Chaotic iterations versus Spread-spectrum: topological-security and stego-security

Christophe Guyeux, Nicolas Friot, and Jacques M. Bahi

Computer Science Laboratory LIFC
University of Franche-Comté
rue Engel Gros, Belfort, France

{christophe.guyeux, nicolas.friot, jacques.bahi}@lifc.univ-fcomte.fr

December 16, 2011

Keywords: Information hiding; Chaotic iterations; Stego-security; Topological-security; Spread-spectrum.

Abstract

A new framework for information hiding security, called topological-security, has been proposed in a previous study. It is based on the evaluation of unpredictability of the scheme, whereas existing notions of security, as stego-security, are more linked to information leaks. It has been proven that spread-spectrum techniques, a well-known stego-secure scheme, are topologically-secure too. In this paper, the links between the two notions of security is deepened and the usability of topological-security is clarified, by presenting a novel data hiding scheme that is twice stego and topological-secure. This last scheme has better scores than spread-spectrum when evaluating qualitative and quantitative topological-security properties. Incidentally, this result shows that the new framework for security tends to improve the ability to compare data hiding scheme.

1 Introduction

Information hiding has recently become a major digital technology [6], [9], especially with the increasing importance and widespread distribution of digital media through the Internet. Spread-spectrum data-hiding techniques have been widely studied in recent years under the scope of security. These techniques encompass several schemes, such as Improved Spread Spectrum (ISS), Circular Watermarking (CW), and Natural Watermarking (NW). Some of these schemes have revealed various security issues. On the contrary, it has been proven in [4] that the Natural Watermarking technique is stego-secure. This stego-security is one of the security classes defined in [4]. In this paper, probabilistic models are used to categorize the security of data hiding algorithms in the Watermark Only Attack (WOA) framework.

We will show that the security level of such algorithms can be studied into a novel framework based on unpredictability, as it is understood in the theory of chaos [5]. To do so, a new class of security will be introduced, namely the topological-security. This new class can be used to study some categories of attacks that are difficult to investigate in the existing security approach. It also enriches the variety of qualitative and quantitative tools that evaluate how strong the security is, thus reinforcing the confidence that can be had in a given scheme.

In addition of being stego-secure, it has been proven in [3] that Natural Watermarking technique is topologically-secure. Moreover, this technique possesses additional properties of unpredictability, namely, strong transitivity, topological mixing, and a constant of sensitivity equal to $\frac{N}{2}$. However NW are not expansive, which is problematic in the Constant-Message Attack (CMA) and Known Message Attack (KMA) setups [3]. In this paper, it is proven by using the new topological-security framework, that a more secure scheme than NW can be found to withstand attacks in these setups. This scheme, introduced in [2], is based on the so-called chaotic iterations. The aim of this work is to prove that this algorithm is stego-secure and topologically-secure, to study its qualitative and quantitative properties of unpredictability, and then to compare it with Natural Watermarking.

The rest of this paper is organized as follows. In Section 2, basic definitions and terminologies in the field of topology, chaos, and security are recalled. In Section 3 the stego-security of chaotic iterations is established in some cases, whereas in Section 4 is studied the topological-security of chaotic iterations. Natural Watermarking and chaotic iterations are then compared in Section 5. The paper ends with a conclusion where our contribution is summarized, and planned future work is discussed.

2 Basic recalls

2.1 Chaotic iterations

In this section, the definition and main properties of chaotic iterations are recalled [1].

2.1.1 Chaotic iterations

In the sequel S^n denotes the n^{th} term of a sequence S and V_i the i^{th} component of a vector V. Finally, the following notation is used: $[\![1;N]\!] = \{1, 2, \dots, N\}$.

Let us consider a *system* of a finite number N of elements (or *cells*), so that each cell has a boolean *state*. A sequence of length N of boolean states of the cells corresponds to a particular *state of the system*. A sequence which elements belong to $[\![1;\mathsf{N}]\!]$ is called a *strategy*. The set of all strategies is denoted by $\mathbb{S}$.

Definition 1 The set $\mathbb{B}$ denoting $\{0, 1\}$, let $f : \mathbb{B}^{\mathsf{N}} \longrightarrow \mathbb{B}^{\mathsf{N}}$ be a function and $S \in \mathbb{S}$ be a strategy. The so-called *chaotic iterations* are defined by $x^0 \in \mathbb{B}^{\mathsf{N}}$ and $\forall (n, i) \in \mathbb{N}^* \times [\![1;\mathsf{N}]\!]$:

$$x_i^n = \begin{cases} x_i^{n-1} & \text{if } S^n \neq i \\ \left(f(x^{n-1})\right)_{S^n} & \text{if } S^n = i. \end{cases}$$

2.1.2 Devaney's chaotic dynamical systems

Consider a metric space $(\mathcal{X}, d)$ and a continuous function f on $\mathcal{X}$. f is said to be *topologically transitive* if, for any pair of open sets $U, V \subset \mathcal{X}$, there exists $k > 0$ such that $f^k(U) \cap V \neq \varnothing$. $(\mathcal{X}, f)$ is said to be *regular* if the set of periodic points is dense in $\mathcal{X}$. f has *sensitive dependence on initial conditions* if there exists $\delta > 0$ such that, for any $x \in \mathcal{X}$ and any neighborhood V of x, there exists $y \in V$ and $n \geqslant 0$ such that $|f^n(x) - f^n(y)| > \delta$. δ is called the *constant of sensitivity* of f. Quoting Devaney in [5],

Definition 1 A function $f : \mathcal{X} \longrightarrow \mathcal{X}$ is said to be *chaotic* on $\mathcal{X}$ if $(\mathcal{X}, f)$ is regular, topologically transitive and has sensitive dependence on initial conditions.

2.1.3 Chaotic iterations and Devaney's chaos

In this section we give outline proofs of the properties on which our secure data hiding scheme is based. The complete theoretical framework is detailed in [1].

Denote by Δ the *discrete boolean metric*, $\Delta(x, y) = 0 \Leftrightarrow x = y$. Given a function f, define the function: $F_f : [\![1;\mathsf{N}]\!] \times \mathbb{B}^{\mathsf{N}} \longrightarrow \mathbb{B}^{\mathsf{N}}$ such that $F_f(k, E) = \left(E_j.\Delta(k, j) + f(E)_k.\overline{\Delta(k, j)}\right)_{j \in [\![1;\mathsf{N}]\!]}$.

Let us consider the phase space $\mathcal{X} = [\![1;\mathsf{N}]\!]^{\mathbb{N}} \times \mathbb{B}^{\mathsf{N}}$ and the map $G_f(S, E) = (\sigma(S), F_f(i(S), E))$, where σ is defined by $\sigma : (S^n)_{n \in \mathbb{N}} \in \mathbb{S} \to (S^{n+1})_{n \in \mathbb{N}} \in \mathbb{S}$, and i is the map $i : (S^n)_{n \in \mathbb{N}} \in \mathbb{S} \to S^0 \in [\![1;\mathsf{N}]\!]$. So the chaotic iterations can be described by the following iterations:

$$X^0 \in \mathcal{X} \text{ and } X^{k+1} = G_f(X^k).$$

We have defined in [1] a new distance d between two points $(S, E), (\check{S}, \check{E}) \in \mathcal{X}$ by $d((S, E); (\check{S}, \check{E})) = d_e(E, \check{E}) + d_s(S, \check{S})$, where:

- $d_e(E, \check{E}) = \displaystyle\sum_{k=1}^{\mathsf{N}} \Delta(E_k, \check{E}_k) \in [\![0;\mathsf{N}]\!]$

- $d_s(S, \check{S}) = \frac{9}{\mathsf{N}} \sum_{k=1}^{\infty} \frac{|S^k - \check{S}^k|}{10^k} \in [0; 1]$.

It is then proven that,

Proposition 1 *G_f is a continuous function on $(\mathcal{X}, d)$.*

In the metric space $(\mathcal{X}, d)$, the vectorial negation $f_0 : \mathbb{B}^N \longrightarrow \mathbb{B}^N, (b_1, \cdots, b_{\mathsf{N}}) \longmapsto (\overline{b_1}, \cdots, \overline{b_{\mathsf{N}}})$ satisfies the three conditions for Devaney's chaos: regularity, transitivity, and sensitivity [1]. So,

Proposition 2 *G_{f_0} is a chaotic map on $(\mathcal{X}, d)$ according to Devaney.*

2.2 Using chaotic iterations as information hiding schemes

2.2.1 Presentation of the scheme

We have proposed in [2] to use chaotic iterations as an information hiding scheme, as follows (see Figure 1). Let:

- $(K, N) \in [0; 1] \times \mathbb{N}$ be an embedding key,
- $X \in \mathbb{B}^{\mathsf{N}}$ be the N least significant coefficients (LSCs) of a given cover media C,
- $(S^n)_{n \in \mathbb{N}} \in [\![1, \mathsf{N}]\!]^{\mathbb{N}}$ be a strategy, which depends on the message to hide $M \in [0; 1]$ and K,
- $f_0 : \mathbb{B}^{\mathsf{N}} \to \mathbb{B}^{\mathsf{N}}$ be the vectorial logical negation.

(a) Original Lena.

(b) Watermarked Lena.

Figure 1: Data hiding with chaotic iterations

So the watermarked media is C whose LSCs are replaced by $Y_K = X^N$, where:

$$\begin{cases} X^0 = X \\ \forall n < N, X^{n+1} = G_{f_0}(X^n). \end{cases}$$

In the following section, two ways to generate $(S^n)_{n\in\mathbb{N}}$ are given, namely Chaotic Iterations with Independent Strategy (CIIS) and Chaotic Iterations with Dependent Strategy (CIDS). In CIIS, the strategy is independent from the cover media X, whereas in CIDS the strategy will be dependent on X. Their stego-security are studied in Section 3 and their topological-security in Section 4.

2.2.2 Examples of strategies

CIIS strategy Let us first introduce the Piecewise Linear Chaotic Map (PLCM, see [7]), defined by:

Definition 2 (PLCM)

$$F(x,p) = \begin{cases} x/p & \text{if} \quad x \in [0;p] \\ (x-p)/(\frac{1}{2}-p) & \text{if} \quad x \in \left[p;\frac{1}{2}\right] \\ F(1-x,p) & \text{else.} \end{cases}$$

where $p \in \left]0;\frac{1}{2}\right[$ is a "control parameter".

Then, we can define the general term of the strategy $(S^n)_n$ in CIIS setup by the following expression: $S^n = \lfloor \mathsf{N} \times K^n \rfloor + 1$, where:

$$\begin{cases} p \in \left[0;\frac{1}{2}\right] \\ K^0 = M \otimes K \\ K^{n+1} = F(K^n,p), \forall n \leq N_0 \end{cases}$$

in which $\otimes$ denotes the bitwise exclusive or (XOR) between two floating part numbers (*i.e.*, between their binary digits representation). Lastly, to be certain to enter into the chaotic regime of PLCM [7], the strategy can be preferably defined by: $S^n = \left\lfloor \mathsf{N} \times K^{n+D} \right\rfloor + 1$, where $D \in \mathbb{N}$.

CIDS strategy The same notations as above are used. We define CIDS strategy as follows: $\forall k \leqslant N$,

- if $k \leqslant \mathsf{N}$ and $X^k = 1$, then $S^k = k$,
- else $S^k = 1$.

In this situation, if $N \geqslant \mathsf{N}$, then only two watermarked contents are possible with the scheme proposed in Section 2.2, namely: $Y_K = (0,0,\cdots,0)$ and $Y_K = (1,0,\cdots,0)$.

3 Evaluation of the stego-security

3.1 Definition of stego-security

Stego-security, defined in the Simmons' prisoner problem [8], is the highest security class in WOA setup [4].

Let $\mathbb{K}$ be the set of embedding keys, $p(X)$ the probabilistic model of N_0 initial host contents, and $p(Y|K_1)$ the probabilistic model of N_0 watermarked contents. We suppose that each host content has been watermarked with the same key K_1 and the same embedding function e.

Definition 3 The embedding function e is stego-secure if and only if:

$$\forall \mathbf{K_1} \in \mathbb{K}, \mathbf{p}(\mathbf{Y}|\mathbf{K_1}) = \mathbf{p}(\mathbf{X})$$

3.2 Evaluation of the stego-security

Let us now study the stego-security of the scheme. We will prove that,

Proposition 3 *CIIS are stego-secure.*

PROOF Let us suppose that $X \sim \mathbf{U}\left(\mathbb{B}^N\right)$ in a CIIS setup. We will prove by a mathematical induction that $\forall n \in \mathbb{N}, X^n \sim \mathbf{U}\left(\mathbb{B}^N\right)$. The base case is immediate, as $X^0 = X \sim \mathbf{U}\left(\mathbb{B}^N\right)$. Let us now suppose that the statement $X^n \sim \mathbf{U}\left(\mathbb{B}^N\right)$ holds for some n. Let $e \in \mathbb{B}^N$ and $\mathbf{B}_k = (0, \cdots, 0, 1, 0, \cdots, 0) \in \mathbb{B}^N$ (the digit 1 is in position k). So $P\left(X^{n+1} = e\right) = \sum_{k=1}^{N} P\left(X^n = e + \mathbf{B}_k, S^n = k\right)$. These two events are independent in CIIS setup, thus: $P\left(X^{n+1} = e\right) = \sum_{k=1}^{N} P\left(X^n = e + \mathbf{B}_k\right) \times P\left(S^n = k\right)$. According to the inductive hypothesis: $P\left(X^{n+1} = e\right) = \frac{1}{2^N} \sum_{k=1}^{N} P\left(S^n = k\right)$. The set of events $\{S^n = k\}$ for $k \in [\![1; N]\!]$ is a partition of the universe of possible, so $\sum_{k=1}^{N} P\left(S^n = k\right) = 1$.

Finally, $P\left(X^{n+1} = e\right) = \frac{1}{2^N}$, which leads to $X^{n+1} \sim \mathbf{U}\left(\mathbb{B}^N\right)$. This result is true $\forall n \in \mathbb{N}$, we thus have proven that,

$$\forall K \in [0; 1], Y_K = X^{N_0} \sim \mathbf{U}\left(\mathbb{B}^N\right) \text{ when } X \sim \mathbf{U}\left(\mathbb{B}^N\right)$$

So CIIS defined in Section 2.2 are stego-secure.

We will now prove that,

Proposition 4 *CIDS are not stego-secure.*

PROOF Due to the definition of CIDS, we have $P(Y_K = (1, 1, \cdots, 1)) = 0$. So there is no uniform repartition for the stego-contents Y_K.

4 Evaluation of the topological-security

4.1 Definition

To check whether an information hiding scheme S is topologically-secure or not, S must be written as an iterate process $x^{n+1} = f(x^n)$ on a metric space $(\mathcal{X}, d)$. This formulation is always possible, as it is proven in [3]. So,

Definition 4 An information hiding scheme S is said to be topologically-secure on $(\mathcal{X}, d)$ if its iterative process has a chaotic behavior according to Devaney.

It can be established that,

Proposition 5 *CIIS and CIDS are topologically-secure.*

Proof It has been proven in [1] that chaotic iterations have a chaotic behavior, as defined by Devaney.

In the two following sections, we will study the qualitative and quantitative properties of topological-security for chaotic iterations. These properties can measure the disorder generated by our scheme, giving by doing so some important informations about the unpredictability level of such a process.

4.2 Quantitative property of chaotic iterations

Definition 5 (Expansivity) A function f is said to be *expansive* if $\exists \varepsilon > 0, \forall x \neq y, \exists n \in \mathbb{N}, d(f^n(x), f^n(y)) \geqslant \varepsilon$.

Proposition 6 G_{f_0} *is an expansive chaotic dynamical system on* $\mathcal{X}$ *with a constant of expansivity is equal to 1.*

Proof If $(S, E) \neq (\check{S}; \check{E})$, then either $E \neq \check{E}$, so at least one cell is not in the same state in E and $\check{E}$. Consequently the distance between (S, E) and $(\check{S}; \check{E})$ is greater or equal to 1. Or $E = \check{E}$. So the strategies S and $\check{S}$ are not equal. Let n_0 be the first index in which the terms S and $\check{S}$ differ. Then $\forall k < n_0, \tilde{G}^k_{f_0}(S, E) = \tilde{G}^k_{f_0}(\check{S}, \check{E})$, and $\tilde{G}^{n_0}_{f_0}(S, E) \neq \tilde{G}^{n_0}_{f_0}(\check{S}, \check{E})$. As $E = \check{E}$, the cell which has changed in E at the n_0-th iterate is not the same as the cell which has changed in $\check{E}$, so the distance between $\tilde{G}^{n_0}_{f_0}(S, E)$ and $\tilde{G}^{n_0}_{f_0}(\check{S}, \check{E})$ is greater or equal to 2.

4.3 Qualitative property of chaotic iterations

Definition 6 (Topological mixing) A discrete dynamical system is said to be topologically mixing if and only if, for any couple of disjoint open set $U, V \neq \varnothing$, $n_0 \in \mathbb{N}$ can be found so that $\forall n \geqslant n_0, f^n(U) \cap V \neq \varnothing$.

Proposition 7 $\tilde{G}_{f_0}$ *is topologically mixing on* $(\mathcal{X}', d')$.

This result is an immediate consequence of the lemma below.

Lemma 1 *For any open ball* B *of* $\mathcal{X}'$, *an index* n *can be found such that* $\tilde{G}^n_{f_0}(B) = \mathcal{X}'$.

Proof Let $B = B((E, S), \varepsilon)$ be an open ball, which the radius can be considered as strictly less than 1. All the elements of B have the same state E and are such that an integer $k\,(= -\log_{10}(\varepsilon))$ satisfies:

- all the strategies of B have the same k first terms,
- after the index k, all values are possible.

Then, after k iterations, the new state of the system is $\tilde{G}'^{k}_{f_0}(E,S)_1$ and all the strategies are possible (all the points $(\tilde{G}'^{k}_{f_0}(E,S)_1, \hat{\mathrm{S}})$, with any $\hat{\mathrm{S}} \in \mathbb{S}$, are reachable from B).

We will prove that all points of $\mathcal{X}'$ are reachable from B. Let $(E',S') \in \mathcal{X}'$ and s_i be the list of the different cells between $\tilde{G}'^{k}_{f_0}(E,S)_1$ and E'. We denote by $|s|$ the size of the sequence s_i. So the point $(\check{E},\check{S})$ of B defined by: $\check{E}=E$, $\check{S}^i = S^i, \forall i \leqslant k$, $\check{S}^{k+i} = s_i, \forall i \leqslant |s|$, and $\forall i \in \mathbb{N}, S^{k+|s|+i} = S'^{i}$ is such that $\tilde{G}^{k+|s|}_{f_0}(\check{E},\check{S}) = (E',S')$. This concludes the proofs of the lemma and of the proposition.

5 Comparison between spread-spectrum and chaotic iterations

The consequences of topological mixing for data hiding are multiple. Firstly, security can be largely improved by considering the number of iterations as a secret key. An attacker will reach all of the possible media when iterating without this key. Additionally, he cannot benefit from a KOA setup, by studying media in the neighborhood of the original cover. Moreover, as in a topological mixing situation, it is possible that any hidden message (the initial condition), is sent to the same fixed watermarked content (with different numbers of iterations), the interest to be in a KMA setup is drastically reduced. Lastly, as all of the watermarked contents are possible for a given hidden message, depending on the number of iterations, CMA attacks will fail.

The property of expansivity reinforces drastically the sensitivity in the aims of reducing the benefits that Eve can obtain from an attack in KMA or KOA setup. For example, it is impossible to have an estimation of the watermark by moving the message (or the cover) as a cursor in situation of expansivity: this cursor will be too much sensitive and the changes will be too important to be useful. On the contrary, a very large constant of expansivity ε is unsuitable: the cover media will be strongly altered whereas the watermark would be undetectable.

Finally, spread-spectrum is relevant when a discrete and secure data hiding technique is required in WOA setup. However, this technique should not be used in KOA and KMA setup, due to its lack of expansivity.schemes, which are expansive.

6 Conclusion and future work

In this paper, the links between stego-security and topological-security has been deepened. The information hiding scheme presented in [2], which is based on chaotic iterations, has been recalled and its level of security has been studied. It has been proven that this algorithm is twice stego and topologically-secure. This was already the case for spread-spectrum techniques, as it has been established in [3]. Moreover, as for spread-spectrum, chaotic iterations possess the qualitative property of topological mixing, which are useful to withstand attacks. However, unlike spread-spectrum, chaotic iterations are expansive, so this scheme is better than spread-spectrum in KOA and KMA setups. Incidentally, this result shows that the new framework for security tends to improve the ability to compare data hiding scheme. In future work, we will give a better understanding of the links between these two security frameworks. Additionally, the comparison between spread-spectrum and chaotic iterations outlined in this paper will be extended. The security of other existing schemes will be studied in the framework of topological-security. Last, but not least, the way to understand these new tools in terms of data hiding aims will be enhanced: this study is required to make topological-security framework truly useful in practice.

References

[1] Jacques Bahi and Christophe Guyeux. Hash functions using chaotic iterations. *Journal of Algorithms & Computational Technology*, 4(2):167–181, 2010.

[2] Jacques Bahi and Christophe Guyeux. A new chaos-based watermarking algorithm. In *SECRYPT 2010, International conference on security and cryptography*, Athens, Greece, 2010. To appear.

[3] Jacques M. Bahi and Christophe Guyeux. A chaos-based approach for information hiding security. arXiv N^o 0034939, April 2010.

[4] Francois Cayre, Caroline Fontaine, and Teddy Furon. Kerckhoffs-based embedding security classes for woa data hiding. *IEEE Transactions on Information Forensics and Security*, 3(1):1–15, 2008.

[5] Robert L. Devaney. *An Introduction to Chaotic Dynamical Systems, 2nd Edition*. Westview Pr., March 2003.

[6] Andrew D. Ker, Tomáš Pevný, Jan Kodovský, and Jessica Fridrich. The square root law of steganographic capacity. In *MMSec '08: Proceedings of the 10th ACM workshop on Multimedia and security*, pages 107–116, New York, NY, USA, 2008. ACM.

[7] Li Shujun, Li Qi, Li Wenmin, Mou Xuanqin, and Cai Yuanlong. Statistical properties of digital piecewise linear chaotic maps and their roles in cryptography and pseudo-random coding. *Proceedings of the 8th IMA International Conference on Cryptography and Coding*, 1:205–221, 2001.

[8] Gustavus J. Simmons. The prisoners' problem and the subliminal channel. In *Advances in Cryptology, Proc. CRYPTO'83*, pages 51–67, 1984.

[9] Fuchun Xie, Teddy Furon, and Caroline Fontaine. Better security levels for 'broken arrows'. In *Proc. of SPIE Electronic Imaging on Media Forensics and Security XII*, San Jose, CA, USA, jan 2010.