

СТЕГАНОГРАФИЧЕСКИЙ МЕТОД НА ОСНОВЕ ИЗМЕНЕНИЯ МЕТОДА ПОДЧЕРКИВАНИЯ СИМВОЛОВ

Савельева М. Г.

*Белорусский государственный технологический университет,
Минск, Беларусь, margarita.saveleva.98@mail.ru*

Стеганография – это наука о способах передачи (хранения) сокрытой информации, при которых скрытый канал организуется на базе и внутри открытого канала с применением особенностей восприятия информации.

Методы текстовой стеганографии можно разделить на три основные категории: основанные на формате, или синтаксические методы, лингвистические, или основанные на обработке естественного языка, методы, а также основанные на случайной и статистической генерации [1].

Среди современных методов компьютерной текстовой стеганографии для электронных документов можно выделить следующие: метод непечатаемых символов (заключается во встраивании секретного сообщения в цвет непечатаемых символов в формате RGB), метод Similar English Font Types (заключается в выборе трех различных похожих шрифтов с последующим кодированием всех букв и пробела как тройки заглавных букв), метод подчеркивания символов (заключается в добавлении невидимых стилей подчеркивания к символам), метод масштабирования символов (заключается в скрытии бит за счет увеличения/уменьшения масштаба текстовых символов на 1 %), метод границ предложения (заключается в использовании возможности добавления левых и правых границ к предложению с цветами, которые незаметны пользователю ввиду особенностей строения человеческого глаза), метод границ абзаца (заключается в добавлении левой и правой границы к абзацу и задания им цветов, неотличимых для человека от фонового цвета документа), метод отслеживания изменений (заключается в использовании средств отслеживания изменений документов MS Word при совместном написании).

В стандартной реализации метода подчеркивания символов авторами предлагается скрывать стеганографическое сообщение в электронном тестовом документе с помощью добавления

подчеркивания определенным символам. Существует 16 различных стилей подчеркивания, которые могут содержать 4 бита, и 2^{24} различных цвета подчеркивания. Поскольку нужно, чтобы подчеркивание оставалось незамеченным пользователем, используется 16 вариантов белого цвета с возможностью переноса 4 бит. Таким образом, можно скрыть 8 бит на символ. Некоторые символы, такие как g, j, p, q и u, имеют заметные изменения во внешнем виде, когда к ним применяется подчеркивание. Из-за этого группа этих символов была исключена из скрытия данных [2].

Предлагается следующая модификация стеганографического метода добавления подчеркивания символов электронного документа – воспользоваться тем, что некоторые символы с нижними вертикальными отростками имеют изменения во внешнем виде при подчеркивании для возможности использования метода не только для электронных документов. В кириллице строчными буквами с нижними вертикальными отростками относятся буквы p, y, ф. При определенных видах подчеркивания цветом в цвет страницы с текстом, подчеркивание становится сложноразличимым для пользователя.

Для демонстрации модификации стеганографического метода используем следующую математическую модель. Зафиксируем множество возможных сообщений $M = \{M_1, M_2, \dots, M_m\}$, множество возможных контейнеров $L = \{L_1, L_2, \dots, L_l\}$, и множество возможных заполненных контейнеров (стеганограмм) $E = \{E_1, E_2, \dots, E_n\}$. Зафиксируем множество отображений:

$$\phi = \{\phi_1, \phi_2, \dots, \phi_k\}, \quad (1)$$

где

$$\phi_i: (M, L) \rightarrow E, i = 1, 2, \dots, k. \quad (2)$$

Определим обратное отображение:

$$\phi_i^{-1}: E \rightarrow (M, L), \quad (3)$$

которое каждому элементу множества E ставит в соответствие элемент множества M и элемент множества L .

Зафиксируем множество ключей $K = \{K_1, K_2, \dots, K_k\}$ так, что для всех $i = 1, 2, \dots, k$ отображение $\phi_i \in \phi$ однозначно задается ключом K_i . Каждое конкретное отображение ϕ_i из множества ϕ соответствует способу встраивания сообщения из множества M в контейнер из множества L при помощи конкретного ключа K_i [3].

В качестве редактора электронных текстовых документов использовался Microsoft Word 2010, однако подчеркивание символов существует и в множестве различных редакторах. Для демонстрации выбрана русскоязычная панграмма, содержащая все буквы русского алфавита и используемая в стандартных средствах Microsoft Windows для просмотра шрифтов.

Microsoft Word поддерживает 16 различных типов подчеркивания. Из них при подчеркивании цветом в цвет страницы с текстом являются трудноразличимыми подчеркивание сплошной волнистой линией и подчеркивание двойной сплошной волнистой линией.

На рисунке 1 представлено использование подчеркивание сплошной волнистой линией и подчеркивание двойной сплошной волнистой линией цвета контрастного к цвету страницы при 100 % масштабе. На рисунке 2 представлено тот же самый пример при 200 % масштабе.

На рисунке 3 представлено использование подчеркивание сплошной волнистой линией и подчеркивание двойной сплошной волнистой линией в цвет страницы при 100 % масштабе. На рисунке 4 представлено тот же самый пример при 200 % масштабе.

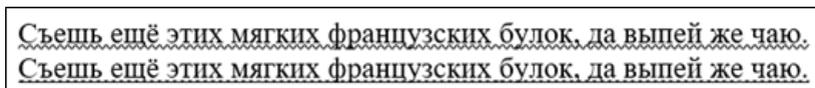


Рисунок 1 – Пример контрастного подчеркивания при 100 % масштабе

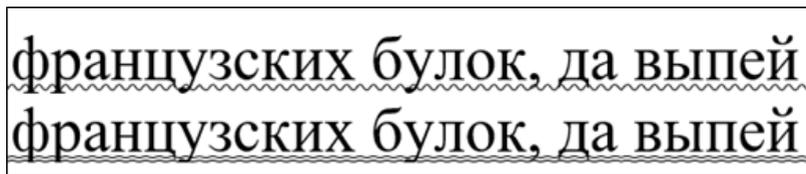


Рисунок 2 – Пример контрастного подчеркивания при 200 % масштабе

Съешь ещё этих мягких французских булок, да выпей же чаю.
Съешь ещё этих мягких французских булок, да выпей же чаю.

Рисунок 3 – Пример подчеркивания в цвет страницы при 100 % масштабе

французских булок, да выпей
французских булок, да выпей

Рисунок 4 – Пример подчеркивания в цвет страницы при 200 % масштабе

Исходя из рассмотренных примеров можно сказать что для пользователя, использующего печатный документ без возможности изменения масштаба, встраивание информации будет трудноразлично. Кроме того, если будет обнаружено какое-то различие в отображении одной и той же буквы в различных местах документа, то можно принять это за погрешности печати.

Количество встраиваемой информации регулируется выбором числа положений подчеркивания. Максимально возможно встроить 3 бит скрываемого сообщения на один символ при использовании трех положений подчеркивания (отсутствие подчеркивания, подчеркивание сплошной волнистой линией, подчеркивание двойной сплошной волнистой линией). Такая разнородность может стать более очевидной для пользователя, поэтому лучше использовать два положения подчеркивания. В этом случае возможно встроить 2 бит на один символ.

Оценка пропускной способности предложенного метода может быть основана на вероятностных характеристиках появления в текстах букв с нижними вертикальными отростками, в русском алфавите это р, у, ф. Согласно НКРЯ (национальный корпус русского языка), частотность встречаемости буквы р – 0,04746 (4,746 %), буквы у – 0,02615 (2,615 %), ф – 0,00267 (0,267 %). Исходя из этого пропускная способность при использовании всех предложенных букв будет равняться 0,07628 (7,628 %) [4].

Стоит отметить, что на встроенную информацию не оказывает влияние редактирование документа. Количество встроенной информации останется прежним при изменении размера, цвета,

шрифта текста, изменении настроек абзаца и т. д. Исключением является лишь сам инструмент подчеркивания.

Ключевым отличием предложенной модификации от исходного метода, является возможность применения метода не только для электронных документов, а также для печатных.

Предложенное изменение стенографического метода подчеркивания символов позволяет применять его не только для защиты интеллектуальной собственности электронных документов, но и для печатных. Кроме использования метода в текстовых документах, можно также применять его для электронного текста в различном виде, не только в виде документа.

СПИСОК ЛИТЕРАТУРЫ

1. Bennett, K.: Linguistic steganography: Survey, analysis, and robustness concerns for hiding information in text. cERIAS Tech Report 2004-13 (2004).

2. Stojanov, I., Mileva, A., Stojanovic, I.: A New Property Coding in Text Steganography of Microsoft Word Documents. In *Securware 2014: The Eighth International Conference on Emerging Security Information, Systems and Technologies*, pp. 25–30 (2014).

3. Кузнецов, А. А. Математическая модель и структурная схема стеганографической системы / А. А. Кузнецов, А. А. Смирнов, Е. В. Мелешко // Збірник наукових праць Кіровоградського національного технічного університету. Техніка в сільськогосподарському виробництві, галузеве машинобудування, автоматизація. – Кіровоград: КНТУ, 2012. – Вип. 25, ч. 1. – С. 273–281.

4. Национальный корпус русского языка. <http://www.ruscorpora.ru/new/corpora-intro.html>.