УДК: 004.056.5 DOI: 10.53816/23061456_2022_3–4_18

МЕТОД ПОВЫШЕНИЯ ЗАЩИЩЕННОСТИ ЦИФРОВОГО МЕДИАКОНТЕНТА НА ОСНОВЕ ИСПОЛЬЗОВАНИЯ ИДЕНТИФИКАЦИОННЫХ НОМЕРОВ

METHOD OF INCREASING THE SECURITY OF DIGITAL MEDIA CONTENT BASED ON THE USE OF IDENTIFICATION NUMBERS

Канд. техн. наук В.А. Лохвицкий, канд. техн. наук С.А. Краснов, канд. техн. наук А.А. Свинарчук

Ph.D. V.A. Lokhvickii, Ph.D. S.A. Krasnov, Ph.D. A.A. Svinarchuk

ВКА им. А.Ф. Можайского

В статье рассмотрена задача встраивания дополнительной информации в видеопоток для защиты авторского (не искаженного) контента. Предложен метод, который основан на внедрении блоков дополнительной информации в области видеопотока, которые остаются неизменными при перекодировании. В основу данного метода заложен принцип наблюдения о влиянии изменения яркостной компоненты на контейнер во время его копирования, извлечения и изменения ключевых кадров. В работе выявлены и описаны расхождения в значениях, извлеченных и вновь сформированных яркостных компонент при обратном формировании цветовой модели RGB. Приведен пример встраивания дополнительной информации в сформированный стегоконтейнер с использованием видеокодеков MPEG-4 и H.264 с помощью разработанного программного средства. Рассмотрены возможные специальные метки-наполнители, которые могут быть встроены в видеопоток.

Ключевые слова: стеганография, цифровой водяной знак, идентификационный номер, дискретное косинусное преобразование, DC-коэффициент.

The article deals with the problem of embedding additional information into a video stream to protect the author's (not distorted) content. The proposed method is based on the introduction of additional information blocks in the area of the video stream, which remain unchanged during transcoding. This method is based on the principle of observing the effect of changing the luminance component on the container during its copying, extracting and changing key frames. The paper identifies and describes the discrepancies in the values of the extracted and newly formed luminance components during the reverse formation of the RGB color model. An example of embedding additional information into a generated stegocontainer using MPEG-4 and H.264 video codecs using the developed software is given. Possible special filler tags that can be embedded in a video stream are considered.

Keywords: steganography, digital water center, identification number, discrete cosine transform, DC coefficient.

Введение

В настоящее время существует проблема обеспечения гарантированной безопасности информации, передаваемой по сетям общей и военной связи [1, 9, 10]. Возникла она из-за стремительного развития и повсеместного внедрения вычислительных, телекоммуникацион-

ных и информационных технологий в различные сферы деятельности, включая военную. Кроме того, постоянно увеличивается пропускная способность телекоммуникационного оборудования, растет объем циркулирующей информации, появляются новые технологии сбора, обработки, хранения и передачи различного, в том числе цифрового мультимедийного контента. Все это приводит к появлению новых уязвимостей и снижению защищенности информации, что особенно недопустимо для сложных автоматизированных и информационных систем военного и двойного назначения [6, 9]. Для обеспечения защищенности контента на различных этапах его существования необходимо постоянно совершенствовать соответствующие способы, методы и алгоритмы защиты информации, в частности для защиты на этапе передачи цифрового медиаконтента — криптографические и стеганографические методы [2, 5, 8]. Можно выделить две основные причины актуальности исследований в области компьютерной стеганографии:

– ограничение на использование криптосредств в ряде стран мира;

– проблема защиты прав интеллектуальной собственности правообладателей и несанкционированного распространения цифрового медиаконтента.

Вторая причина привела к многочисленным исследованиям в области, так называемых цифровых водяных знаков. Цифровой водяной знак (ЦВЗ) — специальная метка, которая незаметно внедряется в цифровой медиаконтент (изображения, звук, видео) с целью контролировать его дальнейшее использование [1, 3].

Технология встраивания идентификационных номеров правообладателей схожа с технологией ЦВЗ. Основное отличие заключается в том, что в этом случае каждая защищенная копия имеет свой уникальный встраиваемый номер или в копию встраивается информация пользователя медиаконтента. Этот идентификационный номер позволяет правообладателю медиаконтента в дальнейшем отслеживать судьбу каждой легальной копии и находить источники их нелегального распространения. В последнее время это стало особенно актуальным, так как с развитием информационной сети Интернет возник вопрос о защите авторских прав правообладателей. В сети сейчас существует множество нелегальных сервисов, которые противозаконно распространяют платный медиаконтент, сфабрикованные (искаженные) материалы, и наносят убытки их правообладателям (официальным государственным, военным и частным организациям). В связи с этим правообладатели медиаконтента заинтересованы в защите своих авторских прав и в возможностях отслеживать источники распространения нелегальных копий своего цифрового медиаконтента [4, 7].

В настоящее время для этих целей активно используются следующие методы встраивания идентификационных меток в цифровой медиаконтент:

1. Метод сокрытия данных, основанный на характерных особенностях форматов файлов;
2. Метод сокрытия данных в векторах движения;
3. Метод, использующий свойство четности координат вектора для сокрытия данных;
4. Метод, который осуществляет сокрытие данных в признаке принадлежности макроблоков к определенным группам слайсов.

Как правило, обобщенный процесс сжатия видео при помощи различных кодеков включает в себя ряд этапов:

1. Цветовая субдискретизация;
2. Блочная компенсация движения;
3. Кодирование с предсказанием;
4. Квантование.

Использование изменения яркостной компоненты для встраивания дополнительной информации в блоки видеопотока

На настоящий момент нет универсального метода, который мог бы использоваться для внедрения дополнительных данных, например в видеопоток. Это обусловлено различными алгоритмами сжатия, внедряемой информацией и емкостью самого контейнера. Так, различные видеокодеки по-разному кодируют информацию, например, в блоках одного кадра могут использоваться разные матрицы квантования, множество других элементов могут сжиматься с потерей, а при попытке вмешаться в один блок весь кадр может кардинально исказиться, и в конечном итоге весь видеоряд начнет накапливать

ошибки. В связи с этим было решено разработать метод встраивания идентификационных сообщений в видеофайлы различных форматов с учетом их специфики.

Предлагаемый метод основан на наблюдении о влиянии изменения яркостной компоненты Y на контейнер во время его копирования, извлечения и изменения ключевых кадров. При извлечении кадра формируется изображение с цветовой моделью RGB, затем на ее основе формируется цветовая модель YCbCr. Было замечено, что при проведении обратной процедуры часть элементов RGB слегка изменяется, даже без внесения каких-либо изменений. Это связано с неточностью перевода из одной цветовой модели в другую.

Также было замечено, что некоторые биты остались неизменными, что означает их устойчивость к перекодированию из одной цветовой модели в другую. Из этого можно сделать вывод, что менее изменившиеся блоки являются более устойчивым к перекодированию и их можно использовать для внедрения дополнительной информации.

Для поиска таких блоков сравниваются блоки одинаковых контейнеров в потоке, над которыми произведено дискретное косинусное преобразование (ДКП) яркости, далее один из этих блоков перекодируется и собирается в новый видеоряд. Если разница коэффициентов, отображающих общую яркость блока (DC-коэффициент) в матрице квантования, равна нулю, значит, данный блок не изменился или изменился минимально и является устойчивым (рисунок).

Следующим критерием для отбора блока является его неоднородность, т.е. необходимо, чтобы DC-коэффициенты значительно отличались друг от друга, так как в случае однородности можно будет легко обнаружить наличие встраивания.

После изменения коэффициентов DC кодировщик переписывает данный блок и возвращает изображение обратно в поток. Для сокрытия данных используется признак четности, то есть при внесении значения 1 необходимо изменять коэффициенты ДКП всей матрицы, при значении 0 оставлять без изменений. Это позволит меньше влиять на изображение и увеличит скорость кодирования.

Для нахождения встроенного сообщения сравниваются пустой и заполненный контейнер. Суммы коэффициентов дискретного косинусного преобразования блоков, в которых есть возможность встроить сообщение, сравниваются, и наличие в них изменений будет говорить о нахождении в этом блоке дополнительной информации.

Таким образом, для реализации предлагаемого метода необходимо:

1. На первом этапе провести покадровое создание копии исходного файла. Вначале получается изображение из *N*-го кадра, которое

	0	1	2	3	4	5	6	7
0	85.50000	47.38007	3.24999	0.42732	-0.75000	-0.26240	0.00680	-0.05203
1	-21.68326	-5.80522	9.50490	1.84207	-0.74703	0.03977	0.12796	-0.49855
2	1.11522	-3.65118	-2.78921	3.02540	0.78858	0.36876	-0.19822	-0.32425
3	-2.31512	0.09738	0.68331	0.48591	0.22807	0.40473	0.08642	-0.34823
4	0.50000	-0.69073	-0.17493	-0.01858	0.25000	0.70279	0.88425	1.55037
5	-0.75433	-0.37144	-0.60691	-0.02205	0.24046	-0.83426	0.07748	0.42060
6	0.07926	-0.28080	0.55178	-0.00208	-0.05604	-0.03509	0.03921	-0.44114
7	-0.39683	-0.42533	0.51808	0.63701	0.28654	0.06174	-0.15290	0.15357

	0	1	2	3	4	5	6	7
0	85.50000	41.49494	-2.23044	0.19588	-1.00000	0.93599	-0.15851	-0.07561
1	-19.07911	-1.62797	10.64288	1.49298	-0.30892	-1.47495	-0.26205	-0.04691
2	1.72887	-2.81068	-3.49632	1.60410	1.80813	1.85752	-0.30178	-0.36399
3	-3.82175	-0.39936	1.21840	0.66865	0.28304	0.41503	0.21154	-1.14127
4	0.75000	-0.27282	-0.40590	0.37170	0.75000	0.21189	0.97992	0.99625
5	-0.28346	-0.43195	0.06212	0.23825	0.25139	0.31879	-0.13066	0.83808
6	0.90746	0.11211	-0.05178	-0.55553	-0.20776	0.18438	0.74632	0.29820
7	0.65283	-0.22369	-0.32366	0.23751	-0.36914	-0.43380	0.07399	0.14053

Рис. Пример блока с минимальными потерями значений DC-коэффициентов

потом переводится в массив RGB, после в массив YCbCr и обратно, в массив RGB. Вновь полученное изображение собирается в кадр и записывается в файл. Далее в пустом и первоначальном контейнере берутся ключевые кадры, и значения коэффициентов DC блоков яркости сравниваются;

2. Из файла с координатами блоков берутся номера кадров и данные о местоположении необходимых блоков, из другого файла берется сообщение, которое необходимо скрыть. Сообщение переводится в двоичный код. Создается новый файл и кадры, которые упоминались в маске, а именно в нем все блоки дискретного косинусного преобразования в них, кроме DC-коэффициентов, модифицируются согласно скрываемому сообщению. После выполняется обратное ДКП блоков и значения в кадре заменяются на измененные значения блока.

Для чтения внедренного сообщения из заполненного контейнера необходимо проделать следующую операцию, взять координаты модифицированных блоков из заранее созданного файла маски, и вычислить сумму ДКП этих блоков. Если сумма не изменилась — присваивается значение 0, иначе присваивается 1. Значения заносятся в отдельную строку, из которой затем берется по 8 элементов. Эти значения переводятся по 8 бит в символ кодировки utf-8 и сохраняются посимвольно в отдельный файл.

Анализ результатов работы программного средства для внедрения дополнительной информации в видеопоток

Для изучения результатов работы предложенного метода было разработано программное средство на языке программирования Python. В качестве опытного образца было взято видео с разрешение 1280×720 px, продолжительностью 13,54 минуты и в него циклично встраивалось сообщение объемом 3 бит на каждые 8 возможных. Тесты проводились на 2 видеокодеках — MPEG-4 и H.264.

По данным табл. 1 можно сделать вывод, что, пустой контейнер на первый взгляд неотличим от контейнера с сообщением, но при более детальном рассмотрении видно, что объем контейнера с сообщением больше на 10 Кб при внесении 590 битов.

По данным табл. 2 можно сделать вывод, что, пустой контейнер на первый взгляд неотличим от контейнера с сообщением, но при более детальном рассмотрении видно, что объем контейнера с сообщением больше на 2 Кб при внесении всего 40 битов.

При сравнении самих изображений также не было замечено каких-либо заметных искажений. Замечено, что объем встраиваемой информации в контейнер ограничен и зависит от кодека — чем лучше метод сжатия, тем меньше объем встраивания.

Таблица 1

Сравнение данных видеофайлов при использовании видеокодека MPEG-4

Тип файла	Скорость потока	Объем файла	Частота кадров
Оригинал	987 кбит/с	98,2 Мб	24
Пустой контейнер	1025 кбит/с	102 Мб	24
Контейнер с сообщением	1025 кбит/с	102 Мб	24

Таблица 2

Сравнение данных видеофайлов при использовании видеокодека H.264

Тип файла	Скорость потока	Объем файла	Частота кадров
Оригинал	1011 кбит/с	100 Мб	24
Пустой контейнер	1013 кбит/с	100 Мб	24
Контейнер с сообщением	1013 кбит/с	100 Мб	24

В качестве информации (специальные метки-наполнители), встраиваемой для защиты авторского медиаконтента, можно использовать:

1. Символьные метки законных правообладателей копий авторского медиаконтента;

2. «Цифровой след» с компьютера, на котором воспроизводится авторский медиаконтент;

В качестве «цифрового следа» можно использовать IP-адрес клиента или Canvas API, WebGL, Audio Context API.

Canvas — это HTML 5 API («Application Programming Interface» — интерфейс программирования приложений, программный интерфейс приложения), который используется для рисования графики и анимации на веб-странице с помощью сценариев JavaScript. Но кроме этого canvas можно использовать в качестве дополнительной энтропии при сборе цифровых отпечатков в веб-браузере. Основанием для этого служит тот факт, что одно и то же изображение, отрисованное при помощи canvas, может отображаться по-разному на разных компьютерах. Это происходит по нескольким причинам:

– на уровне формата изображения веб-браузеры используют разные механизмы обработки изображений, параметры экспорта изображений, уровень сжатия, из-за чего конечные изображения могут иметь разные контрольные суммы, даже если они состоят из одинаковых пикселей;

– на системном уровне — операционные системы имеют разные шрифты, они используют разные алгоритмы и настройки для сглаживания и субпиксельного рендеринга.

Функция Canvas API под названием toDataURL() возвращает закодированные в строку base64 двоичного файла изображения, которой затем хэшируются и используются вместе с другими данными для создания уникального цифрового отпечатка.

WebGL — это JavaScript API для рендеринга интерактивной 3D-графики с помощью видеокарты в любом совместимом веб-браузере без использования плагинов. Приложения WebGL состоят из управляющего кода, написанного на JavaScript, и кода специальных эффектов, который выполняется на графическом процессоре компьютера. Элементы WebGL можно смешивать с другими элементами HTML и комбинировать с другими частями страницы или ее фона.

Audio Context API — это интерфейс Web Audio API, который поддерживают все современные браузеры. Главный принцип заключается в наличии аудиографика, в котором объединен ряд объектов для визуализации характеристик воспроизведения. Результатом является отпечаток, который зависит от параметров системы и конфигурации программного обеспечения.

При попытке расшифровать сообщение выяснилось, что алгоритм искажает около 6 % внесенной информации. Чтобы не потерять важную информацию, необходимо использовать самокорректирующиеся коды и цикличность сообщения, т.е. когда сообщение полностью встраивается, но все еще остались незакодированные, пустые блоки, то запись сообщения в эти блоки повторяется еще раз.

Было отмечено, что объем скрываемых данных и исполнение программы зависят не только от размера файла, но и от частоты опорных кадров.

Данный метод неустойчив к помехам, так как при перекодировании скрытая информация о разности коэффициентов бит становится не актуальной. Это было проверено путем загрузки видео со скрытым сообщением в социальные сети.

Плюсы разработанного метода:

– высокая скорость выполнения алгоритма;

– универсальность;

– сложность в определении наличия скрытого сообщения.

Минусы метода:

– неоднозначность восстановления информации;

– малый объем встраиваемого сообщения в контейнер.

Заключение

Несмотря на отмеченные недостатки, выявленные в ходе экспериментального исследования, предложенный метод может быть применен для защиты прав интеллектуальной собственности правообладателей и предотвращения несанкционированного распространения цифрового медиаконтента в различных организациях, учреждениях, работающих с конфиденциальными документами. За счет достаточно высокой

оперативности предложенного метода метка-наполнитель может крайне быстро встраиваться в контейнер незаметно для пользователя, запросившего медиаконтент.

Литература

1. Грибунин В.Г. Костюков В.Е., Мартынов А.Н. и др. Стеганографические системы. Цифровые водяные знаки. Саров. ФГУП «РФЯЦ-ВНИИЭФ». 2016. 210 с.

2. Грибунин В.Г., Оков Н.Н., Туринцев И.В. Цифровая стеганография. М.: СОЛОН-ПРЕСС. 2009. 272 с.

3. Конахович Г.Ф., Пузыренко А.Ю. Компьютерная стеганография. Теория и практика. — Киев: «МК-Пресс». 2006. 288 с.

4. Aly H.A. Data hiding in motion vectors of compressed video based on their associated prediction error // IEEE Trans. Inform. Forensics Security. 2011. V. 6, № 1. P. 14–18.

5. Макаренко С.И., Ковальский А.А., Краснов С.А. Принципы построения и функционирования аппаратно-программных средств телекоммуникационных систем: учеб. пособие. Ч. 2. Сетевые операционные системы и принципы обеспечения информационной безопасности в сетях. СПб.: Наукоемкие технологии. 2020. 357 с.

6. Бубнов В.П. и др. Модели информационных систем: учебное пособие. М.: Учебно-методический центр по образованию на железнодорожном транспорте. 2015. 188 с.

7. Абазина Е.С., Ерунов А.А. Результаты моделирования метода скрытой передачи информации с кодовым уплотнением в видеоданных // Системы управления, связи и безопасности. 2015. № 2. С. 1–25. URL: journals.intelgr.com/sccs/20152/html (дата обращения 3.10.2021).

8. Набаева К.А. Разработка не обнаруживаемых стегосистем для каналов с шумом: дис. … канд. техн. наук: 05.12.13. СПб.: СПбГУТ. 2014. 176 с.

9. Цветков К.Ю. Методы цифровой стеганографии и их приложения в сетях спутниковой связи // Сборник трудов II ВНК Космических войск. 2004. СПб.: МО РФ. Т. 2. С. 344–349.

10. Макаренко С.И., Блатов И.А., Макаров М.И. Эталонная модель взаимодействия стеганографических систем и обоснование на ее основе новых направлений развития теории стеганографии // Вопросы кибербезопасности. 2014. № 2 (3). С. 24–32.

References

1. Gribunin V.G. Kostyukov V.E., Martynov A.N. and others. Steganographic systems. Digital watermarks. — Sarov. VSUE «RF YC VNIIEF». 2016. 210 p.

2. GribuninV.G., Okov N.N., Turintsev I.V. Digital steganography. M.: SOLON-PRESS. 2009. 272 p.

3. Konakhovich GF, Puzyrenko A.Yu. Computer steganography. Theory and practice. Kiev: «MK-Press». 2006. 288p.

4. Aly H.A. Data hiding in motion vectors of compressed video based on their associated prediction error // IEEE Trans. Inform. Forensics Security. 2011. V. 6, № 1. P. 14–18.

5. Makarenko S.I., Kovalsky A.A., Krasnov S.A. Principles of construction and functioning of hardware and software of telecommunication systems: textbook. allowance. Part 2. Network operating systems and principles of information security in networks. SPb: Science-intensive technologies. 2020. 357 p.

6. Bubnov V.P. and other/ Models of information systems: a tutorial. Moscow: Training and Methodological Center for Education in Railway Transport/ 2015. 188 p.

7. Abazina E.S., Erunov A.A. Modeling results of the method of covert information transmission with code compaction in video data // Control systems, communications and security. 2015. № 2. P. 1–25. URL: journals.intelgr.com/sccs/20152/html (date of treatment 10/03/2021).

8. Nabaeva K.A. Development of undetectable stegosystems for channels with noise: dis. ... Cand. tech. Sciences: 05.12.13. SPb.: SPbGUT. 2014 176 p.

9. Tsvetkov K.Yu. Methods of digital steganography and their applications in satellite communication networks // Proceedings of the II VNK of the Space Troops. 2004. SPb: Ministry of Defense of the Russian Federation. Vol. 2. P. 344–349.

10. Makarenko S.I., Blatov I.A., Makarov M.I. Reference model of interaction of steganographic systems and justification on its basis of new directions of development of the theory of steganography // Issues of cybersecurity. 2014. № 2 (3). P. 24–32.