электронное научно-техническое издание
НАУКА и ОБРАЗОВАНИЕ
Эл № ФС 77 - 30569. Государственная регистрация №0421100025. ISSN 1994-0408

Оценка необнаружимости стеганографических алгоритмов

77-30569/359383
03, март 2012

Ремизов А. В., Филиппов М. В.

УДК.004.318

МГТУ им. Н.Э. Баумана
profitbig@rambler.ru

1. Введение

В связи с повсеместным использованием цифровых носителей и каналов связи актуальна проблема защиты передаваемой и хранимой информации от несанкционированного доступа.

Для защиты собственно информации разработано большое количество криптографических алгоритмов. Однако эти алгоритмы не позволяют скрыть от несанкционированного пользователя факт наличия зашифрованной информации.

Решением данной задачи является стеганография[1] - метод организации связи, который скрывает сообщение в другом, не требующем сокрытия, сообщении. В отличие от криптографии, где противник точно может определить является ли передаваемое сообщение зашифрованным текстом, методы стеганографии позволяют встраивать секретные сообщения в безобидные послания так, чтобы невозможно было заподозрить существование встроенного тайного послания.

По определению, стеганографические алгоритмы должны противодействовать атакам противника, направленным на обнаружение сокрытой информации. В данной статье предлагаются формальные критерии оценки эффективности данного противодействия, а так же методика построения эффективных атак на современные и перспективные стеганографические алгоритмы.

2. Критерий необнаружимости

Задачей сокрытия данных является минимизация вероятности обнаружения скрытых данных. Решения противника о наличии или отсутствии скрытых данных в сообщении всегда сопровождаются ошибками. Программно-аппаратные средства, которыми он располагает, могут также вырабатывать ошибочные посылки, связанные с

естественным несовершенством названных средств, т. е. наличием методических ошибок, носящих случайный характер.

Таким образом, возможны следующие ошибки:

- ошибка "ложной тревоги"

- ошибка необнаружения скрытых данных

Будем обозначать далее событие принятия решения об обнаружении скрытого изображения в подозрительном файле как "ДА", а событие, связанное с необнаружением скрытого изображения - как "НЕТ". Вводим далее следующие обозначения:

Р(ДА/0) - Р(лт) - вероятность ложной тревоги

Р(НЕТ/α) - Р(но) - вероятность необнаружения.

События, связанные с принятием решения о наличии, либо отсутствии скрытой информации в подозрительном файле образуют полную группу, так что

Р(НЕТ/α) + Р(ДА/α) = 1,

Р(НЕТ/0) + Р(ДА/0) = 1.

Тогда вероятность обнаружения Р(обн) определяется зависимостью

Р(обн) = Р(ДА/α) = 1 - Р(НЕТ/α)= 1 - Р(но),

Р(пно) = Р(НЕТ/0) = 1 - Р(ДА/0)= 1 - Р(лт).

Величина Р(обн) - вероятность заключения экспертизы о наличии в подозрительном файле скрытой информации при условии, что скрытая информация действительно в нем присутствует.

Величина Р(пно) - вероятность заключения экспертизы об отсутствии скрытой информации в подозрительном файле, при условии, что ее действительно там нет - вероятность правильного необнаружения.

Таким образом, можно считать, что чем больше значение Р(но) (или чем меньше Р(обн)) при заданном значении Р(лт), тем выше качество системы сокрытия информации.

Вышеперечисленные вероятности в среднем могут быть вычислены опытным путем через частоты принятия противником правильных и ошибочных решений в процессе анализа множества подозрительных сообщений, содержащих (либо нет) данные, сокрытые одним и тем же методом.

Таким образом, оценка необнаружимости основывается в первую очередь на экспериментальных результатах обнаружения сокрытых данных методами обнаружения.

3. Разработка оптимального метода обнаружения

Подавляющее большинство методов обнаружения сокрытых данных основаны на анализе характеристик вероятностного распределения элементов контейнера. Это позво-

ляет прогнозировать действия экспертизы при решении задачи обнаружения скрытых данных. Рассмотрим математическую модель наиболее вероятных действий экспертизы на основе положений теории обнаружения.

Принятие решения экспертизы о наличии скрытых данных в исследуемом контейнере производится не по одному значению какой-то величины, характеризующей содержимое контейнера, а по всему объему контейнера, т.е. по выборке, состоящей из N значений реализации, что позволяет экспертизе более полно использовать априорную информацию и получить наибольший положительный эффект с увеличением объема выборки N.

Таким образом, задача экспертизы по разработке метода обнаружения является задачей оптимизации:

$$\max_{F} P_{обн} при P_{лт} \le P_{лт}^{max}$$

$$P_{обн} = \frac{1}{n}\sum_{i}^{n} F(S(I_i))$$

$$P_{лт} = \frac{1}{n}\sum_{i}^{n} F(I_i)$$

где I набор пустых контейнеров, $S(I)$ функция сокрытия данных, $F(I)$ - функция обнаружения,

$$F(I) = \begin{cases} 1, & решение << ДА >> \\ 0, & решение << НЕТ >> \end{cases}$$

Запишем функцию обнаружения через экспертную оценку $E(I)$:

$$F(I) = \begin{cases} 1, & E(I) \ge E_{крит} \\ 0, & E(I) < E_{крит} \end{cases}$$

$E_{крит}$ - пороговая экспертная оценка.

Так как в качестве контейнеров используются реальные избыточные источники сигнала, контейнер можно разделить на сигнал и шум, где под шумом понимается шум дискретизации, квантовый шум и т.п. искажения, вносимые в "идеальный" сигнал. В случае использования энергонезависимых носителей как контейнеров, под шумом понимаем неиспользуемые блоки файловой системы.

Представим контейнер I как $I = L + G$ где L - очищенный от шума контейнер, G - присутствующий в контейнере шум.

Тогда экспертная оценка контейнера будет определяться взвешенной суммой оценок очищенного контейнера и шума

$$E(I) = w_L E_L(L) + w_G E_G(G),$$

где w_L, w_G - весы соответствующих экспертных оценок [1].

С ростом числа доступных экспертизе пустых потенциальных контейнеров n, точность определения сокрытых в "сигнале" данных увеличивается, так как улучшается модель "сигнала", с которой работает экпертиза. В пределе, имея все возможные пустые контейнеры, экспертиза может абсолютно точно определять сокрытие данных в незашумленной части контейнера.

Исходя из этого, экспертная оценка "сигнала" E_L улучшается следующим образом [1]:

$$\lim_{n\to\infty} E_L(L) = 0,$$

$$\lim_{n\to\infty} E_L(S(L)) = \infty$$

Кроме того, следует отметить, что внедрение сокрытых данных в сигнал L может привести к «видимым» стороннему наблюдателю искажениям. Соответственно, исходя из результатов проведенного анализа и модели стегоканала, практически применимые алгоритмы сокрытия данных размещают скрываемые данные в шуме контейнера:

$$S(I) = L + S(G)$$

Таким образом, оптимальный метод обнаружения скрытой информации строится, основываясь на следующем алгоритме обнаружения:

1. выделяется шум G из предоставленного на экспертизу контейнера I с помощью метода выделения шума N, $G = N(I)$
2. принимается решение "ДА" или "НЕТ" в зависимости от экспертной оценки наличия скрытой информации в выделенном шуме $E_G(G)$

Следовательно задача разработки метода обнаружения может быть представлена как поиск оптимальной функции экспертной оценки E_G и оптимальной функции выделения шума N:

$$\max_{E_G, N} P_{обн} при P_{лт} \le P_{лт}^{max} \qquad (1)$$

Задача построения экспертной оценки E_G является задачей классификации. Для ее решения предлагается использовать метод опорных векторов[2]. Параметры метода опорных векторов, а так же возможные варианты выделения шума и его оценки выбираются предлагается выбирать с помощью генетического алгоритма[3]. Для

построения алгоритма используется учебная выборка, состоящая из набора пустых и заполненных контейнеров.

4. Экспериментальные результаты оценки необнаружимости

Предлагаемая методика была использована при оценке необнаружимости ряда существующих и перспективных алгоритмов сокрытия данных.

В частности, были рассмотрены следующие сценарии сокрытия и соответствующие им алгоритмы:

4.1 Сокрытие на энергонезависимых носителях

Рассматривался метод сокрытия с использованием стеганографической файловой системы (СФС) [4]. Данный метод размещает скрываемые данные в неиспользуемых логических блоках основной файловой системы носителя. В качестве основной файловой системы поддерживаются файловые системы FAT и UDF.

Блоки выбираются случайным образом из числа незанятых основной или скрытой файловыми системами в данный момент. Все записываемые на диск блоки шифруются по описанному в ГОСТ 28147-89 алгоритму шифрования с использованием схемы LRW[5].

При анализе необнаружимости в качестве шума выделялись незанятые основной файловой системой блоки. Для каждого такого блока с помощью критерия согласия Пирсона проверялась гипотеза о том, что значения содержащихся в нем байт соответствуют равномерному распределению, т.е. блок, скорее всего, содержит некие зашифрованные данные.

Все возможные значения байт разбиваются на четыре интервала. Рассчитывается статистика $\chi^2 = N\Sigma\frac{P_i^e - P_i^o}{P_i^e}$, где $P_i^e = \frac{clust\_length}{4}$ - ожидаемое количество попавших в интервал значений, P_i^o- наблюдаемое количество.

Если полученное значение $\chi 2$ меньше квантили закона распределения $\chi 2$ заданного уровня значимости с тремя степенями свободы,

$$P(x) = \frac{\frac{1}{2}^{\frac{3}{2}}}{\Gamma(\frac{3}{2})} x^{\frac{1}{2}} e^{-\frac{x}{2}}$$

кластер считается принадлежащим стеганографической файловой системе.

Для анализа были использованы носители со следующим содержанием:

- программное обеспечение и документы - сценарий типичного использования флеш-брелка;
- изображения в формате JPEG - сценарий типичного использования любительского цифрового фотоаппарата;

- изображения в формате Canon, Nikon, Olympus RAW - сценарий использования профессионального цифрового фотоаппарата;

- зашифрованные данные.

На Рис.1 представлены экспериментальные результаты, иллюстрирующие степень необнаружимости сокрытия данных с помощью стеганографической файловой системы.

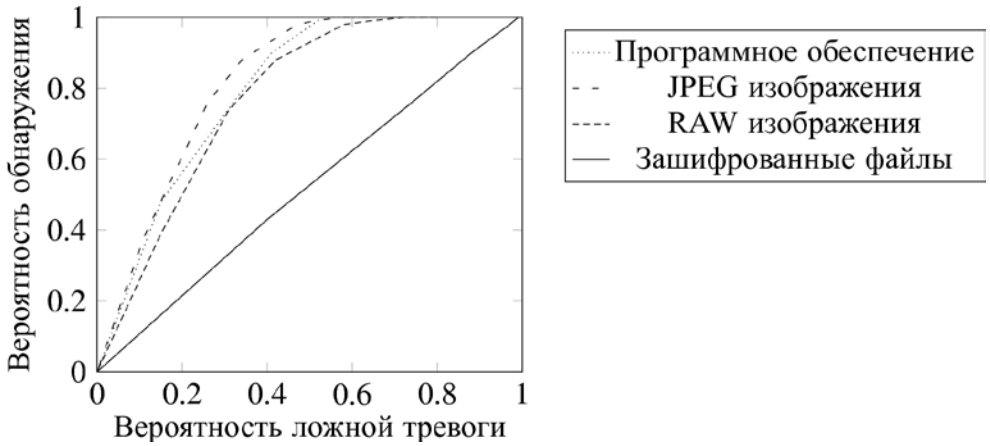

Рис. 1. Необнаружимость в зависимости от содержимого видимой части носителя

Степень необнаружимости увеличивается при дополнительном шифровании данных.

4.2 Сокрытие в сжатых аудиофайлах

Данный сценарий рассматривался на примере сокрытия в файлы формата MP3. Рассматривались классический метод сокрытия mp3stego [6], скрывающий данные в служебной информации кадров MP3 потока, а так же перспективный алгоритм сокрытия в ошибках квантования PQ-MP3 [7], размещающий скрываемые данные путем модификации функции округления коэффициентов модифицированного дискретного косинусного преобразования.

При оценке необнаружимости в качестве возможных характеристик шума рассматривались:

– моменты вейвлет-распределения шума [8];

– моменты распределения посегментной дистанции Хаусдорфа [9];

– распределение значений служебной информации [10];

– вероятности перехода в цепи Маркова 1го порядка, построенной по вторым производным шума[11];

– вероятности перехода в цепи Маркова 1го порядка, моделирующей переход от левого аудиоканала к правому.

Тренировка и тестирование проводились на наборе из 200 ранее не сжимавшихся стереоаудиофайлов из базы [11].

При сокрытие в сжатых аудиофайлах метод mp3stego обнаруживается с близкой к 100% вероятностью. Степень необнаружимости метода PQ-MP3 в зависимости от максимального предельного отклонения/емкости представлена на рис. 2

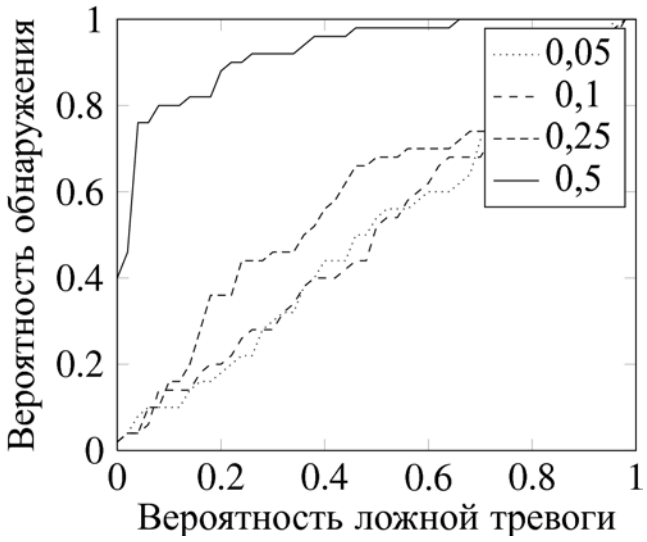

Рис. 2. Необнаружимость метода PQ-MP3 в зависимости от максимального предельного отклонения/емкости

Вероятность обнаружения существенно уменьшается в случае слабого заполнения контейнера скрываемыми данными.

4.3 Сокрытие в сжатых видеофайлах

Для данного сценария рассматривались следующие методы сокрытия: F5[12] скрывающий данные в младших битах коэффициентов косинусного преобразования с применением предварительной фильтрации и матричного кодирования, а так же адаптированный для видеоинформации вариант упомянутого ранее метода PQ-MP3.

При оценке необнаружимости в качестве возможных характеристик шума рассматривались:

– моменты вейвлет-распределения шума [8];

– матрица разницы соседних пикселей, расчитанная на основе цепей Маркова [13];

– Марковские метрики коэффициентов косинусного преобразования [14].

Тренировка и тестирование проводились на наборе из 200 монохромных кадров размером 512x512 пикселов, входящих в базу изображений BOWS2 [15].

На рис.3 представлены экспериментальные данные по необнаружимости при сокрытии информации в сжатых видеофайлах.

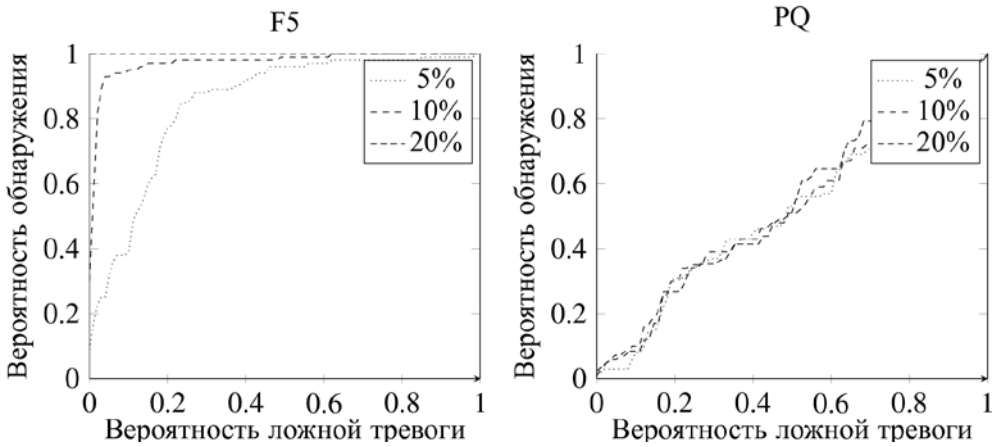

Рис.3. Необнаружимость при сокрытии информации в сжатых видеофайлах.

Наиболее эффективными метриками шума оказались Марковские метрики коэффициентов косинусного преобразования.

6. Список использованной литературы

1. Грибунин В.Г. Цифровая стеганография. - М.: СОЛОН-Пресс, 2002.

2. Cristianini N., Shawe-Taylor J. An introduction to support vector machines and other kernel-based learning methods. - Cambridge University Press, 2000.

3. Koza J. Genetic Programming: On the Programming of Computers by Means of Natural Selection - MIT Press, 1992.

4. Ремизов А.В. Свидетельство о государственной регистрации программы для ЭВМ N2008612801 "Стеганографическая файловая система", 2008.

5. Liskov M., Rivest R., Wagner D. Tweakable Block Ciphers - Proc. 22nd Annual International Cryptology Conference on Advances in Cryptology, 2002.

6. http://www.petitcolas.net/fabien/steganography/mp3stego

7. Ремизов А.В., Филиппов М.В., Чичварин Н.В. Методы защиты информации в звуковых файлах - Информационные технологии, 2009.

8. Fridrich J., Miroslav G. New blind steganalysis and its implications - Proc. SPIE Electronic Imaging, 2006.

9. Liu Y., Chiang K., Corbett C. A Novel Audio Steganalysis Based on High--Order Statistics of a Distortion Measure with Hausdorff Distance - Proc. 11th international conference on Information Security, 2008.

10. Hernandez-Castro J.C., Tapiador J.E. Blind Steganalysis of Mp3stego - J. Inf. Sci. Eng., 2010.

11. Liu Q., Sung A., Qiao M. Novel stream mining for audio steganalysis - Proc. 17th ACM international conference on Multimedia, 2009.

12. Westfeld A. F5 - A Steganographic Algorithm - Information Hiding. - Springer Berlin / Heidelberg, 2001.

13. Pevn´y T., Bas P., Fridrich J. Steganalysis by subtractive pixel adjacency matrix - Trans. Info. For. Sec., 2010.

14. Pevn´y T., Fridrich J. Merging Markov and DCT Features for Multi-Class JPEG Steganalysis - Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX, 2007.

15. http://bows2.gipsa-lab.inpg.fr

electronic scientific and technical periodical
SCIENCE and EDUCATION
EL № FS 77 - 30569. №0421100025. ISSN 1994-0408

Оценка необнаружимости стеганографических алгоритмов

77-30569/359383
03, март 2012

Ремизов А. В., Филиппов М. В.

Bauman Moscow State Technical University
profitbig@rambler.ru

The authors propose formal criteria of estimation of steganographic algorithms in the context of undetectability of hidden information by the opponent. A method of creating optimal detection methods is also presented in the article. The article also contains the results of the analysis of known and prospective stenographic algorithms operation with the usage of methods constructed in this manner.

Publications with keywords: genetic algorithm, steganographic algorithm, undetectability
Publications with words: genetic algorithm, steganographic algorithm, undetectability

References

1. Gribunin V.G., Okov I.N., Turintsev I.V. *Tsifrovaia steganografiia* [Digital steganography]. Moscow, Solon-Press Publ., 2002. 258 p.

2. Cristianini N., Shawe-Taylor J. *An introduction to support vector machines and other kernel-based learning methods*. Cambridge, CUP Publ., 2000. 204 p.

3. Koza J. *Genetic programming: On the programming of computers by means of natural selection*. Cambridge, MIT Publ., 1992. 840 p.

4. Remizov A.V. *Steganograficheskaia failovaia sistema* [Steganographic file system]. Computer program no. 2008612801, 2008.

5. Liskov M., Rivest R., Wagner D. Tweakable Block Ciphers. *Proc. 22nd Annual Int. Cryptology Conf. on Advances in Cryptology (CRYPTO'02)*. Heidelberg, Springer-Verlag, 2002, pp. 31-46.

6. *mp3stego*. Available at: http://www.petitcolas.net/fabien/steganography/mp3stego.

7. Remizov A.V., Filippov M.V., Chichvarin N.V. Metody zashchity informatsii v zvukovykh failakh [Methods of information protection in the acoustic files]. *Informatsionnye tekhnologii*, 2009, no. 10, pp. 10-13.

8. Goljan M., Fridrich J., Holotyak T. New blind steganalysis and its implications. *Proceedings of SPIE*, 2006, vol. 6072, art. no. 607201. DOI: 10.1117/12.643254.

9. Liu Y., Chiang K., Corbett C., Archibald R., Mukherjee B., Ghosal D. A novel audio steganalysis based on high-order statistics of a distortion measure with hausdorff distance**.** *Proc. 11th Int. Conf. on Information Security (ISC-2008).* Taipei, Sept. 15-18, 2008. (*Lecture Notes in Computer Science*, 2008, Vol. 5222, pp. 487-501.). DOI: 10.1007/978-3-540-85886-7_33.

10. Hernandez-Castro, J.C., Tapiador, J.E., Esther, P., Romero-Gonzalez, A. Blind steganalysis of mp3stego**.** *Journal of Information Science and Engineering*, 2010, vol. 26, no. 5, pp. 1787-1799.

11. Liu Q., Sung A.H., Qiao M. Novel stream mining for audio steganalysis. *Proc. 17th ACM Int. Conf. on Multimedia (MM'09).* Beijing, Oct. 19-24, 2009, pp. 95-104. DOI: 10.1145/1631272.1631288.

12. Westfeld A. F5-A steganographic algorithm: High capacity despite better steganalysis. *Proc. 4th Int. Workshop on Information Hiding (IHW '01).* Berlin, Springer, 2001, pp. 289-302.

13. Pevny T., Bas P., Fridrich J. Steganalysis by subtractive pixel adjacency matrix. *IEEE Trans. on Information Forensics and Security*, 2010, vol. 5, no. 2, pp. 215-224, art. no. 5437325. DOI: 10.1109/TIFS.2010.2045842.

14. Pevny T., Fridrich J. Merging Markov and DCT features for multi-class JPEG steganalysis. *Proceedings of SPIE*, 2007, vol. 6505, art. no. 650503.

15. *BOWS-2*. Available at: http://bows2.gipsa-lab.inpg.fr/.