МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение

высшего образования

«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»

Кафедра теоретических основ компьютерной безопасности и криптографии

Стеганография в медиа - файлах

АВТОРЕФЕРАТ

дипломной работы

студента 6 курса 631 группы

специальности 10.05.01 Компьютерная безопасность

факультета компьютерных наук и информационных технологий

Шувалкина Ярослава Павловича

Научный руководитель

д. ф.-м. н., доцент _____ М. Б. Абросимов

22.01.2022 г.

Заведующий кафедрой

д. ф.-м. н., доцент _____ М. Б. Абросимов

22.01.2022 г.

Саратов 2022

ВВЕДЕНИЕ

С появлением современных технологий, методов обработки информации, появились также и угрозы, связанные с потерей и утратой информации. Поэтому защита информации, усовершенствование методов её передачи является важнейшим видом деятельности не только в организациях, но и во всем государстве.

Существуют различные методы тайной передачи информации. Стеганография является одним из них. Названия происходит от греческих слов «στεγανός», что обозначает слово тайна, и «γράφω», что обозначает слово «запись». Стеганография является тайным способом хранения или передачи сообщения. Всего выделяется несколько видов методов стеганографии: классическая, компьютерная и цифровая.

Целями данной работы являлись рассмотрение основных методов стеганографии, а также разработка приложения, реализующего стеганографию в каждом из видов медиафайлов: изображнии, аудио и видео. Данное приложение будет поддерживать внедрение текстовой информации в изображения формата PNG, аудиофайлы формата WAV, видеофайлы формата AVI. Также приложение будет поддерживать внедрение файлов в видеофайл. Кроме того, приложение будет использовать библиотеку FFMPEG в виде её обёртки для языка программирования C#, для извлечения и работы с аудио дорожками видеофайлов, а также библиотеку Accord для чтения и записи видеофайлов.

Дипломная работа состоит из введения, 6 разделов, заключения, списка использованных источников и 1 приложения. Общий объем работы – 58 страницы, из них 40 страниц – основное содержание, включая 27 рисунков и одну таблицу, список использованных источников из 20 наименований.

КРАТКОЕ СОДЕРЖАНИЕ

В первом разделе описываются основные понятия, связанные с классической стеганографией, а также приводятся примеры.

Под классической стеганографией понимают все «некомпьютерные» методы стеганографии, например, симпатические чернила или микрофотоснимки.

Древнегреческий историк Геродот в своих трудах описывал еще один метод сокрытия информации: на обритую голову раба записывалось сообщение, а после того, как у него отрастали волосы, его отправляли в место назначения, где получатель вновь сбривал его волосы и считывал доставленное сообщение.

Примером применения методов классической стеганографии может служить использование невидимых чернил. Текст, который написан при помощи таких чернил проявляется только при определённых условиях, таких как нагревание или определенная химическая реакция.

Во втором разделе описывается общая модель стеганографии, приводятся основные термины. Выделяются главные требования для стеганосистемы.

Впервые модель стеганографического формата из-за запретов на публикации была сформулирована следующим образом. Двое заключенных, Алиса и Боб, находятся в разных камерах. Они могут обмениваться посланиями. Однако их переписка проходит через тюремного надзирателя Уэнди. Заключенные должны разработать план побега, обмениваясь, на первый взгляд, безобидными картинками или текстами. Их сообщения не должны вызывать подозрения у Уэнди. Тайные сообщения, которые передают друг другу заключенные, называются встроенными сообщениями. Предполагается, что Алиса предварительно шифрует сообщение, прежде чем попытаться передать его Бобу. Под скрытым сообщением в таком случае понимается криптограмма. Контейнер – информация различного вида, в которую тайно

внедряется сообщение. Различают аудиоконтейнеры, текстовые контейнеры, видеоконтейнеры и так далее.

Секретный ключ – некоторая общая информация, имеющаяся у заключенных, которую они держат в секрете.

Наличие у Алисы и Боба общего секретного ключа не является необходимым требованием для создания стеганографического канала Совокупность тех средств, которые Алиса и Боб используют для создания стеганографического канала, называется стеганографической системой или просто стеганосистемой.

Ниже приведены требования, которым должна удовлетворять стеганосистема:

1. Свойства контейнера должны быть модифицированы, для того, чтобы изменения невозможно было выявить при визуальном контроле. Это требование определяет качество степени скрытности внедрения сообщения: оно никаким образом не должно привлечь внимание атакующего для беспрепятственного прохождения через систему мониторинга ресурсов.

2. Сообщение должно быть устойчиво к любым искажениям. В процессе передачи контейнер может претерпевать различные изменения, например, увеличение или уменьшение размера, конвертирование в другой формат и так далее.

3. Ради сохранения целостности встраиваемого сообщения необходимо использование кода, исправляющего ошибки.

4. Встраиваемое сообщение должно быть продублировано, для повышения надежности в случае повреждения исходного сообщения.

В третьем разделе описывается компьютерная стеганография, связанные с нею понятия, рассматриваются основные принципы.

Главной целью компьютерной стеганографии является сокрытие сообщения внутри файла-контейнера. Контейнер не должен терять своих

функций, операция должна оставаться незамеченный, обнаружение сокрытой информации должно быть предельно сложным.

Ниже представлены основные требования компьютерной стеганографии:

1. При внедрении секретного сообщения должна сохраняться целостность и аутентичность файла.

2. Допускается, что противнику известны возможные стеганографические методы.

3. Безопасность этих методов основана на сохранении стеганографическим преобразованием основных свойств файла, которые при внедрении в него сообщения и некоторого секретного ключа, всё равно позволяют открыто передавать его.

4. В случае, когда факт внедрения сообщения становится известным противнику, тогда извлечение данного сообщения должно быть сложной вычислительной задачей.

В четвертом разделе рассматривается цифровая стеганография, основные определения, связанные с ней, основные направления.

Цифровая стеганография включает в себя следующие направления:

1. Встраивание информации для скрытой передачи.

2. Встраивание цифровых водяных знаков (ЦВЗ).

3. Встраивание идентификационных номеров.

4. Встраивание заголовков.

Из-за быстрого развития технологий мультимедиа приоритет вопроса защиты интеллектуальной собственности является крайне высоким. Поэтому существует ряд различных мер, направленных на предотвращение посягательств на данный вид собственности, несанкционированного копирования и использования. Одним из способов является встраивание в объект специальных меток, называющихся цифровыми водяными знаками. Цифровые водяные знаки, сокращённо ЦВЗ, являются одним из самых эффективных средств защиты информации.

Кроме того, в этом разделе описывается встраивание информации в контейнер при помощи метода наименее значащих битов.

Младший значащий бит данных в изображении или аудио содержит наименьшее количество информации. По факту он является шумом и его замена не приведет к заметным изменениям, что позволяет использовать его для скрытого внедрения информации. Например, в WAV-файле в области данных младший бит каждого байта аудиофайла можно заменить один бит сообщения, что будет незаметно для человеческого слуха. В изображении можно заменять наименее значащие биты каждого пикселя одного из цветовых каналов, что также будет незаметно для человеческого зрения. Если изменить два младших бита, что является практически незаметным изменением, то можно внедрить для передачи вдвое больший объем данных.

В пятом разделе описываются характеристики форматов данных, используемых в качестве контейнера в практической части работы. Рассматриваются форматы изображений PNG, аудиофайлов WAV, видеофайлов AVI.

Файл PNG (с англ. Portable Network Graphic) относится к растровым изображениям. Данный формат содержит определенную палитру цветов, которые применяются в рисунке. Подобный графический формат довольно часто применяют в сети для вставки изображений на веб страницы.

PNG обладает лучшим уровнем сжатия без потери информации, однако далеко не все реализации полностью используют доступные возможности, а те, что используют, зачастую неправильно их применяют.

В формате PNG поддерживаются следующие основные типы изображения: truecolor, grayscale, а также 8-битное индексированное изображение на основе палитры. Отличительной особенностью является смешивание различных типов изображения в одном PNG файле. Однако нужно учитывать плотность сжатия. При сохранении 8-битного изображения в качестве 24-битного truecolor невозможно получить файл маленького размера.

Многие изображения, предназначенные для веб сайтов, состоят из 256 цветов, иногда даже менее. Таким образом, изменение оригинала с добавлением более 256 цветов также приведет к увеличению размера файла.

Формат аудиофайлов WAV представляет собой две области: область заголовка файла и область данных. В области заголовка находится вся информация о характеристиках файла: его аудио формат, количество каналов, частота дискретезации, размер области данных и т.д. Малейшое изменение байтов области заголовка приводят к повреждению всего файла. В области данных хранятся сами непосредственно WAV данные, отвечающие за звучание аудио. Совокупностью короткого промежутка времени и амплитуды называют сэмпл. Амплитуда выражается числом, которое занимает в файле 8, 16, 24, 32 бита, однако теоретически может выражаться и большим числом. Так как 1 байт состоит из 8 бит, то амплитуда занимает от одного до четырёх байт. Таким образом, чем больше число занимает места в файле, тем больше точность амплитуды и тем шире возможный диапазон значений для этого числа.

Для WAV файлов используются следующие виды разрядностей:

1. 1 байт или 8 бит,
2. 2 байта или 16 бит,
3. 3 байта или 24 бита,
4. 4 байта или 32 бита.

В таблице 1 наглядно показано структура заголовка WAV файла.

Таблица 1 – Структура заголовков файла формата WAV

Местоположение	Поле	Описание
0 – 3 байты	chunkId	Обозначает начало RIFF цепочки
4 – 7 байты	chunkSize	Обозначает размер цепочки, начиная с этой позиции.
8 – 11 байты	format	Содержит специальные символы формата
12 – 15 байты	subchunk1Id	Содержит специальные символы "fmt "
16 – 19 байты	subchunk1Size	Обозначает оставшийся размер подцепочки, начиная с этой позиции
20 – 21 байты	audioFormat	Описывает аудио формат
22 – 23 байты	numChannels	Обозначает количество каналов.
24 – 27 байты	sampleRate	Байты, отвечающие за частоту сэмплирования
28 – 31 байты	byteRate	Обозначает количество байт, переданных за одну секунду воспроизведения
32 – 33 байты	blockAlign	Обозначает количество байт, необходимых для одного сэмпла
34 – 35 байты	bitsPerSample	Обозначает количество бит в сэмпле.
36 – 39 байты	subchunk2Id	Содержит специальные символы «data»
40 – 43 байты	subchunk2Size	Обозначает количество байт в области данных
С 44 байта	data	Область данных

Формат файлов с расширением AVI может содержать сжатые видео и аудио данные, использующие для сжатия различные комбинации кодеков. Это позволяет синхронно воспроизводить видео со звуком. Такой файл, в зависимости от кодека, используемого для сжатия информации, может содержать разные виды данных. AVI файлы поддерживают многопотоковое аудио и видео.

Данный формат изначально предназначался для обмена мультимедийными данными. Он был разработан компанией Microsoft совместно с IBM. Существуют несколько отличающихся друг от друга стандартов файлов формата AVI.

Формат AVI – один из вариантов формата RIF. Файлы, представляющие данный формат, состоят из блоков. Каждый блок также может содержать в себе

еще несколько блоков данных. Самый «верхний» блок RIFF содержит идентификатор формы, который собственно и обозначает, что данный файл является AVI файлом. В нем имеется как минимум два блока: блок заголовков и блок данных.

Блок заголовков содержит всю общую информацию о файле: разрешение, продолжительность воспроизведения, частота кадров, разрешение изображения и т.д. Блок данных содержит в себе непосредственно само видео.

Видео и аудио данные записываются в файл, разбивая потоки данных на множество частей и записывая их в один файл, чередуя друг с другом по очереди. Одна секунда изображения AVI файла в среднем занимает около 2 Мбайт памяти на жестком диске.

В шестом разделе подробно описывается разработанное на языке программирования C# в среде разработки Microsoft Visual Studio программное обеспечение. Программа представляет собой Windows Form приложение, каждая вкладка которой реализует внедрение информации в изображение формата PNG, аудиофайлов формата WAV и видеофайлов формата AVI, а также внедрение файла в видео. Интерфейс приложения представлен на рисунке 1.

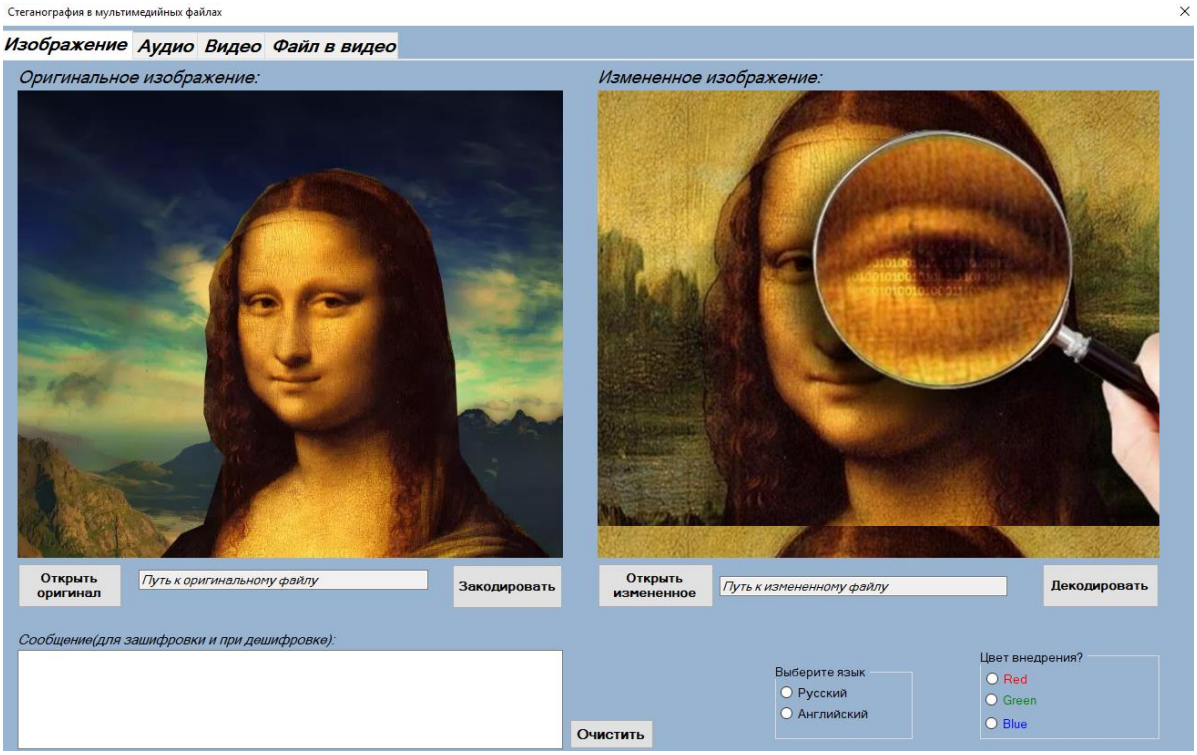

Рисунок 1 – Интерфейс программы

В данном разделе подробно описывается каждая вкладка приложения, приводятся примеры внедрения и извлечения информации из выбранного контейнера.

ЗАКЛЮЧЕНИЕ

В ходе работы были рассмотрены основные виды стеганографии: классическая, компьютерная, цифровая. Были приведены примеры использования каждого вида. Кроме того, были описаны основные понятия, связанные с каждым видом стеганографии. Был рассмотрен метод встраивания битов сообщения в наименее значащие биты контейнера, а также возможность использования в качестве контейнера для тайной передачи информации следующие форматы файлов:

1) Изображения в формате PNG.

2) Аудиофайлы в формате WAV.

3) Видеофайлы в формате AVI.

В ходе практической части работы была реализована программа на языке программирования C# в среде разработки Microsoft Visual Studio, реализующая стеганографию в медиафайлах. Приложение поддерживает скрытие информации в графике формата PNG, аудио формата WAV и видео формата AVI. Каждый из данных форматов является актуальным и часто используется в нынешнее время. Программа использует особенности форматов файлов, реализуя методы компьютерной стеганографии, и для каждого вида контейнеров реализует алгоритм встраивания информации в наименее значащие биты данных, реализуя методы цифровой стеганографии.

Разработанное приложение может использоваться для упрощения тайной передачи информации в среде, не подозревающей о самом факте наличия такой информации.