Секция 3

МАТЕМАТИЧЕСКИЕ МЕТОДЫ КРИПТОГРАФИИ

УДК 519.113.6 DOI 10.17223/2226308X/8/20

ШИФРЫ С ВОДЯНЫМИ ЗНАКАМИ

Г. П. Агибалов

Для защиты конфиденциальности и легальности данных вводится понятие шифра с водяным знаком (называемого также w-шифром). Его основная идея следующая: преобразование открытого текста x композицией операций шифрования и расшифрования с использованием соответствующих ключей приводит к некоторому подходящему тексту x', сохраняющему информацию текста x и содержащему некоторый уникальный водяной знак w, идентифицирующий подлинного владельца x'. Ключи зашифрования и расшифрования в w-шифре должным образом связаны друг с другом и с заданным водяным знаком w. В отличие от шифров, обычно изучаемых в криптографии, функция шифрования в w-шифре не обязательно обратимая. Таким образом, фактически w-шифры не являются шифрами в известном смысле этого слова, но шифры суть w-шифры некоторого частного вида, и все термины, понятия и обозначения, относящиеся к шифрам, полностью применимы к w-шифрам. Показано, как применением w-шифра можно осуществить встраивание водяного знака в данные в процессе зашифрования открытого текста либо в процессе расшифрования шифртекста. Приводятся примеры w-шифров, построенных на базе симметричных поточных шифров.

Ключевые слова: *защита данных, шифрование, водяные знаки, шифры с водяными знаками, поточные шифры.*

Введение

Методы шифрования данных и внедрения в них водяных знаков принадлежат различным областям науки — криптографии [1, 2] и стеганографии [3, 4] соответственно. Первые применяются для защиты конфиденциальности информации, вторые — для защиты информации от её нелегального использования. Как правило, шифрование данных является обратимым преобразованием с ключом расшифрования, неизвестным злоумышленнику, а внедрение в данные водяного знака есть операция сокрытия последнего в данных для идентификации автора нелегальной копии данных. Предполагается, что это сокрытие производится без существенного искажения защищаемых данных, как это делается, например, при внедрении водяного знака в цифровое видео. Рассматривается проблема обеспечения защиты данных от обеих указанных угроз. Есть два тривиальных способа решить эту проблему: сначала внедрить водяной знак в данные и затем зашифровать полученный открытый текст, либо, наоборот, сначала зашифровать данные и затем после расшифрования внедрить в полученный текст нужный водяной знак. Имеются, однако, серьёзные ограничения к применению этих способов на практике [4]. В частности, второй способ предполагает доверенного получателя данных, каковым не является, например, покупатель тех же видеоданных.

В работе вводится понятие шифра с водяным знаком, называемого, для краткости, w-шифр, или, в англоязычной транскрипции, a watermarking cipher. В нём преобразование открытого текста x композицией алгоритмов зашифрования и расшифрования, использующих некоторые, должным образом подобранные ключи зашифрования и расшифрования, порождает текст x', содержащий заданный водяной знак w. Мы показываем, каким образом с помощью w-шифра возможно внедрение водяного знака в данные (в тайне от их получателя, естественно) как в процессе зашифрования данных отправителем, так и в процессе расшифрования данных их получателем. Функции зашифрования и расшифрования в w-шифре не обязательно связаны между собой отношением обратимости, как в криптографических шифрах. Это значит, что в действительности w-шифры не обязаны быть шифрами, но всякий шифр является частным случаем w-шифра (с вырожденным водяным знаком), и все термины, понятия и обозначения, относящиеся к шифрам, вполне уместны и в применении к w-шифрам.

Далее, прежде чем дать общее определение шифров с водяными знаками и описать некоторые конкретные примеры их, мы сформулируем некоторые допущения и предположения, необходимые для того, чтобы сделать это более или менее корректно и понятно.

1. Проблема w-шифрования

Прежде всего, предположим для простоты, что защищаемые данные представлены конечной последовательностью (строкой, словом) символов, являющихся элементами аддитивной группы G с операцией сложения $+$. Например, $G = \mathbb{Z}_n$ или $G = \mathbb{Z}_2^n$ для некоторого $n \geqslant 2$, и т. п. В частности, данные могут быть представлены последовательностью бит (возможно, с некоторой структурой). Для любых $a = a_1a_2 \dots a_r$ и $b = b_1b_2 \dots b_r$ в G^r пусть $a + b = (a_1 + b_1)(a_2 + b_2) \dots (a_r + b_r)$, $-b = (-b_1)(-b_2) \dots (-b_r)$ и $a - b = a + (-b)$.

Предполагается, что водяной знак является некоторой парой $w = (v, \eta)$, где $v = v_1v_2 \dots v_m \in (G \setminus \{0\})^m$ и $\eta = i_1i_2 \dots i_m$, $i_j \in \{1, \dots, l\}$, $j = 1, \dots, m$, $1 \leqslant i_1 < i_2 < \dots < i_m \leqslant l$ для некоторого $l \geqslant 1$. В случае необходимости он обозначается $(v, \eta)_l$. Водяной знак $w' = (v', \eta)$, в котором $v' = -v$, называется инверсией знака w и обозначается $-w$. Очевидно, $-(-w) = w$. В случае $|G| = 2$ считаем, что $G = \mathbb{Z}_2$. В этом случае в любом водяном знаке $w = (v, \eta)$ слово v есть вектор $11 \dots 1$, так что w однозначно определяется набором η, и мы пишем $w = \eta$.

Встраивание водяного знака w в строку данных $x = x_1x_2 \dots x_l \in G^l$ выполняется с помощью операции сложения, определённой на G. Результатом встраивания является строка данных $x' = x'_1x'_2 \dots x'_l$, в которой $x'_j = x_j + v_t$, если $j = i_t \in J = \{i_1, \dots, i_m\}$, и $x'_j = x_j$, если $j \in \{1, \dots, l\} \setminus J$. Говорим, что строка x' есть строка x с внедрённым знаком w, и обозначаем её $x + w$. Мы пишем также $x - w$ вместо $x + (-w)$. Строка v и набор η называются соответственно значением и местоположением знака w в x'. Фактически, числа $i_1, i_2, \dots, i_m$ в η указывают позиции в x для встраивания $v_1, v_2, \dots, v_m$ соответственно из значения v знака w. Набор η называется подходящим местоположением для w в x, если x' получается из x без заметной потери информации. В этом случае мы называем x' производной (или копией) от x с корректно (или приемлемо) встроенным (внедрённым) водяным знаком w.

Например, если x является битовой строкой цифрового видео и $v = 11 \dots 1 \in \mathbb{Z}_2^m$, то встраивание знака w в x состоит в инвертировании бит $x_{i_1}, x_{i_2}, \dots, x_{i_m}$. В этом случае если позиции бит $i_1, i_2, \dots, i_m$ выбраны так, что инверсия этих бит в x заметно не разрушает видео, то полученная битовая строка x' является копией строки x с при-

емлемо внедрённым водяным знаком w, обе x' и x могут равнозначно использоваться как цифровое видео, но x', кроме того, содержит водяной знак для идентификации потенциального злоумышленника.

Мы говорим, что водяной знак и строка данных взаимно подходящие, т. е. w подходит для x и наоборот, если x имеет экспоненциальное количество подходящих местоположений для w в x. Здесь под экпоненциальным количеством подразумевается экспоненциальная функция от длины m знака w. Такое число походящих местоположений предотвращает злоумышленника от атаки грубой силы перечислением всех возможных подходящих местоположений в x'. Так, цифровые аудио- и видеоданные являются двумя примерами битовой строки данных, для которой встраивание водяного знака путём инверсии битов в некоторых позициях является подходящим.

Кроме того, мы предполагаем, что существуют производитель (DP) строки данных x и её покупатели, или клиенты (DC). Производитель DP хочет передать строку x некоторому DC U так, что никто другой не может перехватить x или секретно получить её в своё собственное владение от U. С этой целью DP должен выбрать уникальный и подходящий водяной знак w и некоторый ключ шифрования k_e для некоторого w-шифра C, зашифровать x, применив C и k_e, и послать клиенту U полученный так шифртекст y и нужный ключ расшифрования k_d, построенный таким образом, что расшифрование y на этом ключе даёт в качестве результата некоторую строку данных x', которая является производной от x, с корректно внедрённым знаком w. При этом неважно, на какой стадии, во время шифрования или расшифрования, w встроен в x. Расшифровывая y на ключе k_d, клиент U получает уникальную и приемлемую копию x' данных x. Если U передаст её другому клиенту, то DP сможет однозначно идентифицировать U по значению v из w и его местоположению η в x'.

Поскольку U сам может быть мошенником, ключ расшифрования k_d должен быть связан с водяным знаком w так, что определить w по k_d и шифртексту y вычислительно невозможно за реальное время, т. е. не существует алгоритма вообще или с полиномиальной сложностью (как функции от m), вычисляющего w из k_d и y.

2. Определение w-шифра

Итак, мы приходим к следующему понятию w-шифра: для любых взаимно подходящих водяного знака w и открытого текста x преобразование последнего композицией операций зашифрования и расшифрования на соответствующих ключах, связанных некоторым образом друг с другом и с w, создаёт текст $x' = x + w$, представляющий собой результат внедрения w в x: На этом пути мы вводим два типа w-шифров:

1) w-шифр с w-расшифрованием — открытый текст x зашифровывается в зависимости только от ключа k w-шифра, а полученный шифртекст y расшифровывается в зависимости от k и подходящего водяного знака w. Таким образом, значение k_e ключа зашифрования может быть произвольным, ключ расшифрования k_d должен быть предопределён выбранными k_e и w, т. е. быть функцией от k и w;
2) w-шифр с w-зашифрованием — открытый текст x зашифровывается в зависимости от ключа k w-шифра и подходящего водяного знака w, а полученный шифртекст y расшифровывается в зависимости только от k. Таким образом, ключ зашифрования k_e должен быть функцией от k и w, ключ расшифрования k_d должен быть функцией только от k.

Формально w-шифр определяется набором из шести объектов $C = (X, K, W, h, E, D)$, где X есть множество строк данных, включая открытые тексты, шифртексты и тек-

сты с встроенными водяными знаками, $X = G^*$; K и W суть множества ключей и водяных знаков соответственно; h есть ключевая функция, $h : K \times W \to K$, и E и D суть алгоритмы зашифрования и расшифрования соответственно, являющиеся некоторыми отображениями $E : X \times K \to X$ и $D : X \times K \to X$, такими, что для любых взаимно подходящих $x \in X$ и $w \in W$ и для любого $k \in K$ удовлетворяются следующие условия:

1) в w-шифре с w-расшифрованием —

$$\text{если } E(x,k) = y, \text{ то } D(y, h(k,w)) = x' = x + w;$$

2) в w-шифре с w-зашифрованием —

$$\text{если } E(x, h(k,w)) = y, \text{ то } D(y,k) = x' = x + w.$$

В случае $h(k,w) = k$ для любых $k \in K, w \in W$ мы допускаем вместо k писать $h(k,w)$ в последних выражениях и Λ вместо h в C.

3. Примеры w-шифров

Тривиальный пример w-шифра (X, K, W, Λ, E, D) над G можно построить из симметричного шифра (X, Y, K, E', D') с $X = Y = G^*$ и множества W водяных знаков, положив $E(x,k) = E'(x + w, k)$ и $D(y,k) = D'(y,k)$ или $E(x,k) = E'(x,k)$ и $D(y,k) = D'(y,k) + w$.

Простейшим нетривиальным примером w-шифра является одноразовый блокнот с водяным знаком $C_1 = (X, K, W, h, E, D)$, где $X = K = G^*$. В этом w-шифре с w-расшифрованием для данного водяного знака $w = (v, \eta)$ шифртекст $y = y_1 y_2 \dots y_l \in X$ получается сложением открытого текста $x = x_1 x_2 \dots x_l \in X$ и ключа $k = z_1 z_2 \dots z_l \in K$, т. е. $y = x + k$, и расшифрование y в открытый текст $x' = x'_1 x'_2 \dots x'_l \in X$ с водяным знаком w производится вычитанием другого ключа $k' = k - w = z'_1 z'_2 \dots z'_l \in K$ из y, т. е. $x' = y - k'$.

Этот же w-шифр с w-зашифрованием описывается соотношениями $k' = k + w$, $y = x + k'$, $x' = y - k$.

Непосредственно проверяется, что в обоих случаях $x' = x + w$. В первом случае $k_e = k$, $k_d = h(k,w) = k'$ и знак w автоматически встраивается в x в процессе расшифрования. Во втором случае это делается в процессе зашифрования и $k_e = k' = h(k,w)$, $k_d = k$.

Другими словами, для любых $l \geqslant 1$, $x, k \in G^l$ и $w \in W$

1) в C_1 с w-расшифрованием —

$$E(x,k) = x + k = y,\ h(k,w) = k - w,\ D(y, h(k,w)) = y - h(k,w) = y - k + w = x';$$

2) в C_1 с w-зашифрованием —

$$h(k,w) = k + w,\ E(x, h(k,w)) = x + h(k,w) = x + k + w = y,\ D(y,k) = y - k = x'.$$

Ещё одним примером w-шифра является поточный шифр с водяным знаком $C_A = (X, K, W, h, E, D)$ над конечным полем F с $X = F^*$ и с генератором ключевого потока, являющимся некоторым конечным автономным автоматом A с нелинейной функцией выходов. Автомат A представляется четвёркой объектов $A = (Q, Z, g, f)$, где Q, Z суть множества состояний и выходных символов соответственно, $Q = F^n$, $n \geqslant 1$, $Z = F$ и g, f суть функции переходов и выходов автомата A, $g : Q \to Q$, $f : Q \to Z$. Предполагается, что функция выходов f является непременно частью ключа k w-шифра.

Иногда начальное состояние $q(1)$ автомата A и его функция переходов g могут быть другими частями этого ключа. Далее, для общности, произвольный ключ в K обозначается знаком $k[q(1), g, f]$, подразумевающим f обязательной и $q(1)$ и g опциональными составляющими ключа. Предполагается также, что в A для любого начального состояния $q(1) \in Q$ и целого $l \geqslant 1$ состояния $q(t) = g^{t-1}(q(1))$, $t = 1, 2, \ldots, l$, все различные. В этом случае для любого $w = (v, \eta)_l \in W$ с $v = v_1 v_2 \ldots v_m$ и $\eta = i_1 i_2 \ldots i_m$ можно определить функцию $\delta_{w,q(1),l} : Q \to Z$ таким образом, что для любого $s \in Q$, $\delta_{w,q(1),l}(s) = v_j$, если $s = q(i_j)$, $j \in \{1, 2, \ldots, m\}$, и $\delta_{w,q(1),l}(s) = 0$ в противном случае, т. е. если $s = q(t)$, $t \in \{1, 2, \ldots, l\} \setminus \{i_1, i_2, \ldots, i_m\}$. Ключевая функция h, алгоритмы шифрования и расшифрования E, D и ключи k_e, k_d в C_A определяются в каждом из двух возможных случаев следующим образом:

1) случай w-расшифрования —

$$E(x, k) = E(x_1 x_2 \ldots x_l, k[q(1), g, f]) = y_1 y_2 \ldots y_l = y,$$
$$\text{где } \; y = x + z, \; z = z_1 z_2 \ldots z_l, \; z_t = f(g^{t-1}(q(1))), \; t = 1, 2, \ldots, l;$$
$$h(k, w) = h(k[q(1), g, f], (v, \eta)_l) = k[q(1), g, f_1], \text{ где } \; f_1 = f - \delta_{w,q(1),l};$$
$$D(y, k[q(1), g, f_1]) = D(y_1 y_2 \ldots y_l, k[q(1), g, f_1]) = x'_1 x'_2 \ldots x'_l = x',$$
$$\text{где } \; x' = y - z', \; z' = z'_1 z'_2 \ldots z'_l, \; z'_t = f_1(g^{t-1}(q(1))), \; t = 1, 2, \ldots, l;$$

2) случай w-зашифрования —

$$h(k, w) = h(k[q(1), g, f], (v, \eta)_l) = k[q(1), g, f_2], \text{ где } \; f_2 = f + \delta_{w,q(1),l};$$
$$E(x, h(k, w)) = E(x_1 x_2 \ldots x_l, k[q(1), g, f_2]) = y_1 y_2 \ldots y_l = y,$$
$$\text{где } \; y = x + z', \; z' = z'_1 z'_2 \ldots z'_l, \; z'_t = f_2(g^{t-1}(q(1))), \; t = 1, 2, \ldots, l;$$
$$D(y, k) = D(y_1 y_2 \ldots y_l, k[q(1), g, f]) = x'_1 x'_2 \ldots x'_l = x',$$
$$\text{где } \; x' = y - z, \; z = z_1 z_2 \ldots z_l, \; z_t = f(g^{t-1}(q(1))), \; t = 1, 2, \ldots, l.$$

В обоих случаях непосредственно проверяется, что $x' = x + w$. Кроме того, в первом случае $k_e = k[q(1), g, f]$ и $k_d = k[q(1), g, f_1]$, во втором случае $k_e = k[q(1), g, f_2]$ и $k_d = k[q(1), g, f]$.

Наконец, опишем шифр с водяными знаками $C_R = (X, K, W, h, E, D)$, являющийся конкретизацией w-шифра C_A, в которой автомат $A = (Q, Z, g, f)$ представляет собой нелинейный фильтрующий генератор ключевого потока [2], построенный из регистра сдвига с линейной обратной связью (LFSR) R некоторой длины n с примитивным характеристическим полиномом $c_0 + c_1 u + \ldots + c_{n-1} u^{n-1} - u^n$ в $\mathbb{Z}_2[u]$ и нелинейной булевой фильтрующей функцией f от n переменных. Таким образом, $F = \mathbb{Z}_2$, $X = \mathbb{Z}_2^*$, в любом $w = (v, \eta) \in W$ строка v есть вектор $11\ldots1$, так что $w = \eta = i_1 i_2 \ldots i_m$, $Q = \mathbb{Z}_2^n$, $Z = \mathbb{Z}_2$ и для любого $s = s_0 s_1 \ldots s_{n-1} \in Q$ имеет место $g(s) = s_1 \ldots s_{n-1} s_n$, где $s_n = c_0 s_0 + c_1 s_1 + \ldots + c_{n-1} s_{n-1}$.

Поскольку в $\mathbb{Z}_2$ операции сложения и вычитания совпадают со сложением по mod 2 и сложение с 1 означает инверсию, следующие соотношения верны в C_R: 1) если $q(1) \neq \neq 00\ldots0$ и $l \leqslant 2^n - 1$, то $\delta_{w,q(1),l}(s) = \sum\limits_{j=1}^{m} s^{q(i_j)}$, где для $\sigma = \sigma_0 \sigma_1 \ldots \sigma_{n-1} \in \mathbb{Z}_2^n$ справедливо: $s^\sigma = s_0^{\sigma_0} \wedge s_1^{\sigma_1} \wedge \ldots \wedge s_{n-1}^{\sigma_{n-1}}$, $s_t^{\sigma_t} = \neg s_t$, если $\sigma_t = 0$, и $s_t^{\sigma_t} = s_t$, если $\sigma_t = 1$, $t = 0, 1, \ldots, n-1$; 2) $f_1 = f_2$; 3) алгоритмы зашифрования и расшифрования в случае w-зашифрования являются алгоритмами соответственно расшифрования и зашифрования в случае w-расшифрования.

w-Шифр C_R со встраиванием водяного знака в процессе расшифрования реализован и протестирован на MPEG-видеоданных. Информацию об этом см. в [5, 6].

ЛИТЕРАТУРА

1. *Stinson D. R.* Cryptography. Theory and Practice. CRC Press, 1995. 434 p.
2. *Menezes A., van Oorshot P., and Vanstone S.* Handbook of Applied Cryptography. CRC Press, 1997. 662 p.
3. *Langelaar G. C.* Real-time Watermarking Techniques for Compressed Video Data. Delft: Delft University of Technology, 2000. 155 p.
4. *Mistry D.* Comparison of digital water marking methods // Intern. J. Comp. Sci. Engin. 2010. V. 2. No. 9. P. 2905–2909.
5. *Анжин В. А.* Метод защиты от нелегального копирования в цифровых видеотрансляциях через внедрение водяных знаков при расшифровании // Прикладная дискретная математика. Приложение. 2014. №. 7. С. 73–74.
6. `https://github.com/anjin-viktor/mpeg2decwtrk/` — Method implementation for MPEG2 Video. 2014.

УДК 004.056.55 DOI 10.17223/2226308X/8/21

ПОСТРОЕНИЕ КРИПТОСИСТЕМЫ С ОТКРЫТЫМ КЛЮЧОМ НА ОСНОВЕ ПОЛНОСТЬЮ ГОМОМОРФНОГО ШИФРОВАНИЯ[1]

В. В. Егорова, Д. К. Чечулина

Работа посвящена изучению практической применимости схемы полностью гомоморфного шифрования, созданной в Лаборатории современных компьютерных технологий НИЧ НГУ. Рассмотрено приложение гомоморфного шифрования для построения криптосистемы с открытым ключом, основанной на алгоритме RSA. На примере этой криптосистемы продемонстрирована корректность выполнения арифметических операций над зашифрованными данными, а также отсутствие увеличения размерности зашифрованных сообщений при умножении.

Ключевые слова: *гомоморфное шифрование, криптосистема с открытым ключом, алгоритм RSA.*

В Лаборатории современных компьютерных технологий НИЧ НГУ в рамках проекта «Защищённая база данных» разработана и реализована схема полностью гомоморфного шифрования, позволяющая выполнять операции сложения и умножения над зашифрованными данными. Рассмотрим подробнее эту схему. Пусть требуется шифровать целые числа размера t бит. Для этого необходимо выбрать целое число — модуль m, по которому будут производиться все вычисления в схеме. Модуль является частью секретного ключа. Для того чтобы однозначно восстановить любое зашифрованное число, модуль должен удовлетворять условию $2^t < m$.

Кроме того, для шифрования требуется секретный вектор $k \in \mathbb{Z}^n$, который строится следующим образом. Сгенерируем матрицу W размера $n \times n$, обратимую по модулю m, а также вектор $u \in \mathbb{Z}^n$, компоненты которого по модулю не превосходят m. Вектор k определим как решение системы линейных уравнений

$$(W \cdot k) \bmod m = u,$$

которая всегда разрешима, так как матрица W обратима по модулю m. Таким образом, $k = (W^{-1} \cdot u) \bmod m$. Матрица W и вектор u также являются частью секретного ключа.

Перейдём к описанию алгоритма шифрования. Пусть $p < 2^t < m$ — целое число,

[1]Работа поддержана грантом Минобрнауки РФ, договор № 02.G25.31.0054.