Учреждение образования

«БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ»

Криптографические шифры на основе подстановки символов

Студент: Валдайцев А. Д.

ФИТ 3 курс 5 группа

Преподаватель: Савельева М. Г.

Минск 2023

# Моноалфавитный подстановочный шифр

Одним из самых простых шифров является моноалфавитный шифр подстановки. В данных шифрах каждый символ исходного алфавита заменяется на один и тот же символ также из этого алфавита.

Для наглядной демонстрации шифра простой замены достаточно выписать под заданным алфавитом тот же алфавит, но в другом порядке или, например, со смещением. Записанный таким образом алфавит называют алфавитом замены.

Максимальное количество ключей для любого шифра этого вида не превышает *N!*, где *N* – количество символов в алфавите. Для математического описания криптографического преобразования предполагаем, что зашифрованная буква *ay* (*ay* ∈ *Ci*), соответствующая символу *ax* (*ax* ∈ *Mi*), находится на позиции

(1.1)

где *x*, *y* – индекс (порядковый номер, начиная с 0) символа в используемом алфавите; *k* – ключ. Для расшифрования сообщения *Ci* необходимо произвести расчеты, обратные выражению (1.1), т. е.

(1.2)

Ключом в моноалфавитном подстановочном шифре является число *k*.

# Зашифрование моноалфавитным шифром

В качестве алфавита выбран немецкий язык, а в качестве ключа *k* в формуле (1.1) выбрано число *k* = 7. В качестве открытого текста на исходном языке используется текстовый документ open\_text.txt длиной более 5000 символов, представленный на рисунке 1.1:

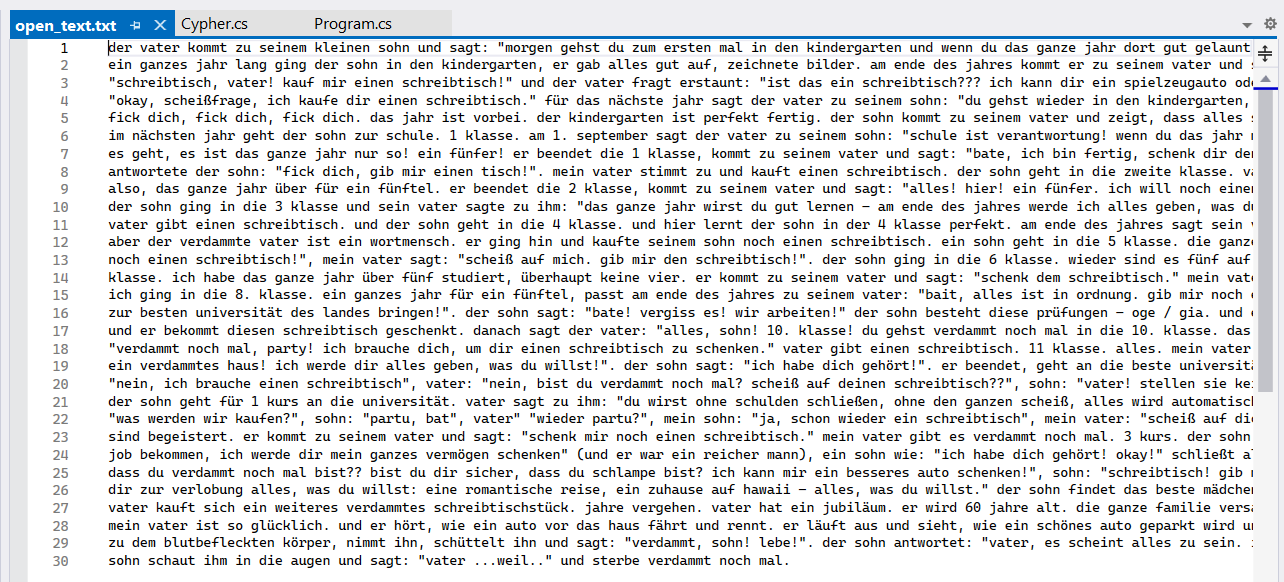


Рисунок 1.1 – Текстовый документ на исходном немецком языке

Далее для зашифрования необходимо получить индексы всех символов исходного текста и подставить индекс символа *x*, ключ *k* = 7 и мощность алфавита *N* = 30 в формулу (1.1). Для зашифрования текста с помощью моноалфавитного подстановочного шифра реализована следующая функция, представленная на рисунке 1.2:

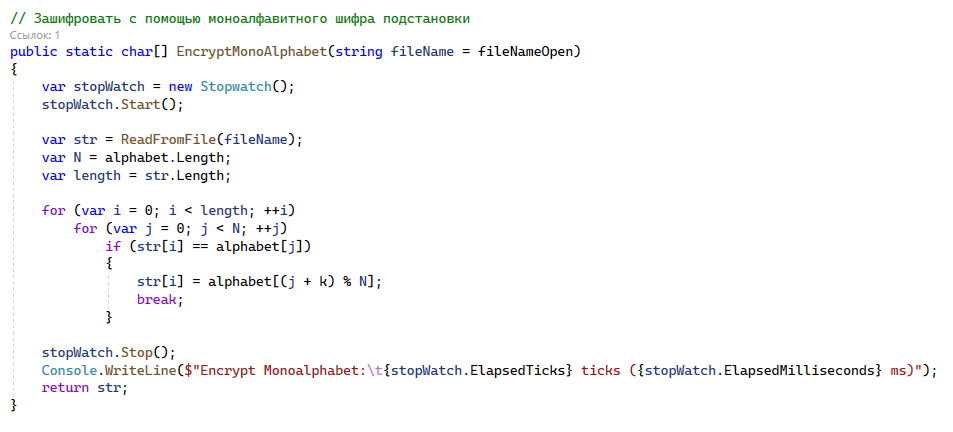


Рисунок 1.2 – Функция зашифрования моноалфавитным шифром

Данная функция записывает полученный шифротекст в файл encrypt\_monoalphabet.txt. При ручной проверке зашифрования можно убедиться, что алгоритм работает корректно: например, символ *d*, имеющий индекс 4, заменяется на символ *k* с индексом 11, а символ *u* с индексом 23 заменяется на символ *a* с индексом 0, что соответствует правильным вычислениям по формуле (1.1). Фрагмент текстового документа с зашифрованным текстом представлен на рисунке 1.3.

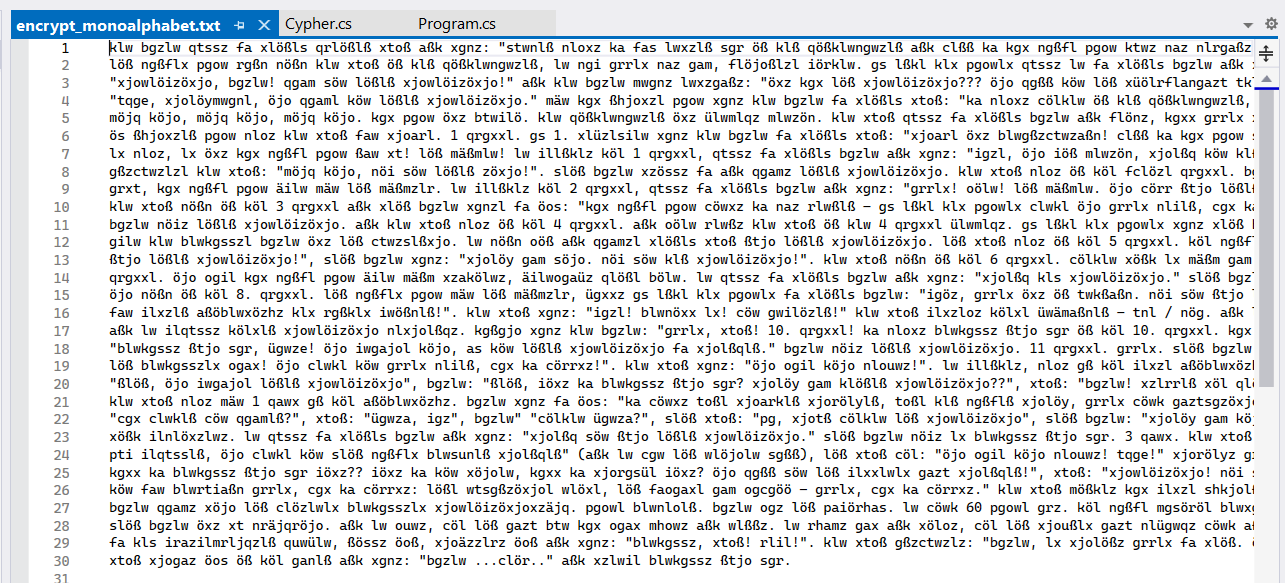


Рисунок 1.3 – Текст, зашифрованный моноалфавитным шифром

# Расшифрование моноалфавитным шифром

Для расшифрования текста, зашифрованного моноалфавитным подстановочным шифром, необходима аналогичная функция, в которой единственным изменением является использование формулы (1.2) вместо формулы (1.1). Функция, считывающая шифротекст из файла, и реализующая алгоритм его расшифрования и записи расшифрованного текста в файл, представлена на рисунке 1.4.

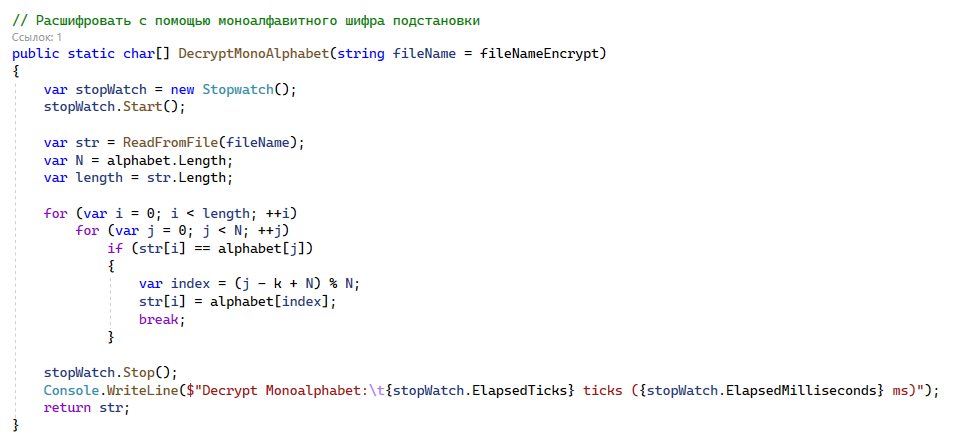


Рисунок 1.4 – Функция расшифрования моноалфавитного шифра

Текстовый документ с расшифрованным текстом представлен на рисунке 1.5.

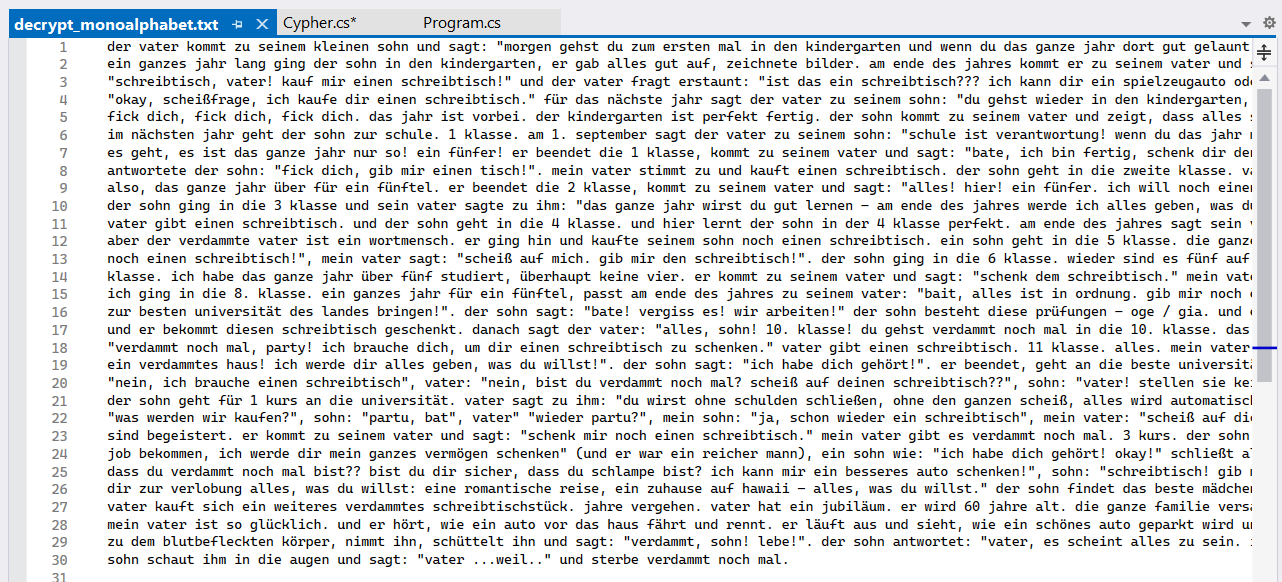


Рисунок 1.5 – Текст, расшифрованный с моноалфавитного шифра

Как видно из рисунков 1.1 и 1.5, исходный открытый текст соответствует расшифрованному тексту, что означает, что алгоритмы зашифрования и расшифрования с помощью моноалфавитного подстановочного шифра работают корректно.

# Таблица Трисемуса

Наряду с моноалфавитными подстановочными шифрами существуют полиалфавитные, которые состоят из нескольких шифров одиночной замены. Одним из полиалфавитных шифров является таблица Трисемуса.

Для зашифрования необходимо сначала заполнить таблицу подстановки, имеющую произвольные количество строк и столбцов, с одним условием: количество ячеек должно равняться мощности алфавита. В качестве алфавита можно взять не только строковые, но и специальные символы, знаки пунктуации и так далее. Саму таблицу можно заполнять либо алфавитом по порядку, либо, что более распространено, с использованием ключевого слова.

Ключевое слово записывается с начала таблицы, без повторения символов. Далее в таблицу записываются все оставшиеся символы.

Алгоритм зашифрования следующий: каждый символ открытого текста заменяется символом, расположенным под ним в таблице подстановки. Символ из последней строки заменяется символом из первой строки.

Для расшифрования необходимо аналогично заменить каждый символ на тот, который расположен над символом в таблице подстановки.

Таким образом, ключом в таблице Трисемуса является ключевое слово и размерность таблицы.

# Формирование таблицы Трисемуса

Для зашифрования необходимо сначала заполнить саму таблицу Трисемуса. Используемый алфавит, а также функция заполнения таблицы Трисемуса, представлены на рисунке 2.1.

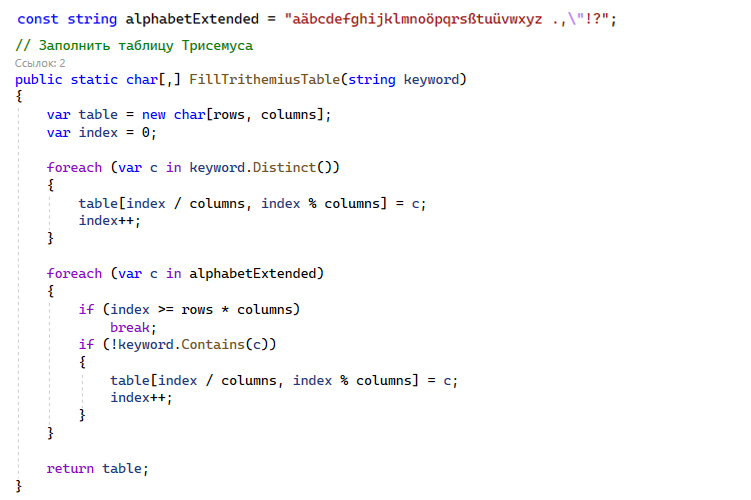


Рисунок 2.1 – Функция заполнения таблицы Трисемуса

При использовании ключевого слова *«enigma»*, размерности таблицы 6×6 и представленного на рисунке 2.1 алфавита получим следующую заполненную таблицу Трисемуса:

Таблица 1 – Таблица Трисемуса

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| e | n | i | g | m | a |
| ä | b | c | d | f | h |
| j | k | l | o | ö | p |
| q | r | s | ß | t | u |
| ü | v | w | x | y | z |
| *<пробел>* | . | , | " | ! | ? |

# Зашифрование таблицей Трисемуса

Для зашифрования необходимо использовать заполненную таблицу подстановки и заменить каждый символ исходного текста из документа open\_text.txt на символ, стоящий в следующей строке. Функция зашифрования таблицей Трисемуса представлена на рисунке 2.2.

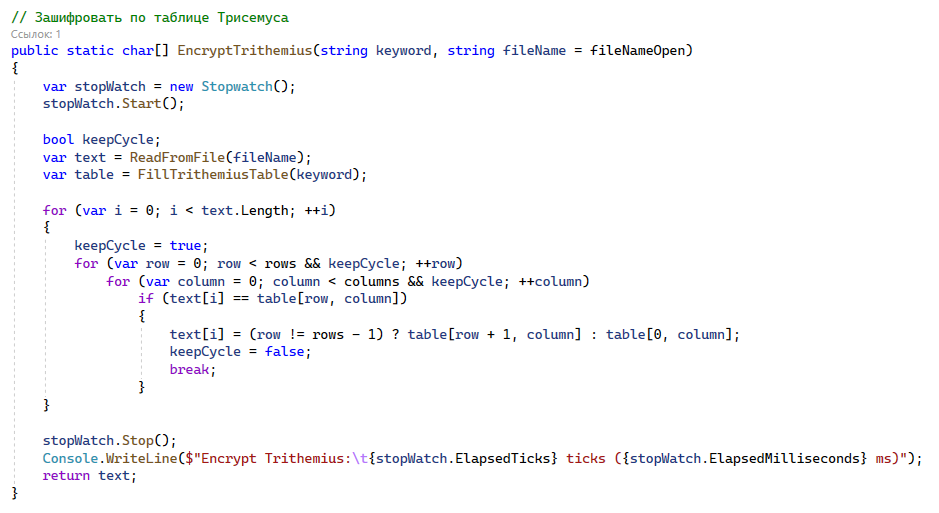


Рисунок 2.2 – Функция зашифрования таблицей Трисемуса

В результате зашифрования получаем текстовый документ encrypt\_trithemius.txt, представленный на рисунке 2.3.



Рисунок 2.3 – Зашифрованный таблицей Трисемуса текст

Также проверим корректность зашифрования таблицей Трисемуса: например, первый символ *d* в данной таблице подстановок заменяется на нижестоящий символ *o*, а символ *e* – на символ *ä*. Как видно из зашифрованного текста с рисунка 2.3, зашифрование происходит корректно.

# Расшифрование таблицы Трисемуса

Для расшифрования текста, зашифрованного с помощью таблицы Трисемуса, необходимо также построить таблицу подстановок той же размерности с таким же алфавитом и ключевым словом, и каждый символ шифротекста заменить на символ в вышестоящей строке. Для реализации расшифрования создана следующая функция, представленная на рисунке 2.4.

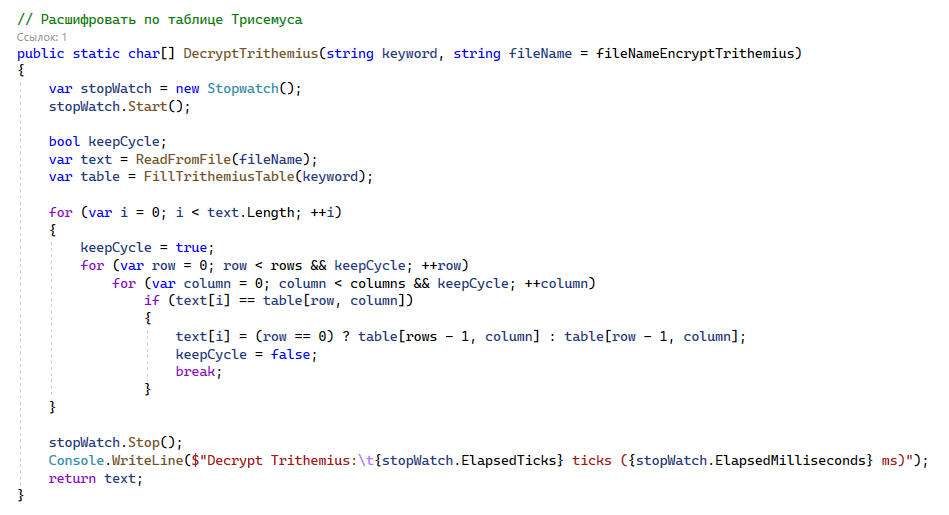


Рисунок 2.4 – Функция расшифрования таблицы Трисемуса

Аналогично при расшифровании, имея данные о ключевом слове, размерности таблицы и исходном алфавите, можно построить таблицу, идентичную таблице 1 и выполнить расшифрование шифротекста. Расшифрованный текст изображён на рисунке 2.5.

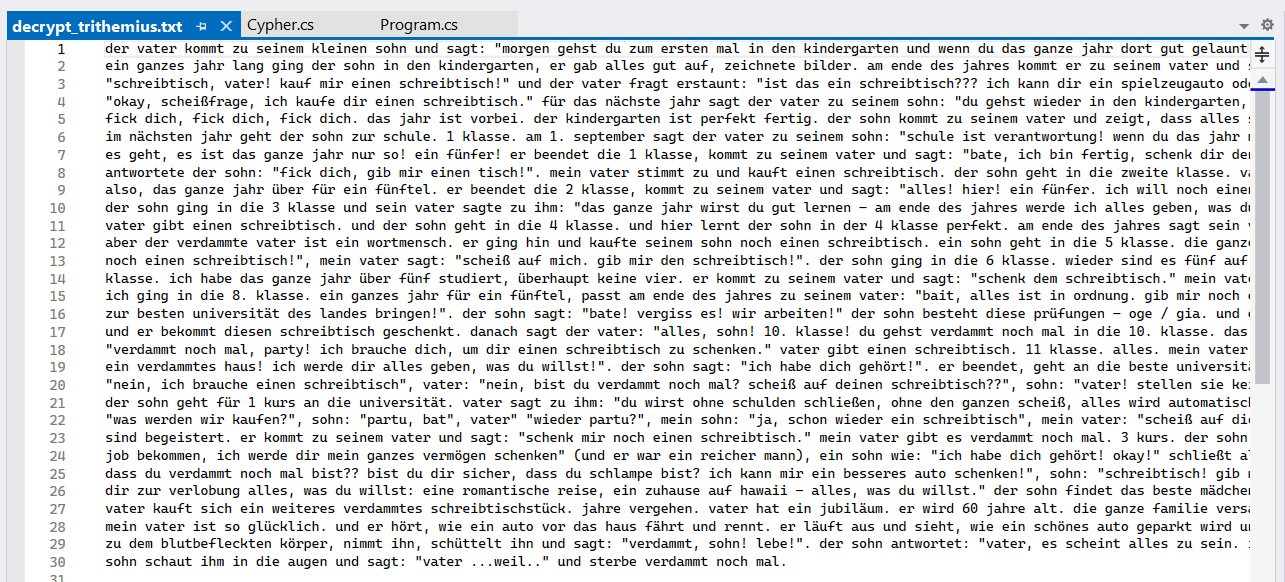


Рисунок 2.5 – Расшифрованный текст таблицы Трисемуса

Как видно из рисунка 2.5, расшифрованный текст идентичен исходному открытому тексту, что свидетельствует о том, что шифрование и расшифрование выполняются корректно.

# Гистограммы частот появления символов

Для определения криптостойкости построим гистограммы частот появления символов в открытом тексте и текстах, зашифрованных моноалфавитным подстановочным шифром и таблицей Трисемуса.

Для вычисления количества появлений символов в тексте используется функция, представленной на рисунке 3.1.

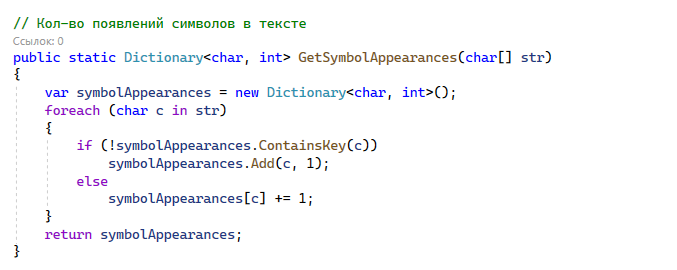


Рисунок 3.1 – Функция подсчёта количества символов

На основании количества появлений символов в тексте при известном общем количестве символов можно построить гистограммы частот появления символов в исходном тексте и шифротекстах, изображённые на рисунке 3.2.

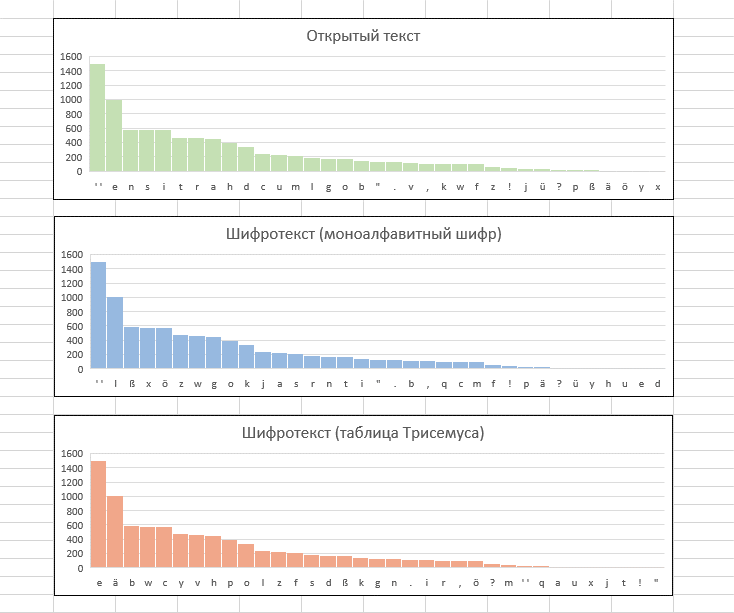


Рисунок 3.2 – Гистограммы частот появления символов в шифротекстах

Как видно из рисунка, все три гистограммы идентичны между собой. Это объясняется тем, что в подстановочных шифрах каждому символу алфавита всегда соответствует только один символ того же алфавита, записанного в другом порядке.

Таким образом, подстановочные шифры являются уязвимыми к частотному криптоанализу. Существует следующая гистограмма статистики частотности букв немецкого языка, основанная на текстовой выборке из почти 180 миллиардов символов, представленная на рисунке 3.3.

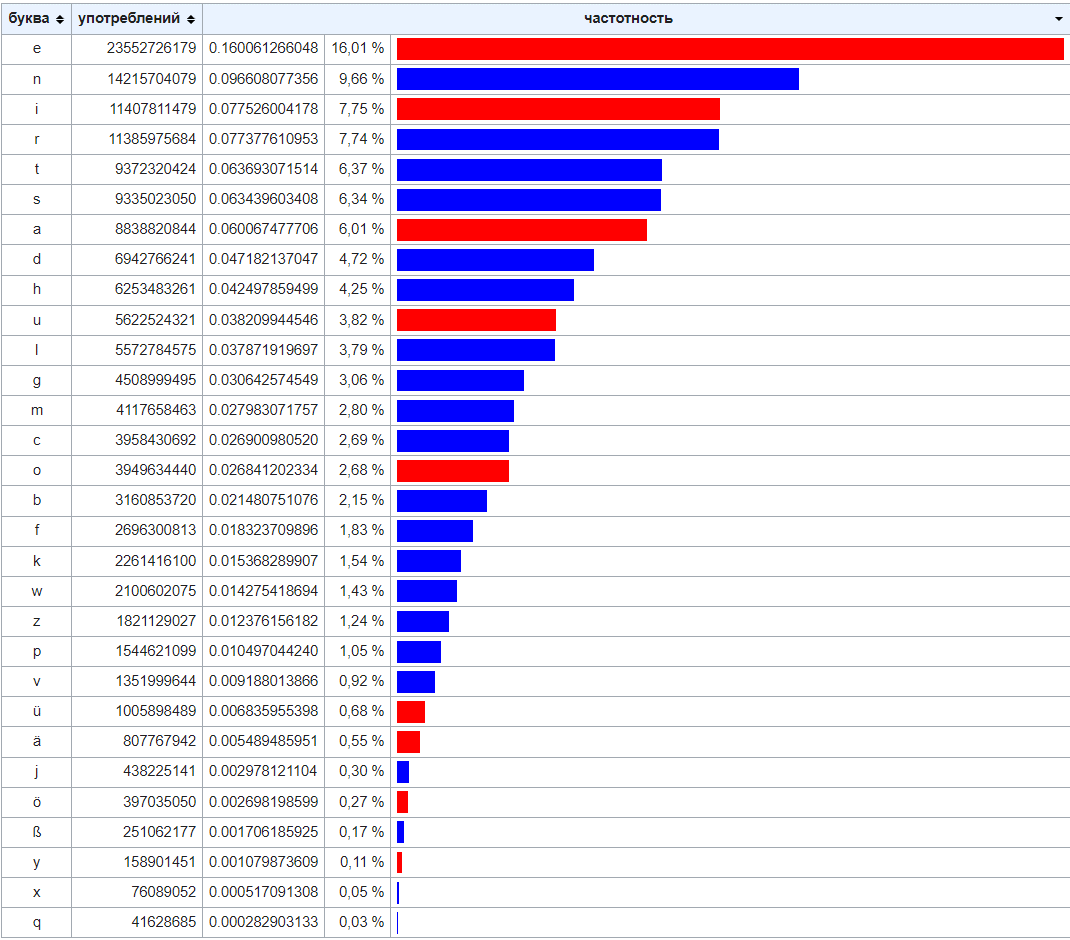


Рисунок 3.3 – Гистограмма частотности букв немецкого языка

Данная гистограмма в большой степени соотносится с гистограммой частотности символов немецкого алфавита, полученной на основании открытого текста. Следовательно, так как соотношение символов в гистограммах не меняется при шифровании подстановочными шифрами, то данные шифры являются уязвимыми к частотному криптоанализу.

# Время выполнения зашифрования и расшифрования

Для оценки времени, затраченного на выполнение операций зашифрования и расшифрования, используется объект Stopwatch, выводящий время выполнения операций в процессорном времени и в миллисекундах.

Вывод функций оценки времени представлен на рисунке 4.1.

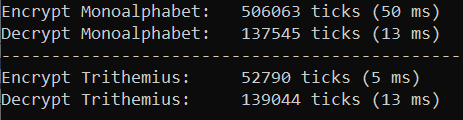


Рисунок 4.1 – Вывод функции оценки времени выполнения

Как видно из рисунка, зашифрование моноалфавитным подстановочным шифром происходит в 10 раз медленнее, чем зашифрование таблицей Трисемуса. Расшифрование обоими способами происходит за 13 миллисекунд.

Для сравнения времени выполнения операций зашифрования и расшифрования различными алгоритмами оценим время для различных входных данных – от 9 тысяч до 900 миллионов символов входного текста. График, показывающий время выполнения операций зашифрования и расшифрования методами моноалфавитного подстановочного шифра и таблицей Трисемуса представлен на рисунке 4.2.

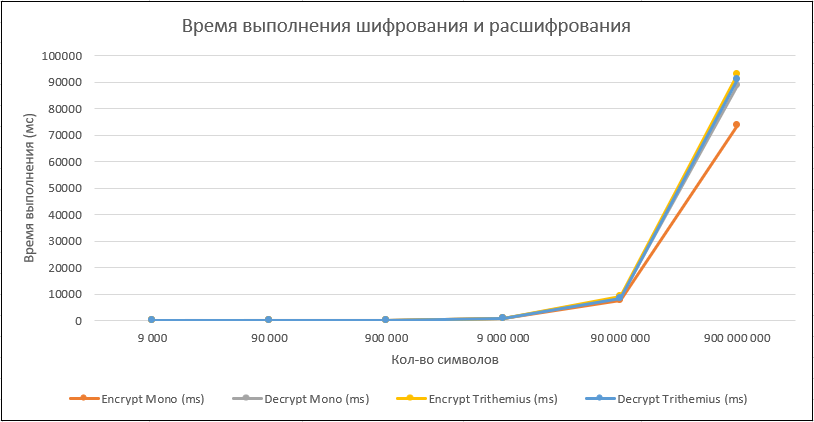


Рисунок 4.2 – Время выполнения для различных входных текстов

Для более подробного описания времени выполнения всех операций, построим таблицу, детально описывающую затраченное время для входных документов различной длины.

Таблица 2 – Время шифрования

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Кол-во символов | Зашифрование моноалфавит-ным шифром (10-3 с) | Расшифрование моноалфавит-ным шифром (10-3 с) | Зашифрова-ние таблицей Трисемуса (10-3 с) | Расшифрова-ние таблицей Трисемуса (10-3 с) |
| 9.000 | 16 | 5 | 2 | 4 |
| 90.000 | 22 | 12 | 9 | 11 |
| 900.000 | 98 | 85 | 82 | 71 |
| 9.000.000 | 751 | 815 | 863 | 758 |
| 90.000.000 | 7533 | 8633 | 9074 | 8529 |
| 900.000.000 | 73695 | 88675 | 92731 | 91068 |

Из данной таблицы можно сделать вывод, что при входных документах, количество символов в которых меньше 1 миллиона, алгоритм зашифрования таблицей Трисемуса быстрее алгоритма моноалфавитной шифровки примерно в 5 раз. Однако, чем больше символов в исходном документе, тем более медленным становится алгоритм шифрования таблицей Трисемуса, и при 1 миллиарде символов зашифрование моноалфавитным шифром быстрее таблицы Трисемуса на 26%.

Расшифрование при количестве символов менее 100 миллионов примерно на 10% быстрее для таблицы Трисемуса. При более длинных входных документах расшифрование таблицы Трисемуса начинает становиться более медленным, и для 1 миллиарда символов уже является медленнее расшифрования моноалфавитного шифра на 3%.

# Вывод

Были рассмотрены два подстановочных шифра – моноалфавитный шифр и полиалфавитный шифр таблицей Трисемуса. С точки зрения криптоанализа, оба шифра не являются криптостойкими, так как уязвимы к частотному анализу ввиду того, что каждому символу исходного текста всегда соответствует только один символ шифротекста.

При сравнении времени выполнения операций зашифрования и расшифрования для обоих шифров выясняется, что зашифрование таблицей Трисемуса быстрее зашифрования моноалфавитным шифром примерно в 5-10 раз при входных документах не длиннее 1 миллиона символов, но начинает замедляться при увеличении количества входных символов. Аналогично, расшифрование таблицы Трисемуса быстрее только при количестве символов, примерно равном 100 миллионам.

Следовательно, таблицу Трисемуса целесообразно использовать для небольших документов (менее 1 миллиона символов).