КОМП'ЮТЕРНІ ЗАСОБИ, МЕРЕЖІ ТА СИСТЕМИ

УДК 004.056

Н.И. АЛИШОВ, В.А. МАРЧЕНКО, С.Г. ОРУДЖЕВА

КОСВЕННАЯ СТЕГАНОГРАФИЯ КАК НОВЫЙ СПОСОБ ЗАЩИТЫ КОМПЬЮТЕРНЫХ ДАННЫХ

Современные информационные системы имеют распределенную архитектуру, зачастую использующие общие сети (Intranet, Internet, сети коммуникационных операторов) в качестве транспортной инфраструктуры для корпоративной сети. С одной стороны это позволяет значительно сократить затраты на разработку и поддержку таких систем, а с другой – доступ к такой общей транспортной сети имеют лица, не являющиеся членами соответствующей информационной системы. Учитывая это, вопросам контроля данных, циркулирующих в таких сетях, уделяется особое внимание как со стороны эксплуатирующих подобные системы, так и со стороны исследователей. Раскрытие или доступность этих данных стороннему пользователю может привести к непредвиденным последствиям, значительным материальным и нематериальным потерям [1]. Одно из основных направлений в информационных технологиях, позволяющее решить поставленную задачу, это применение различных алгоритмов и методов криптографии [2]. Они позволяют контролировать и ограничивать доступ к передаваемым данным только для лиц, имеющих на это право. В современных информационных системах используется множество различных алгоритмов и методов шифрования информации и последующей её передачи получателю [3]. Все они относятся к двум основным направлениям исследований в области связанной с защитой компьютерной информации от несанкционированного использования: компьютерной криптографии и компьютерной стеганографии.

N.I. Alishov, V.A. Marchenko, S.G. Orugeva

INDIRECT STEGANOGRAPHY AS THE NEW METHOD FOR THE PROTECTION OF COMPUTER DATA

Authors offer a new cryptography scrambling technique of the information named «indirect steganography», its theoretical substantiation is given and some features of application in applied systems are described.

Запропонований новий криптографічний метод шифрування інформації який називається «непряма стеганографія», подано його теоретичне обґрунтування й описані деякі особливості застосування в прикладних системах.

Предложен новый криптографический метод шифрования информации называемый «косвенная стеганография», дано его теоретическое обоснование и описаны некоторые особенности применения в прикладных системах.

1. ***Компьютерная криптография*** это метод, в котором информация, подлежащая защите, шифруется с помощью числовых ключей, причем с увеличением разрядности ключей вычислительная сложность преобразования увеличивается.

Общая схема криптографического шифрования показана на рис. 1.

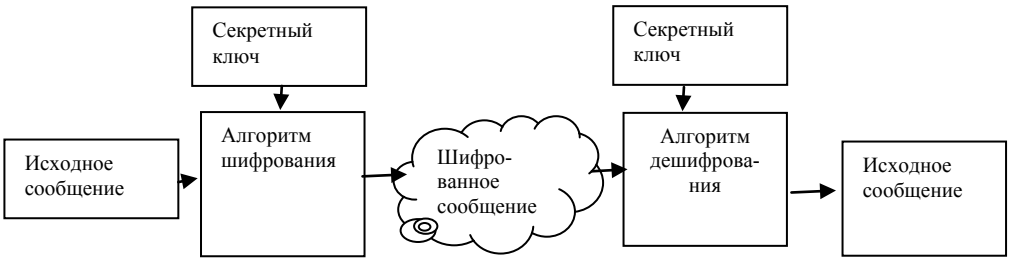

РИС. 1

2. ***Компьютерная стеганография*** это метод, в котором информация, подлежащая защите, смешивается с определенным видом мультимедийной информации (речь, аудио, видео, изображение и т. п.) и передается к законному пользователю. В компьютерной стеганографии трудно реализовать передачу большого объема информации, что очень важно для современных компьютерных сетей.

Общая схема стеганографического преобразования показана на рис. 2.

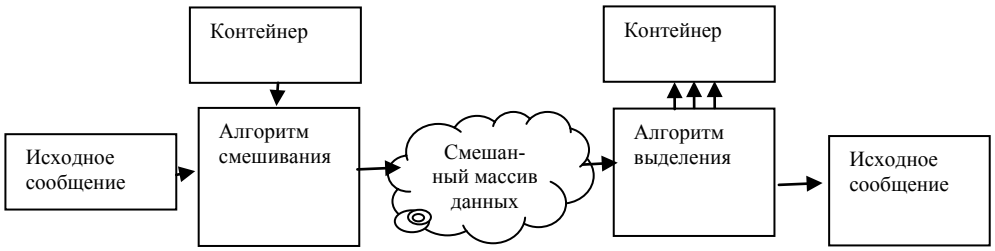

РИС. 2

Формальные определения известных современных стегосистем приведены в [4, 5].

1. Совокупность $Æ = < C, M, D, E >$, где C – множество контейнеров; M – множество секретных сообщений, $|C| \geq |M|$; E: $C \times M \rightarrow C$, D: $C \rightarrow M$ − 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: $C \times M \times K \rightarrow C$, D_k: $C \times K \rightarrow M$ − 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: $C \times M \times K_1 \rightarrow C$, D_k: $C \times K_2 \rightarrow M$ − 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.

Как в компьютерной криптографии, так и в компьютерной стеганографии, в конечном итоге защищаемая информация передается по каналу в шифрованном или смешанном виде, что позволяет криптоаналитику провести соответствующий анализ для взлома шифра и/или выделения полезной информации.

Авторами предлагается новый метод стеганографии, который называется ***косвенной стеганографией.*** Суть метода заключается в следующем. У отправителя и получателя имеются одинаковые массивы данных, которые являются секретными ключами. Байты информации, подлежащие защите, заменяются (по определенному алгоритму) байтами секретного массива данных. Полученный новый массив данных размером исходного сообщения передается адресату. Полученный по каналу массив данных, подвергается обратному преобразованию: байты этого массива данных заменяются байтами секретного массива данных (зеркальный алгоритм).

Общая схема косвенной стеганографии показана на рис. 3.

РИС. 3

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; $E_@$: $C \times M \rightarrow @C$, $D_@$: $C \times @C \rightarrow M$ − 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, то 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 таким, чтобы из них можно было 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*.

В качестве примера рассмотрим простой вариант реализации алгоритма косвенной стеганографии. Сообщение *M* представляет собой компьютерный файл *F*, длина которого равна *l* байтов. Выбираем алгоритм псевдослучайных чисел $\wp(\lambda)$, отвечающий требованиям стойкости генерируемых данных (в настоящее время учеными разработано множество таких алгоритмов [6]. Например, повторяемость алгоритма, описанного в [7], 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 повторяется 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.

Однако проведенные исследования и полученные экспериментальные результаты позволяют судить о его высокой криптоустойчивости.

Следует отметить, что алгоритмы косвенной стеганографии имеют схожие свойства с классом невскрываемых криптоалгоритмов описанных К. Шенноном [8] и достаточно подробно проанализированных позже [9].

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 учитывая, 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).

Использование предложенного метода защиты информации в различных прикладных системах, в зависимости от области применения имеет некоторые особенности. Авторы исследовали применение косвенной стеганографии в компьютерных распределенных сетях в задачах защиты передаваемого трафика, проведения процедур аутентификации и авторизации пользователей, а также в мобильных сетях для решения задач защиты голосового трафика во время разговора между пользователями.

В современных распределённых компьютерных системах для проведения процедур аутентификации и авторизации пользователей в основном используются следующие варианты:

1) использование пары логин\пароль;

2) применение электронных устройств.

Наиболее простой и распространенный метод – использование пары «логин\пароль». Учитывая, что современные вычислительные средства позволяют решать задачу вскрытия простых паролей небольшой длинны за линейное время, используют длинные пароли (длина больше 8–10 символов) которые сложно запоминать. Использование косвенной стеганографии в таких алгоритмах позволяет избежать вышеописанной проблемы, так как передаваемая информация

для системы аутентификации и авторизации пользователей не имеет коррелируемой информации с использованным паролем, в частности его длины и сложности. Таким образом, при пересылке необходимой информации невозможно восстановить пароль по перехваченным сообщениям, так как они будут каждый раз разными. Для согласования ключа контейнера возможным является использование системы «*Диффи – Хеллмана – Меркле*» [10]. В таком случае линии связи должны быть надежно защищены от модификации сообщений при согласовании контейнера-ключа или заданных параметрах его генерации.

Использование электронного устройства при реализации алгоритма косвенной стеганографии позволяет реализовать механизм одноразовых паролей достаточно простым способом. Так как использование этого алгоритма предусматривает наличие контейнера-ключа, который в данном случае будет храниться в памяти электронного устройства, в таком случае он может быть использован как набор одноразовых ключей, для прохождения процедур верификации пользователя. Использования такой реализации процедуры авторизации и аутентификации пользователей не предполагает начального обмена заданными параметрами (контейнера-ключа, начальное зерно генерации и т. п.).

Использование предлагаемого алгоритма защиты передаваемых данных в существующих мобильных сетях сопряжено с решением ряда сложных задач [11]. Такие сети в основной своей массе являются гетерогенными мультисервисными с различными возможностями QoS. Поэтому при внедрении алгоритма защиты передаваемой информации от несанкционированного доступа необходимо обеспечить возможности.

1. *Интероперабельность* алгоритма шифрования и его реализаций. Обеспечивается за счет применения алгоритма на прикладном уровне модели OSI, а также аппаратно-независимым построением схемы шифрования. Таким образом, обеспечивается прозрачность использования алгоритма на разных мобильных сетях.

2. *Потоковый режим* функционирования алгоритма шифрования.

3. *Малые требования к ресурсам,* которые необходимы для эффективного функционирования используемых алгоритмов шифрования.

4. *Высокая криптостойкость и быстрота выполнения операций*. Является наиболее противоречивым требованием, так как в основном в различных криптосистемах криптостойкость системы зависит от сложности решения математической задачи положенной в основу алгоритма или скорости выполнения криптографических операций.

Использование предложенного алгоритма организации защищенного информационного обмена позволяет реализовать быстрые, криптоустойчивые решения, которые могут значительно повысить защищенность информации от несанкционированного доступа. Особенно следует выделить использование данного алгоритма для построения систем идентификации и аутентификации, позволяющие использовать короткие и простые для запоминания пароли.

В предложенном новом методе шифрования исходный файл не шифруется, а вместо него 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 данном способе шифрования никакими методами и средствами нельзя расшифровать перехваченный шифр, если даже криптоаналитику удастся получить предыдущий шифр и предыдущий исходный текст.

1. *Stallings W.* Cryptography and network security: principles and practice. – New York: Prentice Hall, 2006. – 680 p.
2. *Kaufman C., Perlman R., Speciner M.* Network security: private communication in a public world. – Upper Saddle River: Prentice Hall Press, 2002. – 752 p.
3. *Галицкий А.В., Рябко С.Д., Шаньгин В.Ф.* Защита информации в сети – анализ технологий и синтез решений. – М.: ДМК Пресс, 2004. – 616 с.
4. *Хорошко В.А., Чекатков А.А.* Методы и средства защиты информации / Под ред. Ю.С. Ковтанюка. − К.: Юниор, 2003. − 504 с.
5. *Хорошко В.А., Шелест М.Е.* Введение в компьютерную стеганографию.− К.: НАУ, 2002. − 140 с.
6. *Matsumoto M., Nishimura T.* Mersenne twister: a 623-dimensionally equidistributed uniform pseudorandom number generator // ACM Trans. Model. Comput. Simul. 99 9. − N 8. − P. 3−17.
7. *Matsumoto M., Kurita Y.* Twisted GFSR generators // ACM Trans. Model. Comput. Simul. − 1992. − N 2. − P. 179−254.
8. *Шеннон К.* Теория связи в секретных системах // Работы по теории информации и кибернетике. – М.: Изд. иностр. лит., 1963. – С. 333 – 369.
9. *Зубов А.* Совершенные шифры. — М.: Гелиос АРВ, 2003. – 160 с.
10. *Rescorla E.* "Diffie-Hellman Key Agreement Method": RFC2631. – East Palo Alto, RTFM Inc., 1999. – 12 p.
11. *Андрианов В., Соколов А.* Средства мобильной связи. – СПб: BHV, 1999. – 256 с.

Получено 13.08.2009