АЛГОРИТМ ОБНАРУЖЕНИЯ ВНЕДРЕННОЙ В JPEG-ИЗОБРАЖЕНИЯ ИНФОРМАЦИИ

Б. В. Юрьев, А. А. Яковенко

Одесский национальный политехнический университет
Украина, Одесса
nitro.ti161@gmail.com, iakovenko.oleksandr@gmail.com

Разработан алгоритм, с помощью которого можно с высокой вероятностью определить, является ли исследуемое изображение результатом работы алгоритма добавления скрытых данных JSteg. Вывод о факте добавления скрытых данных делается на основе анализа статистики значений компонентов дискретного косинусного преобразования. Представлены результаты работы алгоритма, свидетельствующие о возможности его применения при решении задач стеганографии.

Ключевые слова: стеганография, дискретное косинусное преобразование, сокрытие информации.

Защита и добыча информации никогда не стоит на месте, и с появлением новых способов обработки данных как злоумышленники, так и стражи порядка ищут тайные способы передачи секретных сообщений с помощью современных технологий.

Хотя используемый сейчас формат графических изображений Joint Photographic Experts Group, или JPEG, в ходу уже несколько десятков лет, его эффективность не падает, и поэтому он был выбран для данного исследования. Стоит отметить, что формат кодировки видео Moving Picture Experts Group, или MPEG, основан на тех же технических принципах, поэтому результаты данной работы также применимы к данному формату.

Целью настоящей работы было создание алгоритма, способного с высокой вероятностью определить, является ли исследуемое изображение результатом работы алгоритма добавления скрытых данных JSteg. Вывод о факте добавления скрытых данных делается на основе анализа статистики значений компонентов дискретного косинусного преобразования (ДКП).

Представленная задача является по своей сути стенографической, поскольку при внедрении информации в изображение она уже зашифрована (чаще всего шифром AES), что делает изъятие данных практически невозможным. Передача же данных внутри графического файла без видимого изменения его структуры алгоритмами сокрытия данных в изображениях (такими как JSteg) позволит скрыть сам факт передачи, поскольку данные алгоритмы самые точные из доступных инструментов проверки [1].

После исследования структуры сжатого JPEG-изображения было установлено, что внедрение информации на этапе преобразования файла из цветового пространства RGB в пространство YCbCr и разбиения полученных данных на блоки 8×8 неэффективно. При разбиении значений YCbCr на блоки и дальнейшем их квантовании с помощью ДКП все изображение претерпевает небольшое искажение вследствие особенностей ДКП, что означает потерю внедренной информации. Дальнейшая упаковка квантованных значений ДКП с участием кодов Хаффмана к такому результату не приводит. Следовательно, внедрение информации в квантованные коэффициенты дискретного преобразования (ККДП) стоит ожидать на данном этапе сжатия. Для иллюстрации высокой вероятности этого была исследована эффективность изменения значений ККДП. Было обнаружено, что при изменении значения малозначимого бита коэффициента (или МЗБ), чаще всего хранящего незначительные цветовые изменения графического файла [2], конечный результат мало чем отличается от изменений при естественном сжатии с помощью JPEG.

Рассмотрим основные шаги алгоритма выявления внедрения данных JSteg.

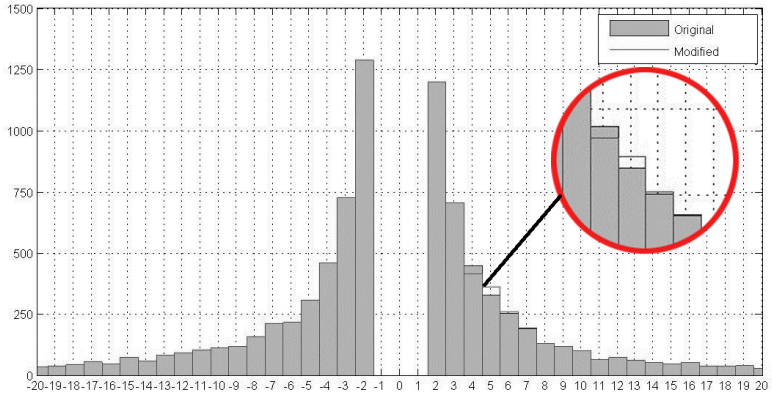
1. *Измерение статистики распределения значений компонентов ДКП в исследуемом изображении.* Этот шаг включает анализ всех jpeg-блоков изображения: для каждого блока производится переход в цветовое пространство YCbCr, ДКП, квантование и регистрация значений, которые принимают ДКП-компоненты. Результатом замера становится гистограмма распределения значений компонентов ККДП H, значение каждого элемента h_i — это вероятность встретить ККДП со значени-

ем *i*. Такая гистограмма несет в себе достаточно информации, чтобы определить факт внедрения JSteg. Поскольку изменение значений ККДП, равных 0 и ±1, не производится, эти компоненты исключатся из гистограммы.

2. *Формирование коэффициента равномерности распределения значений компонентов* k_E проводится таким образом, чтобы значения для изображений с естественной статистикой отличались от значений изображений, статистика которых была изменена путем добавления скрытой информации JSteg. Изменение статистики заключается в том, что при замене младших значащих битов гистограмма H имеет другой вид: значения, которые отличаются только МЗБ, имеет примерно одинаковую вероятность (см. рисунок). Следовательно, коэффициент k_E должен реагировать на разность значений соседних элементов гистограммы:

$$k_E = \sum_i (h_i - h_{i+1}), \quad i = 2,\ 4,\ 6\ ...$$

Коэффициент k_E будет иметь более высокие значения для изображений без внедрения. Чтобы определить порог значений, характерных для изображения без внедрения, нужно собрать статистику таких значений для множества изображений с внедрением и без него. Базируясь на этой статистике, можно подобрать значение порога коэффициента равномерности гистограммы t_E, который позволит выявить факт внедрения с максимальной достоверностью.



Сравнение гистограмм H для изображений с внедрением и без него

На рисунке представлены результаты работы алгоритма обнаружения внедренной информации за счет разности порогов коэффициентов k_E, которые свидетельствует о его работоспособности.

Таким образом, предложенный алгоритм может использоваться для обнаружения скрытых данных в JPEG-изображении и решения задач цифровой безопасности и защиты информации.

ИСПОЛЬЗОВАННЫЕ ИСТОЧНИКИ

1. Конахович Г. Ф., Пузыренко А.Ю. Компьютерная стеганография. — Киев: МК-Пресс, 2006.

2. Chin-Chen Chang, Tung-Shou Chen, Lou-Zo Chung. A steganographic method based upon JPEG and quantization table. // Information Sciences. — 2002. — №141 — P. 123–138.

3. Siwei Lyu. Natural image statistics for digital image forensics / A Thesis Submited to the Faculty in partial fulfillment of the requirements for the degree of Doct. of Philos. in Computer Science, Dartmouth College, 2005

4. Калашніков М., Яковенко О., Кушніренко Н., Чечельницький В. Розробка стеганографічного алгоритму із урахуванням статистики зображення-контейнера // Матеріали IV Міжнародної науково-технічної конференції «Захист інформації і безпека інформаційних систем».— Україна, Львів.— 2015.— С. 129—131.

B. V. Yuriev, A. A. Iakovenko

The algorithm for intrusive information detection in JPEG-images

An algorithm has been developed that can be used to determine with high probability whether the tested image is the result of the JSteg algorithm for adding hidden data. The conclusion about whether hidden data was added is made on the basis of a statistics analysis of the values of the components of the discrete cosine transform. The results of the operation of the algorithm testify to the possibility of its application in solving steganography problems.

Keywords: steganography, discrete cosine transformation, information concealment.