Available online at www.sciencedirect.com
SCIENCE DIRECT®

Physics Letters A 354 (2006) 101–109

PHYSICS LETTERS A

www.elsevier.com/locate/pla

Analysis and improvement for the performance of Baptista's cryptographic scheme

Jun Wei [a,b,*], Xiaofeng Liao [a], K.W. Wong [c], Tsing Zhou [a], Yigui Deng [a]

[a] *Department of Computer Science and Engineering, Chongqing University, Chongqing 400044, PR China*
[b] *Zunyi Medical College, Zhunyi, 563000 Guizhou, PR China*
[c] *Department of Electronic Engineering, City University of Hong Kong, Kowloon Tong, Hong Kong*

Received 16 November 2005; accepted 10 January 2006

Available online 19 January 2006

Communicated by A.R. Bishop

Abstract

Based on Baptista's chaotic cryptosystem, we propose a secure and robust chaotic cryptographic scheme after investigating the problems found in this cryptosystem as well as its variants. In this proposed scheme, a subkey array generated from the key and the plaintext is adopted to enhance the security. Some methods are introduced to increase the efficiency. Theoretical analyses and numerical simulations indicate that the proposed scheme is secure and efficient for practical use.

PACS: 05.45.Ac

Keywords: Cryptosystem; Chaos; Security; Efficiency

1. Introduction

Chaotic systems possess a number of interesting and useful properties such as sensitivity to initial condition and system parameter, ergodicity and mixing (stretching and folding), which are analogous to the confusion and diffusion properties of a good cryptosystem.

In [1], Baptista proposed a chaotic cryptographic scheme governed by the logistic map:

$$F\colon\ x_{i+1} = bx_i(1 - x_i), \tag{1}$$

where $b \in [3.57, 4]$, $x_i \in (0, 1)$, and $i \in \{0, 1, 2, \ldots\}$.

The plaintext is assumed to be composed of a number of m-bit plaintext blocks. Accordingly, a portion of the attractor (usually the interval $[0.2, 0.8]$ as first selected by Baptista) is divided into $N\ (= 2^m)$ equal-width sites. Each site is associated with a plaintext block. Let $\varepsilon = (0.8\text{–}0.2)/N$ be the width of each site, S_i be the site in the range $[0.2 + i\varepsilon, 0.2 + (i + 1)\varepsilon]$ and a_i the corresponding plaintext block. The above association can be represented as follows:

$$\begin{aligned} H: S = {}&\{S_0, S_1, \ldots, S_i, \ldots, S_j, \ldots, S_{N-1}\} \\ &\to A = \{a_0, a_1, \ldots, a_i, \ldots, a_j, \ldots, a_{N-1}\}. \end{aligned} \tag{2}$$

Let $p_i \in A$ denote the ith plaintext block, x_{i-1} be the last chaotic state of Eq. (1) satisfying $x_{i-1} \in S_j$ and $H(S_j) = p_{i-1}$, $F^{c_i}(x_{i-1})$ represent the c_i times chaotic iteration of Eq. (1) departing from x_{i-1}. Then c_i is the corresponding ciphertext unit if it satisfies $F^{c_i}(x_{i-1}) = x_i$ where $x_i \in S_k$ and $H(S_k) = p_i$. It is an integer within the region $[200, 65535]$ as the minimum number of iterations is chosen as 200 and each ciphertext unit is represented by two bytes. For simplicity, in the remaining part of this manuscript, we will use $H(x_i) = p_i$, to represent $x_i \in S_k$ and $H(S_k) = p_i$.

The secret key of this scheme is the association map $H(S)$, the initial condition x_0 and the control parameter b of the logistic map. Since there always exist many possible choices for c_i, a probabilistic coefficient η is introduced in Baptista's scheme to generate different ciphertexts for a fixed plaintext. However,

[*] Corresponding author.
E-mail addresses: weijunfox@sina.com.cn (J. Wei), xfliao@cqu.edu.cn (X. Liao).

0375-9601/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.physleta.2006.01.030

we always set $\eta = 0$ in this Letter as the security is sufficiently high even without incorporating this step.

The chaotic cryptographic scheme proposed in [1] leads to an unbalanced distribution of ciphertext that favors attacks based on statistical knowledge, a modified version [2] is presented to improve it at the expense of efficiency. However, based on some useful information revealed by the ciphertext stream $\{c_i\}$, i.e., the stream of number of iterations, these two chaotic cryptographic schemes are easily cracked by the keystream attacks described in [3–5]. From then on, a variety of countermeasures against the keystream attack have been proposed [6–13]. After studying them, we find that they do not trace back the inherent cause that leads to the keystream attack, but are only proposed for resisting specific keystream attacks. As a result, it is doubtful whether the countermeasures can resist other attacks even resulted from the same cause.

In the remaining part of this Letter, we will first explore the inherent problem of Baptista's scheme, and then propose a more secure and efficient cryptographic scheme that is built on the rectification to Baptista's scheme.

2. Method for enhancing security

Suppose that $P = p_1p_2p_3\cdots$ is the plaintext. For a stream cipher algorithm, the keystream $t_1t_2t_3\cdots$ originating from the secret key is used to encrypt the plaintext according to the rule:

$$C = c_1c_2c_3\cdots = E_{t_1}(p_1)E_{t_2}(p_2)E_{t_3}(p_3)\cdots. \tag{3}$$

It is well known that as long as the keystream is independent of the plaintext in a stream cipher algorithm, the attacks presented in [14, p. 25] become very efficient in breaking the algorithm. Unfortunately, this is just the inherent cause that makes Baptista's scheme very vulnerable to the keystream attacks [3–5]. The orbits for encrypting different plaintexts have no difference if Eq. (1) is iterated from the same initial point x_0 with the same parameter b. Certainly, a natural and effective solution to this problem is to render the orbit dependent on the plaintext as much as possible. However, all the remedial modifications introduced in [6–13] are engaged in complicating the mapping procedures and it is doubtful whether the security loopholes rooted on the invariant chaotic orbit are completely remedied, especially when the schemes presented in [7,8] are broken by the attack in [15] with similar analyzing principle but more intricate procedures than those in [3–5].

All the attacks suggested in [3–5,15] focus on the ciphertext array. This is because the ciphertext array itself not only provides information about the ciphertext distribution, but also leaks the indexing information of each ciphertext unit. In fact, not only the attacks presented in [3–5,15], but also most of the current reasoning attacks or the so-called unpredicted attacks operate on the ciphertext array. For Baptista's scheme and its variants, the ciphertext array associates a certain key essentially for the correct decryption by the intended receiver. Obviously, if we can find a method that not only makes the keystream dependent on the plaintext, but also renders the association mentioned above as indirectly as possible, the security of this scheme will be enhanced substantially.

Based on the above analysis, here we propose a modified version of Baptista's scheme. It is based on updating the value of the control parameter b in Eq. (1). Suppose that the current chaotic state is x_i and its binary representation is $0.b_1b_2\cdots b_j\cdots$. By defining three variables whose binary representation is $x_l = b_1\cdots b_{15}$, $x_m = b_{16}\cdots b_{30}$, $x_h = b_{31}\cdots b_{45}$, respectively, we obtain two equations as well as a judgment:

$$f(x_i) = x_l \oplus x_m \oplus x_h, \tag{4}$$

$$g(x_i) = f(x_1) \oplus f(x_2) \oplus \cdots \oplus f(x_i), \tag{5}$$

$$f(x_i) < \beta \cup nop > 200, \tag{6}$$

where $\oplus$ denotes the bitwise exclusive-OR (XOR) operation, $\cup$ the logical OR operation, variable *nop* is the number of plaintext blocks that have been encrypted since the last update of b, $\beta \in [0, 32767]$ is a coefficient that determines the updating frequency of b. Eq. (4) states that the 15 bits of x_l will be XORed bit-by-bit with those of x_m, i.e., b_1 XOR b_{16}, b_2 XOR b_{17}, etc. The result will then be bitwise XORed with those of x_h to obtain $f(x_i)$. In Eq. (5), the 15-bit value $f(x_i)$ will be XORed bit-by-bit with the 15-bit value accumulated from the first block so as to obtain $g(x_i)$. In Eq. (6), the decimal value of the 15-bit $f(x_i)$ will be compared with β. If its value is smaller or there are already 200 plaintext blocks encrypted since the last update of b, Eq. (6) will then give a TRUE value to update the control parameter b.

The requirement of $nop > 200$ guarantees better dynamical properties of the chaotic system (1) when the cycle length of chaotic orbit is much smaller than 2^L (Assume that the finite precision is L (bits) and fixed-point arithmetic is adopted. For detail, readers may refer to [11].) In the final step of encrypting the ith plaintext block, if the output of Eq. (6) turns out to be true, the control parameter of Eq. (1) will be updated according to the following formula:

$$b_{\text{new}} = b_0 + (b_{\max} - b_0) \times g(x_i)/2^{15}, \tag{7}$$

where b_0 and $b_{\max}$ are the two boundary values of the interval $[b_0, b_{\max}]$ within which the logistic map behaves chaotically. As there exists a period-3 periodic window when the control parameter falls in the interval $[3.83, 3.86]$ of the logistic map, we suggest $b_0 = 3.88$ and $b_{\max} = 4$.

Since b_0 and $b_{\max}$ are fixed, Eq. (7) leads to the case that no matter what the key (b, x_0) and the plaintext are, the real control parameter of Eq. (1) belongs to the fixed set B:

$$B = \{b_{\text{new}} \mid b_0 + (b_{\max} - b_0)/2^{15} \times i,\ i = 0, 1, 2, \ldots, 32767\}.$$

However, the time and the number of occurrence that the elements of B appear in the course of encryption depend on the key (b, x_0) as well as the plaintext.

As periodic windows appear at finite probability in chaotic systems [16], the set B contains elements correspond to periodic windows and should be avoided. Examples are

3.90632507324219, 3.88607940673828,

3.96143920898438, etc.,