International Journal of Modern Physics B
© World Scientific Publishing Company

CRYPTANALYSIS OF A NEW CHAOTIC CRYPTOSYSTEM BASED ON ERGODICITY*

DAVID ARROYO† GONZALO ALVAREZ

Instituto de Física Aplicada, Consejo Superior de Investigaciones Científicas, Serrano 144, 28006 Madrid, Spain

SHUJUN LI‡

FernUniversität in Hagen, Chair of Computer Engineering Universitätsstraße 27, 58084 Hagen, Germany

CHENGQING LI

Department of Electronic Engineering, City University of Hong Kong 83 Tat Chee Avenue, Kowloon Tong, Hong Kong SAR, China

VERONICA FERNANDEZ

Instituto de Física Aplicada, Consejo Superior de Investigaciones Científicas, Serrano 144, 28006 Madrid, Spain

Received 19 June 2008
Accepted 27 November 2008

This paper analyzes the security of a recent cryptosystem based on the ergodicity property of chaotic maps. It is shown how to obtain the secret key using a chosen-ciphertext attack. Some other design weaknesses are also shown.

Keywords: Chaotic encryption, ergodicity, logistic map, Gray ordering number, logistic map, chosen-ciphertext attack, cryptanalysis

1. Introduction

Chaotic maps possess an ergodic behavior which makes them suitable for the design of new cryptosystems. This is the case of the cryptosystem proposed in Ref. 1. This cryptosystem is based on the tent map and has been cryptanalyzed in Ref. 2 and later improved in Ref. 3. In Ref. 4 a new modification on the original scheme described in Ref. 1 was proposed. The authors of this new proposal claim that this

*This paper has been published in *International Journal of Modern Physics B*, vol. 23, no. 6, pp. 651-659, 2009.
†Corresponding author.
‡Current address: Fachbereich Informatik und Informationswissenschaft Universität Konstanz, Fach M697, Universitätsstraße 10, 78457 Konstanz, Germany

modification overcomes all the security problems that were emphasized in Ref. 2, 3. Nevertheless, in this paper we show that the ciphertext still includes enough information to enable a chosen-ciphertext attack based on symbolic dynamics. The rest of the paper is organized as follows. First of all, Sec. 2 gives a brief introduction to the cryptosystem under study. After that, in Sec. 3 the symbolic dynamics based chosen-ciphertext attack is explained. Then some other problems of the cryptosystem under study are discussed in Sec. 4, and finally the last section gives some final comments and conclusions.

2. Description of the cryptosystem

The cryptosystem described in Ref. 4 is based on the transformation of chaotic orbits into binary sequences. These chaotic orbits are generated using a one-dimensional chaotic map defined by

$$x_{n+1} = f(x_n, r), \tag{1}$$

where $f : \mathcal{I} \to \mathcal{I}$ and $0.5 \in \mathcal{I} \subset \mathbb{R}$. If Eq. (1) is iterated N times, then a chaotic orbit will be obtained as

$$\{x_n\}_{n=0}^{N} = \{x_0, x_1, \ldots, x_N\}. \tag{2}$$

The authors of Ref. 4 do not explicitly indicate if x_0 is also included in the chaotic orbit as the first chaotic state. Without loss of generality, in this paper we will assume that this was included.

Finally, the binary counterpart (i.e., the symbolic dynamics based representation) of the original chaotic orbit is given by

$$g_n = g_n(x_0, r) = \begin{cases} 0 \text{ , if } & x_n < 0.5, \\ 1 \text{ , if } & x_n \geq 0.5, \end{cases} \tag{3}$$

for $0 \leq n \leq N$. Henceforth, the binary sequence $\{g_n(x_0, r)\}_{n=0}^{N}$ is noted as $G^N(x_0, r)$ to emphasize its dependency with the initial condition and the control parameter.

The cryptosystem works as follows.

- Step 1) Initialize $i = 0, j = 0$.
- Step 2) For the i-th plain block P_i formed by $b_i = b$ bits, try to find the first b_i-bit segment of $\{g_n\}_{n=j}^{N_{max}+b_i}$ which is equal to P_i; in case a segment is not found, let $b_i = b_i - 1$ and repeat this step[a]. The parameter N_{max} indicates the maximum number of trials in the searching of P_i through the binary sequence.
- Step 3) Denoting by n_i the number of iterations needed to locate the distinguished b_i-bit segment from g_j, output (b_i, n_i) as the i-th cipher-block.

[a]Note that in Ref. 4, there was a typo about $b_i = b_i - 1$, which was published as "$b_i = b_i + 1$".

- Step 4) Set $i = i+1$ and $j = j+n_i+b_i$ [b], then go to Step 2 until the whole plaintext is exhausted.

The decryption process is simpler than the encryption one. In this case, the searching process becomes unnecessary. For the recovery of the i-th plain block, one simply iterates the chaotic map from the current status for $n_i + b_i$ times and record the last b_i chaotic states, which are then transformed into the i-th b_i-bit plain block according to Eq. (3).

In Ref. 4 it is claimed that the secret key of the cryptosystem is composed of the initial condition x_0 and the control parameter r. For a more detailed description of the encryption/decryption procedures, the reader is referred to Ref. 4.

3. Chosen-ciphertext attack

In Ref. 4 it is mentioned that most chaotic systems can be used to implement the above described cryptosystem. Moreover, the resistance of the cryptosystem against the attacks presented in Ref. 2 is assumed without any security analysis. However, this section proves that a wrong selection of the chaotic map allows an estimation of the secret key through a chosen-ciphertext attack.

Among all the possible options, the logistic map was chosen in Ref. 4 as the chaotic system to prove the reliability of the cryptosystem. The logistic map is defined as

$$x_{n+1} = f(x_n, r) = r \cdot x_n \cdot (1 - x_n), \tag{4}$$

for $r \in (3.57148, 4)$ and $x_n \in [0, 1]$. The function $f(x, r)$ for the logistic map is a concave function with only one critical point at 0.5. For this kind of maps the binary sequence referred in Eq. (3) can be interpreted as a Gray code Ref. 5, 6. Moreover, in Refs. 7, 8, 9 it is shown that the family of Gray codes generated using Eq. (3) can be assigned an order according to the initial condition and the control parameter. The existence of this order allows an estimation of the control parameter r and the initial condition x_0 just by analyzing the binary sequence $G^N(x_0, r)$ for a sufficiently large number N. Therefore, as long as one can reconstruct the sequence $G^N(x_0, r)$, one can estimate the secret key of the cryptosystem. This is used to build an attack with three different stages:

(1) Reconstruction of the Gray code derived from the logistic map.
(2) Estimation of the control parameter from the reconstructed Gray code.
(3) Estimation of the initial condition from the reconstructed Gray code and the estimated control parameter.

[b] In Ref. 4 it is not explicitly mentioned how to update the index j. In this paper, we assume that it is updated in such a way that no segment of a chaotic orbit will be reused for encryption of two continuous plain blocks.

3.1. *Reconstruction of the Gray code*

If one has access to the decryption machine, then one can perform a chosen-ciphertext attack (see Page 25 of Ref. 10) to reconstruct $G^N(x_0, r)$, i.e., the Gray code associated to the values of x_0 and r that make up the secret key of the cryptosystem under study. To do so, M ciphertexts are generated as $(b, b \cdot i)$ for $i = 0, 1, 2, \ldots, M$. As an example, let us assume that $x_0 = 0.5$ and $r = 3.78$. In this case, it is satisfied that

$$G^N(0.5, 3.78) = \{1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1 \ldots\}.$$

As a result, if we ask the decryption machine to decrypt $(8, 0)$, then we obtain $\{1, 1, 0, 1, 1, 0, 1, 1\}$. Similarly, the decryption machine will return $\{1, 0, 1, 1, 1, 1, 1, 0\}$ when the input is $(8, 8)$, and $\{1, 1, 1, 1, 0, 1, 1, 1\}$ when the input is $(8, 16)$. In other words, the decryption of the first ciphertext returns the first b bits of $G^N(x_0, r)$, the decryption of the second ciphertext gives the second set of b bits of $G^N(x_0, r)$, and so on.

3.2. *Estimation of the control parameter*

If the binary sequence (i.e., the Gray code) derived from the iteration of the logistic map is known, then it is possible to infer the value of r based on the concept of Gray Ordering Number (GON). The GON was introduced in Ref. 5 as a way to reinterpret the main results of Ref. 11 in a more intuitive way. The calculation of the GON of a binary sequence $G^N(x_0, r)$ involves two steps:

- The binary sequence is transformed into another binary sequence using the next equation:

$$u_i(x_0, r) = \begin{cases} g_i(x_0, r), & \text{if} \quad i = 0, \\ u_{i-1}(x_0, r) \oplus g_i(x_0, r), & \text{if} \quad i > 0, \end{cases} \tag{5}$$

  where $i = \{0, 1, 2, \ldots, N\}$.
- The GON of the original binary sequence is calculated as:

$$GON(G^N(x_0, r)) = 2^{-1} \cdot u_0 + 2^{-2} \cdot u_1 + \cdots + 2^{-N-1} \cdot u_N. \tag{6}$$

According to Ref. 8, for any concave unimodal map with critical point equal to 0.5 it is satisfied that

$$GON(G^N(f(x_0, r), r)) \leq GON(G^N(f(0.5, r), r)) \tag{7}$$

for any value of r in $[3, 4]$ and any value of x_0 in $[0, 1]$. Furthermore, the function $GON(G^N(f(0.5, r), r))$ is an increasing function with respect to r (see Fig. 3(a) of Ref. 8). These two facts are used in Ref. 8 to estimate the value of the control parameter r. First of all, the value $GON(G^n(f(0.5, r), r))$, for $n < N$, is approximated as the maximum value of the GON of M different shift-left sequences obtained from $G^N(x_0, r)$. Afterwards, the monotonic relationship between $GON(G^n(f(0.5, r), r))$ and r is used to obtain an estimation of r through a binary search procedure.

In order to test this algorithm, some simulations have been carried out. The parameter estimation errors for $r = 3.9197398122739102$ are shown in Fig. 1. Different values of x_0 and N were considered, for a fixed length of the subsequences of $n = 100$. Since this method is based on the approximation of the maximum of $GON(G^n(f(0.5, r), r))$ through M different values, it is expected that the exact value of r cannot be obtained unless the value 0.5 is part of the chaotic orbit from which the binary sequence was calculated. Moreover, the characteristic dependency of chaotic maps on the initial condition makes the parameter estimation error depend on the value of x_0, as shows in Fig. 1. Nevertheless, the proposed method allows to obtain an estimation of r which implies a considerable narrowing of the key space and which can be further improved through a trial and error strategy, i.e., a brute force attack on the value of the control parameter in a dramatically reduced key-space.

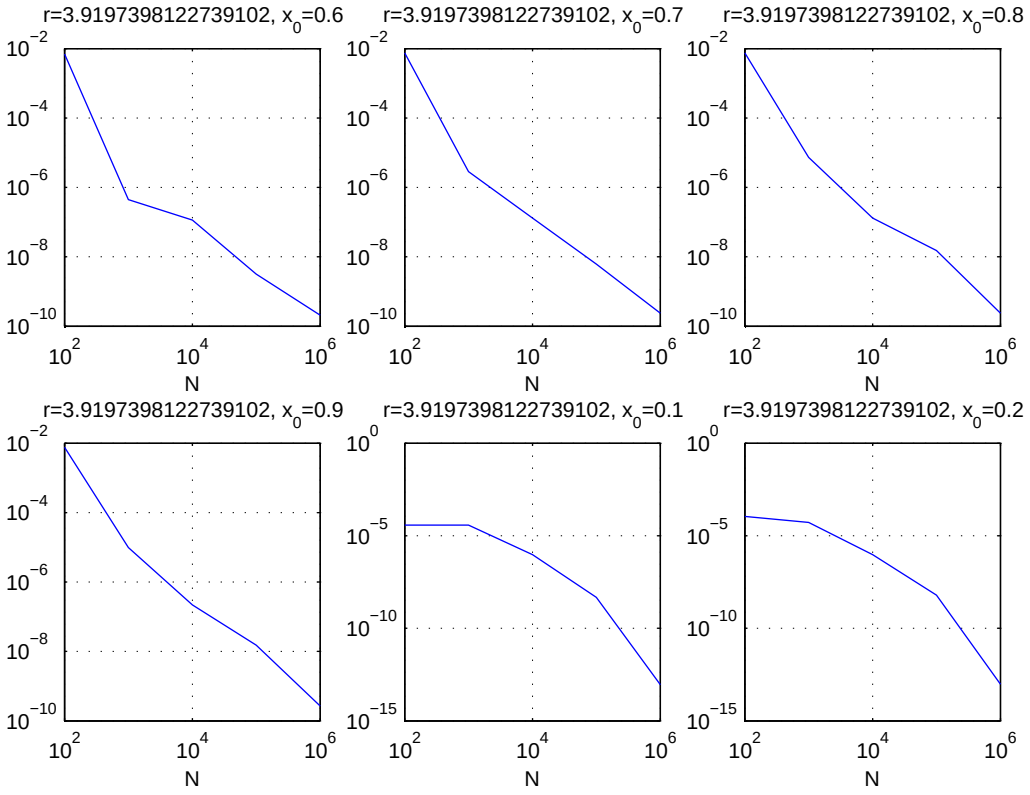

Fig. 1. Parameter estimation errors for $r = 3.9197398122739102$, different values of x_0 and N.

3.3. *Estimation of the initial condition*

In this subsection we will assume that we have obtained the exact value of r by using the algorithm discussed in the last subsection. Indeed, when considering the security of a cryptosystem, a partial knowledge of the key must not lead to the determination of the rest of the key (see Rule 7 in Ref. 12). Therefore, even if we were not able to estimate the value of r and obtain the exact value through a brute-force attack, the recovery of x_0 based on the knowledge of the other subkey r would represent a very important flaw of the cryptosystem under study.

As pointed out in Ref. 8, the GON of $G^N(x_0, r)$ is a monotonic increasing function with respect to x_0 (see Fig. 1 of Ref. 8). This means that one can obtain the value of x_0 through an iterative algorithm similar to that described in the last subsection. This algorithm was used to estimate the value of the initial condition from which $G^N(x_0, r)$ was generated. Different values of r, x_0 and N were considered. The results are shown in Fig. 2. For all analyzed situations, a number of bits greater than 80 implies an estimation error below 10^{-15}. Since all the simulations were performed using double precision, this means that the exact recovery of the initial condition is possible.

4. Other weaknesses

In this section some other problems of the cryptosystem under study are emphasized.

4.1. *Considerations about the chaotic system employed*

In Ref. 4 it is pointed out that most chaotic systems can be used to implement the proposed cryptosystem. However, there is no indication of the requirements that a chaotic system must fulfill to determine a secure cryptosystem according to the proposed encryption/decryption structure. Moreover, in the previous section we proved that at least a family of chaotic maps, i.e., the unimodal chaotic maps with fixed critical point equal to 0.5 cannot be used as long as a high level of security against chosen-ciphertext attack is needed. Furthermore, a different way should be used to generate the binary sequence for the encryption procedure. In the original design, this binary sequence is obtained by comparing each chaotic state included in a chaotic orbit with the fixed threshold value 0.5. Nevertheless, to ensure good statistical characteristics of the binary sequence, the threshold value should be selected according to the dynamics of the underlying chaotic system.

4.2. *Considerations about the chaotic orbit generation*

The characteristics of a cryptosystem should be precisely defined in order to facilitate its implementation (see Rule 1 in Ref. 12). During the encryption step of the cryptosystem under consideration, the plaintext is divided into a set of binary

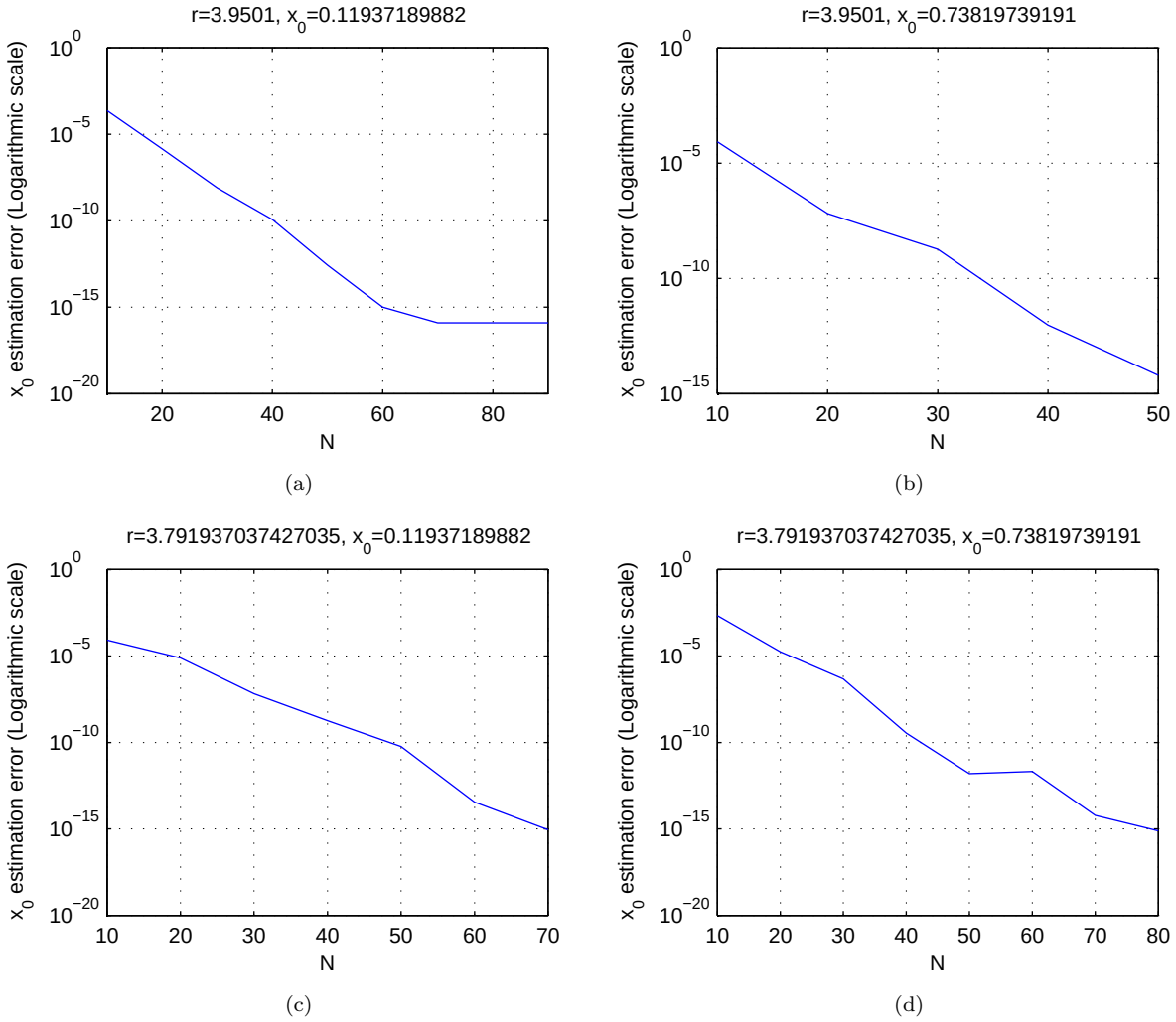

Fig. 2. Initial condition estimation errors for different values of r, x_0 and N.

sequences P_i which are successively located in the binary sequence $G^N(x_0, r)$. It is possible that P_i is not included in $G^N(x_0, r)$. In this case, the length of P_i is progressively decreased until it is found in $G^N(x_0, r)$. Nevertheless, there is no information about the length of $G^N(x_0, r)$, i.e., about the maximum number of iterations N_{max} needed to conclude whether the length of P_i must be decreased. Furthermore, not only the length of $G^N(x_0, r)$ is not explicitly established, but also some interpretation problems concerning the precise way of generating $G^N(x_0, r)$ can be found. Fist of all, in Ref. 4 it is not mentioned whether the first bit of $G^N(x_0, r)$ corresponds to x_0 or to x_1. On the other hand, once the plain block P_i has been encrypted, it is not clear whether the next binary sequence starts from $G^N(x_{n_i}, r)$, $G^N(x_{n_i+1}, r)$

or $G^N(x_{n_i+b}, r)$. Note that we fixed these problems in our description of the cryptosystem given in Sec. 2.

4.3. *Considerations about the key space*

The inadequacy of the logistic map for the implementation of this cryptosystem has been proved by means of a ciphertext attack. However, the selection of this map entails another important problem that suggests not to choose the logistic map as a base of any cryptosystem [13]. This problem concerns the definition of the key space. In Ref. 4 it is claimed that the value of the control parameter r should be selected within the interval $(3.57148...., 4)$ to exhibit a chaotic behavior. However, the existence of periodic windows in this region is well known (see Fig. 3) and so the selection of r should be performed in a more precise manner in order to avoid these (see Rule 5 in Ref. 12).

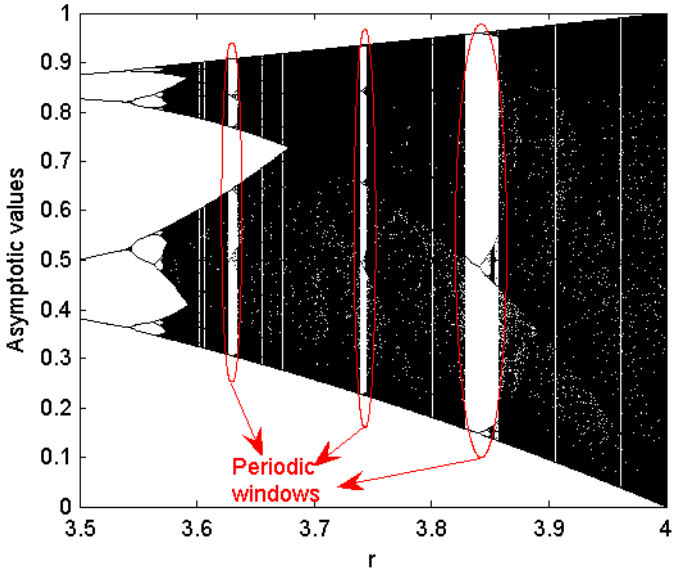

Fig. 3. Bifurcation diagram of the logistic map showing the existence of periodic windows.

5. Conclusions

Some weaknesses of the chaotic cryptosystem described in Ref. 4 have been discussed in this paper. A chosen-ciphertext attack has been described, which can recover the secret key of the cryptosystem by exploiting the theory of symbolic dynamics. Some other problems related to the design of the cryptosystem have also been pointed out. As a result, we recommend not to use this algorithm for secure applications.

Acknowledgments

The work described in this paper was supported by *Ministerio de Educación y Ciencia of Spain*, research grant SEG2004-02418, *Ministerio de Ciencia y Tecnología* of Spain, research grant TSI2007-62657 and *CDTI, Ministerio de Industria, Turismo y Comercio of Spain* in collaboration with Telefónica I+D, Project SEGUR@ with reference CENIT-2007 2004. Shujun Li was supported by a research fellowship from the *Alexander von Humboldt Foundation, Germany.*

References

1. E. Alvarez, A. Fernández, P. García, J. Jiménez, A. Marcano, New approach to chaotic encryption, Physic Letters A 263 (1999) 373–375.
2. G. Alvarez, F. Montoya, M. Romera, G. Pastor, Cryptanalysis of a chaotic encryption system, Physics Letters A 276 (2000) 191–196.
3. S. Li, X. Mou, Y. Cai, Improving security of a chaotic encryption approach, Physics Letters A 290 (3-4) (2001) 127–133.
4. X. Wang, C. Duan, N. Gu, A new chaotic cryptography based on ergodicity, International Journal of Modern Physics B 22 (7) (2008) 901–908.
5. G. Alvarez, M. Romera, G. Pastor, F. Montoya, Gray codes and 1D quadratic maps, Electronic Letters 34 (13) (1998) 1304–1306.
6. T. Cusick, Gray codes and the symbolic dynamics of quadratic maps, Electronic Letters 35 (6) (1999) 468–469.
7. G. Alvarez, F. Montoya, M. Romera, G. Pastor, Cryptanalysis of an ergodic chaotic cipher, Physics Letters A 311 (2003) 172–179.
8. X. Wu, H. Hu, B. Zhang, Parameter estimation only from the symbolic sequences generated by chaos system, Chaos, Solitons and Fractals 22 (2004) 359–366.
9. G. Alvarez, D. Arroyo, J. Nunez, Application of gray code to the cryptanalysis of chaotic cryptosystems, in: 3rd International IEEE Scientific Conference on Physics and Control (PhysCon'2007, 3rd - 7th, September 2007, Potsdam, Germany), IEEE IPACS, Potsdam, Germany, 2007.
URL `http://lib.physcon.ru/?item=1355`
10. D. Stinson, Cryptography: Theory and Practice, CRC Press, 1995.
11. N. Metropolis, M. Stein, P. Stein, On the limit sets for transformations on the unit interval, Journal of Combinatorial Theory, Series A 15 (1) (1973) 25–44.
12. G. Alvarez, S. Li, Some basic cryptographic requirements for chaos-based cryptosystems, International Journal of Bifurcation and Chaos 16 (8) (2006) 2129–2151.
13. D. Arroyo, G. Alvarez, V. Fernandez, On the inadequacy of the logistic map for cryptographic applications, arXiv:0805.4355 (2008).
URL `http://arxiv.org/abs/0805.4355`