Quantitative Evaluation of Chaotic CBC Mode of Operation

Abdessalem Abidi, Qianxue Wang, Belgacem Bouallègue, Mohsen Machhout, and Christophe Guyeux

November 24, 2016

Abstract

In cryptography, block ciphers have a very simple principle. They do not treat the original text bit by bit but they manipulate blocks of text. This manipulation follows various ways which are called block cipher modes of operation. Each one of these modes possesses owns characteristics and its specific security properties. Among them, we quote the Cipher Block Chaining (CBC) mode. This paper presents a complete topological study of the CBC block cipher mode of operation. To do this, we began by proving the chaotic behavior of this mode according to Devaney.

Keywords: Cipher Block Chaining, Mode of operation, Block cipher, Chaos, Devaney's chaos, sensivity, expansivity

1 Introduction

Block ciphers have a very simple principle. They do not treat the original text bit by bit but they manipulate blocks of text – for instance, a block of 64 bits for the DES (Data Encryption Standard) or a block of 128 bits for the AES (Advanced Encryption Standard) algorithm. In fact, the original text is broken into blocks of N bits. For each block, the encryption algorithm is applied to obtain an encrypted block that has the same size. Then we gather all blocks, which are encrypted separately, to obtain the complete encrypted message. For decryption, we proceed in the same way, but this time starting from the cipher text to obtain the original message using the decryption algorithm instead of the encryption function. So, it is not sufficient to put anyhow a block cipher algorithm in a program. We can instead use these algorithms in various ways according to their specific needs. These ways are called the block cipher modes of operation. There are several modes of operation and each mode has owns characteristics and its specific security properties. In this paper, we will consider only one of these modes, which is the cipher block chaining (CBC) mode, and we will study it according to chaos.

The chaos theory we consider in this paper is the Devaney's topological one [7]. In addition to being recognized as one of the best mathematical definition of chaos, this theory offers a framework with qualitative and quantitative tools to evaluate the notion of unpredictability [5]. As an application of our fundamental results, we are interested in the area of information safety and security. In this research work, which is an extension of our previous article [1], the theoretical study of the chaotic behavior for the CBC mode of operation is deepened by evaluating its level of sensibility and expansivity [4].

The remainder of this research work is organized as follows. In Section 2, we will recall some basic definitions concerning chaos and cipher-block chaining mode of operation. Section 3 is devoted to the recall of our previous research works. In Section 4 quantitative topological properties for chaotic CBC mode of operation is studied in detail. This research work ends by a conclusion section in which our contribution is recalled and some intended future work are proposed.

2 Basic recalls

This section is devoted to basic definitions and terminologies in the field of topological chaos and in the one of block cipher mode of operation.

2.1 Devaney's Chaotic Dynamical Systems

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

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

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

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

Definition 3 (Devaney's formulation of chaos [7]) The function f is *chaotic* on $(\mathcal{X}, \tau)$ if f is regular and topologically transitive.

The chaos property is strongly linked to the notion of "sensitivity", defined on a metric space $(\mathcal{X}, d)$ by:

Definition 4 The function f has *sensitive dependence on initial conditions* if there exists $\delta > 0$ such that, for any $x \in \mathcal{X}$ and any neighborhood V of x, there exist $y \in V$ and $n > 0$ such that

$$d\left(f^n(x), f^n(y)\right) > \delta.$$

δ is called the *constant of sensitivity* of f.

Indeed, Banks *et al.* have proven in [6] that when f is chaotic and $(\mathcal{X}, d)$ is a metric space, then f has the property of sensitive dependence on initial conditions (this property was formerly an element of the Devaney's definition of chaos). Additionally, the transitivity property is often obtained as a consequence of the strong transitivity one, which is defined below [9].

Definition 5 f is *strongly transitive* on $(\mathcal{X}, d)$ if, for all point $x, y \in \mathcal{X}$ and for all neighborhood $\mathcal{V}$ of x, it exists $n \in \mathbb{N}$ and $x' \in \mathcal{V}$ such that $f^n(x') = y$.

Finally, a function f has a constant of *expansivity* equal to ε if an arbitrarily small error on any initial condition is *always* magnified until ε [9]. Mathematically speaking,

Definition 6 The function f is said to have the property of *expansivity* if $\exists \varepsilon > 0$, $\forall x \neq y$, $\exists n \in \mathbb{N}$, $d(f^n(x), f^n(y)) \geqslant \varepsilon$.

Then, ε is the *constant of expansivity* of f. We also say that f is ε-expansive.

2.2 CBC properties

Like some other modes of operation, the CBC mode requires not only a plaintext but also an initialization vector (IV) as input. In what follows, we will show how this mode of operation works in practice.

2.2.1 Initialisation vector IV

As what has been already announced, in addition to the plaintext, the CBC mode of operation requires an initialization vector (denoted as IV in what follows) in order to randomize the encryption. This vector is used to produce distinct ciphertexts even if the same plaintext is encrypted multiple times, without the need of a slower re-keying process [10].

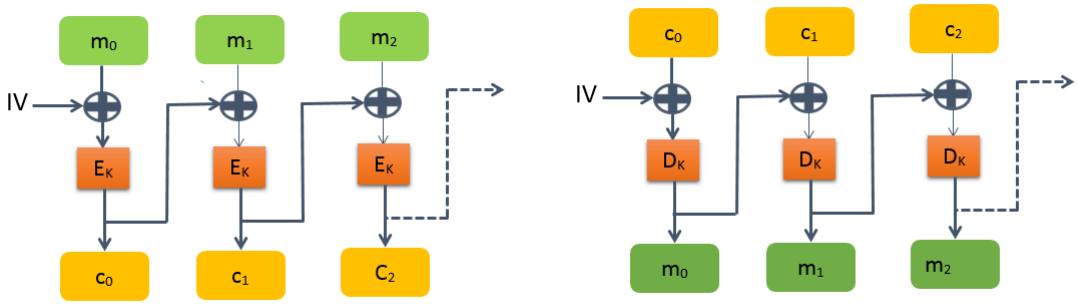
An initialization vector must be generated for each execution of the encryption operation, and the same vector is necessary for the corresponding execution of the decryption operation, see Figure 1. Therefore the IV, or information that is sufficient to calculate it, must be available to each party of any communication. The initialization vector does not need to be secret, so the IV, or information sufficient to determine the IV, may be transmitted with the cipher text. In addition, the initialization vector must be unpredictable: for any given plaintext, it must not be possible to predict the IV that will be associated to the plaintext, in advance to the vector generation [8].

2.2.2 CBC mode characteristics

Cipher block chaining is a block cipher mode that provides confidentiality but not message integrity in cryptography. The CBC mode offers a solution to the greatest part of the problems presented by the ECB (Electronic codebook) for example [12] as, due to the CBC mode, the encryption will depend on the context. Indeed, the cipher text of each encrypted block will depend not only on the initialization vector IV but also on the plaintext of all preceding blocks. Specifically, the binary operator XOR is applied between the current bloc of the plaintext and the previous block of the cipher text, as depicted in Figure 1. Then, we apply the encryption function to the result of this operation. For the first block, the initialization vector takes place of the previous cipher text block.

CBC mode has several advantages. In fact, this mode encrypts the same plaintext differently with different initialization vectors. In addition, the encryption of each block depends on the preceding block and therefore, if the order of the cipher text blocks is modified, the decryption will be impossible and the recipient realizes the problem. Furthermore, if a transmission error affects the encrypted block C_i, then only the blocks m_i and m_{i+1} are assigned, the other blocks will be determined correctly.

The main objective of this series of articles regarding the chaotic topological behavior of the CBC mode of operation is to understand in which extent this mode depends on its inputs. More precisely, is it possible to understand this dependence, in such a way that the effects of a modification of the IV and/or the message can be predicted? If so, this kind of weakness could be considered in the design of specific attacks, while if the converse is proven, that is to say, if the mid-to-long term effects of a slight modification of the input cannot be predicted, that chaotic dependence will make such attacks inefficient.



(a) CBC encryption mode

(b) CBC decryption mode

Figure 1: CBC mode of operation

In what follows, we will summarize the results which proof the chaotic behaviour of the CBC mode of operation.

3 Previously obtained results

In this section, we are interested to recall our previous results. They have been detailed in [1], in which we have proven that some well chosen block ciphers can lead to a chaotic behavior of the CBC mode of operation. Indeed, this mode can be seen as a discrete dynamical system (recurrent sequence), whose evolution can thus be studied using common tools taken from the mathematical analysis [2].

3.1 Modeling the CBC mode as a dynamical system

Our modeling follows a same canvas than what has be done for hash functions [3, 9] or pseudorandom number generation [2]. Let us consider the CBC mode of operation with a keyed encryption function $\varepsilon_k : \mathbb{B}^{\mathsf{N}} \to \mathbb{B}^{\mathsf{N}}$ depending on a secret key k, where N is the size for the block cipher, and $\mathcal{D}_k : \mathbb{B}^{\mathsf{N}} \to \mathbb{B}^{\mathsf{N}}$ is the associated decryption function, which is such that $\forall k, \varepsilon_k \circ \mathcal{D}_k$ is the identity function. We define the Cartesian product $\mathcal{X} = \mathbb{B}^{\mathsf{N}} \times \mathcal{S}_{\mathsf{N}}$, where:

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

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

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

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

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

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

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

This function returns the inputted binary vector x, whose m_j-th components x_{m_j} have been replaced by $f(x)_{m_j}$, for all $j = 1..\mathsf{N}$ such that $m_j = 0$. In case where f is the vectorial negation, this function will correspond to one XOR between the clair text and the previous encrypted state. So the CBC mode of operation can be rewritten as the following dynamical system:

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

For any given $g : [\![0, 2^{\mathsf{N}} - 1]\!] \times \mathbb{B}^{\mathsf{N}} \longrightarrow \mathbb{B}^{\mathsf{N}}$, we denote $G_g(X) = (g(i(X_1), X_2); \sigma(X_1))$ (when $g = \mathcal{E}_k \circ F_{f_0}$, we obtain one cypher block of the CBC, as depicted in Figure 1). So the recurrent relation of Eq.(1) can be rewritten in a condensed way, as follows.

$$X^{n+1} = G_{\mathcal{E}_k \circ F_{f_0}}\left(X^n\right). \tag{2}$$

With such a rewriting, one iterate of the discrete dynamical system above corresponds exactly to one cypher block in the CBC mode of operation. Note that the second component of this system is a subshift of finite type, which is related to the symbolic dynamical systems known for their relation with chaos [11]. We now define a distance on $\mathcal{X}$ as follows: $d((x,m);(\check{x},\check{m})) = d_e(x,\check{x}) + d_m(m,\check{m})$, where:

$$\begin{cases} d_e(x,\check{x}) & = \sum_{k=1}^{\mathsf{N}} \delta(x_k,\check{x}_k) \\ \\ d_m(m,\check{m}) & = \dfrac{9}{\mathsf{N}} \displaystyle\sum_{k=1}^{\infty} \dfrac{\sum_{i=1}^{\mathsf{N}} |m_i - \check{m}_i|}{10^k}. \end{cases}$$

This distance has been introduced to satisfy the following requirements:

- The integral part between two points X, Y of the phase space $\mathcal{X}$ corresponds to the number of binary components that are different between the two internal states X_1 and Y_1.
- The k-th digit in the decimal part of the distance between X and Y is equal to 0 if and only if the k-th blocks of messages X_2 and Y_2 are equal. This desire is at the origin of the normalization factor $\dfrac{9}{\mathsf{N}}$.

3.2 Proofs of chaos

As mentioned in Definition 3, a function f is *chaotic* on $(\mathcal{X},\tau)$ if f is regular and topologically transitive. We have began by stating some propositions that are primarily required in order to proof the chaotic behavior of the CBC mode of operation.

Proposition 1 *Let $g = \varepsilon_k \circ F_{f_0}$, where ε_k is a given keyed block cipher and $f_0 : \mathbb{B}^{\mathsf{N}} \longrightarrow \mathbb{B}^{\mathsf{N}}$, $(x_1,...,x_{\mathsf{N}}) \longmapsto (\overline{x_1},...,\overline{x_{\mathsf{N}}})$ is the Boolean vectorial negation. We consider the directed graph $\mathcal{G}_g$, where:*

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

So if $\mathcal{G}_g$ is strongly connected, then G_g is strongly transitive.

proving by doing so the strong transitivity of G_g on $(\mathcal{X}, d)$.

We have then proven that,

Proposition 2 *If $\mathcal{G}_g$ is strongly connected, then G_g is regular.*

According to Propositions 1 and 2, we can conclude that, if the directed graph $\mathcal{G}_g$ is strongly connected, then the CBC mode of operation is chaotic according to Devaney, as established in our previous research work. In that article and for illustration purpose, we have also given some examples of encryption functions making this mode a chaotic one.

We have previously recalled that the mathematical framework of the theory of chaos offers tools to measure this unpredictable behavior quantitatively. The firsts of these measures are the constants of sensitivity and of expansitivy, recalled in the definitions section. We now intend to investigate these measures.

4 Quantitative measures

Let us firstly focus on the sensitivity property of the CBC mode of operation.

4.1 Sensitivity

Proposition 3 *The CBC mode of operation is sensible to the initial condition, and its constant of sensibility is larger than the length* N *of the block size.*

PROOF Let $X = (x; (m^0, m^1, ...)) \in \mathcal{X}$ and $\delta > 0$. We are looking for $X' = (x'; (m'^0, m'^1, ...)$ and $n \in \mathbb{N}$ such that $d(X, X') < \delta$ and $d(G_g^n(X), G_g^n(X')) > N$.

Let us define $k_0 = \lfloor -log_{10}(\delta) \rfloor + 1$, in such a way that all X' of the form:

$$(X_1, (m^0, m^1, ..., m^{k_0}, m'^{k_0+1}, m'^{k_0+2}, ...))$$

are such that $d(X, X') < \delta$. In other words, all messages m' whose k_0 first blocks are equal to $(m^0, m^1, ..., m^{k_0})$ are δ-close to X.

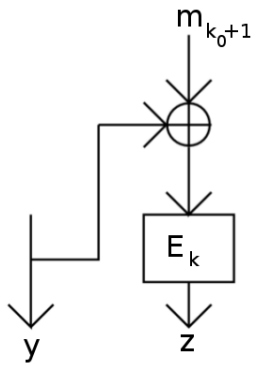


Figure 2: $k_0 + 1$-th iterate of G_g

Let $y = G_g^{k_0}(X)_1$ and $z = G_g^{k_0+1}(X)_1$ as defined in Figure 2. We consider the block message m' defined by:

$$m' = y \oplus \mathcal{D}_k(\overline{z})$$

where $\mathcal{D}_k$ is the keyed decryption function associated to $\mathcal{E}_k$, and $\overline{z}$ is the negation of z. We thus define X' as follow:

- $X'_1 = x$,
- $\forall k \leqslant k_0, m'^k = m^k$,
- $m'^{k_0+1} = m'$,
- $\forall k \geqslant k_0 + 2$, $m'^k = \overline{m^k}$,

so $d(G_g^{k_0+1}(X), G_g^{k_0+1}(X'))$

$$
\begin{aligned}
&= d\left(G_g\left(y;(m_{k_0+1},m_{k_0+2},...)\right),\right.\\
&\qquad \left. G_g\left(y;(m',\overline{m_{k_0+1}},\overline{m_{k_0+2}},...)\right)\right)\\
&= d\left((z;(m_{k_0+2},m_{k_0+3},...)\right),\\
&\qquad \left(E_k(y\oplus m');(\overline{m_{k_0+1}},\overline{m_{k_0+2}},...)\right))\\
&= d\left((z;(m_{k_0+2},m_{k_0+3},...)\right),\\
&\qquad \left(E_k(y\oplus(y\oplus D_k(\overline{z}));(\overline{m_{k_0+1}},\overline{m_{k_0+2}},...)\right))\\
&= d\left((z;(m_{k_0+2},m_{k_0+3},...)\right),\\
&\qquad \left(E_k((y\oplus y)\oplus D_k(\overline{z});(\overline{m_{k_0+1}},\overline{m_{k_0+2}},...)\right))\\
&= d\left((z;(m_{k_0+2},m_{k_0+3},...)\right),\\
&\qquad \left(E_k(0\oplus D_k(\overline{z}));(\overline{m_{k_0+1}},\overline{m_{k_0+2}},...)\right))\\
&= d\left((z;(m_{k_0+2},m_{k_0+3},...)\right),\\
&\qquad \left(E_k(D_k(\overline{z}));(\overline{m_{k_0+1}},\overline{m_{k_0+2}},...)\right))\\
&= d\left((z;(m_{k_0+2},m_{k_0+3},...)\right),(\overline{z};(\overline{m_{k_0+1}},\overline{m_{k_0+2}},...)))\\
&= d_e(z,\overline{z})\\
&\qquad + d_m((m_{k_0+2},m_{k_0+3},...),(\overline{m_{k_0+1}},\overline{m_{k_0+2}},...))\\
&= \mathsf{N}+\frac{9}{\mathsf{N}}\sum_{k=k_0+2}^{\infty}\frac{m_k-\overline{m_k}}{10^k}\\
&= \mathsf{N}+\frac{9}{\mathsf{N}}\sum_{k=k_0+2}^{\infty}\frac{\mathsf{N}}{10^k}\\
&= \mathsf{N}+9\sum_{k=k_0+2}^{\infty}\left(\frac{1}{10^k}\right)=\mathsf{N}+\frac{1}{10^{k_0+1}}>\mathsf{N},
\end{aligned}
$$

which concludes the proof of the sensibility of G_g.

The second important tool that reinforces the chaotic behavior of the CBC mode of operation is the expansivity. The study of this property, which has been recalled in Definition 6, will be regarded below.

4.2 Expansitivity

In this section we offer the proof that:

Proposition 4 *The CBC mode of operation is not expansive.*

Proof Consider for instance two initial vectors $x=(1,0,\ldots,0)$ and $x'=(0,1,0,\ldots,0)$, associated to the messages $m=((0,1,0,\ldots,0),(0,\ldots,0),(0,\ldots,0),\ldots)$ and $m'=((1,0,\ldots,0),(0,\ldots,0),(0,\ldots,0),\ldots)$: all blocks of messages are null in both m and m', except the first block. Let $X=(x,m)$ and $X'=(x',m')$.

Obviously, $x\neq x'$, while $x\oplus m_0=x'\oplus m'_0$. This latter implies that $X_1^0={X'}_1^0$, and by a recursive process, we can conclude that $\forall i\in\mathbb{N}, X_1^i={X'}_1^i$. So the distance between points $X=(x,m)$ and $X'=(x',m')$ is strictly positive, while for all $n>0$, $d\left(G_g^n(X),G_g^n(X')\right)=0$, which concludes the proof of the non expansive character of the CBC mode of operation by the mean of the exhibition of a counter example.

5 Conclusion and future work

In this paper, both expansivity and sensibility of symmetric cyphers are regarded, in the case of the CBC mode of operation. These quantitative topology metrics taken from the mathematical theory of chaos allow to measure in which extent a slight error on the initial condition is magnified during iterations. It is stated that, in addition to being chaotic as defined in the Devaney's formulation, the CBC mode of operation is indeed largely sensible to initial errors or modifications on either the IV or the message to encrypt. Its expansivity has been regarded too, but this property is not satisfied, as it has been established thanks to a counter example.

In future work, we intend to deepen the topological study of the behavior of the CBC mode of operation. We will study whether this mode of operation possesses other qualitative properties of disorder like the topological mixing. Additionally, other quantitative evaluations will be performed, and the level of topological entropy will be evaluated too.

References

[1] Abdessalem Abidi, Qianxue Wang, Belgacem Bouallègue, Mohsen Machhout, and Christophe Guyeux. Proving chaotic behavior of cbc mode of operation. *International Journal of Bifurcation and Chaos*, 2016. Accepted paper, forthcoming.

[2] Jacques Bahi, Xiaole Fang, Christophe Guyeux, and Qianxue Wang. Evaluating quality of chaotic pseudo-random generators. application to information hiding. *IJAS, International Journal On Advances in Security*, 4(1-2):118–130, 2011.

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

[4] Jacques Bahi and Christophe Guyeux. *Discrete Dynamical Systems and Chaotic Machines: Theory and Applications*. Chapman & Hall, CRC Press, jun 2013.

[5] Jacques M Bahi, Raphaël Couturier, Christophe Guyeux, and Pierre-Cyrille Héam. Efficient and cryptographically secure generation of chaotic pseudorandom numbers on gpu. *arXiv preprint arXiv:1112.5239*, 2011.

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

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

[8] Morris Dworkin. Recommendation for block cipher modes of operation. methods and techniques. Technical report, DTIC Document, 2001.

[9] Christophe Guyeux and Jacques Bahi. A topological study of chaotic iterations. application to hash functions. In *CIPS, Computational Intelligence for Privacy and Security*, volume 394 of *Studies in Computational Intelligence*, pages 51–73. Springer, 2012. Revised and extended journal version of an IJCNN best paper.

[10] Kuo-Tsang Huang, Jung-Hui Chiu, and Sung-Shiou Shen. A novel structure with dynamic operation mode for symmetric-key block ciphers. *International Journal of Network Security & Its Applications (IJNSA)*, 5(1):19, 2013.

[11] Douglas Lind and Brian Marcus. *An introduction to symbolic dynamics and coding*. Cambridge University Press, 1995.

[12] Jang Schiltz. Les modes opératoires de la cryptographie symétrique. 2003.