УДК 004.932.2 DOI: 10.26102/2310-6018/2020.29.2.007

Метод обнаружения скрытой передачи данных, использующий стеганографический метод Коха-Жао

Д.Э. Вильховский
Омский государственный университет им. Ф.М. Достоевского
Омск, Россия

Резюме: В статье предлагается алгоритм анализа изображений с встроенным сообщением на основе стеганографического метода Коха-Жао. Объектом исследования являются изображение, в которые было осуществлено встраивание методом Коха-Жао. Ключевая идея заключается в сравнительном анализе пар коэффициентов дискретного косинусного преобразования. Для этого строится зависимость разности коэффициентов от номера блока. Численное дифференцирование данной зависимости даёт возможности определить границы встроенного сообщения. После чего анализ исходной зависимости на выделенном интервале даёт возможность установить применяемые параметры метода Коха-Жао. Научная новизна заключается в разработке алгоритма стеганографического анализа метода Коха-Жао, основанного на анализе коэффициентов дискретного косинусного преобразования, отличающегося тем, что в нем присутствуют модуль автоматического поиска ступенчатых изменений, который позволяет определить параметры встраивания и извлечь сообщение. Выполнен эксперимент на ЭВМ. Установлено, что предлагаемый алгоритм даёт возможность с абсолютной точностью установить размер, содержимое и расположение скрытого сообщения, в случае, когда оно встроено в непрерывную последовательность блоков. Практическая значимость результатов заключается в том, что разработанный программный комплекс позволяет проводить стегоанализ изображений с данными методом Коха-Жао при низком заполнении стегоконтейнера (менее чем 40% битов нулевого битового слоя).

Ключевые слова: анализ коэффициентов ДКП, выявление стеговставок, анализ стегоконтейнера, анализ изображений со вставками, метод Коха-Жао

Для цитирования: Вильховский Д.Э. Метод обнаружения скрытой передачи данных, использующий стеганографический метод Коха-Жао. *Моделирование, оптимизация и информационные технологии.* 2020;8(2). Доступно: https://moit.vivt.ru/wp-content/uploads/2020/05/Vilkhovskiy_2_20_1.pdf DOI: 10.26102/2310-6018/2020.29.2.007

Method of detection of hidden data using steganography method Koch-Zhao

D.E. Vilkhovskiy
Omsk State University F.M. Dostoevsky
Omsk, Russia

Abstract: The article proposes an image analysis algorithm with a built-in message based on the Koch-Zhao steganographic method. The object of research is the image into which the Koch-Zhao embedding was carried out. The key idea is a comparative analysis of pairs of coefficients of a discrete cosine transform. For this, the dependence of the difference of the coefficients on the block number is constructed. Numerical differentiation of this dependence makes it possible to determine the boundaries of the embedded message. After that, the analysis of the initial dependence on the selected interval makes it possible to establish the applicable parameters of the Koch-Zhao method. Scientific novelty lies in the development of the steganographic analysis algorithm of the Koch-Zhao method, based on the analysis of discrete cosine transform coefficients, characterized in that it contains an automatic step search module that allows you to determine the embedding parameters and extract the message. Computer

experiment completed. It is established that the proposed algorithm makes it possible to establish with absolute accuracy the size, content and location of the hidden message, in the case when it is embedded in a continuous sequence of blocks. The practical significance of the results lies in the fact that the developed software package allows stego analysis of images with data using the Koch-Zhao method with low filling of the stegocontainer (less than 40% of bits of the zero bit layer).

Keywords: analysis of DCT coefficients, identification of stego inserts, analysis of a stegocontainer, analysis of images with inserts, Koch-Zhao method.

For citation: Vilkhovskiy D.E. Method of detection of hidden data using steganography method Koch-Zhao *Modeling, Optimization and Information Technology*. 2020;8(2). Available from: https://moit.vivt.ru/wp-content/uploads/2020/05/Vilkhovskiy_2_20_1.pdf DOI: 10.26102/2310-6018/2020.29.2.007(In Russ).

Введение

В общем случае при атаке на стеганографические алгоритмы существует 3 задачи. Первая задача состоит в том, что нужно установить факт встраивания скрытого сообщения. Вторая задача состоит в том, что, нужно установить расположение и размер встраиваемого сообщения в массиве данных изображения-контейнера. Третья задача заключается в том, что нужно максимально точно восстановить встроенное сообщение.

Стеганографический анализ добился максимальных успехов при решении первой задачи. Большая часть методов выявления наличия встроенного сообщения базируется на изучении статистических свойств изображения, применяемого в качестве контейнера. В работах [1,2] опираясь на предположения о случайном распределении битов младших битов в синей компоненте цветного изображения предлагается статистический метод стеганографического анализа, который основан на критерии Хи-квадрат. Указанный метод даёт возможность с высокой вероятностью выявить при равномерном заполнении контейнера факт встраивания сообщения. В работе [3] предлагается метод визуального сопоставления цветовых слоев. Данный подход проявляет свою эффективность, если в изображении-контейнере присутствует достаточно большие области равномерной заливки. В случае большого количества мелких деталей, искажения, которые вносят встраиваемым изображением, визуально никак не заметить. В работах [4-6] применяется моделирование изображения с встроенным сообщением на основе цепей Маркова. Такой подход будет эффективным, когда встроенное сообщение можно охарактеризовать определёнными статистическими характеристиками. В работах [7,8] изучено воздействие стеганографической вставки на степень сжатия изображения-контейнера. Утверждается, что изменения статистических характеристик по причине стеганографического встраивания сообщения уменьшает сжимаемость изображения. При этом уменьшение степени сжатия будет зависеть от размера сообщения.

Цель данной работы заключается в создании алгоритма определения стеганографических вставок в изображении, встраиваемых при помощи метода Коха-Жао [9].

Алгоритм встраивания и постановка задачи

В роли объекта исследования рассмотрим цифровое изображение, о котором нет информации об отсутствии или наличии встроенного сообщения. Известно лишь то, что применён метод встраивания Коха-Жао [11].

Первая задача исследования: нужно установить факт отсутствия или наличия стеганографического встраивания. Вторая задача: если встроенное сообщение присутствует, нужно установить его размеры и положение в изображении-контейнере.

Третья задача: нужно с максимальной точностью выявить встроенное сообщение, если оно есть, при отсутствии какой-либо априорной информации.

Основой для стеганографического метода Коха-Жао [11] является двухмерное дискретное косинусное преобразование (ДКП). Алгоритм встраивания сообщения можно описать так:

1. Первоначальное изображение разделяется на блоки размером 8x8 пикселей.
2. Для каждого блока используется ДКП, результат представлен в виде матрицы коэффициентов D_i $(i = 1, \dots, N;\ N$ – число блоков) размером 8x8.
3. Формируется последовательность блоков, в которых будет выполняться встраивание. По 1 биту информации записывается в каждый блок.
4. В каждом блоке выбираются 2 коэффициента ДКП, которые расположены в среднечастотной области коэффициентов, симметричные относительно главной диагонали ($D_i[3,4]$ и $D_i[4,3]$, $D_i[3,5]$ и $D_i[5,3]$, $D_i[4,5]$ и $D_i[5,4]$).
5. Для передачи бита 0 нужно, чтобы разница модулей пары коэффициентов ДКП превышала определённую положительную величину M_0, для передачи бита 1 разница обязана быть меньше M_0. Т.е., передавая 0 нужно уменьшить модуль второго коэффициента и увеличить модуль первого. Передавая 1 нужно увеличить модуль второго коэффициента и уменьшить модуль первого.
6. Проходим по каждому блоку и исполняем пункты 4 и 5.
7. Для каждого блока осуществляем обратное ДКП.

Выбор среднечастотных коэффициентов ДКП объясняется тем, что нужно минимизировать воздействие встраивания на визуальные свойства измененного изображения. Выбор низкочастотных или высокочастотных коэффициентов приведёт к тому, что появятся эффекты, которые можно будет заметить визуально.

При извлечении сообщения принято считать, что пары изменяемых коэффициентов ДКП известны. Алгоритм извлечения совпадает с алгоритмом встраивания в первых 4-х пунктах:

5. Определяем разность значений модулей для пар коэффициентов, в которые осуществлялось встраивание.
6. Когда разность значений меньше, чем – M_0, то встроен был единичный бит. Когда разность больше M_0, то был встроен бит 0.
7. Последовательно извлекаются биты, встроенные во все блоки.

Анализ алгоритмов извлечения и встраивания указывает на то, что для выполнения успешной атаки на стеганографический метод Коха-Жао нужно установить блоки, в которые встраивались сообщения, пороговое значение M_0 и индексы изменяемых коэффициентов ДКП.

Для того чтобы корректно извлечь сообщения принимающая и отправляющая стороны обязаны иметь общую секретную информацию о параметрах встраивания. Будем опираться на то, что информация о параметрах имеет наименьший размер. В таком случае можно сформулировать 3 предположения. Первое предположение состоит в том, что, встраивание осуществляется в постоянную последовательность блоков. Второе предположение состоит в том, что для всех блоков применяются одни и те же пары коэффициентов ДКП. Третье предположение заключается в том, что, для всех блоков применяется одинаковое значение M_0. Любые отклонения от указанных предположений повышают объем секретной информации.

Алгоритм стеганографического анализа

Для установления встроенного сообщения примем тот факт, что параметр M_0 обязан иметь большое значение, дающее возможность принимающей стороне производить извлечение скрытого сообщения из любого изображения без потерь. Если M_0 будет выбрано недостаточно большим, то в извлекаемом сообщении могут появиться ошибки, обусловленные спецификой изображения-контейнера.

В первую очередь, нужно определить коэффициенты ДКП, в которые выполнялось встраивание. Как и в алгоритме встраивания, для этого, разделим изображение на блоки $B_i(i = 1, \dots, N)$ размером 8x8 пикселей. Применим к каждому блоку B_i $(i = 1, \dots, N)$ ДКП. Как результат получим совокупность матриц коэффициентов $D_i(i = 1, \dots, N)$ размером 8x8. Проведём анализ среднечастотных элементов матриц $D_i(i = 1, \dots, N)$.

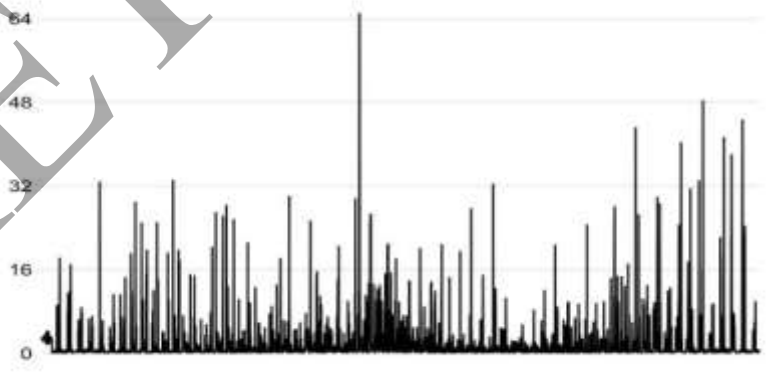
Определим 3 последовательности величин $(i = 1, \dots, N)$:

$$C_i^{(1)} = \left| |D_i[3,4]| - |D_i[4,3]| \right| \quad (1)$$

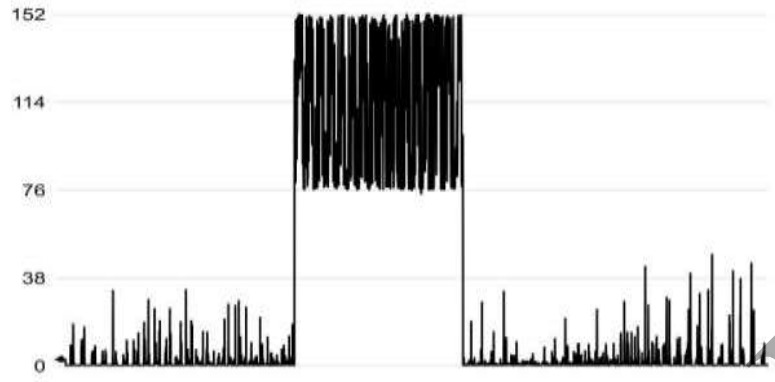
$$C_i^{(2)} = \left| |D_i[3,5]| - |D_i[5,3]| \right| \quad (2)$$

$$C_i^{(3)} = \left| |D_i[4,5]| - |D_i[5,4]| \right| \quad (3)$$

При встраивании сообщения изменяется одна из указанных последовательностей. Построим гистограммы зависимости $C_i^{(j)}$ $(j = 1,2,3;\ i = 1, \dots, N)$ от номера блока i. Встраивание сообщения вызывает изменение одной из последовательностей в виде появления «ступени» высотой M_0. На Рисунке 1 приведён пример подобного изменения.



а)



б)

Рисунок 1 – Гистограмма зависимости $C_i^{(1)}$ от номера блока i: а) изображение без встроенного сообщения, б) изображение с встроенным сообщением.

Figure 1 – The histogram depending $C_i^{(1)}$ on the block number i: a) image without inline message, б) image with inline message.

Проблему выявления встроенного сообщения можно свести к анализу зависимостей $C_i^{(j)}(j = 1,2,3;\ i = 1, \dots, N)$ от номера блока i и поиску ступенчатых изменений. Границы ступеней на гистограмме можно определить при помощи численного дифференцирования на основе разностных схем.

Проведём численное дифференцирование зависимости $C_i^{(j)}\ (j = 1,2,3;\ i = 1, \dots, N)$ по i.

$$dC_i^{(j)} = C_i^{(j)} - C_{i-1}^{(j)} \quad (4)$$

Результатом данной операции ступенчатые изменения выдадут высокие пики, позволяющие установить границы встроенного сообщения. На Рисунке 2 изображена зависимость $dC_i^{(j)}$ от номера блока i для зависимости на Рисунке 1-б.

Чтобы для каждого массива $dC_i^{(j)}$ автоматически установить границы встроенного сообщения определим величины: O_j – среднеквадратичное отклонение для элементов массива $dC^{(j)}$, N_j – среднее значение элементов массива $dC^{(j)}$, M_j – максимальное значение элементов массива $dC^{(j)}$.

Установим величины $R_j = N_j + O_j$. Введём величину Y_j, изменяющаяся в диапазоне значений от R_j до M_j. Подберём значение Y_j так, чтобы было только 2 значения $C_{i_1}^{(j)} > Y_j$ и $C_{i_2}^{(j)} > Y_j$. Значения i_1 и i_2 будут являться границами встроенного сообщения. Для установления значения M_0 нужно найти наименьшее значение $C_i^{(j)}$ на диапазоне от i_1 до i_2.

Алгоритм будет выглядеть так:

1. Разделение изображения на блоки B_i размером 8x8 пикселей.
2. Применим ДКП к каждому блоку B_i. Как результат получатся матрицы коэффициентов ДКП D_i размером 8x8.

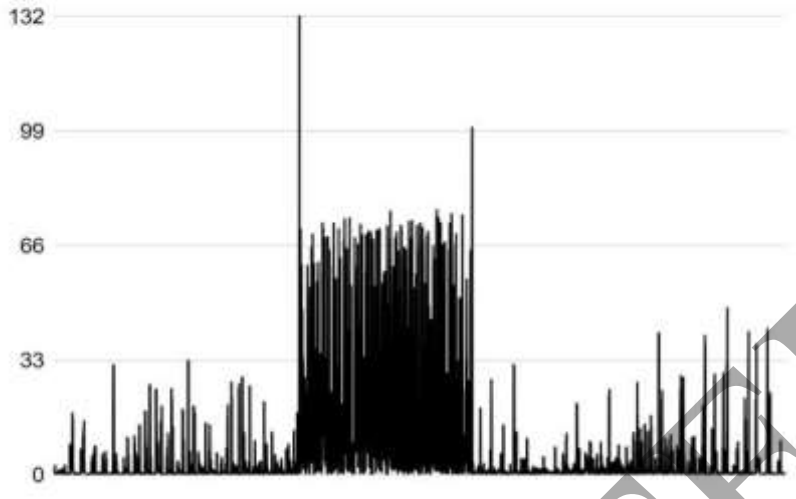


Рисунок 2 – Гистограмма зависимости $dC_i^{(1)}$ от номера блока i.
Figure 2 – The histogram depending $dC_i^{(1)}$ on the block number i.

3. Построим 3 последовательности величин $(i = 1, \dots, N)$:

$$C_i^{(1)} = \left| |D_i[3,4]| - |D_i[4,3]| \right| \quad (5)$$

$$C_i^{(2)} = \left| |D_i[3,5]| - |D_i[5,3]| \right| \quad (6)$$

$$C_i^{(3)} = \left| |D_i[4,5]| - |D_i[5,4]| \right| \quad (7)$$

4. Выполним численное дифференцирование $C_i^{(j)} (j = 1,2,3; i = 1, \dots, N)$ по i:

$$dC_i^{(j)} = C_i^{(j)} - C_{i-1}^{(j)} \quad (8)$$

5. Определим: O_j – среднеквадратичное отклонение для элементов массива $dC^{(j)}$, N_j – среднее значение элементов массива $dC^{(j)}$, M_j – наибольшее значение элементов массива $dC^{(j)}$. Определить величины $R_j = N_j + O_j$.
6. Выполним перебор величины Y_j с шагом dY в диапазоне от R_j до M_j. Установим значение Y_j такое, что есть лишь 2 значения $C_{i_1}^{(j)} > Y_j$ и $C_{i_2}^{(j)} > Y_j$. Если такое значения установить невозможно, то нужно сократить шаг dY. Определим i_1 и i_2.
7. Поиск минимального значения $C_i^{(j)}$ на диапазоне от i_1 до i_2. Присвоим M_0 найденное значение.
8. Извлечём сообщение, применяя найденные параметры.

Компьютерный эксперимент и результаты

Для того, чтобы определить эффективность функционирования представленного в данной статье метода, протестируем его на библиотеке изображений BSDS500. Данная подборка создавалась с целью проверки методов кластеризации и на данный момент содержит 500 изображений в формате JPEG. Выбранная коллекция содержит изображения с различными типами областей заливки и различным содержимым. Более того, формат JPEG также, как и метод Коха-Жао, основывается на ДКП, что исключает дополнительные ошибки при конвертации изображения в другой формат.

Тестирование производилось как с пустым стегоконтейнером (оригинальное изображение), так и с заполненным. Необходимо отметить, что предложенный алгоритм показал свою эффективность при $M_0 > 54$. При более низких величинах M_0 появляются сложности с извлечением встроенных данных. Если алгоритм обнаруживает встроенное сообщение, то в таком случае он может его однозначно извлечь.

Рисунок 3 – Примеры изображений, для которых наблюдаются ложно-позитивные результаты.
Figure 3 – Examples of images for which false positives are observed.

В рамках компьютерного эксперимента на основе обработки 500 изображений коллекции BSDS500 были определены следующие параметры:

TP - процент истинно-положительных результатов, при которых заполненный стегоконтейнер был верно определен и встроенное сообщение было извлечено корректно.

FN – процент ложно-негативных результатов, при которых заполненный стегоконтейнер был верно определен, но встроенное сообщение не удалось извлечь.

TN – процент истинно-негативных результатов, для которых пустой контейнер был верно определен.

FP – процент ложно-позитивных результатов, при которых пустой контейнер был некорректно определен как стегоконтейнер.

Рисунок 4 – Примеры изображений, для которых наблюдаются ложно-негативные результаты.
Figure 4 – Examples of images for which false negative results are observed.

В ходе тестирования предложенного алгоритма были получены следующие результаты:

$$TP = 85{,}5\%,\ FN = 14{,}5\%,\ TN = 77\%,\ FP = 23\%.$$

Как видно, предложенный алгоритм достаточно точно определяет наличие встроенного сообщения. Ошибки при работе данного алгоритма связаны со структурными особенностями изображения. В изображении с пустым стегоконтейнером возможно наличие пиков в последовательности коэффициентов ДКП, которые могут ошибочно приниматься за границы встроенного сообщения. Наличие двух подобных границ приводит к ложно-позитивным результатам. Если в изображении присутствует более двух ложных границ, то это приводит к невозможности определить истинные границы встроенного сообщения.

С точки зрения защиты информации интерес представляют изображения, для которых процесс стегоанализа не позволяет определить факт встраивания данных

.

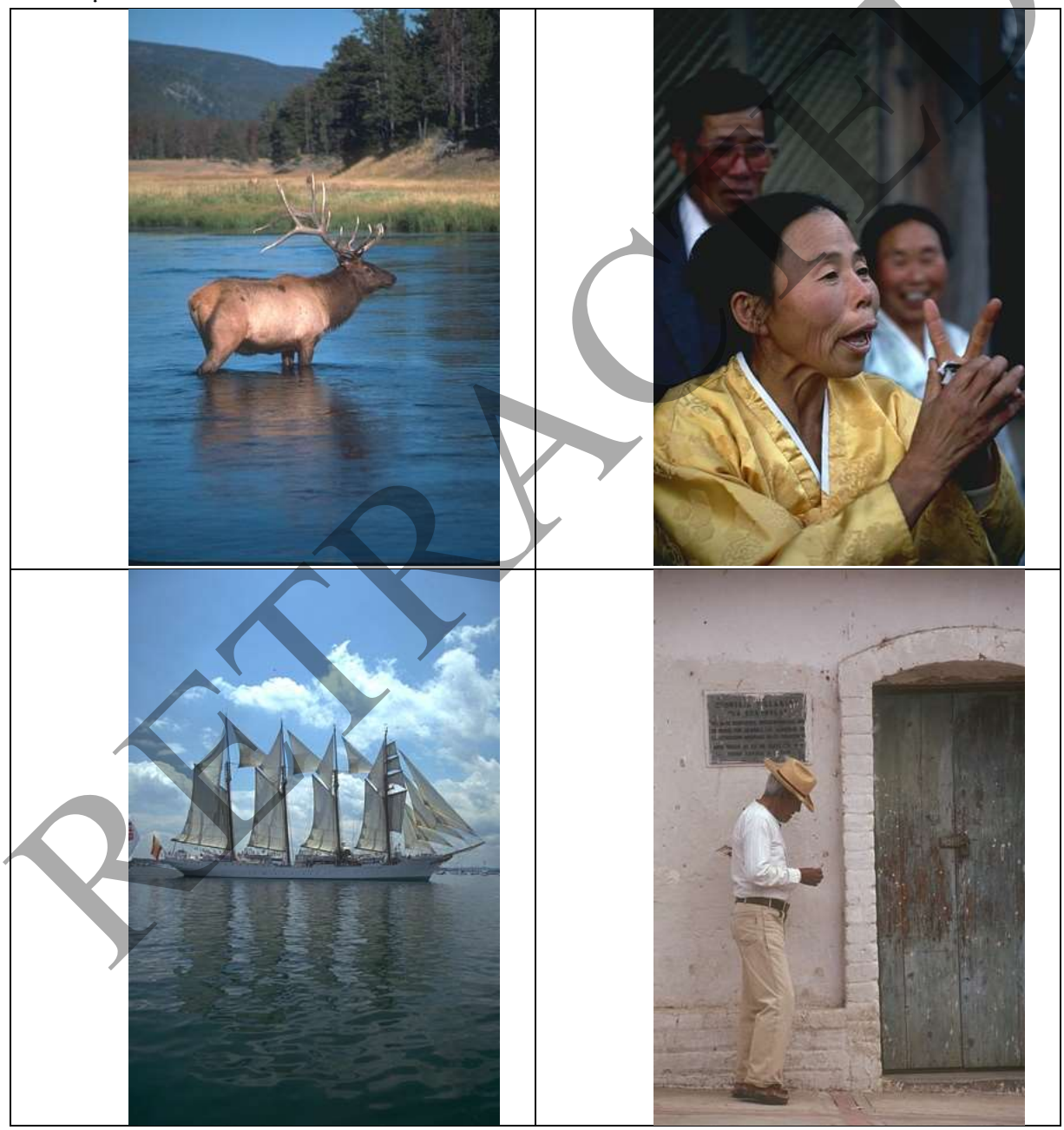


Рисунок 5 – Примеры изображений, для которых наблюдаются истинно положительные и истинно негативные результаты.

Figure 5 – Examples of images for which true positive and true negative results are observed.

На Рисунке 3 представлены примеры изображений, для которых наблюдаются ложно-позитивные результаты. То есть они позволяют ввести в заблуждение злоумышленника и заставить его анализировать пустой стегоконтейнер.

На Рисунке 4 приведены примеры изображений, для которых наблюдаются ложно-негативные результаты. Эти изображения позволяют скрыто передавать встроенные сообщения.

На Рисунке 5 приведены примеры изображений, для которых наблюдаются истинно положительные и истинно негативные результаты. Эти изображения не рекомендуется использовать для передачи скрытых сообщений, так как они легко поддаются стегоанализу.

Обсуждение результатов и выводы

В статье было показано, что стеганографический алгоритм Коха-Жао не является устойчивым к атаке анализа коэффициентов ДКП. Предлагаемый алгоритм на основе анализа коэффициентов дискретного косинусного преобразования даёт возможность абсолютно точно извлекать встроенное сообщение при условии, что оно будет единственным и встроено в непрерывную область. Отступление от указанных предположений увеличивает стойкость стеганографического алгоритма. Тестирование на коллекции изображений показало, что ошибки ложного определения наличия стегановставки в пустом стегоконтейнере составляют 23%. Эффективность обнаружения наличия встроенного сообщения составляет 85,5%.

ЛИТЕРАТУРА

1. Provos N., Honeyman P. *Detecting steganographic content on the internet*. Technical Report CITI 01-1a, University of Michigan. 2001.
2. Westfeld A., Pfitzmann A. *Attacks on Steganographic Systems: Breaking the Steganographic Utilities EzStego, Jsteg, Steganos and Stools and Some Lessons Learned*. 3rd International Workshop on Information Hiding. 2000:61–76.
3. Алиев А.Т. *О применении стеганографического метода LSB к графическим файлам с большими областями монотонной заливки*. Вестник ДГТУ. 2004;4(22):454–460.
4. Барсуков В.С. Романцов А.П. *Оценка уровня скрытности мультимедийных стеганографических каналов хранения и передачи информации.* Специальная Техника. 2000:1.
5. Кустов В.Н., Параскевопуло А.Ю. *Простые тайны стегоанализа*. Защита информации, INSIDE. 2005;4:72–78.
6. Голуб В.А., Дрюченко М.А. *Комплексный подход для выявления стеганографического скрытия в JPEG-файлах*. Инфокоммуникационные технологии. 2009;7(1):44–50.
7. Жилкин М.Ю. *Стегоанализ графических данных в различных форматах*. Доклады ТУСУРа. 2008;2(18):63–64.
8. Монарев В. А. *Сдвиговый метод обнаружения скрытой информации*. Вестник СибГУТИ. 2012;4:62–68.
9. Koch E. *Towards robust and hidden image copyright labeling*. IEEE Workshop on Nonlinear Signal and Image Processing. 1995:452–455.

REFERENCES

1. Provos N., Honeyman P. *Detecting steganographic content on the internet*. Technical Report CITI 01-1a, University of Michigan. 2001.
2. Westfeld A., Pfitzmann A. *Attacks on Steganographic Systems: Breaking the Steganographic Utilities EzStego, Jsteg, Steganos and Stools and Some Lessons Learned*. 3rd International Workshop on Information Hiding. 2000:61–76.
3. Aliev A.T. *On the application of the LSB steganographic method to graphic files with large areas of monotonous fill*. Vestnik DGTU. 2004;4 (22):454–460.
4. Barsukov V.S. Romancov A.P. *Assessment of the stealth level of multimedia steganographic channels for storing and transmitting information*. Special'naja Tehnika. 2000:1.
5. Kustov V.N., Paraskevopulo A.Ju. *Simple secrets of steganalysis*. Zashhita informacii, INSIDE. 2005;4:72–78.
6. Golub V.A., Drjuchenko M.A. *Comprehensive approach for revealing steganographic concealment in JPEG files.* Infokommunikacionnye tehnologii. 2009;7(1):44–50.
7. Zhilkin M.Ju. *Stegoanalysis of graphic data in various formats*. Doklady TUSURa. 2008;2(18):63–64.
8. Monarev V. A. *Shift detection of hidden information*. Vestnik SibGUTI. 2012;4:62–68.
9. Koch E. *Towards robust and hidden image copyright labeling*. IEEE Workshop on Nonlinear Signal and Image Processing. 1995:452–455.

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Вильховский Данил Эдуардович, ассистент кафедры информационной безопасности ФКН в Омском государственном университете им. Ф.М. Достоевского (ОмГУ), Омск, Российская Федерация
e-mail: vilkhovskiy@gmail.com

Danil E. Vilkhovskiy, Assistant at the Department of Information Security of the FCS of Omsk State University, Omsk, Russian Federation