Chaotic block ciphers: From theory to practical algorithms

Article *in* Circuits and Systems I: Regular Papers, IEEE Transactions on · July 2006

DOI: 10.1109/TCSI.2006.874182 · Source: IEEE Xplore

Naoki Masuda
University at Buffalo, The State University of New York
236 PUBLICATIONS **4,580** CITATIONS

Ljupco Kocarev
Macedonian Academy of Sciences and Arts
377 PUBLICATIONS **12,283** CITATIONS

Chaotic Block Ciphers: From Theory to Practical Algorithms

Naoki Masuda, Goce Jakimoski, Kazuyuki Aihara, and Ljupco Kocarev, *Fellow, IEEE*

***Abstract*—Digital chaotic ciphers have been investigated for more than a decade. However, their overall performance in terms of the tradeoff between security and speed, as well as the connection between chaos and cryptography, has not been sufficiently addressed. We propose a chaotic Feistel cipher and a chaotic uniform cipher. Our plan is to examine crypto components from both dynamical-system and cryptographical points of view, thus to explore connection between these two fields. In the due course, we also apply dynamical system theory to create cryptographically secure transformations and evaluate cryptographical security measures.**

***Index Terms*—Block cipher, chaos, cryptography, differential probability, linear probability.**

I. Introduction

OVER THE PAST decade, there has been tremendous interest in behavior of chaotic systems. They are characterized by sensitive dependence on initial conditions, similarity to random behavior, and continuous broadband power spectrum. Chaos has potential applications in several functional blocks of a digital communication system: compression, encryption, and modulation. The pioneering work on chaos synchronization [1] led to several applications in communication, in which chaotic systems with continuous-value signals were used to transmit information. Several schemes have been developed which allow to transform the information signal into a chaotic waveform on the transmitter side and to extract the information signal from the transmitted waveform on the receiver side. The most important among them are: chaotic masking, chaos shift keying, and chaotic modulation. In early days (from 1990 to 1995) the main research goal was to develop schemes in which a single chaotic system is used for both modulation and encryption. This approach eventually evolved into two distinct research areas: chaos-based modulation [2], [3] and chaos-based cryptography [4], [5].

Cryptography is generally acknowledged as the best method of data protection against passive and active fraud. An overview of recent developments in the design of conventional cryptographic algorithms is given in [6], [7]. Three most common cryptographic objects are: block-encryption algorithms (private-key algorithms), pseudorandom number generators (additive stream ciphers), and public-key algorithms. Block ciphers transform a relatively short string (typically 128 or 256 bits) to a string of the same length under control of a secret key.

Previous Work—There are several references to the problem of designing block encryption algorithms using chaotic maps. A block cipher uses a complicated but sufficiently efficient one-to-one transformation on a finite space. For enhanced security, the transformation should be also sensitive to changes in the parameter values that constitute the secret keys. The idea of chaotic block ciphers is centered around the use of modified chaotic maps in this purpose [4], [8]–[15]. To derive a one-to-one cryptographical mapping from a generally many-to-one chaotic map, we must somehow avoid eigenmodes of the mapping with negative expansion exponents. Otherwise, the created cipher is vulnerable to simple cryptanalysis [14]–[18].

Despite these efforts, commercially and theoretically acknowledged ciphers are not associated with chaos or dynamical systems [4], [7]. Popular ciphers are defined with operations on binary spaces or algebraic transformations. On digital computers, they naturally surpass chaotic ciphers, which are derived from continuous maps. Put in another way, most chaotic ciphers have neglected the tradeoff between the operation speed, required memory space, and security. If naively designed, chaotic ciphers are slow on digital computers due to costly nonlinear and chaotic computations [13], [15]. Particularly, multiplications and divisions with floating point arithmetic or many-bit integer arithmetic hamper fast operation [4].

A second point is that, despite attempts to link cryptographical security measures and dynamical characteristics of chaos [14], [15], [19], big gaps still remain between them. Analysis of chaotic ciphers based on their algebraic properties [10] in turn dismisses linkage between the two fields.

Our work—In this paper, we propose a definition of chaotic block cipher. Block ciphers are finite-state maps. We say that a finite-state map, for which the trajectories are always eventually periodic, is chaotic if it approximates a chaotic map (we will give a precise mathematical definition later). We consider two classes of chaotic finite-state maps: key-dependent chaotic S-boxes and chaotic mixing transformations. We suggest two

Manuscript received April 20, 2005; revised August 15, 2005 and December 14, 2005. This research was funded in part by Special Postdoctoral Researchers Program of RIKEN and the National Science Foundation under Grant CCR-0209092. This paper was recommended by Associate Editor Y. Nishio.

N. Masuda is with the Laboratory for Mathematical Neuroscience, RIKEN Brain Science Institute, Wako 351–0198, Japan, and also with the ERATO Aihara Complexity Modeling Project, Japan Science and Technology Agency, Tokyo 151-0064, Japan (e-mail: masuda@brain.riken.jp).

G. Jakimoski is with the Computer Science Department, Florida State University, Tallahassee, FL 32306-4530 USA (e-mail: jakimosk@cs.fsu.edu).

K. Aihara is with the ERATO Aihara Complexity Modeling Project, Japan Science and Technology Agency, Tokyo 151-0064, Japan, and also with the Institute of Industrial Science, University of Tokyo, Tokyo 153-8505, Japan (e-mail: aihara@sat.t.u-tokyo.ac.jp).

L. Kocarev is with the Institute for Nonlinear Science, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0402, USA and also with the Graduate School of Electrical Engineering, University “Kiril i Metodij,” Skopje 1000, Macedonia (e-mail: lkocarev@ucsd.edu).

Digital Object Identifier 10.1109/TCSI.2006.874182

chaotic block ciphers: uniform and Feistel, and exactly estimate bounds for the differential probability and the linear probability to make our ciphers resistant to differential and linear cryptanalysis. In the consequence, we explore potential ties between bit-oriented security analysis and chaoticity. We also evaluate cryptographical quantities with the use of both dynamical-system and cryptographical criteria. The security evaluation of block ciphers consists of three steps [31]. First, one should prove the resistance to differential and linear attacks; second, one should check for the extensions and generalizations of differential and linear attacks; and third, one should take into account several dedicated attacks applicable to cipher with a small number of rounds. However, one should keep in mind that provable security against one or two important attacks *does not imply* that the cipher is secure: other attacks may exist. On the other hand, provable security against certain attacks is certainly a first step in the right direction. In this work we prove that proposed chaotic ciphers are resistant to differential and linear attacks.

In Section II, we briefly review preliminary concepts in cryptography and our previous block cipher, and we specify design principles for the present cipher. In Section III, we define and analyze local transformations called S-boxes, which are derived from the discretized skew tent map. Cryptographical properties of chaos-based and cryptography-based global transformations are examined in Sections IV and V, respectively. The local and global transformations set up in Sections III, IV, and V are combined to define chaotic Feistel ciphers and chaotic uniform ciphers in Section VI. Connection between chaos and cryptography is further discussed in Sections VII and VIII.

II. Preliminaries

A. Product Ciphers

Two general principles guiding the design of block ciphers are *diffusion* and *confusion*. Diffusion means spreading out of the influence of a single plaintext bit over many ciphertext bits so as to hide the statistical structure of the plaintext. An extension of this idea is to spread the influence of a single key bit over many bits of ciphertext. Confusion means use of transformations that complicate dependence of the statistics of ciphertext on the statistics of plaintext. Most ciphers achieve the diffusion and the confusion by means of round repetition. Repeating a single round contributes to cipher's simplicity and ease of implementation.

A very common approach to creating diffusion and confusion is to use a product cipher, i.e., a cipher that is implemented as a composition of simple ciphers. Modern block ciphers consist of four transformations (also called layers): 1) substitution transformation (also called S-box); 2) permutation transformation; 3) linear mixing transformation; and 4) key-adding transformation.

1) Global Structure: Block ciphers in which every input bit is treated in a similar way belong to the class of *uniform networks*, also called *substitution—permutation networks*. An example of such a network is given in Fig. 1. Examples of block ciphers in this class are SAFER, SHARK, Rijndael, and 3-WAY [6], [7], [20]–[23]. An advantage of this approach is inherent parallelism, while a disadvantage is that inverse algorithm, which is required for decryption, may be different from the encryption algorithm.

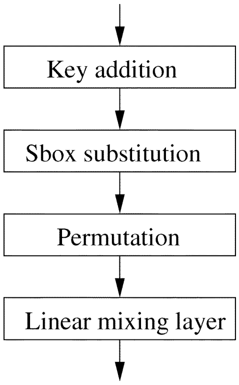


Fig. 1. Uniform structure.

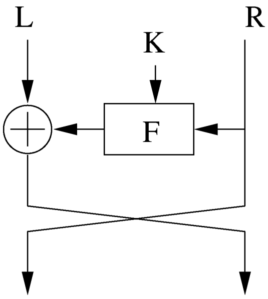


Fig. 2. Feistel structure.

Block ciphers in which the input is divided into two halves, a nonlinear transformation is applied to the right half, the result is added into the left half, and subsequently the left and right halves are swapped, constitute the class of *Feistel networks*. An example of such a network is given in Fig. 2. Examples of block ciphers in this class are data encryption standard (DES), Khufu, CAST, and Twofish [6], [7], [24]. Since the nonlinear part requires most of the computation, two rounds of a Feistel cipher require about the same effort as one round of a uniform cipher. The output of one nonlinear function is input directly to the next one, which decreases the amount of parallelism but increases the propagation of local changes. An advantage of this approach is that the same round function can be used for both encryption and decryption, while a round function itself need not be invertible. Since DES is a Feistel network, more cryptanalytic experience is available on Feistel ciphers than on any other general structure.

Other variants and extensions of uniform and Feistel networks have been proposed [6], [7]. For example, Snefru is an extension of Feistel cipher in which the input is divided into more parts. Such ciphers are called general unbalanced Feistel networks. RC5 is almost a Feistel cipher. MISTY2 is a variant of Feistel cipher that increases parallelism so that the calculation of two consecutive rounds can be carried out at the same time. IDEA [30] is another variant of Feistel cipher.

2) Nonlinear Transformations (S-Boxes): A nonlinear transformation is essential for every strong encryption algorithm. Nonlinear transformations are often implemented as lookup tables or S-boxes. An S-box with p input bits and q output bits is denoted by $p \rightarrow q$. The DES uses eight different 6 → 4 S-boxes. Byte level S-boxes (8 → 8) are suited for software implementation on 8-bit processors. Examples of block ciphers with 8 → 8 S-boxes include SAFER, SHARK, and Rijndael. For modern processors with 32-bit or 64-bit words, S-boxes with more output bits provide higher efficiency. Thus, for example, Snefru, Blowfish, CAST, and SQUARE use 8 → 32 S-boxes. The S-boxes can be selected at random as in Snefru, can be computed using a chaotic map (as suggested in this paper), or can contain some mathematical structure such as an exponential over a finite field. Examples of the last approach include SAFER, SHARK, and Rijndael.

Nonlinear transformations can be also realized using the processor instructions. FEAL uses addition and rotation, IDEA employs addition and multiplication, RC5 uses addition, exclusive OR, and data-dependent rotations, while 3-WAY uses a simple Boolean function, which can be also considered as a 3 → 3 S-box. Cryptanalytic results indicate that resistance against linear or differential cryptanalysis is not a property of S-boxes or of global structure, but rather of the combination of both. However, for a similar global structure, increasing the size of the S-boxes usually enhances the resistance.

3) Linear Mixing Transformations: In order to decrease the complexity of implementation and increase the speed of computations, usually nonlinear transformations are applied only to small parts of the block data and linear transformations are used to spread local changes. The simplest linear functions are a bit permutation, used, for example, in DES, and a rotation, used in Khufu. More general linear transformations are pseudo-Hadamard transformation, used in SAFER, and MDS transformation: the diffusion operation based on maximum distance separable (MDS) linear codes, used in Twofish and Rijndael.

4) Wide Trail Strategy: The well-known differential cryptanalysis exploits the difference propagation from the round input to the round output. An m-round *differential trail* consists of a chaining of difference propagations. A chaining means that difference in a round input somehow propagates to difference in an output, which is regarded as the difference in the input to the next round, and so on. An input subblock or an S-box in a differential trail is said to be *active* if its input difference is nonzero. Similarly, the linear cryptanalysis takes advantage of the correlation between the input and output bits of a round. An m-round *linear trail* consists of a chaining of m round transformation correlations. These are called the (linear) steps of the trail. A subblock is active with respect to a linear trail if its output selection vector is nonzero for that linear trail. We define that the *weight* of a given trail is the number of active S-boxes. The weight of a given vector of bytes is the number of active bytes.

There are two ways to fight against differential and linear cryptanalyzes. One is to build S-boxes with low differential and linear probabilities. These quantities are actually evaluated for our cryptosystem in Section III-B. The other, which the wide trail strategy [20] is concerned to, is to design the round transformation so that only trails with many active S-boxes occur. The round transformation must be designed in such a way that differential (respectively, linear) steps with few active S-boxes are followed by differential (respectively, linear) steps with many active S-boxes. The sum of the number of active input S-boxes and that of active output S-boxes, minimized over the input space, is called *branch number*. A secure round transformation should have a large branch number.

B. Chaotic Block Ciphers

Let $f : S \rightarrow S$ be an N-dimensional chaotic map. For simplicity, we assume that S is either an N-dimensional cube $[0,1]^N$ or an N-dimensional torus. Let $f_M : \{m_0, m_1, \ldots m_{M-1}\}^N \rightarrow \{m_0, m_1, \ldots m_{M-1}\}^N$ be a map induced by f when S is discretized (quantized) with $\{m_0, m_1, \ldots m_{M-1}\}^N$. We assume that as the discretization becomes finer, or as M goes to infinity, f_M approaches f; in this case, we say f_M approximates f. For example, $S = [0,1]$ can be uniformly quantized as $\{0/M, 1/M, \ldots (M-1)/M\}$. Clearly, the map f_M induces a map $F_M : \{0, 1, \ldots M-1\}^N \rightarrow \{0, 1, \ldots M-1\}^N$ in a natural way.

Definition: We say that a block cipher is chaotic if its S-box and diffusion transformations are, respectively, bijections F_M and G_M induced by f_M and g_M, which are approximations of chaotic maps f and g.

1) Example: Finite-State Tent Map: Previously, we proposed a chaotic block cipher with space-discretized skew tent maps [14], [15]. We also analyzed security from dynamical system viewpoints, including the sensitive dependence on plaintexts and on keys, and the exponential decay of correlation between plaintexts and ciphertexts. As we discuss in Sections III and IV, these security properties in terms of the Euclidean distance may be related to cryptographical security associated with the Hamming distance. Let us briefly review the finite-state tent map.

For a positive integer $M \geq 2$, the rescaled skew tent map $F_A : [0, M] \rightarrow [0, M]$, $0 < A < M$ is defined by

$$F_A = \begin{cases} X/A, & 0 \leq X \leq A \\ (M-X)/(M-A), & A < X \leq M. \end{cases}$$

The map F_A is one-dimensional (1-D) $(N = 1)$, exact, and therefore mixing and ergodic. The Lyapunov exponent λ is given by

$$\lambda = -\frac{A}{M}\log\frac{A}{M} - \frac{M-A}{M}\log\frac{M-A}{M}. \tag{1}$$

To convert the two-to-one map F_A to a one-to-one map for cryptographical purpose, we restrict the domain and the range to an identical finite plaintext space defined by

$$P' = \{X; X = 1, 2, \ldots, M\}. \tag{2}$$

The modified skew tent map on P' is represented by

$$\tilde{F}_A(X) \equiv \begin{cases} \left\lceil \frac{M}{A}X \right\rceil, & 1 \leq X \leq A \\ \left\lfloor \frac{M}{M-A}(M-X) \right\rfloor + 1, & A < X \leq M. \end{cases} \tag{3}$$

The parameter A takes values from the following set:

$$K' = \{A; A = 1, 2, \ldots, M\}.$$

The inverse of $\tilde{F}_A$ is calculated as

$$\tilde{F}_A^{-1}(Y) = \begin{cases} X_1, & m(Y) = Y, \quad \frac{X_1}{A} > \frac{M-X_2}{M-A} \\ X_2, & m(Y) = Y, \quad \frac{X_1}{A} \leq \frac{M-X_2}{M-A} \\ X_1, & m(Y) = Y+1 \end{cases} \tag{4}$$

where

$$X_1 = \lfloor M^{-1}AY \rfloor \quad X_2 = \lceil (M^{-1}A - 1)Y + M \rceil \tag{5}$$

$$m(Y) = Y + \left\lfloor \frac{AY}{M} \right\rfloor - \left\lceil \frac{AY}{M} \right\rceil + 1. \tag{6}$$

The encryption function and the decryption function are $\tilde{F}_A^n$ and $\tilde{F}_A^{-n}$, respectively, where n is the iteration number. When $M/3 \leq A \leq 2M/3$ and $n \cong 2.39 \log_2 M$, the sensitive dependence on plaintexts and on keys is ensured. If A is coprime to M, we can reduce n to $n \cong 1.66 \log_2 M$. Increases in n and M enhance security at the expense of operation speed and consumed memory.

2) Design Principle: As reviewed in Section I, most current chaotic ciphers are less powerful than broadly used ones. To fill this gap in and put forward the utility of chaotic ciphers, we lay following design principles.

- *General construction scheme:* We combine four building blocks explained in Section II-A-1.
- *Links:* We explore links between chaos and cryptography in the course of designing and analyzing chaotic ciphers. This includes creating cryptographically secure transformations based on chaos and applying dynamical system theory to cryptographical security analysis. We neglect advanced cryptographical techniques such as initial permutations [7, p. 271] and software/hardware optimizations, which obscure the main points of this work.
- *Small code:* We use (3) and (4) as chaotic S-boxes for local nonlinear transformation and avoid devoting the memory space to storing S-boxes as lookup tables. Another merit of the chaotic S-boxes is security analysis guided by dynamical system theory [15], [19].
- *Key-dependent S-boxes:* Generally speaking, S-boxes whose form essentially depends on key values are slower but more secure than key-independent ones [7, p. 298]. Use of key-independent chaotic S-boxes was pioneered in [10], in which the S-box is constructed with a transformation given by $\overline{F}((X+K) \bmod M)$, where K is the key. Here we propose key-dependent S-boxes since $\tilde{F}_A$ is naturally defined so.

Under these constraints, we design a chaotic Feistel cipher and a chaotic uniform cipher. One block is defined to be 128 bits (16 bytes) with 8 → 8 S-boxes.

III. Chaotic Nonlinear Transformations

A. Construction

Our key-dependent S-box is a one-to-one transformation on

$$P_S \equiv \{X; X = 0, 1, \ldots, M-1\}, \qquad M = 256. \tag{7}$$

To comply with binary implementations, we include 0 and exclude $M = 256$ from the domain, which makes P_S slightly but not essentially different from P' in (2). We define the S-box by

$$S_A(X) = \tilde{F}_A^n(X+1) - 1, \qquad A \in K_S. \tag{8}$$

We specify K_S and the values of n in Section III-B. The inverse needed for uniform ciphers is given by

$$S_A^{-1}(X) = \tilde{F}_A^{-n}(X+1) - 1, \qquad A \in K_S. \tag{9}$$

B. Security Analysis and Parameter Setup

Numerical analysis of a whole cipher is formidable since the plaintext space, the ciphertext space, and the key space are huge. In this section we perform security analysis of the S-box. The results reviewed in Section II-B-1 tell us how to determine n and K_S from dynamical theory perspective. From a cryptographical viewpoint, however, they must be chosen so that the generated S-boxes are strong against differential cryptanalysis and linear cryptanalysis [7], [10]. Resistance to these cryptanalyses is evaluated, respectively, by the differential probability (DP) and the linear probability (LP). We define

$$\mathrm{DP} = \max_{\Delta X \neq 0, \Delta Y} \frac{|\{X \in P' | S_A(X) \oplus S_A(X \oplus \Delta X) = \Delta Y\}|}{256} \tag{10}$$

and

$$\mathrm{LP} = \max_{a \neq 0, b \neq 0} \left(2 \cdot \frac{|\{X \in P' | X \cdot a = S_A(X) \cdot b\}|}{256} - 1 \right)^2 \tag{11}$$

where $\oplus$ denotes XOR, and $x_1 \cdot x_2$ denotes the parity of the bitwise product of x_1 and x_2.

The discretized map $\tilde{F}_A$ does not necessarily preserve mathematical properties of the tent map. Therefore, analytical evaluation of (10) or (11) with the use of dynamical system theory seems difficult. On the other hand, this fact may mean that the chaotic S-boxes merit from the absence of algebraic structure that could benefit attackers. Nevertheless, we can partially evaluate DP if we suppose, as in most chaotic ciphers [8], [14], [15], that the maximum in (10) is realized when $\Delta X = 1$. Using (3), we obtain

$$\tilde{F}_A(X+1) - \tilde{F}_A(X) = \begin{cases} 1, & \text{for } 2A - M - 1 \\ & \quad \text{points in } \{1, 2, \ldots, A-1\} \\ 2, & \text{for } M - A \\ & \quad \text{points in } \{1, 2, \ldots, A-1\} \\ k, & \text{for } kM - (k+1)A + 2 \\ & \quad \text{points in } \{A, A+1, \ldots, M\} \\ k+1 & \text{for } kA - (k-1)M - 2 \\ & \quad \text{points in } \{A, A+1, \ldots, M\} \end{cases} \tag{12}$$

where $M/2 \leq A \leq M$, and the condition

$$(k-1)M/k \leq A < kM/(k+1) \tag{13}$$

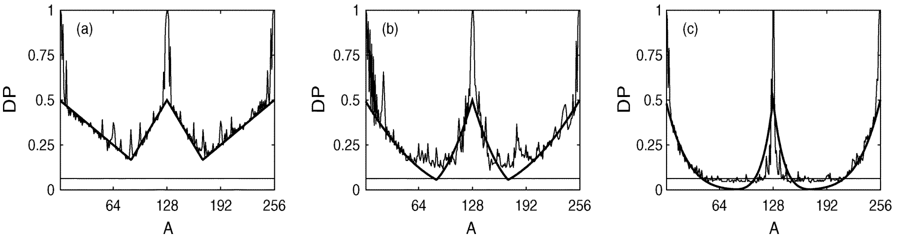


Fig. 3. Numerically obtained DP (thin lines) for $\tilde{F}_A^n$ and its theoretical estimation DP_d (thick lines). The iteration numbers are (a) $n = 1$, (b) $n = 2$, and (c) $n = 5$. The reference level $\mathrm{DP} = 2^{-4}$ is also shown.

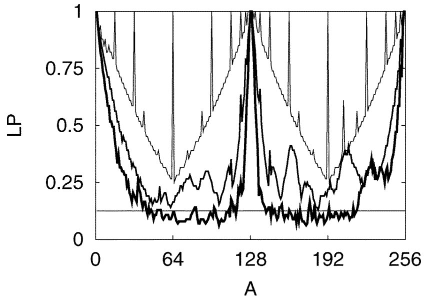


Fig. 4. Numerically obtained LP for $\tilde{F}_A^n$. The iteration numbers are $n = 1$ (thinnest line), 2, and 5 (thickest line). The reference level $\mathrm{LP} = 2^{-3}$ is also shown.

uniquely determines a $k \geq 2$ [25]. The expression for $0 \leq A \leq M/2$ is a symmetric copy of (12) with respect to $A = M/2$. Let us assume that $\tilde{F}_A(X+1) - \tilde{F}_A(X)$ is randomly distributed with the density naturally derived from (12). Then, a rough lower bound DP_d of DP is estimated as

$$\mathrm{DP}_d = \begin{cases} \frac{(1-2A/M)^n}{2}, & 0 < A \leq \frac{M}{3} \\ \frac{(4A/M-1)^n}{2}, & \frac{M}{3} < A \leq \frac{M}{2} \\ \frac{(3-4A/M)^n}{2}, & \frac{M}{2} < A \leq \frac{2M}{3} \\ \frac{(2A/M-1)^n}{2}, & \frac{2M}{3} < A \leq M. \end{cases} \tag{14}$$

Fig. 3 shows that, for a wide range of A, DP_d resembles numerically obtained DP. Fig. 3 and (14) indicate that we should avoid A close to 1, 128, or 256, for which $\Delta X = 1$ does not achieve the maximum in (10). Actually, $\tilde{F}_A$ with $A \cong 1$ or 256 is nearly the identity map, whereas $\tilde{F}_A$ with $A \cong 128$ approximates the orthodox tent map, that is, the one bit shift. With these values of A excluded from K_S, $n = 5$ guarantees $\mathrm{DP} \leq 2^{-4}$, which is required for sufficient security [10].

We note that ignorance of discretization effects leads to a wrong estimate. For the continuous tent map, we would have

$$\mathrm{DP}'_d = (\max\{A/M, 1 - A/M\})^n \tag{15}$$

which deviates from (14) even if M is sufficiently large. Here, discretization drastically affects the security of chaotic ciphers, although they are often neglected in both chaotic ciphers [9], [13] and numerical simulations of chaotic orbits [26].

Numerically obtained LP is shown in Fig. 4. Again, $A \cong$ 1, 128, or 256 is weak due to the strong linearity of $\tilde{F}_A$. The security level $\mathrm{LP} \leq 2^{-3}$ [10] is achieved with the same values of n and A as those for DP.

Accordingly, we set $n = 5$ and choose A satisfying $\mathrm{DP} \leq 2^{-4}$ and $\mathrm{LP} \leq 2^{-3}$. In addition, A is assumed to be odd since the security is enhanced when A and $M = 256$ are coprime [15]. We define K_S with $|K_S| = 64$ by

$$K_S = \{51, 53, 55, \ldots, 117, 139, 141, 143, \ldots, 201\}. \tag{16}$$

With A in K_S defined in (16), MX/A is not an integer provided $1 \leq X < A$. Then we can rewrite (3) without ceiling functions; floor functions are more tractable for digital computers. As a result, $\tilde{F}_A$ is represented by

$$\tilde{F}_A(X) = \begin{cases} \left\lfloor \frac{256}{A} X \right\rfloor + 1, & 1 \leq X < A) \\ 256, & X = A \\ \left\lfloor \frac{256}{256-A}(256 - X) \right\rfloor + 1, & A < X \leq 256. \end{cases} \tag{17}$$

The values $256/A$ and $256/(256 - A)$ $(A \in K_S)$ up to 8 bits should be stored, which consumes 128 bytes of the memory. This is to replace computationally precious integer divisions by integer multiplications. Equation (4) is also simplified because $m(Y)$ in (6) is equal to Y unless $Y = 256$. Accordingly, we have

$$\tilde{F}_A^{-1}(Y) = \begin{cases} X_1, & X_1(256 - A) > A(256 - X_2) \\ X_2, & X_1(256 - A) \leq A(256 - X_2) \end{cases} \tag{18}$$

where

$$X_1 = \lfloor M^{-1} A Y \rfloor \quad X_2 = M - \lfloor (1 - M^{-1} A) Y \rfloor . \tag{19}$$

IV. Mixing Transformations (Diffusion Layers)

In this section we suggest several chaos-based one-to-one mixing transformations. Then, we evaluate their mixing properties in two ways. One is in terms of the Hamming distance and concerned to cryptographical security. The other is in terms of the Euclidean distance, which is linked to linear dissipation of dynamical systems with positive Lyapunov exponents. Thus, we will be able to argue the transition from chaos to cryptosystems.

A. 1-D Map

We consider permutations G of the set $\{0, 1, \ldots, M-1\}$. Let us denote the Euclidean distance between two integers i and j by $d(i, j) \equiv |i - j|$. Then, $|G(i+1) - G(i)|$ measures the divergence of two trajectories evolving in one iteration, starting from two "slightly" different initial conditions: an initial point i and its neighbor $i + 1$. In cryptography, one looks for permutations for which $|G(i + 1) - G(i)|_H > a$ for some a and all $i = 0, \ldots m - 2$, where $|\cdot|_H$ is the Hamming distance. Maximizing this distance is closely related to finding an MDS matrix. Here, let us focus on the Euclidean distance.

Let $M = 2m$ and let m be an even number (we usually take $M = 256$ and $m = 128$). Consider the map $G : \{0, 1 \ldots, M - 1\} \to \{0, 1 \ldots, M - 1\}$ defined as

$$G(x) = \begin{cases} x + m(\mathrm{mod} M), & \text{for } x: \text{ even} \\ x, & \text{for } x: \text{ odd.} \end{cases}$$

It is easy to see that $|G(i+1) - G(i)| \geq m - 1$. In terms of the Euclidean distance, this G is the best 1-D mixing transformation. However, $|G(i+2) - G(i)|_H = 1$ while $|(i+2) - i|_H = 1$. The map G is linear and 1-D, which is a partial reason why G does not provide mixing in terms of the Hamming distance. Transition from the Euclidean view to the Hamming view will be seen in the following examples.

B. Two-Dimensional Cat Map

Let us consider a family of two-dimensional (2-D) cat maps [12], [26]–[28], or the Anosov diffeomorphism, defined as

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} g+1 & g \\ 1 & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \mod 256 \tag{20}$$

where $X_1, X_2, Y_1, Y_2 \in P_S$. For dynamical systems in general, information dissipation is measured by quantities such as the Lyapunov exponents and the entropy, and by presence or absence of ergodicity, mixing, and exactness. The eigenvalues of the cat map represented by (20) are

$$\lambda = g + 2 \pm \sqrt{(g+2)^2 + 1} \tag{21}$$

and information dissipation is guaranteed because the largest eigenvalue is greater than one.

The cat map serves as a diffusion layer because its inverse is well-defined on the integer space on which cryptographical transformations are based. A special case $g = 1$ is known as the pseudo-Hadamard transform (PHT). The PHT is used in various cryptosystems because it requires only two additions in a digital processor [10], [21], [22], [24]. Next, let $g = 128$. Then, the 2-D Euclidean distance between, for example, (7,9) and all its neighbors, or (7,8), (7,10), (6,9), and (8,9), becomes more than $g-1$ after applying (20). This holds for any point because the cat map is linear. Among all cat maps, (20) with $g = 128$ gives best mixing in terms of the Euclidean distance. Cryptographically, this map obviously has branch number 2 (see Section II-A-4 for branch number), meaning that an input with one active byte results in an output with at least one active byte. Actually, this is true for any 2-D mixing transformations, which are one-to-one by definition. The next example accommodates more nontrivial branch-number arguments of chaos-driven maps.

C. Four-Dimensional Torus Map

For stronger ties between chaos and cryptography, we examine a family of four-dimensional (4-D) torus maps: $G = (g_{ij})$, $0 \leq g_{ij} \leq 255$ $(1 \leq i, j \leq 4)$. Such multi-dimensional torus maps appear in the context of, for example, quantum chaos [29]. We write the input–output relation as follows:

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix} = G \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{pmatrix} \mod 256 \tag{22}$$

where $X_i, Y_i \in P_S$ $(1 \leq i \leq 4)$. The following lemma is a first relation between G based on dynamical systems and cryptographical security.

Lemma 1: The mixing layer G has branch number at most 4.

Proof: Let us suppose that G has branch number 5, which is the possible maximum. Then, each input with one active byte yields an output with all bytes active. The inputs

$$\begin{pmatrix} 128 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 128 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 128 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 128 \end{pmatrix} \tag{23}$$

respectively yield

$$\begin{pmatrix} 128g_{11} \\ 128g_{21} \\ 128g_{31} \\ 128g_{41} \end{pmatrix}, \begin{pmatrix} 128g_{12} \\ 128g_{22} \\ 128g_{32} \\ 128g_{42} \end{pmatrix}, \begin{pmatrix} 128g_{13} \\ 128g_{23} \\ 128g_{33} \\ 128g_{43} \end{pmatrix}, \begin{pmatrix} 128g_{14} \\ 128g_{24} \\ 128g_{34} \\ 128g_{44} \end{pmatrix} \tag{24}$$

mod 256. All the elements of each vector in (24) are nonzero mod 256. Consequently, g_{ij} $(1 \leq i, j \leq 4)$ is odd. Then, the determinant of G is even, and G does not have the inverse whose elements are integers. This contradicts the assumption that G is one-to-one. ■

Based on Lemma 1, it is impossible to obtain 4-D maps with branch number 5 unless we resort to some discrete-math machineries used in other ciphers [20], [24]. However, the following theorem facilitates semi-optimal mixing transformations originating from linear chaotic maps. The obtained branch number 4 is larger, and hence more resistant to differential cryptanalysis, than ones obtained by the Hadamard shuffle (=2) and the Armenian shuffle (=3) used in the SAFER family [21], [22].

Theorem 1: The mixing transformation G has branch number 4 if and only if the following conditions are satisfied.

(a) Each column of G has exactly three odd elements.

(b) Each row of G has exactly three odd elements.

(c) The determinant of G is 1 or -1.

Proof: Suppose that G has branch number 4. Then, any input vector with one active byte induces an output vector with at least three active bytes. The inputs given in (23) and the corresponding outputs represented by (24) suggest that each column of G has at least three odd elements. Let us next consider the following input–output relation:

$$\begin{pmatrix} 128(g_{11}+g_{12}) \\ 128(g_{21}+g_{22}) \\ 128(g_{31}+g_{32}) \\ 128(g_{41}+g_{42}) \end{pmatrix} = G \begin{pmatrix} 128 \\ 128 \\ 0 \\ 0 \end{pmatrix} \mod 256. \tag{25}$$

Since the branch number is 4 and two input vectors are active, at least two elements of the output vector: $128(g_{i1} + g_{i2})$

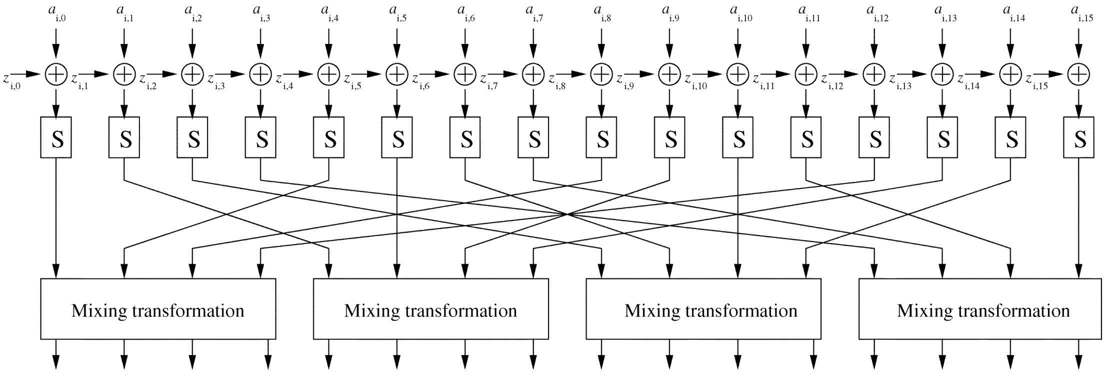


Fig. 5. Round function of the 128-bit uniform cipher. Each $a_{1,i}0 \leq i \leq 15$, is a byte, and the mixing transformation has branch number 4.

$(1 \leq i \leq 4)$ are odd. Without losing generality, we assume that $g_{11} + g_{12}$ and $g_{21} + g_{22}$ are odd. Then, $(g_{11}, g_{12}), (g_{21}, g_{22}) = (\mathrm{odd}, \mathrm{even})$ or (even, odd). Since at most one even number is allowed for each column of G, it follows that

$$\begin{pmatrix} g_{11} & g_{21} \\ g_{12} & g_{22} \\ g_{13} & g_{23} \\ g_{14} & g_{24} \end{pmatrix} = \begin{pmatrix} \text{even} & \text{odd} \\ \text{odd} & \text{even} \\ \text{odd} & \text{odd} \\ \text{odd} & \text{odd} \end{pmatrix} \text{ or } \begin{pmatrix} \text{odd} & \text{even} \\ \text{even} & \text{odd} \\ \text{odd} & \text{odd} \\ \text{odd} & \text{odd} \end{pmatrix}. \tag{26}$$

By applying the same logic to the inputs

$$\begin{pmatrix} 128 \\ 0 \\ 128 \\ 0 \end{pmatrix}, \begin{pmatrix} 128 \\ 0 \\ 0 \\ 128 \end{pmatrix}, \begin{pmatrix} 0 \\ 128 \\ 128 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 128 \\ 0 \\ 128 \end{pmatrix}, \text{and} \begin{pmatrix} 0 \\ 0 \\ 128 \\ 128 \end{pmatrix} \tag{27}$$

we verify that the third and fourth columns also contain exactly one even element, and hence (a) is necessary. In addition, even elements of different columns have to be located in different rows so that all the output vectors corresponding to (27) have two active bytes. As a result, (b) is required. With (a) and (b) at hand, (c) must hold for the inverse to be well-defined on P_S^4. It is easy to see that these conditions are sufficient. ■

For instance

$$G = \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \tag{28}$$

satisfies the conditions of Theorem 1. The eigenvalues are

$$\lambda = 1, -1, 2 \pm \sqrt{5} \tag{29}$$

and $\lambda = 2 + \sqrt{5} > 1$ guarantees the diffusive property as a dynamical system as well as a cryptographical object.

The minimal stretching of the Euclidean distance after applying G becomes

$$\min_{x,y \in P_S^4} d(Gx, Gy) = \min_{x \in P_S^4} d(Gx, 0) \tag{30}$$

$$= \min_i \sqrt{g_{1i}^2 + g_{2i}^2 + g_{3i}^2 + g_{4i}^2} \tag{31}$$

$$= \sqrt{3} \tag{32}$$

which is indeed larger than 1 but far from the best. To improve on it, we note that addition of an even multiple of a column (respectively, row) of G to another column (respectively, row) respects the conditions in Theorem 1. Also, it is not so difficult to independently construct a proper G with large g_{ij} $(1 \leq i, j \leq 4)$. For example

$$G \equiv \begin{pmatrix} 130 & 129 & 129 & 130 \\ 129 & 130 & 129 & 129 \\ 129 & 129 & 128 & 129 \\ 129 & 129 & 129 & 128 \end{pmatrix} \tag{33}$$

has branch number 4, and the eigenvalues are

$$\lambda = 1, -1, 258 \pm \sqrt{258^2 + 1}. \tag{34}$$

Now, $\lambda = 258 + \sqrt{258^2 + 1}$ is much larger than (29). In accordance with this, we obtain (35), shown at the bottom of the page, which is much larger than (32) and close to the possible maximum: $\sqrt{4} \times 128 = 256$.

This G serves as an example that bridges the concepts of diffusion in terms of large branch numbers and diffusion in terms of the Euclidean norm.

V. Analysis of Global Structures

Consider a 128 bit uniform cipher given in Fig. 5 for which the mixing transformation has branch number 4. This global

$$\min_{x,y \in P_S^4} d(Gx, Gy) = \sqrt{(256 - 130)^2 + (256 - 129)^2 + (256 - 129)^2 + (256 - 129)^2} \tag{35}$$

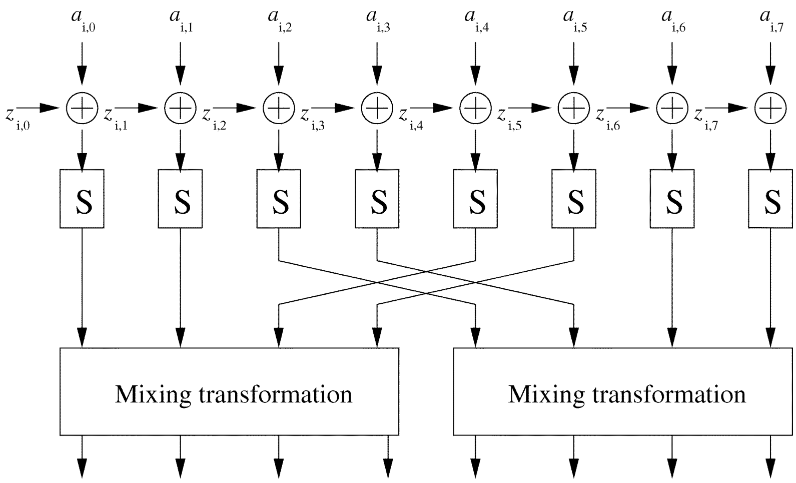


Fig. 6. The F function of the 128-bit Feistel cipher. Each $a_{i,k}$, $0 \leq k \leq 7$, is a byte, and the mixing transformation has a branch number 4.

structure with branch number of the mixing transformation equal to 5 is used in Rijndael [20].

Theorem 2: Every 4-round differential trail of the uniform cipher has at least 16 active S-boxes.

Proof: There is at least one active mixing transformation in the second round. The sum of the active input and output bytes of the active mixing transformation is at least 4. Therefore, the sum of the active mixing transformations in the first and third round is at least 4, and the sum of the active S-boxes of any four round differential trail is at least $4 \times 4 = 16$. ■

Now we consider a 128-bit Feistel cipher whose half-round function F is shown in Fig. 6. The branch number of the mixing transformation is again 4. We denote by a_i and b_i the vectors of eight input bytes to F and eight output bytes of F in the ith half-round, respectively. Let us consider the number of active S-boxes $W(a_i)$ in the ith half-round.

Lemma 2:

(a) If $W(a_i) = W(a_{i+1}) = 0$, then $W(a_j) = 0$ for any half-round j.

(b) If $W(a_{i-1}) = W(a_{i+1}) = 0$, then $W(a_j) = 0$ for any half-round j.

Proof:

(a)

$$\begin{aligned} W(a_i) &= 0 \Rightarrow\ W(b_i) = 0 \\ W(a_{i+1}) = W(b_i) &= 0 \Rightarrow\ W(a_{i-1}) = 0 \\ W(a_{i+1}) &= 0 \Rightarrow\ W(b_{i+1}) = 0 \\ W(a_i) = W(b_{i+1}) &= 0 \Rightarrow\ W(a_{i+2}) = 0 \end{aligned}$$

Continuing in the same way, we find that $W(a_j) = 0$ for any half-round j.

(b)

$$W(a_{i-1}) = W(a_{i+1}) = 0 \Rightarrow\ W(b_i) = 0 \Rightarrow\ W(a_i) = 0$$

The rest follows from (a).

■

Lemma 3:

(a) If $W(a_i) = 0, W(a_{i-1}) \neq 0$, and a_{i+1} exists, then $W(a_{i+1}) \neq 0$.

(b) If $W(a_i) = 0, W(a_{i+1}) \neq 0$, and a_{i-1} exists, then $W(a_{i-1}) \neq 0$.

Proof: Suppose $W(a_{i+1}) = 0$. Then $W(a_{i+1}) = W(a_i) = 0$ and according to Lemma 2, $W(a_j) = 0$ for all possible a_j. This contradicts with $W(a_{i-1}) \neq 0$. We can prove (b) similarly. ■

Theorem 3: Every 4-round differential trail of the Feistel cipher has at least 10 active S-boxes.

Proof: Let us consider a sequence of four half-rounds i, $i+1$, $i+2$ and $i+3$. According to Lemma 3, $W(a_{i+1}) = W(a_{i+2}) = 0$ cannot occur. Assume that $W(a_{i+1}) \neq 0$. This means that there is at least 1 active mixing transformation in round $i+1$, and $W(a_i) + W(a_{i+1}) + W(a_{i+2}) \geq 4$. Similarly, if $W(a_{i+2}) \neq 0$, we get $W(a_{i+1}) + W(a_{i+2}) + W(a_{i+3}) \geq 4$. Hence, the number of active S-boxes over four half-rounds is at least 4, and the number of active S-boxes over eight half-rounds (four rounds) is at least eight. It is not hard to verify that the only possible cases in which there are exactly 4 active S-boxes over four half-rounds are when $W(a_i) = W(a_{i+3}) = 0$ and $W(a_{i+1}) = W(a_{i+2}) = 0$. The only candidate distribution of active S-boxes such that there are 8 active S-boxes over eight half-rounds is (0,2,2,0,0,2,2,0). However, this is not possible since there are two consecutive zero weights. Assume now that there is a weight distribution over eight half-rounds such that there are just 9 active S-boxes. We first consider the case in which there are 5 active S-boxes in the first four half-rounds and 4 active S-boxes in the last four half-rounds. The weight distribution for the last four half-rounds must be (0,2,2,0). Since there is no active S-box in the fifth half-round, there is at least 1 active mixing transformation in the fourth round, and $W(a_3) + W(a_4) = 4$ and $W(a_1) + W(a_2) = 1$ must hold. There are two possibilities: $W(a_1) = 1$, $W(a_2) = 0$ or $W(a_1) = 0$, $W(a_2) = 1$. If $W(a_1) = 0$ and $W(a_2) = 1$, then $W(a_3) = 3$ and $W(a_4) = 4$. If $W(a_1) = 1$ and $W(a_2) = 0$, then $W(a_3) = 1$, $W(a_4) = 3$ and $W(a_5) = 4$. Both cases are in contradiction with our assumption that $W(a_3) + W(a_4) = 4$ and $W(a_5) \neq 4$. Therefore, there is no differential trail with 5 active S-boxes in the first four half-rounds and 4 active S-boxes in the last four half-rounds. Similarly, one can show that there is no differential trail with 4 active S-boxes in the first four half-rounds and

5 active S-boxes in the last four half-rounds. Hence, there is no differential trail over eight half-rounds with less than 10 active S-boxes. ■

VI. Chaotic Block Ciphers

We construct a 128-bit chaotic block cipher with the S-boxes defined in Section III and chaotic mixing transformation proposed in Section IV-C, for which the branch number is 4. As calculated in Section III, the values of DP and LP for the chaotic S-box are $\mathrm{DP} \le 2^{-4}$ and $\mathrm{LP} \le 2^{-3}$, respectively. The length of the key is 128 bits, and therefore, the size of the searchable key space is 2^{128}. We suggest that the cipher has at least 16 rounds. With the help of Theorems 2 and 3, we can estimate the values of DP and LP for the whole cipher.

- *Chaotic uniform cipher*—For the uniform cipher with block diagram shown in Fig. 5, we have $\mathrm{DP} \le 2^{-256}$ and $\mathrm{LP} \le 2^{-192}$.
- *Chaotic Feistel cipher*—For the Feistel cipher with block diagram shown in Fig. 2, where the F function is given in Fig. 6, we have $\mathrm{DP} \le 2^{-160}$ and $\mathrm{LP} \le 2^{-120}$.

Because of the modular structure, the operation speed is much higher than, for example, the cipher proposed in [15], which applies computationally demanding nonlinear transformations to a whole block. Nevertheless, byte multiplications used by $\tilde{F}_A$ and $\tilde{F}_A^{-1}$ make our cipher still slower than commercially or academically appreciated ciphers based on bit calculations. Actually, IDEA [30] is considered slow because of byte multiplications [31].

Remark 1: For an $8 \to 8$ S-box one has $\mathrm{DP} \ge 2^{-7}$ and $\mathrm{LP} \ge 2^{-8}$. We did not attempt to optimize the values of DP and LP for a chaotic S-box and used $\mathrm{DP} \le 2^{-4}$ and $\mathrm{LP} \le 2^{-3}$. However, different approaches yield chaos-based S-boxes with $\mathrm{DP} \le 2^{-5}$ and $\mathrm{LP} \le 2^{-5}$ [32].

Remark 2: A 4-D mixing transformation which has the largest possible branch number 5 is MDS transformation. Theorems corresponding to Theorems 2 and 3, when the mixing transformation is MDS, are proven in the Appendix. If MDS transformation is used, the number of active S-boxes over a trail generally increases. For uniform (respectively, Feistel) cipher any trail over four rounds has at least 25 (respectively, 18) active S-boxes.

Remark 3: The time complexity of a differential (respectively, linear) cryptanalysis attack is slightly greater than $1/\mathrm{DP}$ (respectively, $1/\mathrm{LP}$) block encryptions. In our case, we have

- For the uniform cipher: the time complexity of a DC attack is about 2^{160} block encryptions, and the complexity of an *LC* attack is about 2^{192} block encryptions.
- For the Feistel cipher: the corresponding complexities are 2^{160} and 2^{120}, respectively.

Remark 4: Figs. 5 and 2 (where the F function is given in Fig. 6) show block diagrams of the chaotic block ciphers. We denote the key z as $z = z_0 \| z_1 \| \dots \| z_{16}$, where z_i are bytes and $\|$ denotes concatenation. The S-boxes proposed in this paper depend on the parameter $A \in K_s$, see the discussion in the section Section III-B. The total number of all possible $8 \to 8$ S-boxes is clearly equal to 256!. Out of them, we suggest in this paper $|K_s| = 64$ different S-boxes, which can be described by simple equation; for all of them we have $\mathrm{DP} \le 2^{-4}$ and $\mathrm{LP} \le 2^{-3}$. The suggested estimated values of DP and LP for the whole cipher clearly do not depend of the value of A. Note that we did not explicitly define the map $f : \{0, 1, \dots 255\} \to K_s$, or in other words, we did not specify the dependence of A on z_i.

VII. Further Exploration of Connection Between Chaos and Cryptography

To bridge local and global structures from a dynamical viewpoint, let us consider a toy model of 8-byte chaotic Feistel cipher whose half-round function acts on the 4-byte space. A half round consists of four chaotic S-boxes derived from $\tilde{F}_A$, the PHT mixing layer introduced in Section IV-B, namely

$$H_2 \equiv \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \tag{36}$$

and a 4-byte Hadamard-type permutation given by

$$R_4 \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{37}$$

Following SAFER [21], [22], we interleave H_2 and R_4 by

$$\begin{aligned} \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix} &= H_4 R_4 H_4 \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{pmatrix} \\ &= \begin{pmatrix} 4 & 2 & 2 & 1 \\ 2 & 1 & 2 & 1 \\ 2 & 2 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{pmatrix} \bmod 256 \end{aligned} \tag{38}$$

where

$$H_4 \equiv \begin{pmatrix} H_2 & 0 \\ 0 & H_2 \end{pmatrix}. \tag{39}$$

Actually, $\tilde{F}_A^n$ and F_A^n are distant in the Sobolev sense, particularly when n is large [15], [25]. However, we approximate the expansion exponent of the chaotic S-box by $n\lambda$ where λ is given by (1) with $M = 256$, since $n = 5$ is not so large.

Based on (21), H_2 is chaotic with eigenvalues $\lambda_{\mathrm{PHT1}} = (3 + \sqrt{5})/2$ and $\lambda_{\mathrm{PHT2}} = (3 - \sqrt{5})/2$. Four eigenvalues of R_4 are equal to 1. However, the eigenvalues of $H_4 R_4 H_4$ do not coincide with the products of the eigenvalues of H_2 and those of R_4 because the eigenmodes of H_2 are different from those of R_4. The eigenvalues of $H_4 R_4 H_4$ are given by

$$\lambda_1 = 1, \lambda_2 = -1, \lambda_3 = \frac{7 + 3\sqrt{5}}{2}, \lambda_4 = \frac{7 - 3\sqrt{5}}{2}. \tag{40}$$

The Feistel network introduces another complication. If we denote by $\overline{S}_A$ the S-box layer as a half-block transformation, the half-round is represented by

$$(X, Y) \to (Y + F(X), X), \quad F = H_4 R_4 H_4 \overline{S}_A \tag{41}$$

where X and Y are half blocks. To understand the Feistel structure as a dynamical system, we replace F with the skew tent

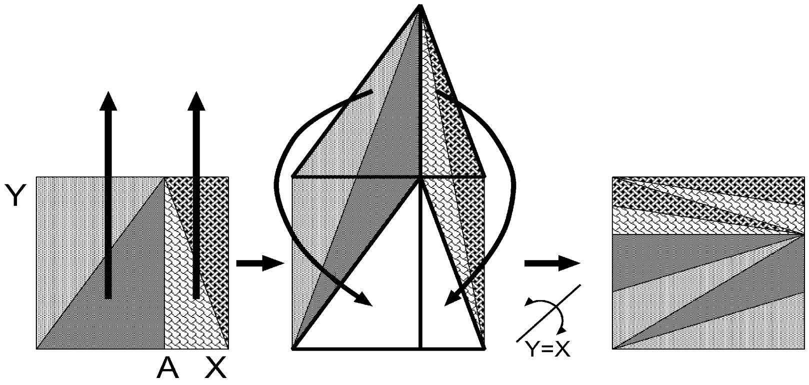


Fig. 7. Half-round of the chaotic Feistel cipher.

map. Equation (41) and chaoticity of F cause hyperbolicity and a horseshoe, as schematically shown in Fig. 7. Then, existence of chaotic attractors is strongly suggested [26, p. 26], [27, p. 255, p. 306], [33, p. 18]. Chaos also emerges from the Feistel network on the p-adic norm space derived from another sort of space discretization. For example, if F is the p-adic logistic map, the Feistel network is homeomorphic to the Smale horseshoe in the p-adic sense [34].

Once again, the Lyapunov exponents of the entire half-round are different from straightforward combination of λ, λ_{PHT1}, λ_{PHT2}, and λ_i $(1 \leq i \leq 4)$. Using (41), they are evaluated as

$$\begin{aligned}\lambda''_{i,1} &= \log\left(\frac{e^{n\lambda}\lambda_i + \sqrt{e^{2n\lambda}\lambda_i^2 + 4}}{2}\right)\\ \lambda''_{i,2} &= \log\left(\frac{e^{n\lambda}\lambda_i - \sqrt{e^{2n\lambda}\lambda_i^2 + 4}}{2}\right), \qquad 1 \leq i \leq 4. \end{aligned} \quad (42)$$

Since

$$\sum_{i=1}^{4}\lambda''_{i,1} + \sum_{i=1}^{4}\lambda''_{i,2} = 0 \quad (43)$$

and $\lambda''_{i,1}$, $\lambda''_{i,2} \neq 0$ $(1 \leq i \leq 4)$, the chaotic Feistel cipher is a hyperbolic, chaotic, and mixing automorphism.

VIII. Conclusion

We have proposed chaotic cryptographical objects such as S-boxes and mixing transformations, a chaotic Feistel cipher, and a chaotic uniform cipher. We have also examined diffusion and confusion properties in terms of both the Euclidean geometry, which dynamical systems are based on, and the bit and byte spaces, which modern cryptography relies on.

Although we have established some links between chaos and cryptography, the connection is still very weak in some points. For example, application of the theory of continuous chaotic maps to their discretization is still in its infancy at an S-box level. Indeed, discrete-space counterparts of chaoticity and ergodicity (as identified by long periodic orbits), the sensitive dependence on initial conditions, positive Lyapunov exponents and entropy have been explored [5], [15], [19], [25], [35]. However, every orbit of a discretized map on a finite space is periodic and hence not chaotic, and space discretization may significantly alter dynamical systems [19], [25], [26], [36]. Moreover, a discretized map is ambiguously consistent with infinitely many continuous-state maps because infinite kinds of extrapolation are possible. Protocols implemented on analog hardware are free from these difficulties. However, they suffer from thermal noise and inevitable insecurity due to local stability required for reliable communications [37].

For binary functions, LP measures nonlinearity, and its properties can be mathematically proved in some cases [23], [38]. In contrast, it seems that dynamical system theory lacks measures for nonlinearity [4]; the Lyapunov exponents, the entropy, and the mixing property measure linear spreading. Fig. 4 demonstrates this fact; $A \cong M/2$ is among the weakest in terms of DP and LP although the Lyapunov exponent of $\tilde{F}_{A=M/2}$, or $\lambda = \log 2$, is the larger than with any other $\tilde{F}_A$. By the same token, a skew tent map with a specific A share identical values of topological entropy and metric entropy with the logistic map, which is even qualitatively different from the skew tent map.

Another missing link is difference in the concept of complexity or randomness in cryptography and that in dynamical system theory. Complex dynamical systems are those with high algorithm complexity. Consequently, the degree of randomness is associated with the minimal length of the program necessary to generate a sequence. However, cryptographical randomness means computational complexity required for cryptanalysis, or capability as pseudo random-number generators [4], [39]. These topics are warranted for future work.

Appendix

We have used mixing transformations with branch number 4 in favor of dynamical-system approaches. In this appendix, we prove some theorems for the MDS mixing transformations with branch number 5. Although MDS matrices do not have dynamical-system origins, as proved in Theorem 1, they are cryptographically better and can be used in combination with chaotic crypto objects to produce hybrid ciphers.

A. Uniform Structure

For the uniform network shown in Fig. 5, the following results are known with mixing transformations that have branch number 5 [20].

Theorem 4: The weight of a two-round trail with Q active columns at the input of the second round is lower-bounded by $5Q$.

Lemma 4: In a two-round trail, the sum of the number of active columns at its input and the number of active columns at its output is at least 5. In other words, the sum of the column weights of a_0 and a_2 is at least 5.

Theorem 5: Any trail over four rounds has at least 25 active S-boxes.

B. Feistel Structure

We consider a Feistel cipher with a half-round function F as shown in Fig. 6. Here we assume that the branch number is 5.

Let us consider four rounds. If the weight $W(a_i)$ of the vector a_i at the input of the ith half-round function is greater than one for all eight half-rounds, then the number of active S-boxes is obviously greater than or equal to sixteen. If there is a trail over four rounds with less than sixteen active S-boxes, then $W(a_j)$ must be less than two for a half-round $\mathrm{HR_j}$. The following lemmas focus on this situation (i.e., the cases $W(a_j) = 1$ and $W(a_j) = 0$ for some half-round $\mathrm{HR_j}$). To spare space, we omit the proofs.

Lemma 5: $W(a_{i-1}) + W(a_i) + W(a_{i+1}) \geq 5m$, where m is the number of active mixing transformations in the ith half-round HR_i.

Lemma 6: If $W(a_j) > 0$ for $j = i-4, \ldots, i+4$ and $W(a_i) = 1$, then the following are true.

1. At least one of HR_{i-4}, HR_{i-3}, HR_{i-2} and HR_{i-1}, has two active mixing transformations.
2. At least one of HR_{i-3}, HR_{i-2}, HR_{i-1} and HR_{i+1}, has two active mixing transformations.
3. At least one of HR_{i+4}, HR_{i+3}, HR_{i+2} and HR_{i+1}, has two active mixing transformations.
4. At least one of HR_{i+3}, HR_{i+2}, HR_{i+1} and HR_{i-1}, has two active mixing transformations.

Lemma 7: If $W(a_i) = 0$, then, the following are true.

1. At least one of HR_{i-3}, HR_{i-2} and HR_{i-1}, has two active mixing transformations.
2. If both HR_{i-2} and HR_{i-1} have only one active mixing transformations, then $W(a_{i-1}) > 2$.
3. At least one of HR_{i+3}, HR_{i+2} and HR_{i+1}, has two active mixing transformations.
4. If both HR_{i+2} and HR_{i+1} have only one active mixing transformations, then $W(a_{i+1}) > 2$.

Based on the analysis above, one can conclude that the number of active S-boxes in a trail over four rounds is less than sixteen only for the following two distributions of active S-boxes over eight half-rounds: (1,0,1,4,5,2,0,2) and (2,0,2,5,4,1,0,1). Obviously, these trails have fifteen active S-boxes.

Theorem 6: Any trail over four rounds has at least fifteen active S-boxes.

ACKNOWLEDGMENT

The authors wish to thank K. Umeno for helpful discussions on this work.

REFERENCES

[1] L. M. Pecora and T. L. Carroll, "Synchronization in chaotic systems," *Phys. Rev. Lett.*, vol. 64, pp. 821–824, 1990.

[2] G. Kolumban, M. P. Kennedy, G. Kis, and Z. Jako, "FM-DCSK: A novel method for chaotic communications," in *Proc. IEEE Int. Symp. ISCAS'98*, , 1998, vol. 4, pp. 477–480.

[3] M. Sushchik, N. Rulkov, L. Larson, L. Tsimring, H. Abarbanel, K. Yao, and A. Volkovskii, "Chaotic pulse position modulation: A robust method of communicating with chaos," *IEEE Commun. Lett.*, vol. 4, no. 4, pp. 128–130, Apr. 2000.

[4] L. Kocarev, "Chaos-based cryptography: A brief overview," *IEEE Circuits Syst. Mag.*, vol. 1, pp. 6–21, 2001.

[5] F. Dachselt and W. Schwarz, "Chaos and cryptography," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 48, no. 12, pp. 1498–1509, Dec. 2001.

[6] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone, *Handbook of Applied Cryptography*. Boca Raton, FL: CRC, 1997.

[7] B. Schneier, *Applied Cryptography: Protocols, Algorithms, and Source Code in C*. New York: Wiley, 1996.

[8] J. Fridrich, "Symmetric ciphers based on two-dimensional chaotic maps," *Int. J. Bifurc. Chaos*, vol. 8, pp. 1259–1284, 1998.

[9] T. Habutsu, Y. Nishio, I. Sasase, and S. Mori, "A secret key cryptosystem by iterating chaotic map," in *Proc. EUROCRYPT'91*, 1991, vol. 547, LNCS, pp. 127–140.

[10] G. Jakimoski and L. Kocarev, "Chaos and cryptography: Block encryption ciphers based on chaotic maps," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 48, no. 1, pp. 163–169, Jan. 2001.

[11] L. Kocarev, G. Jakimoski, T. Stojanovski, and U. Parlitz, "From chaotic maps to encryption schemes," in *Proc. IEEE Int. Symp. ISCAS'98*, 1998, vol. 4, pp. 514–517.

[12] L. Kocarev, M. Sterjev, and P. Amato, "RSA Encryption algorithm based on torus automorphisms," in *Proc. IEEE Int. Symp. ISCAS'04*, 2004, vol. IV, pp. 577–580.

[13] Z. Kotulski and J. Szczepanski, "Discrete chaotic cryptography," *Ann. Phys.*, vol. 6, pp. 381–394, 1997.

[14] N. Masuda and K. Aihara, "Chaotic cipher by finite-state Baker's map," (in Japanese) *Trans. IEICE*, vol. J82-A, pp. 1038–1046, 1999.

[15] ——, "Cryptosystems with discretized chaotic maps," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 49, no. 1, pp. 28–40, Jan. 2002.

[16] T. Beth, D. E. Lazic, and A. Mathias, "Cryptanalysis of cryptosystems based on remote chaos replication," in *Proc. CRYPTO'94*, 1994, vol. 839, LNCS, pp. 318–331.

[17] E. Biham, "Cryptanalysis of the chaotic-map cryptosystem suggested at EUROCRYPT'91," in *Proc. EUROCRYPT'91*, 1991, vol. 547, LNCS, pp. 532–534.

[18] G. Jakimoski and L. Kocarev, "Analysis of some recently proposed chaos-based encryption algorithms," *Phys. Lett. A*, vol. 291, pp. 381–384, 2001.

[19] L. Kocarev, P. Amato, D. Ruggiero, and I. Pedaci, "Discrete Lyapunov exponent for Rijndael block Cipher," in *Proc. NOLTA'04*, 2004, pp. 609–612.

[20] J. Daemen and V. Rijmen, *The Design of Rijndael: AES—The Advanced Encryption Standard*. Berlin, Germany: Springer Verlag, 2002.

[21] J. L. Massey, "SAFER K-64: A byte-oriented block-ciphering algorithm," in *Fast Software Encryption, 1993*. Cambridge, U.K.: Springer-Verlag, 1994, vol. 809, LNCS, pp. 1–17.

[22] J. Massey, G. Khachatrian, and M. Kuregian, *Nomination of SAFER+ as Candidate Algorithm for the Advanced Encryption Standard (AES)*. Sunnyvale, CA: Cylink, 1998.

[23] V. Rijmen, J. Daemen, B. Preneel, A. Bosselaers, and E. De Win, "The Cipher SHARK," in *Fast Software Encryption, 1996*. Cambridge, U.K.: Springer-Verlag, 1997, vol. 1039, LNCS, pp. 99–111.

[24] B. Schneier, J. Kelsey, D. Whiting, D. Wagner, C. Hall, and N. Ferguson, *Twofish: A 128-Bit Block Cipher*. : Counterpane Internet Security, Inc., 2000 [Online]. Available: http://www.counterpane.com/

[25] N. Masuda and K. Aihara, "Dynamical characteristics of discretized chaotic permutations," *Int. J. Bifurc. Chaos*, vol. 12, pp. 2087–2103, 2002.

[26] A. E. Jackson, *Perspective of Nonlinear Dynamics*. Cambridge, U.K.: Cambridge Univ. Press, 1989.

[27] C. Robinson, *Dynamical Systems; Stability, Symbolic Dynamics, and Chaos*. Boca Raton, FL: CRC, 1995.

[28] Y. G. Sinai, Ed., *Encyclopaedia of Mathematical Sciences, Dynamical Systems II*. New York: Springer-Verlag, 1989.

[29] A. M. F. Rivas, M. Saraceno, and A. M. Ozorio de Almeida, “Quantization of Multidimensional Cat Maps,” *Nonlinearity*, vol. 13, pp. 341–376, 2000.

[30] X. Lai and J. Massey, “A proposal for a new block encryption standard,” in *Proc. EUROCRYPT’90*, 1991, vol. 473, LNCS, pp. 389–404.

[31] B. Preneel, V. Rijmen, and A. Bosselaers, “Recent developments in the design of conventional cryptographic algorithms,” in *COSIC’97 Course*, 1998, vol. 1528, LNCS, pp. 105–130.

[32] J. Szczepanski, J. M. Amigo, T. Michalek, and L. Kocarev, “Cryptographically secure substitutions based on the approximation of mixing maps,” *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 52, no. 2, pp. 443–453, Feb 2005.

[33] J. Guckenheimer, J. Moser, and S. E. Newhouse, *Dynamical Systems: C.I.M.E. Lectures Bressanone, Italy, June 1978*. Boston, MA: Birkhäuser, 1980.

[34] C. F. Woodcock and N. P. Smart, “p-adic chaos and random number generation,” *Exp. Math.*, vol. 7, pp. 334–342, 1998.

[35] L. Kocarev and J. Szczepanski, “Finite-space Lyapunov exponents and pseudochaos,” *Phys. Rev. Lett.*, vol. 93, p. 234 101, 2004.

[36] M. Blank, “Pathologies generated by round-off in dynamical systems,” *Phys. D*, vol. 78, pp. 93–114, 1994.

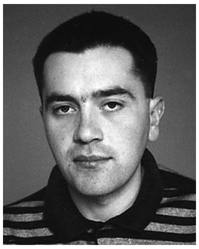
[37] K. M. Short, “Unmasking a modulated chaotic communications scheme,” *Int. J. Bifurc. Chaos*, vol. 6, pp. 367–375, 1996.

[38] J. H. Cheon, S. Chee, and C. Park, “S-boxes with controllable nonlinearity,” in *Proc. EUROCRYPT’99*, 1999, vol. 1592, LNCS, pp. 286–294.

[39] L. Kocarev and G. Jakimoski, “Pseudorandom bits generated by chaotic maps,” *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 50, no. 1, pp. 123–126, Jan. 2003.

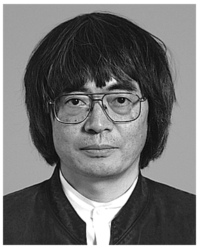
Naoki Masuda received the B.E., Master’s, and Ph.D. degrees in mathematical engineering and information physics from the University of Tokyo, Tokyo, Japan, in 1998, in 2000, and in 2002, respectively.

Currently, he is a Postdoctoral Fellow at RIKEN Brain Science Institute, Wako, Japan. His research interests include chaotic cryptosystems, computational neuroscience, complex networks, and interacting particle systems.



Goce Jakimoski received the M.Sc. degree in electrical engineering from the Ss. Cyril and Methodius University, Skopje, Macedonia, in 1998. He is working toward the Ph.D. degree in computer science, at the Florida State University, Tallahassee, since 2001.

His research interests are in the area of cryptography, network and computer security.



Kazuyuki Aihara received the B.E. degree in electrical engineering and the Ph.D. degree in electronic engineering from the University of Tokyo, Tokyo, Japan, in 1977 and 1982, respectively.

Currently, he is a Professor in Institute of Industrial Science and Graduate School of Information Science and Technology, the University of Tokyo. He is also Director of Aihara Complexity Modeling Project, ERATO, JST. His research interests include mathematical modeling of complex systems, parallel distributed processing with chaotic neural networks, and nonlinear time series analysis.

Ljupco Kocarev (F’06) is a Research Scientist at the Institute for Nonlinear Science, University of California San Diego, and Professor at the Graduate School of Electrical Engineering, University “Kiril i Metodij,” Skopje, Macedonia. His scientific interests include nonlinear systems and circuits, coding and information theory, networks and networks on chip, and cryptography. He has coauthored more than 100 journal papers in 18 different international peer-reviewed journals ranging from mathematics to physics and from electrical engineering to computer sciences.

Dr. Kocarev is a foreign member of the Macedonian Academy of Sciences and Arts. According to the *Science Citation Index*, his work has been cited more than 2500 times.