УДК 004.056:003.26

С.С. Барильник, Н.Е. Герасимов, И.В. Минин, О.В. Минин

Новые подходы к программно-технической защите сетевых публикаций

Рассматриваются методы программно-технической защиты сетевых публикаций от копирования и нелегального использования средствами новых стелсографических алгоритмов защиты авторского права, разработанные авторами.

Согласно [1] «мультимедийный» web-документ, строго говоря, представляет собой совокупность различных работ, защищаемых законом об авторском праве. Кроме того, web-документ, содержащий базу данных, может подпадать под другие законы. В то же время существует ряд технических особенностей сети, которые существенно осложняют защиту авторских и смежных прав. Например, легкость создания копий в неограниченном количестве, а также легкость записи на жесткий диск персонального компьютера частей Интернет-сайта (что является нарушением права на воспроизведение) делает каждого пользователя сети потенциальным нарушителем законодательства.

Вопрос о том, какие из исключительных прав владельца защищенной авторским правом работы могут быть нарушены в результате распространения работ через сеть Интернет, был предметом теоретических споров. Совершенно очевидно, что нарушается право копирования, поскольку распространение подразумевает, что:

- для того чтобы появиться на сайте, работа представляется на языке разметки гипертекста (HTML) (что может рассматриваться как «воспроизведение в материальной форме» либо как нарушение права автора на адаптацию или перевод);
- работа хранится в виде компьютерной программы на главном сервере (в Великобритании, например, определение копирования включает в себя «хранение в электронной форме на любом виде носителей»);
- работа воспроизводится на компьютерах всех «посетителей» web-сайта. Надо заметить, что в данном случае «разрешение» на нарушение, данное третьей стороне, является первичной формой нарушения, таким образом, человек, который предоставляет третьим сторонам доступ к работе, может считаться санкционирующим копирование, которое, в техническом смысле, происходит каждый раз, когда посетители web-сайта достигают соответствующей страницы.

Начальник бюро специальных технических мероприятий (БСТМ) МВД России Борис Мирошников недавно заявил, что в России «достаточно... оригинальным считается преступление, которому быстро нашелся специальный термин — «фишинг» (уже давно используется мошенниками по всему миру). Это размещение в Интернете сайтов-близнецов известных компаний и банков с целью «выуживания» информации о счетах и платежных реквизитах [2]. В этой связи проблема «распознавания» поддельных страниц сайтов приобретает весьма актуальный характер. «Мы рекомендуем службам информационной безопасности ежедневно вести активный поиск в Интернете подобных двойников, их организации и уведомлять нас о подобных находках», — подчеркнул руководитель БСТМ.

Для доказательства авторства спорной работы используется несколько механизмов. К основным способам защиты сетевых публикаций можно отнести [3]:

- публикацию статьи на бумажном носителе (самый простой и надежный вариант при условии, конечно, что дата публикации более ранняя, чем дата появления контрафактного экземпляра);
- засвидетельствование у нотариуса даты создания статьи (этот способ защиты осуществляется путем нотариального заверения распечатки подготовленной статьи с указанием даты и автора произведения);
- иной способ удостоверения факта существования статьи на определенную дату (например, можно отправить самому себе обычное письмо, почтовый штемпель на конверте и будет подтверждением);
- программно-техническую защиту (подразумевается использование программы, предназначенной для защиты в сети прав и законных интересов авторов электронных публикаций, путем отображения публикаций способом, исключающим их копирование и/или иное несанкционированное размножение, модификацию).

В вопросе защиты контента (содержимое) сайта от копирования существует несколько отличных друг от друга программно-технических подходов [4]. Применение специальных стелсографических технологий, позволяющих размещать согласно ч. 4 ГК РФ внутри web-документа «невидимых» для постороннего сообщений позволит частично решить эту проблему.

Нами была разработана новая стелсографическая технология размещения скрытых от постороннего взгляда сообщений, произвольной длины и содержания в HTML-кодах, не изменяющих внешний вид, содержание и дизайн собственно этой страницы [5–7]. В принципе, данную технологию можно рассматривать как аналог цифровых водяных знаков.

Данный способ передачи скрытой информации относится к нестандартным способам передачи информации по легальным каналам (так называемые потайные каналы subliminal channels). Данные каналы используют тогда, когда имеется легальный коммуникационный канал, но политика безопасности запрещает передавать по нему определенную информацию; иными словами, информацию передавать можно, но она не должна выглядеть подозрительно (в соответствии с некими, обычно не очень четкими критериями). Алгоритм работы реализованной программы следующий.

Скрытие сообщения:

1. Чтение исходного HTML-файла и сообщения.
2. Кодирование сообщения.
3. Оптимизация исходного HTML-файла и подсчет количества мест для вставки закодированного сообщения.
4. Разбиение закодированного сообщения на части.
5. Выделение памяти для HTML-файла со скрытым сообщением.
6. Вставка частей закодированного сообщения в HTML-файл.
7. Запись результатов в файл.

Извлечение сообщения:

1. Чтение HTML-файла со скрытым сообщением.
2. Поиск, извлечение и декодирование сообщения.

Анализ основных свойств разработанного алгоритма сводится к следующему:

- стегоканал не обнаруживается визуально, и его можно определить только при сравнении объема контейнера V_0, свободного от скрываемой информации, объемом заполненного контейнера V_1 (так как $V_1 > V_0$), если заранее известен V_0. Таким образом, если V_0 не известен, то система обладает стегостойкостью;
- сжатие методами стандартного архивирования (RAR, ZIP и т.д.) заполненного контейнера не приводит к искажению скрываемой информации;
- ограничений по использованию типа встраиваемого сообщения нет (в качестве такого были использованы текстовый файл, звуковой файл формата MP3, видеосигнал формата MPEG4, графическое изображение, сжатое алгоритмом сжатия JPEG);
- в отличие от известных алгоритмов стеганографии на основе методов преобразования текста, позволяющих скрывать сообщения в HTML-файлах (например, реализованный в программе Steganos for Window), в описанном методе передачи скрытых сообщений нет ограничений на объем контейнера в зависимости от объема скрываемого сообщения;
- в отличие от известных алгоритмов стеганографии на основе методов преобразования текста, в описанном методе передачи скрытых сообщений нет ограничений и на содержание контейнера;
- передача сообщений осуществляется анонимно, т.е. получатель сообщения может находиться в любом, заранее не известном месте, достаточно иметь доступ к закодированной странице;
- скрытие информации в HTML-коде может проводиться многократно, т.е. поверх одного сообщения скрываться другое, при этом все свойства алгоритма остаются неизменными;
- разработанные меры и реализация алгоритма дополнительно позволяют снизить вероятность визуального обнаружения факта наличия скрытого сообщения.

Следует отметить, что разработанная методика позволяет в принципе скрывать в HTML-контейнере сообщения произвольного размера. Более того, как уже указывалось, вопрос защиты контента (содержимое) сайта от копирования может быть основан на использовании возможностей языка JavaScript. Однако и сами программы на этом языке нуждаются в защите. Предлагаемая технология защиты авторских прав работоспособна и для скриптов, написанных на JavaScript.

Для непосредственного скрытия информации в web -странице можно использовать один из стеганографических алгоритмов, адаптированных для HTML:

1. Биты скрываемой информации представляются в виде непечатаемых символов. Такими символами являются «Пробел» и «Горизонтальная табуляция». Таким образом, можно представить биты в виде символов: «1» – «Пробел», «0» – «Горизонтальная табуляция». Каждый байт скрываемой информации преобразуется в последовательность таких символов, где каждому символу соответствует бит скрываемого байта. Например, скрываемый байт 0Ч43 = 0100 0011 => « || | | | |||», где || – «Пробел», | | – «Горизонтальная табуляция». Далее полученная последовательность помещается в конец строки и становится «невидима». По такому принципу можно скрыть один байт информации в одной строке [5].
2. В Windows для перевода строки используются два символа: 0Ч0D и 0Ч0A. В современных Unix операционных системах для этого достаточно одного символа: 0Ч0A. Большинство текстовых редакторов понимают и правильно отображают оба формата перевода строк. Пользуясь этой особенностью, можно «прятать» биты скрываемой информации: «0» – 0Ч0A, «1» – 0Ч0D 0Ч0A, т.е. наличие 0Ч0D является «1». По такому принципу можно скрыть один бит информации в одной строке [5].

В случае защиты web-документа контейнером является HTML-файл. Для увеличения стегостойкости информации будем вставлять полученные последовательности не в конец каждой строки, а только в конце строк, заканчивающихся на тэг (<html>, </titled>, </body>,
, <td> и т.д.). Это

позволяет не отображать скрытых пробелов и горизонтальных табуляторов на странице при работе с первым алгоритмом.

Второй алгоритм имеет свои особенности. Если злоумышленник будет знать, что данные передаются, и догадается, что они скрываются таким образом (наличие 0Ч0D в конце строки, так как в конце разных строк будут разные окончания), то с целью маскировки спрятанных бит информации от их визуального анализа в Hex-редакторе, можно записать в конец каждой строки текста разные окончания (случайным образом), а считывать только необходимые окончания (в данном случат, только строки, заканчивающиеся на тэг). Поэтому прочитать такие биты, не зная маски, невозможно, что увеличивает стойкость к одной из самых распространенных атак – DOS-атак – анализ статистических данных, путем анализа всего документа.

В статье [5] также приведен сравнительный анализ предложенных алгоритмов, который показал, что второй алгоритм обладает большей стегостойкостью, но его пропускная способность гораздо меньше первого, так как он более требователен к размеру контейнера (больше 100 Кб).

Следует отметить, что описанные алгоритмы и их практическая реализация внедрены на официальном сайте фирмы, занимающейся разработкой программного обеспечения ООО «Графические программные системы», в качестве защиты авторского права на проект «МирКибер» в июне прошлого года.

Таким образом, в работе предложена и исследована новая стелсографическая технология скрытой передачи информации произвольного содержания в HTML-файлах. Разработана демонстрационная версия программы. Определены основные свойства алгоритма, при этом если заранее не известен размер файла-контейнера, то алгоритм обладает стегостойкостью. Предлагается использовать данный механизм скрытой передачи данных для защиты авторских прав в web-документах и контроля их целостности. Дальнейшие работы в этом направлении будут направлены на оптимизацию предложенного алгоритма и улучшения его свойств.

Перспективными областями применения данного механизма передачи скрытой информации, по мнению авторов, являются:

- Скрытая аннотация документов (в том числе мультимедийные базы данных).
- «Цифровой сертификат», скрытый в HTML-коде страницы,.
- Контроль целостности документов.
- Скрытая связь (особенно включая случаи, когда криптографические методы использовать нельзя).
- Аутентификация (электронная коммерция, электронная почта, электронное конфиденциальное делопроизводство).
- Защита от копирования.
- Деловой документооборот с конфиденциальной информацией, включая «ноу-хау» и т.п.

Литература

1. Информационные технологии в бизнесе / Под ред. М. Желены. – СПб: Питер, 2002. It-курьер, №5, 2007.

2. Серго Антон. Некоторые вопросы защиты авторского права в Интернете (теория и практика). Режим доступа http://www.internet-law.ru/articles/ap-is.htm, свободно.

3. Белов Вячеслав. Защита контента. Режим доступа http://kis.pcweek.ru/Year2004/N5/CP1251/TematicReviews/chapt1.htm, свободно.

4. Барильник С.С., Минин И.В., Минин О.В. Адаптация алгоритмов текстовой стеганографии для HTML // 8 междунар. Сиб. школа-семинар по электронным приборам и материалам EDM'2007. Novosibirsk: NSTU, 2007. – P. 225–228.

5. Барильник С.С., Минин И.В., Минин О.В., Щетинин Ю.В. Текстовая стеганография в HTML: реализация скрытых каналов передачи данных // Вторая междунар. Науч.-практ. конф. «Виртуальные и интеллектуальные системы», 2007. (Ползуновский Альманах). Барнаул: АГТУ. 2007. С. 28–29.

6. Минин И.В., Минин О.В., Герасимов Н.Е. Стелсографическая защита интеллектуальной собственности на на документы в WWW // Восьмой международной симп. ТЕХНОМАТ 2007 «Материалы, Методы и Технологии». – Болгария. 28 мая – 1 июня.

Барильник Станислав Сергеевич
Новосибирский Государственный Технический Университет
Факультет автоматики и вычислительной техники, кафедра защиты информации, студент 4-го курса
Эл. адрес: krutoystas@mail.ru

Герасимов Николай Евгеньевич
Новосибирский Государственный Технический Университет, аспирант.
Эл. адрес: nikolay.gerasimov@gmail.com.

Минин Игорь Владиленович
Новосибирский Государственный Технический Университет
Доктор технических наук, профессор кафедры Защиты Информации
Эл. адрес: prof.minin@gmail.com.

Минин Олег Владиленович
Новосибирский Государственный Технический Университет
Доктор технических наук, профессор кафедры Защиты Информации
Эл. адрес: prof.minin@gmail.com.

S.S. Barilnik, N.E. Gerasimov, I.V. Minin, O.V. Minin
New approaches to software-technical protection of network publications

In given article are considered the methods of program-technical protection of network publications publications by means of new stelthographic algorithms of the copyright protection against copying and illegal using, which was written authors.