УДК 003.042

АТАКИ НА СТЕГАНОГРАФИЧЕСКУЮ СИСТЕМУ ЦИФРОВЫХ ВОДЯНЫХ ЗНАКОВ И ПРОТИВОДЕЙСТВИЕ ИМ

И.В. ИСАКОВ
(***Представлено: канд. физ.-мат. наук, доц. Ю.Ф. ПАСТУХОВ***)

В данной работе исследуем перечень возможных атак на стегосистему цифровых водяных знаков, а способы противодействия этим атакам.

Классификация атак на ЦВЗ. ЦВЗ должны удовлетворять противоречивым требованиям визуальной (аудио) незаметности и робастности к основным операциям обработки сигналов. В дальнейшем без потери общности будем предполагать, что в качестве контейнера используется изображение.

Обратимся к системе встраивания сообщений путем модификации младшего значащего бита (LSB) пикселов. Практически любой способ обработки изображений может привести к разрушению значительной части встроенного сообщения. Например, рассмотрим операцию вычисления скользящего среднего по двум соседним пикселам $\frac{(a+b)}{2}$, являющуюся простейшим примером низкочастотной фильтрации. Пусть значения пикселов *a* и *b* могут быть четными или нечетными с вероятностью $p = \frac{1}{2}$. Тогда и значение младшего значащего бита изменится после усреднения в половине случаев. К тому же эффекту может привести и изменение шкалы квантования, скажем, с 8 до 7 бит. Аналогичное влияние оказывает и сжатие изображений с потерями. Более того, применение методов очистки сигналов от шумов, использующих оценивание и вычитание шума, приведет к искажению подавляющего большинства бит скрытого сообщения.[1] Существуют также и гораздо более губительные для ЦВЗ операции обработки изображений, например, масштабирование, повороты, усечение, перестановка пикселов. Ситуация усугубляется еще и тем, что преобразования стегосообщения могут осуществляться не только нарушителем, но и законным пользователем, или являться следствием ошибок при передаче по каналу связи.

Сдвиг на несколько пикселов может привести к не обнаружению ЦВЗ в детекторе. Аналоговые видеомагнитофоны, как правило, несколько сдвигают изображение из-за неравномерности вращения двигателя лентопротяжного механизма или изнашивания ленты. Сдвиг может быть незаметен для глаза, но привести к разрушению ЦВЗ.

Теперь рассмотрим атаки, специфичные для систем ЦВЗ. Можно выделить следующие категории атак против таких стегосистем[3].

1. Атаки против встроенного сообщения — направлены на удаление или порчу ЦВЗ путем манипулирования стего. Входящие в эту категорию методы атак не пытаются оценить и выделить водяной знак. Примерами таких атак могут являться линейная фильтрация, сжатие изображений, добавление шума, выравнивание гистограммы, изменение контрастности и т. д.

2. Атаки против стегодетектора — направлены на то, чтобы затруднить или сделать невозможной правильную работу детектора. При этом водяной знак в изображении остается, но теряется возможность его приема. В эту категорию входят такие атаки, как аффинные преобразования (то есть масштабирование, сдвиги, повороты), усечение изображения, перестановка пикселов и т. д.

2. Атаки против протокола использования ЦВЗ — в основном связаны с созданием ложных ЦВЗ, ложных стего, инверсией ЦВЗ, добавлением нескольких ЦВЗ.

4. Атаки против самого ЦВЗ — направлены на оценивание и извлечение ЦВЗ из стегосообщения, по возможности без искажения контейнера. В эту группу входят такие атаки, как атаки сговора, статистического усреднения, методы очистки сигналов от шумов, некоторые виды нелинейной фильтрации [4] и другие.

Надо заметить, что рассматриваемая классификация атак не является единственно возможной и полной. Кроме того, некоторые атаки (например, удаление шума) могут быть отнесены к нескольким категориям.

В соответствии с этой классификацией все атаки на системы встраивания ЦВЗ могут быть разделены на четыре группы:

1) атаки, направленные на удаление ЦВЗ;
2) геометрические атаки, направленные на искажение контейнера;
3) криптографические атаки;
4) атаки против используемого протокола встраивания и проверки ЦВЗ.

Методы противодействия атакам на системы ЦВЗ.

В простейших стегосистемах ЦВЗ при встраивании используется псевдослучайная последовательность, являющаяся реализацией белого гауссовского шума и не учитывающая свойства контейнера. Такие системы практически неустойчивы к большинству рассмотренных выше атак. Для повышения робастности стегосистем можно предложить ряд улучшений[2].

В робастной стегосистеме необходим правильный выбор параметров псевдослучайной последовательности. Известно, что при этом системы с расширением спектра могут быть весьма робастными по отношению к атакам типа добавления шума, сжатия и т. п. Так считается, что ЦВЗ должен обнаруживаться при достаточно сильной низкочастотной фильтрации (7x7 фильтр с прямоугольной характеристикой). Следовательно, база сигнала должна быть велика, что снижает пропускную способность стегоканала. Кроме того, используемая в качестве ключа ПСП должна быть криптографически безопасной.

Причиной нестойкости систем ЦВЗ с расширением спектра к подобным атакам объясняется тем, что используемая для вложения последовательность обычно имеет нулевое среднее. После усреднения по достаточно большому количеству реализаций ЦВЗ удаляется. Известен специальный метод построения водяного знака, направленный против подобной атаки. При этом коды разрабатываются таким образом, чтобы при любом усреднении всегда оставалась не равная нулю часть последовательности (статическая компонента). Более того, по ней возможно восстановление остальной части последовательности (динамическая компонента). Недостатком предложенных кодов является то, что их длина увеличивается экспоненциально с ростом числа распространяемых защищенных копий. Возможным выходом из этого положения является применение иерархического кодирования, то есть назначения кодов для группы пользователей. Некоторые аналогии здесь имеются с системами сотовой связи с кодовым разделением пользователей (CDMA).

Различные методы противодействия предлагались для решения проблемы прав собственности. Первый способ заключается в построении необратимого алгоритма ЦВЗ. ЦВЗ должен быть адаптивным к сигналу и встраиваться при помощи однонаправленной функции, например, хэш-функции. Хэш-функция преобразует 1000 бит исходного изображения V в битовую последовательность b_i, $i = 1 \ldots 1000$. Далее, в зависимости от значения b_i используется две функции встраивания ЦВЗ. Если $b_i = 0$, то используется функция $v_i(1 + aw_i)$, если $b_i = 1$, то функция $v_i(1 + aw_i)$, где v_i - i-й коэффициент изображения, w_i - i-й бит встраиваемого сообщения. Предполагается, что такой алгоритм формирования ЦВЗ предотвратит фальсификацию.

Второй способ решения проблемы прав собственности заключается во встраивании в ЦВЗ некоторой временной отметки, предоставляемой третьей, доверенной стороной. В случае возникновения конфликта лицо, имеющее на изображении более раннюю временную отметку, считается настоящим собственником.

Один из принципов построения робастного ЦВЗ заключается в адаптации его спектра. В ряде работ показано, что огибающая спектра идеального ЦВЗ должна повторять огибающую спектра контейнера. Спектральная плотность мощности ЦВЗ, конечно же, намного меньше. При такой огибающей спектра винеровский фильтр дает наихудшую оценку ЦВЗ из возможных: дисперсия значений ошибки достигает дисперсии значений заполненного контейнера. На практике адаптация спектра ЦВЗ возможна путем локального оценивания спектра контейнера. С другой стороны, методы встраивания ЦВЗ в области преобразования достигают этой цели за счет адаптации в области трансформанты.[4]

Для защиты от атак типа аффинного преобразования можно использовать дополнительный (опорный) ЦВЗ. Этот ЦВЗ не несет в себе информации, но используется для «регистрации» выполняемых нарушителем преобразований. В детекторе ЦВЗ имеется схема предискажения, выполняющая обратное преобразование. Здесь имеется аналогия с используемыми в связи тестовыми последовательностями. Однако, в этом случае атака может быть направлена именно против опорного ЦВЗ. Другой альтернативой является вложение ЦВЗ в визуально значимые области изображения, которые не могут быть удалены из него без существенной его деградации. Наконец, можно разместить стего в инвариантных к преобразованию коэффициентах. Например, амплитуда преобразования Фурье инвариантна к сдвигу изображения (при этом меняется только фаза).

Другим методом защиты от подобных атак является блочный детектор. Модифицированное изображение разбивается на блоки размером 12x12 или 16x16 пикселов, и для каждого блока анализируются все возможные искажения. То есть пикселы в блоке подвергаются поворотам, перестановкам и т. п. Для каждого изменения определяется коэффициент корреляции ЦВЗ. Преобразование, после которого коэффициент корреляции оказался наибольшим, считается реально выполненным нарушителем. Таким образом появляется возможность как бы обратить внесенные нарушителем искажения. Возможность такого подхода основана на предположении о том, что нарушитель не будет значительно искажать контейнер (это не в его интересах).

Заключение. Подведя итоги, можно с уверенностью сказать, что от каждой атаки можно защититься зная все тонкости данной атаки, а также всегда стоит закрывать уязвимости своей системы и проводить атаки против неё тестируя на устойчивость к ним. Проводя атаки на свою же ЦВЗ мы анализируем её уязвимость к данной атаке которую в итоге мы можем закрыть. Полностью обеспечить защищенность своей ЦВЗ невозможно, ведь пока мы закрываем одну уязвимость, злоумышленник ищет новую, и эта гонка между злоумышленником и специалистом по компьютерной безопасности будет вечной. Самый верный способ защиты это «нападение», если мы сами умеем выявлять уязвимости, то мы и успешно сможем их закрыть, это остается главным аспектом деятельности специалистов КБ – изучение и практика атак, и разработка систем защиты.

ЛИТЕРАТУРА

1. Садов, В.С. Компьютерная стеганография (Конспект лекций)/В.С. Садов – Минск: БГУ, 2010. – 211 с.
2. Шнайер, Б. Секреты и ложь. Безопасность данных в цифровом мире / Б. Шнайер – СПб.: Питер, 2003. – 368 с.
3. Грибунин, В.Г. Цифровая стеганография/ В.Г. Грибунин – Москва: «Академия», 2009. – 265 с.
4. Шелухин, О.И. Стеганография. Алгоритмы и программная реализация/ О.И. Шелухин /С.Д. Канаев – Москва:«Горячая линия - Телеком», 2017. – 592 с.