МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»
(СГУ)

Кафедра теоретических основ компьютерной безопасности и криптографии

Исследование медиафайлов на предмет скрытой информации

АВТОРЕФЕРАТ

дипломной работы

студента 6 курса 631 группы
специальности 10.05.01 Компьютерная безопасность
факультета компьютерных наук и информационных технологий
Лукьянова Дмитрия Владиславовича

Научный руководитель
к. п. н., доцент _____ А. С. Гераськин
18.01.2019 г.

Заведующий кафедрой
д. ф.-м. н., доцент _____ М. Б. Абросимов
18.01.2019 г.

Саратов 2019

ВВЕДЕНИЕ

Методы стеганографии были известны человечеству еще задолго до появления компьютеров. Как полагают историки, еще в Древнем Египте для секретной передачи послания брили голову раба, писали на ней сообщение, после чего, когда волосы отрастали, отправляли раба получателю сообщения. С появлением фотографии стали доступны новые способы — например, в ничего не значащие снимки добавлялись микроточки. Подобные методы передачи секретных сообщений активно использовались во время Второй мировой войны. Можно найти еще массу примеров использования стеганографии в прошлом [1].

Вопрос обеспечения защиты передаваемой информации актуален и в наши дни. При этом во многих странах существуют законы, запрещающие применение стойких алгоритмов криптографии. На применение стеганографии подобных ограничений пока нет, что делает привлекательным ее использование как дополнительного способа защиты информации. В отличие от криптографии, основной задачей стеганографии является сокрытие самого факта передачи сообщения.

В связи с развитием и распространением компьютерных технологий, в качестве контейнера могут служить разнообразные данные: текстовый документ, музыка, изображения, видео и др. Значительный размер медиафайлов, их информационная избыточность, а также возросшая популярность передачи их через сеть делают аудио, видео и графические файлы особенно привлекательными для использования в качестве стегоконтейнера [1].

Цель данной работы является разработка и реализация алгоритма для исследования медиафайлов на предмет скрытой информации.

Задачами диплома являются:

- анализ методов сокрытия информации в медиафайлах;
- анализ методов применяемого для поиска скрытой информации в медиафайлах;

- анализ программного обеспечения используемого для поиска стегановложений в медиафайлах;
- разработка алгоритма для проверки медиафайла на предмет наличия скрытой информации, его реализация и тестирование.

Дипломная работа состоит из введения, 6 разделов, заключения, списка использованных источников и 13 приложений. Общий объем работы – 104 страницы, из них 52 страницы – основное содержание, включая 27 рисунков и 2 таблицы, список использованных источников из 30 наименований.

КРАТКОЕ СОДЕРЖАНИЕ

Первый раздел дипломной работы посвящен рассмотрению методов стеганографии в медиафайлах. Он начинается с краткого анализа частей файла, которые могут быть использованы для сокрытия информации, в зависимости от применяемого типа стеганоалгоритма. Далее следуют два подраздела. Первый из них посвящен рассмотрению методов сокрытия информации в изображениях, второй — методам сокрытия информации в аудиофайлах. В результате рассмотрения наиболее распространенных методов стеганографии в медиафайлах были выявлены основные их недостатки, которые позже были использованы для разработки алгоритма стеганоанализа.

Второй раздел работы посвящен рассмотрению методов стеганоанализа медиафайлов. Он начинается с рассмотрения сигнатурного и схемного методов анализа файла. Далее следуют два подраздела. В первом из них рассматриваются методы стеганоанализа изображений, в том числе RS-анализ изображений и статистическая атака на основе критерия χ -квадрат, которые в дальнейшем были разработанном использованы в алгоритме анализа изображений на предмет наличия скрытой информации. Во втором подразделе были рассмотрены методы стеганоанализа аудиофайлов. В заключение этого раздела был сделан вывод о том, что универсального метода для анализа медиафайлов на предмет скрытой информации, выявляющего применение большинства стеганоалгоритмов среди рассмотренных нет, что подталкивает к разработке собственного.

В третьем разделе задача стеганоанализа медиафайла сводится к задаче классификации и формулируется ее решение с помощью нейронных сетей. А именно описывается сверточная нейронная сеть, ее структура (слои используемые в ней, их порядок, входные и выходные данные), доказывается корректность подхода и возможность применения для разных типов файла.

В четвертом разделе проводится анализ существующих методик анализа медиафайла на предмет наличия в нем скрытой информации. По причине отсутствия универсальной (существуют методики анализа только изображений

или только аудиофайлов) разрабатывается собственный алгоритм, основанный на методе стеганографии медиафайла с помощью нейронных сетей, который был описан в третьем разделе.

Пятый раздел содержит в себе анализ существующего программного обеспечения, предназначенного для поиска скрытой информации в медиафайлах. В заключении раздела был сделан вывод о том, что на данный момент не существует на рынке программного обеспечения продукта, выполнявшего анализ медиафайла на предмет наличия в нем скрытой информации, сочетающего в себе следующие качества: высокая точность обнаружения стеганоконтейнеров, работа с различными типами файлов (изображения, аудиофайлы, видео, …), работа с большим количеством форматов медиафайлов(.jpg, .bmp, .wav, .mp3, .avi, …), не высокая стоимость и понятный «обычному» пользователю графический интерфейс.

Шестой раздел содержит в себе описание реализованного в ходе выполнения дипломной работы программного комплекса, его составных частей, требований к установленным сторонним библиотекам и модулям. Так же продемонстрировано поведение программы при различных входных данных и выбранных режимах работы. В итоге даются результаты тестирования программного комплекса. Тестирование можно считать успешно пройденным. Количество ложно отрицательных вердиктов составляет 2.5%, а ложно положительных — 21%. Стоит заметить что столь высокий процент ошибок второго типа можно проигнорировать из-за того, что он обусловлен высокой «чувствительностью» анализатора служебных полей файла.

ЗАКЛЮЧЕНИЕ

В ходе написания диплома были изучены основные методы стеганографии и стеганоанализа в медиафайлах: метод сокрытия данных в области преобразования; метод сокрытия данных в коэффициентах дискретного косинусного преобразования; метод сокрытия данных в наименьших значащих битах; метод сокрытия данных с помощью четного кодирования; метод сокрытия данных с помощью фазового кодирования; сокрытия данных с помощью метода расширенного спектра; сигнатурный метод анализа файлов; схемный метод анализа файлов метод выявление скрытой информации при помощи визуального анализа битовых срезов; метод выявление скрытой информации при помощи оценки числа переходов значений младших битов в соседних элементах изображения; RS-анализ изображений; метод выявление скрытой информации при помощи оценки частот появления k -битовых серий в потоке LSB элементов контейнера; метод выявление скрытой информации при помощи при помощи статистической атаки на основе критерия χ -квадрат. метод поиска скрытой информации в фазовой области аудиоданных. Были рассмотрены общие сведения о нейронных сетях и на основе светрточной нейронной сети был разработан метод анализа медиафайлов на предмет наличия скрытой информации.

Так же были изучены методики поиска скрытой информации в аудиофайлах и изображениях, на основе которых был создан собственный алгоритм обнаружения стегановставок во всех областях файла. Разработанный алгоритм был реализован в виде программного комплекса и протестирован. Его тестирование дало положительные результаты.

Помимо этого был проведен анализ существующих аналогов реализованной программы. По результатам анализа был сделан вывод что все существующие средства уступают программе, разработанной в данной работе. В следствии чего, можно рекомендовать использовать, реализованный в ходе написания дипломной работы, программный комплекс для анализа аудиофайлов и изображений на предмет наличия в них скрытой информации.

Кроме того, существует теоретическая возможность модификации разработанного алгоритма для анализа видеофайлов.

Все поставленные цели были достигнуты, задачи выполнены.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Кузнецов, А.И. Двоичная тайнопись (по материалам открытой печати) // КомпьютерПресс. 2004. № 4. С. 38–41.

2. Методы сокрытия информации в графических изображениях [электронный ресурс]: статья, открытый доступ. URL: https://ru.bmstu.wiki/Методы_сокрытия_информации_в_графических_изображениях (дата обращения 20.09.2018) Загл. с экрана. Яз. рус.

3. Carandall, R. Some Notes on Steganography. [электронный ресурс]: статья, открытый доступ. URL: http://dde.binghamton.edu/download/Crandall_matrix.pdf (дата обращения 10.10.2018) Загл. с экрана. Яз. англ.

4. Fridrich, J. Feature-Based Steganalysis for JPEG Images and its Implications for Future Design of Steganographic Schemes // Lecture Notes in Computer Science (LNCS). — Vol. 3200. — 2005. — С. 67-81.

5. Иваненко, В.Г.; Ушаков, Н.В. Защита изображений формата jpeg при помощи цифровых водяных знаков. Безопасность информационных технологий, том 4, вып. 25, С. 106 - 113.

6. Гераськин, А.С.; Стрельникова, С.Ю.; Завенягин, М.П. Исследование возможности улучшения реализации алгоритма метода Коча для встраивания цифровых водяных знаков в изображения. Безопасность информационных технологий, том 4, вып. 25, С. 87-95.

7. Васина, Т.С. Обзор алгоритмов стеганографии [электронный ресурс]: статья, открытый доступ. URL: http://technomag.edu.ru/doc/370605.html (дата обращения 15.10.2018) Загл. с экрана. Яз. рус.

8. Кокорин, П.П. О методах стегоанализа в аудиофайлах, Тр. СПИИРАН, 4 (2007), 239–246

9. Сокрытие данных методами стеганографии [электронный ресурс]: статья, открытый доступ. URL: https://ru.bmstu.wiki/Сокрытие_данных_методами_стеганографии (дата обращения 07.09.2018) Загл. с экрана. Яз. рус.

10. Нигматуллин, Э.В.; Ковырзина, К.С.; Соколова, А.В. Обзор методов цифровой аудио стеганографии «Научное сообщество студентов XXI столетия. Технические науки»: Электронный сборник статей по материалам XLII студенческой международной научно-практической конференции. – Новосибирск: Изд. АНС «СибАК». – 2016. – № 5 (41)/ [Электронный ресурс] – Режим доступа. – URL: http://www.sibac.info/archive/Technic/5(41).pdf.

11. Гизунов, Д.С. Методика автоматизированного обнаружения скрытой информации в компьютерных файлах / Д.С. Гизунов, О.А. Демченко, Е.И. Никутин // Известия ТРТУ. – 2006. – Т. 71, № 16. – С. 49-53.

12. Provos, N Detecting steganographic content on the internet / N. Provos, P. Honeyman. // Technical Report CITI 01-1a, University of Michigan, 2001.

13. Алиев, А.Т. О применении стеганографического метода LSB к графическим файлам с большими областями монотонной заливки / А.Т. Алиев // Вестник ДГТУ. – Ростов-на-Дону, 2004. – Т. 4, № 4 (22). – С. 454-460.

14. Барсуков, В.С. Оценка уровня скрытности мультимедийных стеганографических каналов хранения и передачи информации / Барсуков В. С.; Романцов А.П. // Специальная Техника. – 2000. – № 1.

15. Кустов, В.Н.; Параскевопуло, А.Ю. Простые тайны стегоанализа / Кустов В. Н.; Параскевопуло А. Ю. // Защита информации, INSIDE. – 2005. – № 4. – С. 72-78.

16. Иванов, М.А. Теория, применение и оценка качества генераторов псевдослучайных последовательностей / Иванов, М.А.; Чугунков, И.В. – М.: КУДИЦ-ОБРАЗ, 2003.

17. Дрюченко М.А. Алгоритмы выявления стеганографического скрытия информации в JPEG-файлах [электронный ресурс]: статья, открытый доступ. URL: http://www.vestnik.vsu.ru/pdf/analiz/2007/01/2007-01-04.pdf (дата обращения 03.11.2018) Загл. с экрана. Яз. рус.

18. Chen, W. Study of steganalysis methods //Department of Electrical and Computer Engineering, New Jersey 2005

19. Волосатова, Т.М., Чичварин, Н.В. Методика стеганоанализа аудиофайлов. Современные тенденции развития науки и технологий. № 8-2 (2016), С. 9-16

20. Радько, П. Основы ИНС [электронный ресурс]: статья, открытый доступ. URL: https://neuralnet.info/chapter/основы-инс (дата обращения 05.10.2018) Загл. с экрана. Яз. рус.

21. Романов, А.А. Сверточные нейронные сети [электронный ресурс]: статья, открытый доступ. URL: https://scientificresearch.ru/images/PDF/2018/21/svertochnye.pdf (дата обращения 05.12.2018) Загл. с экрана. Яз. рус.

22. Jessica, J.F.; Jan K. Rich models for steganalysis of digital images. IEEE Trans. Information Forensics and Security 7, 3 (2012), 868–882.

23. Min, L.; Qiang, C.; Shuicheng, Y. Network in network. CoRR abs/1312.4400 (2013)

24. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. CoRR abs/1409.1556 (2014)

25. Ioffe S.; Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the 32nd International Conference on Machine Learning, ICML 2015, Lille, France, 6-11 July 2015. 448–456

26. Wang, Y.; Yang, K.; Yi X.; Zhao X.; Xu Z. YCNN-based Steganalysis of MP3 Steganography in the Entropy Code Domain [электронный ресурс]: статья, открытый доступ. URL: http://www.media-security.net/wp-content/uploads/1527558074547225.pdf (дата обращения 23.11.2018) Загл. с экрана. Яз. англ.

27. OutGuess – Steganography Detection [электронный ресурс]: официальная страница продукта, открытый доступ. URL: http://www.outguess.org/detection.php (дата обращения 15.04.2015) Загл. с экрана. Яз. англ.

28. StegoHunt [Электронный ресурс] // WetStone Technologies. Cortland, NY, USA., открытый доступ. URL: https://www.wetstonetech.com/products/stegohunt/ (дата обращения 29.12.2018) Загл. с экрана. Яз. англ.

29. Newman, J. Can stego image from a mobile phone stego app be detected? [Электронный ресурс]: статья, открытый доступ. URL: https://www.nist.gov/sites/default/files/documents/2017/08/23/jennifernewmanwednesdayafternoonsession.pdf

30. BackBone. Руководство пользователя программы StegAlyerAS [электронный ресурс]: статья, открытый доступ. URL: www.computer-forensics-lab.org/getartdoc.php?af_id=127