Стеганография

История стеганографии

Стеганография (от греч. στεγανοσ — скрытый и греч. γραφω — пишу, буквально «тайнопись») — это наука о скрытой передаче информации путём сохранения в тайне самого факта передачи. В отличие от криптографии, которая скрывает содержимое секретного сообщения, стеганография скрывает само его существование. Стеганографию обычно используют совместно с методами криптографии, таким образом, дополняя её.

Первые упоминания о стеганографии встречаются в трудах древнегреческого историка Геродота. В первом способе на обритую голову раба записывалось необходимое сообщение, а когда его волосы отрастали, он отправлялся к адресату, который вновь брил его голову и считывал доставленное сообщение. Второй способ заключался в следующем: сообщение наносилось на деревянную дощечку, а потом она покрывалась воском, и, тем самым, не вызывала никаких подозрений. Потом воск соскабливался, и сообщение становилось видимым.

Одним из наиболее распространенных методов классической стеганографии является использование симпатических (невидимых) чернил. Текст, записанный такими чернилами, проявляется только при определенных условиях (нагрев, специальное освещение, химический проявитель и т. д.). Изобретенные ещё в I веке н. э. Филоном Александрийским, они продолжали использоваться как в средневековье, так и в новейшее время.

Существуют также чернила с химически нестабильным пигментом. Текст, написанный этими чернилами, первое время выглядит, как текст, написанный обычной ручкой, но через некоторое время нестабильный пигмент, входящий в состав чернил, разлагается, и от текста не остаётся ни следа.

Другие стеганографические методы:

- микроточки — микроскопические фотоснимки, вклеиваемые в текст писем;
- запись на боковой стороне колоды карт, расположенных в условленном порядке;
- замена смысла слов в тексте;
- трафареты, которые, будучи положенными на текст, оставляют видимыми только значащие буквы;
- узелки на нитках и т. д.

Основные понятия

В 1983 году Симмонс предложил т. н. «проблему заключенных». Её суть состоит в том, что есть человек на свободе (Алиса), в заключении (Боб) и охранник Вилли. Алиса хочет передавать сообщения Бобу без вмешательства охранника. В этой модели сделаны некоторые допущения. Предполагается, что перед заключением Алиса и Боб договариваются о кодовом символе, который отделит одну часть текста письма от другой, в которой скрыто сообщение. Вилли имеет

право читать и изменять сообщения. В 1996 году на конференции Information Hiding: First Information Workshop была принята единая терминология:

Стеганографическая система (стегосистема) — объединение методов и средств, используемых для создания скрытого канала для передачи информации. При построении такой системы условились о том, что:

- враг представляет работу стеганографической системы. Неизвестным для противника является ключ с помощью которого можно узнать о факте существования и содержания тайного сообщения;
- при обнаружении противником наличия скрытого сообщения он не должен суметь извлечь сообщение до тех пор пока он не будет владеть ключом;
- противник не имеет технических и прочих преимуществ.

Сообщение — это термин, используемый для общего названия передаваемой скрытой информации, будь то надпись молоком, сообщение на голове раба или цифровой файл.

Контейнер — так называется любая информация, используемая для сокрытия тайного сообщения. Пустой контейнер — контейнер, не содержащий секретного послания. Заполненный контейнер (стегоконтейнер) — контейнер, в который записано секретное сообщение.

Стеганографический канал (стегоканал) — канал передачи стегоконтейнера.

Ключ (стегоключ) — секретный ключ, нужный для сокрытия стегоконтейнера. Ключи в стегосистемах бывают двух типов: секретные и открытые. Если стегосистема использует секретный ключ, то он должен быть создан или до начала обмена сообщениями, или передан по защищённому каналу. Стегосистема, использующая открытый ключ, должна быть устроена таким образом, чтобы было невозможно получить из него закрытый ключ. В этом случае открытый ключ мы можем передавать по незащищённому каналу.

Компьютерная стеганография

Компьютерная стеганография — направление классической стеганографии, основанное на особенностях компьютерной платформы. В качестве примера можно упомянуть о стеганографической файловой системе StegFS. Приведём ещё несколько примеров:

Использование зарезервированных полей компьютерных форматов файлов. Суть метода состоит в том, что многие форматы файлов имеют зарезервированные поля, или поля расширений, которые были введены при стандартизации формата, но сейчас не используются, и, в подавляющем большинстве случаев, по умолчанию заполняется нулями. Мы можем использовать эту «неиспользуемую» часть для записи своих данных. Недостатками этого метода является низкая степень скрытности и крайне малый объем передаваемой информации.

Метод скрытия информации в неиспользуемых местах дисков. При использовании этого метода информация записывается в неиспользуемые части диска (например, на нулевую дорожку). Недостатки этого способа такие же, как и у предыдущего: малая степень скрытности и небольшой объём сообщений.

Метод использования особых свойств форматов, которые не отображаются на экране. Яркий пример этого метода белый текст на белом фоне в текстовом редакторе. С помощью этого метода можно закодировать практически неограниченный объём информации, но обнаружить эту информация по-прежнему довольно легко.

Использование особенностей файловых систем. При хранении на жестком диске файл почти всегда занимает целое число кластеров. К примеру, в файловой системе $FAT32$ стандартный размер кластера — 4 Кб. Соответственно для хранения 1 Кб информации на диске выделяется 4 Кб информации, из которых используется только 1 Кб, а остальные 3 можно использовать для хранения скрытой информации. Этот метод тоже довольно легко обнаружить.

Цифровая стеганография

Цифровая стеганография — направление классической стеганографии, основанное на сокрытии или внедрении дополнительной информации в цифровые объекты (контейнеры), вызывая при этом некоторые искажения этих объектов. Как правило, контейнеры являются мультимедиа-объектами: изображениями, видео файлами, аудиозаписями и т.д. И в эти объекты вносятся изменения, лежащие ниже порога чувствительности среднестатистического человека. Кроме того, в оцифрованных объектах изначально присутствует шум квантования, при воспроизведении этих объектов на современных устройствах появляется дополнительный шум цифро-аналогового преобразования. Всё это способствует большей незаметности сокрытой информации.

Алгоритмы

Все алгоритмы встраивания скрытой информации можно разделить на несколько групп:

- алгоритмы, работающие с самим цифровым сигналом (например, метод LSB);
- алгоритмы, использующие наложение скрываемой информации поверх оригинала (часто используется для встраивания цифровых водяных знаков);
- алгоритмы, использующие особенности форматов файлов (сюда можно отнести запись информации в метаданные или в зарезервированные поля файла).

По способу встраивания информации стегоалгоритмы можно разделить на три группы:

- линейные (аддитивные). Алгоритмы аддитивного внедрения информации заключаются в линейной модификации исходного изображения, а извлечение информации в декодере производится корреляционными методами;
- нелинейные. В нелинейных методах встраивания информации используется скалярное либо векторное квантование;
- другие. Среди других методов определенный интерес представляют методы, использующие идеи фрактального кодирования изображений.

Метод LSB

Суть метода LSB (Least Significant Bit, наименьший значащий бит) заключается в следующем: мы заменяем младшие биты в байтах контейнера битами секретного сообщения. При этом мы стараемся добиться того, чтобы искажения в контейнере были минимальными. Рассмотрим конкретный пример реализации метода LSB. Пусть у нас есть изображение, закодированное в

формате bmp, мы будем менять биты в байтах, отвечающих за кодирование цветов. Допустим, очередной байт нашего секретного сообщения – 11001011, а байты в изображении – … $11101100\ 01001110\ 01111100\ 0101100111$ …. В этом случае кодирование будет выглядеть так: мы разобьём байт секретного сообщения на 4 двухбитовые части: $11, 00, 10, 11$, и заменим полученными фрагментами младшие биты байтов исходного изображения: … $11101\mathbf{11}\ 01001\mathbf{0}0\ 01111\mathbf{1}0\ 0101100111$ …. Такая замена в общем случае не заметна человеческому глазу. Более того, многие старые устройства вывода, даже не смогут отобразить такие незначительные перемены.

Понятно, что можно менять не только 2 младших бита, но и любое их количество. Тут есть следующая закономерность: чем большее количество бит мы меняем, тем больший объём информации мы можем спрятать, и тем большие помехи в исходном изображении это вызовет. Для примера вот два изображения:

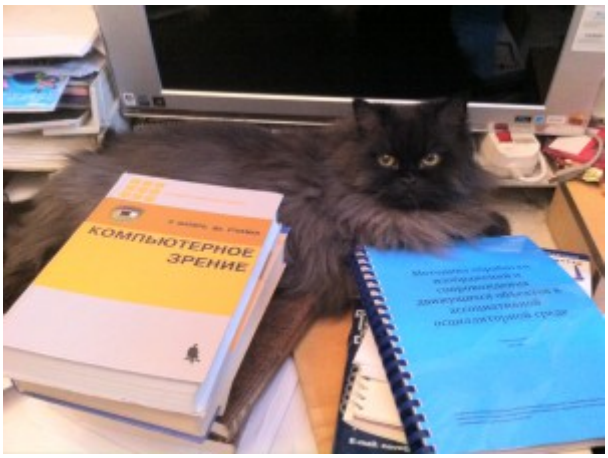

Рисунок 1 Исходное изображение

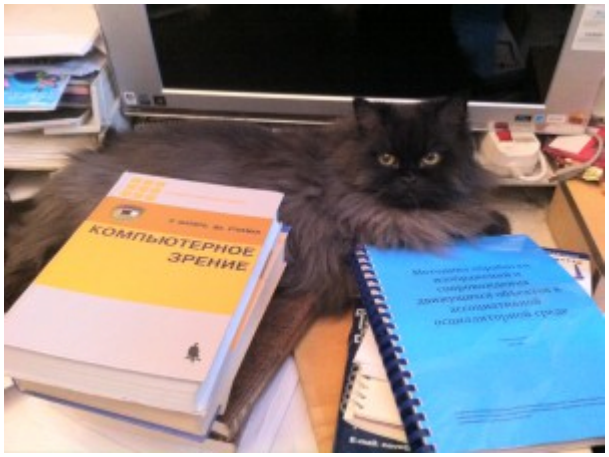

Рисунок 2 Изображение с записанными данными

Во втором изображении записан весь текст лекции от её начала до текущего момента, при этом искажения, внесённые в изображение практически невозможно заметить.

Метод LSB очень прост для реализации, но в тоже время он является весьма неустойчивым: изменение изображения скорее всего уничтожит записанную информацию, поэтому применять этот метод стоит при использовании помехозащищённого стегоканала. Обнаружить -кодирование можно по аномальным характеристикам распределения значений младших бит отсчётов цифрового сигнала. Для того, чтобы усложнить анализ можно изменять младшие биты во всём

изображении целиком, вне зависимости от того, какая его часть использовалась для сокрытия информации.

Другие методы скрытия информации в графических файлах ориентированы на форматы файлов с потерями, к примеру, $JPEG$. В отличие от LSB они более устойчивы к геометрическим преобразованиям.

Программная реализация метода LSB

При рассмотрении программной реализации предполагаем, что данные представлены в виде строки str, размер сообщения хранится в переменной $MsgSize$, файл с изображением открыт как файловый поток $f: TFileStream$. Тогда стеганографические операции описываются следующими фрагментами кода.

Запись данных	Чтение данных

for i:=1 to length(str) do
begin
  l1:=byte(str[i]) shr 6;
  l2:=byte(str[i]) shl 2; l2:=l2 shr 6;
  l3:=byte(str[i]) shl 4; l3:=l3 shr 6;
  l4:=byte(str[i]) shl 6; l4:=l4 shr 6;
 	
  f.ReadBuffer(tmp,1);
  f.Position:=f.Position-1;
  tmp:=((tmp shr 2) shl 2)+l1;
  f.WriteBuffer(tmp,1);
 	
  f.ReadBuffer(tmp,1);
  f.Position:=f.Position-1;
  tmp:=((tmp shr 2) shl 2)+l2;
  f.WriteBuffer(tmp,1);
 	
  f.ReadBuffer(tmp,1);
  f.Position:=f.Position-1;
  tmp:=((tmp shr 2) shl 2)+l3;
  f.WriteBuffer(tmp,1);
 	
  f.ReadBuffer(tmp,1);
  f.Position:=f.Position-1;
  tmp:=((tmp shr 2) shl 2)+l4;
  f.WriteBuffer(tmp,1);
end;
	for i:=1 to MsgSize do
	begin
	  f.ReadBuffer(tmp,1);
	  l1:=tmp shl 6;
	
	  f.ReadBuffer(tmp,1);
	  l2:=tmp shl 6; l2:=l2 shr 2;
	
	  f.ReadBuffer(tmp,1);
	  l3:=tmp shl 6; l3:=l3 shr 4;
	
	  f.ReadBuffer(tmp,1);
	  l4:=tmp shl 6; l4:=l4 shr 6;
	  str:=str+char(l1+l2+l3+l4);
	end;

Эхо-методы

Эхо-методы применяются в цифровой аудиостеганографии и используют неравномерные промежутки между эхо-сигналами для кодирования последовательности значений. При наложении ряда ограничений соблюдается условие незаметности для человеческого восприятия. Эхо характеризуется тремя параметрами: начальной амплитудой, степенью затухания, задержкой. При достижении некоего порога между сигналом и эхом они смешиваются. В этой точке человеческое ухо не может уже отличить эти два сигнала. Наличие этой точки сложно определить, и она зависит как от качества исходной записи, так и от слушателя. Для обозначения логического нуля и логической единицы используются различные задержки, но они обе должны быть меньше, чем порог чувствительности слушателя.

Фазовое кодирование

Фазовое кодирование (phase coding, фазовое кодирование) — также применяется в цифровой аудиостеганографии. Происходит замена исходного звукового элемента на относительную фазу, которая и является секретным сообщением. Фаза подряд идущих элементов должна быть добавлена таким образом, чтобы сохранить относительную фазу между исходными элементами. Фазовое кодирование является одним из самых эффективных методов скрытия информации.

Метод расширенного спектра

Метод встраивания сообщения заключается в том, что специальная случайная последовательность встраивается в контейнер. Затем с помощью специального согласованного фильтра, данная последовательность детектируется. Данный метод позволяет встраивать большое количество сообщений в контейнер, и они при этом не будут создавать помехи друг другу.

Использование особенностей форматов JPEG и RAR

Особенностью формата $jpeg$ является то, что все программы просмотра игнорируют любую информацию, записанную после изображения и некоторых служебных данных. Программы, работающие с -архивами наоборот игнорируют любую информацию до начала заархивированной части. Два этих эффекта позволяют объединять между собой -изображение и rar-архив с полным сохранением функциональности. Если полученный объединённый файл имеет расширение jpg $(jpeg)$, то по умолчанию он будет открыт, как изображение в соответствующей программе. Если же файл имеет расширение rar, то он будет открыт архиватором и часть с изображением будет просто проигнорирована.

Для использования этого метода не нужны никакие специальные программы. Его можно реализовать в консоли $Windows$ с помощью следующей команды. Если изображение называется $img.jpg$, а архив $arj.rar$, то соответствующая команда выглядит так: $copy\ /b\ img.jpg + arj.rar\ img\_s.jpg$. Полученный файл $img\_s.jpg$ может быть открыт и как архив, и как изображение.

Описанный метод очень легко обнаружить по увеличению размера исходного файла.

Применение цифровой стеганографии

Помимо передачи секретных сообщений, цифровая стеганография может использоваться для встраивания цифровых водяных знаков (ЦВЗ). ЦВЗ являются основой для систем защиты авторских прав. Методы этого направления предназначены для встраивания скрытых маркеров, устойчивых к различным преобразованиям контейнера. Встроив такой ЦВЗ в изображение, автор всегда сможет доказать свои права на изображение.