КОСВЕННАЯ СТЕГАНОГРАФИЯ

Надир Алишов

***Аннотация**: Описывается оригинальный метод шифрования и дешифрования на основе способа, называемого косвенной стеганографией. Рассматривается вариант реализации устройства для формирования нераскрываемых шифров файлов и документов, что позволит обеспечить требуемый уровень защиты от несанкционированного доступа, реализовать электронную цифровую подпись, а также новую технологию организации безопасности информационных ресурсов в корпоративных сетях компьютеров.*

***Ключевые слова**: криптография, компьютерная стеганография, стегосистема, защита информации, случайные и псевдослучайные числа, контейнер, секретный ключ.*

***ACM Classification Keywords**: D.4.6 Security and Protection*

***Conference**: The paper is selected from XVth International Conference "Knowledge-Dialogue-Solution" KDS 2009, Varna, Bulgaria, June-July 2009*

Введение

Известны *два направления исследований*, связанных с защитой компьютерной информации от несанкционированного использования.

1. ***Компьютерная криптография.*** Информация, подлежащая защите, шифруется с помощью числовых ключей, причем с увеличением разрядности ключей возрастает вычислительная сложность преобразования. Существует множество вариантов реализации компьютерной криптографии, но общую схему такого шифрования можно представить следующим образом (рис. 1).

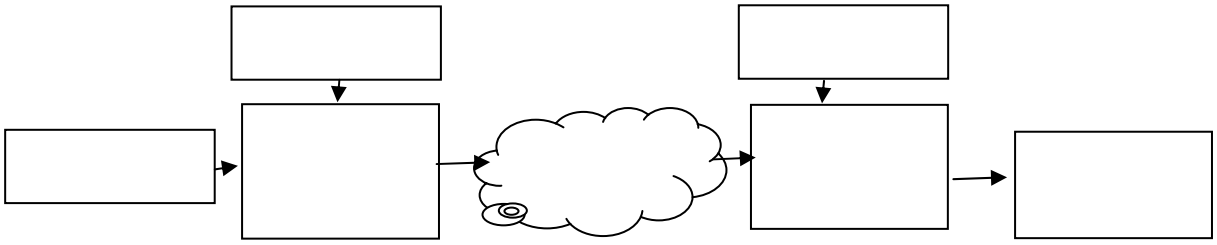

Рис. 1. Концепция компьютерной криптографии

2. ***Компьютерная стеганография.*** Данные, подлежащие защите, смешиваются с определенным видом мультимедийной информации (речь, аудио, видео, изображение и т.п.) и передаются законному пользователю. Разработано множество вариантов реализации этого метода. При этом общая схема стеганографического преобразования выглядит следующим образом (рис. 2).

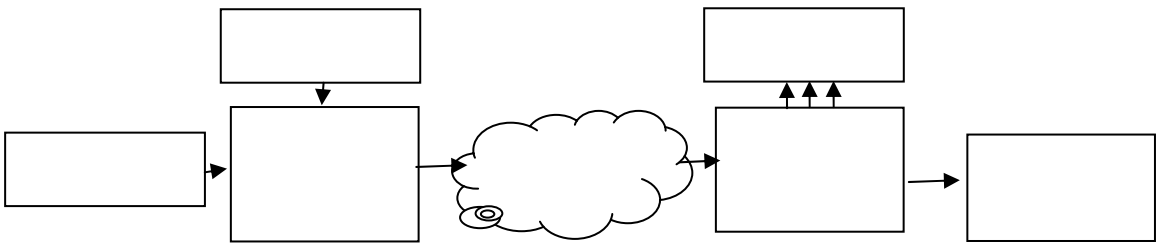

Рис. 2. Концепция компьютерной стеганографии

Главный недостаток указанных методов состоит в том, что в конечном итоге защищаемая информация передается по каналу в зашифрованном или смешанном виде, что позволяет криптоаналитику провести соответствующий анализ для взлома шифра и/или выделения полезных данных. К тому же при компьютерной стеганографии трудно реализовать передачу большого объема информации, а это очень важно для современных компьютерных сетей.

Метод решения поставленной задачи

***При косвенной стеганографии* (*предлагаемый метод*)** полезная информация вообще не передается по каналу. Суть метода заключается в следующем. У отправителя и получателя имеются одинаковые файлы, которые являются секретными ключами. Байты информации, подлежащей защите, заменяются (*по определенному алгоритму*) байтами секретного файла. Новый файл такого же размера, как и исходное сообщение, передается адресату и при получении подвергается обратному преобразованию: его байты заменяются байтами секретного файла (*зеркальный алгоритм*).

Общая схема косвенной стеганографии выглядит следующим образом (рис. 3).

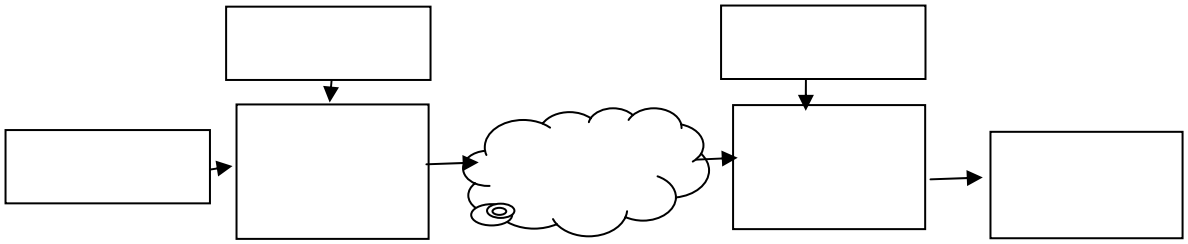

Рис. 3. Концепция косвенной стеганографии

Формальные определения известных современных стегосистем приведены в [1, 2].

1. Совокупность $Æ = <C, M, D, E>$, где C – множество контейнеров; M – множество секретных сообщений, $|C| \geq |M|$; $E: C \times M \to C$, $D: C \to M$ − функции сокрытия и извлечения сообщения из контейнера C, причем $D(E(c, m)) = m$ для любых $m \in M$ и $c \in C$, представляет собой бесключевую стегосистему.

2. Совокупность $Æ = <C, M, K, D, E>$, где C – множество контейнеров; M – множество секретных сообщений, причем $|C| \geq |M|$; K – множество секретных ключей; $E_k: C \times M \times K \to C$, $D_k: C \times K \to M$ − стеганографические преобразования со свойством $D_k(E_k(c, m, k), k) = m$ для любых $m \in M$, $c \in C$ и $k \in K$, представляет собой стегосистему с секретным ключом.

3. Совокупность $Æ = <C, M, K, D, E>$, где C – множество контейнеров, M – множество секретных сообщений, причем $|C| \geq |M|$; $K = (K_1, K_2)$ − множество пар стегоключей; $E_k: C \times M \times K_1 \to C$, $D_k: C \times K_2 \to M$ − стеганографические преобразования со свойством $D_k(E_k(c, m, k_1), k_2) = m$ для любых $m \in M$ и $c \in C$, представляет собой стегосистему с открытым ключом.

Для косвенной стегосистемы дадим следующее определение.

Определение. Совокупность $Œ = <C, @C, M, D, E>$, где C – множество контейнеров-ключей, $@C$ – множество параметров элементов множества C (множество косвенных контейнеров), M – множество секретных сообщений, $E_@: C \times M \to @C$, $D_@: C \times @C \to M$ − стеганографические преобразования со свойством $D_@(E_@(c, m), @c) = m$ для любых $m \in M$, $c \in C$ и $@c \in @C$, представляет собой ***косвенную стегосистему***.

Согласно этому определению, множество *C* представляет собой секретный (или личный) ключ, используемый для шифрования и дешифрования исходных сообщений (секретных данных). Кроме того, требование $|C|\geq|M|$ не является строгим.

В отличие от классических формальных стегосистем, где криптоаналитику доступно множество контейнеров, косвенная стегосистема предусматривает возможность доступа лишь к параметрам элементов контейнера. Кроме того, если в классических системах скрыты (от криптоаналитика) либо алгоритмы преобразования (например, E: $C\times M \rightarrow C$, D: $C \rightarrow M$), либо ключи шифрования, либо и то и другое, в предлагаемой системе секретной информацией является содержимое самого контейнера, что позволяет разрабатывать достаточно высокоустойчивые системы защиты информационных ресурсов в компьютерных системах и сетях.

В качестве параметров элементов контейнера могут быть использованы адреса размещения элементов, их цветовые гаммы, форматы, корреляционные показатели и т.п. Для упрощения дальнейшего изложения будем рассматривать адресные параметры элементов контейнера-ключа *C*. Следует иметь в виду, что независимо от того, как заданы значения параметров − прямо или косвенно, они должны адекватно отражать значения элементов множества *M*.

Пусть задан алфавит с конечным множеством букв. Будем считать, что контейнер-ключ *C* формируется из букв алфавита. Расположение букв алфавита в контейнере должно быть произвольным (например, псевдослучайным) с возможностью их многократного вхождения. Каждой букве алфавита ставится в соответствие значение адресного пространства. Совокупность значений адресного пространства составляет множество @*C* (косвенный контейнер). Количество букв алфавита должно быть таковым, что из них можно составить любое сообщение *M*. Таким образом, сообщения подобны контейнеру-ключу *C* в том смысле, что они составляются из одинаковых букв с разным количеством их повторений и разным месторасположением.

Положим, что необходимо передать секретное сообщение *M* по каналу связи. Для этого произвольным образом выбирается первоначальный адрес расположения какого-либо элемента (буквы) в контейнере-ключе. Начиная с этого адреса (в любом направлении) осуществляется поиск первого элемента (буквы) сообщения в массиве элементов (букв) контейнера-ключа. Так как контейнер обязательно содержит все буквы алфавита и каждая буква повторяется в произвольном порядке в массиве элементов контейнера многократно, то поиск завершится успешно. Первая буква сообщения заменяется адресом найденного элемента (буквы) контейнера. Далее в массиве-контейнере осуществляется поиск второго элемента (буквы) сообщения, который замещается адресом найденного элемента (буквы). Процесс повторяется до полного формирования множества адресов. Сформированное таким образом множество адресных данных представляет собой косвенный контейнер @*C*, который отправляется адресату по открытому каналу. Адресат имеет такой же секретный массив *C* (контейнер-ключ), как у отправителя. В отправленном косвенном контейнере также содержатся стеганографические образы начального (стартового) значения адреса поиска, параметры массива сообщений, временного штампа и т.п., т.е. алгоритм расшифровки сообщений является «зеркальным» отображением алгоритма шифровки: по значению первого элемента косвенного контейнера @*C* в контейнере-ключе осуществляется поиск буквы (элемента), адрес которого записан в первом элементе @*C*. Содержимое найденного адреса замещает первый элемент косвенного контейнера @*C*. Далее осуществляется поиск второй буквы и т.д. В конечном итоге буквы множества @*C* будут совпадать с буквами множества *M*.

Рассмотрим конкретный вариант реализации алгоритма косвенной стеганографии. В качестве исходного сообщения *M* будем брать компьютерный файл *F* длиной *L* байтов. Выбираем алгоритм псевдослучайных чисел $\wp(\lambda)$, отвечающий требованиям стойкости генерируемых данных (в настоящее время учеными

разработаны множество таких алгоритмов [3]. Например, повторяемость алгоритма, описанного в [4], составляет примерно 6000 десятичных знаков). Назначаем стартовое число $\lambda = \lambda_0$ для $\wp(\lambda)$. Выбор можно осуществлять либо наугад, либо с помощью простого генератора случайных чисел разового пользования. В первой версии реализованного алгоритма косвенной стеганографии генерация псевдослучайных чисел выполняется следующим образом. Генератор $\wp(\lambda)$, начиная со стартовой точки $\lambda = \lambda_0$, генерирует 2^{20} строк. Каждая строка состоит из 256 байтов. В каждой строке содержатся все двоичные числа от 0 до 255, расположенные случайным образом по закону генератора $\wp(\lambda)$, который гарантирует генерацию неодинаковых чисел в каждой строке. Кроме того, гарантируется отсутствие одинаковых строк в выбранной длине генерируемого массива чисел. Таким образом, формируется двумерный массив случайных чисел *C*(*i, j*), где $i = 256, j = 4096$.

Процесс шифрования файла *F* заключается в следующем. С помощью простого случайного генератора выбирается строка $j = \varpi$ в массиве *C*(*i, j*) (номер строки ϖ также подлежит шифрованию для отправки получателю). Содержимое первого байта [1] файла *F* представляется как адрес байта @[1] в строке $j = \varpi$. Содержимое байта [@[1]] записывается в первый байт файла *F*, т.е. [1]:= [@[1]]. Затем содержимое второго байта [2] файла *F* представляется как адрес байта @[2] в строке $j = \varpi \pm 1$. Содержимое байта [@[2]] в строке $j = \varpi \pm 1$ записывается во второй байт файла *F*, т.е. [2]:= [@[2]]. Процесс повторяется до замещения последнего байта значением массива случайных чисел по выбранному адресу. Таким же способом замещаются значения ряда служебных данных, в том числе значение ϖ. В случае, когда *L*> 2^{20}, процесс может повторяться по кругу.

В реализованной для задач реального времени версии алгоритма косвенной стеганографии, так называемом «алгоритме на лету», нет необходимости повторять процесс по кругу, так как количество генерируемых неповторяемых чисел намного больше, чем объем отправляемых любых файлов по сети. Этот же алгоритм может быть использован не только для задач реального времени, но и для обычных блоковых шифруемых данных.

Безусловно, научное обоснование криптоустойчивости алгоритма косвенной стеганографии требует еще глубокого анализа со стороны криптоаналитиков [5], однако проведенные исследования и полученные экспериментальные результаты позволяют судить о его высокой криптоустойчивости.

Практические особенности реализации косвенной стеганографии таковы.

1. *Проблема распространения ключа* (*передача контейнера C*). Поскольку эта проблема актуальна для всех методов и технологий криптографии с ключами, можно использовать самые передовые алгоритмы и способы распространения ключей. Существенным является тот факт, что, в отличие от других методов, в данном случае требуется разовая гарантия доставки ключа, так как после гарантированного получения ключа адресатом можно при первом же сеансе изменить содержимое контейнера *C*. Поэтому, например, содержимое контейнера можно передать с помощью открытых ключей, длина которых заведомо гарантирует невозможность дешифровки содержимого контейнера (2048, 4096). Безусловно, при этом потребуется намного больше времени для шифрования и дешифрования содержимого контейнера, но в связи с тем, что соответствующие вычисления выполняются один раз, такой способ является оправданным.

2. *Вероятность восстановления содержимого контейнера C по известному криптоаналитику шифру @C.* Прежде всего следует обратить внимание на то, что длина ключа-контейнера *C*, по сравнению с известными методами шифрования с использованием ключей, несравнимо большая ($|C| \geq |M|$). Поэтому восстановление контейнера по значениям становится невозможным. Например, в программно реализованном варианте шифрования компьютерных файлов количество вариантов перебора равно 256! (около 2^{1700} вариантов).

Рис. 4. Косвенный стеганограф

На рис. 4 показано фотографическое изображение варианта реализации алгоритмов компьютерной стеганографии на базе микропроцессорного устройства-ключа с USB-интерфейсом. Ряд модификаций этого устройства позволяет реализовать санкционированный доступ к компьютерам, файлам, а также охраняемым объектам. Кроме того, возможно шифрование «на лету» потоковой информации (речь, аудит, видео), что позволит обеспечить защиту от неавторизованного доступа к таким ресурсам.

Заключение

В предложенном новом методе шифрования, называемом косвенной стеганографией, исходный файл не шифруется, а вместо этого передаются по сети признаки шифруемого файла. Вычислительная сложность алгоритма минимальна, так как шифрование файлов предполагает только замещение байтов исходного файла байтами специально организованного файла-ключа. При этом способе шифрования никакими методами и средствами нельзя расшифровать перехваченный шифр, если даже криптоаналитику удастся получить предыдущий шифр и предыдущий исходный текст.

Литература

1. *Хорошко В.А., Чекатков А.А.* Методы и средства защиты информации / Под ред. Ю.С. Ковтанюка. − К.: Юниор, 2003. − 504 с.
2. *Хорошко В.А., Шелест М.Е.* Введение в компьютерную стеганографию. − К.: НАУ, 2002. − 140 с.
3. *Matsumoto M., Nishimura T.* Mersenne twister: a 623-dimensionally equidistributed uniform pseudorandom number generator // ACM Trans. Model. Comput. Simul. −1994. − N 8. − P. 3−17.
4. *Matsumoto M., Kurita Y.* Twisted GFSR generators // ACM Trans. Model. Comput. Simul. − 1992. − N 2. − P. 179−254.
5. *Шнайер Б.* Прикладная криптография. Протоколы, алгоритмы, исходные тексты на языке Си. − М.: Триумф, 2003. − 815 с.

Информация об авторе

Надир Алишов *– доктор технических наук, ведущий научный сотрудник, Институт кибернетики им. В.М. Глушкова НАН Украины, пр. Глушкова, 40, Киев-03187, Украина; e-mail: anio@ukrtel.com*