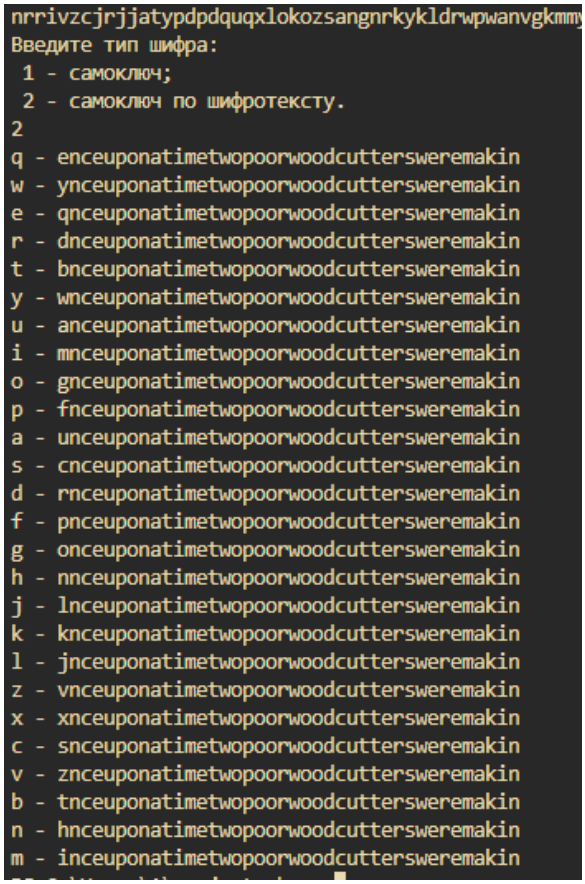

Рисунок 13.

В данном случае ключом является g, то есть тот же что был использован при шифровании. С полным кодом программы, использованной при криптоанализе можно ознакомиться по ссылке: https://github.com/il3241/crypto_pr3_github.

6 Выводы о проделанной работе

В данной практической работе мы рассмотрели три вариации шифра: классический шифр Виженера, шифр Виженера с самоключом, шифр Виженера с самоключом по шифротексту. Данные шифры обладают следующими плюсами и минусами.

Классический шифр Виженера.

Плюсы:

1) Легко хранить и передавать ключ;
2) Защищён от взлома перебором ключей;
3) Защищен от простого частотного криптоанализа.

Минусы:

1) Легко взламывается криптоанализом при установлении длины ключа (самый оптимальный вариант).

Шифр Виженера с самоключом.

Плюсы:

1) Легко хранить и передавать ключ.
2) Защищен от простого частотного криптоанализа

Минусы.

1) Легко поддается взлому перебором.

Шифр Виженера с самоключом по шифротексту.

Плюсы.

1) Легко хранить и передавать ключи.
2) Защищен от простого частотного криптоанализа

Минусы:

1) Легко поддается взлому перебором.

7 Список использованных источников

1. Индекс совпадений – URL: https://ru.wikipedia.org/wiki/Индекс_совпадений

2. Оскар Уайльд “Мальчик-звезда” – URL: http://www.poperechny.net/english/the-star-child-na-angliyskom-yazyke-malchik-zvezda.html

3. С.М. Авдошин, канд. техн. наук. проф., А.А. Савельева “Криптоанализ: современное состояние и перспективы развития” – URL: https://www.hse.ru/data/712/315/1234/Авдошин.Савельева_Криптоанализ.pdf