

УДК 003.26+347.78

Н. П. Шутько

Белорусский государственный технологический университет

**ЗАЩИТА И ПЕРЕДАЧА ТЕКСТОВОЙ ИНФОРМАЦИИ
НА ОСНОВЕ ИЗМЕНЕНИЯ КЕРНИНГА**

Описан и проанализирован новый метод текстовой стеганографии для скрытой передачи данных и охраны прав интеллектуальной собственности на текстовые документы (контейнеры). Метод основан на модификации одного из пространственно-геометрических параметров текста-контейнера – кернинга. Метод предусматривает изменение значения кернинга любых кернинговых пар символов в документе-контейнере. Модификация указанных параметров позволяет осаждать тайное сообщение (авторскую информацию) в процессе прямого стеганографического преобразования защищаемого текста-контейнера либо передавать эту информацию по стегаканалу. В первом случае решается задача защиты права интеллектуальной собственности, во втором – обеспечения повышенного уровня конфиденциальности передаваемых сообщений между абонентами. Учитывая специфическую особенность процессора MS Word, состоящую в том, что кернинг применяется к символам, размер (кегель) которых не ниже заданного специальной опцией в диапазоне от 1 до 1638 пт, и такая опция может применяться независимо к любым парам знаков и даже единичным знакам текста, предложены два варианта практической реализации метода. Метод характеризуется более высокой эффективностью в сравнении с другими методами текстовой стеганографии.

Ключевые слова: текстовая стеганография, авторское право, параметры шрифта, кернинг.

N. P. Shutko

Belarusian State Technological University

**PROTECTION AND TRANSFER OF TEXT INFORMATION
ON THE BASIS OF KERNING CHANGING**

A new method of text steganography for hidden data transmission and protection of intellectual property rights for text documents (containers) is described and analyzed. The method is based on the modification of one of the spatial-geometric parameters of the text-container – kerning. The method involves changing of the kerning value of any kerning pairs of characters in the document-container. Modification of these parameters allows to embed a secret message (author's information) during the process of direct steganographic transformation of the protected text-container or transmit this information by a stegochannel. In the first case, the problem of protection of intellectual property rights is solved, in the second one – ensuring of an increased level of confidentiality of transmitted messages between abonents. Considering the specific feature of the MS Word processor, which consists in the fact that kerning is applied to symbols whose size (font size) is not lower than specified by a special option in the range from 1 to 1638 pt, and such option can be applied independently to any pairs of characters and even single characters of the text, two variants of practical implementation of the method are proposed. The method is characterized by a higher efficiency in comparison with other methods of text steganography.

Key words: text steganography, copyright, font parameters, kerning.

Введение. Проблема защиты информации и защиты авторских прав на текстовые документы приобретает все большую актуальность. Основными объектами авторского права, относящимися к информационным технологиям, в том числе издательским, являются: бумажные и электронные версии различных текстовых или иных документов, базы данных, компьютерные программы. Это означает, что объекты интернета также относятся к интеллектуальной собственности. Среди множества предлагаемых методов защиты текстовой информации ни один не дает полной гарантии. В последние годы предпринимаются попытки создания инст-

рументов защиты контента от несанкционированного использования, особенно через глобальные информационные сети, на основе стеганографии [1–2]. При этом защита или передача информации производится путем ее тайного размещения (осаждения) в текстовый документ, называемый контейнером. Процесс осаждения подразумевает изменение одного или нескольких параметров контейнера. В случае решения задачи по защите авторских прав именно этот документ-контейнер и является объектом авторского права, т. е. объектом защиты. А размещенная в нем информация используется для доказательств авторского права.

В настоящей статье предлагается и анализируется метод текстовой стеганографии, основанный на осаждении/извлечении тайной информации в документ-контейнер путем модификации кернинга.

Сущность метода. Для понимания сущности метода кратко поясним специфику используемого здесь параметра шрифта.

В текстовых документах встречаются такие сочетания знаков, которые образуют визуальные «дыры» либо «сгущения». Например, в текстах на основе кириллицы – это такие сочетания: ГА, ТА, АТА, БТ и т. п., на основе латиницы – АУ, АВ, Т, ff, а на основе греческого алфавита – ΘΑ, ΔΟ, λκ. Под кернингом обычно понимается межсимвольное расстояние между отдельными парами символов (именно фактор парности отличает кернинг от апроша). Таким образом, технология кернинга, появившаяся в полиграфии после внедрения фотонабора (а затем и компьютерного набора), включает подбор межбуквенных интервалов для конкретных пар букв с целью улучшения внешнего вида и удобочитаемости текста. Такой избирательный подбор позволяет компенсировать неравномерности визуальной плотности текста, получаемой при использовании стандартных апрошей для каждой буквы.

На рис. 1 приведен пример применения кернинга для некоторых пар символов. Создается визуальное впечатление, что промежуток между «А» и «W» гораздо больше, чем между «W» и «Е» (в первом случае), хотя формально они одинаковы.



Рис. 1. Пример применения кернинга

В данном случае сочетание «AW» как раз и является кернинговой парой. После применения автоматического (во втором случае) и ручного (в третьем случае) кернинга визуальное восприятие текста улучшилось.

С появлением цифрового фотонабора стало возможным хранить такие критические сочетания знаков (кернинговые пары) для некоторого

условного шрифта, общее число которых мы обозначим N_k , в памяти компьютера с указанием величины (σ), на которую необходимо сдвинуть символы, чтобы визуально выровнять буквенные просветы.

Как правило, текстовые редакторы или процессоры содержат встроенные средства настройки кернинга, которые определяют стандартный межбуквенный интервал для того или иного шрифта. При этом σ устанавливается в соответствии со значениями из таблицы кернинговых пар, встроенной в вышеуказанный файл со шрифтом. Такая настройка позволяет выровнять шрифт и является стандартной. В некоторых шрифтах сейчас количество пар доходит до нескольких тысяч. Считается, что для качественного набора достаточно учитывать 300–700 пар [3].

Значение кернинга может быть как положительным (когда знаки раздвигаются, разреживаются, $\sigma > 0$), так и отрицательным (когда сдвигаются, $\sigma < 0$). Эта величина в программах верстки устанавливается в процентах от ширины символа пробела используемого шрифта.

Кернинг, как параметр, измеряется в тысячных долях круглой шпации (E_m) – единицы измерения, которая определяется относительно текущего размера шрифта и равна ширине символа «М». Например, для шрифта размером 6 пунктов одна круглая шпация равна 6 пунктам, а для шрифта размером 10 пунктов – 10 пунктам. Таким образом, размер кернинга σ строго пропорционален текущему размеру шрифта. Сдвиги букв относительно автоматически установленного межсимвольного расстояния (измеряемого, например, апрошем) можно производить с различным шагом: от 0,01 до 0,04 величины E_m , в зависимости от нужной точности.

На основе анализа общих принципов задания размера кернинга и управления этим размером, предлагаемый метод основывается на принудительном применении кернинга, не зависящем от установок параметров текста-контейнера, созданных средствами текстового процессора или иного специализированного текстового редактора. Здесь есть одна важная особенность. Изменять значения кернинга можно лишь в программах верстки (например, таких, как InDesign) или в программах, предназначенных для работы с векторной или растровой графикой (например, CorelDraw, Photoshop). В текстовом процессоре MS Word значения кернинга автоматически установлены в таблицах кернинговых пар каждого шрифта, доступа к которым нет.

Поскольку нас в большей степени интересует MS Word, как основная среда работы с текстовыми документами, и возможности мо-

дификации кернинга значительно ограничены, наш дальнейший анализ будет привязан именно к MS Word. Здесь возможности пользователя практически связаны лишь с указанием минимального размера шрифта, для которого можно применять кернинг. Это означает, что, если текст набран на основе шрифта размером, например, 14 пт, а мы в специальной опции (Главная/Шрифт) установили минимальный размер в 15 пт, при котором будет выполняться кернинг (см. поясняющую иллюстрацию на рис. 2), то для нашего текста эта процедура не будет выполнена.

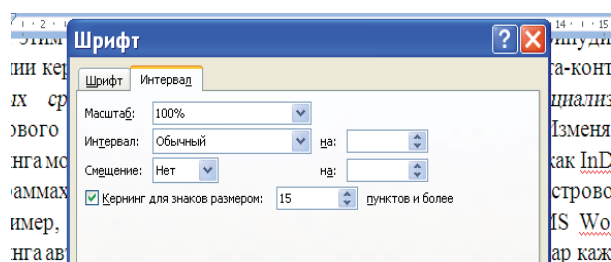


Рис. 2. Пример установки параметров для применения кернинга в MS Word

Таким образом, предлагаемый метод может быть интерпретирован так, что само значение кернинга мы программно не изменяем, а изменяем лишь размер символов, к которым будет применен кернинг в результате осаждения секретной информации.

Следует принять во внимание также следующую важную особенность. Текстовый процессор MS Word позволяет изменять параметры кернинга для знаков определенного размера. Размер символов, к которым может быть применен кернинг, должен быть в диапазоне между 1 и 1638 пт; в частном случае этот размер может быть указан в последней строке окна на рис. 2 (а сам размер шрифта можно устанавливать с точностью до 0,5 пт). Исходя из приведенных возможностей и ограничений, предлагается реализация стеганографического метода в двух вариантах.

Вариант 1. Предполагает осаждение в одной кернинговой паре одного бита сообщения M_i , переведенного в двоичный вид при соответствующем объеме N_{2c} : например, «0», если кернинг для этой пары не применяется, и «1» – если применяется. Идея основывается на том, что автоматическое установление кернинга приводит к изменению определенного межсимвольного расстояния на величину $\Delta\sigma$ для шрифта с установленным размером (кеглем) E_m (например, 14,5 пт; см. опцию на рис. 2) не будет применяться процессором по отношению к любым парам символов меньшего раз-

мера ($E_m - (\Delta E_m)$). Таким уменьшенным размером шрифта будет написан текст M_i (например, это будет размер 14 пт). Отклонение размера в 0,5 практически визуально не ЗАМЕТНО (средняя часть – «аметн» – последнего слова набрана размером 14,5 пт, основной текст – 14 пт; здесь $\Delta E_m = 0,5$ пт). Приведенный простой пример означает, что если мы устанавливаем в соответствии с рис. 2 кернинг для знаков размером 14,5 пт, то эта операция не будет распространяться на основной текст.

Отклонение размера $(E_m)_t$ некоторой t -й кернинговой пары символов на величину ΔE_m от установленной E_m означает, что в ней осажден «0» (кернинга нет) и «1» – в противном случае:

$$\Delta\sigma_t = 0; \text{ при } E_m - (E_m)_t = \Delta E_m, \quad (1)$$

$$\Delta\sigma_t \neq 0; \text{ при } E_m - (E_m)_t = 0, \quad (2)$$

где $t = 1, 2, \dots, Z_k$, Z_k – количество кернинговых пар в документе-контейнере.

Таким образом, для данного метода стеганографические ключи [4] определяют методы выбора очередности кернинговых пар символов текста-контейнера и метод применения либо неприменения кернинга для выбранной пары символов.

Легко заметить, что максимальный объем (бит) осажденного сообщения в анализируемом варианте не может превысить число выявленных кернинговых пар в тексте-контейнере Z_k .

Вариант 2. Отличается от предыдущего варианта методом кодирования осаждаемого сообщения. С учетом того, что, как отмечено выше, размер символов, к которым может быть применен кернинг, должен быть в диапазоне между 1 и 1638 пт, наша идея основывается на том, что сам факт применения или неприменения кернинга значения не имеет, как не имеет значения и размер шрифта документа. Имеет значение лишь указанный для произвольной (t -й) кернинговой пары текста-контейнера параметр $(E_m)_t$, который может выбираться из указанного диапазона.

Такой «индивидуальный подход» применен для примера к следующей последовательности символов «ГАТААТАБТ»: для первых двух символов $E_m = 11$ пт, для двух последующих $E_m = 72$ пт, для 5–7 символов $E_m = 1000$ пт, для двух последних $E_m = 1$ пт. Понятно, что собственно кернинг применен к двум первым и двум последним символам.

Принимая во внимание число 1638 (максимально возможное значение E_m), отмечаем, что достаточно просто в любой кернинговой паре осаждать до 10 бит сообщения M_i ; для этого нам достаточно иметь и $E_{m\max} = 1024$ пт. Таким образом, по определению данный вариант, с

точки зрения информативности, на порядок превышает предыдущий. Исходя из этого, значению $(E_m)_i = 1$ пт может соответствовать осаждаемая последовательность 0000000001, а $(E_m)_i = 15$ пт – 0000001111 и т. д.

Как видно из выполненного анализа, предлагаемый метод по уровню сложности алгоритма его реализации сопоставим с известными алгоритмами текстовой стеганографии, например, Line-Shift Coding или Word-Shift Coding.

Если выполняется условие $Z_k \geq N_c$, где N_c – число символов в исходном сообщении M_i , и, если исходное сообщение не подвергается дополнительным преобразованиям, то целесообразно принять $E_{m\max} = 2^8 = 256$ пт, что даст возможность осаждать в любой кернинговой паре по одному символу (в кодах ASCII) сообщения M_i .

В других случаях нужно предусматривать возможность удлинения последней осаждаемой последовательности до стандартного размера (например, 10 бит).

Заключение. Предложенный и проанализированный на примерах метод тайной передачи информации в тексте-контейре или защиты авторского права на этот текстовый документ основан на реализации текстовой стеганографии путем изменения такого пространственно-геометрического параметра текста-контейнера, как кернинг. По уровню сложности сопоставим с методами на основе модификации междустрочного интервала (Line-Shift Coding) и пробельного расстояния между словами (Word-Shift Coding), если принять, что в последних может применяться псевдослучайный выбор модифицируемых элементов.

Вместе с тем, предлагаемый метод характеризуется большей эффективностью, поскольку число символов и их кернинговых пар в текстах всегда значительно превышает число строк или число слов. Необходимо также отметить, что стегознаками являются все символы документа, в том числе специальные знаки и символы.

Литература

1. Шутько Н. П. Защита авторских прав на электронные текстовые документы методами стеганографии // Труды БГТУ. 2013. № 6: Физ.-мат. науки и информатика. С. 131–134.
2. Шутько Н. П., Романенко Д. М., Урбанович П. П. Математическая модель системы текстовой стеганографии на основе модификации пространственных и цветовых параметров символов текста // Труды БГТУ. 2015. № 6: Физ.-мат. науки и информатика. С. 152–156.
3. Шрифты. Разработка и использование / Барышников Г. М. [и др.]. М.: ЭКОМ, 1997. 288 с.
4. Шутько Н. П. Алгоритмы реализации методов текстовой стеганографии на основе модификации пространственно-геометрических и цветовых параметров текста // Труды БГТУ. 2016. № 6: Физ.-мат. науки и информатика. С. 160–165.

References

1. Shutko N. P. Copyright protection on the electronic text documents by methods of steganography. *Trudy BGTU* [Proceedings of BSTU], 2013, no. 6: Physical-mathematical sciences and informatics, pp. 131–134 (In Russian).
2. Shutko N. P., Romanenko D. M., Urbanovich P. P. Mathematical model of the text steganography on the base of modifying the spatial and color settings of text characters. *Trudy BGTU* [Proceedings of BSTU], 2015, no. 6: Physical-mathematical sciences and informatics, pp. 152–156 (In Russian).
3. Baryshnikov G. M., Bizyaev A. U., Efimov V. V., Moiseev A. A. *Shrifty. Razrabotka i ispolzovanie* [Fonts. Development and using]. Moscow, ECOM Publ., 1997. 288 p.
4. Shutko N. P. The algorithms of realization of text steganography methods based on the modification of the geometric and color text parameters. *Trudy BGTU* [Proceedings of BSTU], 2016, no. 6: Physical-mathematical sciences and informatics, pp. 160–165 (In Russian).

Информация об авторе

Шутько Надежда Павловна – кандидат технических наук, старший преподаватель кафедры информационных систем и технологий. Белорусский государственный технологический университет (220006, г. Минск, ул. Свердлова, 13а, Республика Беларусь). E-mail: shutko_bstu@mail.ru

Information about the author

Shutko Nadezhda Pavlovna – PhD (Engineering), Senior Lecturer, the Department of Information Systems and Technologies. Belarusian State Technological University (13a, Sverdlova str., 220006, Minsk, Republic of Belarus). E-mail: shutko_bstu@mail.ru

Поступила 22.04.2017