УДК 004.056.53

К. В. Защелкин, канд. техн. наук,
Е. Н. Иванова, А. А. Ищенко

ПОВЫШЕНИЕ СТОЙКОСТИ СТЕГАНОГРАФИЧЕСКИЙ СИСТЕМЫ К JPEG-АТАКАМ ЗА СЧЕТ УСТРАНЕНИЯ ЯВЛЕНИЯ ПЕРЕРОЖДЕНИЯ БЛОКОВ СТЕГО-КОНТЕЙНЕРА

Аннотация. *Рассмотрена задача стеганографического скрытия данных. Показана проблема реализации стеганографических методов, выполняющих скрытие данных в частотную область растрового изображения. Такие методы обычно позиционируется как стойкие к JPEG-сжатию изображения-контейнера. Показано, что в ряде случаев JPEG-сжатие приводит к перерождению пригодных блоков в непригодные и наоборот. Это делает невозможным извлечение данных из изображения. В работе предлагается подход к решению данной проблемы. Описана программная реализация и экспериментальное исследование предлагаемого подхода.*

Ключевые слова: *стеганография, защита информации, скрытие данных, секретная передача данных, JPEG-сжатие, дискретно-косинусное преобразование, внедрение данных, графический стего-контейнер*

K. V. Zashcholkin, PhD.,
E. N. Ivanova, A. A. Ishchenko

PERSISTENCE INCREASE OF STEGANOGRAPHIC SYSTEM AGAINST JPEG-ATTACKS BY ELIMINATING PHENOMENA OF METAMORPHOSIS OF BLOCKS OF STEGO-CONTAINERS

Abstract. *The problem of information steganography hiding was reviewed. The problem steganographic methods implementation, which performs steganography data hiding in bitmap image frequency domain was shown. Such methods usually are positioned to be persistant to JPEG-image compression container. It is shown that in some cases JPEG-compression results in the transformation of usable into unusable units and vice versa. This makes data extraction from the image impossible. An approach to solving this problem is proposed in this paper. The software realization and experimental study of the proposed approach was described.*

Keywords: *steganography, data protection, data hiding, secret data transfer, JPEG-compression, discrete cosine transform, information embedding, graphics stego-container*

К. В. Защолкін, канд. техн. наук,
О. М. Іванова, А. О. Іщенко

ПІДВИЩЕННЯ СТІЙКОСТІ СТЕГАНОГРАФІЧНОЇ СИСТЕМИ ДО JPEG-АТАК ЗА РАХУНОК УСУНЕННЯ ЯВИЩА ПЕРЕРОДЖЕННЯ БЛОКІВ СТЕГО-КОНТЕЙНЕРА

Анотація. *Розглянуто задачу стеганографічного приховування даних. Показано проблему реалізації стеганографічних методів, які виконують приховування даних в частотну область растрового зображення . Такі методи зазвичай позиціонується як стійкі до JPEG-стиску зображення-контейнера. Показано, що в ряді випадків JPEG-стиснення призводить до переродження придатних блоків в непридатні і навпаки. Це робить неможливим вилучення даних з зображення. В роботі пропонується підхід до вирішення даної проблеми. Описана програмна реалізація та експериментальне дослідження запропонованого підходу.*

Ключові слова: *стеганографія, захист інформації, приховування даних, секретна передача даних, JPEG-стиск, дискретно-косинусне перетворення, вбудовування даних, графічний стего-контейнер*

Введение. Цифровая стеганография является одним из эффективных направлений защиты информации в компьютерных системах. В ее основе лежит принцип скрытия факта существования защищаемой информации [1]. Это принципиально отличает стенографические методы от криптографического подхода. Стеганографические методы дают возможность встраивать дополнительную скрытую информацию в стегоконтейнеры, не нарушая их информационной целостности.

Подавляющее большинство современных стеганографических методов основано на встраивании секретной информации в мультимедийные контейнеры, информация в которых изначально имеет аналоговую природу: графические, звуковые, видеофайлы [2].

Центральной проблемой теории цифровой стеганографии является проблема противодействие атакам на стеганографическую систему. Одним из существенных видов атак выступает активная атака, которая заключается во внесении атакующей стороной в стего-контейнер искажений, не нарушающих восприятия контейнера, однако разрушаю-

щих находящуюся в нем встроенную информацию. Применительно к графическим стего-контейнерам такое искажение может состоять в небольшом изменении яркости или контрастности изображения; незначительном повороте изображения; применении методов сжатия изображения с потерями. Наибольший интерес приставляют именно методы противодействия активным стего-атакам, основанным на сжатии с потерями [3, 4]. Это обусловлено чрезвычайной распространенностью использования сжатия изображений с потерями в современных информационных технологиях и сложностью противодействия атакам, базирующимся на таком сжатии [5]. Далее в данной работе рассматриваются элементы противодействия активным стего-атакам, основанным на применении наиболее используемого на текущий момент метода сжатия с потерями – JPEG [6].

Анализ проблемы. Основной группой стего-методов, ориентированных на противодействие JPEG-атакам, являются методы, оперирующие особенностями JPEG-сжатия: метод Коха-Жао [7], метод Бенгама-Мемона-Эо-Юнг [8] и производные от них. Эти методы используют для внедрения стего-информации частотную область изображения и учитывают такие основные стадии JPEG сжатия, как разбиение изображения на блоки, дискретно-косинусное преобразование (ДКП), квантование результатов ДКП [9, 10]. В рамках таких методов изображение разбивается на блоки размером 8x8 пикселей, для каждого из которых выполняется процедура ДКП, подобно тому, как это делается в методе JPEG. В один ДКП-блок изображения внедряется один бит стего-информации. Для этого осуществляется изменение значений блока, приводящее к выполнению неравенства (конкретная форма которого обусловлена методом) между несколькими заданными ДКП-коэффициентами в блоке. Расположение этих коэффициентов в блоке обычно устанавливает ортогональную зависимость между стойкостью стего-информации к JPEG-искажениям и степенью визуальной деградации изображения в результате внедрения этой информации в контейнер. Для уменьшения такой зависимости блоки подвергают классификации на пригодные и непригодные для встраивания в них стего-информации. Пригодными считаются блоки, для которых данная зависимость проявляется минимально в силу особенностей области изображения, описываемой ими. Так например, часто используемый на практике стего-метод Бенгама-Мемона-Эо-Юнг [8] (далее, метод БМЭЮ) и методы, производные от него, численно определяют, что пригодными считаются блоки, одновременно удовлетворяющие двум требованиям в пространственной области изображения:

1) пригодные блоки не должны иметь резких перепадов яркости;

2) пригодные блоки не должны быть слишком монотонными.

Метод БМЭЮ рекомендует выполнять анализ указанных требований посредством исследования значений блоков в пространстве ДКП (рис. 1).

Рис. 1. Стандартное разбиение блока изображения в пространстве ДКП

Это исследование состоит в следующем.

1) Блоки, не отвечающие первому требованию, характеризуются наличием больших значений низкочастотных коэффициентов ДКП. Для численного разграничения низкочастотных коэффициентов ДКП данного блока с большими и малыми значениями вводится порог P_L. Сумма низкочастотных ДКП коэффициентов Σ_L (суммирование производится по всем низкочастотным ДКП коэффициентам блока за исключением DC-

коэффициента, расположенного в левом верхнем углу блока) сравнивается с порогом P_L, в результате чего принимается решение о том, имеет ли блок резкие перепады яркости.

2) Для блоков, не отвечающих второму требованию, характерно равенство нулю большинства высокочастотных коэффициентов ДКП. Принятие решение по этому критерию производится при помощи порога P_H, с которым сравнивается сумма Σ_H высокочастотных ДКП коэффициентов блока.

Блок считается пригодным для встраивания в случае выполнения следующего составного условия

$$(\Sigma_L < P_L) \,\&\, (\Sigma_H > P_H) \ . \qquad (1)$$

Пороги P_L и P_H устанавливаются на стороне внедрения секретной информации в контейнер и являются частью стего-ключа, необходимого для извлечения информации.

Метод БМЭЮ предполагает, что на стороне извлечения информации из контейнера, производится подсчет значений Σ_L и Σ_H и выполняется аналогичная классификация блоков на такие, в которых может содержаться встроенная стего-информация, и такие, в которых такая информация содержаться не может.

Постановка цели работы. Проведенное исследование практической реализации метода БМЭЮ [8] и методов, основанных на нем, позволило выявить проблему видоизменения блоков в результате JPEG-атаки на стего-контейнер. Обнаружены случаи, при которых атака данного вида, даже при малой степени JPEG-сжатия, переводит блоки, имеющие встроенную стего-информацию (и соответственно классифицированные как пригодные), в класс непригодных для встраивания. Имеет место и обратное явление, при котором блоки, классифицированные на этапе внедрения информации как непригодные, после сжатия могут быть отнесены к множеству пригодных и содержащих внедренную стего-информацию. Такие явления приводят к нарушению целостности информации, извлекаемой из стего-контейнера, и в конечном итоге снижают стойкость стего-системы к JPEG-атакам.

Цель данной работы состоит в повышении стойкости к JPEG-атакам стего-систем, выполняющих встраивание информации в частотную область изображения при помощи БМЭЮ-подобных методов. Поставленная цель достигается за счет модификации указанных методов, приводящей к устранению явления перерождения блоков изображения-контейнера в результате JPEG-сжатия.

Модификация метода БМЭЮ. Предлагается модификация метода БМЭЮ, устраняющая проблему перерождения блоков изображения и повышающая стойкость метода к JPEG-атакам. Модификация состоит в выполнении апостериорной классификации блоков путем применения JPEG-сжатия на этапе встраивания стего-информации в контейнер. Рассмотрим предлагаемую процедуру апостериорной классификацией блоков (рис. 2).

Исходные данные процедуры (рис. 2, блок 2):

– множество блоков B, на которые разбито исходное изображение-контейнер, представленных в пространстве ДКП;

– внедряемая в изображение-контейнер стего-информации, представленная двоичной последовательностью M;

– p – вектор параметров JPEG-сжатия контейнера [6], определяющий степень выраженности JPEG-атаки на заполненный стего-контейнер;

– пороги P_L и P_H, численно определяющие допустимую суммарную величину низкочастотной и высокочастотной составляющей ДКП блока.

Для определенности будем считать, что количество блоков изображения и длина внедряемой в изображение двоичной последовательности совпадают. В общем случае длина последовательности не должна превышать количество блоков.

Рассматриваемые далее действия применяются ко всем блокам изображения-контейнера и всем разрядам внедряемой в изображение последовательности (рис. 2, цикл – блок 2 и блок 14).

Шаг 1 (рис. 2, блок 4): независимо от того, относится ли текущий обрабатываемый блок B_i изображения к классу блоков, пригодных для встраивания или нет, выполняется внедрение в него в соответствии с мето-

дом БМЭЮ очередного разряда m_i двоичной последовательности M. Процедура внедрения обозначена на рис. 2 как $BMYY(\,)$. В результате такого внедрения получается модифицированный блок B_i^*.

Шаг 2 (рис. 2, блок 5): выполняется сохранение исходного блока B_i в JPEG-формат с применением параметров сжатия p. В результате получается сжатый блок $B_{i\,JPEG}$.

Шаг 3 (рис. 2, блок 6): выполняется сохранение блока B_i^*, содержащего встроенную информацию, в JPEG-формат с применением тех же параметров сжатия, что и на предыдущем шаге. В результате получается блок $B_{i\,JPEG}^*$.

Шаг 4 (рис. 2, блок 7): выполняется подсчет суммы низкочастотных и высокочастотных ДКП коэффициентов для сжатого исходного блока $B_{i\,JPEG}$.

Шаг 5 (рис. 2, блок 8): выполняется подсчет суммы низкочастотных и высокочастотных ДКП коэффициентов для сжатого блока $B_{i\,JPEG}^*$, содержащего встроенную стего-информацию.

Шаг 6 (рис. 2, блок 9): осуществляется проверка выполнения условия (1) для блока $B_{i\,JPEG}^*$. Функция проверки выполнения условия (1), обозначенная $\delta(\,)$, возвращает значение “истина” и “ложь”, выражающее выполнение или невыполнение условия. На рис. 2, значения, возвращаемые функцией $\delta(\,)$, для наглядности, показаны как “истина” – “пригоден” и “ложь” – “непригоден”, что выражает пригодность или непригодность блока для встраивания в него информации.

Если условие (рис. 2, блок 9) выполняется, то это означает, что блок $B_{i\,JPEG}^*$, содержащий встроенную стего-информацию, апостериорно признан пригодным и на стороне извлечения будет предпринята попытка извлечь информацию из этого блока.

В этом случае блок $B_{i\,JPEG}^*$ помещается в выходной стего-контейнер (рис. 2, блок 13), содержащий результат применения метода. После этого осуществляется переход к обработке следующего блока и следующего разряда встраиваемой стего-информации.

Если условие (рис. 2, блок 9) не выполняется, то осуществляется переход к следующему шагу.

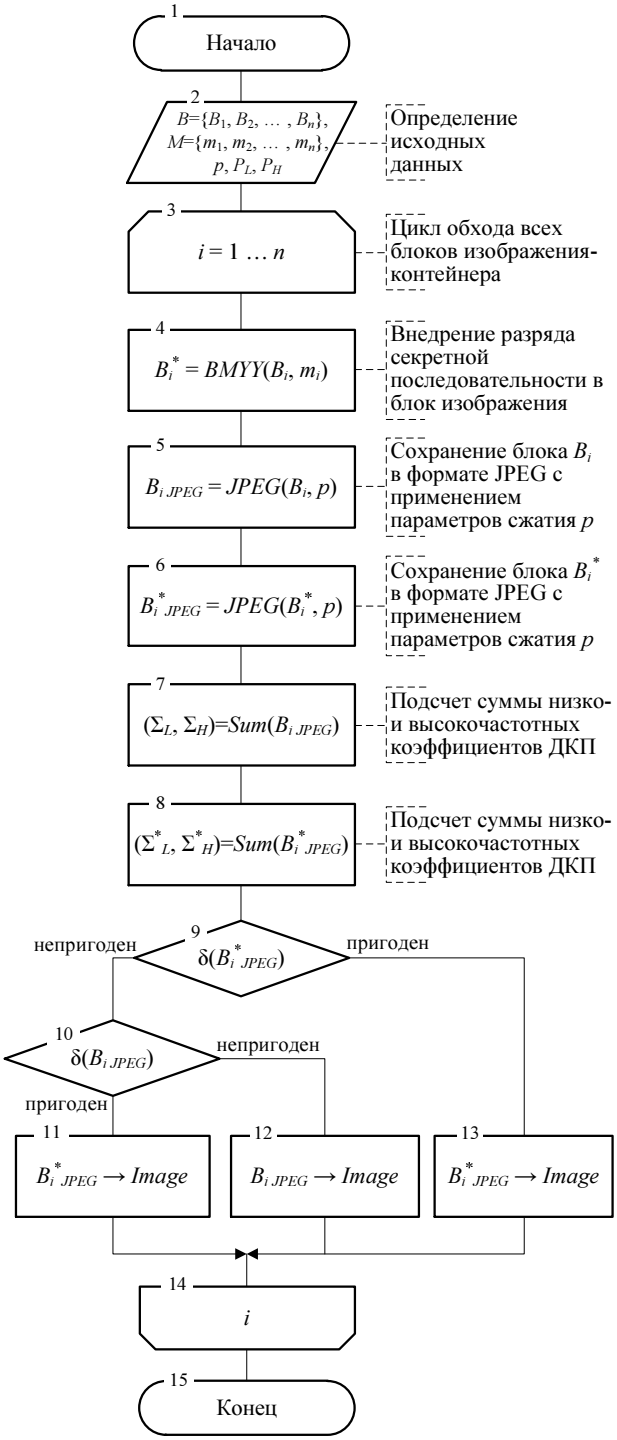

Рис. 2. Блок-схема предлагаемой апостериорной процедуры классификации блоков при встраивании

Шаг 7 (рис. 2, блок 10): осуществляется проверка выполнения условия (1) для блока $B_{i\,JPEG}$. Если условие не выполняется, то это означает, что блок апостериорно признан непригодным и на стороне извлечения не будет предприниматься попытка извлечь информацию из данного блока. Это является правильным решением, так как внедренной стего-информации в блоке $B_{i\,JPEG}$ нет. В этом случае блок $B_{i\,JPEG}$ помещается в выходной стего-контейнер и осуществляется переход к обработке следующего блока и следующего разряда стего-информации.

Если условие (рис. 2, блок 10) выполняется, то это означает, что на стороне извлечения будет предпринята попытка извлечь информацию из блока $B_{i\,JPEG}$ при ее отсутствии в нем. Следовательно, в выходное изображение необходимо поместить блок, который на стороне извлечения будет признан не пригодным и не содержащим встроенной информации. В этом случае в выходное изображение помещается блок $B^{*}_{i\,JPEG}$ (рис. 2, блок 11), поскольку для текущей ветки блок-схемы на предыдущем шаге (рис. 2, блок 9) данный блок был признан непригодным для встраивания. Этот блок визуально не отличается от блока $B_{i\,JPEG}$, но на стороне извлечения не будет предприниматься попытка извлечь из него информацию.

После этого осуществляется переход к обработке следующего блока и следующего разряда встраиваемой стего-информации.

Предложенная процедура классификации блоков предполагает выполнение JPEG-сжатия контейнера на этапе внедрения информации. При этом фактически имитируется JPEG-атака на заполненный контейнер и принимается апостериорное решение о классификации блока.

Программная реализация и экспериментальное исследование предложенной модификации метода. Для экспериментального исследования предложенного усовершенствованного метода БМЭЮ было разработано программное обеспечение, выполняющее встраивание данных по классическому методу БМЭЮ и встраивание по методу БМЭЮ с учетом предложенной модификации.

Исходные данные эксперимента:

1) множество $Im=\{im_1, im_1, \ldots, im_{100}\}$, состоящее из 100 растровых изображений, различающихся:

– природой их происхождения (фотоснимки и синтетические изображения);

– размером;

– различными долями областей сплошной заливки и областей, содержащих мелкие контрастные детали;

2) множество $T=\{T_1, T_2, \ldots, T_{50}\}$, состоящее из 50 текстовых сообщений длиной от 50 до 250 символов.

Методика проведения эксперимента:

1) из множества *Im* случайным образом выбиралось изображение $im_k \in Im$;

2) из множества *T* случайным образом выбиралось текстовое сообщение $t_q \in T$;

3) сообщение t_q встраивалось в изображение im_k в соответствии с классическим вариантом метода БМЭЮ;

4) случайным образом выбирался вектор параметров JPEG-сжатия p_j, дающий степень качества сжатого изображения в диапазоне от 75 до 100 по стобалльной шкале (данный диапазон качества обычно задействован при выполнении стеганографических JPEG-атак).

5) сообщение t_q встраивалось в изображение im_k в соответствии с предложенной модификацией метода БМЭЮ с применением параметра p_j для апостериорной классификации;

6) выполнялось JPEG-сжатие (стего-атака) изображений, полученных в пунктах 3 и 5. При этом применялись параметры сжатия, дающие равную или большую степень качества изображения по отношению к параметрам p_j;

7) из сжатых изображений, полученных в предыдущем пункте, в соответствии с методом БМЭЮ, извлекалась встроенная информация;

8) результаты эксперимента оценивались на основе наличия ошибок извлечения, вызванных неправильной классификацией пригодных блоков как непригодных и непригодных как пригодных.

Оценка стойкости стего-системы к атакам, выраженная степенью нарушения целостности

встроенной информации, численно оценивалась при помощи коэффициента ошибочных разрядов *BER* (Bit Error Rate) [5]:

$$BER(M, M^*) = \left(\sum_{i=1}^{N} p_i \right) / N , \qquad (2)$$

где M – двоичная последовательность, выражающая оригинальное сообщение, встраиваемое в стего-контейнер;

M^* – двоичная последовательность, извлеченная из стего-контейнера;

N – общее количество разрядов, встроенных в стего-контейнер;

$$p_i = \begin{cases} 1, \ if \ \ M_i \neq M_i^*; \\ 0, \ if \ \ M_i = M_i^*; \end{cases}$$

M_i и M_i^* – i-й разряд последовательности M и M^* соответственно.

Серия из 50 экспериментов, проведенных по указанной методике, показала, что:

– средний коэффициент *BER* для традиционного метода БМЭЮ на диапазоне качества сжатого изображения от 75 до 100 составил около 0,044, а для предложенной модификации – около 0,01;

– предложенная модификация метода БМЭЮ позволяет правильно извлекать внедренные сообщения из фрагментов изображений-контейнеров, на которых традиционный метод БМЭЮ давал ошибку извлечения по причине неправильной классификации блоков, вызванной JPEG-атакой;

– стойкость модифицированного метода к активным стенографическим атакам, не связанным с JPEG-сжатием (незначительным размытием и поворотом изображения-контейнера) осталась неизменной по сравнению с исходным вариантом метода.

Предлагаемая модификация метода увеличивает его вычислительную сложность на этапе внедрения информации, так как требует дополнительных процедур JPEG-сжатия для каждого из блоков. Однако этот недостаток компенсируется устранением проблемы неверной классификации блоков на этапе извлечения информации из изображения-контейнера, подвергшегося JPEG-атаке.

Заключение. В данной работе предложен подход к повышению стойкости стего-системы к JPEG-атакам. В частности, выполнено усовершенствование процедуры классификации блоков, характерной для БМЭЮ-подобных методов, осуществляющих внедрение данных в частотную область растрового изображения. Усовершенствование состоит в реализации апостериорной классификации блоков изображения путем применения JPEG сжатия на этапе встраивания стего-информации в контейнер.

За счет введения предложенных модификаций были устранены ошибки извлечения, вызванные преобразованием блоков в результате JPEG-атаки на изображение-контейнер. А именно, были устранены ошибки, вызванные неправильной классификацией блоков как непригодных для встраивания при их реальной пригодности и пригодных для встраивания при их реальной непригодности.

Стойкость усовершенствованного метода к активным стенографическим атакам не связанным с JPEG-сжатием, при этом осталась на уровне классического метода БМЭЮ.

Список использованной литературы

1. Конахович Г. Ф. Компьютерная стеганография [Текст] / Г. Ф. Конахович, А. Ю. Пузыренко. – К. : МК-Пресс, 2006. – 288 с.

2. Fridrich J. Steganography in Digital Media [Text] / J. Fridrich. – New York : Cambridge University Press, 2010. – 448 p.

3. Грибунин В.Г. Цифровая стеганография [Текст] / В.Г. Грибунин. – М. : Салон-пресс, 2002. – 344 с.

4. Аграновский А.В. Стеганография, цифровые водяные знаки и стегоанализ [Текст] / А. В. Аграновский, А. В. Балакин, В. Г. Грибунин. – М. : Вузовская книга, 2009. – 220 с.

5. Михайличенко О. В. Резистивность цифровых водяных знаков к JPEG-преобразованию [Текст] / О. В. Михайличенко, Н. Н. Прихожев, А. Г. Коробейников // Научно-технический вестник информационных технологий, механики и оптики. – Санкт-Петербург, 2007. – № 40. – С. 248 – 251.

6. Гонсалес Р. Цифровая обработка изображений. Издание 3-е. [Текст] / Р. Гонсалес, Р. Вудс. – М. : Техносфера, 2012. – 1104 с.

7. Zhao J., and. Koch E. Embedding Robust Labels into Images for Copyright Protection [Text], *Proceedings of the International Congress on Intellectual Property Rights for Specialized In-*

formation, Knowledge and New Techniques, (1995), German, Munich, pp. 242 – 251.

8. Benham D., Memon N., Yeo B., and Yeung M. Fast, Watermarking of DCT-based Compressed Images [Text], *Proceedings of the International Conference on Image Science, Systems and Technology,* (1997), USA, Las Vegas, pp. 243 – 252.

9. Shih F. Watermarking, Steganography, and Forensics [Text], (2012), New York, *CRC Press*, – 424 p.

10. Cox I., Miller M., Bloom J., and Fridrich J.. Digital Watermarking and Steganography [Text], (2008), Burlington, *Morgan Kaufmann Publishers*, 592 p.

Получено 28.04.2014

References

1. Konahovich G.F., and Puzirenko A.U. Kompyuternaya steganografiya [Computer Steganography], (2006), Kiev, Ukraine, *MK-Press Publ.*, 288 p. (In Russian).

2. Fridrich J. Steganography in Digital Media, (2010), New York, *Cambridge University Press*, 448 p. (In English).

3. Gribunyn V.G. Tsifrovaya steganografiya [Digital Steganography], (2002), Moscow, Russian Federation, *Salon-Press Publ.*, 344 p. (In Russian).

4. Agranovsky A.V., Balkin A.V., and Gribunyn V.G. Steganografiya, tsifrovye vodyanye znaki i stegoanaliz [Steganography, Digital Watermarks and Stegoanalysis], (2009), Moscow, Russian Federation: *University Book Publ.*, 220 p. (In Russian).

5. Mikhaylichenko O.V., Prikhozhev N.N., and Korobeinikov A.G. Rezistivnost tsifrovykh vodyanykh znakov k JPEG-preobrazovaniyu [Digital Watermarks Resistance to JPEG-transformation], (2007), *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, St. Petersburg, Russian Federation, No 40, pp. 248 – 251 (In Russian).

6. Gonzalez R., and Woods R. Tsifrovaya obrabotka izobrazheniy. Izdanie 3-e. [Digital Image Processing, 3rd Edition], (2012), Moscow, Russian Federation, *Technosphere Publ.*, 1104 p. (In Russian).

7. Zhao J., and Koch E. Embedding Robust Labels into Images for Copyright Protection, (1995), *Proceedings of the International Congress on Intellectual Property Rights for Specialized Information, Knowledge and New Techniques*, German, Munich, pp. 242 – 251 (In English).

8. Benham D., Memon N, Yeo B, and Yeung M. Fast Watermarking of DCT-based Compressed Images, (1997), *Proceedings of the International Conference on Image Science, Systems and Technology*, USA, Las Vegas, pp. 243 – 252. (In English).

9. Shih F. Watermarking, Steganography, and Forensics, (2012), New York, *CRC Press Publ.*, 424 p. (In English).

10. Cox I., Miller M., Bloom J., and Fridrich J. Digital Watermarking and Steganography, (2008), Burlington, *Morgan Kaufmann Publishers*, 592 p. (In English).

Защелкин Константин Вячеславович, канд. техн. наук, доц. каф. компьютерных интеллектуальных систем и сетей Одесского нац. политехн. ун-та,
тел.: (048) 734-83-22.
E-mail: const-z@te.net.ua

Иванова Елена Николаевна, ст. преподаватель каф. компьютерных систем Одесского нац. политехнич. ун-та,
тел.: (048) 734-83-91.
E-mail: enivanova@ukr.net

Ищенко Артем Александрович, студент каф. компьютерных интеллектуальных систем и сетей Одесского нац. политехн. ун-та,
тел.: (048) 734-83-22.
E-mail: rtemkaxxx@gmail.com