Encrypting Multimedia Data Using Modified Baptista's Chaos-Based Algorithm

Octaviana Datcu(✉), Radu Hobincu, Mihai Stanciu, and Radu Alexandru Badea

University "Politehnica" of Bucharest, Bucharest, Romania
od@elcom.pub.ro

Abstract. One of the easiest to implement, yet complex, symmetric key chaos-based ciphers is the one proposed by Baptista in 1998. It has attracted much interest from scholars, who underlined its deficiencies and proposed different methods to enhance it. The present paper proposes an additional step in the encryption procedure - a modulo two sum between the binary representations of Baptista's cryptograms and that of the value of the chaotic logistic map at that very iteration. This results in a flat distribution of the cryptograms. Thus, one of the major drawbacks of Baptista's cryptosystem, the exponential decay of the repartition of the cyphertext values, is surmounted. The original Baptista's algorithm is described, the proposed method is exemplified on a short message and its results are discussed when applied on multimedia files.

Keywords: Multimedia data encryption · Chaos cipher
Baptista-type algorithms

1 Introduction

Since Baptista's proposal of the symmetric key chaos-based cipher in [1] it has attracted much interest from scholars. The cryptosystem has been cryptanalyzed by researchers and several works proposed methods to enhance it. Some algorithms can be found in [2–7], where one can see that Baptista's cipher has also inspired hashing and compression schemes. An overview of existing chaos-based encryption techniques is presented in 2015, by [8].

To overcome one of the major draw-backs of Baptista's cryptosystem, the exponential decay of the repartition of the cyphertext values, the present work proposes an additional step in the encryption procedure. The binary representations of the cryptograms resulted from Baptista's method and those of the corresponding values of the logistic map are added modulo two without carry (bitwise XOR operation). This results in a flat distribution of the cryptograms.

Section 2 briefly describes Baptista's original algorithm. The main contribution of this paper is given in Sect. 3, where the proposed method is exemplified on a short message. Both Baptista's algorithm and the modified one are applied on multimedia files and their results are discussed and analyzed in Sect. 4. Section 5 concludes the paper, underlying the main advantage of the proposed encryption.

O. Fratu et al. (Eds.): FABULOUS 2017, LNICST 241, pp. 185–190, 2018.
https://doi.org/10.1007/978-3-319-92213-3_27

2 Baptista's Chaos-Based Algorithm

We briefly describe the original algorithm proposed by Baptista [1]:

- The domain of the simplest chaotic discrete-time system, the logistic map (1), is considered for randomness generation. The values of the chaotic map at iteration n, $X(n)$, are in (0,1) and bifurcation parameter b in (0,4].

$$X(n+1) = b \cdot X(n) \cdot [1 - X(n)], \tag{1}$$

Remark. The aperiodic behavior, sensitive to perturbations, the map exhibits for parameters $b > 3.57$ [9], enables a pseudo-random dynamic, described by some probability density functions as the one depicted on the right side in Fig. 1, where the range $[X_{min}; X_{max}] = [0.2; 0.8]$ was divided into 256 equal-length subintervals, and the probability that the values of the logistic map fall within was computed. In Fig. 1 (left), for $b < 3.57$, one can see that the logistic map takes values within four subintervals only.

- Considering the ASCII association, messages such as text, image or sound can be represented as integer numbers. For each of the 256 ASCII characters it is assigned a subinterval of the logistic map range, a site of $\varepsilon = (X_{max} - X_{min})/256$ length.
- The encryption key is chosen to be the parameter b and the initial value of the logistic map, X_1. The parameter b is chosen after an analysis of the distribution it engenders. Figure 1 depicts this repartition for $b = 3.56$ and $b = 3.92$. More results are given on the authors' website[1].
- The chaotic map is iterated 250 times to get it in the stationary regime.
- Starting from the 251-st value of the considered map, the site (subinterval) assigned to the first plain character, I_{m1}, is searched.
- When the amplitude of the logistic map is within I_{m1}, at iteration k_1, the cryptogram for m_1 is generated as the value of that iteration, $C(m_1) = k_1$.
- The logistic map is reinitialized: $X_1 = X(k_1)$ and the assigned site for the next plain character, I_{m2}, is searched, and so on, until the entire plain message is parsed.

We have enciphered a 216255 characters text using the key $(b; X_1) = (4; 0.223860125802667)$. The occurrence of each ASCII character within the plain text is given in Fig. 2. The bifurcation parameter b was chosen to be 4 because the known probability density function of the logistic map for this value [1]. The initial condition X_1 was randomly generated from the range (0,1) and truncated to 15 significant digits equivalent to 64-bits double precision floating point used in the implementation. All digits are relevant, given the high sensitivity of chaotic systems to initial conditions [1]. The length of the plaintext was chosen such that the quantity of encrypted data is sufficient to draw some conclusions regarding the efficiency of the algorithm.

The biased distribution of the cryptograms, as the one obtained in Fig. 2, exposes Baptista's cipher [1] to attacks as the one in [6]. Thus, we propose a method to obtain a uniform distribution for the enciphered versions of the plain messages, such that it

[1] http://ham.elcom.pub.ro/~od/cercetare/Fabulous_2017/.

offers no information about the time needed to reach each corresponding subinterval assigned to the characters to be encrypted.

3 Modified Algorithm

For Baptista's algorithm, described in Sect. 2, the iteration k at which the amplitude of the logistic map is within I_{mj}, is the cryptogram of the plain character m_j; the index $j = 1, \ldots, L$, with L the length of the plain message. Thus, $C = k$.

Prior to sending that value through the communication channel, one more step is implemented, in the modification of Baptista's cipher we propose in this paper. A modulo 2 sum without carry (bitwise XOR) between the value of the iteration k and that of the state of the logistic map, $X(k)$ is added.

For a short example, the plain message '*cipher*', Table 1 gives the ASCII codes, m, the limits of the ε-length subinterval assigned to each ASCII code, the cryptograms which would result from enciphering with Baptista's algorithm [1], C, and the corresponding values of the chaotic system (1) at those iterations, $X(C)$. The values of C are represented on 16 bits, as in Baptista's original algorithm. The proposed step is explained in Table 2, where the 64-bits double precision values $X(C)$ are represented in hexadecimal notation, resulting in four 16-bits words. Further, the results of the bitwise XOR between the cryptograms from Baptista's cipher and the 16-bits words corresponding to the value of the logistic map are shown in the last column of Table 2, in base 10. It can be observed that the most significant bits are not suitable for encryption, as they do not change.

Results are obtained for the key $(b, X_1) = (4, 0.223860125802667)$ and the above mentioned text in Fig. 3. An original and an encrypted image, along with their corresponding histograms are shown in Fig. 4, for the same secret key.

Table 1. Baptista's cipher: message 'cipher' and key $(b, X_1) = (4, 0.223860125802667)$

ASCII code, m	$I_m = [x_{min} + \varepsilon \cdot (m - 1); x_{min} + \varepsilon \cdot m)$	C	$X(C) \in I_m$
99	[0.42968750; 0.43203125)	1123	0.432005039722298
105	[0.44375000; 0.44609375)	1679	0.445796898468479
112	[0.46015625; 0.46250000)	769	0.460193739139050
104	[0.44140625; 0.44375000)	1150	0.441488974614396
101	[0.43437500; 0.43671875)	641	0.434819475135869
114	[0.46484375; 0.46718750)	968	0.465731114802751

Table 2. Proposed encryption: message 'cipher' and key $(b, X_1) = (4, 0.223860125802667)$

C	X(C)	typecast(X(C), 'uint16')				XOR(C, X(C))			
1123	0.432005039722298	1ffd	7754	a5f8	3fdb	7070	29495	41371	15288
1679	0.445796898468479	247a	b6e5	87ef	3fdc	8949	45162	33120	14675
769	0.460193739139050	4675	70db	73d0	3fdd	17780	29658	28881	15580
1150	0.441488974614396	db9a	f8e0	415a	3fdc	57316	64670	17700	15266
641	0.434819475135869	d621	1057	d415	3fdb	54432	4822	54932	15706
968	0.465731114802751	ad0e	e0b3	ce89	3fdd	44742	58235	52545	15381

For an answering machine sound, and the key $(b, X_1) = (4, 0.3)$, one can listen to the encrypted.wav file on authors' website. Also, to test the sensitivity of the proposed cipher to a slight change in the secret key, the encrypted sound was deciphered with $(b, X_1) = (4, 0.3 + 10^{-15})$.

4 Discussion and Analysis

Roughly speaking, the size of the key space is $2^{52} \cdot 2^{52} \approx 2^{108}$, because X_1 and b are represented as 64-bits double precision numbers, with 52-bits mantissa. Nevertheless, more thorough analysis, like the one the bifurcation diagrams in Fig. 5 show, is worth being performed to better approximate the length of the interval parameter b lies in for a good encryption. The distribution of the cryptograms is of great importance. When it is uniform it hides the redundancy of the plain image as it can be observed from the histograms in Figs. 3 and 4. Information entropy is the most important feature of randomness and it has been computed for original and encrypted images in Fig. 5,

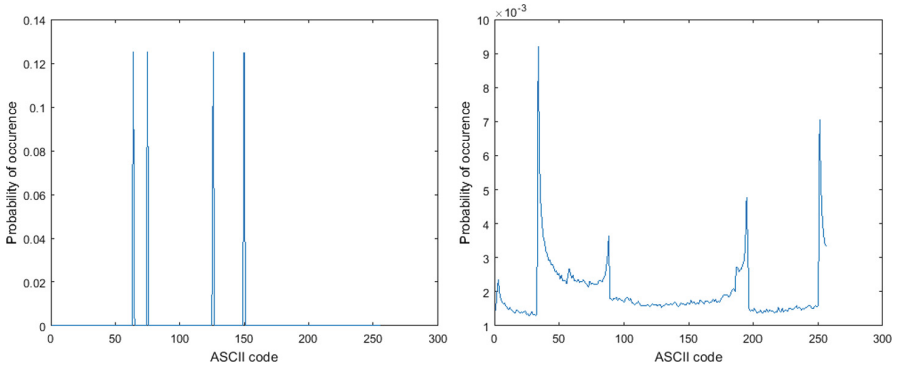

Fig. 1. Probability of occurrence of the cryptograms in the cyphertext for $b = 3.56$ *(left)* and $b = 3.92$ *(right)*; $X_1 = 0.2$ using Baptista's original algorithm.

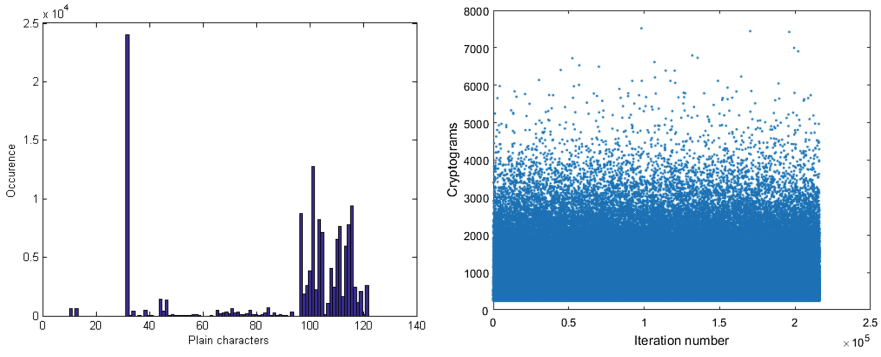

Fig. 2. Histogram for the plain characters and the distribution of their encrypted versions using the original Baptista's algorithm. The plain message is text.

obtaining the values 4.871 and, respectively 7.994bits. The entropy of an image representing noise on 256 gray levels is ideally 8 [10]. If the entropy of the ciphered grayscale image would be less than 8, there exists a certain level of predictability, which indicates a weakness of the cryptosystem. A more detailed analysis of the security level of the proposed cipher is given on the authors' website.

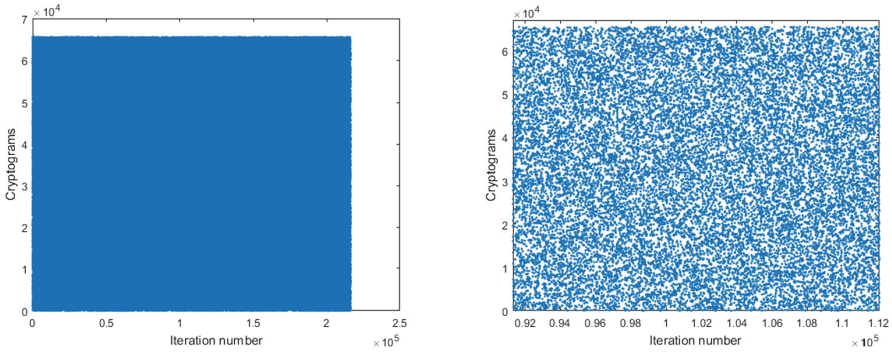

Fig. 3. The distribution of cryptograms using the proposed modified algorithm on text (left) and zoom (right). The secret key: $(b, X_1) = (4, 0.223860125802667)$.

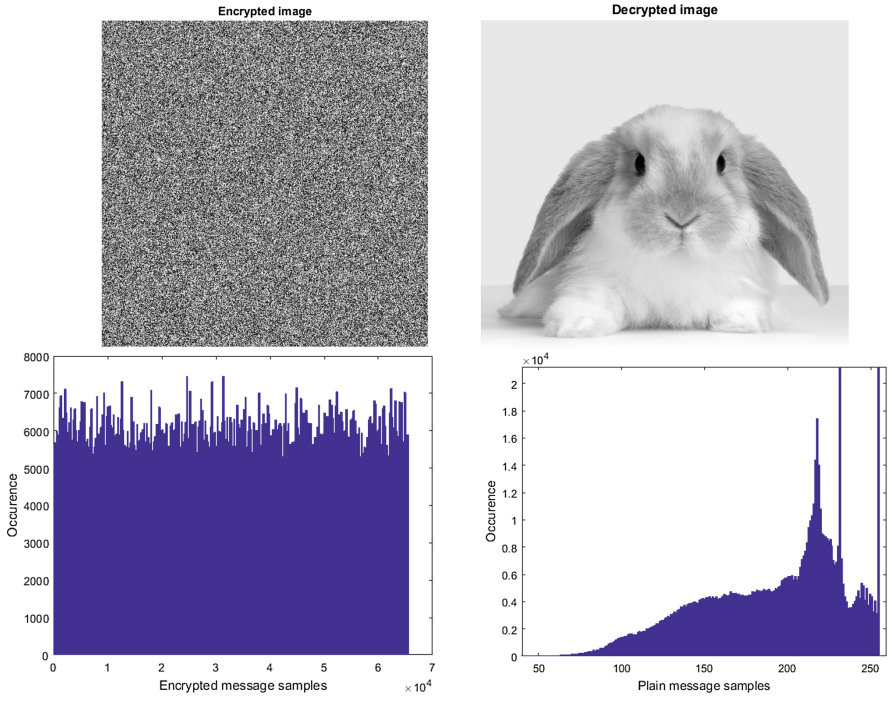

Fig. 4. The distribution of cryptograms with the modified algorithm, applied to an image identical to the decrypted one. The secret key: $(b, X_1) = (4, 0.223860125802667)$.

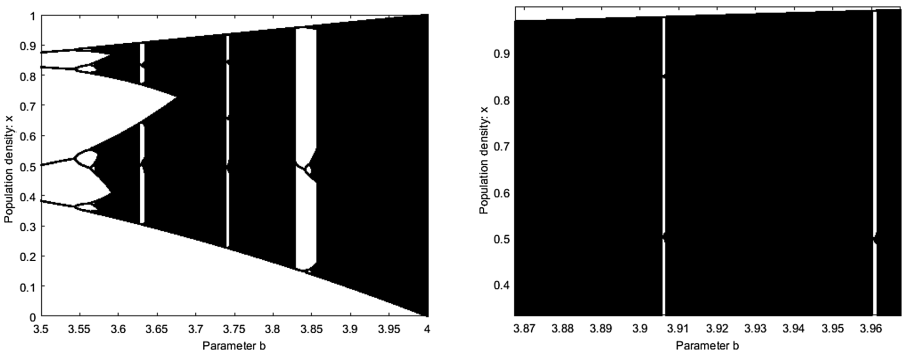

Fig. 5. Bifurcation diagram of the logistic map for parameter *b* in [3.5; 4] (left) and b in [3.87; 3.97] (right).

5 Conclusions

Aiming to obtain a flat distribution of the cryptograms output by one of the best-known chaos-based enciphering scheme, Baptista's algorithm, the present paper adds a step in the original method. Thus, a bitwise XOR is computed between the value of the 16-bits Baptista's cryptograms and the least significant 16-bits from the binary representation of the double precision floating point values of the corresponding logistic map amplitudes. While solving the exponential decay in the distribution of Baptista's cryptograms, the proposed method does not overcome the increase in the size of the cyphertext.

References

1. Baptista, M.S.: Cryptography with chaos. Phys. Lett. A **240**, 50–54 (1998)
2. Jokimoski, G., Kocarev, L.: Analysis of some recently proposed chaos-based encryption algorithm. Phys. Lett. A **291**, 381–384 (2001)
3. Li, S., Chen, G., Wong, K.W., Mou, X., Cai, Y.: Baptista-type chaotic cryptosystems: problems and countermeasures. Phys. Lett. A **332**, 5–6 (2004)
4. Nitharwal, B., Rani, M., Saini, H.C.: Improving security of the Baptista's cryptosystem using two-step logistic map. Int. J. Comput. Netw. Inf. Secur. **7**(5), 34 (2015)
5. Wong, K.W.: A combined chaotic cryptographic and hashing scheme. Phys. Lett. A **307**, 292–298 (2003)
6. Álvarez, G., Montoya, F., Romera, M., Pastor, G.: Cryptanalysis of dynamic look-up table based chaotic cryptosystems. Phys. Lett. A **326**, 211–218 (2004)
7. Chen, Y., Liao, X.: Cryptanalysis on a modified Baptista-type cryptosystem with chaotic masking algorithm. Phys. Lett. A **342**, 389–396 (2005)
8. Shukla, P.K., Khare, A., Rizvi, M.A., Stalin, S., Kumar, S.: Applied cryptography using chaos function for fast digital logic-based systems in ubiquitous computing. Entropy **17**(3), 1387–1410 (2015)
9. Eckmann, J.-P.: Roads to turbulence in dissipative dynamical systems. Rev. Mod. Phys. **53**, 643 (1981)
10. Gonzalez, R.C., Woods, R.E., Eddins, S.L.: Digital Image Processing Using MATLAB. Prentice Hall, New Jersey (2003). Chap. 11