УДК 004.415.24

И.В. Швидченко

КРИПТОСТЕГАНОГРАФИЧЕСКИЙ АЛГОРИТМ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ СЕГМЕНТАЦИИ

Использование телекоммуникационных систем связи и глобальной информационной сети Интернет вызвало в последнее время активное развитие как криптографических, так и стеганографических методов защиты информации. Криптографическая защита обеспечивает гарантируемое блокирование несанкционированного доступа к информации, а стеганографическая защита обеспечивает сокрытие самого факта передачи информации. Интерес к стеганографии можно подтвердить тем фактом, что в настоящее время на рынке программных продуктов предлагается большое количество стеганографических программ, востребованных обычными пользователями и коммерческими организациями. Однако отсутствие средств сертификации и методов проверки надежности стеганографических систем приводит к тому, что использование последних не может гарантировать целостность и конфиденциальность информации без использования с ними в комплексе криптографических средств защиты.

Известно [1], что построение реальных ключевых стеганосистем — достаточно сложная задача, требующая разработки легко перестраиваемых алгоритмов сокрытия информации, что в большинстве случаев либо нерационально, либо просто невозможно в связи с предъявляемыми требованиями к уровню скрытности и ограничениями на полосу передачи данных. Поэтому на практике преимущество отдается бесключевым стеганосистемам вида $\Sigma = (C, M, S, H, R)$, где h и r — независящие от ключа правила встраивания и извлечения сообщений. Безопасность таких стеганосистем основывается только на секретности используемых стеганографических преобразований H и R, что противоречит принципу Керхгоффса для систем защиты информации [2]. Поэтому для повышения безопасности безключевых систем одним из возможных выходов может быть их совместное использование с проверенными на стойкость криптографическими алгоритмами и протоколами. В случае информированности нарушителя о стеганографическом преобразовании успешная атака на системы данного вида возможна в случае успешной атаки на криптографические алгоритмы, т.е. стойкость системы к атакам пассивного противника будет определяться стойкостью к взлому криптографической части. Организация внутренней криптозащиты в структуре стеганографической системы позволит, во-первых, обезопасить информацию пользователя от раскрытия ее содержания в случае извлечения сообщения, во-вторых, изменить статистические характеристики сообщения — достичь равномерности распределения секретного сообщения по контейнеру. Ключ шифрования в данном случае выступает параметром ключевой информации стеганографической системы.

Систему скрытой передачи информации в открытых каналах связи, основанную на совместном применении криптографических алгоритмов, стеганографических методов, а также алгоритмов согласования входных и выходных данных указанных алгоритмов и методов, назовем криптостеганографической.

В качестве алгоритма шифрования можно взять популярный в настоящее время алгоритм Advanced Encryption Standard (AES), также известный как Rijndael, разработанный в 2000 г. двумя криптографами из Бельгии — В. Бриджменом и Дж. Даймен [3]. Алгоритм представляет собой симметричный блочный

© И.В. ШВИДЧЕНКО, 2010

шифр, который имеет ряд преимуществ: высокую скорость шифрования, минимальные требования к вычислительным ресурсам, стойкость к подавляющему большинству атак по времени выполнения и потребляемой мощности и возможность использования любых комбинаций размеров блока и длин ключа, кратных 32 бит. Стойкость Rijndael к дифференциальному и линейному методам криптоанализа эквивалентна сложности простого перебора ключей. Алгоритм не защищен патентами и доступен для свободного использования в любых продуктах.

На сегодняшний день широко применяемым методом сокрытия информации в изображениях является метод замены наименее значимых битов (НЗБ), предложенный Е. Адельсоном в 1990 г. [4]. Идея метода заключается в замене от одного до четырех младших битов в байтах цветового представления точек исходного изображения битами скрываемого сообщения. Возможность такой замены обусловлена наличием в изображениях структурной избыточности.

Наиболее удобными для использования контейнерами являются растровые изображения в режиме True Color, т.е. полноцветные изображения, в которых под хранение информации о цвете каждой точки отводится по три байта. Каждый содержит информацию об интенсивности одной из трех составляющих цвета (палитра RGB): красной, синей, зеленой. Так как под хранение каждой из составляющих отводится по одному байту, то соответствующие им интенсивности цветов изменяются в пределах от 0 до 255. Таким образом, общее число возможных цветов равно $256^3 \approx 1{,}6 \cdot 10^7$. Известно [5], что человеческий глаз способен различить лишь порядка четырех тысяч цветов, а для хранения такого количества достаточно всего четырех битов. При применении метода НЗБ предполагается, что цвет точки практически в полной мере определяется старшими четырьмя битами каждого из трех байтов представления RGB, следовательно, оставшиеся четыре младших бита можно использовать для внедрения скрываемой информации.

Исследования особенностей человеческого зрения [6] показали, что порог чувствительности глаза к изменению освещенности при средних ее значениях составляет $\Delta I = 0{,}01 - 0{,}03I$ или 1–3 % (рис. 1).

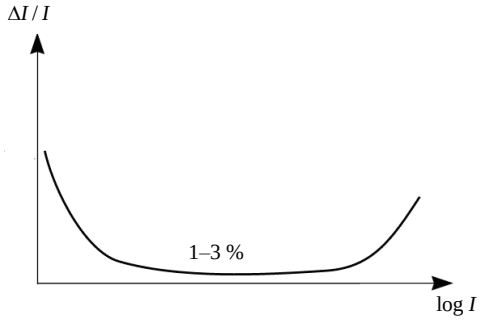


Рис. 1

Использование для внедрения информации четырех младших разрядов в байтах исходного изображения может привести к изменению интенсивности порядка 6 %, что в два раза превышает порог чувствительности человеческого глаза. Даже замена трех младших битов (~3 %) вносит заметные для человеческого глаза искажения. Изменение яркости в пределах 1–1,5 % остается незаметным. Следовательно, для того чтобы внедрение в изображение дополнительной информации оказалось незаметным для человеческого глаза, оптимальным будет подвергнуть модификации либо первый и второй разряды (максимальное искажение 1,17 %), либо только третий разряд (искажение 1,56 %) каждого из трех байтов, отвечающих за цвет точки [7].

Поскольку результат внедрения должен остаться незаметным, то немаловажным является выбор подходящего контейнера, от этого зависит объем внедряемых

данных и стойкость стеганоконтейнера к методам пассивного стеганоанализа, т.е. вероятность обнаружения наличия скрытого сообщения нарушителем. Выбор контейнера-оригинала необходимо рассматривать с точки зрения метода внедрения данных, так как именно он определяет биты, которые будут модифицированы на биты сообщения. Из-за достаточно большого количества методов сокрытия данных проблема выбора подходящего контейнера до сих пор не решена. В работах [8, 9] предложены основные критерии выбора, которые в достаточной мере учитывают особенности контейнера, необходимые для получения стеганоконтейнера, устойчивого к визуальному стеганоанализу в случае использования метода замены младших битов.

Контейнер должен удовлетворять таким критериям: зашумленность, отсутствие плавных переходов цветов и монотонных областей, «пестрота» и большое число перепадов яркости. При выборе не следует использовать общеизвестные изображения (например, картину «Джоконда») и изображения, конвертированные из JPEG-формата в формат BMP. JPEG-компрессия оставляет «след» в изображении, который может быть обнаружен. После встраивания информации этот «след» сохраняется, и существует способ определить, что данное изображение не могло быть получено лишь в результате декомпрессии JPEG-изображения, что вызовет подозрения. Лучше использовать изображения, полученные с помощью фотоаппарата или сканера, а не с помощью графических редакторов.

После внедрения данных в подходящий контейнер важно оценить стойкость модифицированного контейнера к известным методам стеганоанализа. В реальных условиях наиболее типичным видом атаки нарушителя является атака на стеганоконтейнер, так как исходный контейнер ему, как правило, неизвестен. В этих условиях обнаружение скрытого сообщения возможно на основе выявления нарушений зависимостей, присущих естественным контейнерам.

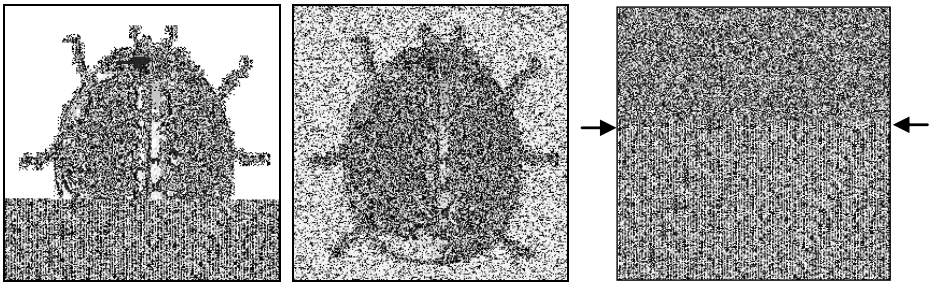
Наиболее эффективным методом стеганоанализа изображений, в которых информация внедрялась в младшие разряды элементов контейнера, является *метод визуального анализа битовых срезов*. Основная идея метода заключается в сравнении изображения в целом с изображениями его битовых срезов [7]. С помощью несложной программы изображение просматривают по слоям, т.е. берутся битовые срезы изображения, затем стеганоаналитик сначала смотрит на изображение, отмечает для себя некоторые его элементы и пытается сопоставить их с элементами изображения конкретного битового среза. Учитывая, что интенсивность каждого цвета определяется одним байтом, всего необходимо просмотреть восемь таких срезов. Для каждого из трех цветов первый срез — это изображение, построенное самыми младшими битами, второй срез — изображение, построенное вторыми битами и т.д. Далее полученное изображение битового среза просматривают и визуально сравнивают с анализируемым изображением.

Известно несколько методов встраивания данных в изображение-контейнер [10], которые применяются в настоящее время в реализациях стеганопрограмм, предлагаемых на рынке программных продуктов.

1. *Метод последовательной замены.* Младший бит цветовой компоненты каждого пиксела, начиная от начала изображения, последовательно заменяется соответствующим битом скрываемого сообщения. При этом встраиваемое сообщение, не использующее максимально возможный объем, оставляет часть контейнера нетронутым. В этом случае достаточно легко определить, в какой именно части контейнера находится сообщение (рис. 2, *а*).

2. *Метод распределенного внедрения*. Сообщение распределяется на весь контейнер. Нет четкой границы между частью изображения, не тронутой при внедрении сообщения, и частью, измененной стеганографическим преобразованием (рис. 2, *б*).

3. *Метод внедрения с заполнением.* Встраиваемое сообщение любой длины дополняется до длины контейнера (до числа пикселов изображения), при этом битовый срез стеганоконтейнера не содержит следов исходного изображения. Следует отметить, если скрываемое сообщение до встраивания не шифруется, то его вероятностные характеристики не совпадают с вероятностными характеристиками младших битов используемого контейнера, в отличие от предварительно зашифрованного сообщения, где каждый его бит практически равновероятен и независим от соседних битов. Визуально это можно выявить, сопоставляя изображения битового среза с внедренным шифрованным и нешифрованным сообщением [11]. На рис. 2, *в* (метод внедрения с заполнением) видно, что нижняя часть изображения содержит нешифрованные скрытые данные.



а *б* *в*

Рис. 2

Для примера использовалось изображение с небольшим количеством цветов и монотонными областями (рис. 3). Как видно из рис. 2, для метода визуального анализа битовых срезов большое значение имеет то, как именно осуществляется запись скрываемой информации, если она записывается в подряд идущие биты или равномерно распределяет биты сообщения (на основе генератора псевдослучайных чисел) по всему пространству изображения, то факт сокрытия может быть безошибочно установлен. Данный пример демонстрирует, что выбор контейнера такого рода непригоден для сокрытия информации одним из указанных методов.

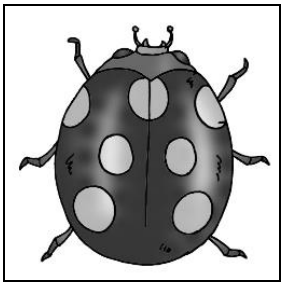


Рис. 3

Найдем подход сокрытия информации в подобных контейнерах, после применения которого невозможно было бы обнаружить наличие внедренных данных методами визуального и статистического анализа.

Воспользуемся методами сегментации, которые применяются и играют важную роль в задачах обработки изображений и компьютерного зрения.

Пусть C обозначает всю область изображения. Сегментацией называется процесс разбиения области C на n подобластей $C_1, C_2, \ldots, C_n$ таких, что:

а) $\bigcup_{i=1}^{n} C_i = C;$

б) C_i является связной областью, $i = 1, 2, \ldots, n;$

в) $C_i \cap C_j = \varnothing$ для всех $i,\ j = 1, 2, \ldots, n,\ \ i \neq j$;

г) $P(C_i) = \text{TRUE}$ при $i = 1, 2, \ldots, n$;

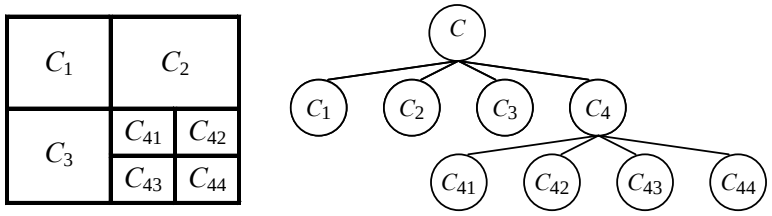
д) $P(C_i \cup C_j) = \text{FALSE}$ для любых двух смежных областей C_i и C_j.

Здесь $P(C_i)$ — некоторый логический предикат, определенный на точках множества C_i, а $\varnothing$ — пустое множество. Логическим предикатом называется функция, принимающая только два значения: истина (TRUE) и ложь (FALSE).

Условие а) означает, что сегментация должна быть полной, т.е. каждый пиксел изображения принадлежит какой-то области. Условие б) подтверждает, что точки каждой области должны быть связными в некотором оговоренном смысле (например, они 4- или 8-связны). Условие в) указывает на то, что области не должны пересекаться. Условие г) относится к свойствам, которые должны соблюдаться для всех пикселов из одной области, например $P(C_i) = \text{TRUE}$, если все пикселы в C_i имеют одинаковую яркость. Условие д) означает, что прилегающие области C_i и C_j различаются в смысле предиката P.

Рассмотрим один из методов сегментации — разделение/слияние областей. Пусть вся область изображения обозначена C и выбран предикат P. Один из подходов к сегментации C состоит в том, чтобы последовательно разбивать эту область на все более и более мелкие квадратные подобласти C_i, пока выполняется условие $P(C_i) = \text{TRUE}$. Процесс начинается со всей области изображения. Если $P(C) = \text{FALSE}$, то изображение делится на четверти вертикальной и горизонтальной прямыми, проходящими через середину. Если для какой-то четверти предикат P принимает значение FALSE, то она аналогичным способом делится на более мелкие четверти и т.д. (рис. 4, *а*).

Такой метод разбиения удобно представлять в форме так называемого квадродерева — дерева, у которого вершины, не являющиеся листьями, имеют в точности четыре потомка (рис. 4, *б*). Подизображения, которые соответствуют узлам квадродерева, иногда называются квадрообластями, или квадроизображениями. Отметим, что корень дерева соответствует целому изображению, а каждая вершина — какой-то его подобласти [12].



а *б*

Рис. 4

Если использовать только операцию разделения, то в окончательном разбиении изображения могут присутствовать соседние области с одинаковыми свойствами. Этот недостаток устраняется операцией слияния. Для соблюдения ограничений а)–д) требуется, чтобы слиянию подвергались только соседние области, пикселы которых в совокупности удовлетворяют предикату P. Другими словами, две соседних области C_j и C_k сливаются только в том случае, если $P(C_j \cup C_k) = \text{TRUE}$.

Изложенное выше можно кратко суммировать в виде процедуры, на каждом шаге которой выполняются следующие действия:

1) любая область C_j, для которой $P(C_j) = \text{FALSE}$, разделяется на четыре непересекающиеся четверти;

2) любые две соседние области C_j и C_k, для которых $P(C_j \cup C_k) = \text{TRUE}$, объединяются в одну;

3) если невозможно выполнить ни одной операции слияния или разделения, то процедура заканчивается.

Приведем пошаговое описание криптостеганографического алгоритма, в котором используется описанный выше метод сегментации.

Алгоритм.

1. Пусть отправитель имеет сообщение m, которое необходимо скрытно передать по открытому каналу связи, длиной N битов.

2. С помощью алгоритма псевдослучайных чисел, отвечающих требованиям стойкости генерируемых данных, генерируется ключ k_1 и выполняется шифрование сообщения одним из стойких криптографических алгоритмов, $m' = E(m, k_1)$.

3. С использованием общего с получателем генератора ключей формируется сеансовый ключ k_2, с помощью которого выбирается исходный контейнер c в цветовом пространстве RGB.

4. Выделяются соответствующие каждому цвету массивы (цветовые составляющие), записываются в единый массив и осуществляется сегментация методом разделения/слияния областей (критерием однородности выступает равенство всех пикселов области). В результате получается набор однородных областей различного размера. Для встраивания сообщения используются области, состоящие из одного пиксела. Количество единичных областей с указанным критерием однородности равно Q.

5. Если необходимо передать сообщение m' большего размера, чем количество единичных областей Q, то отправитель должен выбрать другой контейнер или подвергнуть модификации и второй разряд каждого из трех байтов.

6. Для большей защищенности информации перед скрытием можно временно изменить порядок цветовых матриц. Встраивание выполняется по столбцам последовательно, поэтому если $m' << Q$, то будет заполнена часть изображения, а часть будет не тронута. Для распределения сообщения по изображению необходимо подобрать такой критерий однородности, чтобы после сегментации количество единичных областей было близким к количеству битов в скрываемом сообщении. Данный критерий выступает в качестве ключевой информации стеганографической системы. Отправитель устанавливает предикат P, который может быть основан как на свойствах яркости, так и на свойствах текстуры, цвета, глубины. В результате выполнения сегментации с указанным предикатом получается количество единичных областей R, близкое к количеству битов в скрываемом сообщении.

7. С помощью метода замены НЗБ строится стеганоконтейнер $s = h(c, m')$. В предлагаемой реализации замене подвергается только первый наименее значимый бит.

8. Криптостеганоконтейнер s передается по открытому каналу связи, а ключевая информация k_1, k_2 и P — по закрытому.

9. Принимающая сторона, определив по ключу k_2 исходный контейнер, выделяет соответствующие каждому цвету массивы, переводя значения цветовых

составляющих каждого пиксела изображения в числовые матрицы. При необходимости выставляет их в нужной последовательности. Используя предикат P, осуществляет сегментацию.

10. Зная стеганопреобразование и местонахождение единичных блоков, получатель извлекает из стеганоконтейнера зашифрованное сообщение $m' = r(s)$.

11. С помощью секретного ключа k_1 и известного криптографического алгоритма расшифровывает $m = D(m', k_1)$ и читает скрытое сообщение.

Так как объем файла в BMP формате достаточно велик, то для передачи стеганоконтейнера его можно конвертировать в формат PNG, использующий алгоритм обратимого сжатия без потерь LZ77.

Рассмотрим пример и оценим стойкость предложенного криптостеганографического алгоритма к пассивным атакам.

В качестве сообщения возьмем файл формата TXT с количеством символов $m = 3745$. Используя реализацию алгоритма AES с длиной ключа и блока 128 бит, и числом раундов равным 10, выполним шифрование сообщения. Количество символов в сообщении после шифрования $m' = 3920$. Количество НЗБ, необходимое для сокрытия сообщения, $N' = 8 \times 3920 = 31360$ бит.

Выбираем изображение-контейнер размером 256×256 пикселов (см. рис. 3). Выделяем цветовые составляющие и размещаем в общем массиве. Проводим сегментацию данного массива, где в качестве критерия однородности берем равенство всех пикселов области. Получаем количество единичных областей, равное $Q = 84056$. $Q >> N'$, поэтому для распределения сообщения по изображению устанавливаем такой критерий однородности (предикат), чтобы после сегментации количество единичных областей было близким к количеству битов в скрываемом сообщении. В данном примере область будет считаться однородной, если разница между максимальным и минимальным значением пикселов области меньше некоторого параметра t (данный параметр подбирался экспериментально). При $t = 0{,}035$ количество единичных областей равно $R = 32652$. На рис. 5 представлен результат сегментации с указанным параметром. Скрываем сообщение.

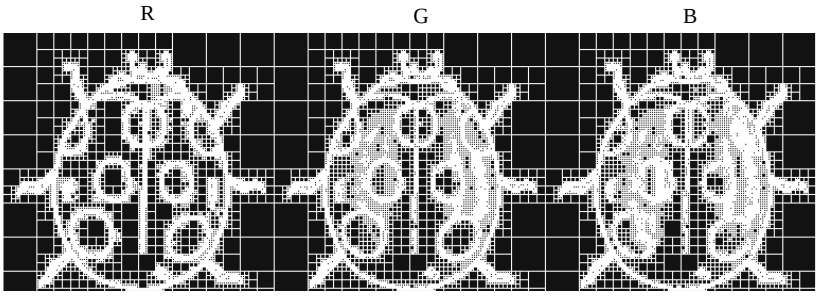
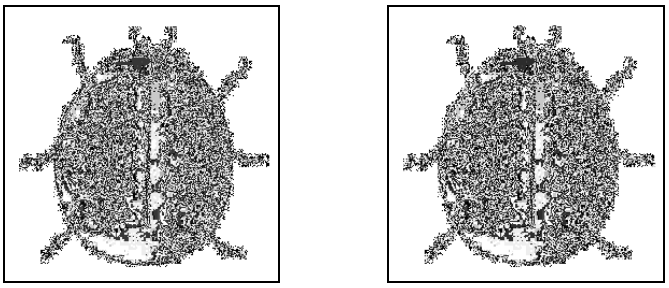


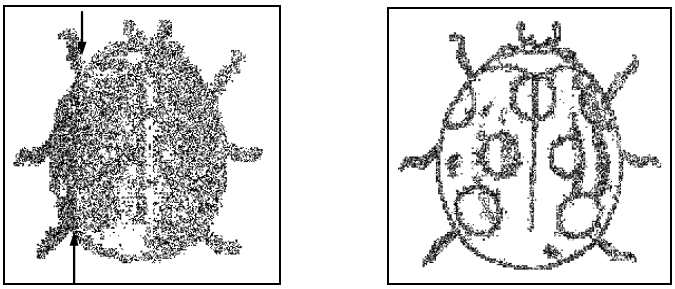
Рис. 5

При визуальном анализе контейнеров, представляющих собой графические файлы формата BMP, в качестве элементов анализируемой последовательности следует выбирать все наименее значащие биты. Поскольку в предложенном алгоритме сокрытия данные внедрялись в последние биты, анализировался битовый срез по последнему биту. Из рис. 6 видно, что битовый срез по последнему биту исходного изображения (*а*) визуально ничем не отличается от битового среза по последнему биту стеганоконтейнера (*б*). На рис. 7 показана разница между исходным контейнером и стеганоконтейнером с разными предикатами (*а* — значения яркости всех пикселов области равны, *б* — установленный отправителем).



а *б*

Рис. 6



а *б*

Рис. 7

Статистические методы базируются на понятии «естественного» контейнера. Суть методов заключается в оценивании вероятности существования стеганографического вложения с неизвестной стеганосистемой на основе критерия оценки близости исследуемого контейнера к «естественному». В работе использовался один из известных статистических критериев для оценки стойкости стеганосистемы против атак пассивного противника — критерий Хи-квадрат [10]. Как показали результаты, статистическая атака не обнаруживает следы вложения информации.

К преимуществам статистических методов относится практически неограниченная область применения. Основным же недостатком является само предположение о существовании «естественного» контейнера. Файл с изображением в формате BMP и глубиной 24 бит может содержать произвольные данные, что вполне естественно, но отличаться от известных моделей «естественного» контейнера.

В системе Matlab написана программа, реализующая описанный алгоритм. Использование предложенного способа сокрытия информации позволит получить стеганоконтейнер, устойчивый к атакам пассивного противника и к геометрической атаке типа «обрезание краев».

І.В. Швідченко

КРИПТОСТЕГАНОГРАФІЧНИЙ АЛГОРИТМ З ВИКОРИСТАННЯМ МЕТОДІВ СЕГМЕНТАЦІЇ

Запропоновано новий стеганографічний підхід приховання інформації в зображеннях методом найменших значущих бітів. Основна ідея запропонованого підходу — використання для вкраплення інформації областей, отриманих в результаті сегментації зображення. Підхід забезпечує стеганостійкість до візуального і статистичного аналізу і стійкість до витягування змісту повідомлення за рахунок використання в структурі стеганографічної системи блокового криптографічного алгоритму.

I.V. Shvidchenko

CRYPTOSTEGANOGRAPHIC ALGORITHM USING METHODS OF SEGMENTATION

A new steganographic approach of data hiding in images by the method of least significant bits (LSB) is offered. The main concept of the proposed approach is the use for embedding information of areas, got as the result of image segmentation. The approach provides steganographic robustness to the visual and statistical analysis, and hardness of message content, due to the use in the structure of the steganographic system of block cipher algorithm.

1. *Алиев А.Т., Аграновский А.В.* Вопросы построения криптостеганографических систем. Модель стеганографического канала передачи данных // Информационное противодействие угрозам терроризма. — 2006. — № 8. — С. 79–91.
2. *Зубов А.Ю.* Совершенные шифры. — М. : Гелиос АРВ, 2003. — 160 с.
3. *RIJNDAEL* description. Submission to NIST by J. Daemen, V. Rijmen. — (http://csrc.nist.gov/CryptoToolkit/aes/).
4. *Patent* 9395/5 USA / E. Adelson. Digital signal encoding and decoding apparatus. — 1990.
5. *Кустов В.Н., Федчук А.А.* Методы встраивания скрытых сообщений // Защита информации. Конфидент. — 2002. — № 3. — С. 34–37.
6. *Girod B.* The information theoretical significance of spatial and temporal masking in video signals // Proc. of the SPIE Human Vision, Visual Processing, and Digital Display. — 1989. — **1077**. — P. 178–187.
7. *Алиев А.Т.* О применении стеганографического метода LSB к графическим файлам с большими областями монотонной заливки // Вестн. ДГТУ. — 2004. — **4**, № 4 (22). — С. 454–460.
8. *Чиркова С.В., Бородин Г.А.* Об эффективном методе выбора контейнеров для стеганографии // Электронный журнал «Вычислительные сети. Теория и практика». — М., 2007. — № 2 (11):13.5. — http://network-journal.mpei.ac.ru/.
9. *Чиркова С.В., Бородин Г.А.* Метод выбора контейнера для стеганографии // Там же. — 2009. — № 2 (15):11.2. — http:// network-journal.mpei.ac.ru/.
10. *Грибунин В.Г., Оков И.Н., Туринцев И.В.* Цифровая стеганография. — М. : СОЛОН-Пресс, 2002. — 261 с.
11. *Швидченко И.В.* Анализ криптостеганографических алгоритмов // Проблемы управления и информатики. — 2007. — № 4. — С. 149–155.
12. *Павлидис Т.* Алгоритмы машинной графики и обработки изображений: Пер. с англ. — М. : Радио и связь, 1986. — 400 с.

Получено 05.05.2010