УДК 003.26

СТЕГАНОГРАФИЧЕСКОЕ ПРОГРАММНОЕ СРЕДСТВО

Чкоидзе О. Д., студент группы 861401

Белорусский государственный университет информатики и радиоэлектроники

г. Минск, Республика Беларусь

Зельманский О. Б. – канд. техн. наук

Аннотация. Проанализированы методы сокрытия данных в цифровых изображениях, а именно метод наименьшего значащего бита и метода двумерного косинусного преобразования, разработано специализированное программное средство, реализующее данные методы.

Ключевые слова. Стеганография, дискретное косинусное преобразование, наименее значимый бит, гаммирование.

На сегодняшний день существует множество способов сокрытия информации в цифровых изображениях, однако многие из них не являются эффективными в части сокрытия самого факта передачи информации, так, например, склеивание цифрового изображения с архивом или запись важной информации в метаданные являются легко обнаружимыми способами передачи информации посредством цифровых изображений, поскольку анализ структуры файла в шестнадцатеричном редакторе позволяет обнаружить скрываемую информацию.

Эффективными методами сокрытия информации в цифровых изображениях являются те методы, которые работают непосредственно с цветовыми информационными каналами изображения. Каналы являются изображениями в градациях серого, которые содержат информацию различного типа. Цветовые информационные каналы создаются автоматически при открытии нового изображения. Например, в изображении RGB есть канал для каждого цвета (красный, зелёный и синий). Также существует альфа-канал, который несёт в себе информацию о прозрачности пикселя. Стоит отметить, что в зависимости от формата изображения, количество каналов может быть разным, так в изображениях формата “Bitmap” существует только 3 цветовых канала, а в изображениях формата “Portable Network Graphics” существует 4 цветовых канала.

К эффективным методам сокрытия информации можно причислить такие методы, как метод наименьшего значащего бита (НЗБ) и множество вариаций метода двумерного косинусного преобразования (ДКП).

В рамках настоящего проекта, был разработан программный модуль, который позволяет работать с различными стеганографическими методами и анализировать их влияние на исходное изображение. Во время разработки была использована библиотека, позволяющая получить из картинки любого формата информацию о цветовых каналах. Сам модуль написан на C++ с помощью Visual Studio 2022.

Метод, который позволяет скрыть наиболее большой объём двоичных данных в пределах одной картинки, является НЗБ (LSB – Least Significant Bit). Суть данного метода заключается в том, что произвольно выбирается один или несколько цветных каналов в которых будет производится скрытие информации, далее определяется первое значение цветового сигнала, одного из выбранных каналов, в формате целого неподписанного числа, то есть числа имеющего значение от 0 до 255, в некоторых случаях это может быть число с плавающей запятой, которое будет варьироваться от 0.0 до 1.0, после чего это число конвертируется в двоичный формат и на него накладывается маска с помощью побитового логического оператора И (AND), которая обнуляет последние два бита данного числа, далее, посредством логического оператора ИЛИ (OR), в освобождённое место, записывается два бита двоичных данных, которые необходимо скрыть в исходном изображении. Этот процесс повторяется до тех пор, пока не будут записаны все входные данные. Стоит учитывать, что, если в одном канале не хватает места для записи двоичных данных в полном объёме, то они записываются в следующий канал, однако в разработанном программном модуле изначально производится фрагментация входных данных с учётом количества выбранных цветовых каналов для записи, это позволило значительно уменьшить последствия изменения цветовых сигналов в одном цветовом канале. Таким образом, исходные данные равномерно распределяются по всем выбранным каналам, сохраняя исходное состояние некоторых пикселей, тем самым уменьшая шанс возможных визуальных дефектов.

Для наглядности рассмотрим изображение, преобразованное данным методом (рис.1, 2). Входными данными в этом случае выступала строчка с цифрами от нуля до тридцати (81 байт двоичной информации), а запись производилась равномерно по всем цветовым каналам. Исходное изображение имело формат PNG (Portable Network Graphics) и размеры 512 на 512 пикселей.

Рисунок 1 – Исходное изображение

Рисунок 2 – Преобразованное изображение

Как можно видеть, визуальных дефектов не наблюдается, а следовательно факт передачи информации остался в тайне.

Размер двоичных данных (I), которые потенциально можно скрыть в данном изображении рассчитывается по следующей формуле:

$$I = \frac{((W \times H) \times 8) \times N}{2},$$

где W – ширина исходного изображения;

H – высота исходного изображения;

N – количество цветовых каналов.

После произведения расчётов, получаем размер в 3145728 бит или 384 килобайт, не учитывая канал прозрачности. Необходимо понимать, что цена за большой объём двоичной информации, которую позволяет скрыть этот метод, является его неустойчивость к повреждениям или изменениям преобразованного изображения.

Метод ДКП (DCT – Discrete Cosine Transform), вариация Коха-Жао. Преимущество данного метода заключается в его стойкости к повреждениям или изменениям преобразованного изображения, однако, как и с предыдущим методом у этого преимущества есть своя цена, заключается она в малом потенциальном объёме данных, которые может быть скрыто в изображении.

Суть метода заключается в том, чтобы фрагментировать выбранные цветовые каналы на блоки 8 на 8 сигнальных коэффициентов каждый и применить к ним формулу ДКП. После этого получаются матрицы размером 8 на 8, в которых будут хранится коэффициенты ДКП. Далее пользователь должен произвольно выбрать случайное число и два индекса среднечастотных коэффициентов ДКП с помощью которых будет выполнятся интеграция скрываемых данных в исходное изображение. Следующим шагом будет циклический проход по каждой из матриц ДКП, учитывая размер входных данных, во время которого будут произведены изменения над выбранными пользователем двумя коэффициентами ДКП по следующим правилам:

1) Находится модуль каждого из двух коэффициентов, и вычисляется их разность.
2) При записи бита равного единице эта разность должна быть больше заданного пользователем случайного числа. В противном случае эти коэффициенты изменяются так, чтобы это правило выполнялось. В разработанном программном модуле эти коэффициенты обмениваются значениями.
3) При записи бита равного нулю эта разность должна быть меньше, заданного пользователем, случайного числа. В противном случае выполняется то, что было оговорено ранее.

После завершения цикла к каждой из изменённых матриц применяется формула обратного двумерного косинусного преобразования (ОДКП). Значения коэффициентов результирующей матрицы ОДКП будут представлять новые значения цветовых сигналов, которые в дальнейшем будут записаны в соответствующий цветовой канал (рис. 3). Необходимо учитывать, что выходные коэффициенты ОДКП могут выходить за пределы значений цветового сигнала, в связи с этим, каждый коэффициент должен пройти через функцию нормирования, при этом может произойти потеря или искажение входных данных, исходя из этого, можно сделать вывод, что, чем больше случайное значение, выбранное пользователем, тем вероятнее то, что после прохождения через функцию нормирования данные останутся без искажений.

Для чтения данных цветовые каналы фрагментируются на блоки 8 на 8 сигнальных коэффициентов, на их основе вычисляются матрицы с коэффициентами ДКП, после чего сравниваются значения коэффициентов с учётом выбранных ранее индексов в матрице. Если первый коэффициент больше второго, то в буфер данных записывается бит со значением 1, в противном случае записывается бит со значением 0.

Рисунок 3 – Преобразованное изображение

Для расчёта ДКП используется следующая формула [1]:

$$\mathrm{DCT}(u, v) = \frac{1}{\sqrt{2N}} C(u) C(v) \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \mathrm{signal}[i][j] \cos\frac{(2i+1)u\pi}{2N} \cos\frac{(2j+1)v\pi}{2N}$$

$$C(x) = \frac{1}{\sqrt{2}} \text{ if } x \text{ is } 0, \text{else } 1 \text{ if } x > 0,$$

где u – индекс смещения по ширине в матрице ДКП;

v – индекс смещения по высоте в матрице ДКП;

N – размерность матрицы ДКП;

signal – матрица цветовых сигналов;

i – индекс смещения по ширине в матрице цветовых сигналов;

j – индекс смещения по высоте в матрице цветовых сигналов.

Для расчёта ОДКП используется следующая формула [3]:

$$\mathrm{IDCT(i,j)} = \frac{1}{\sqrt{2\mathrm{N}}}\sum_{\mathrm{u=0}}^{\mathrm{N-1}}\sum_{\mathrm{v=0}}^{\mathrm{N-1}} \mathrm{C(u)C(v)DCT[u][v]}\cos\frac{(2\mathrm{i}+1)\mathrm{u}\pi}{2\mathrm{N}}\cos\frac{(2\mathrm{j}+1)\mathrm{v}\pi}{2\mathrm{N}}$$

$$C(\mathrm{x}) = \frac{1}{\sqrt{2}} \text{ if x is } 0, \text{else } 1 \text{ if x} > 0,$$

Для расчёта размера двоичных данных (I), которые потенциально возможно записать в коэффициенты ДКП используется следующая формула:

$$\mathrm{I} = ((\frac{\mathrm{W}}{\mathrm{L}} \times \frac{\mathrm{H}}{\mathrm{L}})) \times \mathrm{N},$$

где L – размерность матрицы.

После произведения расчёта размера получается число в 12288 бит или 1,536 килобайт (на 3 канала). Таким образом, при устойчивости данного алгоритма к повреждениям преобразованного изображения, он имеет малый объём информации на один цветной канал.

Предварительное использование алгоритма шифрования на блоке входных данных является важной составляющей сокрытия факта передачи информации по стегоканалу, так как представленные стеганографические алгоритмы являются публичными и довольно известными. Однако, возникает очевидная проблема, связанная с возможным повреждением или изменением преобразованного изображения при передаче через физический канал или сжатию, поэтому ассиметричные шифры могут привести к полной потере скрытой информации. Для решения данной проблемы применяется симметричный шифр, а именно – гаммирование, смысл которого заключается в наложении последовательности случайных чисел на буфер с входными данными. Данная последовательность называется гамма-последовательностью и используется для шифрования и дешифрования данных. Клод Шеннон доказал, что при определённых свойствах гаммы этот метод шифрования является абсолютно стойким. Следующие требования к гамма-последовательности обязательно должны быть соблюдены [2]:

1) Для шифрования каждого нового сообщение нужно использовать новую гамму.
2) Для формирования гаммы нужно использовать аппаратные генераторы случайных чисел, основанные на физических процессах.
3) Длина гаммы должна быть не меньше длины защищаемого сообщения.

Рассмотрим результаты дешифрования данных с наложением искусственных повреждений при использовании ранее рассмотренного стеганографического метода ДКП в вариации Коха-Жао (рис. 4).

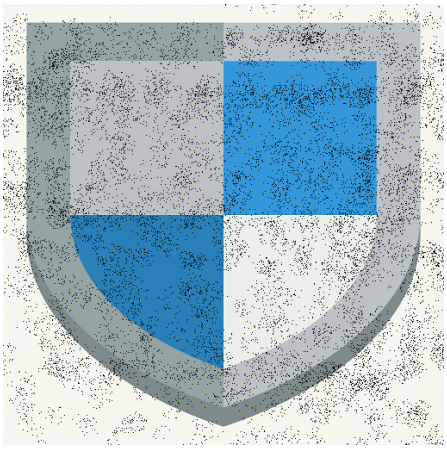

Рисунок 4 – Преобразованное изображение с наложением искусственного шума

Исходные данные: “1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30”

Результат дешифрования данных: “1 2 3 4 5 6 7 8 9 10 11 12 11 14 15 16 17 18 19 20 21 22 23 24дг5 26 27 28 29 3р”.

Исходя из результатов, можно сделать вывод, что при повреждении преобразованного изображения, из стегоканала возможно частичное извлечение данных, хоть и с определённым количеством ошибок.

В рамках данной работы было разработано программное средство, позволяющее анализировать и работать со стеганографическими методами. На сегодняшний день передача данных с помощью стегоканалов остаётся довольно актуальной и может быть использована как для доставки вредоносного кода, так и для передачи засекреченных сообщений. Предложенное программное средство может быть усовершенствовано путем добавления дополнительных стеганографических методов с целью увеличения объема передаваемой информации и стойкости шифрования, а также с целью изучения и сравнительного анализа различных методов сокрытия и шифрования информации, так как базовый код полностью основан на структурах и классах, предназначенных для работы с пикселями изображений и с информацией, разбивая её на битовые буферы.

Список использованных источников:

1. Lossy Data Compression: JPEG / Standford University [Электронный ресурс]. – Режим доступа: https://cs.stanford.edu/people/eroberts/courses/soco/projects/data-compression/lossy/jpeg/dct.htm - Дата доступа: 04.04.2022
2. Гаммирование / Wikipedia [Электронный ресурс] – Режим доступа: https://ru.wikipedia.org/wiki/Гаммирование - Дата доступа: 04.04.2022
3. Computer graphics / Stanford University [Электронный ресурс] – Режим доступа: http://nifty.stanford.edu/2003/pests/2002/lectures/08.2_graphics/TwoDimensions.html?CurrentSlide=13 Дата доступа: 05.04.2022

UDC 003.26

STEGANOGRAPHY SOFTWARE TOOL

Chkoidze O. D., student of the group 861401

Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

Zelmansky O.B. - PhD

Annotation. The methods of hiding data in digital images, namely the method of the least significant bit and the method of two-dimensional cosine transform, have been analyzed, a specialized software tool has been developed that implements these methods.

Keywords. Steganography, discrete cosine transform, least significant bit, additive cipher.