2005 International Symposium on Nonlinear
Theory and its Applications (NOLTA2005)
Bruges, Belgium, October 18-21, 2005

Discrete Chaos and Cryptography

J.M. Amigó(1), J. Szczepanski(2) and L. Kocarev(3)

(1)Centro de Investigación Operativa, Universidad Miguel Hernández
Avda. de la Universidad, 03202 Elche, Spain,
(2) Institute of Fundamental Technological Research, Polish Academy of Science
Swietokrzyska 21, 00-049 Warsaw, Poland,
(3) Institute for Nonlinear Science, University of California San Diego
9500 Gilman Drive, La Jolla CA 92093-0402, USA,
Email: jm.amigo@umh.es, jszczepa@ippt.gov.pl, lkocarev@ucsd.edu.

Abstract—We propose definitions of discrete Lyapunov exponent and discrete entropy for permutations on a finite set. We justify our definitions by proving that in the 'infinite limit', i.e., when the cardinality M of the set goes to infinity, the discrete concepts converge to their continuous counterparts for a large class of chaotic maps. Consequently, we say that a discrete-time dynamical system on a finite-state phase space is discretely chaotic if its discrete Lyapunov exponent tends to a positive number (or to ∞) when $M \to \infty$. Possible applications of discrete chaos to cryptography are also discussed.

1. Introduction

What makes chaotic systems so attractive both for theoreticians and practitioners is their random-like behavior — in spite of being deterministic. As way of illustration, let us mention that, already in 1949, C. Shannon [1] proposed this kind of transformations to construct secure cryptosystems. It is thus no surprise that, when chaos theory flourished in the nineteen-eighties and -nineties, several cryptosystems were proposed based on the discretization of chaotic maps. Viewing how the resulting permutations mix the pixels of digital pictures [2], one cannot but admit that their 'confusion' and 'diffusion' properties are seemingly unsurpassed —in spite of being periodic. The examples could be multiplied with the same message: there must be some sense in which discrete maps may be also called chaotic.

The authors of this communication have tried to come to grips with the concept of discrete chaos by proposing a first tool to measure it, namely, the discrete Lyapunov exponent. As its continuous counterpart, the discrete Lyapunov exponent measures the local (i.e., between neighboring points) average spreading of the discrete-time discrete-space dynamical system considered. Discrete chaos plays an important role in numerical computation, cryptography, digital electronics and communications and, potentially, whenever a complex continuous phenomenon is implemented on a finite-state machine. In most modern block ciphers including both the former and current standards for commercial encryption DES and AES, the confusion-diffusion strategy proposed by Shannon is implemented, roughly speaking, by means of bit permutations with strong nonlinearity (S-boxes) on subblocks of the input block and permutations with fast spreading factor on whole blocks, respectively. This being the case, the security of all these ciphers relies ultimately on such permutations delivering the right mixing and propagation properties. Here is where discrete chaos comes in: it provides tools like the discrete Lyapunov exponent and entropy to quantify the said properties. The design and certification of special-purpose permutations is just an example of possible and interesting applications of discrete chaos to cryptography. Others include the design of cryptologic algorithms, hash functions and the like.

2. The Tools of Discrete Chaos

2.1. The discrete Lyapunov Exponent

Let $\mathcal{S} = \{s_1, ..., s_M\}$ be a linearly ordered finite set endowed with a metric $d(\cdot,\cdot)$, and $F : \mathcal{S} \to \mathcal{S}$ be a bijection or equivalently, an M-permutation. We define the discrete Lyapunov exponent (DLE) of F as

$$\lambda_F = \frac{1}{M-1}\sum_{i=1}^{M-1} \log \frac{d\left(F(s_{i+1}), F(s_i)\right)}{d\left(s_{i+1}, s_i\right)}.$$

Following the tradition, we will use natural logarithms to calculate λ_F. Observe that λ_F depends on the order and on the metric d but is invariant under rescaling and, furthermore, has the same invariances as d.

In the examples and applications we will consider below, F will be a permutation on a subset $\mathcal{S}$ of $\mathbb{R}$ (for instance, $\mathcal{S}$ = $\{0, ..., M-1\} \equiv \mathbb{Z}_M$) endowed with the Euclidean distance $d(s_i, s_j) = \left|s_i - s_j\right|$. The case $\mathcal{S} = \mathbb{Z}_2^l$, with $\mathbb{Z}_2^l \equiv \mathbb{Z}_2\times...\times\mathbb{Z}_2$ (l times) being the set of binary strings of length k endowed with Hamming distance, has also interest for cryptographic applications, but it will not be considered here. Observe that if $\mathcal{S} = \mathbb{Z}_M$ or $\mathcal{S} = \mathbb{Z}_2^l$ lexicographically ordered, then $\lambda_F \geq 0$.

Example 1. Suppose that $M = 2m$ and define

$$F_{\max}(s) = \begin{cases} m+k & \text{if } s = 2k \quad 0 \leq k \leq m-1 \\ k & \text{if } s = 2k+1 \quad 0 \leq k \leq m-1 \end{cases}$$

on $\mathbb{Z}_M$. The DLE of $F_{\max}$ is

$$\lambda_{F_{\max}} = \frac{m}{2m-1}\ln m + \frac{m-1}{2m-1}\ln(m+1). \qquad (1)$$

It can be shown [3] that $F_{\max}$ has the largest DLE among all permutations of the set $\{0, ..., M-1\}$.

Let $z_{j+1} = f(z_j)$, $j = 0, 1, ..., M-1$, be a typical trajectory of length M of a one-dimensional chaotic map $f : [0,1] \to [0,1]$, such that $z_{j+1} \neq z_j$ for all j and $|z_{M-1} - z_0| < \varepsilon$. We define $f(z_{M-1}) = z_0$ and order z_j according to the metric to obtain x_j, that is, $x_0 < x_1 < ... < x_{M-1}$, so that x_i and x_{i+1} are neighbors in the metric sense. Define $m_i = \lfloor x_i N \rfloor$, where N is chosen such that $m_i \neq m_j$ for all i and j. The map f induces then the obvious permutation $F_M : \{m_0, ..., m_{M-1}\} \to \{m_0, ..., m_{M-1}\}$ with $F(m_i) = m_j$ when $f(x_i) = x_j$. The following theorem justifies calling λ_{F_M} a discrete Lyapunov exponent.

Theorem 1 [3]: In the above setting, $\lim_{M\to\infty} \lambda_{F_M} = \lambda_f$, where λ_f is the Lyapunov exponent of f.

Example 2. For the general *tent map*,

$$f(x) = \begin{cases} \frac{x}{a} & 0 \le x \le a \\ \frac{x-1}{a-1} & a < x \le 1 \end{cases},$$

the induced family of M-permutations can be given in closed form, namely,

$$F_M(s) = \begin{cases} 1 & s = 0 \\ \mathrm{Cl}\left(\frac{M}{A}s\right) & 0 < s \le A \\ \mathrm{Fl}\left[\frac{M}{M-A}(M-s)\right] + 1 & A < s \le M-1 \end{cases}$$

where $\mathrm{Cl}(x)$ and $\mathrm{Fl}(x)$ denote, as usual, the ceiling and floor of x, $a = A/M$ and $s \in \{0, 1, ..., M-1\}$. The following table shows the values of λ_{F_M} and λ_f for the symmetric tent map ($a = 0.5$, $\lambda_f = \ln 2 = 0.693147$).

M	λ_F	$\lvert\lambda_f - \lambda_F\rvert$
256	0.690440	2.7076×10^{-3}
1024	0.692470	6.7690×10^{-4}
34012	0.693127	2.0379×10^{-5}

Example 3. For the *logistic map* $f(x) = 4x(1-x)$, $0 \le x \le 1$, we get the following table ($\lambda_f = \ln 2$).

M	λ_F	$\lvert\lambda_f - \lambda_F\rvert$
64	0.681690	0.011457
128	0.687300	0.005710
512	0.691484	0.001663

Definition: We say that an M-permutation F_M on $\{0, 1, ..., M-1\}$ has a *perfect nonlinearity* if the differences $|F(i+1) - F(i)|$, $i = 0, 1, ..., M-2$, take all possible values $1, 2, ..., M-1$.

Note that this definition is weaker than the usual one for Boolean functions on binary blocks. The following example shows the existence of maps with perfect nonlinearity.

Example 4. Let $M = 2m$ and define

$$F_{non}(s) = \begin{cases} k & \text{if } s = 2k & 0 \le k \le m-1 \\ M-1-k & \text{if } s = 2k+1 & 0 \le k \le m-1 \end{cases}$$

on $\mathbb{Z}_M$. The DLE of F_{non} is

$$\lambda_{F_{non}} = \frac{1}{M-1}\ln(M-1)!. \qquad (2)$$

2.2. The Discrete Entropy

Our definition of discrete entropy is inspired by the so-called *permutation entropy*, that is a concept introduced by Bandt and Pompe in [4] as a practical complexity measure for experimental time series. The basic idea calls for comparing the values $x_1 = f(x_0), x_2 = f(x_1), ...$ on a typical orbit of length M of a one-dimensional dynamical system and keeping track of the order patterns appearing with increasing M. Bandt *et al.* proved in [5] that permutation metric and topological entropy of piecewise monotone interval maps coincide, respectively, with Kolmogorov-Sinai (KS) and topological entropy. The result concerning permutation metric entropy and KS entropy was generalized to ergodic information sources and higher dimensional ergodic interval maps in [6].

Let F be a bijection on $\mathcal{S} = \{s_1, ..., s_M\}$ with the linear ordering $<$ (possibly induced by a metric). Note that seemingly more general situations, like bijections on product sets endowed with the product (or lexicographical) ordering, can be reduced to the case we are envisaging here. Thus, F is an M-permutation on a linearly ordered set.

For $2 \le r \le M$ and π the r-permutation $[\pi(0), \pi(1), ..., \pi(r-1)]$ (shorthand for $0 \mapsto \pi(1)$, $1 \mapsto \pi(1)$,...), we define (i)

$$P_\pi = \left\{ s \in \mathcal{S} : F^{\pi(0)}(s) < ... < F^{\pi(r-1)}(s) \right\}, \qquad (3)$$

i.e., P_π is the set of points $s \in \mathcal{S}$ such that the orbit segment $\{s, F(s), ..., F^{r-1}(s)\}$ is ordered under $<$ as π (alternatively, we say that s 'defines' the r-permutation π), and (ii)

$$p_\pi = \frac{|P_\pi|}{\sum_{\pi\in\sigma_r} |P_\pi|},$$

where $|\cdot|$ stands here for cardinality. We will refer to F also as a *substitution*, reserving the word permutation for π. We define the *discrete entropy rate* of F of rank r, $2 \le r \le M$, (based on the corresponding rate of the permutation entropy) as

$$\bar{H}_\Pi^{(r)}(F) = -\frac{1}{r-1}\sum_{\pi\in\sigma_r} p_\pi \log p_\pi. \qquad (4)$$

For entropy and entropy rate, logarithms to base 2 (and corresponding units in bits and bits per symbol, respectively) are in general preferred to other bases.

Example 5. For the right shift modulo M,

$$F_1 = [1, 2, ..., M-1, 0],$$

we get

$$\bar{H}_{\Pi}^{(r)}(F_1) = \frac{M-r+1}{M(r-1)}\log_2\frac{M}{M-r+1} + \frac{1}{M}\log_2 M$$

for $2 \leq r \leq M$. In particular, for $M = 4$ we have

$$\bar{H}_{\Pi}^{(2)}(F_1) = 0.81;\ \bar{H}_{\Pi}^{(3)}(F_1) = 0.75;\ \bar{H}_{\Pi}^{(4)}(F_1) = 0.67.$$

As for

$$F_2 = [M, 0, M+1, 1, M+2, 2, ..., 2M-1, M-1],$$

the substitution on $\{0, 1, ..., 2M-1\}$ with maximum discrete Lyapunov exponent (Example 1), we get for $M = 2$,

$$\bar{H}_{\Pi}^{(2)}(F_2) = 1;\ \ \bar{H}_{\Pi}^{(3)}(F_2) = 1;\ \ \bar{H}_{\Pi}^{(4)}(F_2) = 0.67.$$

We see that $\bar{H}_{\Pi}^{(r)}(F_1) \leq \bar{H}_{\Pi}^{(r)}(F_2)$ for $r = 2, 3, 4$. In particular, the smaller ranks $r = 2, 3$ show that F_2 is more random than F_1.

Suppose finally that F has non-zero entropy rates of ranks $r = 2, 3, ..., q$. The possibly simplest way to encapsulate in a single number the information contained in the whole hierarchy $\bar{H}_{\Pi}^{(2)}(F)$,..., $\bar{H}_{\Pi}^{(q)}(F)$ (without having to dissect F into cycles) consists of taking the arithmetic mean of it:

$$h_{\Pi}(F) = \frac{1}{q-1}\sum_{r=2}^{q}\bar{H}_{\Pi}^{(r)}(F).$$

In this way $h_{\Pi}(F)$ takes into account both high and, most importantly, low and middle ranks on an equal footing; indeed, although the number of summands in $\bar{H}_{\Pi}^{(r)}(F)$ grows as $r!$ (see (4)), the sum of the non-zero terms (before getting multiplied by $1/(r-1)$) actually scales linearly in r, rendering the different entropy rates of comparable sizes. Moreover, if we let formally $q \to \infty$, we recover the usual definition $h_{\Pi}(F) = \lim_{r\to\infty}\bar{H}_{\Pi}^{(r)}(F)$, since a convergent sequence and the arithmetic mean of their successive terms have the same limit. We call $h_{\Pi}(F)$ the (metric) *discrete entropy* (or just the entropy) of F.

Theorem 2 [7]: Let I be an n-dimensional interval and $f : I \to I$ an ergodic map. Let F_M be a substitution on M elements obtained from f, after discretizing I, in a way similar to Sect. 2.1. Then $\lim_{M\to\infty} h_{\Pi}(F_M) = h_{KS}(f)$, where $h_{KS}(f)$ is the KS entropy of f.

3. Discrete Chaos

For $\mathcal{S} = \{s_1, ..., s_M\}$ let us consider an arbitrary map $F : \mathcal{S} \to \mathcal{S}$ (not necessarily bijective). We say that the fixed point s_i (i.e., $F(s_i) = s_i$) is an *eventually fixed point* for s_j if there exists $n \geq 1$ such that $F^n(s_j) = s_i$.

Definition: We say that s_i is a *stable fixed point* for the map F if $F(s_i) = s_i$ and s_i is an eventually fixed point for at least one of its neighbor points $s_{i\pm1}$. In a similar way, one can define *stable* periodic orbits of period p using F^p instead of F.

We say that a periodic orbit is *unstable* if it is not stable. In particular, all periodic orbits (cycles) of permutations are unstable.

A set $\mathcal{L}_M = \{l_i \in \mathbb{R} : i = 1, 2, ..., M\}$ is said to be a (one-dimensional) *finite lattice* if $l_{i+1} = l_i + \Delta = l_1 + i\Delta$ for $1 \leq i \leq M-1$, where $\Delta > 0$ can be chosen (by rescaling the Euclidean metric) to be 1. For example, $\mathbb{Z}_M = \{0, 1, ..., M-1\}$ is a finite lattice. Given $\mathcal{A} = \{a_1, ..., a_m\} \subset \mathcal{L}_M$ for every $M \geq M_0$, we define $\partial\mathcal{A} = \{a_1 \pm 1, a_2 \pm 1, ..., a_m \pm 1\}$ to be the *neighboring set* of $\mathcal{A}$ in $\mathcal{L}_M$ (if $a_1 = l_1$ or $a_m = l_M$, then the neighboring points are $l_1 + 1$ and $l_M - 1$, respectively). Given now a map F_M on $\mathcal{L}_M$, $M \geq M_0$, we say that $\mathcal{A}$ is an *invariant set* of F_M if $F_M(\mathcal{A}) = \mathcal{A}$.

Definition: We say $\mathcal{A}$ is an *attractor* of F_M, if $\mathcal{A}$ is invariant under F_M and there exists $l \in \partial\mathcal{A}$ such that $F_M(l) \in \mathcal{A}$.

Let $\mathcal{A}_M \subset \mathcal{L}_M$ be an invariant set under the action of the map $F_M : \mathcal{L}_M \to \mathcal{L}_M$, $M \geq M_0$, such that F_M restricted to $\mathcal{A}_M$ is a bijection and write $G_M = F_M|_{\mathcal{A}_M}$.

Definition: (i) We say that the map F_M is *discretely chaotic* on the set $\mathcal{A}_M$ if $\lim_{M\to\infty}\lambda_{G_M} > 0$. (ii) We say that $\mathcal{A}_M$ is a *discretely chaotic attractor* for F_M if $\mathcal{A}_M$ is an attractor of F_M and $\lim_{M\to\infty}\lambda_{G_M} > 0$.

In particular, if F_M is an M-permutation, then F_M is discretely-chaotic if $\lim_{M\to\infty}\lambda_{F_M} > 0$. The permutations $F_{\max}$ and F_{non} from Examples 1 and 4 are discretely chaotic.

Observe that, in strict sense, the concepts of discretely chaotic map and attractor refer to a family of maps rather than to a single map. In most applications, F_M is certainly obtained via phase space discretization and truncation of the orbits of a continuous map f, as in the proofs of Theorems 1 and 2, and therefore it belongs to a family of maps (generated by f) by construction. Otherwise, if $\mathcal{S} = \mathbb{Z}_M$ or a translate, one can always compare λ_{F_M} to $\lambda_{F_{\max}}$ and gauge in this way the 'distance' from F_M to $F_{\max}$ —the most discretely chaotic permutation on $\mathbb{Z}_M$.

4. Applications to Cryptography

We will now focus on the cryptographic applications of discrete chaos and, more concretely, on the quality assessment and performance comparison of S-boxes.

4.1. Discrete Lyapunov exponents

As of this writing, we have analyzed the S-boxes of Rijndael cipher (the winner of the Advanced Encryption Standard –AES– competition) by means of the discrete Lyapunov exponent. The cipher is designed for 128, 192 and 256 bit block lengths but, for simplicity, we consider here the first implementation only. Rijndael applies the following transformations:

i) The *ByteSub* transformation $S(x)$ is a byte-level S-box (thus, $S : \mathbb{Z}_2^8 \to \mathbb{Z}_2^8$) defined as

$$S(x) = Bx^{-1} + b,$$

where $x^{-1} \in GL(2^8)$ is the multiplicative inverse of x if $x \neq 0$ or 0 if $x = 0$, B is an 8×8 binary matrix A obtained by successively rotating the bits of its first row $B_{1j} = (1,0,0,0,1,1,1,1)$ to the right, and $b = (1,1,0,0,0,1,1,0)^{transpose}$. The ByteSub transformation defines a permutation F on $\{0,...,255\}$ with $\lambda_F = 4.01$, while $\lambda_{\max} = 4.86$ (see (1)) and $\lambda_{non} = 4.55$ (see (2)). The role of the ByteSub transformation is to mix in a strong nonlinear way the input information.

ii) Let $b_{0,0},...,b_{0,3},...,b_{3,0},...,b_{3,3}$ be the 16 bytes (128 bits) of the input block. The *ShiftRow transformation* takes the words

$$\begin{aligned} w_0 &= (b_{0,0}, b_{0,1}, b_{0,2}, b_{0,3}) \\ w_1 &= (b_{1,0}, b_{1,1}, b_{1,2}, b_{1,3}) \\ w_2 &= (b_{2,0}, b_{2,1}, b_{2,2}, b_{2,3}) \\ w_3 &= (b_{3,0}, b_{3,1}, b_{3,2}, b_{3,3}) \end{aligned} \tag{5}$$

and returns $w_i >>> C_i$, $i = 0,1,2,3$, where $w >>> C$ is the rotation of the sequence w of bytes to the right by C bytes. The values of C_i are $C_i = i$, $i = 0,1,2,3$. The role of the ShiftRow permutation is just to permute all 16 bytes of the input block, thus it is a permutation on $\{0,1,...,15\}$. Its DLE turns out to be 0.93, which is substantially smaller than the maximum one (for $M = 16$) 2.13.

iii) Given an input block in the form (5), the *MixColumn transformation* can be viewed as a linear transformation in $GF(2^8)^4$. In fact, if $c_j = (b_{0,j}, b_{1,j} b_{2,j}, b_{3,j})$, $0 \leq j \leq 3$, is the jth column of (5), then MixColumn is

$$c_j \mapsto \begin{pmatrix} 02 & 03 & 01 & 01 \\ 01 & 02 & 03 & 01 \\ 01 & 01 & 02 & 03 \\ 03 & 01 & 01 & 02 \end{pmatrix} c_j,$$

where the matrix entries are pair of hexadecimal numbers representing bytes in the usual way. Therefore, MixColumn induces a permutation on $\{0,1,...,2^{32}-1\}$. We have found the Lyapunov exponent of MixColumn to be 21.49 ($\lambda_{\max} \approx 22.18$).

In addition to the analysis of the single transformations, the behavior of their composition (i.e., of the Rijndael cipher) has been evaluated. To this aim, we assign to each 128 bit block an integer in $\{0,1,...,2^{128}-1\}$ via its binary representation. The computation of the DLE has been performed on 7000 iterations of the Rijndael map obtaining 87.04, to be compared to $\lambda_{\max} = 88.72$ (see (1) with $M = 2^{128}$).

4.2. Discrete entropy

As for the applications of discrete entropy, let us illustrate them with an example taken from [8]. The 4×4 S-boxes

$$S_1 = [15,12,2,1,9,7,10,4,6,8,5,11,0,3,13,14]$$

$$S_2 = [8,2,4,13,7,14,11,1,9,15,6,3,5,0,10,12]$$

(the 4-bit number $b_1b_2b_3b_4$ being identified, as usual, with the decimal number $b_12^3 + b_22^2 + b_32^1 + b_4$) are 0/1 balanced, nonlinear and fulfill the maximum entropy criterion. But from the discrete entropy point of view, they are quite different. S_1 consists of two cycles of length 7 and two fixed points. Its discrete entropies are:

$$\begin{aligned} &\bar{H}_\Pi^{(2)}(S_1) = 0.99; \quad \bar{H}_\Pi^{(3)}(S_1) = 1.04; \quad \bar{H}_\Pi^{(4)}(S_1) = 0.96; \\ &\bar{H}_\Pi^{(5)}(S_1) = 0.84; \quad \bar{H}_\Pi^{(6)}(S_1) = 0.70; \quad \bar{H}_\Pi^{(7)}(S_1) = 0.58; \end{aligned}$$

and $h_\Pi(S_1) = 0.85$. S_2 consists of two cycles of lengths 12 and 4, with

$$\begin{aligned} &\bar{H}_\Pi^{(2)}(S_2) = 0.99; \; \bar{H}_\Pi^{(3)}(S_2) = 1.08; \; \bar{H}_\Pi^{(4)}(S_2) = 1.17; \\ &\bar{H}_\Pi^{(r)}(S_2) = 3.59/(r-1) \text{ for } r = 5,...,12 \end{aligned}$$

thus $h_\Pi(S_2) = 0.68$. As expected, the discrete entropy of S_1 is higher and, consequently, it generates more pseudorandomness that S_2.

Conclusion

The basic conceptual framework of discrete chaos has been presented and potential applications to cryptography have been illustrated with the analysis of some S-boxes — mainly, those of AES.

References

[1] C.E. Shannon, “Communication theory of secrecy systems,” *Bell Syst. Techn. J.* **28**, pp. 656-715, 1949.

[2] J. Fridrich, “Symmetric ciphers based on two dimensional maps,” *Int. J. Bif. Chaos* **8**, pp. 1259-1284, 1998.

[3] L. Kocarev, J. Szczepanski, J.M. Amigó, I. Tomovski and P. Amato, “Discrete Chaos – Part I: Theory” (submitted).

[4] C. Bandt and B. Pompe, “Permutation entropy: A natural complexity measure for time series,” *Phys. Rev. Lett.* **88**, p. 174102, 2002.

[5] C. Bandt, G. Keller and B. Pompe, “Entropy of interval maps via permutations,” *Nonlinearity* **15**, pp. 1595-602, 2002.

[6] J.M. Amigó, M.B. Kennel and L. Kocarev, “The permutation entropy rate equals the metric entropy rate for ergodic information sources and ergodic dynamical system,” Physica D (in press).

[7] J.M. Amigó, M.B. Kennel and L. Kocarev, “Discrete entropy” (submitted).

[8] M. Adámyová, “A construction of S-boxes based on Boolean fuctions with maximum entropy,” *Proceedings of ELITECH* (Bratislava), 1998.