УДК 003.26+004.56

МЕТОД НА ОСНОВЕ АТРИБУТОВ ТЕКСТА В МНОГОКЛЮЧЕВОЙ СТЕГАНОГРАФИЧЕСКОЙ СИСТЕМЕ

Берников В.О.
БГТУ, г. Минск, Республика Беларусь; vladbernikovronaldo@gmail.com

Реферат. В докладе кратко описываются разработанный стеганографический метод на основе атрибутов текста и программное средство для осаждения и обратного извлечения стегосообщений на базе многоключевой модели стеганографической системы. Многоключевая модель стеганосистемы предполагает совместное использование методов стеганографии, криптографии, помехоустойчивого кодирования или других преобразований для повышения стеганографической стойкости в целом. Программное средство может использоваться как в научных исследованиях, так и в учебном процессе.

В настоящее время проблема обеспечения должной защиты авторского права на электронные документы-контейнеры текстового типа приобретает все большую актуальность в связи с развитием информационных технологий [1]. Одно из средств для нейтрализации данной проблемы является применение новых эффективных методов стеганографии. При этом должен обеспечиваться требуемый уровень защищенности стеганосистемы перед несанкционированным использованием осажденной в контейнер информации. Использование многоключевой модели информационной системы обеспечивает эффективное решение данной проблемы.

Формально многоключевую стеганосистему можно представить следующим образом:

$$X = \{M, C, K, S, f, \mu\} \quad (1),$$

где: M – множество сообщений, C – множество контейнеров, K – множество ключей, S – множество секретных сообщений, f – функция осаждения информации, μ – функция извлечения информации [2, 5].

Для анализа эффективности разработанного стеганографического метода написано специализированное программное средство, которое использует описанную выше модель информационной системы. Контейнерами для внедрения стегосообщений являются электронные документы формата *.doc или *.docx.

В данном программном средстве представлены возможности дополнительного выбора преобразования информации, выбора документа, в котором осаждается секретное сообщение, выбор формата сохранения полученного документа, а также сокрытие и извлечение секретного сообщения. Для разбора электронных документов используется библиотека Aspose.Words, которая содержит необходимые методы для работы с документами. Программное средство написано на языке C# в среде Visual Studio. В качестве графического интерфейса используется интерфейс программирования приложений – Windows Presentation Foundation.

Структуру программного средства можно представить следующим образом:

– алгоритм преобразования символов сообщения в Unicode;

– алгоритм преобразования Unicode в битовую последовательность;

– алгоритм сокрытия битовой последовательности в стеганоконтейнере текстового типа на основе атрибутов;

– алгоритм извлечения скрытой битовой последовательности из стеганоконтейнера текстового типа;

– алгоритм получения скрытого сообщения из битовой последовательности.

Атрибуты представляют собой пару «ключ–значение». Количество этих атрибутов зависит от того, насколько большой объем информации требуется скрыть. Каждое значение атрибута имеет определенную длину, которая регулируется размером осаждаемой информации. В качестве ключа хранится название и номер блока, а в значении – сами символы, входящие в этот блок.

Стеганографический метод на основе атрибутов текста работает следующим образом: сначала нужно выбрать документ (стеганоконтейнер), в котором осаждается информация. Далее используются алгоритмы преобразования исходного формата сообщения в формат, допустимый для реализации сокрытия. Предполагается дополнительное шифрование стегосообщений при использовании симметричных (AES и TwoFish) и асимметричных (RSA и Эль-Гамаль) криптосистем соответственно (первый ключ рассматриваемой модели стеганографической системы). Также может быть использовано кодирование секретной информации (второй ключ) на основе классического и модифицированного кодов Хемминга, а также циклических кодов). Дополнительно, для проверки того, что информация успешно извлекается, может применяться хеширование сообщений (третий ключ) при использовании алгоритмов SHA512 и MD5. По результатам преобразования, получается битовая последовательность, состоящая из нулей и единиц. С помощью специального алгоритма последовательно выбираются псевдослучайные символы в тексте, которые зависят от общего количества абзацев, а также длины сообщения. Стоит отметить, что может быть использован и четвертый ключ для псевдарандомизации секретных бит по всему электронному документу [3,4]. Демонстрация работы данного метода осаждения секретной информации представлена на рисунке 1.

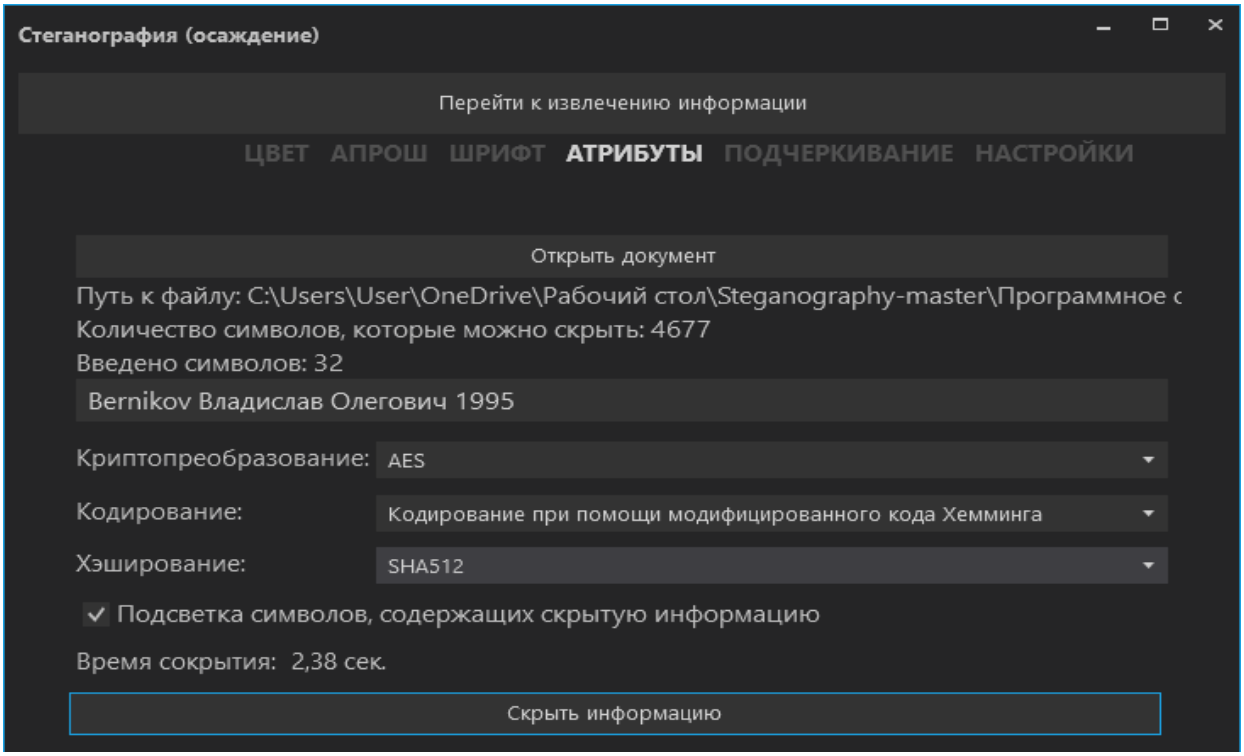

Рисунок 1 – Диалоговое окно для осаждения информации на основе атрибутов текста

Выбирается документ, куда будет помещена секретная информация. Производится автоматический подсчет символов, которые можно скрыть в выбранном электронном документе. Непосредственно вводится само секретное сообщение, а также выбираются соответствующие ключи многоключевой модели информационной системы. В нашем примере выбрано криптографическое преобразование секретной информации на основе алгоритма AES, кодирование на основе модифицированного кода Хемминга, а также хеширование стегосообщений на основе алгоритма SHA512 для проверки целостности внедряемой информации. Выбрана подсветка символов стегосообщения для наглядности сокрытия секретных бит информации в электронном документе на основе разработанного стеганографического метода.

Далее стоит убедиться, что секретная информация действительно внедрена в электронный документ. Продемонстрируем стеганоконтейнер после осаждения секретной информации (рисунок 2).

Рисунок 1.1 – Общая классификация методов компьютерной стеганографии

Компьютерная стеганография базируется на двух принципах. Первый заключается в том, что если файлы, содержащие оцифрованное изображение или звук, могут быть до некоторой степени видоизменены без потери функциональности, то текстовые документы, коды программ или базы данных требуют абсолютной точности при обратных преобразованиях. Это обстоятельство чрезвычайно важно, если, например, текстовый документ-контейнер с осажденной информацией претерпевает конвертацию на основе иного стиля (шрифта, кегля и т.п.) или при его архивации. Понятно, что сам документ при этом не должен измениться.

Все многообразие методов текстовой стеганографии подразделяется на синтаксические, которые не затрагивают семантику текстового сообщения, и

лингвистические методы, которые основаны на эквивалентной трансформации текстовых файлов, сохраняющей смысловое содержание текста, его семантику.

1.1 Синтаксические методы текстовой стеганографии

Рисунок 2 – Стеганоконтейнер после осаждения на основе атрибутов текста

Как видно из рисунка, секретная информация успешно осаждена псевдарандомизированным образом по всему электронному документу (отдельный ключ многоключевой модели информационной системы). Розовый цвет показывает, что в данных символах текста хранится стегосообщение.

Далее продемонстрируем процесс извлечения стегосообщения из электронного документа (рисунок 3). Сначала выбирается документ, в который уже предварительно спрятана информация. Выбираются те же ключи рассматриваемой модели системы, которые были выбраны при осаждении стегосообщения. Предварительно открываются файлы с тайным ключом и хешем, так как использовались соответствующие алгоритмы шифрования и хеширования для нашей секретной информации.

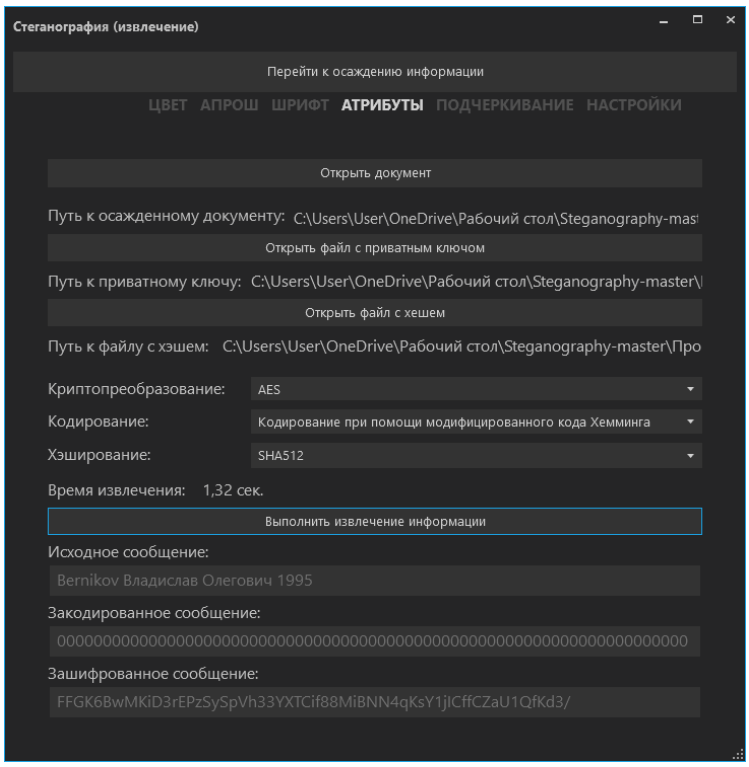

Рисунок 3 – Диалоговое окно для извлечения информации на основе атрибутов текста

Как видно из рисунка, стегосообщение успешно извлечено из электронного

документа. Далее откроем временный *xml*-файл со скрытым секретным сообщением в атрибутах текста (рисунок 4).

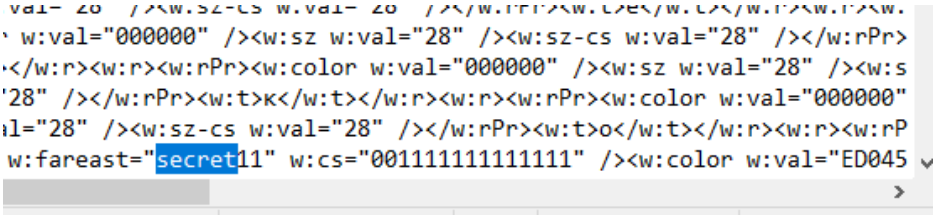

Рисунок 4 – Фрагмент *xml*-файла с осажденной информацией

Как видно из рисунка, информация действительно спрятана в атрибутах текста по соответствующему ключу и с определенным значением (в данном случае это часть закодированного стегосообщения). Стоит отметить, что если стегосообщение только зашифровано, но не закодировано дополнительно, то в атрибутах текста *xml*-файла оно будет храниться в зашифрованном виде, и соответственно, если не выбран ни один ключ рассматриваемой модели информационной системы – то в открытом виде. Инструменты Microsoft Word-а не позволяют обнаружить факт наличия скрытого сообщения, поэтому данный метод можно считать эффективным.

Описанное программное средство реализовано на основе модели информационной системы, которая подразумевает применение практически неограниченного числа ключей. Представлен процесс внедрения и извлечения стегосообщения при использовании контейнера текстового типа формата DOCX и DOC на основе разработанного стеганографического метода при использовании атрибутов текста. Разработанное средство используется также в учебном процессе при изучении студентами дисциплин «Защита информации и надежность информационных систем» и «Криптографические методы защиты информации».

Список литературы:

1. Урбанович, П.П. Защита информации методами криптографии, стеганографии и обфускации/ П.П. Урбанович. – Минск: БГТУ, 2016. – 220 с.

2. Pavel Urbanovich, Nadzeya Shutko. Theoretical Model of a Multi-Key Steganography System, in: Recent Developments in Mathematics and Informatics, Contemporary Mathematics and Computer Science Vol. 2, Ed. A. Zapała. – Wydawnictwo KUL, Lublin, 2016, Part II, Chapter 11. – P. 181-202.

3. Берников, В.О. Разработка стеганографических методов на основе многоключевой модели информационной системы/ В.О. Берников // Новые математические методы и компьютерные технологии в проектировании, производстве и научных исследованиях. – Гомель: ГГУ им. Ф. Скорины. – 2018. – С. 192-193.

4. Берников, В.О. Анализ стеганографической стойкости текстового документа-контейнера в многоключевой стеганосистеме // 69-я НТК студентов и магистрантов: сб. науч. работ: в 4-х ч. 17-22 апреля 2018 г. – Минск: БГТУ, 2018. – Ч. 4. – С.14-17.

5. Берников, В. О. Математическое моделирование стеганографической стойкости многоключевой системы / В. О. Берников, П. П. Урбанович // Информационные технологии : материалы 83-й научно-технической конференции профессорско-преподавательского состава, научных сотрудников и аспирантов (с международным участием), Минск, 4-15 февраля 2019 г. / отв. за изд. И. В. Войтов; УО БГТУ. – Минск : БГТУ, 2019. – С. 31-33.