УДК 004.056

Д. М. Андрущенко, Г. Л. Козина, Л. М. Карпуков
Запорожский национальный технический университет
ул. Жуковского, 64, 69063 Запорожье, Украина

Защита информации от несанкционированного распространения в сети Интернет без ссылки на первоисточник

Рассмотрена проблема нарушения авторских прав при размещении информации в сети Интернет. Предложен способ, позволяющий создавать препятствия для нарушения авторского права при публикации информации в сети Интернет.

Ключевые слова: *текстовая информация, защита, авторское право, Интернет.*

Введение

Наиболее распространенный способ представления текстовой информации в электронном виде в сети Интернет путем использования стандартного языка разметки документов HTML [1] имеет ряд недостатков. Основной из них — это возможность беспрепятственного копирования и автоматизированного распространения информации без согласия автора (авторов) и без ссылки на первоисточник. Очевидно, что на получение информации, создание интеллектуального продукта и представление результатов работы тратятся значительные средства. Поэтому для авторов актуально иметь возможность создания препятствий для дублирования и распространения без ссылки на первоисточник продукта их труда.

В последнее время распространяется тенденция перевода фондов библиотек в цифровую форму. Например, в фондах Национальной библиотеки Украины имени В.И. Вернадского насчитывается 53 тыс. документов в электронной форме [2]. Но легкость копирования материалов в цифровом формате допускает возможность многочисленного нарушения авторских прав.

Таким образом, является актуальным совершенствование и разработка надежных методов защиты текстовой информации от несанкционированного копирования и распространения с нарушением авторских прав.

Анализ существующих методов

Одним из подходов решения данной проблемы является использование системы защищенного просмотра документов DefView [3, 4], разработанной компа-

нией «Шофт». Согласно этому способу несанкционированное копирование страниц документа предотвращает клиентское приложение, запрещающее снимок экрана и кэширования страниц в постоянной памяти компьютера пользователя. Недостатком этого способа является незащищенность от известных альтернативных методов захвата изображения экрана, например, с помощью известной программы Screenshot Captor [5]. В этом случае, путем использования программ оптического распознавания символов, информация легко может быть представлена в любом незащищенном формате и беспрепятственно распространяться.

Известны также способы защиты файлов формата PDF (Portable Documents Format) от редактирования, печати и распространения, которые заключаются в том, что информацию шифруют, а для расшифровки необходимо знать секретный ключ [6, 7]. Но, с другой стороны, известны также способы взлома защиты такого вида [8]. Кроме того, и в этом случае возможно применение методов захвата изображения экрана и распознавания символов.

В [9] рекомендуется способ защиты программной продукции от несанкционированного использования и копирования, предназначенный для создания технологических препятствий нарушению авторских и смежных прав при распространении и использовании программной продукции. Способ заключается в том, что наиболее ценная часть программного продукта, которая интерпретирует главный алгоритм, размещается на удаленном сервере и предоставляется в режиме хостинга. Недостатком такого подхода является низкая эффективность, особенно в том случае, если ценность представляют не алгоритмы обработки данных, а именно исходная информация. Ведь публикуемые сведения могут быть перехвачены на одном из двух этапов — при передаче ее от удаленного сервера к части программного обеспечения, которая принимает эти данные, либо на этапе, когда данные сформированы в виде, приспособленном к восприятию пользователем.

Известен [10, 11] способ представления тестовой информации в комбинации с сеткой, который позволяет защитить от подделок изделия издательско-полиграфического дела (банкноты, ценные бумаги, этикетки и т.д.). Согласно данному способу, строится защитная сетка, состоящая из большого количества очень тонких линий (40–70 мкм), построенных на основе аналитических выражений, обычно окрашенных в разные цвета, плавно изменяя оттенок из одного в другой. Основная цель этого способа — создание сетки, которую очень трудно воспроизвести без специального оборудования. Этот способ предназначен, прежде всего, для защиты от подделок документов, которые выступают носителями информации, поэтому практически не пригоден для защиты информации, представленной в самом документе, от дублирования: 1) сетка не защищена от возможности ее удаления из электронного документа по одному из признаков — толщина линий или их цвет; 2) алгоритм построения сетки достаточно сложен, поэтому требует сравнительно много времени при наложении на большие объемы текстовой информации.

Цель данной работы состоит в разработке более эффективного способа защиты информации, предоставляемой в электронном виде, от несанкционированного распространения без ссылки на первоисточник.

Метод защиты путем наложения сетки

Поставленная задача достигается тем, что строят графические элементы и, объединяя их, образуют сетку и накладывают ее на информацию, представленную в виде, предназначенном для зрительного восприятия человеком [12].

Идея метода.

1. Выбирают базовые точки $(x_i; y_i)$ согласно уравнениям $x_i = f_1(i,h)$, $y_i = f_2(i,h)$, $i = 1,...,N$, где $f_1(u,v)$, $f_2(u,v)$ — произвольные функции от двух переменных (u,v); h — хэш данных, которые защищаются; N — произвольное целое число, $N > 10$.

2. Через базовые точки проводят график интерполяционной функции, толщина линии и цвет сетки совпадает с толщиной и цветом букв представленной информации.

3. Накладывают название первоисточника информации

4. Комбинированную информацию воспроизводят в растровом формате и генерируют новый файл в формате jpg, djvu и т.п.

5. При необходимости в растровое изображение с защищенной информацией встраивают невидимый водяной знак с идентификационной информацией о пользователе, которому предоставлен доступ к материалам и датой предоставления доступа.

Случайный выбор базовых точек и интерполяционной функции исключает возможность автоматического удаления сеток. Если базовые точки выбирают согласно уравнениям $x_i = f_1(i,h)$, $y_i = f_2(i,h)$, то при генерации сеток для идентичных блоков информации они полностью совпадают. Это исключает возможность отделения сеток от информации. Идентичность толщины линий сетки с толщиной линий символов и цвета сетки с цветом символов не дает возможность автоматического удаления сеток по этим признакам. Наложение вместе с сеткой первоисточника информации исключает возможность распространения информации без ссылки на первоисточник. Идентификатор в виде невидимого цифрового водяного знака (ЦВЗ) дает возможность идентифицировать правонарушителя в случае несанкционированного распространения информации.

Пример использования метода

Например, необходимо защитить электронный текстовый документ, который подан в формате Open Office с любой информацией. Документ разбивают на страницы и преобразуют в изображения размером 2500×3000 пикселей. Пусть хэш документа $h = 25143228238232323$. Для построения базовых точек возьмем $N = 30$, $f_1(i,h) = 80 \cdot (i-1)$, $f_2(i,h) = 480 \cdot (i-1) + r(h)$, $i = 1,2,...,30$, $r = -150 + (3450368945694567 \cdot h + 85056975543689) \bmod 150$.

Для построения сетки через базовые точки проводят интерполяционную функцию сплайнами третьего порядка. Сетку накладывают на документ и воспроизводят в растровом формате. В качестве данных, характеризующих авторство информации, используется название высшего учебного заведения «ЗНТУ». При необходимости в полученные изображения могут быть встроены невидимые циф-

ровые водяные знаки по методу [13], содержащие информацию о лице, которому представлен доступ и о дате и времени предоставления доступа, например: «Иванов И.П., 01.02.2010, 10:55:16». Генерируют выходные файлы doc1.jpg (рис. 1), doc2.jpg и т.д.

Сетки, наложенные на разные страницы документа, существенно отличаются. На рис. 2,*а* показан фрагмент изображения с защитной сеткой без невидимого водяного знака, на рис. 2,*б* — изображение, содержащее цифровой водяной знак. Изображения не существенно отличается одно от другого.

Пример реализации функции на языке программирования C# представлен в листинге.

Листинг. Фрагмент программы на языке программирования C#:

```
            void DrawingCurve(Pen pen,Graphics gr)
            {
              Point[] apt = new Point[25];
              System.Drawing.Image nameImage =
System.Drawing.Image.FromFile("name.png");
      //          g.ScaleTransform((float)objImage.Width / WatermImage.Width,
(float)objImage.Height / WatermImage.Height);
              Random rand = new Random();
              Rectangle destRect1 = new Rectangle(0, 0, nameImage.Width,
nameImage.Height);
              GraphicsUnit units = GraphicsUnit.Pixel;
              Random r = new Random();
              for (int y = -1000; y < 2000; y += 200)
              {
                apt = new Point[204];
                int inext = 0;

                for (int x = 4000; x >= 0; x -= 40)
              {
                apt[inext] = new Point(x, x + y + r.Next(130));
                inext++;
                apt[inext] = new Point(x, x + y - r.Next(130));
                inext++;
              }
              gr.DrawCurve(pen, apt, 3, 195, 0.8f);
            for (int y = -1000; y < 2000; y += 200)
            {
              int inext = 0;
              for (int x = 0; x <= 4000; x += 40)
              {
                apt[inext] = new Point(x, 1000 - x + y + r.Next(30));
                inext++;
                apt[inext] = new Point(x, 1000 - x + y - r.Next(30));
```

```
      inext++;
    }
    gr.DrawCurve(pen, apt, 3, 195, 0.8f);
  }
```

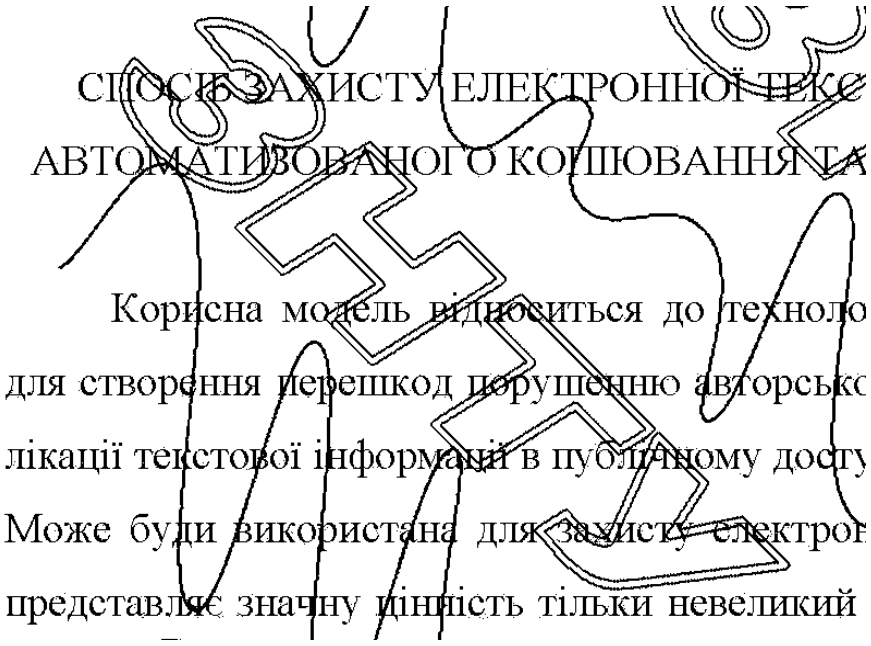

Рис. 1. Фрагмент защищенной текстовой информации

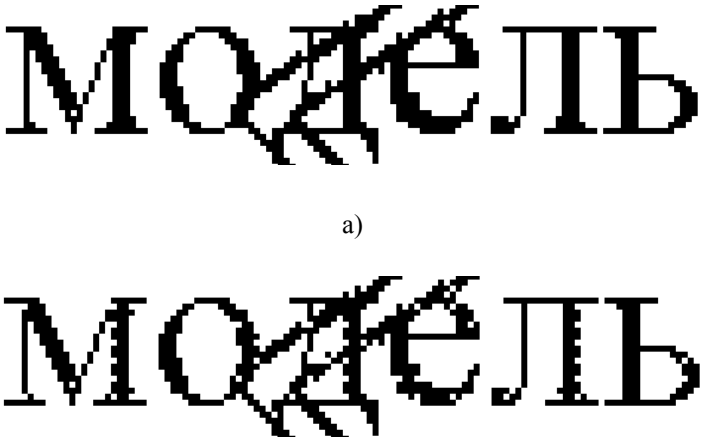

Рис. 2. Увеличенный фрагмент защищенной текстовой информации: а) без ЦВЗ; б) с ЦВЗ

Выводы

Предложенный способ может быть использован для защиты электронной или печатной информации, которая состоит из предложений, формул, таблиц, графиков и рисунков от автоматизированного дублирования без ссылки на первоисточ-

ник в том случае, когда перепечатку информации путем восприятия и воспроизведения человеком можно считать неэффективной. Достигнута низкая эффективность попыток автоматического распознавания текста, защищенного сеткой. Количество ошибок при распознавании известной программой Adobe Fine Reader 9.0 документа с защитной сеткой на рис. 1 достигает 50 %, без защитной сетки — 2 %. Распространение информации вместе с защитной сеткой сопровождается информацией о первоисточнике и идентификатором правонарушителя.

1. *Язык* разметки гипертекста [Электронный ресурс]. — Режим доступа: http://ru.wikipedia.org/wiki/HTML

2. *Склад* електронного фонду Національної бібліотеки України імені В.І. Вернадського [Електронний ресурс]. — Режим доступу: http://www.nbuv.gov.ua/eb/pub.html

3. *Система* защищенного просмотра документов DefView [Электронный ресурс]. — Режим доступа: http://diss.rsl.ru/?menu=infoblockru/infoblockru42/

4. *Пример* внедрения программы DefView в Российскую государственную библиотеку [Электронный ресурс]. — Режим доступа: http://www.shoft.ru/clients/case_rsl/

5. *Screenshot* Captor [Электронный ресурс]. — Режим доступа: http://www.donationcoder.com/Software/Mouser/screenshotcaptor/

6. *Protect* PDF Files and Documents [Электронный ресурс]. — Режим доступа: http://www.adobe.com/products/acrobatpro/protect-pdf-files-documents.html

7. *Secure* PDF File Viewer for Complete PDF File Security [Электронный ресурс]. — Режим доступа: http://www.locklizard.com/pdf_security_viewer.htm

8. *Скляров Д.В.* Искусство защиты и взлома информации [Электронный ресурс] / Д.В Скляров. — СПб.: БХВ-Петербург, 2004. — 288 с.

9. *Пат.* 5194 Україна, МПК G06F12/14,G06K13/24. Спосіб захисту програмної продукції від несанкціонованого використання та копіювання [Електронний ресурс] / Г.М. Дашкієв. — № 20040706182; заявл. 26.07.2004; опубл. 15.02.2005, Бюл. № 2, 2005 р. — 6 с.: іл.

10. *Пат.* 38479 Україна, МПК G06K15/22. Спосіб захисту тестової, табличної та графічної інформації / М.А. Заркевич. — № u200810199; заявл. 08.08.2008; опубл. 12.01.2009, Бюл. № 1, 2009 р. — 7 с.: іл.

11. *Технология* полиграфического метода защиты документов [Электронный ресурс]. — Режим доступа: www.securesoft.ru

12.*Заявка* U201008671 Україна. Спосіб захисту інформації від несанкціонованого розповсюдження без посилання на першоджерело / Д.М. Андрущенко, Л.М. Карпуков, Г.Л. Козіна. — Заявл. 12.07.2010.

13. *Коржик В.И.* Строгая аутентификация двоичных изображений без внесения искажений / В.И. Коржик, М.А. Зубарев // Проблемы информационной безопасности. Компьютерные системы. — ФГУП «Политехника». — Санкт-Петербург, 2008. — № 1. — С. 63–69.

Поступила в редакцию 13.04.2011