ОЛИМПИАДА ШКОЛЬНИКОВ «ШАГ В БУДУЩЕЕ»

НАУЧНО-ОБРАЗОВАТЕЛЬНОЕ СОРЕВНОВАНИЕ «ШАГ В БУДУЩЕЕ, МОСКВА»

1334

регистрационный номер

Радиоэлектроника и лазерная техника
название факультета

ЛАЗЕРНЫЕ И ОПТИКО-ЭЛЕКТРОННЫЕ СИСТЕМЫ
название кафедры

Встраивание скрытой информации в цифровые изображения
название работы

Автор: ***Гаврилова Дарья Вадимовна***

фамилия, имя, отчество

Лицей №1581, 11 класс
наименование учебного заведения, класс

Научный руководитель: ***Трофимов Николай Евгеньевич***
фамилия, имя, отчество
МГТУ им. Н.Э. Баумана, каф. РЛ-2
место работы
Старший преподаватель
звание, должность

_____
подпись научного руководителя

Москва – 2019

Аннотация

Работа посвящена сокрытию данных в цифровых изображениях. В настоящее время проблема обеспечения конфиденциальности хранимых и пересылаемых данных стала чрезвычайно важна. В работе предложен метод использующий ряды Фурье (дискретное преобразование Фурье). Был проведен анализ качества предложенного метода сокрытия данных. Для демонстрации метода была разработана программа на языке программирования Python.

Оглавление

Введение

Информация является одним из важнейших предметов современной жизни. В настоящее время происходит активное развитие сетевых технологий. Как следствие этого большое количество информации передается по сетям. При этом возрастает процент атак злоумышленников и попыток несанкционированного доступа к передаваемой информации. Поэтому появилась необходимость защиты информации от несанкционированного доступа. Сокрытие данных очень важно, оно гарантирует защиту вашей информации.

Цели и Задачи:

Цель:

Провести анализ методов сокрытия информации в цифровых изображениях.

Для достижения поставленной цели были решены следующие задачи:

- Изучить историю методов сокрытия информации
- Реализовать метод, использующий дискретное преобразование Фурье на Python
- Провести оценку метода, использующего дискретное преобразование Фурье

1. Сокрытие информации

1.1. Стеганография и Криптография

Стенографировать можно как вручную, так и при помощи специальных инструментов. Информация обычно прячется в графических и видео файлах. Там, где можно подрезать цветовую гамму и на глаз разницы не увидишь [3].

Стеганография является наукой, обеспечивающей обмен информацией таким образом, что скрывается сам факт существования секретной связи. Слово "стеганография" в переводе с греческого буквально означает "тайнопись" (steganos – секрет, тайна; graphy – запись). Стеганография не заменяет криптографию (шифрование данных), а дополняет ее еще одним уровнем безопасности. При обработке данных стеганографическими методами происходит скрытие передаваемой информации в других объектах (файлах, дисках и т. п.) таким образом, чтобы постороннее лицо не догадывалось о существовании скрытого секретного сообщения.

Внесенные искажения должны быть ниже уровня чувствительности средств распознавания. Кроме скрытой передачи сообщений, стеганография применяется для аутентификации и маркировки авторской продукции. При этом, часто в качестве внедряемой информации используются дата и место создания продукта, данные об авторе, номер лицензии, серийный номер и др. Эта информация обычно внедряется как в графические аудио произведения, так и в защищаемые программные продукты. Стеганографическая система или стегосистема – совокупность средств и методов, которые используются для формирования скрытого канала передачи информации [9].

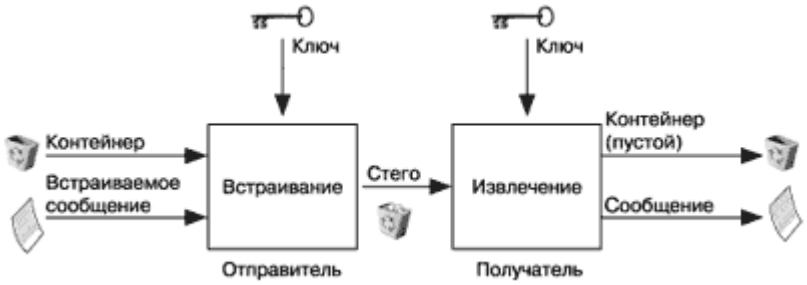


Рис.1 Обобщённая модель стеганосистемы

Криптография — наука о методах обеспечения конфиденциальности, целостности данных, аутентификации, а также невозможности отказа от авторства [2].

Компьютерная стеганография — направление классической стеганографии, основанное на особенностях компьютерной платформы. Примеры — стеганографическая файловая система StegFS для Linux, скрытие данных в неиспользуемых областях форматов файлов, подмена символов в названиях файлов, текстовая стеганография и т. д.

В некоторых задачах без Стенографии и Криптографии не обойтись.

1.2. История возникновения

О важности сохранения информации в тайне знали уже в древние времена, когда с появлением письменности появилась и опасность прочтения ее нежелательными лицами. Более того, первоначально письменность сама по себе была криптографической системой, так как в древних обществах ею владели только избранные. С широким распространением письменности криптография стала формироваться как самостоятельная наука. В документах древних цивилизаций - Индии, Египта, Месопотамии - есть сведения о системах и способах составления шифрованных писем.

Точное время возникновения этих способов обмена тайной информацией теряется в глубине веков, и установить его невозможно. Историки полагают, что первые протокриптографические приемы появились в Древнем Египте около 4 тыс. лет назад. Писцы, составлявшие жизнеописания правителей, стремились придать стандартным иероглифам необычный вид на монументах и гробницах, чтобы сообщить надписям менее обыденный и более почтительный стиль. Жрецы пользовались этим же приемом при переписывании религиозных текстов, чтобы те выглядели для мирян загадочнее и внушительнее. Такие «переводы» становились все менее понятными простому люду, который в результате оказывался во все большей зависимости от жрецов.

По мере развития египетской цивилизации ширилось использование иероглифов. С увеличением количества надписей, высекавшихся на стенах храмов, люди теряли к ним интерес. Египтологи считают, что писцы тогда стали еще больше видоизменять некоторые знаки в стремлении пробудить любопытство и привлечь внимание населения. Эти модификации никоим образом не были кодами или шифрами, но они заключали в себе два основных принципа криптологии, а именно: изменение письма и сокрытие его смысла. Бесспорных доказательств, указывающих на широкое использование модификаций иероглифов для сокрытия дипломатических, торговых или военных планов в Древнем Египте, нет.

Более явные криптологические примеры дошли до нас от цивилизаций Месопотамии — от вавилонян, ассирийцев, халдеев, использовавших особую систему письма — клинопись. В 1500 г. до н. э. на глиняной табличке был записан тщательно охраняемый рецепт глазури для гончарных изделий. Знаки, определяющие необходимые ингредиенты, были намеренно перемешаны. Таким образом, мы имеем право утверждать, что эта табличка является самой ранней известной секретной записью.

Примерно с 500 г. до н. э. в Индии также широко применялись секретные записи, в частности в донесениях шпионов и текстах, предположительно использовавшихся Буддой. Методы засекречивания включали в себя фонетическую замену, когда согласные и гласные менялись местами, использование перевернутых букв и запись текста под случайными углами. Различные индийские трактаты, ярким примером которых является «Артхашастра» (около 321—300 гг. до н. э.), показывают, что индийцы были хорошо знакомы со способами сокрытия информации. Согласно «Камасутре», классическому произведению об эротике и других видах наслаждений, написанному Ватсьяяной около IV в. до н.э., владение приемами составления секретного письма (наряду с музыкой, кулинарией и шахматами) является одним из 64 искусств, которыми должны владеть женщины [5].

Симпати́ческие (невидимые) черни́ла — чернила, записи которыми являются изначально невидимыми и становятся видимыми только при определённых условиях (нагрев, освещение, химический проявитель и т. д.) .

Одним из наиболее распространённых методов классической стеганографии является использование симпатических чернил. Обычно процесс записи осуществляется следующим образом: первый слой — наносится важная запись невидимыми чернилами, второй слой — ничего не значащая запись видимыми чернилами [1].

Микрото́чка — изображение, уменьшенное до такой степени, что неосведомлённый наблюдатель не сможет его ни прочесть, ни даже обнаружить.

Обычно «микроточки» имеют не более миллиметра в диаметре. Своё название получили от сходства с типографской точкой. Технология изготовления микроточек является одной из разновидностей стеганографии [6].

2. Показатели качества систем сокрытия данных

2.1. Скрытность

Скрытность, то есть незаметность внесенной скрытой информации.

Незаметность наличия скрытого сообщения определяется свойствами зрительной системы человека (ЗСЧ). Информация должна быть встроена таким образом, чтобы среднестатистический человек был не в состоянии отличить изображение с встроенной информацией от исходного изображения-контейнера [12].

2.2. Объем скрываемых данных.

Основной информационной характеристикой системы сокрытия данных является третий показатель. Чтобы оценивать количество полезной передаваемой информации используют BER.

Bit error rate (BER) - частота появления ошибочных битов. Отношение числа неверно принятых битов (0 вместо 1 и наоборот) к полному числу переданных битов при передаче по каналу связи [8].

BER =E/N,

где E – количество ошибочных битов,

N – общее число переданных битов.

BER в пределе эквивалентно понятию вероятности ошибки.

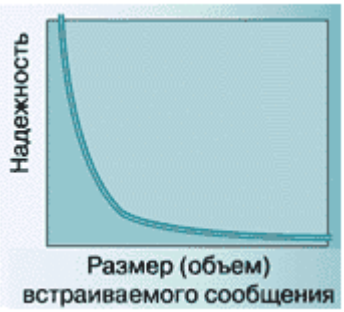


Рис.2

На Рис.2 зависимость показывает, что при увеличении объема встраиваемых данных снижается надежность системы. Таким образом, используемый в стегосистеме контейнер накладывает ограничения на размер встраиваемых данных [10].

2.3. Устойчивость

Устойчивость (робастность), то есть устойчивость к искажениям.

Стойкость системы сокрытия информации определяется с одной стороны способностью противостоять стеганоанализу (undetectability), и с другой стороны устойчивостью к искажениям стегано-объекта, направленных на уничтожение, модификацию или подмену сообщения. При анализе стойкости системы сокрытия данных к злонамеренным действиям пользуются принципом, аналогичным принципу Керкгоффса, используемому в криптографии. Он состоит в предположении, что в засекреченном виде держится только определённый набор параметров алгоритмов, называемый ключом, а сами алгоритмы встраивания и извлечения сообщения должны быть открытыми. Другими словами, при оценке стойкости системы сокрытия

данных необходимо предполагать, что противник знает об используемой системе всё, кроме применяемых ключей [12].

2.4. Формат JPEG

Процесс сжатия по схеме JPEG состоит из нескольких шагов. На первом шаге производится преобразование изображения из цветового пространства RGB в пространство YUV, основанное на характеристиках яркости и цветности. Вся дальнейшая работа производится именно с этим цветовым пространством, которое благодаря некоторым своим характеристикам позволяет получать нам столь большие степени сжатия.

Что же такого необычного в YUV представлении цвета по сравнению с RGB? А то, что оно наиболее близко к "естественному", тому, которое неосознанно выполняет человек. Y-компонента, или яркость, тесно связана с качеством картинки. Точнее сказать Y - это и есть картинка, только черно-белая. Компоненты U и V содержат информацию о цвете и позволяют нам раскрашивать Y-картинку.

На следующем после преобразования шаге изображение разделяется на квадратные участки размером 8x8 пикселей. После этого над каждым участком производится т.н. дискретное косинус-преобразование (ДКП). При этом выполняется анализ каждого блока, разложение его на составляющие цвета и подсчет частоты появления каждого цвета.

Человеческий глаз устроен таким образом, что наиболее чувствителен именно к яркостной составляющей изображения (Y-компонента) и наименее к цветовым.

Анализируя частотную информацию о появлении цветов, удается избавиться от части информации уже в процессе квантования. При этом цвета в верхней части спектра исключаются, что практически не сказывается на зрительном восприятии образа. Также исключается часть яркостной информации. Грубо говоря, JPG просто отбрасывает от яркостной составляющей половину полезного сигнала, а от цветовой 3/4 [11].

Предназначенные для использования в электронных средствах массовой информации не анимированные полутоновые и полноцветные фотографии обычно сохраняют в цифровых форматах PNG и JPEG. К достоинствам этих форматов можно отнести кроссплатформенность, возможность обработки практически во всех графических редакторах, хорошие показатели качества изображений. Изображения в формате JPEG за счёт возможности их сжатия с потерями имеют меньший размер по сравнению с аналогичными, сохранёнными в формате PNG. Следовательно, использование формата JPEG для сохранения и передачи изображений является более предпочтительным [3].

3. Методы сокрытия информации в изображениях

Эхо-методы

Эхо-методы применяются в цифровой аудиостеганографии и используют неравномерные промежутки между эхо-сигналами для кодирования последовательности значений. При наложении ряда ограничений соблюдается условие незаметности для человеческого восприятия. Эхо характеризуется тремя параметрами: начальной амплитудой, степенью затухания, задержкой. При достижении некоего порога между сигналом и эхом они смешиваются. В этой точке человеческое ухо не может уже отличить эти два сигнала. Наличие этой точки сложно определить, и она зависит от качества исходной записи, слушателя. Чаще всего используется задержка около 1/1000, что вполне приемлемо для большинства записей и слушателей. Для обозначения логического нуля и единицы используется две различных задержки. Они обе должны быть меньше, чем порог чувствительности уха слушателя к получаемому эху.

Эхо-методы устойчивы к амплитудным и частотным атакам, но неустойчивы к атакам по времени [4].

Фазовое кодирование

Фазовое кодирование (phase coding, фазовое кодирование) — так же применяется в цифровой аудиостеганографии. Происходит замена исходного

звукового элемента на относительную фазу, которая и является секретным сообщением. Фаза подряд идущих элементов должна быть добавлена таким образом, чтобы сохранить относительную фазу между исходными элементами. Фазовое кодирование является одним из самых эффективных методов скрытия информации [4].

Метод расширенного спектра

Метод встраивания сообщения заключается в том, что специальная случайная последовательность встраивается в контейнер, затем, с использованием согласованного фильтра, данная последовательность детектируется. Данный метод позволяет встраивать большое количество сообщений в контейнер, и они не будут создавать помехи друг другу при условии ортогональности применяемых последовательностей. Преимуществом данного метода является противодействие геометрическим преобразованиям, удалению части файла и тд. Метод заимствован из широкополосной связи [4].

Широкополосные методы

Широкополосные методы используются при передаче данных, обеспечивая высокую помехоустойчивость и препятствуя процессам их перехвата. Отличительной особенностью широкополосного метода от других является расширение диапазона частот сигнала за счет кода, на который не влияют передаваемые данные. Необходимая информация рассредоточена по всей полосе частот и, в случае потери сигнала, данные могут быть восстановлены из других полос частот. Подобный подход к сокрытию сигналов значительно усложняет процесс выявления зашифрованной информации, а также ее удаления. Поэтому широкополосные методы устойчивы к любым атакам [4].

Существует два основных метода расширения спектра:

Метод псевдослучайной последовательности. В данном методе используется секретный сигнал, модулируемый псевдослучайным сигналом.

Метод прыгающих частот. В данном методе частота несущего сигнала должна меняться по определенному псевдослучайному закону [4].

Статические методы

Статический метод — метод сокрытия данных, при котором изменяются определенные статистические характеристики изображения, при этом получатель способен распознать видоизмененное изображение от исходного [4].

Методы искажения

Методы искажения — методы сокрытия данных, при которых, в зависимости от секретного сообщения, выполняются последовательные преобразования контейнера. В данном методе важно знать первоначальный вид контейнера. Зная различия между первоначальным контейнером и стеганограммой, можно восстановить исходную последовательность преобразований и извлечь скрытые данные. Следует отметить, что при применении данного метода важно соблюдать правило: распространение набора первоначальных контейнеров осуществляется только через секретные каналы доставки. В случае несоблюдения данного правила, противник тоже сможет завладеть набором первоначальных контейнеров, что приведет к вскрытию тайной переписки [4].

Структурный метод

Структурный метод — метод сокрытия данных, при котором формируется скрываемый текст, посредством осуществления последовательных модификаций частей изображения. Данный метод позволяет не только модифицировать изображение, в котором будет скрыто послание, но и создавать изображение по секретному сообщению. Структурный метод весьма устойчив против атак [4].

Метод замены наименее значащего бита или LSB

Суть метода замена наименее значащего бита (Least Significant Bits - LSB) заключается в сокрытии информации путем изменения последних битов изображения, кодирующих цвет на биты скрываемого сообщения. Разница

между пустым и заполненным контейнерами должна быть не ощутима для органов восприятия человека [7].

4. Метод, использующий дискретное преобразование Фурье

Суть метода, использующий ряды Фурье, заключается в сокрытии информации путем разложения скрывающей информации на ряды Фурье и наложения ее на изображение. Разница между пустым и заполненным контейнерами должна быть не ощутима для органов восприятия человека.

Ряд Фурье — в математике — способ представления произвольной сложной функции суммой более простых. В общем случае количество таких функций может быть бесконечным, при этом чем больше таких функций учитывается при расчете, тем выше оказывается конечная точность представления исходной функции. В большинстве случаев в качестве простейших используются тригонометрические функции синуса и косинуса, в этом случае ряд Фурье называется тригонометрическим, а вычисление такого ряда часто называют разложением на гармоники [7].

Принцип сокрытия информации

Принцип сокрытия данных в изображении и декодирования представлен на рис.3.

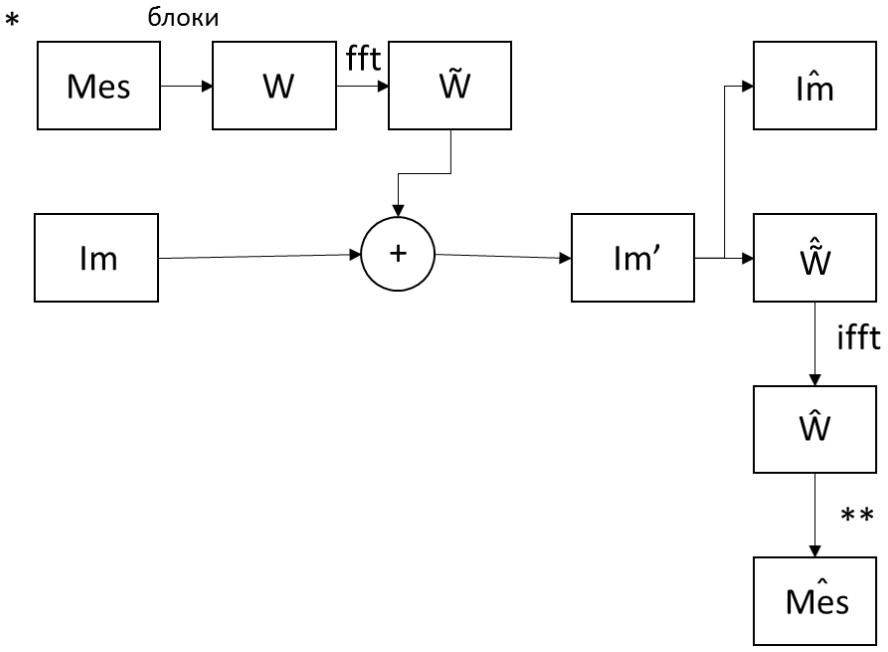
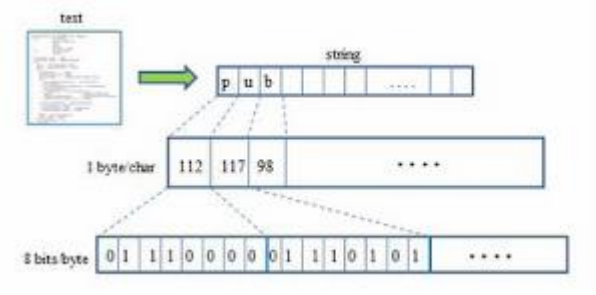


Рис.3 Принцип сокрытия данных и декодирования



*Вводим текст с клавиатуры. Преобразуем введенный текст в числа (с помощью кодировочной таблицы) и переводим в бинарный код (Преобразование текста в байтовую последовательность). Например, бинарный код *abc* выглядит следующим образом: 011000010110001001100011

Рис.4 Преобразование текста в байтовую последовательность

Затем создаем матрицу из нулей размерами с половину исходной картинки, в которую блоками 5×5 мы будем записывать бинарный код.

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Рис.5 Сокрытие информации в изображении

011000010110001001100011

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Рис.6 Сокрытие информации в изображении

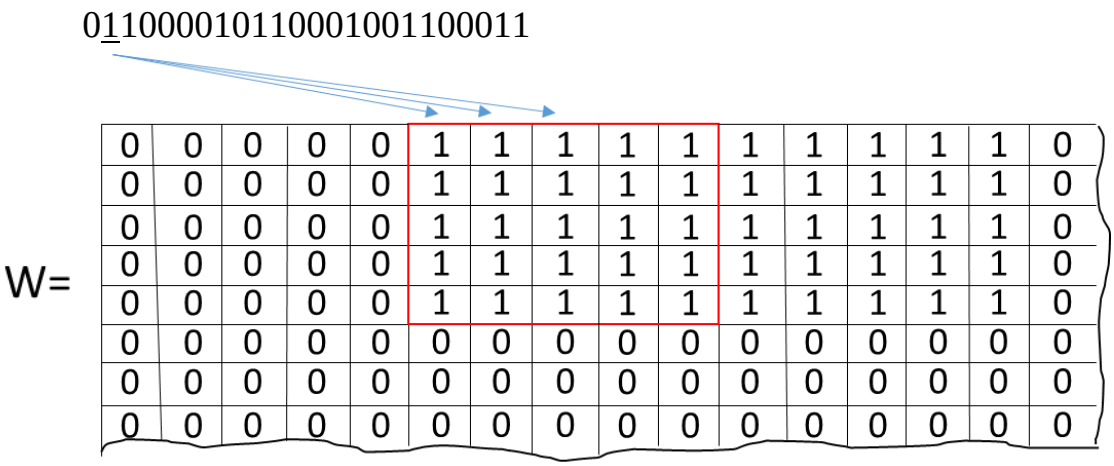


Рис.7 Сокрытие информации в изображении

Затем мы применяем преобразование Фурье к нашей матрице. И складываем данное изображение с матрицей как в формуле (1):

$$Im' = Im + \tilde{W} \qquad (1)$$

**По блокам вытаскиваем значения пикселей. Получаем длинную строку из нулей и единиц. Берем первые 25 значений, ищем их среднее арифметическое и округляем. Делаем так совсем список и в итоге получаем бинарный код сообщения. Остаётся только преобразовать его в текст.

Достоинства и недостатки метода

Достоинства метода это конечно же устойчив к большинству видов атак. Но также имеются и недостатки и самый главный из них это маленький объем скрываемых данных.

5. Реализация метода, использующего дискретное преобразование Фурье на Python

5.1. Сокрытие

При реализации данной задачи на Python, была использована библиотека PIL и numpy:

```
from PIL import Image

import numpy as np
```

Также выводилась формула расчета ширины и длины зоны информации и номера блоков:

k - длинна битовой цепочки

p*q - размер блока

width – длина картинки

height – ширина картинки

```
n1=ceil((k*p*q/a)**0.5 / p) * p  #ширина зоны с информацией

m1=ceil((k*p*q)/n1/q)*q  #длина зоны с информацией

    for i in range(height//2):

      for j in range(width):

          R=(i+1)//p # номер блока по вертикали

          C=(j+1)//q # номер блока по горизонтали

            u = int(C + R * m1/p) # номер бита в битовой цепочке
```

Читаем изображение и текст:

```
im1 = Image.open("тигр.jpg")

text=str(input("Введите текст: "))
```

Преобразуем текст в числа (с помощью кодировочной таблицы) и переводим в бинарный код (Преобразование текста в байтовую последовательность):

```
for nchar in range (len(text)):

   g = ord(text[nchar])

   b=[0]*8

   n=len(b)

   for i in range(n):
```

```
        y = int(g % 2)
        b[n-i-1]=y
        g = int(g / 2)
    d+=b
```

В значения каждого из каналов по блокам встраиваем, переведённый в бинарный код, текст:

```
if k<=(m1*n1):
    for i in range(height//2):
        for j in range(width):
            R=(i+1)//p
            C=(j+1)//q
            u = int(C + R * m1/p)
            x=j
            y=i
            newR = d[u%k]
            newG = d[u%k]
            newB = d[u%k]
            im.putpixel((x,y),(newR, newG, newB))
```

Делаем преобразование Фурье и сохраняем изменения:

```
im3 = Image.fromarray(im3, 'RGB')
im3. save ("out1.png")
```

5.2. Обратная программа (Декадирование)

При реализации данной задачи на Python, была использована библиотека PIL и numpy:

```
from PIL import Image

import numpy as np
```

Читаем изображение:

```
im1 = Image.open("out1.png")
```

Узнаем размеры изображения:

```
(width, height) = im.size
```

Делаем обратное преобразование Фурье:

```
im1 = np.fft.ifft(im1)
```

Берем значения пикселей, вытаскиваем их по блокам, формируем из этих значений массив и разбиваем по 25, по каждым 25 значениям ищем средне арифметическое и формируем из них список, затем разбиваем уже этот список по 8:

Переводим из двоичной системы в десятичную:

```
m=2**7*b[0]+2**6*b[1]+2**5*b[2]+2**4*b[3]+2**3*b[4]+2**2*b[5]+2**1*b[6]+2**0*b[7]
```

Преобразуем числа в буквы и запоминаем:

```
l.append(chr(m))
```

Выводим расшифрованный текст:

```
print(*l)
```

Заключение

В результате проведенного исследования была изучена история методов сокрытия информации, реализован метод использующий ряды Фурье (дискретное преобразование Фурье) на Python, проверен этот метод на устойчивость и дана оценка данному методу.

Список литературы

[1]. Википедия свободная энциклопедия [Электронный ресурс] – Режим доступа: https://ru.wikipedia.org/wiki/Симпатические_чернила, свободный.

[2]. Википедия свободная энциклопедия [Электронный ресурс] – Режим доступа: https://ru.wikipedia.org/wiki/Криптография, свободный.

[3]. CyberSafeRus [Электронный ресурс] – Режим доступа: https://habrahabr.ru/company/cybersafe/blog/213543/, свободный.

[4]. Википедия свободная энциклопедия [Электронный ресурс] – Режим доступа: https://ru.wikipedia.org/wiki/Стеганография, свободный.

[5]. Google Сайты [Электронный ресурс] – Режим доступа: https://sites.google.com/site/anisimovkhv/learning/kripto/lecture/tema2, свободный.

[6]. Википедия свободная энциклопедия [Электронный ресурс] – Режим доступа: https://ru.wikipedia.org/wiki/Микроточка, свободный.

[7]. Википедия свободная энциклопедия [Электронный ресурс] – Режим доступа: http://ru.math.wikia.com/wiki/Ряд_Фурье, свободный.

[8]. Молодежный научно-технический вестник, автор: Кобецкая Е. А. [Электронный ресурс] – Режим доступа: http://sntbul.bmstu.ru/doc/849819.html, свободный.

[9]. Файловый архив студентов, 1042 вуза [Электронный ресурс] – Режим доступа: https://studfiles.net/preview/3676281/page:7/, свободный.

[10]. О. В. Генне, ООО "Конфидент" [Электронный ресурс] – Режим доступа: file:///G:/ОСНОВНЫЕ%20ПОЛОЖЕНИЯ%20СТЕГАНОГРАФИИ.html, свободный.

[11]. Mike Melnikov [Электронный ресурс] – Режим доступа: http://citforum.ru/internet/webd/article_21.shtml, свободный.

[12]. Автор: Трофимов Николай Евгеньевич [Электронный ресурс] – Режим доступа: thesis.pdf, CD-ROM.