

УДК 004.056.53

И.Р. Файзуллин, Р.Т. Файзуллин

Аппаратно-эффективный алгоритм формирования маркера начала сообщения

Предложен аппаратно, но не программно, эффективный криптографический алгоритм для формирования маркера начала передачи информации. Предложена модификация алгоритма, позволяющая передавать информацию в виде специально встроенных ошибок в маркер. Показано, что задача, на которой основана криптографическая стойкость алгоритма, не принадлежит классу NP. В наилучшем для атакующего случае дана оценка числа операций, необходимых для нахождения ключа.

Ключевые слова: мастер-ключ, маркер, криптографическая стойкость, автомат без памяти.

Введение

Постоянное увеличение скорости передачи данных и пропускной способности каналов связи приводит к качественному изменению подходов к проблеме защиты информации. Стеганография [1], скрытые каналы [2–3], сегментирование данных [4] из экзотических инструментов становятся одними из наиболее перспективных способов практической защиты информации. В связи с этим остро стоит вопрос о вычислительной эффективности шифроалгоритмов, их доказуемой стойкости и желательном преимуществе аппаратных реализаций алгоритмов перед программными.

В настоящей работе предлагается алгоритм защиты информации, одновременно совмещающий в себе черты стандартных блочных алгоритмов и стеганографии. Оценена стойкость алгоритма и предложен эффективный метод реализации.

Рассмотрим поток данных, состоящий из бит, полученных или с помощью физического генератора псевдослучайных чисел, или с помощью программного датчика псевдослучайных чисел. Задача состоит в том, чтобы встроить в этот поток маркер сообщения, который будет содержать само сообщение. Схожая задача была рассмотрена в работе [5], но там маркер предваряет сообщение и зависит от него, как хэш-код, что представляется не совсем эффективным, т.к. требует офф-лайн – обработки больших объемов данных.

Рассмотрим последовательность байтов Q^1, \dots, Q^N , каждый из которых состоит из K бит, и соответствующий им кортеж бит q^1, \dots, q^N . Будем называть объединение этих кортежей мастер-ключом.

Передающая сторона, или Алиса, генерирует кортеж X^1, \dots, X^N с условиями:

$$\begin{aligned} X^j &\geq Q^j, \text{ если } q^j = 0, \\ X^j &< Q^j, \text{ если } q^j = 1 \end{aligned} \quad (1)$$

и маркирует начало сообщения операцией побитового сложения байтов

$$X^j, Q^j: Z_i^j = X_i^j \oplus Q_i^j \quad i=1, \dots, K.$$

Получатель сообщения, или Боб, производим аналогичную операцию XOR над байтом Z^j и байтом Q^j , получая в итоге \bar{X}^j . Если некоторая последовательность битов $\bar{X}^m, \dots, \bar{X}^{m+N-1}$ удовлетворяет условиям (1), тот Боб делает вывод о том, что с вероятностью $1-2^{-N}$ эта последовательность байтов маркирует начало сообщения.

Но заметим, что Алиса может специально генерировать некоторые из X^j так, что условия (1) для них не будут выполняться. Если число таких «ошибок» будет относительно мало, например не более четверти, и окаймлено заранее заданным числом «неошибок», то Боб, проверяя входной поток, может сделать хорошо обоснованный вывод о наличии специально встроенных Алисой символов «ошибки». Боб должен выбрать из двух гипотез: он получает случайный поток данных или поток данных, созданный Алисой. Учитывая, что случайный поток данных, например созданный с помощью генератора псевдослучайных чисел BBS, хорошо описывается схемой Бернулли, а при увеличении N хорошо описывается нормальным законом, то выделение кортежа

$\bar{X}^m, \dots, \bar{X}^{m+N-1}$ не представляет большого труда, с учетом, например, того, что доверенное лицо проводит тест на открытый текст для наиболее вероятного фрагмента потока данных.

В этом случае Боб может интерпретировать маркер как битовую строку \bar{q}^j , $j=1, \dots, N$, где нулями являются те из \bar{q}^j , которые совпадают с q^j , а единицами – места «ошибок». Длина слова в этом случае будет равна N , а «полезное пространство сообщений» 2^{N-l} , где l сравнимо с N .

Насколько стойкой является предложенная схема шифрования?

Предположим, что аналитик каким-то образом смог выделить несколько кортежей Z^{js} , $j=1, \dots, N$, $s=1, \dots, M$, $M \geq 2$, отвечающих маркерам, и ему известно, что в маркерах нет «ошибок».

С учетом независимости блоков аналитик может рассматривать только один блок j и отвечающие ему неизвестные $q^j, Q^j, X^{j1}, \dots, X^{jM}$.

В этом случае мы получаем систему линейных уравнений в Z_2 и неравенств в R :

$$z^{js}_i = x^{js}_i \oplus Q^{js}_i \quad i=1, \dots, K, \quad s=1, \dots, M,$$

$$(X^{js} - Q^j)(-1)^{q^j} \geq 0.$$

Предполагая, что, например, $q^s = 0$ получим:

$$(X^{js} - Q^{js}) \geq 0.$$

Попытка решить эту систему прямыми методами обречена на неудачу, т.к. ассоциированные системы линейных алгебраических уравнений имеют ядро размерности, сравнимой с K , что делает задачу теоретически неразрешимой.

Предположим, что мы точно знаем расположение M кортежей Z^j в потоке данных и тем самым знаем $z^{js}_i = x^{js}_i \oplus Q^{js}_i$ $i=1, \dots, K$, $s=1, \dots, M$, $j=1, \dots, N$ для каждого j, s . Зафиксируем некоторое j и оценим трудоемкость определения битов Q^j_1, \dots, Q^j_K при известном q .

Оказывается, что в этом случае верна следующая лемма.

Лемма 1:

В наилучшем для атакующего случае, трудоемкость определения мастер-ключа Q^j_1, \dots, Q^j_N и q^1, \dots, q^N не меньше чем $CNK2^K$.

Рассмотрим все возможные значения Q^j , которые, возможно, являются битами маркера без ошибок, очевидно, что их число оценивается величиной 2^K . Для каждого такого Q^j по известному Z^j однозначно определяется X^j и проверяется, например, условие $(X^j - Q^j) \geq 0$. Заметим, что априори нельзя выделить множество таких пробных Q^j , для которых всегда будет выполняться только условие $(X^j - Q^j) \geq 0$, кроме, конечно, тривиального случая нулевого вектора. Поэтому мы вынуждены проверять условие в $2^K - 1$ случаях.

По результатам проверки Q^j заносится в одно из множеств $A = \{Q^j, X^j - Q^j \geq 0\}$ или $B = \{Q^j, X^j - Q^j < 0\}$.

Выберем из M кортежей другой байт Q^j и также образуем множества $A' = \{Q^j, X^{j1} - Q^j \geq 0\}$, $B' = \{Q^j, X^{j1} - Q^j < 0\}$. Предположим, что все пересечение $A \cap A'$ состоит только из одного H^j , а пересечение $B \cap B'$ пусто.

В этом случае H^j будет искомым байтом, а искомым бит q^j будет определяться однозначно.

Даже в этом простейшем случае мы вынуждены решить системы два раза, для каждого кортежа, и число операций побитового сложения для определения X^j пропорционально K , т.е. длине байта.

Циклический сдвиг кортежа q^1, \dots, q^N и встраивание «ошибок» или сообщения решают эту проблему практически, но лемма по существу остается верной. Мы с вероятностью 0,25 «успешно» попадаем в отрезок Голомба [6] длины больше чем единица. Отметим, что мы рассматриваем идеальный случай, в котором выборка и распределение данных в памяти производятся мгновенно.

Из леммы следует другой практический вывод о том, что гарантированная на сегодняшний момент стойкость, равная 2^{80} , без учета операций записи и считывания из памяти, будет обеспечиваться выбором длины байта K , большей или равной 64.

Также можно сформулировать очевидную лемму о классе сложности, к которому принадлежит задача криптоанализа данного алгоритма.

Лемма 2:

Задача определения мастер-ключа при известных кортежах Z^j не принадлежит классу NP.

Допустим, что нам известны несколько кортежей Z^j и необходимо проверить, что наборы Q_1^j, \dots, Q_N^j и q^1, \dots, q^N являются ключом. Покажем, что даже если все $X_i^j = Z_i^j \oplus Q_i^j$ $i=1, \dots, K$ удовлетворяют (1) это не означает, что Q_1^j, \dots, Q_N^j и q^1, \dots, q^N мастер-ключ. Допустим, что Z^j и Q^j на нескольких общих позициях имеют совпадающие нулевые биты. Тогда \bar{Q}^j , полученное произвольной расстановкой единиц в этих позициях, так же позволит сгенерировать $\bar{X}_i^j = Z_i^j \oplus \bar{Q}_i^j$ $i=1, \dots, K$ удовлетворяющие неравенствам (1). Единственный выход – это увеличение числа M в общем случае до величины 2^r , где r – это число нулевых бит, стоящих в общих позициях у Z^j и Q^j .

Только в этом случае можно выбрать единственное Q^j из всех возможных \bar{Q}^j . Доказательство завершено.

Заметим, что реализация кодирования информации, или создание последовательности X^1, \dots, X^N , может быть осуществлена инвертированием бит, например, применением операций И, ИЛИ с битами байта Q и содержимым некоего регистра сдвига с линейной обратной связью. Если, например, $q^j = 0$, то некоторые нулевые биты Q^j необходимо инвертировать на единичные (ИЛИ), если же $q^j = 1$, то, наоборот, необходимо инвертировать единичные биты Q^j (И).

Принимающая сторона может определить набор q^1, \dots, q^N , пропуская его через регистр сдвига и тестируя байты на предмет проверки $(X^{js} - Q^j)(-1)^{q^j} \geq 0$, также с помощью автомата без памяти (конъюнктивную нормальную форму КНФ), предложенного в [7].

Обратим внимание на то немаловажное обстоятельство, что аппаратное шифрование и дешифрование происходит за $O(1)$ время, необходимое для прохождения сигнала через автомат, реализующий КНФ, а программная реализация требует $СКN$ числа операций. Это позволяет сделать вывод о перспективности использования данного подхода в практике корпораций и неэффективности программных реализаций частными лицами.

Компьютерное моделирование формирования и распознавания маркера показало, что система работоспособна и не возникает гонок автоматов.

Можно сделать вывод о том, что разработан эффективный алгоритм формирования маркера начала сообщения, доказана его криптографическая стойкость, основанная на решении задачи из класса EXP , и определены нижние пределы применимости (по длине ключа). Проведено компьютерное моделирование работы протокола и предложены схемы аппаратной реализации.

Литература

1. Грибунин В.Г. Цифровая стеганография / В.Г. Грибунин, И.Н. Оков, И.В. Туринцев. – М.: Солон-Пресс, 2002.
2. Столлингс В. Криптография и защита сетей: принципы и практика. – 2-е изд. – М.: Вильямс, 2003.

3. Методы защиты информации от атак с помощью скрытых каналов и враждебных программно-аппаратных агентов в распределенных системах / А.А. Грушо, Н.А. Грушо, Е.Е. Тимонина // Вестник РГГУ. – 2009. № 10. – С. 33–45.
4. Коначович Г.Ф., Пузыренко А.Ю. Компьютерная стеганография. Теория и практика. – К.: МК-Пресс, 2006. – 288 с.
5. Файзуллин И.Р. Крипто-стеганографический алгоритм с использованием хэш-функции для маркировки начала сообщения // Научная сессия МИФИ–2007. – Т. 16. Компьютерные науки. Информационные технологии. – М.: МИФИ, 2007. – С. 149–150.
6. Golomb S.W. Run-length encodings // IEEE Trans. Inf. Theor. – 1996. – IT-12, № 3. – P. 399–401.
7. Дулькейт В.И. Алгоритм минимизации функционала, ассоциированного с задачей 3-SAT, и его практические применения // В.И. Дулькейт, Р.Т. Файзуллин, И.Г. Хныкин // Компьютерная оптика. – 2008. – Т. 32, № 1. – С. 68–73.

Файзуллин Ильдар Рашитович

Начальник отдела функциональной разработки Департамента SAP,
Компания ЗАО «ТД Перекресток»
Тел.: +7 (495) 662-88-88
Эл. почта: blackildar@list.ru

Файзуллин Рашит Тагирович

Проф. каф. средств связи и информационной безопасности
Омского государственного технического университета
Тел.: +7(381-2) 73-03-34
Эл. почта: r.t.faizullin@mail.ru

Faizullin I.R., Faizullin R.T.

Hardware effective algorithm for mark formation of the data transmission beginning

Hardware (not software) effective cryptographic algorithm for mark formation of the data transmission beginning. A modification of the algorithm, which allows the data transmitting in the form of errors specially built-in marker, is suggested. It is shown that the problem, on which the algorithm cryptographic firmness is based, does not belong to NP class. The estimation of a number of operations needed for finding a key is given for the best (in the viewpoint of an attacker) case.

Keywords: master-key, marker, cryptographic firmness, memoryless automaton.
