УДК 003.26

Е. А. Блинова
Белорусский государственный технологический университет

СТЕГАНОГРАФИЧЕСКИЙ МЕТОД НА ОСНОВЕ ИЗМЕНЕНИЯ МЕЖДУСТРОЧНОГО РАССТОЯНИЯ НЕОТОБРАЖАЕМЫХ СИМВОЛОВ СТРОК ЭЛЕКТРОННОГО ТЕКСТОВОГО ДОКУМЕНТА

Описывается модификация известного метода текстовой стеганографии, основанного на изменении междустрочного расстояния электронного документа (line-shift coding). С его помощью предлагается скрывать тайное сообщение в изменении высоты междустрочных интервалов. Сущность модификации метода состоит в том, чтобы использовать в качестве контейнера электронный документ и изменять междустрочное расстояние не всей строки или абзаца, а только неотображаемых символов (пробелов, знаков абзаца, табуляций и т. д.). Анализ абзаца средствами редактора электронных документов не показывает наличия смещения строк. Изменение начертания и/или размера шрифта не выявляет присутствия скрытого сообщения. Приводятся примеры отображения текстов с измененным междустрочным расстоянием для неотображаемых символов, анализируется скрытность встраиваемого сообщения. Метод может быть применен в различных версиях редакторов электронных документов Microsoft Word и Adobe InDesign. Оценка пропускной способности метода основывается на вероятностных характеристиках текстовых контейнеров для русского и английского языка. Существует возможность увеличения скрытности для внесения стеганографических меток. Предлагаются варианты использования стеганографического метода в электронных документах для внесения цифровых водяных знаков с целью защиты от несанкционированного копирования и распространения.

Ключевые слова: стеганография, электронный документ, изменение междустрочного расстояния.

E. A. Blinova
Belarusian State Technological University

STEGANOGRAPHIC METHOD BASED ON THE LINE-SHIFT CODING METHOD ON NON-DISPLAYED SYMBOLS OF THE ELECTRONIC TEXT DOCUMENT

The modification of a known method of text steganography based on the change line spacing of an electronic document (line-shift coding), which proposes to hide a secret message into a change of the height of the interline spacing is given. The modification of the method is to use an electronic document as a container and change the interline spacing not of the entire line or paragraph, but only of non-displayed characters (spaces, paragraph marks, tabs, etc.) An analysis by regular means does not show the presence of the line-shift coding. Changing the font size does not reveal the presence of a hidden message. Some examples of texts with modified interline spacing for non-displayed symbols were given, the secrecy of the embedded message is analyzed. The method can be applied in various versions of Microsoft Word and Adobe InDesign. The estimation capacity of the method is based on the probability characteristics of text containers for Russian and English. Some variations to increase secrecy for steganographic labels are given. The method can be used in electronic documents for embedding the digital watermarks to protect unauthorized copying and distribution.

Key words: steganography, electronic document, line-shift coding.

Введение. Стеганография – наука о способах передачи скрытого сообщения (стегосообщения) в открытой среде или о хранении скрытого сообщения в открыто хранящемся объекте, при котором не возникает подозрений о наличии скрытого сообщения.

Секретное сообщение при помощи стеганографического метода встраивается в не привлекающий внимания объект, который направляется адресату или размещается в общедоступной области [1]. Получатель сообщения, зная стеганографический ключ, расшифровывает сообщение. Следует отметить, что стеганографический ключ не шифрует данные, а скрывает место и порядок их встраивания. В литературе часто предполагается, что скрытое сообщение может быть предварительно зашифровано криптографическими методами для дополнительной защиты данных.

Для скрытия информации используются различные виды контейнеров (или файлов-контейнеров – в электронной стеганографии): текстовые документы, файлы HTML, изображения, звуковые и видеофайлы. Для

файлов-контейнеров разработано большое количество разнообразных методов, как общих, так и основанных на специфических свойствах контейнера [1, 2].

Основными направлениями применения стеганографических методов являются скрытая передача данных по открытому каналу, скрытое хранение данных в открытых данных, внесение одинаковых стеганографических меток во все копии документа (Watermaking), внесение различных стеганографических меток в каждую копию электронного документа (Digital Fingerprint).

Математическая модель стеганографической системы может быть представлена в виде соотношения:

$$S = Emb(C, M, K), \quad (1)$$

$$M = Ext(S, K), \quad (2)$$

где S, M – множество скрытых сообщений; $Emb()$ и $Ext()$ – функции встраивания скрытого сообщения в файл-контейнер и извлечения из файла-контейнера соответственно; C – множество всех контейнеров; K – множество стеганографических ключей.

В связи с широким распространением электронные документы формата Microsoft Office часто используются в качестве файлов-контейнеров. Для них применяются методы, которые используют наравне с классическими методами текстовой стеганографии и методы, свойственные контейнеру, такие как особое форматирование и смещение текста, размещение диакритических знаков, наличие истории редактирования и прочей служебной информации, что позволяет добиться увеличения скрытности и пропускной способности. Среди используемых классических методов текстовой стеганографии – метод изменения междустрочного расстояния, или line-shift coding, в применении к электронным текстовым документам. Методы произвольного интервала текстовой стеганографии основаны на манипулировании свободным местом в тексте, таких как высота междустрочного интервала, наличие дополнительных пробелов и различные виды смещений [1, 2].

Основная часть. Метод изменения междустрочного расстояния, или line-shift coding, успешно применялся с целью маркирования технической документации для предотвращения недобросовестного использования со стороны допущенных к ней специалистов [3, 4]. В его стандартной реализации предлагается скрывать стеганографическое сообщение в изменении высоты междустрочных интервалов. Для каждой копии документа выбирался свой набор междустрочных интервалов, что позволяло выявить источник несанкционированных копий. Пример текстового документа с различными междустрочными интервалами приведен на рис. 1 (строки с измененным междустрочным интервалом отмечены метками).

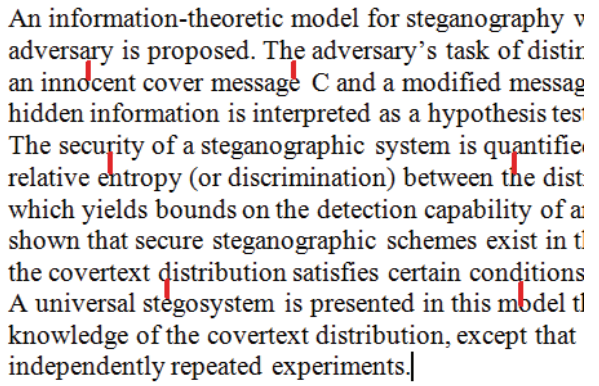

An information-theoretic model for steganography v
adversary is proposed. The adversary's task of distir
an innocent cover message C and a modified messag
hidden information is interpreted as a hypothesis tes
The security of a steganographic system is quantifie
relative entropy (or discrimination) between the dist
which yields bounds on the detection capability of a
shown that secure steganographic schemes exist in t
the covertext distribution satisfies certain conditions
A universal stegosystem is presented in this model t
knowledge of the covertext distribution, except that
independently repeated experiments.

Рис. 1. Фрагмент документа с различными междустрочными интервалами

Однако такой метод имеет несколько существенных недостатков: он обладает малой пропускной способностью и может быть выявлен как для электронного документа путем изменения параметров размера и начертания шрифта (рис. 2), так и для его напечатанной копии, которая может быть обработана как изображение (отсканирована и проанализирована на различные высоты междустрочных интервалов).

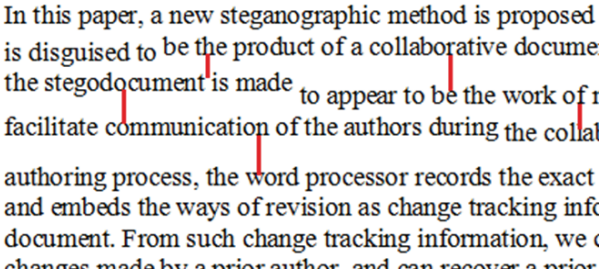

In this paper, a new steganographic method is proposed
is disguised to be the product of a collaborative docume
the stegodocument is made to appear to be the work of
facilitate communication of the authors during the collat
authoring process, the word processor records the exact
and embeds the ways of revision as change tracking info
document. From such change tracking information, we

Рис. 2. Документ с различными междустрочными интервалами после изменения размера шрифта

Предлагается следующая модификация стеганографического метода изменения междустрочного расстояния электронного документа – производить смещение не всей строки, а только неотображаемых символов (пробелов, табуляций, знаков переноса строки, неразрывных пробелов, абзацев и т. д.).

В качестве редактора электронных текстовых документов использовался Microsoft Word 2010, однако изменение высоты строки как для полной строки, так и для отдельных символов существует и в других редакторах. В Microsoft Word 2010 такое смещение выполняется при помощи команд *Шрифт/Интервал/Смещение*.

На рис. 3 изображен текст со смещением некоторых символов – пробелов и знаков абзаца.

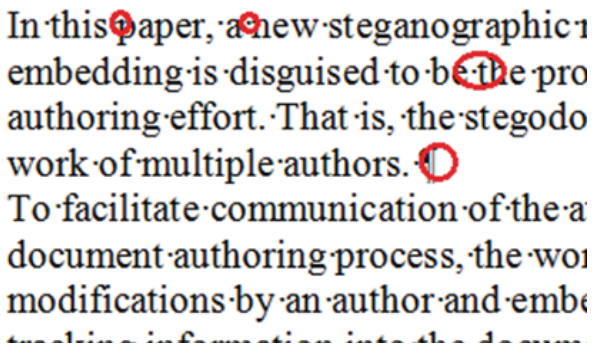


Рис. 3. Текст со смещенными неотображаемыми символами

Визуально незаметное смещение может производиться в диапазоне ±2 пункта, что позволяет встроить 5 бит скрытого сообщения на один неотображаемый символ. Однако находящиеся последовательно смещения на +2 пункта и –2 пункта могут быть заметны.

Поэтому предлагается вносить скрываемую информацию не в каждый неотображаемый символ, а через один (рис. 4). Это дает 5 бит скрытого сообщения на два неотображаемых символа. Можно ограничиться смещением в ±1 пункт, что дает 6 бит скрытого сообщения на два неотображаемых символа.

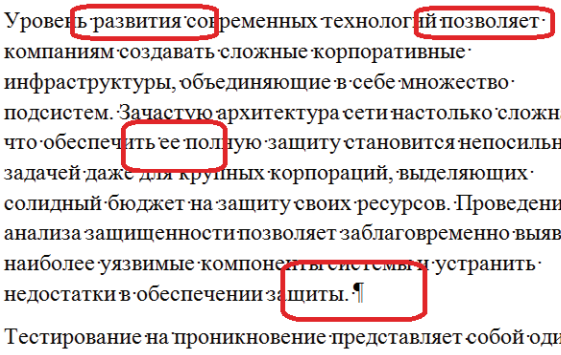


Рис. 4. Текст с последовательно смещенными на две позиции неотображаемыми символами

Отметим, что изменение начертания и размера шрифта не влияет на отображение электронного текста (рис. 5).

Можно предложить некоторые возможные модификации метода. Изменения смещения могут производиться по заранее согласованной маске, например, только по четным абзацам, только по каждой третьей строке, исключая первый и последний абзацы и т. д. Разумеется, пропускная способность будет уменьшена, но для небольших по объему скрытых меток это безразлично.

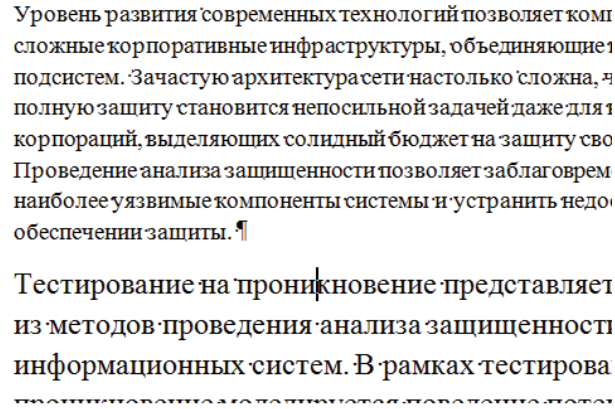


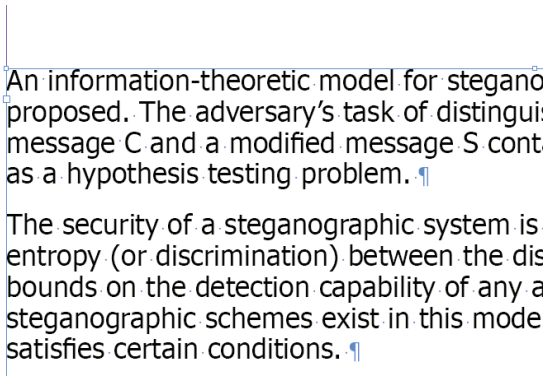
Рис. 5. Документ с измененными междустрочными интервалами неотображаемых символов после изменения размера шрифта

Также отметим, что стандартными средствами редактора различные высоты смещения символов текста не определяются в отличие от других свойств формата (размера, начертания и пр.).

Оценка пропускной способности данного стеганографического метода может быть основана на вероятностных характеристиках появления в текстах неотображаемых символов (пробелов, табуляций, знаков абзаца и пр.). На частоту встречаемости неотображаемых символов были проанализированы 100 русско- и англоязычных научных и художественных текстов. Для русского языка частота появления пробела в текстах составляет 0,163, для английского языка – 0,179. Частота встречаемости для других неотображаемых символов зависит от размера и особенностей файла-контейнера. Таким образом, пропускная способность предлагаемого метода намного выше пропускной способности классических методов текстовой стеганографии, оценка которых дана в [1].

При переносе текста между различными редакторами электронных текстовых документов смещение неотображаемых символов переносится только в некоторых редакторах электронных документов. Было протестировано внедрение скрытой информации в некоторые, наиболее часто применяющиеся, редакторы электронных документов: Microsoft Word (версии от 2000 до 2010), Adobe InDesign версии CS5 и ранее, Corel версии X6 и ранее. Установлено, что при переносе в Microsoft Word и Adobe InDesign наличие скрытой информации остается в тексте-контейнере, при вставке в Corel смещение не переносится, таким образом, скрытое сообщение теряется (рис. 6).

При переносе в Adobe Acrobat изменение междустрочного интервала неотображаемых символов, к сожалению, невозможно из-за особенностей экспорта в формат .pdf.


An information-theoretic model for stegano
proposed. The adversary's task of distingui
message C and a modified message S cont
as a hypothesis testing problem. ¶

The security of a steganographic system is
entropy (or discrimination) between the dis
bounds on the detection capability of any a
steganographic schemes exist in this mode
satisfies certain conditions. ¶

Рис. 6. Документ с различными междустрочными интервалами при вставке в Adobe InDesign

Заключение. Рассмотренная модификация стеганографического метода изменения междустрочного интервала неотображаемых символов может быть применена для внесения цифрового водяного знака в электронные документы с целью защиты авторского права на интеллектуальную собственность и подтверждения целостности документа. Также метод может быть использован для размещения различных скрытых стеганографических меток в каждую копию электронного документа, например, в файлы верстки электронных книг и журналов для выяснения канала несанкционированного копирования и распространения.

Литература

1. Грибунин В. Г. Цифровая стеганография. М.: Солон-Пресс, 2002. 272 с.

2. Shutko N., Blinova E. The use of aprosh and kerning in text steganography // New Electrical and Electronic Technologies and their Industrial Implementation: 9-th Conf., Zakopane, Poland, june 23–26, 2015. Zakopane, 2015. P. 77.

3. Electronic marking and identification techniques to discourage document copying / J. Brassil [et al.] // IEEE Journal on Selected Areas in Communications. 1995. Vol. 13, no. 8. P. 1495–1504.

4. Document Marking and Identification using Both Line and Word Shifting / S. H. Low [et al.]. Boston: Infocom, 1995. 8 p.

References

1. Gribunin V. G. *Tsifrovaya steganografiya* [Digital Steganography]. Moscow, Solon-Press Publ., 2002. 272 p.

2. Shutko N., Blinova E. The use of aprosh and kerning in text steganography. *New Electrical and Electronic Technologies and their Industrial Implementation: 9-th Conf.* Zakopane, 2015, p. 77.

3. Brassil J., Low S. H., Maxemchuk N. F. O'Gorman L. Electronic marking and identification techniques to discourage document copying. *IEEE Journal on Selected Areas in Communications*, 1995, vol. 13, no. 8, pp. 1495–1504.

4. Low S. H., Maxemchuk N. F., Brassil J. T., O'Gorman L. Document Marking and Identification using Both Line and Word Shifting. Boston, Infocom, 1995. 8 p.

Информация об авторе

Блинова Евгения Александровна – аспирант. Белорусский государственный технологический университет (220006, г. Минск, ул. Свердлова, 13а, Республика Беларусь). E-mail: evgenia.blinova@belstu.by

Information about the author

Blinova Evgeniya Aleksandrovna – PhD student. Belarusian State Technological University (13a, Sverdlova str., 220006, Minsk, Republic of Belarus). E-mail: evgenia.blinova@belstu.by

Поступила 11.03.2016