УДК 004.056.55:004.932.2

# ЗАЩИТА ОТ НЕПРАВОМЕРНОГО ИСПОЛЬЗОВАНИЯ ГРАФИЧЕСКОЙ ИНФОРМАЦИИ В СОЦИАЛЬНЫХ СЕТЯХ

**Земцов А.Н., Цыбанов В.Ю.**

*ФГБОУ ВО «Волгоградский государственный технический университет», Волгоград, e-mail: ecmsys@yandex.ru*

Рассматриваются вопросы разработки методики защиты от неправомерного использования графической информации в социальных сетях с помощью стеганографических методов, встраивающих и скрывающих цифровые водяные знаки в частотную область вейвлет-преобразования графических изображений в сочетании с сингулярным разложением. Социальные сети являются мощным инструментом межличностной коммуникации, охватывают огромную аудиторию и стали неотъемлемой частью жизни современного общества. Проблема неправомерного использования медиаконтента является одной из наиболее актуальных. Традиционные методы не решают проблему защиты от пиратства, так как после покупки в зашифрованном виде и с использованием электронной цифровой подписи изображения некоторым юридическим или физическим лицом, ничто не мешает ему использовать его в своих корыстных целях. Успешно решены задачи, связанные с установлением характеристик цифровых водяных знаков, разработкой методики защиты, позволяющей однозначно идентифицировать владельца защищаемого медиаконтента, а также процедурой верификации цифровых водяных знаков. На основе оценки меры пикового отношения уровня сигнала к уровню шума проведен анализ результатов экспериментов в части вносимых методом маркирования искажений в эталонные изображения. Встраивание информации производится в низкочастотной области вейвлет-спектра. Приводится обоснование эффективности предлагаемой методики маркирования изображений. Предлагаемая методика допускает совместное использование с современными методами обработки изображений, в том числе методами сжатия на основе дискретного вейвлет-преобразования.

**Ключевые слова: социальные сети, средства массовой информации, стеганография, скрытие данных, кратномасштабный анализ, вейвлет-преобразование, цифровой водяной знак, защита информации**

# PROTECTION AGAINST ILLEGAL USE OF IMAGES ON SOCIAL NETWORKS

**Zemtsov A.N., Tsybanov V.Yu.**

*Volgograd State Technical University, Volgograd, e-mail: ecmsys@yandex.ru*

The issues of developing a methodology for protecting against the unlawful use of graphic information in social networks using steganographic methods that embed and hide digital watermarks in the frequency domain of graphic images in combination with a singular decomposition are considered. Social networks are a powerful tool for interpersonal communication, encompass a huge audience, and have become an integral part of the life of modern society. The problem of misuse of media content is one of the most relevant. Traditional methods do not solve the problem of protection against piracy, because after buying in encrypted form and using an electronic digital signature of the image by some legal or natural person, nothing prevents him from using it for his own selfish purposes. Successfully solved the problems associated with establishing the characteristics of digital watermarks, developing a protection methodology that allows you to uniquely identify the owner of the protected media content, as well as the verification procedure for digital watermarks. Based on the assessment of the measure of the peak ratio of the signal level to the noise level, the analysis of the experimental results in part of the distortions introduced by the method of marking into the reference images is carried out. Information is embedded in the low-frequency region of the wavelet spectrum. The substantiation of the effectiveness of the proposed method for marking images is given. The proposed technique allows sharing with modern image processing methods, including compression methods and based on a discrete wavelet transform.

**Keywords: social networks, media, steganography, data hiding, multi-resolution analysis, wavelet transform, digital watermark, information security**

В последние годы наблюдается значительный рост заинтересованности процессом межличностной коммуникации. Как следствие, социальные сети, являясь мощным инструментом межличностной коммуникации, стали неотъемлемой частью жизни современного общества. Позволяя достигать наибольшей эффективности общения, социальные сети охватывают аудиторию, превышающую население отдельных государств. В подобных условиях интенсивного роста обмена информацией особо остро стоит проблема правомерности заимствования медиаконтента как результата интеллектуальной деятельности. Решение данной проблемы может осуществляться различными способами. Одним из наиболее эффективных является использование стеганографических методов [1].

Цифровой водяной знак – это данные, встроенные в мультимедийный объект таким образом, что впоследствии водяной знак может быть обнаружен и извлечен с целью подтверждения наличия права на результат интеллектуальной деятельности. Информационный медиаконтент, ко-

торый подвергается процедуре защиты стегоалгоритмами, обычно называют оригиналом или произведением. Цифровой водяной знак представляет собой информацию, которая встраивается в защищаемый стегоалгоритмами оригинал.

Цифровые водяные знаки и методы маркирования ими можно классифицировать по нескольким категориям с учетом различных особенностей. К одной из классификационных характеристик исследователи относят область внедрения, в которую производится встраивание водяного знака. Например, цифровой водяной знак может быть встроен в пространственную область стегоконтейнера. Альтернативная возможность предусматривает маркирование водяными знаками в частотной области, что позволяет достичь устойчивости к внешним по отношению к стегоконтейнеру воздействиям [2].

Маркирование цифровыми водяными знаками можно разделить на следующие три категории по виду маркированного медиаконтента в формате аудио, видео и изображения. По способу восприятия наблюдателем цифровые водяные знаки принято разделять на четыре категории: видимые, невидимые устойчивые к воздействию, невидимые неустойчивые и двойственные.

Наблюдатель может фиксировать видимый водяной знак при осмотре изображения, вследствие чего такой подход имеет ограниченное применение. Невидимые устойчивые водяные знаки встраиваются таким образом, что изменения, внесенные в пиксель или группу пикселей изображения, не могут быть замечены непосредственным восприятием с помощью зрения, а также характеризоваться устойчивостью к различным внешним воздействиям, например преднамеренным атакам, стандартным операциям обработки сигналов, и удовлетворять требованию восстановления только с помощью заданного механизма декодирования. Невидимые, неустойчивые к внешнему воздействию цифровые водяные знаки встраиваются так, что любое изменение маркированного изображения приводит к разрушению или искажению водяного знака. Двойственные цифровые водяные знаки имеют признаки обеих категорий: видимых и невидимых, причем невидимый используется в качестве резервной копии видимого.

Цель исследования заключается в разработке методики защиты от неправомерного использования графической информации в социальных сетях с помощью стеганографических методов, встраивающих и скрывающих цифровые водяные знаки в частотную область графических изображений.

*Предлагаемая методика защиты графической информации*

Для достижения поставленной цели необходимо решить ряд задач, связанных с установлением характеристик цифровых водяных знаков, разработать методику и алгоритмы, позволяющие однозначно идентифицировать владельца медиаконтента, а также процедуру верификации цифровых водяных знаков.

Цифровой водяной знак может характеризоваться рядом существенных показателей:

1. Лояльность определяет возможность алгоритма маркирования встраивать цифровой водяной знак так, чтобы вносимые алгоритмом искажения не оказывали значимого влияния на качество исходного изображения. Если наблюдатель не может различить исходное и защищенное изображения, процедура защиты изображения считается незаметной. Тем не менее малозначительное искажение маркированного изображения может стать визуально заметным при сравнении с исходным изображением. Изменения в исходном изображении представляется невозможным заметить, если у пользователя отсутствует возможность сравнения с исходным изображением.

2. Полезная нагрузка – количество информации, которая может быть записана в стегоконтейнер.

3. Робастность обеспечивает устойчивость водяного знака к изменению и возможному удалению из стегоконтейнера. Цифровой водяной знак должен быть невосприимчив к стандартным непреднамеренным и преднамеренным воздействиям, т.е. должен характеризоваться устойчивостью к различным распространенным методам обработки сигналов, таким как сжатие, квантование и др., общим геометрическим преобразованиям, например масштабированию, вращению и т.д. В [3] предлагается каскадная стегосистема многобитовых голографических цифровых водяных знаков, которая позволяет обеспечить робастность цифрового водяного знака к широкому спектру преднамеренных и непреднамеренных искажений.

4. Достоверность – водяной знак должен однозначно идентифицировать владельца.

Решение доверенного арбитра, основанное на протоколе верификации цифрового водяного знака, должно указывать, что автор результата интеллектуальной деятельности является единственным владельцем изображения, а иные лица не могут правомерно использовать полученный автором результат интеллектуальной деятельности

без его согласия. Тем не менее существуют некоторые проблемы с протоколом верификации разрешения права собственности, поскольку широко распространено такое явление, как пиратство. Приведем обобщенную классификацию проблем, связанных с верификацией права собственности результата интеллектуальной деятельности:

1) тупик владения, когда право собственности не может быть установлено до тех пор, пока пират обладает возможностью представить доказательство права собственности, которое является в той же степени убедительным, что и представленное владельцем подтверждение права на результат интеллектуальной деятельности;

2) контрафактное владение, когда пират представляет доказательство права собственности, которое является даже более убедительным, чем фактическое доказательство владельца права на результат интеллектуальной деятельности;

3) кража собственности, когда пират получает стегоконтейнер, в том числе легальным путем, и встраивает в него новый цифровой водяной знак.

С позиции доказательности общие требования к цифровым водяным знакам можно сформулировать следующим образом:

1) цифровой водяной знак должен характеризоваться незаметностью при малой вносимой ошибке;

2) размер полезной нагрузки должен позволять осуществить полную идентификацию владельца права собственности. К полезной нагрузке условно можно отнести и ключ цифрового водяного знака, который может быть предоставлен группе правообладателей либо использоваться как приватный;

3) в связи с тем, что предполагается использование цифрового водяного знака в юридических процессах, то целесообразно потребовать высокого значения вероятности его верификации. Верификация цифрового водяного знака невозможна без требования устойчивости его обнаружения;

4) устойчивость к непреднамеренным воздействиям, таким как преобразование из одного формата в другой, сжатие изображения и т.п., а также возможность совместного использования с несколькими водяными знаками, как правило, используется до трех других встроенных водяных знаков;

5) устойчивость к злонамеренным воздействиям (атакам), направленным на уничтожение цифрового водяного знака, атакам синхронизации, а также атакам замещения, при которых злоумышленник пытается заменить оригинальный цифровой водяной знак контрафактным;

6) устойчивость к криптографическим атакам, защита от атаки сговором.

На рис. 1 представлена предлагаемая схема маркирования изображения, где $I$ – исходное изображение, $W$ – цифровой водяной знак, $I_W$ – маркированное изображение, и $K$ – ключ. Функция встраивания $E_{mb}$ осуществляет маркирование изображения $I$ цифровым водяным знаком $W$ с использованием ключа $K$, в результате которого функцией $E_{mb}$ генерируется модифицированное изображение $I_W$. Использование ключа $K$ позволяет повысить защищенность стегосистемы. Предварительно исходное изображение $I$ преобразуется либо в пространственную область, либо маркирование может осуществляться в частотной области, для чего выполняется некоторое прямое спектральное преобразование. Для восстановления изображения $I$ по набору спектральных коэффициентов выполняется обратное преобразование.

На практике стегоконтейнер, содержащий водяной знак, может быть намеренно или случайно искажен. В обоих случаях стегосистема должна обеспечивать возможность обнаружения и извлечения цифрового водяного знака после атаки.

Использование цифровых водяных знаков для защиты права на результат интеллектуальной деятельности требует создания эффективной процедуры верификации извлекаемых цифровых водяных знаков [2].

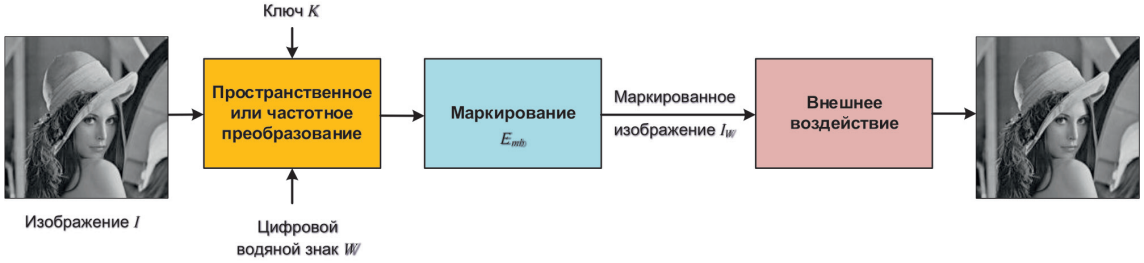



*Рис. 1. Схема маркирования изображения*

Протокол верификации цифровых водяных знаков представляет собой последовательность этапов, в которых принимает участие одно или несколько лиц, претендующих на право владения медиаконтентом. Протокол верификации цифровых водяных знаков предназначен для получения достоверного суждения относительно правомерности заявленных требований на результат интеллектуальной деятельности. Предлагаемую концепцию проектирования процедуры верификации извлекаемых цифровых водяных знаков для подтверждения права собственности представим в виде следующих основных этапов:

1) автор создает изображение, которое является исходным для процедуры защиты изображением и называется оригинальным изображением или стегоконтейнером;

2) автор помечает цифровыми водяными знаками оригинальное изображение и получает маркированное изображение с встроенным в него цифровым водяным знаком. Стегоконтейнер хранит доказательства права на результат интеллектуальной деятельности, которые могут быть использованы при последующем оспаривании этого права;

3) автор публикует в социальной сети защищенное фотографическое произведение;

4) автор обнаруживает, что фотографическое произведение, правом собственности на которое он обладает, используется лицом, которое не внесло никакого вклада в его создание;

5) чтобы реализовать право владения фотографическим произведением, автор должен представить значимые доказательства доверенному арбитру, в качестве которого может выступить суд;

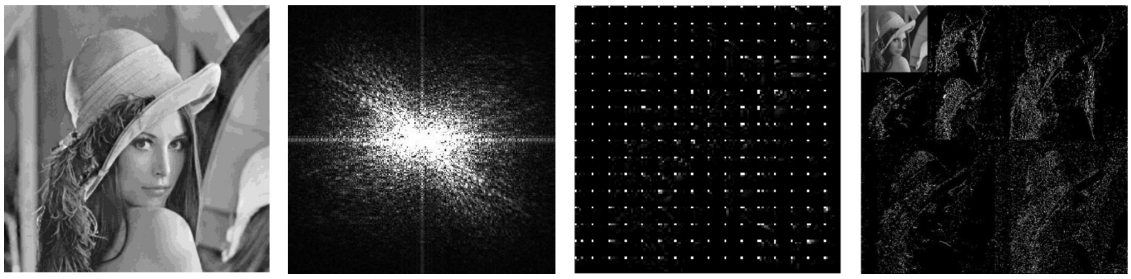
6) оценивая доказательства раздельно и их взаимосвязь в совокупности, применяя методику проведения экспертизы путем верификации цифровых водяных знаков, доверенный арбитр принимает решение относительно того, кто может использовать соответствующий результат интеллектуальной деятельности.

В большинстве отечественных социальных сетей используются изображения, основанные на стандартах Международной электротехнической комиссии и Международного консультационного комитета по телефонии и телеграфии, таких как JPEG [4], JPEG-LS [5] и JPEG 2000 [6]. Таким образом, методика аппроксимации изображений с помощью спектрального преобразования является основным этапом построения набора алгоритмов анализа и постобработки изображений. Концепция вейвлет-преобразования была впервые разработана И. Мейером [7] и получила дальнейшее развитие в работах С. Малла [8, 9].

Достоинством пространственных методов является простота использования, но такие методы обеспечивают степень скрытности и робастность хуже, чем спектральные, являющиеся более стойкими к искажениям [11]. На рис. 2 показаны некоторые наиболее популярные частотные преобразования, реализованные в разрабатываемом проекте. Предлагаемая методика защиты от неправомерного использования графической информации в социальных сетях с помощью стеганографических методов основана на многоуровневом дискретном вейвлет-преобразовании [2] в сочетании с сингулярным разложением.

*Результаты экспериментов*

Для контроля искажений, вносимых методом защиты изображения, представляется целесообразным использовать весовой коэффициент силы встраивания α. Для анализа границ применимости предлагаемого метода защиты изображения необходимо выполнить эксперименты по встраиванию цифрового водяного знака, а также анализ устойчивости метода к различным преднамеренным искажениям посредством последующей верификации.



*а)* *б)* *в)* *г)*

*Рис. 2. Частотное преобразование стегоконтейнера: а) исходный стегоконтейнер; б) дискретное преобразование Фурье; в) дискретно-косинусное преобразование в алгоритме JPEG [10]; г) дискретное вейвлет-преобразование*

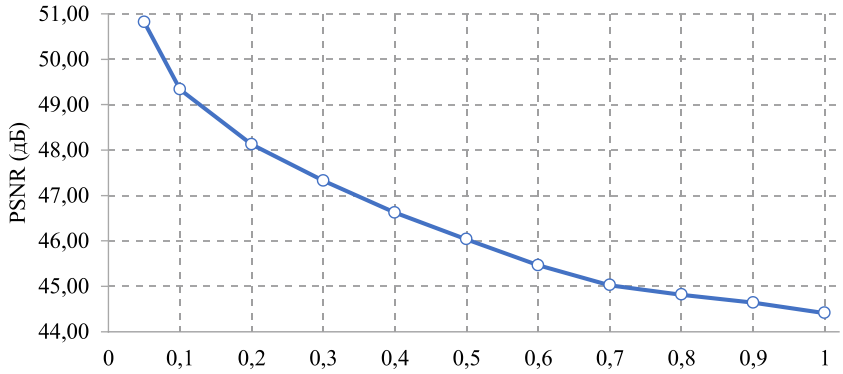



*Рис. 3. Зависимость PSNR от весового коэффициента силы встраивания α*

В качестве меры искажений, вносимых методом маркирования, используются различные метрики, такие как среднеквадратическое отклонение

$$MSE = \frac{\sum_{x,y}(I_{x,y} - I_{Wx,y})^2}{N_1 N_2}.$$

Здесь $N_1$, $N_2$ – размеры изображения в пикселях, $I_{x,y}$ – яркость пикселя $x$, $y$ в исходном изображении, $I_{Wx,y}$ – яркость пикселя в маркированном изображении. Визуальное качество маркированного стегоконтейнера должно быть как можно выше, что подразумевает незаметность вносимых искажений операцией встраивания цифрового водяного знака. Пиковое отношение уровня сигнала к уровню шума является разновидностью среднеквадратического отклонения и измеряется в децибелах:

$$PSNR = 10\log_{10}\left(N_1 N_2 \frac{\max\limits_{x,y} I_{x,y}^2}{\sum_{x,y}(I_{x,y} - I_{Wx,y})^2}\right).$$

Для обеспечения высокой робастности метод обнаружения и верификации водяных знаков должен быть устойчивым к изменениям в исходном изображении, вызванным как непреднамеренными, так и преднамеренными искажениями, которые не всегда направлены на то, чтобы полностью удалить или уничтожить водяной знак, чаще отключить его обнаружение. Вносимые методом маркирования искажения в стегоконтейнер считаются приемлемыми, если $PSNR \geq 28 \div 30$ дБ. На основании полученных экспериментальных данных для датасета изображений были получены зависимости PSNR от весового коэффициента силы встраивания α, график для отдельного эталонного изображения приведен на рис. 3.

## Заключение

В настоящей работе предложена методика защиты от неправомерного использования графической информации в социальных сетях с помощью стеганографических методов, встраивающих и скрывающих цифровые водяные знаки в частотную область вейвлет-преобразования графических изображений в сочетании с сингулярным разложением.

Вносимые алгоритмом маркирования искажения не вызывают значимой деградации исходного защищаемого изображения. В противном случае маркированный медиаконтент с встроенным водяным знаком является малопригодным для его практического использования. В маркированное изображение, содержащее цифровой водяной знак, могут вноситься случайные или преднамеренные изменения. Предлагаемая методика защиты от неправомерного использования графической информации в социальных сетях характеризуется высокой робастностью, что подтверждается экспериментальными результатами.

## Список литературы

1. Shih F.Y. Digital Watermarking and Steganography: Fundamentals and Techniques, Second Edition, CRC Press, 2017. 270 p.

2. Земцов А.Н., Аль-Макреби И.М. Об оценке вносимых искажений методом маркирования в низкочастотной области вейвлет-спектра изображения // Инженерный вестник Дона. 2015. № 2–2 (36). [Электронный ресурс]. URL: http://ivdon.ru/ru/magazine/archive/n2p2y2015/2962 (дата обращения: 28.05.2020).

3. Кочкарев А.И. Исследование и разработка многобитовых систем цифровых водяных знаков в условиях возможных атак: дис. … канд. техн. наук. Санкт-Петербург, 2019. 140 с.

4. Pennebaker W.B., Mitchell J.L. JPEG Still Image Data Compression Standard, 1st edition, Kluwer Academic Publishers, 1992. 638 p.

5. ISO/IEC 10918-1 ITU-T Rec. T. 81, Information Technology – Digital Compression and Coding of Continuous-tone Still Images, 1992. 182 p.

6. ISO/IEC 15444-1 ITU-T Rec. T. 800, Information Technology – JPEG 2000 Image Coding System: Core Coding System, 2019. 196 p.

7. Olkkonen H. Discrete Wavelet Transforms: Algorithms and Applications, IntechOpen, 2011. 308 p.

8. Baleanu D. Advances in wavelet theory and their applications in engineering, physics and technology, Books on Demand, 2012. 648 p.

9. Lokenath D., Firdous A.S. Wavelet Transforms and Their Applications, 2nd edition, Birkhauser Basel, 2015. 553 p.

10. Земцов А.Н. Робастный метод цифровой стеганографии на основе дискретного косинусного преобразования // Известия Волгоградского государственного технического университета. 2011. № 11 (84). С. 141–144.

11. Бахрушина Г.И., Коржавин В.А. Использование дискретных преобразований при разработке устойчивых алгоритмов цифрового маркирования изображений // Ученые заметки ТОГУ. 2016. № 4–1. [Электронный ресурс]. URL: http://pnu.edu.ru/media/ejournal/articles-2016/TGU_7_176_1.pdf (дата обращения: 25.05.2020).