ИНФОРМАТИКА И УПРАВЛЕНИЕ В ТЕХНИЧЕСКИХ И СОЦИАЛЬНЫХ СИСТЕМАХ

УДК 004.056.5 ***DOI: 10.46960/1816-210X_2023_2_7***

МОДИФИКАЦИЯ СТЕГАНОГРАФИЧЕСКОГО МЕТОДА КВАНТОВАНИЯ ИЗОБРАЖЕНИЯ ДЛЯ УВЕЛИЧЕНИЯ ОБЪЕМА ВСТРАИВАЕМОГО СООБЩЕНИЯ

А.Д. Белов
ORCID: **0000-0003-3845-2320** e-mail: **anton98belov@mail.ru**
Нижегородский государственный технический университет им. Р.Е. Алексеева
Нижний Новгород, Россия

В.Ю. Карпычев
ORCID: **0000-0001-8527-2600** e-mail: **kavlyr@yandex.ru**
Нижегородский государственный технический университет им. Р.Е. Алексеева,
Приволжский институт повышения квалификации ФНС России
Нижний Новгород, Россия

Разработан алгоритм модификации стеганографического метода квантования изображения для увеличения объема скрываемого сообщения. Для достижения цели предлагается использовать методы алгоритма сжатия сообщения CCITT Group 3, основанного на группировке цепочек одинаковых бит. При этом модифицированный алгоритм предполагает несколько итераций сжатия сообщения выбранным методом для поиска оптимального числа сжатий, позволяющих с максимальным коэффициентом достигнуть результата. Разработанный алгоритм применим также и к ряду стеганографичеких методов, в результате которых будут изменены байты пикселей исходного изображения-контейнера. Наибольшая эффективность достигается при сокрытии сообщений, длины подстрок которых имеют примерно одинаковую длину при одной или нескольких итерациях сжатия. В перспективе предложенный алгоритм может быть модифицирован для сокрытия сообщений, состоящих из подстрок, всегда имеющих сильно различающиеся длины при любом числе итераций сжатия.

Метод, основанный на предполагаемом алгоритме, может быть применен в областях, где необходимо скрытно передавать или хранить большие сообщения в наиболее малом файле-контейнере, например, при общении абонентов.

Ключевые слова: защита информации, стеганография, сокрытие данных, стегосистема, квантование изображения, изображение, сжатие, метод Хаффмана, CCITT Group 3.

ДЛЯ ЦИТИРОВАНИЯ: Белов, А.Д. Модификация стеганографического метода квантования изображения для увеличения объема встраиваемого сообщения / А.Д. Белов, В.Ю. Карпычев // Труды НГТУ им. Р.Е. Алексеева. 2023. № 2. С. 7-13. DOI: 10.46960/1816-210X_2023_2_7

MODIFICATION OF IMAGE QUANTIZATION STEGANOGRAPHY TECHNIQUE FOR VOLUME ENLARGEMENT

A.D. Belov
ORCID: **0000-0003-3845-2320** e-mail: **anton98belov@mail.ru**
Nizhny Novgorod State Technical University n.a. R.E. Alekseev
Nizhny Novgorod, Russia

V.Yu. Karpychev
ORCID: **0000-0001-8527-2600** e-mail: **kavlyr@yandex.ru**
Nizhny Novgorod State Technical University n.a. R.E. Alekseev,
Federal State Institution of Advanced Finance Professional Training
Federal Tax Service Institute
Nizhny Novgorod, Russia

Abstract. Developed is an algorithm for modification of image quantization steganography technique for volume enlargement of encapsulated message. For that purpose, the authors propose to use CCITT Group 3 message compression algorithm techniques based on identical bit chain clusterization. At that, the modified algorithm assumes multiple iterations of message compression using the chosen technique to find the optimal number of compressions to compress the message to a maximum degree. The developed algorithm is also applicable to a number of steganography techniques providing alteration to pixel bytes of an original carrier image. This algorithm is most effective in encapsulating messages with almost identical substring lengths in one or more compression iterations. In the long term, the proposed algorithm can be modified to encapsulate messages with highly different substring lengths at any number of compression iterations. A technique based on the proposed algorithm can be applied in fields where large messages need to be shared or stored within the smallest carrier file, e.g. at subscribers' communication.

***Key words**:* information security, steganography, data encapsulation, stegosystem, image quantization, image, compression, Huffman coding, CCITT Group 3.

FOR CITATION: A.D. Belov, V.Yu. Karpychev. Modification of image quantization steganography technique for volume enlargement. Transactions of NNSTU n.a. R.E. Alekseev. 2023. № 2. Pp. 7-13.
DOI: 10.46960/1816-210X_2023_2_7

Введение

Обеспечение конфиденциальности коммуникаций является одной из целей защиты информации, для которой используются криптографические и стеганографические методы скрытия. Применение криптографии обеспечивает бóльшую стойкость, однако при определенных условиях стеганография также широко используема в цифровых коммуникациях. При этом существующие методы компьютерной стеганографии имеют ряд недостатков: не обеспечивают полную скрываемость сообщения и не рассчитаны на большой объем данных.

Анализ известных стеганографических методов, предназначенных для скрытия и незаметной передачи конфиденциальных данных, показал, что максимальный их объем не превышает 25 % размера файла контейнера [1]. Это определяет актуальность исследований, направленных на повышение указанного значения объема скрываемых данных.

Теоретический анализ

Разрабатываемый алгоритм модификации применяется к методу квантования изображения; также он применим к большинству методов, меняющих байты пикселей изображения, при небольших изменениях этапа записи бит. Данный метод основывается на зависимости между пикселями [2-3], которая описывается функцией $\Delta=\Theta(x)$. Простейшим случаем является использование разницы c_i-c_{i+1} как функции Θ, где c_i – значение цвета пикселя. Встраивание информации происходит путем изменения разницы Δ. Бит сообщения под номером i скрывается при помощи вычисления Δi. Если значение встраиваемого бита не соответствует значению из стеганоключа, Δi заменяется на Δ, значение которой равняется значению встраиваемого бита. Для этого меняются значения цвета пикселя.

Для построения стеганоключа используется псевдослучайная последовательность: каждому возможному значению Δ ставится в соответствие значение бита, сгенерированного с помощью генератора псевдослучайной последовательности (ГПСП). Такой метод имеет большой недостаток: для данного алгоритма размер сообщения мал по сравнению с размером файла контейнера. Например, в изображении-контейнере, имеющем 1000 пикселей, где каждый пиксель представлен в виде трех цветов RGB, принимающих значения от 0 до 255,

можно зашифровать сообщение с максимальной длиной (1000-1)*3=2997 бит, а отношение максимального размера сообщения к размеру файла-контейнера будет при этом составлять 2997/24000=~0,125.

В работе предлагается для решения задачи применить методы, основанные на сжатии сообщения методом CCITT Group 3 [4]. Стандартный метод CCITT Group 3 имеет следующий алгоритм:

1) чтение строки одинаковых бит;
2) подсчет длины полученной строки;
3) замена строки на строку вида <бит, длина строки>.

Часто сжатая строка имеет фрагменты, в которых есть последовательности сжатых строк одинаковых бит. Для таких случаев и предназначается обозначение бита в сжатой строке. Однако, если метод будет считывать строки одинаковых бит вплоть до противоположного бита, сжатая строка имеет избыточность: обозначение бита будет излишним. Поэтому такой подход при опускании обозначений бит позволит получить меньшую по длине сжатую строку. Обозначение длин строк будет также преобразовано в битовые строки, однако каждая битовая строка будет иметь различную длину. При сохранении в файл-контейнер будет очень затратно хранить длину каждой строки, поэтому необходимо выбрать некоторую статическую ее длину, позволяющую наиболее оптимально сжать изначальное сообщение. Полученную строку можно также свернуть методом *CCITT Group 3*, сохраняя длины строк в некоторую таблицу на всех итерациях сжатии. Также полученная строка должна иметь некоторую добавочную информацию, позволяющую точно определить длину подстрок и всего сообщения.

Метод

Общий модифицированный алгоритм сжатия может быть реализован следующим образом (рис. 1).

1. *Получение всех длин строк.* Для этого система «пробегает» по всему исходному файлу и добавляет запись <длина строки, количество строк> в таблицу длин строк.

2. *Расчет наиболее подходящей длины строк.* При этом система проходит по таблице длин подстрок, рассчитывает длины сжатых строк и выбирает оптимальную длину строки.

Для расчета длины сжатой строки выведена формула:

$$L = \left(num_{bitLen \le optBitLen} + \sum_{bitLen > optBitLen} \left(\frac{len}{optlen} \right) * 2 - 1 \right) * bit , \qquad (1)$$

где $num_{bitLen \le optBitLen}$ – число подстрок, не превосходящих при сжатии размер текущей сжатой подстроки из таблицы длин, $\sum_{bitLen > optBitLen} \left(\frac{len}{optlen} \right)$ – число подстрок сжатого сообщения, получившихся при разбиении подстрок исходного сообщения (превосходящих при сжатии размер текущей сжатой подстроки из таблицы длин) по размеру текущей сжатой подстроки из таблицы длин, bit – размер текущей сжатой подстроки из таблицы длин.

3. *Сравнение длины полученной оптимальной строки с 1.* Если длина равна 1, выполняется преобразование методом разностей бит.

4. *Сжатие строки.* Для этого строки с меньшей или равной по числу бит длиной преобразуются в число, которое затем преобразуется в последовательность бит, по длине равную оптимальной длине строки. Строки с большей длиной разбиваются на последовательность подстрок. Для сохранения принципа чередования бит в строках между подстроками ставятся подстроки противоположных бит с нулевой длиной.

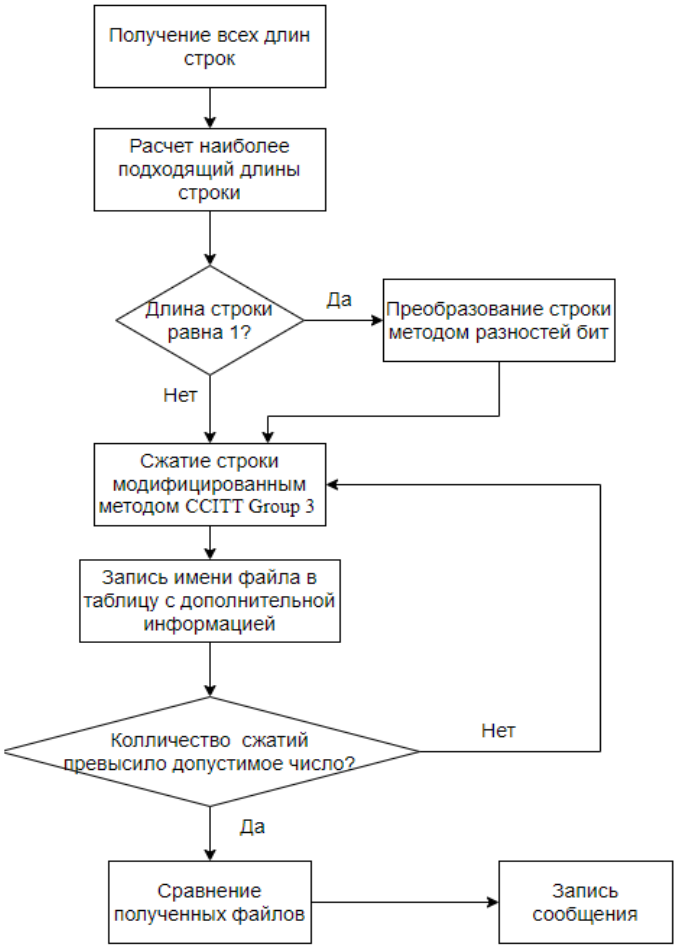

Рис. 1. Алгоритм сжатия сообщения

Fig. 1. Message compression algorithm

5. *Запись файла в таблицу с дополнительной информацией.* При этом алгоритм предпринимает попытку встраивания сообщения в файл и получает смещение – разность между позициями бит, и количества позиций от последнего бита до конца слоя цвета. Для этого система использует метод записи сообщения (рассмотренный позже) во временный файл-копию исходного изображения.

6. *Проверка на ограничение сжатий.* Поскольку добавочная информация помещается в первую строку бит изображения, а длины строк при каждом сжатии необходимо сохранять, помещая их в добавочную информацию, количество сжатий ограничено оставшимися незанятыми битами первой строки каждого слоя.

7. *Сравнение полученных файлов.* Алгоритм пробегается по таблице файлов и ищет файл с наименьшей строкой и наибольшей степенью сжатия.

8. *Запись сжатого сообщения в файл.* Отличительной особенностью алгоритма квантования изображения является встраивание сообщения в виде бит в разности между соседними байтами. Однако стандартный подход малоэффективен при встраивании сжатого сообщения, следовательно, его также необходимо модифицировать. Для этого требуется преобразовать байты цветов изображения в 8-битовую последовательность и записывать биты сообщения в разности соседних нулевых и первых бит. При этом для наибольшего объема изображения контейнера требуется определить размерность изображения и, если длина меньше высоты, записывать сообщение в столбцы, в противном случае – в строки. Для встраивания сообщения используется следующий алгоритм (рис. 2).

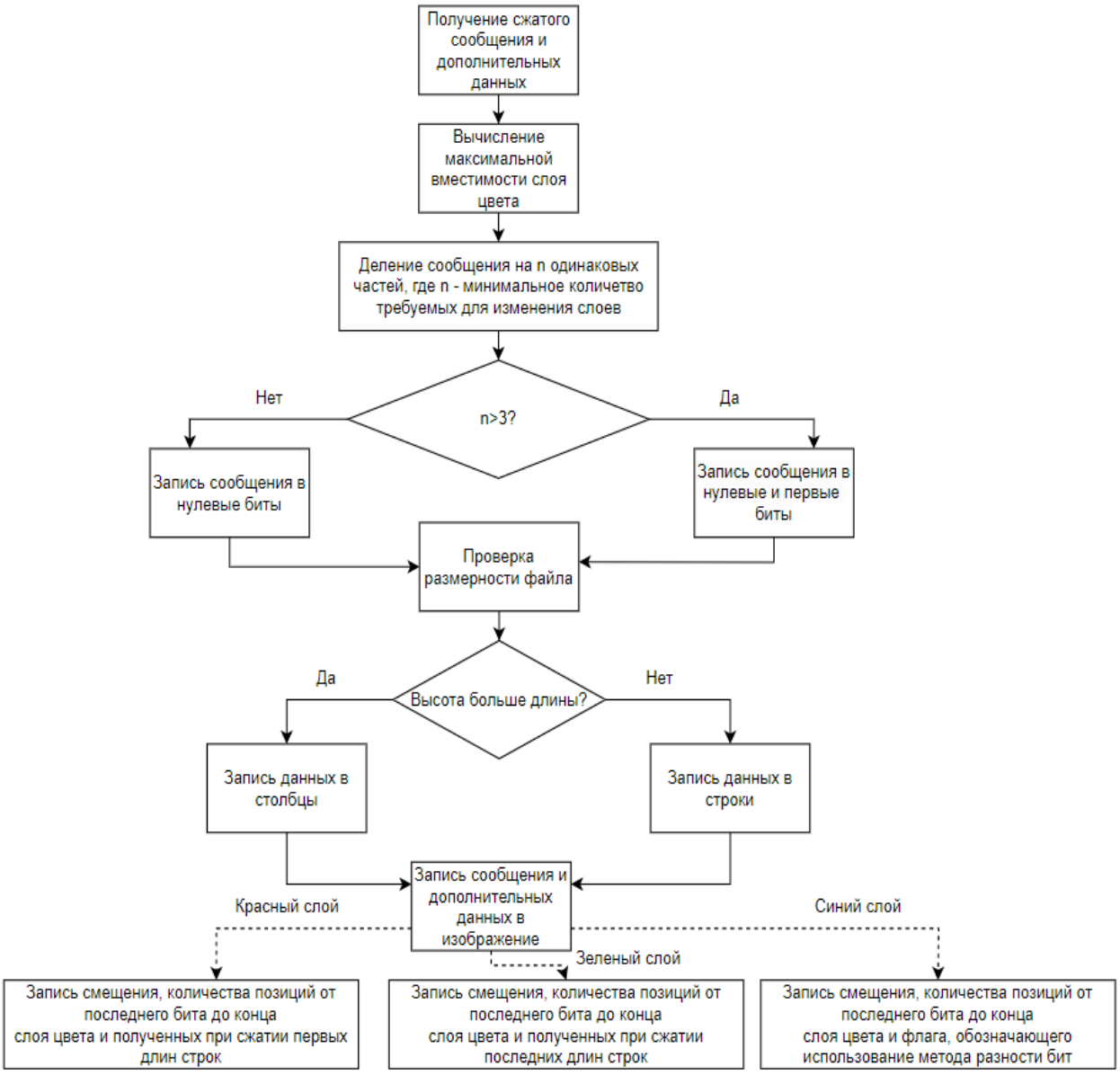

Рис. 2. Алгоритм встраивания сообщения

Fig. 2. Message embedding algorithm

1. *Получение сжатого сообщения и дополнительных данных.* На этом этапе алгоритм получает данные из части сжатия сообщения.

2. *Вычисление максимальной вместимости слоя цвета.* Максимальная вместимость вычисляется по формуле:

$$cap_{max} = \left(height_{image} - 1\right) * \left(width_{image} - 1\right) \quad (2)$$

3. *Деление сообщения на n одинаковых частей,* где n – минимальное количество требуемых для изменения слоев. При этом, если слоев меньше трех, запись сообщения происходит в нулевые биты.

4. *Проверка размерности файла.* Данный этап требуется для определения типа записи сообщения: если длина меньше высоты, запись сообщения происходит в столбцы, в противном случае – в строки.

5. *Запись сообщения в изображение.* Сообщение встраивается с выбранными на предыдущих этапах условиями.

6. *Запись дополнительных данных в изображение.* Сначала в каждый слой записывается бит использования слоя. Затем в нулевые строки/столбцы красных и синих слоев записываются смещение, количество позиций от последнего бита до конца слоя цвета и полученные при сжатии длины строк. В нулевую строку/столбец зеленых слоев записываются смещение, количество позиций от последнего бита до конца слоя цвета и флаг использования метода разности бит. Эти строки записываются в нулевые биты слоев. В первые биты записываются нули и единицы в соответствии со следующими шагами:

1) если соответствующий первому биту нулевой бит – конец строки, то необходимо записать значение 3 бита;

2) если соответствующий первому биту нулевой бит – не конец строки, то необходимо записать значение противоположное значению 3 бита.

Такой алгоритм исключает вероятность того, что большая часть первых бит будет равна 0. На данном этапе сама запись бит происходит следующим образом: если значение цвета пикселя не равно 0 или 255, при этом нужно изменить оба первых бита, тогда алгоритм вычитает/прибавляет 1 в соответствии со следующей таблицей (табл. 1).

Таблица 1.
Изменение первых бит пикселя

Table 1.
Alteration of initial pixel bits

Исходный байт	Операция	Полученный байт
…0100	-1	…0011
…0101	+1	…0110
…0110	-1	…0101
…0111	+1	…1000

Данный способ позволит уменьшить максимальное отклонение от исходного изображения, что также позволит увеличить незаметность встроенного сообщения.

Заключение

Предложенный алгоритм основывается на методе сжатия сообщения *CCITT Group 3* (со средней степенью сжатия 2) [5] и методе квантования изображения (с вместимостью контейнера 12,5 %). За счет этих двух факторов и изменения алгоритма сокрытия, позволяющего скрывать пару битов в зависимости между двумя пикселями, теоретическая вместимость контейнера составляет 2 * 12,5% * 2 = 50 %.

Разработанный способ сокрытия основан на алгоритме, меняющем младшие биты изображения, следовательно, предложенная методика модификации может быть применена и к ряду других методов, основанных на изменениях младших бит файла-контейнера.

Поскольку длины сжатых подстрок в данном алгоритме должны быть одинаковы, такой алгоритм наиболее эффективно применяется при сокрытии сообщений, длины подстрок которых имеют примерно одинаковую длину при одной или нескольких итераций сжатия. В противном случае алгоритм будет менее эффективен за счет добавления незначащих нулей к коротким сжатым подстрокам и увеличения длин этих подстрок.

В дальнейших исследованиях предполагается практическое подтверждение вместимости контейнера в разработанном методе. Также предложенный алгоритм может быть модифицирован для сокрытия сообщений, состоящих из подстроки, всегда имеющих сильно различающиеся длины при большом числе итераций сжатия.

Библиографический список

1. **Абазина, Е.С**. Сравнительный анализ и классификация методов цифровой и компьютерной стеганографии, и перспективные направления ее развития / Е.С. Абазина, А.А. Ерунов // Труды Военно-космической академии имени А.Ф. Можайского. 2016. № 655. С. 5-16.
2. **Котцов, В.А.** Стеганографическое использование структуры сигнала цифрового изображения / В.А. Котцов, П.В. Котцов // Цифровая обработка сигналов. 2021. № 1. С. 44-50.
3. Метод квантования изображения [Электронный ресурс] // Режим доступа: https://studfile.net/preview/7379018/page:32/ (дата обращения 13.01.2023).
4. **Ватолин, Д.** Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео / Д. Ватолин, А. Ратушняк, М. Смирнов, В. Юкин. – М.: ДИАЛОГ-МИФИ, 2003. – 384 с.
5. Методы сжатия данных: Сжатие изображений. Часть 2 [Электронный ресурс] // Режим доступа: http://www.compression.ru/book/part2/part2__2.htm (дата обращения 20.01.2023).

Дата поступления
в редакцию: 17.04.2023

Дата принятия
к публикации: 01.06.2023