Applied Mathematics and Computation 332 (2018) 123–135

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Substitution box generation using Chaos: An image encryption application

V.M. Silva-García [a], R. Flores-Carapia [a,*], C. Rentería-Márquez [b], B. Luna-Benoso [c], M. Aldape-Pérez [a]

[a] Instituto Politécnico Nacional, CIDETEC, Security, 07700, México
[b] Instituto Politécnico Nacional, ESFM, Code Theory, 07738, México
[c] Instituto Politécnico Nacional, ESCOM, Programming, 07738, México

ARTICLE INFO

Keywords:
Image encryption
Chaos
Walsh function
Advanced Encryption Standard
Randomness

ABSTRACT

There are procedures to encrypt images; however, sometimes there is a loss of information in the decryption process or the key set size is not specifically mentioned. In this research, substitution boxes are built for the Advanced Encryption Standard (AES) cryptosystem using Chaos, and generated by a non-linear differential equation. The boxes' non-linearity is quantified using the Walsh function. One thousand twenty four boxes are chosen with a non-linearity of 106. To generate a pseudorandom permutation over 256 elements, an algorithm that defines a bijective function is employed. The AES utilized in this article uses 128 bit keys and applies a box in each round; that is, using an array of 10 boxes for each plaintext block of 128 bits. An encryption application for color images is presented. The degree of the encrypted images' randomness is measured to quantify the cipher quality. Image encryption is performed without information loss. The aim in future is to design a device to encrypt video in a robust manner and in real time without loss of information.

1. Introduction

At present, sensitive information is transmitted on public networks, such as the Internet. For this reason, different types of encryption techniques have emerged, particularly for images [1–3]. Furthermore, several aspects should be considered in the images encryption, namely: the first aspect is the proposed system complexity, which leads directly to the elements number of the keys set, for example, there are cryptosystems that have as maximum 2^{199} keys [4,5], although, the proposal system can have as a maximum 2^{256}. As a second aspect, it is pointed out that there are investigations that do not perform randomness tests on encrypted or pre-encrypted images with instruments such as Entropy and Correlation [6], but in this investigation randomness evaluations are carried out using these tools.

The third aspect refers to recent attacks that have been made to the AES cryptosystem which will be mentioned later. However, there are important works that are not indicated [7]. Regarding the fourth point about the encryption process;

* Corresponding author.
E-mail addresses: vsilvag@ipn.mx (V.M. Silva-García), rfloresca@ipn.mx (R. Flores-Carapia), crenteri@esfm.ipn.mx (C. Rentería-Márquez), blunab@ipn.mx (B. Luna-Benoso), maldape@ipn.mx (M. Aldape-Pérez).

https://doi.org/10.1016/j.amc.2018.03.019

that is, the substitution boxes are important since these give non-linearity to the encryption procedure, Nevertheless, investigations have been carried out that do not apply substitution boxes [8]. Also, there are investigations where the substitution boxes non-linearity is lower than in this research [2].

The fifth point refers to the results, Entropy particularly. This work reports results that improve the Entropy of other works [9,10]. A sixth issue to consider is the following: investigations have been made that do not apply randomness measurements to the encryption image that are included in the NIST 800-22 standard [11]. However, in this work the Discrete Fourier Transform is applied [12].

The seventh aspect refers to the data loss in the information encryption [13]. In this research, to encrypt the information without loss is proposed, since there are countries, for example México, whose regulation does not permit this [14].

An algorithm that defines a bijective function [15] was developed that is used to obtain boxes of 16×16 elements, i.e., a permutation over a 256 number array is built. To generate this permutation on a 256 element arrangement, 256 constants are required; these constants are obtained using Chaos. This procedure is described in detail below.

Chaos is generated by non-linear differential Eq. (1).

The boxes' non-linearity is measured using the Walsh function $W_{f_i} = \rho(f_i \oplus D_a)$ where f_i is a Boolean function that defines a correspondence $f_i : \mathbb{Z}_2^8 \longrightarrow \mathbb{Z}_2$, with $\mathbb{Z}_2 = \{0, 1\}$ [16].

To measure the randomness degree of an encrypted figure, four tests are applied, namely: entropy, correlation coefficient (or simply "correlation", discrete Fourier transform (DFT), and a goodness-of-fit test using the chi-square distribution [17]. For many studies that perform randomness analysis of encrypted figures, entropy and correlation tests are utilized [18]. Furthermore, some of the figures used in this research ("Plane", "Lena", "Jet" and "Giraffe") are employed in other encryption investigations [19,20].

This paper is organized as follows: In the present section a short description of the state of the art was carried out; in the second, a short explanation of the tools used is presented. In the third section a description and proof of the algorithm that defines the bijective function is addressed; in the fourth, boxes are generated and non-linearity is measured. In the fifth section, a box with a non-linearity of 106 is presented. In the sixth section, the randomness results are presented. In the seventh the results are analysed, and in the final section the conclusions are shown.

2. Encryption tools

2.1. Chaos

Let us have a non-linear differential equation

$$\frac{dP(t)}{dt} = f(t, P(t)), \tag{1}$$

where $f(t, P(t)) = aP(t) - bP^2(t)$ with $a, b > 1$. Moreover, the time variable is discretized; put another way, taking the values $t_0, t_1 \ldots t_n$ and for these quantities the following values can be written: $P(t_0), P(t_1), \ldots P(t_n)$.

Then, the Euler algorithm [21] gives the following:

$$P(t_{n+1}) = P(t_n) + f(t_n, P(t_n)) \times \triangle t, \tag{2}$$

where $\triangle t$ is a very small amount.

The parameters r, s are defined as follows:

$$r = 1 + a \times \triangle t, \tag{3}$$

$$s = b \times \triangle t \tag{4}$$

In addition, the consideration that the $P(t)$ function takes on the point t_n, the value $P(t_n) = \frac{r}{s} t_n$ [21].

Substituting this result in Eq. (2) gives:

$$t_{n+1} = r \times t_n(1 - t_n) \tag{5}$$

Thus, the limit of the t_n variable when $n \to \infty$ can be written for Eq. (5). That is,

$$\lim_{n\to\infty} t_{n+1} = \lim_{n\to\infty} r \times t_n(1 - t_n). \tag{6}$$

In practical cases, say $n = 2000$, the t value is stabilized when the limit exists. For example, in Table 1 some values of t appear for $n = 2000$, $t_0 = 0.71828182845$, and various values of r.

Table 1
The t value when $n = 2000$ and $t_0 = 0.71828182845$.

r	1.2	1.4	1.6	1.8	2
t	0.16666	0.285714285	0.375	0.4444	0.5