МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ

Учреждение образования «БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ

ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ»

Факультет информационных технологий

Кафедра информационных систем и технологий

Специальность Информационные системы и технологии

ОТЧЕТ ПО ЛАБОРАТОРНОЙ РАБОТЕ №2 НА ТЕМУ:

Исследование криптографических шифров на основе подстановки (замены) символов

Выполнила:

Студентка 3 курса 1 группы ФИТ

Шимчёнок Елизавета Константиновна

**Цель:** изучение и приобретение практических навыков разработки и использования приложений для реализации подстановочных шифров.

**Задачи:**

1. Закрепить теоретические знания по алгебраическому описанию, алгоритмам реализации операций зашифрования/расшифрования и оценке криптостойкости подстановочных шифров.

2. Ознакомиться с особенностями реализации и свойствами различных подстановочных шифров на основе готового программного средства (L\_LUX).

3. Разработать приложение для реализации указанных преподавателем методов подстановочного зашифрования/расшифрования.

4. Выполнить исследование криптостойкости шифров на основе статистических данных о частотах появления символов в исходном и зашифрованном сообщениях.

5. Оценить скорость зашифрования/расшифрования реализованных способов шифров.

6. Результаты выполнения лабораторной работы оформить в виде описания разработанного приложения, методики выполнения экспериментов с использованием приложения и результатов эксперимента.

**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:

**{АхС}b ∩ {АхС}n = 0**.

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 сгенерировать 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. Приведенные утверждения справедливы для следующих типов подстановочных шифров:

⦁ моноалфавитных (шифры однозначной замены или простые подстановочные);

⦁ полиграммных;

⦁ омофонических (однозвучные шифры или шифры многозначной замены);

⦁ полиалфавитных.

Одним из существенных недостатков моноалфавитных шифров является их низкая криптостойкость. Зачастую метод криптоанализа базируется на частоте встречаемости букв исходного текста.

Данная лабораторная работа посвящена анализу одного из разделов практической криптографии. В связи с этим здесь будет уместно охарактеризовать противоположность криптографии – криптоанализ. Данный термин введен американским криптографом Уильямом Ф. Фридманом в 1920 г. Еще раз вспомним, что криптоанализ – это раздел криптологии, занимающийся методами взлома шифров или методами организации криптографических атак на шифры.

**Ход работы**

В соответствии с 12 вариантом было произведено зашифрование/расшифрование текстового документа на основе русского языка. Шифры подстановки: шифр Цезаря с ключевым словом “безопасность”, таблица Трисемуса с ключевым словом “безопасность”.

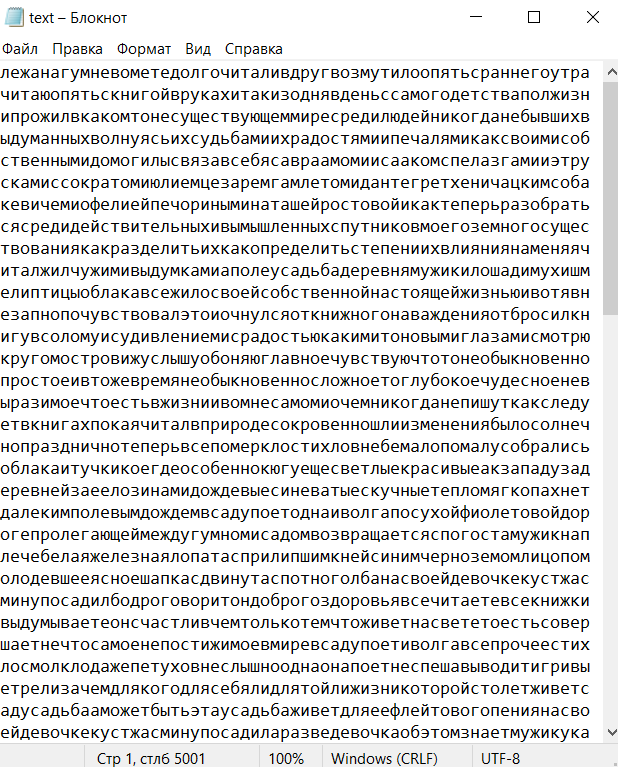


Рисунок 1 – Файл с входными данными на русском языке

Алгоритм шифра Цезаря заключается в том, чтобы записать алфавит языка, перед ним вставить ключевое слово (в данном случае “безопасность”) и удалить все повторяющиеся символы, оставляя каждого по одному. В результате получится строка с тем же количеством букв, что и в алфавите (в данном случае 33), только немного перемешанная (алфавит подстановки). Затем поиск каждого символа сообщения осуществляется в полученной строке в соответствии с его индексом в алфавите.

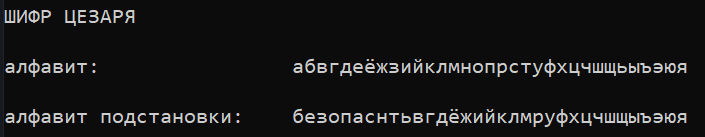


Рисунок 2 – Алфавит и алфавит подстановки

Результаты зашифровки и расшифровки записываются в файлы.

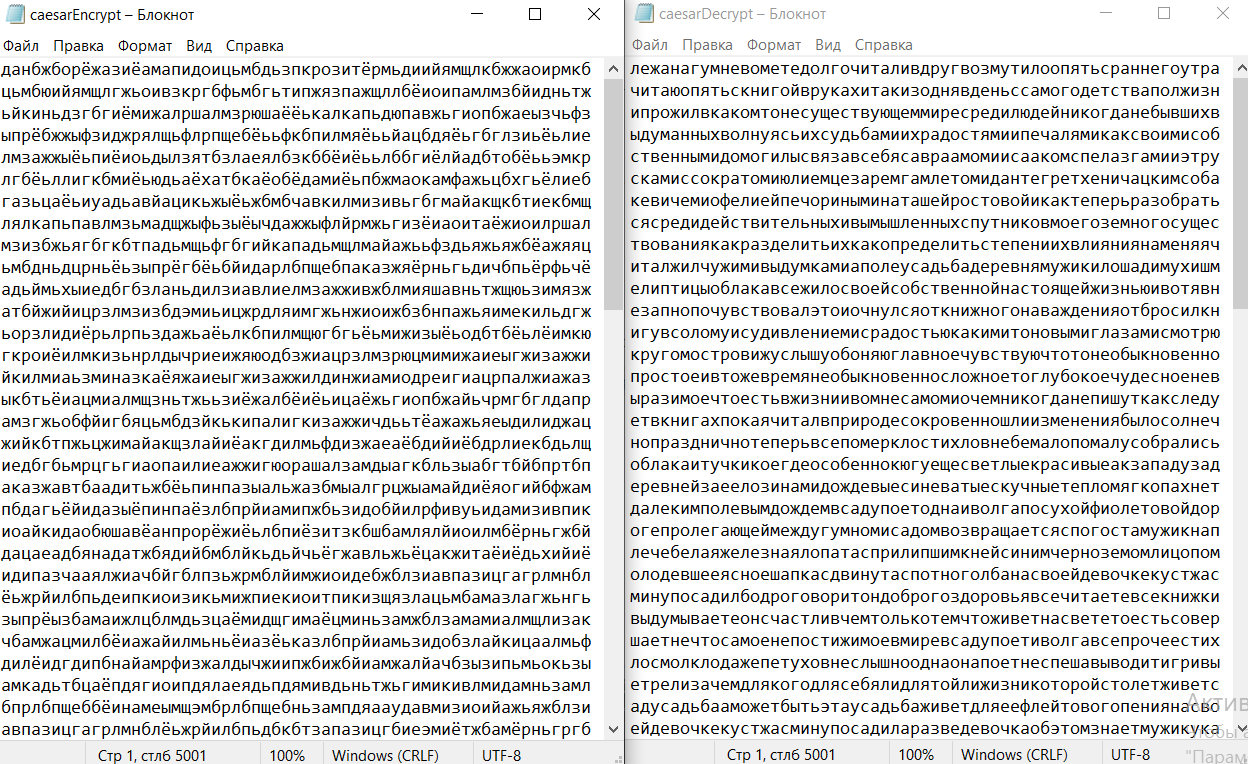


Рисунок 3 – Результаты зашифровки/расшифровки

Таблица Трисемуса представляет собой таблицу 8х4 или 4х8, в которую буквы записываются в таком же порядке, как и в алгоритме шифра Цезаря. Зашифровка сообщения проводится путем поиска требующегося символа и взятием буквы, путем спуска вниз на один элемент.

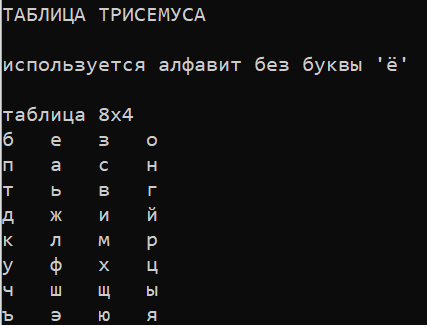


Рисунок 4 – Созданная таблица Трисемуса

Результаты зашифровки и расшифровки записываются в файл.

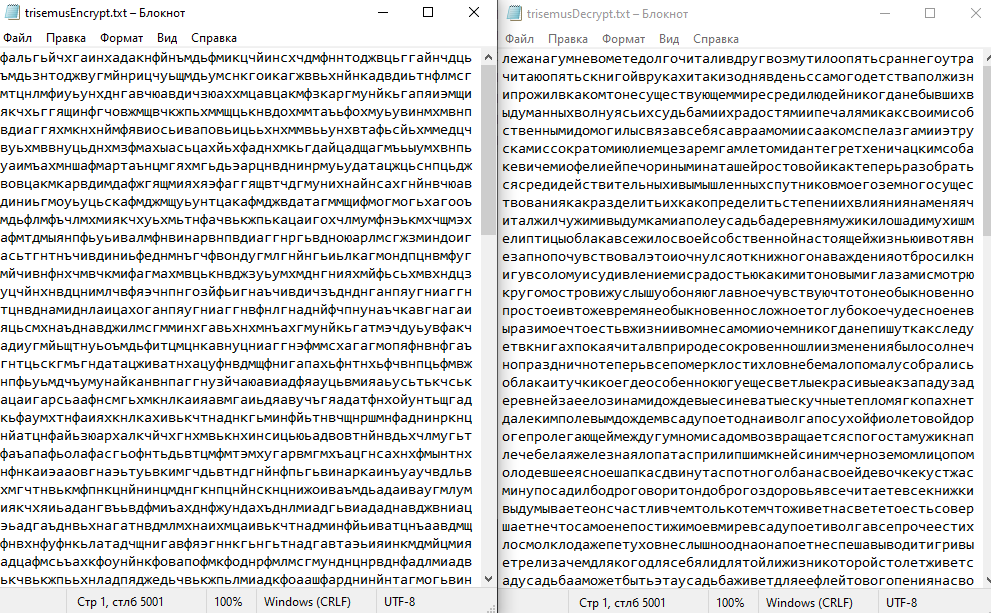


Рисунок 5 – Результаты зашифровки/расшифровки

Результат частотного анализа появления символов для исходного и зашифрованных сообщений продемонстрирован на гистограммах:

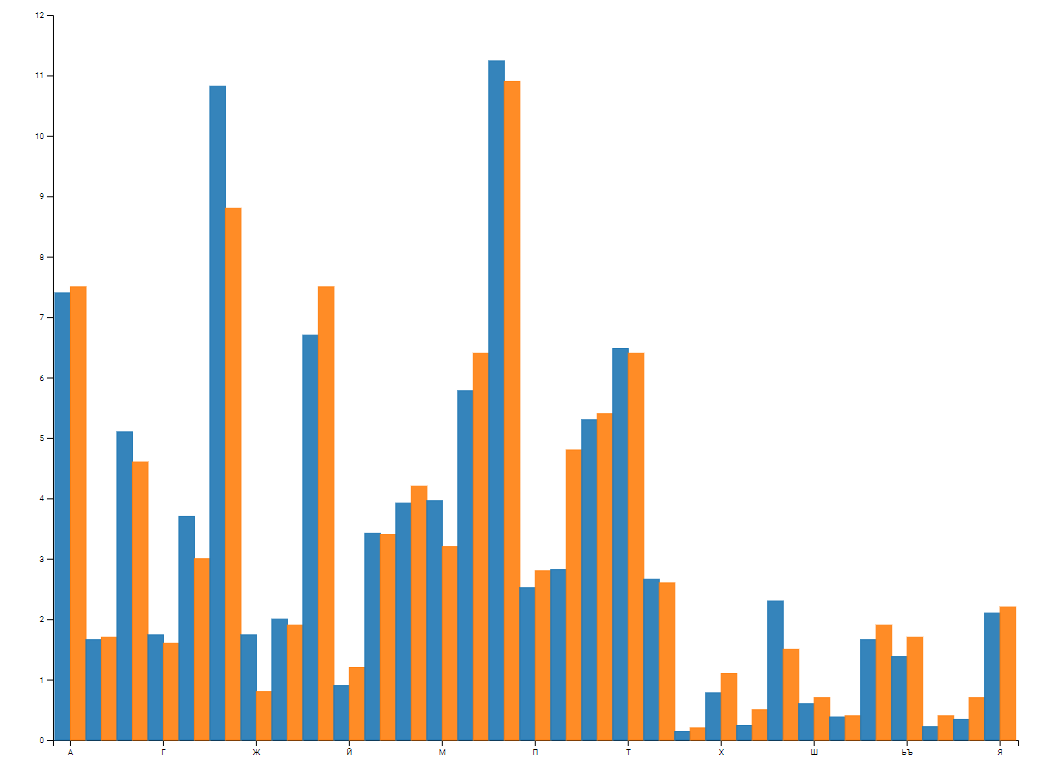


Рисунок 6 – Гистограмма для текста исходного документа

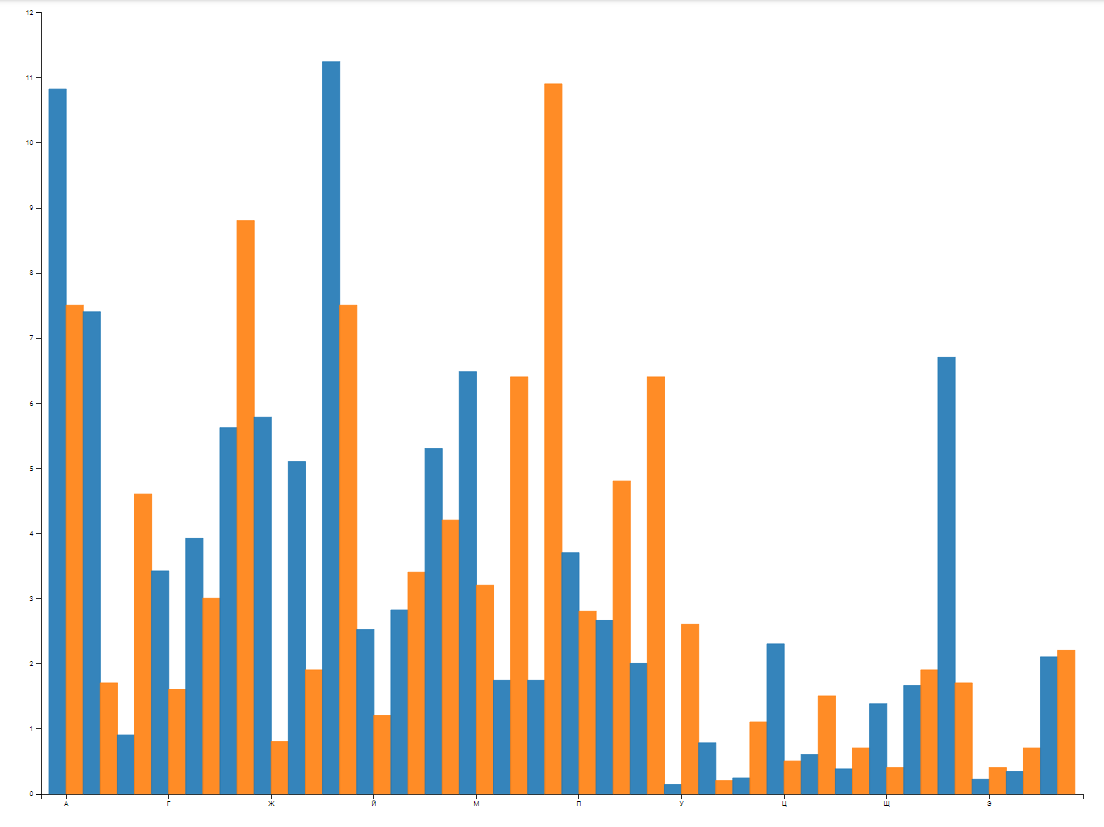


Рисунок 7 – Гистограмма для зашифрованного текста методом Цезаря

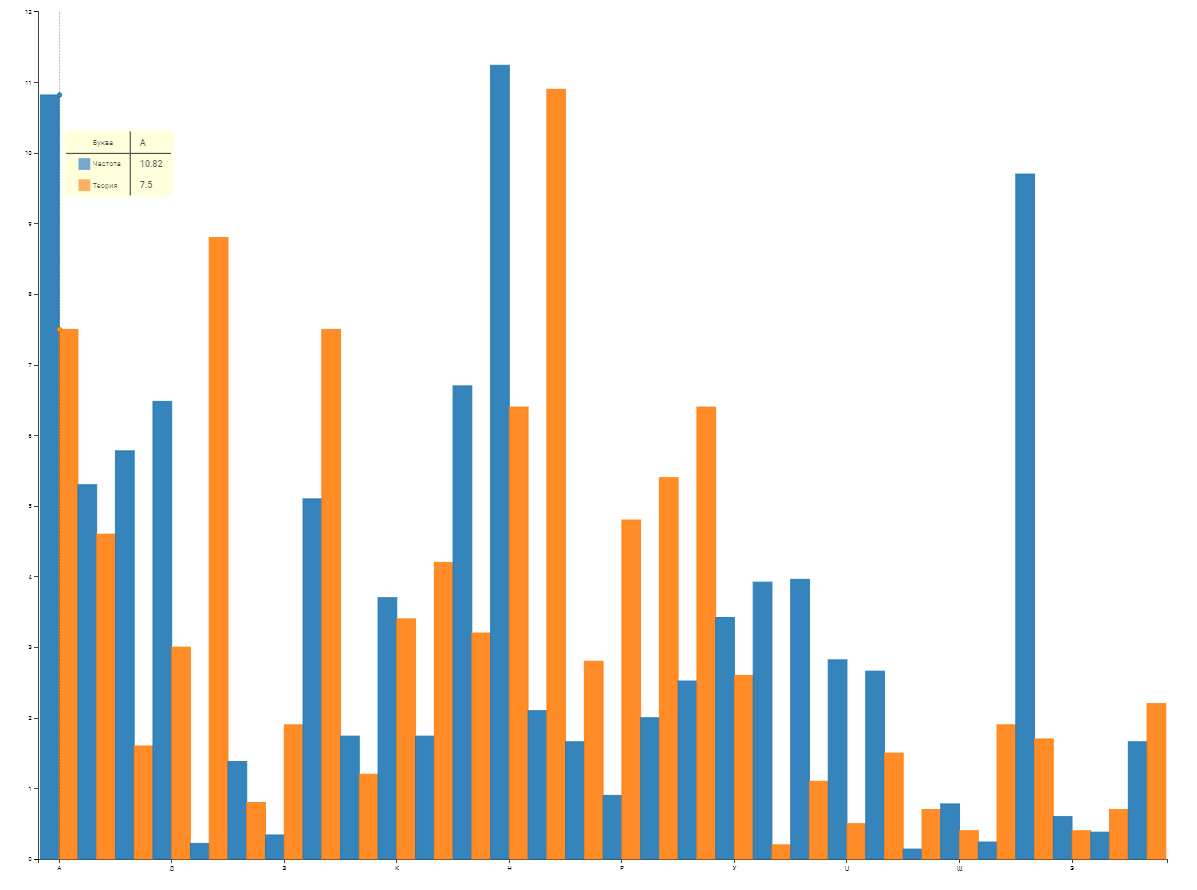


Рисунок 8 – Гистограмма для зашифрованного текста таблицей Трисемуса

Время выполнения операций зашифрования/расшифрования были замерены при помощи класса *StopWatch*:

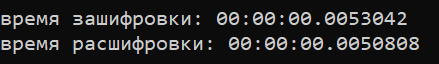


Рисунок 9 – Время для шифра Цезаря



Рисунок 10 – Время для таблицы Трисемуса

**Ответы на контрольные вопросы**

1. **В чем заключается основная идея криптографических преобразований на основе шифров замены?**

Одни символы (или группа символов) заменяются на другие по какому-либо алгоритму.

1. **Привести классификационные признаки и дать сравнительную характеристику разновидностям подстановочных шифров.**

Подстановочные шифры делятся на моноалфавитные, полиграммные, омофонические и полиалфавитные. В моноалфавитных символы сообщения заменяются на другие символы того же алфавита. Для полиграммных шифров характерна замена группы символов и на символ того же или другого алфавита. В омофонических шифрах производят замену сходных по звучанию частей сообщения. В полиалфавитных для шифрования используется 2 и более алфавита.

Наиболее сложными для реализации, но также наиболее криптостойкими, являются омофонические шифры, так как сложно создать алгоритм, определяющий созвучие слов или групп символов. Наименее криптостойкими являются моноалфавитные шифры, так как легко проанализировать частоту встречаемости символов, за счёт чего сообщение и можно расшифровать.

1. **Сколько разновидностей шифров, подобных шифру Цезаря, можно составить для алфавитов русского и белорусского языков?**

Существует 4 разновидности шифра Цезаря, но так как мы можем комбинировать несколько шифров, необходимо найти сумму сочетаний, умноженную на 2, так как 2 алфавита:

1. **Можно ли использовать в качестве ключевого в шифре Виженера слово, равное по длине открытому тексту? Обосновать ответ.**

Да.

1. **По какому признаку можно определить, что текст зашифрован шифром Плейфера?**

Так как шифр Плейфера работает с биграммами, в зашифрованном сообщении будут в среднем чаще встречаться двойные повторения символов (например, AA или VV). Частотный анализ большого шифра также покажет, что распределение вероятностей встречи символов по числам значительно отличается от алфавитного, что не позволяет установить соответствие между символом шифра и символом алфавита по вероятности его встречи в тексте.

1. **Имеются ли предпочтения в выборе размеров таблицы Трисемуса для виртуального алфавита мощностью 40: 4×10; 10×4; 5×8; 8×5; 2×20; 20×2?**

Разницы нет, но стоит учитывать, что при выборе разных размерностей таблиц получатся разные шифры.

1. **Охарактеризовать основные виды атак на шифры.**

Атака с известным шифротекстом – известен алгоритм, имеет образец шифра, неизвестен ключ.

Атака с выбором шифротекста – имеется большое количество шифротекста, из которого можно выбрать только нужную часть.

Адаптивная атака с выбором шифротекста - имеется возможность выбирать новые шифрограммы для расшифрования с учетом того, что известна некоторая информация из предыдущих сообщений.

Атака с известным открытым шифротекстом - то же, что и предыдущая, но для некоторых шифрограмм имеются соответствующие им открытые тексты.

Атака с выбором открытого текста – имеется несколько открытых текстов и соответствующих им шифротекстов.

Адаптивная атака с выбором открытого текста – имеется возможность выбирать новые шифрограммы и соответствующий им открытый текст, при этом известна некоторая информация из предыдущих сообщений.

Атака на основе связанных ключей – неизвестны ключи, но известно соотношение или связь между ними.

Атака с выбором ключа – часть ключа задаётся при взломе.

1. **Сравнить криптостойкость шифра Цезаря и шифра Виженера.**

Шифр Цезаря имеет меньшую криптостойкость, так как его легко взломать брутфорсом (прямом перебором значений) или с помощью частотного анализа, если имеется большой образец зашифрованной инфоомации.

1. **Охарактеризовать основные методы взлома подстановочных шифров.**

Брутфорс – прямой перебор всех возможных вариантов.

Частотный анализ – установка однозначного соответствия между символами шифра и алфавита на основе схожего числа вероятности встречи символа в тексте.

Анализ встречаемости триграмм – лучше всего подходит для зашифрованного английского текста. Так как триграммы THE, AND, ING и другие встречаются в английском языке достаточно часто, необходимо установить соответствие между буквами триграмм и буквами шифра (как правило на основе частотного анализа). Соответствие между оставшимися символами устанавливается на основе уже установленных.

Поиск восхождением к вершине – берётся часть шифротекста (называемая ключом), и оставшаяся часть расшифровывается с его помощью. Затем вычисляется коэффициент вероятности принадлежности расшифрованного текста к естественному языку. В ключе производятся какие-либо изменения (например, перестановка первых двух букв местами) и снова расшифровывается текст и вычисляется коэффициент. Так повторяется, пока коэффициент не перестанет изменяться.

**Вывод:** в ходе лабораторной работы было разработано авторское приложение с целью изучения и приобретения практических навыков разработки и использования приложений для реализации подстановочных шифров. Были реализованы алгоритмы шифра Цезаря и таблицы Трисемуса с ключевым словом.