МАТЕМАТИЧЕСКИЕ МЕТОДЫ КРИПТОГРАФИИ

UDC 519.113.6 DOI 10.17223/20710410/31/6

WATERMARKING CIPHERS

G. P. Agibalov

National Research Tomsk State University, Tomsk, Russia

In order to protect both data confidentiality and legality, a concept of a watermarking cipher (also called a w-cipher) is defined. The main idea of this cjncept is as follows: the transformation of a plaintext x by the composition of encryption and decryption operations using some encryption and decryption keys yields a proper text x' containing a unique watermark w. The encryption and decryption keys in the w-cipher are connected with each other and with the given watermark w in some way. In contrast with the ciphers usually studied in cryptography, the encryption function in a w-cipher is not compulsorily invertible. Thus in fact w-ciphers are not ciphers in the known sense of the word, but the ciphers are w-ciphers of a certain partial type, and all terms, notions and notations related to ciphers are quite applicable to w-ciphers. It is shown how data watermarking can be performed by applying a w-cipher in such a way that the concealment of a watermark into a plaintext is accomplished by this w-cipher either in the encryption or in the decryption processes. Some examples of w-ciphers constructed on the basis of symmetric stream ciphers are presented in the paper.

Keywords: *data protection, encryption, watermarks, watermarking ciphers, stream ciphers.*

Introduction

The methods for data encryption and data watermarking belong to different subjects of science — to cryptography [1, 2] and steganography [3, 4], respectively. The first methods are used for protecting information confidentiality, the second — for protecting information from its illegal usage. As a rule, data encryption is an invertible transformation of the data by a cipher with a decryption key unknown to malefactors, and data watermarking is a concealment of a unique watermark into the data for identifying the author of a data illegal copy. The problem under consideration is to provide data protection from both of the mentioned threats. The evident ways to solve the problem is to watermark the data and then to encrypt the watermarked data or, on the contrary, to encrypt the data and then, after decrypting, to watermark the received plaintext. There are some limitations to application of these ways [4]. Particularly, the second way implies a trusted data receiver.

In this paper, we define a concept of a watermarking cipher (called, in short, a w-cipher) according to which the transformation of a plaintext x by the composition of the encryption and decryption transformations using some encryption and decryption keys yields a watermarked text x' containing a watermark w. We show how data watermarking can be performed by applying a w-cipher in such a way that the concealment of a watermark into a plaintext is accomplished by this w-cipher either in the encryption, or in the decryption processes. It should be said that the encryption and decryption functions in it don't compulsorily satisfy the invertibility relation connecting these functions in the

encrypting ciphers usually defined in cryptography. Thus in fact w-ciphers are not ciphers in the known sense of the word, but the ciphers are a certain partial type of w-ciphers, and all terms, notions and notations related to ciphers are quite appropriate to w-ciphers. Before we give a general definition of watermarking ciphers and describe some particular examples of them, we will state some assumptions and suppositions urgent to make this actions more or less correctly and clear.

1. Watermarking problem

First of all, we suppose to the simplicity that any protected data is a string of characters being elements of an additive group G with the addition operation "+". For example, $G = \mathbb{Z}_n$ or $G = \mathbb{Z}_2^n$ for some $n \geqslant 2$, and so on. In particular, a data may be represented by a bitstring (with a certain structure perhaps). For any $a = a_1a_2 \dots a_r$ and $b = b_1b_2 \dots b_r$ in G^r, let $a + b = (a_1 + b_1)(a_2 + b_2) \dots (a_r + b_r)$, $-b = (-b_1)(-b_2) \dots (-b_r)$, and $a - b = a + (-b)$.

A *watermark* w is supposed to be a pair (v, η), where $v = v_1v_2 \dots v_m \in (G \setminus \{0\})^m$ and $\eta = i_1i_2 \dots i_m$, $i_j \in \{1, \dots, l\}$, $j = 1, \dots, m$, $1 \leqslant i_1 < i_2 < \dots < i_m \leqslant l$ for some integers l, m, $l > m \geqslant 1$. In the case of necessity, it is denoted by $(v, \eta)_l$. The number m is called the *length* of w. A watermark (v', η) in which $v' = -v$ is called the *inversion* of w and denoted by $-w$. Evidently, $-(-w) = w$. In the case of $|G| = 2$, we assume $G = \mathrm{Z}_2$. In this case, in any watermark $w = (v, \eta)$, the string v is the vector $11 \dots 1$, thus w is uniquelly determined by the string η and we write $w = \eta$.

The *concealment* of w into a data string $x = x_1x_2 \dots x_l \in G^l$ is fulfilled with the help of the addition operation defined in G. The resulting data string is $x' = x'_1x'_2 \dots x'_l$ in which $x'_j = x_j + v_t$ if $j = i_t \in J = \{i_1, \dots, i_m\}$ and $x'_j = x_j$ if $j \in \{1, \dots, l\} \setminus J$. The string x' is said to be x *marked by* w and denoted by $x + w$. We also agree to write $x - w$ instead of $x + (-w)$. The strings v and η are called, respectively, the *value* and the *abode* of w in x'. In fact, numbers $i_1, i_2, \dots, i_m$ in η indicate positions in x for concealing the components $v_1, v_2, \dots, v_m$, respectively, in the value of w. The string η is said to be a *proper abode* for w in x if x' is obtained from x without an appreciable loss of information. In this case, we call x' a *derivative* (or *copy*) *of* x *properly watermarked by* w.

For example, if x is a digital video bitstring and $v = 11 \dots 1 \in \mathbb{Z}_2^m$, then the concealment of w into x consists in inverting bits $x_{i_1}, x_{i_2}, \dots, x_{i_m}$. In this case, if the bit positions $i_1, i_2, \dots, i_m$ are selected so that the inversion of these bits in x does not noticeably distort the video, then the resulting bitstring x' is a properly watermarked copy of x and both x' and x can be equally used as digital video, but x', besides, contains a watermark for identifying a potential malefactor.

The watermark w and data string x are said to be *mutually proper*, that is, w is proper for x and vice versa if x has an exponential number of proper abodes for w in x. Here, by the exponential number is meant an exponential function of the length m of w. Such a number of proper abodes prevents a malefactor from brute-force attack by enumerating all possible proper abodes in x'. For instance, digital audio and video data are two examples of data bitstring for which watermarking by bit inversion in some positions is proper.

Besides, we suppose there exist a data string producer (DP) and a data string customers, or clients (DC). The DP needs to transmit a data string x to a DC U so that nobody else could intercept x or secretly receive it in his possession from U. With this objective, the DP should like to select a unique proper watermark w and an encryption key k_e for a w-cipher C, to encrypt x by applying C and k_e and to send to U the resulting ciphertext y and an appropriate decryption key k_d constructed thus that the decryption of y using this key results in a data string x', which is a derivative of x properly watermarked by w.

It doesn't matter in what stage w is inserted into x — in the decryption or encryption process. By decrypting y on the key k_d, the client U obtains a unique and proper copy x' of x. If U hands over it to another client, the DP can uniquely identify U by the value v of w and its abode η in x'.

Since U may himself be a malefactor, the decryption key k_d should be connected with the watermark w so that it is computationally infeasible to determine w given k_d and the cyphertext y, that is, there is no an algorithm either at all or of polynomial complexity (as a function of m) computing w from k_d and y.

2. Watermarking cipher definition

Thus we come to the following concept of a w-cipher: for any mutually proper watermark w and plaintext x, the transformation of x by the composition of the encryption and decryption transformations using any encryption and decryption keys connected in some way with each other and with w yields a watermarked text $x' = x + w$. In this way, we introduce two types of watermarking ciphers.

1. A w-cipher with watermarking decryption — a plaintext x is encrypted depending on only a cipher key k, the resulting ciphertext y is decrypted depending on both k and a proper watermark w; thus the encryption key k_e may be arbitrary, the decryption key k_d should be predetermined by the chosen encryption k_e and w, that is, should be a function of k and w.

2. A w-cipher with watermarking encryption — a plaintext x is encrypted depending on both a cipher key k and a proper watermark w, the resulting ciphertext y is decrypted dependening on only k; thus the encryption key k_e should be a function of k and w, the decryption key k_d should be a function of only k.

Formally a *w-cipher* is defined by a 6-tuple $C = (X, K, W, h, E, D)$, where X is the set of data strings including the *plaintexts, ciphertexts*, and *watemarked texts*, $X = G^*$; K and W are the sets of *keys* and *watermarks*, respectively; h is the *key function*, $h : K \times W \to K$, and E and D are the *encryption* and the *decryption algorithms* being some mappings $E : X \times K \to X$ and $D : X \times K \to X$ such that, for any mutually proper $x \in X$ and $w \in W$, for any $k \in K$, the following conditions are satisfied:

1) in the *w-cipher with watermaking decryption* —

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

2) in the *w-cipher with watermaking encryption* —

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

In the case of $h(k, w) = k$ for any $k \in K, w \in W$, we assume to write k instead of $h(k, w)$ in the last expressions and Λ instead of h in C.

3. Watermarking cipher examples

A trivial example of a w-cipher (X, K, W, Λ, E, D) over G may be constructed out of a symmetric cipher (X, Y, K, E', D') with $X = Y = G^*$ and the set W of watermarks as follows: $E(x, k) = E'(x + w, k)$, $D(y, k) = D'(y, k)$ or $E(x, k) = E'(x, k)$, $D(y, k) = D'(y, k) + w$.

A simplest non-trivial example of a w-cipher is the watermarking one-time pad $C_1 = (X, K, W, h, E, D)$ with $X = K = G^*$. In this w-cipher with the watermarking decryption, for a given watermark $w = (v, \eta)$, the ciphertext $y = y_1 y_2 \dots y_l \in X$ is obtained by the addition of a plaintext $x = x_1 x_2 \dots x_l \in X$ and a key string $k = z_1 z_2 \dots z_l \in K$, that is,

$y = x + k$ and the decryption of y resulting in a watermarked plaintext $x' = x'_1 x'_2 \dots x'_l \in X$ is carried out by the subtraction of another key string $k' = k - w = z'_1 z'_2 \dots z'_l \in K$ from y, that is, This w-cipher with watermarking encryption is described by the relations: $k' = k + w$, $y = x + k'$, $x' = y - k$. It is directly verified that in both cases $x' = x + w$. In the first case $k_e = k$, $k_d = h(k, w) = k'$ and the watermark w is automatically concealed into x in the decryption process. In the second case this is done in the encryption process and $k_e = k' = h(k, w)$, $k_d = k$.

In other words, for any $l \geqslant 1$, $x, k \in G^l$, and $w \in W$

1) in C_1 with watermarking decryption —

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

2) in C_1 with watermarking encryption —

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

Another example of a w-cipher is a watermarking stream cipher $C_A = (X, K, W, h, E, D)$ over a finite field F with $X = F^*$ and the keystream generator being a finite autonomous automaton A with a nonlinear output function. The automaton A is represented by a 4-tuple $A = (Q, Z, g, f)$, where Q, Z are the sets of states and output symbols, respectively, $Q = F^n$, $n \geqslant 1$, $Z = F$ and g, f are the transition and output functions of A, $g : Q \to Q$, $f : Q \to Z$. It is also supposed that the output function f is a part of the cipher key, without fail. Sometimes the initial state $q(1)$ of the automaton A and its transition function g may be other parts of the key. Further to the generality, an arbitrary key in K is denoted by the sign $k[q(1), g, f]$ implying f to be compulsory and $q(1)$ and g to be optional. It is also supposed that in A, for any initial state $q(1) \in Q$ and integer $l \geqslant 1$ under consideration, the states $q(t) = g^{t-1}(q(1)), t = 1, 2, \dots, l$, are all different. In this case, for any $w = (v, \eta)_l \in W$ with $v = v_1 v_2 \dots v_m$ and $\eta = i_1 i_2 \dots i_m$, we may define a function $\delta_{w,q(1),l} : Q \to Z$ in such a way that, for any $s \in Q$, $\delta_{w,q(1),l}(s) = v_j$ if $s = q(i_j)$, $j \in \{1, 2, \dots, m\}$, and $\delta_{w,q(1),l}(s) = 0$ otherwise, that is, if $s = q(t)$, $t \in \{1, 2, \dots, l\} \setminus \{i_1, i_2, \dots, i_m\}$. The key function h, the encryption and decryption algorithms E, D and keys k_e, k_d in C_A are particularly defined in the each of two possible cases as follows:

1) watermarking decryption case —

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

2) watermarking encryption case —

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

In both cases, it is immediately verified that $x' = x + w$. Besides, in the first case, $k_e = k[q(1), g, f]$ and $k_d = k[q(1), g, f_1]$; in the second case, $k_e = k[q(1), g, f_2]$ and $k_d = k[q(1), g, f]$.

Finally, we describe a watermarking cipher $C_R = (X, K, W, h, E, D)$ being a particular instance of the w-cipher C_A, in which the automaton $A = (Q, Z, g, f)$ is a nonlinear filter keystream generator [2] constructed using a maximum-length linear feedback shift register (LFSR) R of a length n with a primitive connection polynomial $c_0 + c_1 u + \ldots + c_{n-1} u^{n-1} - u^n$ in $\mathbb{Z}_2[u]$ and with a nonlinear Boolean filtering function f in n variables. Thus $F = \mathbb{Z}_2$, $X = \mathbb{Z}_2^*$, in any $w = (v, \eta) \in W$ the string v is a vector of 1's, thus $w = \eta = i_1 i_2 \ldots i_m$, $Q = \mathbb{Z}_2^n$, $Z = \mathbb{Z}_2$, and for $s = s_0 s_1 \ldots s_{n-1} \in Q$, $g(s_0 s_1 \ldots s_{n-1}) = s_1 \ldots s_{n-1} s_n$, where $s_n = c_0 s_0 + c_1 s_1 + \ldots + c_{n-1} s_{n-1}$.

Since in $\mathbb{Z}_2$ the addition and subtraction operations coincide with the addition modulo 2 and adding to 1 means the inversion, the following relations hold in C_R: 1) if $q(1) \neq 00 \ldots 0$ and $l \leqslant 2^n - 1$, then $\delta_{w,q(1),l}(s) = \sum\limits_{j=1}^{m} s^{q(i_j)}$, where for $\sigma = \sigma_0 \sigma_1 \ldots \sigma_{n-1} \in \mathbb{Z}_2^n$, $s^\sigma = s_0^{\sigma_0} \wedge s_1^{\sigma_1} \wedge \ldots \wedge s_{n-1}^{\sigma_{n-1}}$, $s_t^{\sigma_t} = \bar{s}_t$ if $\sigma_t = 0$, $s_t^{\sigma_t} = s_t$ if $\sigma_t = 1$, $t = 0, 1, \ldots, n-1$; 2) $f_1 = f_2$; 3) the encryption and decryption algorithms in the watemarking encryption case are obtained by permutating them in the watemarking decryption case.

The w-cipher C_R with the watermarking in the process of decryption was implemented and tested for MPEG compressed video data. The implementation is available in [5, 6].

Evidently, the all constructions above remain valid if the additive group in them is replaced by a multiplicative one.

REFERENCES

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, vol. 2, no. 9, pp. 2905–2909.
5. *Anjin V. A.* Metod zashchity ot nelegal'nogo kopirovaniya v tsifrovykh videotranslyatsiyakh cherez vnedrenie vodyanykh znakov pri rasshifrovanii [Content protection with bitstream watermarking at decryption stage]. Prikladnaya Diskretnaya Matematika. Prilozhenie, 2014, no. 7, pp. 73–74. (in Russian)
6. `https://github.com/anjin-viktor/mpeg2decwtrk/` — Method implementation for MPEG2 Video. 2014.