Randomness and disorder of chaotic iterations. Applications in information security field

Xiaole Fang [§], Christophe Guyeux[†], Qianxue Wang[‡], and Jacques M. Bahi[†]

November 28, 2016

Abstract

Design and cryptanalysis of chaotic encryption schemes are major concerns to provide secured information systems. Pursuing our previous research works, some well-defined discrete chaotic iterations that satisfy the reputed Devaney's definition of chaos have been proposed. In this article, we summarize these contributions and propose applications in the fields of pseudorandom number generation, hash functions, and symmetric cryptography. For all these applications, the proofs of chaotic properties are outlined.

1 Introduction

Applying chaotic systems to construct cryptosystems has been extensively investigated since 1990s, and this field of research has attracted more and more attention in the near decades. Some researchers have pointed out that there exists tight relationship between chaos and randomness, thus it is a natural idea to use chaos to enrich the design of cryptographic applications. However, almost all current researches of chaotic systems consider real domain. Since all operations (iterations) are on the real numbers, Real Domain Chaotic Systems (RDCSs) realized in a computer or a digital device will inevitably lead to finite precision effects, and may result in consequent dynamic degradation, such as short cycle-length, non-ideal distribution and correlation, low linear complexity, and so on. Chaotic iterations (CIs), for its part, refers to chaotic systems defined on integer domain. They have been deeply studied in our previous collaborative works, in order to solve degradation of chaotic dynamic properties by finite precision effects on traditional RDCSs. In this research work, we intend to deepen the theoretical and practical knowledge already obtained on CIs. More general theoretical designs and applications for security will be done to further investigate and learn more about the CIs framework.

The remainder of this research work is organized as follows. The basic recalls of CIs are given in Section 2, while results of Devaney's chaos are provided in Section 3. Sections 4, 5, and 6 show applications to pseudorandom number generation, hash functions, and symmetric cryptography respectively. This article ends by a conclusion section in which the article is summarized.

2 Basic recalls

2.1 Devaney's theory of chaos

In the remainder of this article, S^n denotes the n^{th} term of a sequence S while $\mathcal{X}^{\mathbb{N}}$ is the set of all sequences whose elements belong to $\mathcal{X}$. V_i stands for the i^{th} component of a vector V. $f^k = f \circ ... \circ f$ is for the k^{th} composition of a function f. $\mathbb{N}$ is the set of natural (non-negative) numbers, while $\mathbb{N}^*$ stands for the positive integers $1, 2, 3, \ldots$ Finally, the following notation is used: $[\![1; N]\!] = \{1, 2, \ldots, N\}$.

Consider a topological space $(\mathcal{X}, \tau)$ and a continuous function $f : \mathcal{X} \to \mathcal{X}$ on $(\mathcal{X}, \tau)$.

Definition 1. *The function f is* topologically transitive *if, for any pair of open sets $U, V \subset \mathcal{X}$, there exists an integer $k > 0$ such that $f^k(U) \cap V \neq \varnothing$.*

Definition 2. *An element x is a* periodic point *for f of period $n \in \mathbb{N}$, $n > 1$, if $f^n(x) = x$. f is* regular *on $(\mathcal{X}, \tau)$ if the set of periodic points for f is dense in $\mathcal{X}$: for any point x in $\mathcal{X}$, any neighborhood of x contains at least one periodic point.*

Definition 3 (Devaney's formulation of chaos [8]). *The function f is* chaotic *on $(\mathcal{X}, \tau)$ if f is regular and topologically transitive.*

Banks *et al.* have proven in [7] that, when the topological space is a metric one $(\mathcal{X}, d)$, chaos implies sensitivity, defined below:

Definition 4. *The function 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 exist $y \in V$ and $n > 0$ such that $d\left(f^n(x), f^n(y)\right) > \delta$.*
δ is called the constant of sensitivity *of f.*

2.2 Chaotic Iterations

Define by $\mathcal{S}_X$ the set of sequences whose elements belong in $X \subset \mathbb{N}, X \neq \varnothing$, that is, $\mathcal{S}_X = X^{\mathbb{N}}$.

Definition 5. *The set $\mathbb{B}$ denoting $\{0, 1\}$, let $\mathsf{N} \in \mathbb{N}^*$, $f : \mathbb{B}^{\mathsf{N}} \longrightarrow \mathbb{B}^{\mathsf{N}}$ be a function, and $S \in \mathcal{S}_{[\![1,\mathsf{N}]\!]}$ be a sequence of integers between 1 and N. The so-called* chaotic iterations *are defined by $x^0 \in \mathbb{B}^{\mathsf{N}}$ and*

$$\forall n \in \mathbb{N}^*, \forall i \in [\![1;\mathsf{N}]\!], x_i^n = \begin{cases} x_i^{n-1} & \textit{if } S^n \neq i \\ \left(f(x^{n-1})\right)_{S^n} & \textit{if } S^n = i. \end{cases} \tag{1}$$

In other words, at the n^{th} iteration, only the S^n–th component of the vector x^n is updated. Note that in a more general formulation, each S^n can be a subset of $\{1, 2, \ldots, \mathsf{N}\}$. Let us remark that the term “chaotic”, in the name of these iterations, has *a priori* no link with the mathematical theory of chaos, recalled before.

3 Chaos results about chaotic iterations

We now recall how to define a suitable metric space where chaotic iterations are continuous. For further explanations, see, *e.g.*, [5]. Let δ be the *discrete Boolean metric*, $\delta(x, y) = 0 \Leftrightarrow x = y$. Given a function f, define the function $F_f : [\![1;\mathsf{N}]\!] \times \mathbb{B}^{\mathsf{N}} \longrightarrow \mathbb{B}^{\mathsf{N}}$ by:

$$(k, E) \quad \longmapsto \quad \left(E_j.\delta(k, j) + f(E)_k.\overline{\delta(k, j)}\right)_{j\in[\![1;\mathsf{N}]\!]}$$

where + and . are the Boolean addition and product operations. Consider the phase space: $\mathcal{X} = [\![1;\mathsf{N}]\!]^{\mathbb{N}} \times \mathbb{B}^{\mathsf{N}}$, and the map defined on $\mathcal{X}$ by:

$$G_f(S, E) = \left(\sigma(S), F_f(i(S), E)\right), \tag{2}$$

where σ is the *shift* function defined by $\sigma : (S^n)_{n\in\mathbb{N}} \in [\![1,\mathsf{N}]\!]^{\mathbb{N}} \longrightarrow (S^{n+1})_{n\in\mathbb{N}} \in [\![1,\mathsf{N}]\!]^{\mathbb{N}}$ and i is the *initial function* $i : (S^n)_{n\in\mathbb{N}} \in [\![1,\mathsf{N}]\!]^{\mathbb{N}} \longrightarrow S^0 \in [\![1;\mathsf{N}]\!]$. Then the chaotic iterations proposed in Definition 5 can be described by the following discrete dynamical system, whose topological chaos can now be studied:

$$\begin{cases} X^0 \in \mathcal{X} \\ X^{k+1} = G_f(X^k). \end{cases} \tag{3}$$

To do so, a relevant distance between two points $X = (S, E), Y = (\check{S}, \check{E}) \in \mathcal{X}$ has been introduced in [5] as follows: $d(X, Y) = d_e(E, \check{E}) + d_s(S, \check{S})$, where

$$\begin{cases} d_e(E, \check{E}) & = & \displaystyle\sum_{k=1}^{\mathsf{N}} \delta(E_k, \check{E}_k), \\ d_s(S, \check{S}) & = & \displaystyle\frac{9}{\mathsf{N}} \sum_{k=1}^{\infty} \frac{|S^k - \check{S}^k|}{10^k}. \end{cases} \tag{4}$$

It has been established in [5] that,

Proposition 1. *G_f is continuous in the metric space $(\mathcal{X}, d)$.*

The chaotic property of G_f has been firstly established for the vectorial Boolean negation $f_0(x_1, \ldots, x_{\mathsf{N}}) = (\overline{x_1}, \ldots, \overline{x_{\mathsf{N}}})$ [5]. To obtain a characterization, we have secondly introduced the notion of asynchronous iteration graph recalled thereafter [2]. Let f be a map from $\mathbb{B}^{\mathsf{N}}$ to itself. The *asynchronous iteration graph* associated with f is the directed graph $\Gamma(f)$ defined by: the set of vertices is $\mathbb{B}^{\mathsf{N}}$; for all $x \in \mathbb{B}^{\mathsf{N}}$ and $i \in [\![1;\mathsf{N}]\!]$, the graph $\Gamma(f)$ contains an arc from x to $F_f(i, x)$. We have then proven in [2] that,

Theorem 1. *Let $f : \mathbb{B}^{\mathsf{N}} \to \mathbb{B}^{\mathsf{N}}$. G_f is chaotic (according to Devaney) if and only if $\Gamma(f)$ is strongly connected.*

Finally, we have established in [2] that,

Theorem 2. *Let $f : \mathbb{B}^n \to \mathbb{B}^n$, $\Gamma(f)$ its iteration graph, $\check{M}$ its adjacency matrix and M a $n \times n$ matrix defined by $M_{ij} = \frac{1}{n}\check{M}_{ij}$ if $i \neq j$ and $M_{ii} = 1 - \frac{1}{n}\sum\limits_{j=1, j\neq i}^{n} \check{M}_{ij}$ otherwise.*

If $\Gamma(f)$ is strongly connected, then the output of the chaotic iterations follows a law that tends to the uniform distribution if and only if M is a double stochastic matrix.

These results of topological chaos and uniform distribution have led us to study the possibility of building a pseudorandom number generator (PRNG) based on chaotic iterations. As G_f, defined on the domain $[\![1;\mathsf{N}]\!]^{\mathbb{N}} \times \mathbb{B}^{\mathsf{N}}$, is built from Boolean networks $f : \mathbb{B}^{\mathsf{N}} \to \mathbb{B}^{\mathsf{N}}$, we can preserve the theoretical properties on G_f during implementations (due to the discrete nature of f).

4 Application to pseudorandom number generation

Let $\mathsf{N} \in \mathbb{N}^*$, $f : \mathbb{B}^{\mathsf{N}} \to \mathbb{B}^{\mathsf{N}}$, and $\mathcal{P} \subset \mathbb{N}^*$ a non empty and finite set of integers. Any couple $(u, v) \in \mathcal{S}_{[\![1,\mathsf{N}]\!]} \times \mathcal{S}_{\mathcal{P}}$ defines a "chaotic iterations based" PRNG, which is denoted by $CIPRNG_f^2(u, v)$ [11]. It is defined as follows:

$$\left\{\begin{array}{l} x^0 \in \mathbb{B}^{\mathsf{N}} \\ \forall n \in \mathbb{N}, \forall i \in [\![1,\mathsf{N}]\!], x_i^{n+1} = \left\{\begin{array}{ll} f(x^n)_i & \text{if } i = u^n \\ x_i^n & \text{else} \end{array}\right. \\ \forall n \in \mathbb{N}, y^n = x^{v^n}. \end{array}\right. \tag{5}$$

The outputted sequence produced by this generator is $(y^n)_{n\in\mathbb{N}}$.

The formerly proposed $CIPRNG_f^1(u)$ [3, 6] is equal to $CIPRNG_f^2\left(u, (1)_{n\in\mathbb{N}}\right)$, where $(1)_{n\in\mathbb{N}}$ is the sequence that is uniformly equal to 1. It has been proven as chaotic for the vectorial Boolean negation $f_0 : \mathbb{B}^{\mathsf{N}} \longrightarrow \mathbb{B}^{\mathsf{N}}$, $(x_1, \dots, x_{\mathsf{N}}) \longmapsto (\overline{x_1}, \dots, \overline{x_{\mathsf{N}}})$ in [3] and for a larger set of well-chosen iteration functions in [2] but, as only one bit is modified at each iteration, this generator is not able to pass any reasonable statistical tests. The *XOR CIPRNG*(S), for its part [4], is defined as follows: $x^0 \in \mathbb{B}^{\mathsf{N}}$, and $\forall n \in \mathbb{N}, x^{n+1} = x^n \oplus S^n$ where $S \in \mathcal{S}_{[\![1,\mathsf{N}]\!]}$ and $\oplus$ stands for the bitwise *exclusive or* (xor) operation between the binary decomposition of x^n and S^n. This is indeed a $CIPRNG_{f_0}^2(u, v)$ generator: for any given $S \in \mathcal{S}_{[\![1,\mathsf{N}]\!]}$, v^n is the number of 1's in the binary decomposition of S^n while $u^{v^n}, u^{v^n+1}, \dots, u^{v^{n+1}-1}$ are the positions of these ones. The *XOR CIPRNG* has been proven chaotic and it is able to pass all the most stringent statistical batteries of tests [4], namely the well-known DieHARD, NIST, and TestU01. Furthermore, the output sequence is cryptographically secure when S is cryptographically secure [4]. Following the same canvas than in the previous section, we have then characterized which $CIPRNG_f^2(u, v)$ is chaotic according to Devaney.

Denote by $\mathcal{X}_{\mathsf{N},\mathcal{P}} = \mathbb{B}^{\mathsf{N}} \times \mathbb{S}_{\mathsf{N},\mathcal{P}}$, where $\mathbb{S}_{\mathsf{N},\mathcal{P}} = \mathcal{S}_{[\![1,\mathsf{N}]\!]} \times \mathcal{S}_{\mathcal{P}}$. We then introduce a directed graph $\mathcal{G}_{f,\mathcal{P}}$ as follows.

- Its vertices are the 2^{N} elements of $\mathbb{B}^{\mathsf{N}}$.
- Each vertex has $\displaystyle\sum_{i=1}^{\mathsf{p}} \mathsf{N}^{p_i}$ arrows, namely all the $p_1, p_2, \dots, p_{\mathsf{p}}$ tuples having their elements in $[\![1,\mathsf{N}]\!]$.
- There is an arc labelled $a_1, \dots, a_{p_i}$, $i \in [\![1,\mathsf{p}]\!]$ between vertices x and y if and only if $y = F_{f,p_i}(x, (a_1, \dots, a_{p_i}))$.

We have finally proven that

Theorem 3. *The pseudorandom number generator $CIPRNG_f^2$ is chaotic on $\mathcal{X}_{\mathsf{N},\mathcal{P}}$ if and only if its graph $\mathcal{G}_{f,\mathcal{P}}$ is strongly connected.*

5 Application to hash functions

For the interest to add chaos properties to an hash function, among other things regarding their diffusion and confusion, reader is referred to [1]. Recall that, among other cryptographical properties, an hash function must be resistant to collisions: an adversary should not be able to find two distinct messages m and m' such that $h(m) = h(m')$. Furthermore, an hash function must be second-preimage resistant, that is to say: an adversary given a message m should not be able to find another message m' such that $m \neq m'$ and $h(m) = h(m')$.

Let us now give a post-operative mode that can be applied to a cryptographically secure hash function without loosing the cryptographic properties recalled above.

Definition 6. *Let*

- $k_1, k_2, n \in \mathbb{N}^*$,
- $h : (k, m) \in \mathbb{B}^{k_1} \times \mathbb{B}^* \longmapsto h(k, m) \in \mathbb{B}^n$ *a keyed hash function,*
- $S : k \in \mathbb{B}^{k_2} \longmapsto \left(S(k)^i\right)_{i\in\mathbb{N}} \in [\![1, n]\!]^{\mathbb{N}}$*:*
  - *either a cryptographically secure pseudorandom number generator (PRNG),*
  - *or, in case of a binary input stream $m = m^0 \| m^1 \| m^1 \| \dots$ where $\forall i, |m^i| = n$, $\left(S(k)^i\right)_{i\in\mathbb{N}} = \left(m^k\right)_{i\in\mathbb{N}}$.*
- $\mathcal{K} = \mathbb{B}^{k_1} \times \mathbb{B}^{k_2} \times \mathbb{N}$ *called the* key space,
- *and $f : \mathbb{B}^n \longrightarrow \mathbb{B}^n$ a bijective map.*

We define the keyed hash function $\mathcal{H}_h : \mathcal{K} \times \mathbb{B}^* \longrightarrow \mathbb{B}^n$ by the following procedure

```
Inputs:  $k = (k_1, k_2, n) \in \mathcal{K}$
         $m \in \mathbb{B}^*$
Runs:    $X = h(k_1, m)$, or $X = h(k_1, m^0)$ if m is a stream
         for $i = 1, \dots, n$ :
             $X = G_f(S^i, X)$
         return X
```

$\mathcal{H}_h$ is thus a chaotic iteration based post-treatment on the inputted hash function h. The strategy is provided by a secured PRNG when the machine operates in a vacuum whereas it is redetermined at each iteration from the input stream in case of a finite machine open to the outside world. By doing so, we obtain a new hash function $\mathcal{H}_h$ with h, and this new one has a chaotic dependence regarding the inputted stream. Furthermore, we have stated that [10],

Theorem 4. *If h satisfies the collision resistance property, then it is the case too for $\mathcal{H}_h$. And if h satisfies the second-preimage resistance property, then it is the case too for $\mathcal{H}_h$.*

Finally, as $\mathcal{H}_h$ simply operates chaotic iterations with strategy $\mathcal{S}$ provided at each iterate by the media, we have [10]:

Theorem 5. *In case where the strategy $\mathcal{S}$ is the bitwise XOR between a secured PRNG and the input stream, the resulted hash function $\mathcal{H}_h$ is chaotic.*

6 Application to symmetric cryptography

Let us now present our last discoveries in the field of chaotic iteration based security. We have recently proven that the well-known Cipher Block Chaining (CBC) mode of operation, invented by IBM in 1976, can behave chaotically. The demonstration of this result is outlined thereafter, while details regarding the CBC mode of operation can be found in Patent [9].

Let us consider $\mathcal{X} = \mathbb{B}^{\mathsf{N}} \times \mathcal{S}_{\mathsf{N}}$, where:

- N is the size for the block cipher,
- $\mathcal{S}_{\mathsf{N}} = [\![0, 2^{\mathsf{N}} - 1]\!]^{\mathbb{N}}$, the set of infinite sequences of natural integers bounded by $2^{\mathsf{N}} - 1$, or the set of infinite N-bits block messages,

in such a way that $\mathcal{X}$ is constituted by couples of internal states of the mode of operation together with sequences of block messages. Let us consider the initial function:

$$\begin{array}{cccc} i: & \mathcal{S}_{\mathsf{N}} & \longrightarrow & [\![0, 2^{\mathsf{N}} - 1]\!] \\ & (m^i)_{i \in \mathbb{N}} & \longmapsto & m^0 \end{array}$$

that returns the first block of a (infinite) message, and the shift function:

$$\begin{array}{cccc} \sigma: & \mathcal{S}_{\mathsf{N}} & \longrightarrow & \mathcal{S}_{\mathsf{N}} \\ & (m^0, m^1, m^2, ...) & \longmapsto & (m^1, m^2, m^3, ...) \end{array}$$

which removes the first block of a message. Let m_j be the j-th bit of integer, or block message, $m \in [\![0, 2^{\mathsf{N}} - 1]\!]$ expressed in the binary numeral system, and when counting from the left. We define:

$$\begin{array}{cccc} F_f: & \mathbb{B}^{\mathsf{N}} \times [\![0, 2^{\mathsf{N}} - 1]\!] & \longrightarrow & \mathbb{B}^{\mathsf{N}} \\ & (x, m) & \longmapsto & \left(x_j m_j + f(x)_j \overline{m_j}\right)_{j=1..\mathsf{N}} \end{array}$$

This function returns the inputted binary vector x, whose m_j-th components x_{m_j} have been replaced by $f(x)_{m_j}$, for all $j = 1..\mathsf{N}$ such that $m_j = 1$. So the CBC mode of operation can be rewritten as the following dynamical system:

$$\left\{ \begin{array}{ll} X^0 = & (IV, m) \\ X^{n+1} = & \left(\mathcal{E}_k \circ F_{f_0}\left(i(X_1^n), X_2^n\right), \sigma(X_1^n)\right) \end{array} \right. \qquad (6)$$

where IV is the input vector, m the message to encrypt, and $\mathcal{E}_k$ the keyed symmetric cypher which has been selected. For $g : [\![0, 2^{\mathsf{N}} - 1]\!] \times \mathbb{B}^{\mathsf{N}} \longrightarrow \mathbb{B}^{\mathsf{N}}$, we denote $G_g(X) = (g(i(X_1), X_2); \sigma(X_1))$. So the reccurent relation of Eq.(6) can be rewritten in a condensed way, as follows.

$$X^{n+1} = G_{\mathcal{E}_k \circ F_{f_0}}(X^n). \qquad (7)$$

Using all this material, we have proven that:

Theorem 6. *Let $g = \varepsilon_k \circ F_{f_0}$, where ε_k is a given keyed block cipher and $f_0 : \mathbb{B}^{\mathsf{N}} \longrightarrow \mathbb{B}^{\mathsf{N}}$, $(x_1, ..., x_{\mathsf{N}}) \longmapsto (\overline{x_1}, ..., \overline{x_{\mathsf{N}}})$ is the Boolean vectorial negation. We consider the directed graph $\mathcal{G}_g$, where:*

- *vertices are all the N-bit words.*
- *there is an edge $m \in [\![0, 2^{\mathsf{N}} - 1]\!]$ from x to $\check{x}$ if and only if $g(m, x) = \check{x}$.*

So if $\mathcal{G}_g$ is strongly connected, then G_g is strongly transitive, and so the CBC mode of operation is chaotic.

7 Conclusion

In this article, the research works we have previously done in the field of chaotic iterations are summarized and clarified. Applications for pseudorandom number generation, hash function, and symmetric cryptography have then been outlined. Both theoretical analysis and experimental results confirm the feasibility of this approach.

References

[1] Jacques Bahi, Jean-François Couchot, and Christophe Guyeux. Quality analysis of a chaotic proven keyed hash function. *International Journal On Advances in Internet Technology*, 5(1):26–33, 2012.

[2] Jacques Bahi, Jean-François Couchot, Christophe Guyeux, and Adrien Richard. On the link between strongly connected iteration graphs and chaotic boolean discrete-time dynamical systems. In *FCT'11, 18th Int. Symp. on Fundamentals of Computation Theory*, volume 6914 of *LNCS*, pages 126–137, Oslo, Norway, August 2011.

[3] Jacques Bahi, Christophe Guyeux, and Qianxue Wang. A novel pseudo-random generator based on discrete chaotic iterations. In *INTERNET'09, 1-st Int. Conf. on Evolving Internet*, pages 71–76, Cannes, France, August 2009.

[4] Jacques M. Bahi, Raphaël Couturier, Christophe Guyeux, and Pierre-Cyrille Héam. Efficient and cryptographically secure generation of chaotic pseudorandom numbers on GPU. *CoRR*, abs/1112.5239, 2011.

[5] Jacques M. Bahi and Christophe Guyeux. Topological chaos and chaotic iterations, application to hash functions. In *WCCI'10, IEEE World Congress on Computational Intelligence*, pages 1–7, Barcelona, Spain, July 2010. Best paper award.

[6] Jacques M. Bahi, Christophe Guyeux, and Qianxue Wang. Improving random number generators by chaotic iterations. Application in data hiding. In *ICCASM 2010, Int. Conf. on Computer Application and System Modeling*, pages V13–643–V13–647, Taiyuan, China, October 2010.

[7] J. Banks, J. Brooks, G. Cairns, and P. Stacey. On Devaney's definition of chaos. *Amer. Math. Monthly*, 99:332–334, 1992.

[8] Robert L. Devaney. *An Introduction to Chaotic Dynamical Systems*. Addison-Wesley, Redwood City, CA, 2nd edition, 1989.

[9] W.F. Ehrsam, C.H.W. Meyer, J.L. Smith, and W.L. Tuchman. Message verification and transmission error detection by block chaining, February 14 1978. US Patent 4,074,066.

[10] Christophe Guyeux, Qianxue Wang, Xiaole Fang, and Jacques Bahi. Introducing the truly chaotic finite state machines and theirs applications in security field. In *NOLTA 2014, 24th International Symposium on Nonlinear Theory and its Applications*, pages ***–***, Luzern, Switzerland, September 2014. Short paper. To appear.

[11] Qianxue Wang, Jacques Bahi, Christophe Guyeux, and Xiaole Fang. Randomness quality of CI chaotic generators. application to internet security. In *INTERNET'2010. The 2nd Int. Conf. on Evolving Internet*, pages 125–130, Valencia, Spain, September 2010. IEEE Computer Society Press. Best Paper award.