Министерство образования Республики Беларусь

Учреждение образование

«Белорусский государственный технологический университет»

Кафедра информационных систем и технологий

**Отчет к лабораторной работе**:

«Исследование криптографических шифров на основе подстановки (замены) символов»

Выполнил:

студент 3 курса 8 группы ФИТ

Ярмолик Николай Сергеевич

Проверил:

Берников В. О.

Минск 2020

**1 Теоретические сведения**

Сущность подстановочного шифрования состоит в том, что, исходный текст (из множества М) и зашифрованный текст (из множества С) основаны на использовании одного и того же или разных алфавитов, а тайной или ключевой информацией является алгоритм подстановки. Если исходить из того, что используемые алфавиты являются конечными множествами, то в общем случае каждой букве ax алфавита AM (ax∈AM) для создания сообщения Мi (Мi ∈M) соответствует буква ay или множество букв{АхС} для создания шифртекста Сi (Сi ∈С). Важно, чтобы во втором случае любые два множества (например, {АхС}b и {АхС}n, b≠n, 1 ≤ b, n, x, y ≤ N, N – мощность алфавита), используемые для замены разных букв открытого текста, не пересекались.

Если в сообщении Мi содержится несколько букв ax, то каждая из них заменяется на символ ay либо на любой из символов {АхС}. За счет этого с помощью одного ключа можно генерировать различные Сi для одного и того же Мi. Так как множества {АхС}b и {АхС}n попарно не пересекаются, то по каждому символу Сi можно однозначно определить, какому множеству он принадлежит, и, следовательно, какую букву открытого сообщения Мi он заменяет. В силу этого открытое сообщение восстанавливается из зашифрованного однозначно.

**Моноалфавитные шифры подстановки**

В данных шифрах операция замена производится только над каждым одиночным символом сообщения Мi. 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∈Сi), соответствующая символу aх (aх∈Мi), находится на позиции y∈x + k (mod N), где x, y – индекс (порядковый номер, начиная с 0) символа в используемом алфавите, k – ключ.

Для расшифрования сообщения Сi необходимо произвести расчеты обрат-ные, т. е.: х ∈ у – k (mod N).

Классический шифр подстановки – шифр Цезаря, шифр Цезаря с ключевым словом.

**Полиграммные шифры**

В таких шифрах одна подстановка соответствует сразу нескольким символам исходного текста.

Первым известным шифром этого типа является *шифр Порты*

**Омофонические шифры**

*Омофонические шифры* (*омофоническая замена*) или *однозвучные шифры подстановки* создавались с целью увеличить сложность частотного анализа шифртекстов путем маскировки реальных частот появления символов текста с помощью *омофонии*.

***Полиалфавитные шифры***

*Полиалфавитные* (или *многоалфавитные*) шифры состоят из нескольких шифров однозначной замены. Выбор варианта алфавита для зашифрования од-ного символа зависит от особенностей метода шифрования.

***Таблица Трисемуса***. В 1518 году в развитии криптографии был сделан важ-ный шаг благодаря появлению в Германии первой печатной книги по крипто-графии. Аббат Иоганнес Трисемус, настоятель монастыря в Вюрцбурге, напи-сал книгу «Полиграфия», в которой он описал ряда шифров, один из которых развивает идею *многоалфавитной подстановки*. Зашифрование осуществля-ется так: заготавливается *таблица подстановки* (так называемая «*таблица Трисемуса*» – таблица со стороной равной *N*, где *N* – мощность алфавита), где первая строка – это алфавит, вторая – алфавит, сдвинутый на один символ, и т. д. При зашифровании первая буква открытого текста заменяется на букву, стоящую в первой строке, вторая – на букву, стоящую во второй строке, и т.д. После использования последней строки вновь возвращаются к первой.

Еще раз вспомним, что криптоанализ – это раздел криптологии, занимающийся методами взлома шифров или методами организации криптографических атак на шифры.

***Атака с известным шифртекстом*** (ciphertext only attack). Предполага-ется, что противник знает алгоритм шифрования, у него имеется набор пере-хваченных шифрограмм, но он не знает секретный ключ.

Разновидности такой атаки:

*полный перебор ключей* (лобовая атака, brute force attack);

*атака по словарю*, перебор ключей по словарю (dictionary attack); применяется часто для взлома паролей;

*частотный криптоанализ* – метод взлома шифра, основывающийся на предположении о существовании зависимости между частотой появления символов алфавита в открытых сообщениях и соответствующих шифрозамен в шифрограммах.

***Атака с выбором шифртекста*** (chosen ciphertext attack). Криптоаналитик имеет возможность выбрать необходимое количество шифрограмм и получить соответствующие им открытые тексты. Он также может воспользоваться устройством расшифрования один или несколько раз для получения шифртек-ста в расшифрованном виде. Используя полученные данные, он может попы-таться восстановить секретный ключ.

***Адаптивная атака с выбором шифртекста*** (adaptive chosen ciphertext attack). Криптоаналитик имеет возможность выбирать новые шифрограммы для расшифрования с учетом того, что ему известна некоторая информация из предыдущих сообщений. Поскольку в некоторых криптографических прото-колах при получении шифрограммы, несоответствующей стандарту (содержа-щей ошибки), отправитель получает ответное сообщение, иногда с детализи-рованным описанием этапа проверки и причины возникновения ошибки. Криптоаналитик может использовать эту информацию для последовательной посылки и уточнения параметров криптосистемы.

***Атака с известным открытым текстом*** (known plaintext attack). То же, что и предыдущая, но противник для некоторых шифрограмм получает в свое распоряжение соответствующие им открытые тексты.

***Атака с выбором открытого текста*** (chosen plaintext attack). Криптоана-литик обладает некоторыми открытыми текстами и соответствующими шиф-ртекстами. Кроме того, он имеет возможность зашифровать несколько предва-рительно выбранных открытых текстов (до начала атаки).

Разновидности:

*атака на основе получения временного неконфиденциального доступа к шифрующему устройству* для получения пар открытых и тайных текстов (известны случаи реализации таких атак спецслужбами);

*атака на основе использования информации о структуре сообщений или стандартных фразах* – криптоаналитики из Блетчли-Парка (Bletchley Park) могли определить открытый текст сообщений Энигмы в зависимости от того, когда эти сообщения были посланы и как они подписывались;

*перебор ключей по словарю* (dictionary attack) – криптоаналитик шифрует слова и фразы, наличие которых предполагается в шифрограмме, с использованием различных ключей; совпадение зашифрованных слов и фраз с частями шифрограммы может говорить о взломе ключа.

***Адаптивная атака с выбором открытого текста*** (adaptive chosen plaintext attack). Криптоаналитик имеет возможность выбирать новые откры-тые тексты с учетом того, что ему известна некоторая информация из преды-дущих сообщений – он может получить пары «открытое сообщение – шифро-грамма», в т.ч. и после начала атаки.

Разновидности:

*провоцирование противника на использование в сообщениях определенных слов или фраз*; придуман англичанами: Королевские военно-воздушные силы Великобритании минировали определенные участки Северного моря, этот процесс был назван «Садоводством» (англ. «gardening»); практически сразу после этого немцами посылались зашифрованные сообщения, включающие слово «мины» и названия мест, где они были сброшены;

*дифференциальный криптоанализ* – метод вскрытия симметричных блочных шифров (и других криптографических примитивов, в частности, хеш-функций) предложен в 1990 г. израильскими специалистами Эли Бихамом и Ади Шамиром и основан на изучении разностей между шифруемыми значе-ниями на различных раундах для пары подобранных открытых сообщений при их зашифровании одним и тем же ключом;

*интегральный криптоанализ* – аналогичен дифференциальному крипто-анализу, но в отличие от него рассматривает воздействие алгоритма не на пару, а сразу на множество открытых текстов; предложен в 1997 г. Ларсом Кнудсе-ном;

*линейный криптоанализ* – предложен японским криптологом Мицуру Мацуи в 1993 г.; основан на использовании некоторых *линейных приближений*, которые означают, например, следующее: если выполняется операция XOR над некоторыми битами открытого текста, затем – над некоторыми битами шифртекста, а затем над результатами, то получается бит (или биты), который представляет собой XOR некоторых битов ключа; это и есть линейное прибли-жение, которое может быть верным с некоторой вероятностью;

*использование открытых ключей в асимметричных системах* – крипто-аналитик имеет возможность получить шифртекст, соответствующий выбранному сообщению, на основе открытого ключа.

***Атака на основе связанных*** ключей (related key attack). Криптоаналитик знает не сами ключи, а некоторые различия (соотношения) между ними; ре-альные криптосистемы используют разные ключи, связанные известным соотошением, например, для каждого нового сообщения предыдущее значение ключа увеличивается на единицу, или преобразуется на основе операции сдвига.

***Атака с выбором ключа*** (chosen key attack). Криптоаналитик задает часть ключа, а на оставшуюся часть ключа выполняет атаку на основе связанных ключей.

В основе современной криптографии лежит теория чисел. Теория чисел или высшая арифметика – раздел математики, изучающий натуральные числа и иные похожие величины. В зависимости от используемых методов в теории чисел рассматривают несколько направлений.

**2 Практическая часть**

В данной лабораторной работе необходимо было разработать авторское приложение в соответствии с целью лабораторной работы. Приложение должно

выполнять зашифрование/расшифрование текстовых документов (объемом не менее 5 тысяч знаков) созданных на основе алфавита языка в соответствии с нижеследующей таблицей вариантов задания, оценивать время выполнения операций зашифрования/расшифрования, формировать гистограммы частот появления символов для исходного и зашифрованного сообщений, при этом согласно варианту, номер 12 следует использовать шифры подстановки, а именно: Цезаря с ключевым словом, а также таблицу Трисемуса. Алфавит Русский, ключевое слово – безопасность.

Общий интерфейс программного средства приведен на рисунке 2.1.

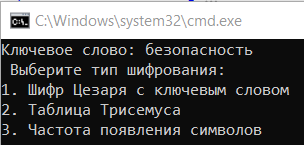


Рисунок 2.1 *–* Общий интерфейс

Далее приведен скриншот шифрования Цезаря с ключевым словом (рисунок 2.2). Ключ равен 6, ключевое слово безопасность:

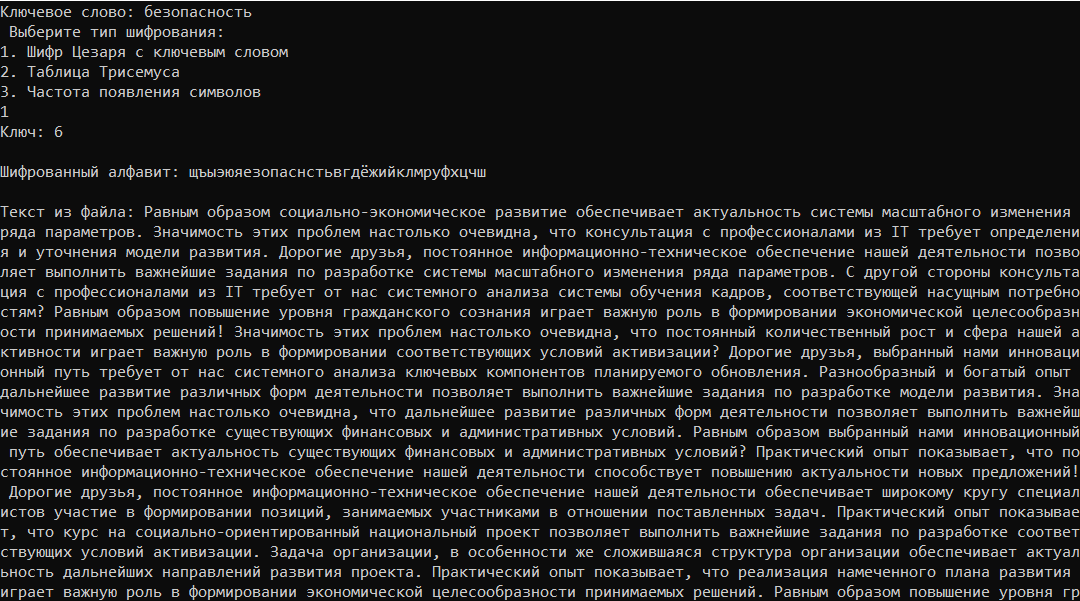


Рисунок 2.2 – Шифрование Цезарем с ключевым словом

Далее приведен скриншот шифрования текста таблицей Трисемуса с ключевым словом безопасность (Рисунок 2.3)

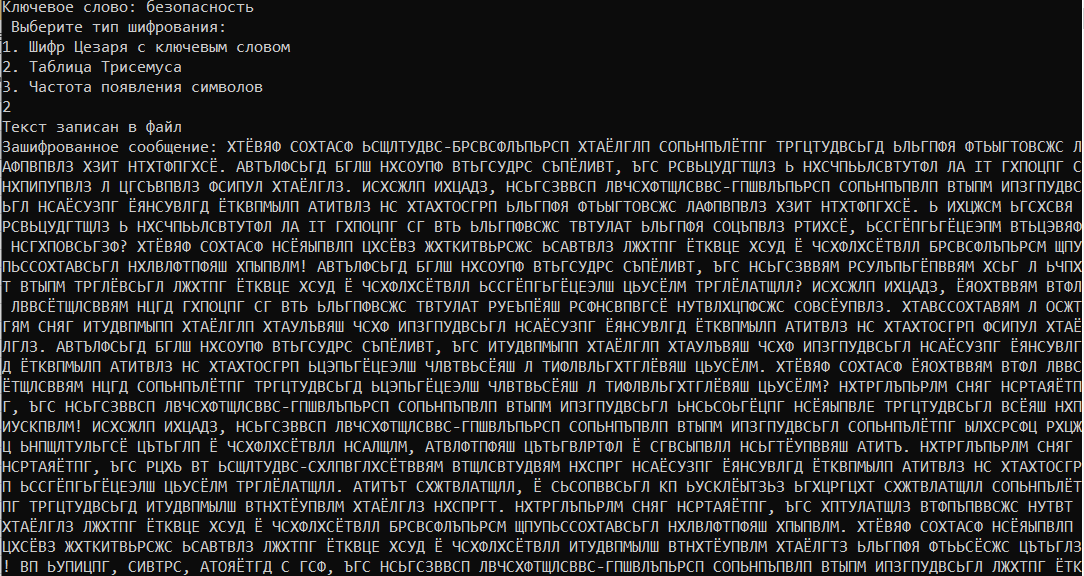


Рисунок 2.3 – Шифрование текста таблицей Трисемуса

Также в рамках лабораторной работы требовалось построить гистограммы частоты появления символов в исходном тексте (Рисунок 2.4), а также в тексте зашифрованной шифром цезаря с ключевым словом (Рисунок 2.5) и таблицей Трисемуса с ключевым словом (Рисунок 2.6)

Теперь построим гистограмму частоты символов исходного сообщения (рисунок 2.4)

Рисунок 2.4 – Гистограмма появления символов исходного текста

Далее приведена гистограмма появления символов в тексте, зашифрованном при помощи шифра Цезаря с ключевым словом (Рисунок 2.5)

Рисунок 2.5 – Текст, зашифрованный шифром Цезаря с КС

Также приведена гистограмма появления символов в тексте, зашифрованном таблицей Трисемуса с КС (Рисунок 2.6)

Рисунок 2.6 – Текст, зашифрованный таблицей Трисемуса с КС

Также в рамках данной лабораторной работы необходимо было вычислить время, требуемое для операций шифрования/расшифрования данных. Для шифра Цезаря с КС и ключом 6 потребовалось 1531мс, для шифрования и расшифрования таблицей Трисемуса понадобилось 1893мс.

Вывод: В данной лабораторной работе были получены практические навыки программной реализации алгоритмов шифрования Цезаря с КС, а также таблицей Трисемуса с КС, вычислено время шифрования/расшифрования текста согласно данным алгоритмам, а также построены гистограммы появления символов в тексте, для исходного текста и шифрованного текста