Hash Functions Using Chaotic Iterations
Jacques Bahi, Christophe Guyeux

To cite this version:

HAL Id: hal-00563113
https://hal.archives-ouvertes.fr/hal-00563113
Submitted on 7 Feb 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Hash Functions Using Chaotic Iterations

Jacques M. BAHI, Christophe GUYEUX
University of Franche-Comte, IUT Belfort
2 rue Engel Gros, 90016 Belfort, France
Email: jacques.bahi@univ-fcomte.fr, christophe.guyeux@univ-fcomte.fr

Received 01/04/2009; Accepted 02/06/2009

ABSTRACT

In this paper, a novel formulation of discrete chaotic iterations in the field of dynamical systems is given. Their topological properties are studied: it is mathematically proven that, under some conditions, these iterations have a chaotic behavior as defined by Devaney. This chaotic behavior allows us to propose a way to generate new hash functions. An illustrative example is detailed in order to show how to use our theoretical study in practice.

Keywords: Discrete dynamical systems, chaotic iterations. Devaney’s chaos, hash functions.

1. INTRODUCTION

Chaotic iterations have been introduced on the one hand by Chazan and Miranker [1] in a numerical analysis context and on the other hand by Robert [2] in the discrete dynamical systems framework. The goal was to derive sufficient conditions ensuring the convergence (or the stability) of such iterations. In this paper, a new point of view is presented: the goal here is to study the conditions under which these iterations admit a chaotic behavior. Contrary to the previous studies, convergence or stability are not sought.

This article presents the research results related to this question. We prove that under some conditions, discrete chaotic iterations produce chaos, precisely, they produce topological chaos as described by Devaney. This topological chaos is a rigorous and well studied framework in the field of mathematical theory of chaos. Behind the theoretical interest connecting the field of the chaotic discrete iterations and the one of topological chaos, our study gives a framework making it possible to create hash functions that can be mathematically evaluated and compared. A hash function is a transformation that takes a variable-size input and returns a fixed-size string, which is called the hash value. They are used to speed up table lookup or data comparison tasks and for digital signature. Hash functions, such as MD5 or SHA-256, can be described by discrete iterations on a finite set. In this paper, the elements of this finite set are called cells. These cells
represent the blocks of the text to which the hash function will be applied. Some required qualities for hash functions such as the avalanche effect, resistance to collisions, and unpredictability can be mathematically described by notions from the theory of topological chaos, namely, sensitivity, transitivity, entropy, and expansivity [3], [4], [5]. These concepts are approached but non deepened in this article. More detailed studies will be carried out in forthcoming articles.

This study is the first of a series we intend to carry out. We think that the mathematical framework in which we are placed offers interesting new tools allowing the conception, the comparison, and the evaluation of new algorithms in computer security framework, not only hash functions.

The rest of the paper is organized as follows. The first next section is devoted to some recalls on two distinct domains: topological chaos and discrete chaotic iterations. The third and fourth sections constitute the theoretical study of the present paper. The topological framework is defined and the proof that chaotic iterations have a topological chaos behavior is given. Section 5 details how it is possible to apply chaotic results in the computer science framework. Its following section contains the application to hash functions and an illustrative example. The paper ends by some discussions and future work.

2. BASIC RECALLS
This section is devoted to basic definitions and terminologies in the field of topological chaos and in the one of chaotic iterations.

2.1 Devaney’s Chaotic Dynamical Systems
Consider a metric space \((X, d)\) and a continuous function \(f : X \to X\).

**Definition 1** \(f\) is said to be topologically transitive if, for any pair of open sets \(U, V \subset X\), there exists \(k > 0\) such that \(f^k(U) \cap V \neq \emptyset\).

**Definition 2** An element (a point) \(x\) is a periodic element (point) for \(f\) of period \(n \in \mathbb{N}\), if \(f^n(x) = x\). The set of periodic points of \(f\) is denoted \(\text{Per}(f)\).

**Definition 3** \((X, f)\) is said to be regular if the set of periodic points is dense in \(X\),

\[
\forall x \in X, \forall \varepsilon > 0, \exists p \in \text{Per}(f), d(x, p) \leq \varepsilon.
\]

**Definition 4** \(f\) has sensitive dependence on initial conditions if there exists \(\delta > 0\) such that, for any \(x \in X\) and any neighborhood \(V\) of \(x\), there exists \(y \in V\) and \(n \geq 0\) such that \(|f^n(x) - f^n(y)| > \delta\). \(\delta\) is called the constant of sensitivity of \(f\).
Let us now recall the definition of a chaotic topological system, as described by Devaney [4]:

**Definition 5** \( f : \mathcal{X} \to \mathcal{X} \) is said to be *chaotic* on \( \mathcal{X} \) if,

1. \( f \) has sensitive dependence on initial conditions,
2. \( f \) is topologically transitive,
3. \((\mathcal{X}, f)\) is regular. \( \square \)

Therefore, quoting Robert Devaney: “A chaotic map possesses three ingredients: unpredictability, indecomposability and an element of regularity. A chaotic system is unpredictable because of the sensitive dependence on initial conditions. It cannot be broken down or decomposed into two subsystems, because of topological transitivity. And, in the midst of this random behavior, we nevertheless have an element of regularity, namely the periodic points which are dense.” Fundamentally different behaviors are then possible and occur with an unpredictably way.

### 2.2 Chaotic Iterations

In the sequel \( S^n \) denotes the \( n^{th} \) term of a sequence \( S \), \( V_i \) denotes the \( i^{th} \) component of a vector \( V \), and \( f^k = f \circ \ldots \circ f \) denotes the \( k^{th} \) composition of a function \( f \). Finally, the following notation is used: \([1;N] = \{1, 2, \ldots, N\}\). Let us consider a *system* of a finite number \( N \) of *cells*, so that each cell has a boolean *state*. Then a sequence of length \( N \) of boolean states of the cells corresponds to a particular *state* of the system. A sequence which elements belong in \([1;N]\) is called a *strategy*. The set of all strategies is denoted by \( \mathbb{S} \).

**Definition 6** Let \( S \in \mathbb{S} \). The *shift* function 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 the *initial function* \( i \) is the map which associates to a sequence, its first term: \( i : (S^n)_{n \in \mathbb{N}} \in \mathbb{S} \to S^0 \in [1;N] \). \( \square \)

**Definition 7** The set \( \mathbb{B} \) denoting \{0,1\}, let \( f : \mathbb{B}^N \to \mathbb{B}^N \) be a function and \( S \in \mathbb{S} \) be a strategy. Then, the so-called *chaotic iterations* are defined by

\[
\begin{align*}
x^0 & \in \mathbb{B}^N, \\
\forall n \in \mathbb{N}^*, \forall i \in [1;N], & x^n = \begin{cases} x^{n-1}_i & \text{if } S^n \neq i. \\
(f(x^{n-1}))_{S^n} & \text{if } S^n = i. \end{cases} \quad (1)
\end{align*}
\]
In other words, at the \( n \)-th iteration, only the \( S_n \)-th cell is “iterated”. Note that in a more general formulation, \( S_n \) can be a subset of components and \( f(x^n \{ n-1 \}) \), can be replaced by \( f(x^k) \), where \( k < n \), describing for example delay in transmissions (see, e.g., [6], or [7]). For the general definition of such chaotic iterations, see, e.g., [2].

3. THE NEW TOPOLOGICAL SPACE

In this section we will put our study in a topological context by defining a suitable metric space where chaotic iterations are continuous.

3.1 Defining the Iteration Function and the Phase Space

Denote by \( \delta \) the discrete boolean metric, \( \delta(x, y) = 0 \Leftrightarrow x = y \). Given a function \( f \), define the function

\[
F_f : \{1; N\} \times \mathbb{B}^N \rightarrow \mathbb{B}^N
\]

\[
(k, E) \rightarrow (E_j \cdot \delta(k, j) + f(E)_j \cdot \overline{\delta(k, j)})_{j \in 1:N},
\]

where + and . are the boolean addition and product operations. Consider the phase space:

\[
\mathcal{X} = \{1; N\} \times \mathbb{B}^N
\]

and the map defined on \( \mathcal{X} \):

\[
G_f(S, E) = (\sigma(S), F_f(i(S), E)).
\]

Then the chaotic iterations defined in (1) can be described by the following iterations

\[
\begin{cases}
X^0 \in \mathcal{X} \\
X^{k+1} = G_f(X^k).
\end{cases}
\]

3.2 Cardinality of \( \mathcal{X} \)

By comparing \( \mathbb{S} \) and \( \mathbb{R} \), we have the result.

**Proposition 1** The phase space \( \mathcal{X} \) has, at least, the cardinality of the continuum.

**Proof.** Let \( \varphi \) be the map which transforms a strategy into the binary representation of an element in \([0,1]\), as follows. If the \( n \)-th term of the strategy is 0, then the \( n \)-th associated digit is 0, else it is equal to 1.
With this construction, \( \varphi : [1;N]^N \rightarrow [0,1] \) is onto. But \( ]0,1[ \) is isomorphic to \( \mathbb{R} \) (\( x \in ]0,1[ \mapsto \tan(\pi(x-\frac{1}{2})) \) is an isomorphism), so the cardinal of \([1;N]^N\) is greater or equal than the cardinality of \( \mathbb{R} \). So the cardinal of the Cartesian product \( \mathcal{X} = [1;N]^N \times \mathbb{B}^N \) is greater or equal to the cardinality of \( \mathbb{R} \).

\[ \square \]

Remark 1. This result is independent on the number of cells of the system. \( \square \)

### 3.3 A New Distance

We define a new distance between two points \( X = (S,E), Y = (\tilde{S},\tilde{E}) \in \mathcal{X} \) by

\[
d(X,Y) = d_s(E,\tilde{E}) + d_s(S,\tilde{S}),
\]

where

\[
\begin{align*}
    d_s(E,\tilde{E}) &= \sum_{k=1}^N \delta(E_k,\tilde{E}_k), \\
    d_s(S,\tilde{S}) &= \frac{9}{10^k} \sum_{k=1}^N |S_k,\tilde{S}_k| \\
\end{align*}
\]

If the floor value \( \lfloor d(X, Y) \rfloor \) is equal to \( n \), then the systems \( E, \tilde{E} \) differ in \( n \) cells. In addition, \( d(X, Y) - \lfloor d(X, Y) \rfloor \) is a measure of the differences between strategies \( S \) and \( \tilde{S} \). More precisely, this floating part is less than \( 10^{-k} \) if and only if the first \( k \) terms of the two strategies are equal. Moreover, if the \( k^{th} \) digit is nonzero, then the \( k^{th} \) terms of the two strategies are different.

### 3.4 Continuity of the Iteration Function

To prove that chaotic iterations are an example of topological chaos as defined by Devaney [4], \( G_f \) must be continuous on the metric space \((\mathcal{X}, d)\).

Theorem 1. \( G_f \) is a continuous function. \( \square \)

**Proof.** We use the sequential continuity.

Let \((S^n, E^n)_{n \in \mathbb{N}}\) be a sequence of the phase space \( \mathcal{X} \), which converges to \((S, E)\). We will prove that \((G_f(S^n, E^n))_{n \in \mathbb{N}}\) converges to \((G_f(S, E))\). Let us recall that for all \( n \), \( S^n \) is a strategy, thus, we consider a sequence of strategies (i.e. a sequence of sequences).

As \( d((S^n, E^n); (S, E)) \) converges to 0, each distance \( d_s(E^n, E) \) and \( d_s(S^n, S) \) converges to 0. But \( d_s(E^n, E) \) is an integer, so \( \exists n_0 \in \mathbb{N}, d_s(E^n, E) = 0 \) for any \( n \geq n_0 \).
In other words, there exists a threshold \( n_0 \in \mathbb{N} \) after which no cell will change its state:

\[
\exists n_0 \in \mathbb{N}, \quad n \geq n_0 \implies E^n = E.
\]

In addition, \( d_s(S^n, S) \to 0 \), so \( \exists n_1 \in \mathbb{N}, \quad d_s(S^n, S) < 10^{-1} \) for all indexes greater than or equal to \( n_1 \). This means that for \( n \geq n_1 \), all the \( S^n \) have the same first term, which is \( S^0 \):

\[
\forall n \geq n_1, \quad S^n_0 = S_0.
\]

Thus, after the \( \max(n_0, n_1) \)th term, states of \( E^n \) and \( E \) are identical, and strategies \( S^n \) and \( S \) start with the same first term.

Consequently, states of \( G_f(S^n, E^n) \) and \( G_f(S, E) \) are equal, so, after the \( \max(n_0, n_1) \)th term, the distance \( d \) between these two points is strictly less than 1.

We now prove that the distance between \( (G_f(S^n, E^n)) \) and \( (G_f(S, E)) \) is convergent to 0. Let \( \varepsilon > 0 \).

- If \( \varepsilon \geq 1 \), we have seen that distance between \( (G_f(S^n, E^n)) \) and \( (G_f(S, E)) \) is strictly less than 1 after the \( \max(n_0, n_1) \)th term (same state).
- If \( \varepsilon < 1 \), then \( \exists k \in \mathbb{N}, \quad 10^{-k} \geq \varepsilon \implies 10^{-(k+1)} \). But \( d_s(S^n, S) \) converges to 0, so

\[
\exists n_2 \in \mathbb{N}, \quad \forall n \geq n_2, \quad d_s(S^n, S) < 10^{-(k+2)},
\]

thus after \( n_2 \), the \( k + 2 \) first terms of \( S^n \) and \( S \) are equal.

As a consequence, the \( k + 1 \) first entries of the strategies of \( G_f(S^n, E^n) \) and \( G_f(S, E) \) are the same (\( G_f \) is a shift of strategies) and due to the definition of \( d_s \), the floating part of the distance between \( (S^n, E^n) \) and \( (S, E) \) is strictly less than \( 10^{-(k+1)} \) \( \geq \varepsilon \).

In conclusion,

\[
\forall \varepsilon > 0, \quad \exists N_0 = \max(n_0, n_1, n_2) \in \mathbb{N}, \quad \forall n \geq N_0, \quad d(G_f(S^n, E^n); G_f(S, E)) \leq \varepsilon.
\]

\( G_f \) is consequently continuous. \( \diamond \)

In this section, we proved that chaotic iterations can be modeled as a dynamical system in a topological space. In the next section, we show that chaotic iterations are a case of topological chaos, following the definition of Devaney.
4. DISCRETE CHAOTIC ITERATIONS AS TOPOLOGICAL CHAOS

To prove that we are in the framework of Devaney’s topological chaos, we have to check the regularity, transitivity and sensitivity conditions. We will prove that the vectorial logical negation function

$$ f_0(x_1, ..., x_N) = (\bar{x}_1, ..., \bar{x}_N) \quad (3) $$

satisfies these hypotheses.

4.1 Regularity

**Proposition 2** Periodic points of \( G_{f_0} \) are dense in \( \mathcal{X} \).

**Proof.** Let \((\bar{S}, \bar{E}) \in \mathcal{X}\) and \( \epsilon > 0 \). We are looking for a periodic point \((\bar{S}, \bar{E})\) satisfying \( d((\bar{S}, \bar{E}); (\bar{S}, \bar{E})) < \epsilon \).

As \( \epsilon \) can be strictly lesser than 1, we must choose \( \bar{E} = \bar{E} \). Let us define \( k_0(\epsilon) = [\log_{10}(\epsilon)] + 1 \) and consider the set

$$ S_{\bar{S}, \bar{E}_0(\epsilon)} = \{ S \in \mathcal{S} / S^k = \bar{S}^k, \forall k \leq k_0(\epsilon) \}. $$

Then, \( \forall S \in S_{\bar{S}, \bar{E}_0(\epsilon)}, d((S, E); (\bar{S}, \bar{E})) < \epsilon \). It remains to choose \( \bar{S} \in S_{\bar{S}, \bar{E}_0(\epsilon)} \) such that \((\bar{S}, \bar{E}) = (\bar{S}, \bar{E})\) is a periodic point for \( G_{f_0} \).

Let \( \mathcal{J} = \{ i \in \{1, 2, ..., N\} / E_i \neq \bar{E}_i \}, \) where \((S, E) = G_{f_0}^i(\bar{S}, \bar{E})\), \( i_0 = \text{card}(\mathcal{J}) \) and \( j_1 < j_2 < ... < j_m \) the elements of \( \mathcal{J} \). Then, \( \bar{S} \in S_{\bar{S}, \bar{E}_0(\epsilon)} \) defined by

- \( \bar{S}^k = \bar{S}^k \), if \( k \leq k_0(\epsilon) \),
- \( \bar{S}^k = \bar{S}_{k-k_0(\epsilon)} \), if \( k \in \{ k_0(\epsilon) + 1, k_0(\epsilon) + 2, ..., k_0(\epsilon) + i_0 \} \),
- and \( \bar{S}^k = \bar{S}^j \), where \( j \leq k_0(\epsilon) + i_0 \) is satisfying \( j \equiv k \mod (k_0(\epsilon) + i_0) \), if \( k > k_0(\epsilon) + i_0 \),

is such that \((\bar{S}, \bar{E})\) is a periodic point, of period \( k_0(\epsilon) + i_0 \), which is \( \epsilon \)-close to \((\bar{S}, \bar{E})\).

As a conclusion, \((\mathcal{X}, G_{f_0})\) is regular. \( \diamond \)

4.2 Transitivity

**Proposition 3** \((\mathcal{X}, G_{f_0})\) is topologically transitive.

**Proof.**
PROOF 4. Let us define $\mathcal{E} : \mathcal{X} \to \mathbb{B}^\mathbb{N}$, such that $\mathcal{E}(S, E) = E$. Let $B_A = B(X_A, r_A)$ and $B_B = B(X_B, r_B)$ be two open balls of $\mathcal{X}$, with $X_A = (S_A, E_A)$ and $X_B = (S_B, E_B)$. We are looking for $\tilde{X} = (\tilde{S}, \tilde{E})$ in $B_A$ such that $\exists n_0 \in \mathbb{N}$, $G_{f_0}^{n_0}(X) \in B_B$. $\tilde{X}$ must be in $B_A$ and $r_A$ can be strictly lesser than 1, so $\tilde{E} = E_A$. Let $k_0 = \left\lfloor \log_{10}(r_A) + 1 \right\rfloor$. Then $\forall S \in S$, if $S^k = S_A^k$, $\forall k \leq k_0$, then $(S, \tilde{E}) \in B_A$. Let us notice $(\tilde{S}, \tilde{E}) = G_{f_0}^{k_0}(S_A, E_A)$ and $c_1, \ldots, c_{k_1}$ the elements of the set $\{i \in [1, N] / \tilde{E}_i \notin \mathcal{E}(X_B_i)\}$. So any point $X$ of the set

$$(S, E_A) \in \mathcal{X} / \forall k \leq k_0, S^k = S_A^k \text{ and } \forall k \in [1, k_1], S^{k_0+k} = c_k$$

is satisfying $X \in B_A$ and $\mathcal{E}(G_{f_0}^{k_0+k_1}(X)) = E_B$.

Lastly, let us define $k_2 = \left\lfloor \log_{10}(r_B) + 1 \right\rfloor$. Then $\tilde{X} = (\tilde{S}, \tilde{E}) \in \mathcal{X}$ defined by:

1. $\tilde{X} = E_A$,
2. $\forall k \leq k_0, \tilde{S}^k = S_A^k$,
3. $\forall k \in [1, k_1], \tilde{S}^{k_0+k} = c_k$,
4. $\forall k \in \mathbb{N}^*, \tilde{S}^{k_0+k} = S_B^k$,

is such that $\tilde{X} \in B_A$ and $G_{f_0}^{k_0+k_1}(\tilde{X}) \in B_B$. $\diamond$

4.3 Sensitive Dependence on Initial Conditions

**Proposition 4** $(\mathcal{X}, G_f)$ has sensitive dependence on initial conditions. $\boxdot$

**PROOF 5.** Banks *et al.* proved in [8] that having sensitive dependence is a consequence of being regular and topologically transitive. $\diamond$

4.4 Devaney’s Chaos

In conclusion, $(\mathcal{X}, G_f)$ is topologically transitive, regular and has sensitive dependence on initial conditions. Then we have the result.

**Theorem 2** $G_f$ is a chaotic map on $(\mathcal{X}, d)$ as defined by Devaney. $\boxdot$

**Remark 2** We have proven that the set of the iterate functions $f$ such that $(\mathcal{X}, G_f)$ is chaotic (in the meaning of Devaney), is a nonempty set. In a future work, we will give a characterization of this set.
5. CHAOS IN COMPUTER SCIENCE

It is worthwhile to notice that even if the set of machine numbers is finite, we deal with the infinite set of strategies that have a finite but unbounded lengths. Indeed, it is not necessary to store all the terms of the strategy in the memory, only the $n^{th}$ term (an integer less than or equal to $N$) of the strategy has to be stored at the $n^{th}$ step, as it is illustrated in the following example. Let us suppose that a given text is input from the outside world into the computer character by character and that the current term of the strategy is computed from the ASCII code of the current stored character. Then, as the set of all possible texts of the outside world is infinite and the number of their characters is unbounded, we have to deal with an infinite set of finite but unbounded strategies.

Of course, the previous example is a simplistic illustrating example. A chaotic procedure should be introduced to generate the terms of the strategy from the stream of characters.

Then in the computer science framework, we also have to deal with a finite set of states of the form $B^N$ and as stated before an infinite set $S$ of strategies. The sole difference with the previous study is that, instead of being infinite the sequences of $S$ are finite with unbounded length.

The proofs of continuity and transitivity are independent of the finiteness of the length of strategies (sequences of $S$). In addition, it is possible to prove the sensitivity property in this situation. So even in the case of finite machine numbers, we have the two fundamental properties of chaos: sensitivity and transitivity, which respectively implies unpredictability and indecomposability (see [4], p.50). The regularity supposes that the sequences are of infinite lengths. To obtain the analogous of regularity in the context of finite sets, we define below the notion of periodic but finite sequences.

**Definition 8** A strategy $S \in S$ is said to be periodic but finite if $S$ is a finite sequence of length $n$ and if there exists a divisor $p$ of $n$, $p \neq n$, such that $\forall i \leq n - p$, $S^i = S^{i+p}$. A point $(E, S) \in X$ is said to be periodic but finite, if its strategy $S$ is periodic but finite.

For example, $(1, 2, 1, 2, 1, 2, 1, 2)$ ($p = 2$) and $(2, 2, 2)$ ($p = 1$), are periodic but finite. This definition can be interpreted as the analogous of periodicity on finite strategies. Then, following the proof of regularity (section 4.1), it can be proven that the set of periodic but finite points is dense on $X$, hence obtaining a desired element of regularity in finite sets, as quoted by Devaney ([4], p.50): “two points arbitrary close to each other could have completely different behaviors, the one could have a cyclic behavior as long as the system iterates..."
while the trajectory of the second could ‘visit’ the whole phase space”. It should be recalled that the regularity was introduced by Devaney in order to counteract the effects of sensitivity and transitivity: two points close to each other can have fundamentally different behaviors.

In conclusion, even in the computer science framework our previous theory applies. In what follows, an example of the use of chaotic iterations in the field of computer science is given.

6. HASH FUNCTIONS BASED ON TOPOLOGICAL CHAOS

6.1 Introduction

The use of chaotic map to generate hash algorithm is a recent idea. In [9] for example, a digital signature algorithm based on elliptic curve and chaotic mappings is proposed to strengthen the security of an elliptic curve digital signature algorithm. Other examples of the generation of an hash function using chaotic maps can be found in [10] and [11].

We define in this section a new way to construct hash functions based on chaotic iterations. As a consequence of the previous theory, generated hash functions satisfy the topological chaos property. Thus, this approach guarantees to obtain various desired properties in the domain of hash functions. For example, the avalanche criterion is closely linked to the sensitivity property.

The hash value will be the last state of some chaotic iterations: initial state $X_0$, finite strategy $S$ and iterate function must then be defined.

6.2 Initial State

The initial condition $X_0 = (S, E)$ is composed by:

- A N = 256 bits sequence $E$ obtained from the original text.
- A chaotic strategy $S$.

In the sequel, we describe in detail how to obtain this initial condition.

6.2.1 How to Obtain $E$

The first step of our algorithm is to transform the message in a normalized 256 bits sequence $E$. To illustrate this step, we take an example, our original text is: The original text

Each character of this string is replaced by its ASCII code (on 7 bits). Then, we add a 1 to this string.
So, the binary value (1111000) of the length of this string (120) is added, with another 1:

10101001 10100011 00101010 00001101 11111100 10110100
11100111 11010011 10111011 00001110 11000100 00011101
00110010 11110000 11101001

The whole string is copied, but in the opposite direction. This gives:

10101001 10100011 00101010 00001101 11111100 10110100
11100111 11010011 10111011 00001110 11000100 00011101
00110010 11111000 11101001 11110001
00011111 00101110
00111110 10011001 01110000 01001110 01100001 10111011
10010111 11001110 01011010 01111011 01100000 10101001
10001011

So, we obtain a multiple of 512, by duplicating enough this string and truncating at the next multiple of 512. This string, in which the whole original text is contained, is denoted by \( D \).

Finally, we split our obtained string into blocks of 256 bits and apply to them the exclusive-or function, obtaining a 256 bits sequence.

11111010 11000101 01111110 00010110 00000101 11011101
00101000 01110100 11001101 00010011 01001100 00100111
01010111 00001001 01110101 00010011 00100001 01110010
01000011 10101011 10010000 11001011 00100010 11001100
10111000 01010010 11011110 10000001 10100001 11111010
10011101 01111101

So, in the context of subsection (1), \( N = 256 \) and \( E \) is the above obtained sequence of 256 bits.

We now have the definitive length of our digest. Note that a lot of texts have the same string. This is not a problem because the strategy we will build will depend on the whole text.

Let us build now the strategy \( S \).
6.2.2 How to Choose \( S \)

To obtain the strategy \( S \), an intermediate sequence \( (u^n) \) is constructed from \( D \), as follows:

1. \( D \) is split into blocks of 8 bits. Then \( u^n \) is the decimal value of the \( n^{th} \) block.
2. A circular rotation of one bit to the left is applied to \( D \) (the first bit of \( D \) is put on the end of \( D \)). Then the new string is split into blocks of 8 bits another time. The decimal values of those blocks are added to \( (u^n) \).
3. This operation is repeated again 6 times.

It is now possible to build the strategy \( S \):

\[
S^0 = u^0, \quad S^n = (u^n + 2 \times S^{n-1} + n) \pmod{256}.
\]

\( S \) will be highly dependent to the changes of the original text, because \( \theta \rightarrow 2\theta \pmod{1} \) is known to be chaotic as described by Devaney [4].

6.2.3 How to Construct the Digest

To construct the digest, chaotic iterations are done with initial state \( X^0 \),

\[
f: \quad \{E_1, \ldots, E_{256}\} \rightarrow \{E_1, \ldots, E_{256}\},
\]

as iterate function and \( S \) for the chaotic strategy.

The result of these iterations is a 256 bits vector. Its components are taken 4 per 4 bits and translated into hexadecimal numbers, to obtain the hash value:

63A88CB6AF0B18E3BE828F9BDA4596A6A13DFE38440AB9557DA1C0C6B1EDBBBD

As a comparison if instead of considering the text “The original text” we took “the original text”, the hash function returns:

33E0DFB5BB1D88C924D2AF80B14FF5A7B1A3DEF9D0E831194BD814C8A3B948B3

In this paper, the generation of hash value is done with the vectorial boolean negation \( f_0 \) defined in eq. (3). Nevertheless, the procedure remains general and can be applied with any function \( f \) such that \( G_f \) is chaotic.

In the following subsection, a complete example of the procedure is given.
6.3 Application Example
Consider the following message [12]:

Wanderers in that happy valley,
    Through two luminous windows, saw
Spirits moving musically,
    To a lute's well-tuned law,
Round about a throne where, sitting
    (Porphyrogene!)
In state his glory well befitting,
    The ruler of the realm was seen.

And all with pearl and ruby glowing
    Was the fair palace door,
Through which came flowing, flowing,
    And sparkling evermore,
A troop of Echoes, whose sweet duty
    Was but to sing,
In voices of surpassing beauty,
    The wit and wisdom of their king.

Its hash value is:

FF51DA4E7E50FBA7A8DC6858E9EC3353BDE2E465E1A6A1B03BEAA12A4AD694FB

As a comparison, if an additional space is put before "Was the fair palace door," the hash value will be:

03ABFA49B834D529669CFC1AECC13E14EA5FFD2349582380BCBDBF8400017445

and if "Echoes" is replaced by "echoes" in the original text:

FE54777C52D373B7AED2EA5ACAD422B5B563BB3B91E8FCB48AAE9331DAC54A9B

Those examples give an illustration of the avalanche effect obtained by this algorithm. A more complete study of the properties possessed by our hash functions and resistance under collisions will be studied in a future work.
7. CONCLUSION

We proved that discrete chaotic iterations behave as Devaney’s topological chaos if the iteration function is the vectorial boolean negation function. We applied these results to the generation of new hash functions. The vectorial boolean negation function has been chosen here, but the process remains general and other iterate functions $f$ can be used. The sole condition is to prove that $G_f$ satisfies Devaney’s chaos property.

By considering hash functions as an application of our theory, we shown how some desirable aspects in computer security field such as unpredictability, sensitivity to initial conditions, mixture, and disorder can be mathematically guaranteed and even quantified by mathematical tools.

Theory of chaos recalls us that simple functions can have, when iterated, a very complex behavior, while some complicated functions could have foreseeable iterations. This is why it is important to have tools for evaluating desired properties. Our simple function may be replaced by other “chaotic” functions which can be evaluated with quantitative tools, like the constant of sensitivity. Another important parameter is the choice of the strategy $S$. We proposed a particular strategy that can be easily improved by multiple ways.

Much work remains to be made. For example we are convinced that the good comprehension of the transitivity property, enables to study the problem of collisions in hash functions.

In future work we plan to investigate other forms of chaos such as Li-York [3] or Knudsen [5] chaos. Other quantitative and qualitative tools like expansivity or entropy (see e.g. [13]) will be explored and the domain of applications of our theoretical concepts will be enlarged.

REFERENCES


