Федеральное государственное автономное образовательное учреждение высшего образования

«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ «ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

Московский институт электроники и математики им. Тихонова

Департамент электронной инженерии

Отчёт о практической работе №3

по дисциплине «Математические основы защиты информации»

«Шифры гаммирования»

Студент гр. БИБ202

Шадрунов А. С.

«25» мая 2022 г.

Руководитель

Заведующий кафедрой информационной безопасности киберфизических систем

канд. техн. наук, доцент

О. О. Евсютин

«___» _____ 2022 г.

Москва, 2022

Содержание

1. Задание на практическую работу

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

В рамках практической работы необходимо выполнить следующее:

1. Написать программную реализацию 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;
3. Провести криптоанализ данных шифров;
4. Подготовить отчет о выполнении работы.

2. Краткая теоретическая часть

2.1. Описание шифров

Шифры гаммирования – разновидность шифров, принцип которых заключается в наложении на открытый текст особой последовательности символов той же длины, вырабатываемая на основе секретного ключа шифрования. Такая последовательность называется гаммой.

В основе работы рассматриваемых в данной работе шифров лежит процесс гаммирования. Обычно под этим термином подразумевается сложение символов открытого текста с символами гаммы по модулю соответствующего алфавита, однако в современных шифрах правило, по которому вычисляются символы шифртекста, может быть более сложным [1].

Один из известных представителей шифров гаммирования – шифр Виженера. Рассмотрим его далее.

2.1.1. Шифр Виженера

Опишем алгоритм работы шифра Виженера [2]. Предварительно необходимо составить гамму на основе ключа шифрования; этот процесс разберём далее.

Алгоритм зашифрования:

1. Представить гамму γ и открытый текст x в виде элементов кольца классов вычетов по модулю мощности алфавита m:

$$\gamma = (\gamma_1, ..., \gamma_l), \gamma_i \in \mathbb{Z}_m$$
$$x = (x_1, ..., x_l), x_i \in \mathbb{Z}_m$$

2. Сложить номер символа гаммы и номер символа открытого текста по модулю мощности алфавита:

$$y_i = (x_i + \gamma_i) \; mod \, m$$

3. Представить шифртекст y в виде символов алфавита:

$$y = (y_1, ..., y_l), y_i \in \mathbb{Z}_m$$

Для **расшифрования** нужно проделать обратную операцию: из символов шифртекста вычесть символы гаммы: $x_i = (y_i - \gamma_i) \; mod \; m$.

2.1.2. Способы выработки гаммы

Существуют различные способы выработки гаммы по ключу шифрования.

- **Повторение ключа (лозунга)**. Гамма формируется циклическим повторением обычно короткого ключа (лозунга) до тех пор, пока размер гаммы не сравняется с размером текста.

- **Самоключ Виженера по открытому тексту**. В этом режиме в качестве начального ключа мы выбираем только один символ, к нему добавляем все символы открытого текста, за исключением последнего, и таким образом формируем гамму:

$$x = (x_1, ..., x_l), \; k = k_1, \; \gamma = (k_1, x_1, ..., x_{l-1})$$

- **Самоключ Виженера по шифртексту**. В этом режиме так же ключ состоит из одного символа, который формирует первый символ гаммы и используется для зашифрования первого символа открытого текста. Далее полученный символ шифртекста становится вторым символом гаммы, затем процесс повторяется.

$$x = (x_1, ..., x_l), \; k = k_1,$$
$$\gamma = (k_1, y_1, ..., y_{l-1})$$
$$y = (y_1, ..., y_l)$$

2.2. Методы криптоанализа шифров

2.2.1. Нахождение длины ключа

Криптоанализ шифра Виженера осложняется тем, что злоумышленник не знает длину ключа. Без этой информации нельзя применить полный перебор или другие техники атаки на шифр. Однако при малом размере ключа можно узнать его размер перебором. Для этого необходимо предположить, что длина ключа равняется n, и затем выбрать из шифртекста каждую n-ю букву. Если n угадано, то полученные символы образуют текст, зашифрованный, по сути, шифром простой замены, дешифровать который не составляет труда.

Чтобы узнать, правильно ли подобрано n, можно воспользоваться **индексом совпадений** [3]. Индекс совпадений — число, показывающее вероятность того, что две произвольно выбранные из текста буквы окажутся одинаковыми. Для любого текста индекс совпадений вычисляется по формуле:

$$I = \sum_i \frac{f_i(f_i - 1)}{n(n-1)},$$

где f_i – количество появлений i-й буквы алфавита в тексте, а n – количество букв в тексте. Для английского языка индекс совпадений имеет значение 0.0667, в то время как для случайного набора букв этот показатель равен 0.038. Для текста, зашифрованного с помощью моноалфавитного шифра, индекс совпадений также равен 0.0667, что и позволяет отыскать необходимый параметр.

Пример этой атаки будет продемонстрирован далее.

В заключение можно отметить, что атака брутфорс в данном шифре даже теоретически невозможна, так как путём полного перебора гаммы получим все возможные открытые тексты данной длины, в том числе все осмысленные тексты. Выбрать верный из них невозможно [4].

3. Примеры шифрования

Зашифруем слово "Калининград" с помощью шифра Виженера.

$$X = \text{КАЛИНИНГРАД}$$

3.1. Повторение ключа

Воспользуемся методом повторения ключа для выработки гаммы. В качестве ключа будем использовать слово **"шифр"**. Представим текст и ключ в виде элементов кольца классов вычетов $\mathbb{Z}_m$:

$$X = (11\ 0\ 12\ 9\ 14\ 9\ 14\ 3\ 17\ 0\ 4)$$

$$K = (25\ 9\ 21\ 17)$$

Получим гамму повторением ключа:

$$\gamma = (25\ 9\ 21\ 17\ 25\ 9\ 21\ 17\ 25\ 9)$$

3.1.1. Зашифрование

Для зашифрования необходимо сложить по модулю $m = 33$ значения открытого текста и гаммы:

$$\begin{array}{r} X = 11\ 0\ 12\ 9\ 14\ 9\ 14\ 3\ 17\ 0\ 4 \\ \gamma = 25\ 9\ 21\ 17\ 25\ 9\ 21\ 17\ 25\ 9\ 21 \\ \hline Y = 3\ 9\ 0\ 26\ 6\ 18\ 2\ 20\ 9\ 9\ 25 \end{array}$$

В текстовом виде получаем:

$$Y = \text{ГИАЩЁСВУИИШ}$$

3.1.2. Расшифрование

Расшифруем последовательность "ГИАЩЁСВУИИШ" с помощью шифра Виженера. Для этого из значений шифртекста вычитаем гамму:

Получаем исходное слово "Калининград".

$$
\begin{array}{r}
Y = \ 3\ 9\ 0\ 26\ 6\ 18\ 2\ 20\ 9\ 9\,25 \\
\gamma = \ 25\,9\,21\ 17\,25\ \ 9\ \ 21\ 17\,25\,9\,21 \\
\hline
X = 11\ 0\ 12\ \ 9\ \ 14\ \ 9\ \ 14\ \ 3\ \ 17\ 0\ \ 4
\end{array}
$$

3.2. Самоключ Виженера по открытому тексту

В данном режиме гамма представляет собой конкатенацию односимвольного ключа и открытого текста, за исключением его последнего символа. В качестве ключа выберем, к примеру, символ **"К"**. Представим текст и ключ в виде элементов кольца классов вычетов $\mathbb{Z}_m$:

$$X = (11\ 0\ 12\ 9\ 14\ 9\ 14\ 3\ 17\ 0\ 4)$$

$$K = (11)$$

Получим гамму:

$$\gamma = (11\ 11\ 0\ 12\ 9\ 14\ 9\ 14\ 3\ 17\ 0)$$

3.2.1. Зашифрование

Зашифрование происходит аналогично предыдущему случаю:

$$
\begin{array}{r}
X = 11\ \ 0\ \ 12\ \ 9\ \ 14\ \ 9\ \ 14\ \ 3\ \ 17\ \ 0\ \ 4 \\
\gamma = \ 11\ 11\ \ 0\ \ 12\ \ 9\ \ 14\ \ 9\ \ 14\ \ 3\ \ 17\,0 \\
\hline
Y = 22\ 11\ 12\ 21\ 23\ 23\ 23\ 17\ 20\ 17\ 4
\end{array}
$$

В текстовом виде получаем:

$$Y = \text{ХКЛФЦЦЦРУРД}$$

3.2.2. Расшифрование

Расшифруем последовательность "ХКЛФЦЦЦРУРД". Так как для построения гаммы необходимо знание открытого текста, постепенно расшифровываем шифртекст и сразу подставляем в гамму:

$$X_1 = Y_1 - \gamma_1 = Y_1 - K$$

$$X_i = Y_i - \gamma_i = Y_i - X_{i-1}, i = \overline{2, l}$$

Действуя таким образом, получаем:

$$
\begin{array}{l}
Y = 22\ 11\ 12\ 21\ 23\ 23\ 23\ 17\ 20\ 17\ 4 \\
\gamma = \ 11 \\
\hline
X = \mathbf{11}
\end{array}
$$

$$
\begin{array}{l}
Y = 22\ 11\ 12\ 21\ 23\ 23\ 23\ 17\ 20\ 17\ 4 \\
\gamma = \ 11\ \mathbf{11} \\
\hline
X = 11
\end{array}
$$

$$\dots$$

$$
\begin{array}{l}
Y = 22\ 11\ 12\ 21\ 23\ 23\ 23\ 17\ 20\ 17\ 4 \\
\gamma = \ 11\ 11\ \mathbf{0}\ 12\ 9\ 14\ 9\ 14\ 3\ 17\ 0 \\
\hline
X = 11\ \mathbf{0}\ 12\ 9\ 14\ 9\ 14\ 3\ 17\ 0\ 4
\end{array}
$$

Получили исходное слово "Калининград".

3.3. Самоключ Виженера по шифртексту

В данном режиме гамма представляет собой конкатенацию односимвольного ключа и шифртекста, за исключением его последнего символа. Ясно, что при зашифровании гамма вырабатывается непосредственно в ходе процесса. При расшифровании же, напротив, гамма находится тривиально.

В качестве ключа выберем тот же символ **"К"**. Представим текст и ключ в виде элементов кольца классов вычетов $\mathbb{Z}_m$:

$$X = (11\ 0\ 12\ 9\ 14\ 9\ 14\ 3\ 17\ 0\ 4)$$

$$K = (11)$$

3.3.1. Зашифрование

При зашифровании начинаем процесс с первого символа:

$$
\begin{array}{l}
X = 11\ 0\ 12\ 9\ 14\ 9\ 14\ 3\ 17\ 0\ 4 \\
\gamma = \ 11 \\
\hline
Y = \mathbf{22}
\end{array}
$$

$$
\begin{array}{l}
X = 11\ \ 0\ \ 12\ 9\ 14\ 9\ 14\ 3\ 17\ 0\ 4 \\
\gamma = \ 11\ \mathbf{22} \\
\hline
Y = 22
\end{array}
$$

$$\dots$$

$$
\begin{array}{l}
X = 11\ \ 0\ \ 12\ \ 9\ \ 14\ \ 9\ \ 14\ \ 3\ \ 17\ 0\ 4 \\
\gamma = \ 11\ 22\ \mathbf{22}\ \ 1\ \ 10\ 24\ \ 0\ \ 14\ 17\ 1\ 1 \\
\hline
Y = 22\ \mathbf{22}\ \ 1\ \ 10\ 24\ \ 0\ \ 14\ 17\ \ 1\ \ 1\ 5
\end{array}
$$

В текстовом виде получаем:

$$Y = \text{ХХБЙЧАНРББЕ}$$

3.3.2. Расшифрование

Расшифруем последовательность "ХХБЙЧАНРББЕ". При расшифровании сформировать гамму можно сразу:

$$\gamma = (K, Y_1, ..., Y_{l-1})$$

$$\gamma = \text{КХХБЙЧАНРББ}$$

Далее вычитаем из символов шифртекста символы гаммы по модулю:

$$
\begin{array}{l}
Y = 22\ 22\ \ 1\ \ 10\ 24\ \ 0\ \ 14\ 17\ \ 1\ \ 1\ 5 \\
\gamma = \ 11\ 22\ 22\ \ 1\ \ 10\ 24\ \ 0\ \ 14\ 17\ 1\ 1 \\
\hline
X = 11\ \ 0\ \ 12\ \ 9\ \ 14\ \ 9\ \ 14\ \ 3\ \ 17\ 0\ 4
\end{array}
$$

Получаем исходное слово "Калининград".

4. Программная реализация шифра Виженера

Опишем особенности программной реализации шифра Виженера.

Реализация на языке Python **доступна по ссылке** на GitHub.

4.1. Шифр Виженера с повторением ключа

На вход программа получает строку, состоящую из символов английского или русского алфавитов, а также ключ (короткое слово того же алфавита). Прочие символы не допускаются. Производится проверка пользовательского ввода: все символы входной последовательности и ключа принадлежат выбранному алфавиту. Пробелы в начале и в конце строк не учитываются. Все символы приводятся к верхнему регистру. Примеры зашифрования и расшифрования представлены на рисунках 1-4.

Видно, что после расшифрования зашифрованных перед этим последовательностей программа вывела исходный текст, что является признаком правильной работы. Также полученные шифротексты совпадают с теми, что были составлены вручную в разделе 3.

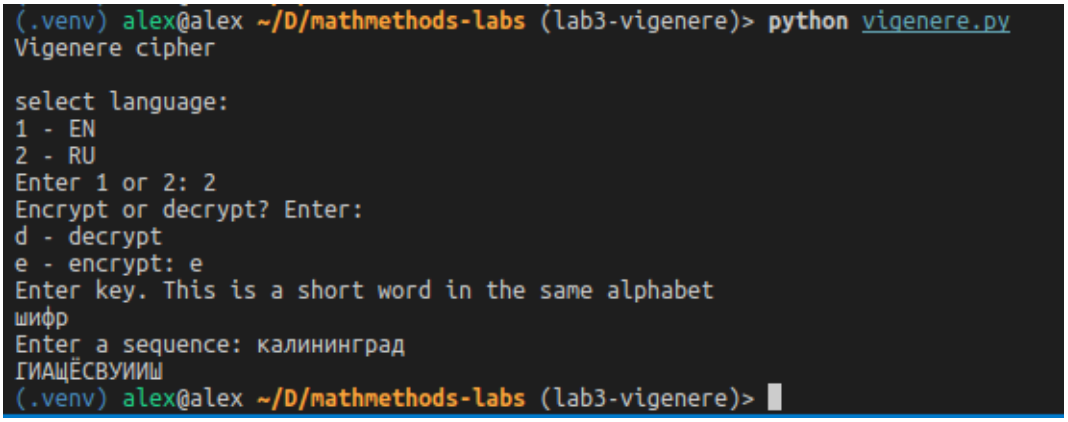

Рис. 1 – Пример зашифрования шифром Виженера с повторением ключа

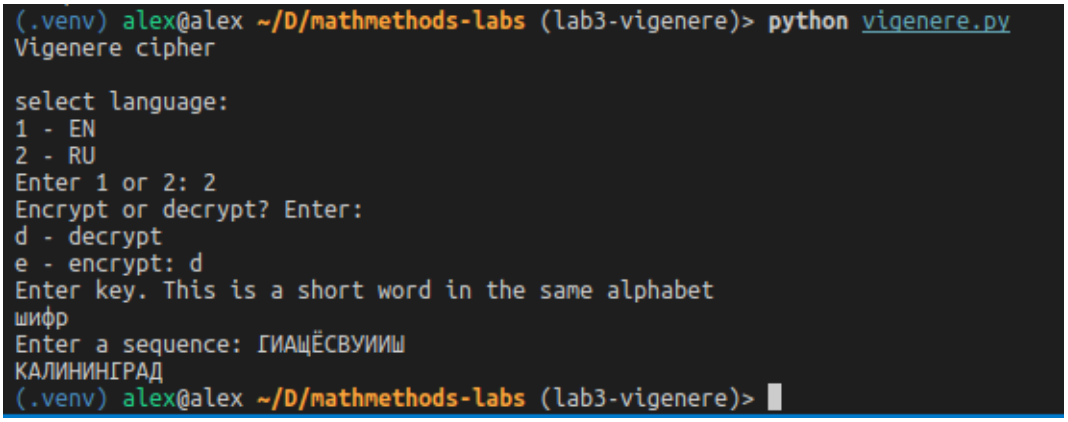

Рис. 2 – Пример расшифрования шифром Виженера

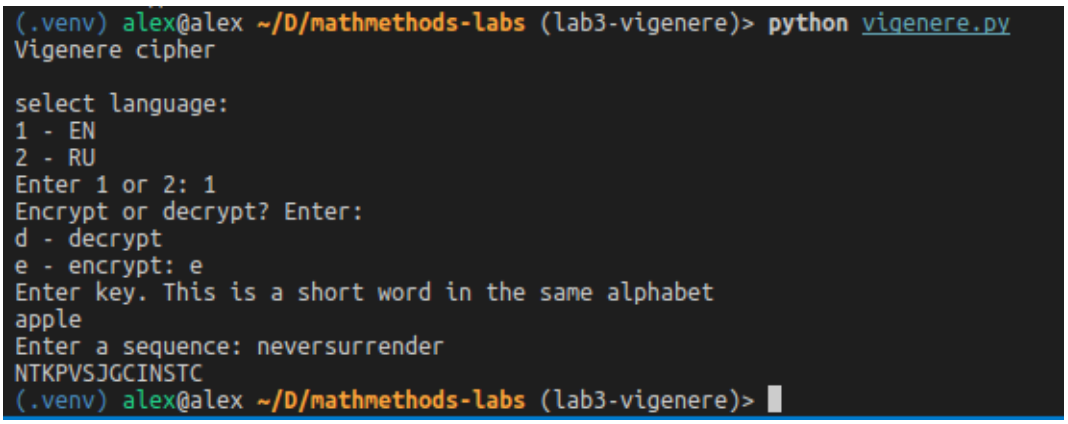

Рис. 3 – Пример зашифрования шифром Виженера на английском языке

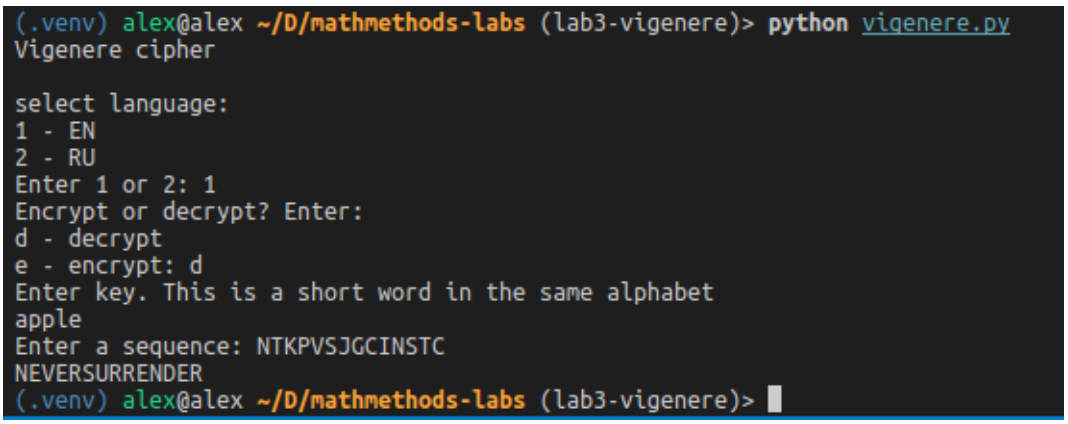

Рис. 4 – Пример расшифрования шифром Виженера на английском языке

4.2. Самоключ Виженера по открытому тексту

Примеры зашифрования и расшифрования представлены на рисунках 5-6.

На вход программа получает строку, состоящую из символов английского или русского алфавитов, а также ключ (символ того же алфавита). Прочие символы не допускаются. Производится проверка пользовательского ввода: все символы входной последовательности принадлежат выбранному алфавиту. Пробелы в начале и в конце строк не учитываются. Все символы приводятся к верхнему регистру.

Видно, что после расшифрования зашифрованных перед этим последовательностей программа вывела исходный текст, что является признаком правильной работы. Также полученные шифротексты совпадают с теми, что были составлены вручную в разделе 3.

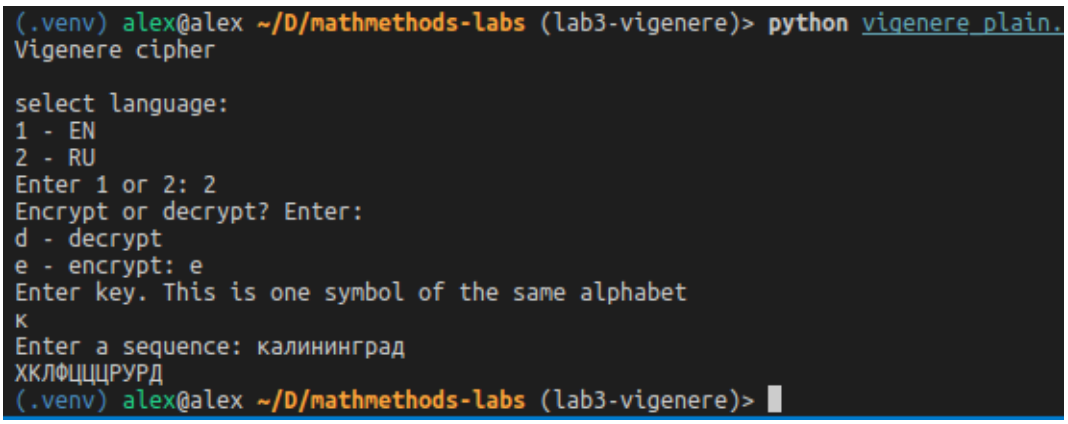

Рис. 5 – Пример зашифрования методом самоключа Виженера по открытому тексту

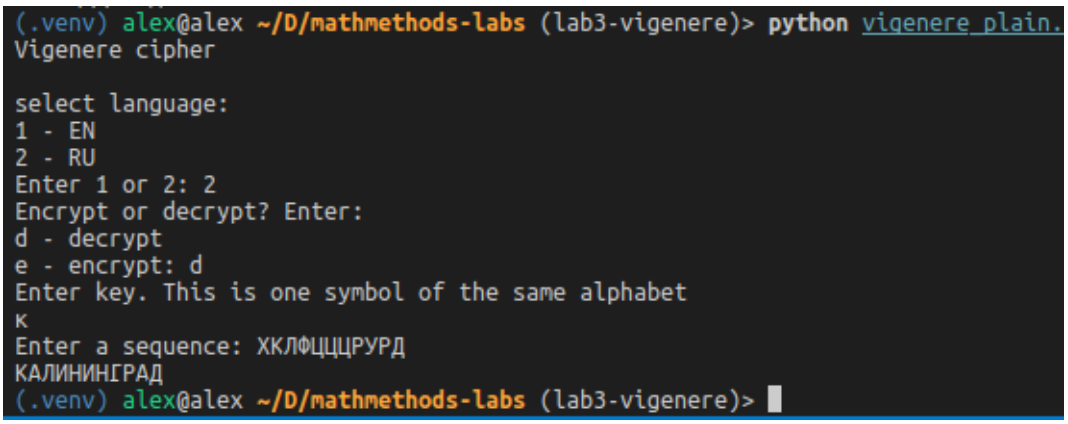

Рис. 6 – Пример расшифрования методом самоключа Виженера по открытому тексту

4.3. Самоключ Виженера по шифртексту

Примеры зашифрования и расшифрования представлены на рисунках 7-8.

На вход программа получает строку, состоящую из символов английского или русского алфавитов, а также ключ (символ того же алфавита). Прочие символы не допускаются. Производится проверка пользовательского ввода: все символы входной последовательности принадлежат выбранному алфавиту. Пробелы в начале и в конце строк не учитываются. Все символы приводятся к верхнему регистру.

Видно, что после расшифрования зашифрованных перед этим последовательностей программа вывела исходный текст, что является признаком правильной работы. Также полученные шифротексты совпадают с теми, что были составлены вручную в разделе 3.

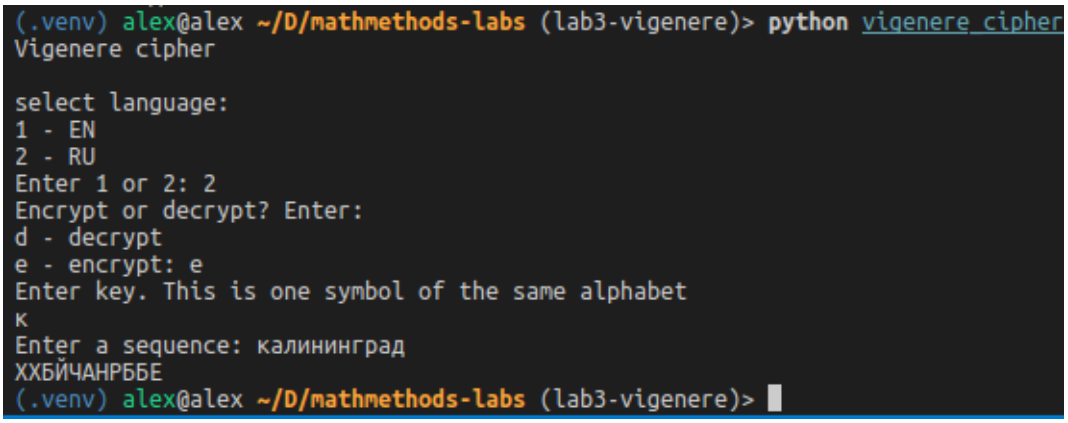

Рис. 7 – Пример зашифрования методом самоключа Виженера по шифртексту

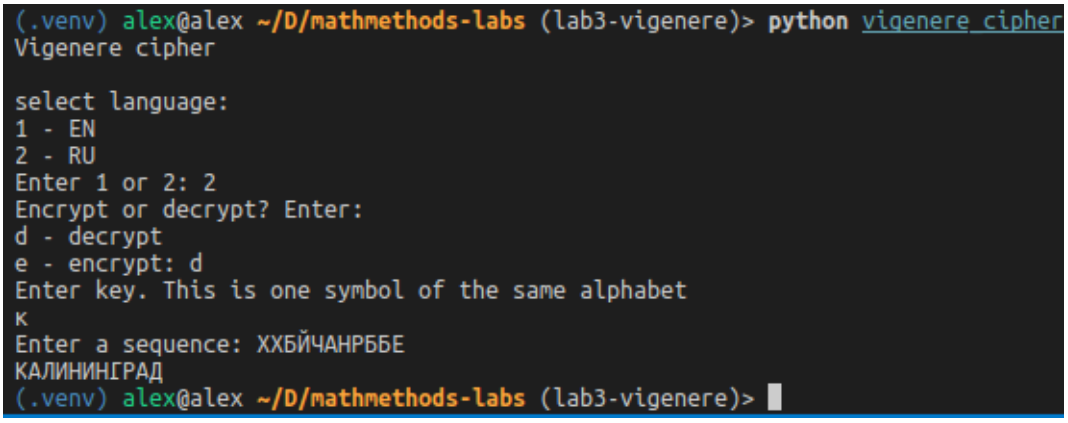

Рис. 8 – Пример расшифрования методом самоключа Виженера по шифртексту

5. Примеры криптоанализа

Попробуем взломать шифр Виженера методом индекса совпадений.

Выберем текст: *Невидима и свободна! Невидима и свободна! Пролетев по своему переулку, Маргарита попала в другой, пересекавший первый под прямым углом. Этот заплатанный, заштопанный, кривой и длинный переулок с покосившейся дверью нефтелавки, где кружками продают керосин и жидкость от паразитов во флаконах, она перерезала в одно мгновение и тут усвоила, что, даже будучи совершенно свободной и невидимой, все же и в наслаждении нужно быть хоть немного благоразумной. Только каким-то чудом затормозившись, она не разбилась насмерть о старый покосившийся фонарь на углу. Увернувшись от него, Маргарита покрепче сжала щетку и полетела помедленнее, вглядываясь в электрические провода и вывески, висящие поперек тротуара.* <...> М. Булгаков.

Удалим из текста пробелы и прочие символы, оставим лишь буквы русского алфавита: НЕВИДИМАИСВОБОДНАНЕВИДИМАИСВОБОДНАПРОЛЕТЕВПОСВОЕМУ **...**

В качестве ключа выберем слово **"шифр"**.

Шифртекст: ЁНЦЩЬСБРБЪЦЯЩЧШЮШЦЩТБМЭЭШСЁТЖЙГФЁИДБЖФЩГЭКДЯЙКГХ **...**

Процесс зашифрования показан на рисунке 9.

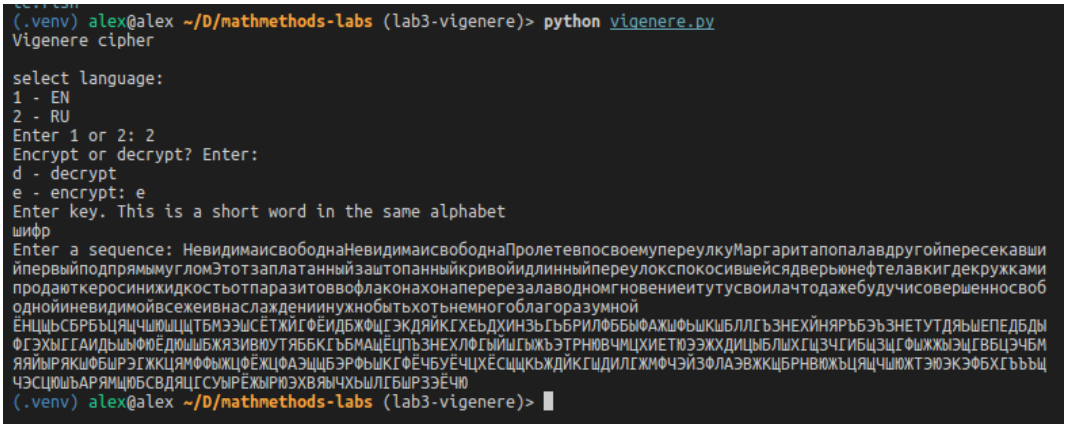

Рис. 9 – Зашифрование большого текста

Теперь будем применять индекс совпадения для определения длины ключа. Известно, что для текста на русском языке (в том числе зашифрованного моноалфавитным шифром) индекс совпадений составляет 0.0553, а для случайного набора символов – 0.03125 [5]. Реализуем алгоритм, перебирающий длину ключа и рассчитывающий индекс совпадений в каждом случае. Результат работы приведён на рисунке 10.

Видно, что при любой длине ключа значение индекса совпадений не превышает 0.39, лишь в случае длины ключа 4 получаем значение 0.423. Отсюда делаем вывод, что **длина ключа равна четырём**.

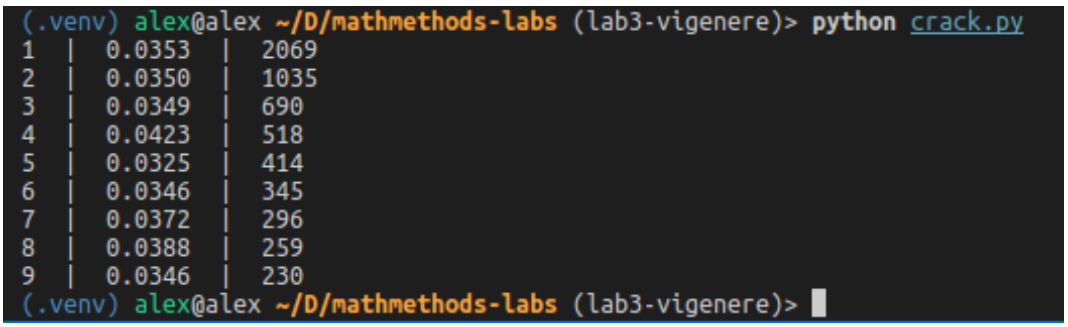

Рис. 10 – Результат работы подбирающего алгоритма. *Названия колонок: длина ключа, индекс совпадений, количество символов в выбранной подстроке*

Остаётся взломать четыре сдвиговых шифра. Так как длина ключа мала, просто переберём все возможные варианты, расшифровывая, к примеру, первые 12 символов шифртекста. Далее находим среди полученных текстов осмысленный (рисунок 11).

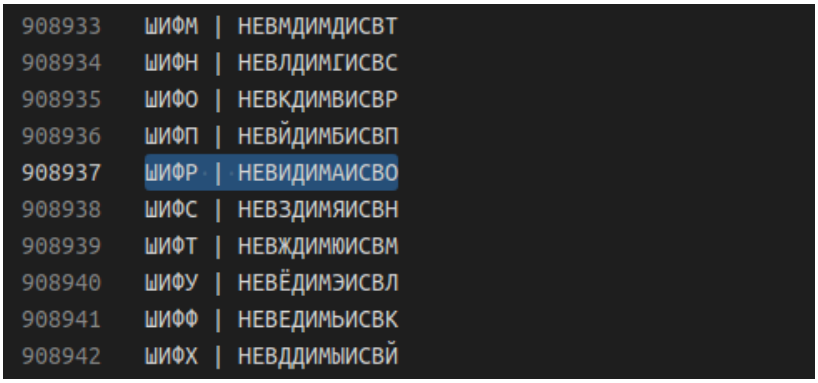

Рис. 11 – Поиск открытого текста

5.1. Шифр Виженера с самоключом

Атаке, проиллюстрированной выше, подвержен шифр Виженера в режиме повторения ключа. Две другие разновидности этого шифра таким способом взломать не удастся. Однако при более внимательном рассмотрении можно заметить, что разновидности шифра Виженера с самоключом взламываются даже проще, чем было рассмотрено ранее.

В случае самоключа Виженера по открытому тексту злоумышленник имеет только шифртекст. Для восстановления гаммы по шифртексту необходимо знать ключ, который представляет собой символ алфавита. Таким образом, злоумышленнику достаточно перебрать все символы алфавита, чтобы успешно дешифровать переданное сообщение.

Ещё более тривиальным для криптоанализа оказывается самоключ Виженера по шифртексту. злоумышленник, получив шифртекст, может сразу же восстановить всё сообщение, кроме первого символа. Обратимся к примеру из пункта 3.3 для иллюстрации.

Шифртекст $Y = (22\ 22\ 1\ 10\ 24\ 0\ 14\ 17\ 1\ 1\ 5)$, гамма $\gamma = (...\ 22\ 22\ 1\ 10\ 24\ 0\ 14\ 17\ 1\ 1)$

$$
\begin{array}{r}
Y = 22\ 22\ \ 1\ \ 10\ 24\ \ 0\ \ 14\ 17\ \ 1\ \ 1\ 5 \\
\gamma = \ ...\ 22\ 22\ \ 1\ \ 10\ 24\ \ 0\ \ 14\ 17\ 1\ 1 \\
\hline
X = ...\ \ 0\ \ 12\ \ 9\ \ 14\ \ 9\ \ 14\ \ 3\ \ 17\ 0\ 4
\end{array}
$$

Видим, что даже без перебора ключа злоумышленник способен восстановить почти всё слово: "...АЛИНИНГРАД".

6. Выводы о проделанной работе

В данной работе мы подробно изучили шифры гаммирования на примере шифра Виженера. Для выработки гаммы рассматривались три режима: повторение ключа, самоключ Виженера по открытому тексту и самоключ Виженера по шифртексту. Мы научились реализовывать алгоритмы зашифрования и расшифрования с помощью шифра Виженера на языке Python в каждом из режимов, а также успешно применили к шифру методы криптоанализа (в том числе индекс совпадений).

В результате было обнаружено, что шифр Виженера подвержен атаке с индексом совпадений, особенно при малом размере ключа, что позволяет свести дешифрование к взлому небольшого количества сдвиговых шифров. Две другие разновидности шифра Виженера оказались нестойкими.

На основе проведённого криптоанализа можно сделать вывод, что шифр Виженера в рассматриваемой реализации не является достаточно криптостойким для защиты ценной информации. Однако значение шифра Виженера и шифров гаммирования в целом нельзя недооценить: на их основе был разработан шифр Вернама, обладающий абсолютной криптографической стойкостью.

Список использованных источников

[1] Википедия. Гаммирование — Википедия, свободная энциклопедия. — 2020. — [Онлайн; загружено 18 мая 2022]. Access mode: https://ru.wikipedia.org/?curid=389823&oldid=111027819.

[2] Bruen Aiden A., Forcinito Mario. Cryptography, information theory, and error-correction: a handbook for the 21st century. — Hoboken, N.J : Wiley-Interscience, 2005. — ISBN: 9780471653172.

[3] @NeverWalkAloner Алексей. Классический криптоанализ. — Habr, 2015. — Nov. — Режим доступа: https://habr.com/ru/post/271257/.

[4] Полянин Миша. Почему нельзя взломать шифр Вернама - Журнал "код". — 2022. — Apr. — Режим доступа: https://thecode.media/vernam/.

[5] Википедия. Индекс совпадений — Википедия, свободная энциклопедия. — 2021. — [Онлайн; загружено 25 мая 2022]. Access mode: https://ru.wikipedia.org/?curid=2700934&oldid=118191993.