Chaotic Modeling and Simulation (CMSIM) 2: 273–280, 2013

Analysis of FIPS 140-2 Test and Chaos-Based Pseudorandom Number Generator

Lequan Min, Tianyu Chen, and Hongyan Zang

Mathematics and Physics School, University of Science and Technology Beijing, Beijing 100083 China
(E-mails: minlequan@sina.com, zhy_lixiang@126.com, cty_furmosi@sina.com)

Abstract. Pseudo random numbers are used for various purposes. Pseudo random number generators (PRNGs) are useful tools to provide pseudo random numbers. The FIPS 140-2 test issued by the American National Institute of Standards and Technologyhas been widely used for the verifications the statistical properties of the randomness of the pseudo random numbers generated by PRNGs.

First this paper analyzes the FIPS 140-2 test. The results show that

- The required interval of the FIPS140-2 Monobit Test corresponds to the confident interval with significant level $\alpha = 0.0001(1 - \alpha)$.
- The required interval of the FIPS140-2 Pork Test corresponds to χ^2 test with significant level $\alpha = 0.0002(1$ - $\alpha)$.
- The required intervals of the FIPS140-2 Run Test correspond to the confident interval with significant level $\alpha = 0.00000016(1 - \alpha)$.

Second this study considers a novel chaotic map (NCM), whose prototype is the Lorenz three-dimensional Lorenz chaotic map. A NCP -based CPRNG is designed. Using the FIPS 140-2 test measures the 1000 keystreams randomly generated by the RC4 algorithm, and the 1000 keystreams generated by the CPRNG with perturbed randomly initial conditions in the range $|\epsilon| \in [10^{-16}, 10^{-4}]$. The results show that the statistical properties of the randomness of the sequences generated via the CPRNG and the RC4 do not have significant differences. The results confirm once again that suitable designed chaos-based PRNGs may generate sound random sequences, in particular for a replacement for the one-time pad system.
Keywords: FIPS 140-2 Test, Analysis in required intervals, Chaos-based pseudorandom number generator, RC4, Randomness comparison..

1 Introduction

Pseudorandom numbers are important in applications such as in simulations of physical systems[1], in cryptography[2], in Entertainment[3], and in protecting computer systems. John von Neumann was the first contributor in computer-based random number generators. Today algorithmic pseudorandom number

Received: 28 June 2012 / Accepted: 10 February 2013

ISSN 2241-0503

generators (PRNGs) have replaced almost random number tables and hardware random number generators in practical uses.

A algorithmic PRNG is an algorithm for generating sequences of numbers that approximate the properties of random numbers. A poor PRNG will lead to weak or guessable its keys, and leak the information which is prevented. There are many designed tests for measuring the randomness quantities of the sequences of numbers generated via PRNGs. The FIPS 140-2 test[4], the SP800-22 test[5], and the Diehard Battery test[6] are popular tests to be used in evaluating the randomness quantities of the sequence numbers deriving from PRNGs.

Since Lorenz's influential article[7] and Li and York's pioneer paper [8], the study of chaos has been rapidly developed. Matthews has first derived a chaotic encryption algorithm and shown that it may be suitable for a replacement for the one-time pad system[9].

G*á*mez-Guzm*á*n et al. have considered a modified Chua's circuit generator of 5-scroll chaotic attractor and shown that it may have a potential application to transmit encrypted audio and image information[11]. Stojanovski and Kocarev [10] have analyzed the application of a chaos-based PRNG. Li et al.[12] have reported that using only 120 consecutive known plain-byres can broken the whole secret key of a multiple one-dimensional chaotic map -based PRNG. Yu et al[13] have introduced and analyzed a quadric polynomial chaotic map based PRNG by the FIPS PUB 140-2 test.

This paper analyzes the standards of the randomness criteria of the FIPS 140-2 test, introduces a novel chaotic map (NCM), designs a NCM-based PRNG. Using the FIPS 140-2 test measures and compares the randomness performances of the NCM-based PRNG and the RC4 algorithm – a famous algorithm PRNG used in computer prevent.

The rest of this paper is organized as follows. Section 2 discusses the standards of the randomness criteria of the FIPS 140-2 test. Section 3 introduces the NCM, stimulates numerically its dynamic orbits, designed the NCM-based PRNG. Section 4 compares the randomness quantities of the NCM-based PRNG and the RC4 PRNG. Section 5 gives concluding remarks.

2 Analysis of FIPS 140-2 Test

The FIPS 140-2 Test issued by the National Institute of Standard and Technology consists of four tests: Monobit test, Pork test, Run test and Long Run test. Each test needs a single stream of 20,000 one and zero bits from keystream generation. Any failure in the test means the sequence of stream must be rejected. The four test are listed as for follows:

(1) Monobit test: Count the numbers N of "0" and "1" in the 20,000 bitstream, respectively. The test is passed if the N is fallen into the required interval given in the second column in Table 1.
(2) Poker test: Divide a sequence of 20,000 into 5,000 consecutive 4-bit segments. Denote $f(i)$ to be the number of each 4-bit valve i where $0 < i < 15$.

Then calculate the following:

$$N = \frac{16}{5000} \sum_{i=1}^{16} f(i)^2 - 5000. \quad (1)$$

The test is passed if the N is fallen into the required interval given in the second column in Table 1.

(3) Run test: Run is defined as maximal sequence of consecutive bits of either all $'1'$ or all $'0'$ that is the part of a 20,000 bitstream. Count and store the run bits with ≥ 1. The test is passed if the length of each run is fallen into the required interval listed in the second column in Table 1.

(4) Long Run test: The test is passed if there are no runs of length 26 or more.

Table 1. The required intervals of the FIPS 140-2 Monobit Test Pork Tests and Run Test, and the calculated confident intervals of random sequences with different significant level $\alpha's$. Here MT, PT, and RT represent the Monobit Test, the Pork Test and the Run Test; k represents the length of the run of a tested sequence.

	FIPS 140-2 Standard	$\alpha = 10^{-4}$	Golomb's
	Required Interval	Confident Interval	Postulates
MT	9,725∼10,275	9,725∼10,275	10000
		$\alpha = 2 \times 10^{-4}$	
PT	2.16∼46.17	2.41∼44.26	16.01
RT	FIPS 140-2 Standard	$\alpha = 1.6 \times 10^{-7}$	Golomb's
k	Required Interval	Confident Interval	Postulates
1	2,315∼2,685	2,315∼2,685	2,500
2	1,114∼1,386	1,119∼1,381	1,250
3	527∼723	532∼718	625
4	240∼384	247∼378	313
5	103∼209	110∼203	156
6+	103∼209	110∼203	156

Golomb has proposed three postulates on the randomness that pseudorandom sequences should satisfy [14]:

1. **Balance Property**. In one period of a pseudorandom sequence, If the period p is even, then the number of ones is equal to the number of zeros, otherwise they differ only by one.
2. **Run Distribution Property**. In one period of a pseudorandom sequence, the frequency of runs of length k is $\frac{1}{2^k}$. The numbers of the same length one run and zero run are the same.
3. **Ideal Autocorrelation Property**. The autocorrelation function $AC(k)$ has two values for a period. Explicitly:

$$AC(k) = \frac{1}{p} \sum_{i=1}^{p} s_i s_{i+k} = \begin{cases} 1 & for\ k = np \\ \frac{-1}{p} & otherwise \end{cases}$$

where 0's of the sequence are replace by 1's and 1's by -1's, $s_i s_j$ denote the multiplication of two bits s_i and s_j.

According to Golomb's postulates (1) and (2), the ideal values of the N's of the Monobit test and the Run test should be those listed in the 4th column in Table 1.

1. **Monobit test analysis**: Let $\epsilon = \epsilon_1 \epsilon_2 \cdots \epsilon_n$ be an one and zero bit sequence where n is the length of the bit string. Denote $X_i = 2\epsilon_i - 1$, then $S_n = X_1 + X_2 + \cdots X_n = 2(\epsilon_1 + \epsilon_2 + \cdots + \epsilon_n) - n$. If ϵ is a sequence of independent identically distributed Bernoulli random variables, then[5]

$$\frac{S_n}{\sqrt{n}} \sim N(0,1)$$

where $N(0,1)$ is a standard normal distribution.
The confident interval of $S'_n = \epsilon_1 + \epsilon_2 + \cdots \epsilon_n$ with significant level α is given by

$$\frac{n}{2} - \frac{\sqrt{n}}{2} Z_{\frac{\alpha}{2}} \leq S'_n \leq \frac{n}{2} + \frac{\sqrt{n}}{2} Z_{\frac{\alpha}{2}}$$

where $Z_{\frac{\alpha}{2}}$ (Matlab command $norminv(1-\alpha/2)$) is the inverse of the normal cumulative distribution function. In the case $n = 20,000$ and $\alpha = 0.0001$, the calculated result is given in the second column in Table 1 which is the same as the required interval given by the FIPS 140-2 test.

2. **Run test analysis**. Pick up the runs of length k from an one and zero bitstream and construct a new bit stream. Replace each one run of length k by 1, and zero run of length k by 0. Then we obtain an one and zero bit sequence $\epsilon' = \epsilon'_1 \epsilon'_2 \cdots \epsilon'_{n'}$ where n' is the length of the new bit string. Assume ϵ' is a sequence of independent identically distributed Bernoulli random variables, then similar to the analysis in the case of the Monobit test, we obtain

$$\frac{S_{n'}}{\sqrt{n'}} \sim N(0,1)$$

The confident interval of $S'_{n'} = \epsilon'_1 + \epsilon'_2 + \cdots \epsilon'_{n'}$ with significant level α is given by

$$\frac{n'}{2} - \frac{\sqrt{n'}}{2} Z_{\frac{\alpha}{2}} \leq S'_{n'} \leq \frac{n'}{2} + \frac{\sqrt{n'}}{2} Z_{\frac{\alpha}{2}}$$

For an ideal 20,000 one and zero bit pseudorandom stream, the length n' of a bit sequence ϵ' generated via the runs of length k should equal to $10000/2^k$. Let $\alpha = 1.6 \times 10^{-7}$, the calculated confident intervals are listed in the second column in Table 1 which are almost the same as the required intervals given by the FIPS 140-2 test.

3. **Poker test analysis**. Assume the the 4-bit segments are distributed independently and identically. Then the statistic quality

$$N = \frac{16}{5000}\sum_{i=1}^{16} f(i)^2 - 5000$$
$$= \sum_{i=1}^{16} \frac{5000}{1/16}(\frac{f(i)}{5000} - \frac{1}{16})^2$$

obeys χ^2 distribution. Hence the confident interval of the statistic quality of N with significant level α is given by

$$\chi^2_{1-\frac{\alpha}{2}}(15) \le N \le \chi^2_{\frac{\alpha}{2}}(15),$$

where $\chi^2_\alpha(15)$ (Matlab command chi2inv(α,15)) is the inverse of the χ^2 cumulative distribution function with free degree 15.

Let $\alpha = 0.0002$. The calculated confirmation interval is given in Table 1 which is similar to the one given by the FIPS 140-2 test.

3 New Chaotic Map and Pseudorandom Number Generator

we consider a novel chaotic map (NCM), whose prototype is the three-dimensional Lorenz chaotic map [15].

$$\begin{cases} X(n+1) = k_1X(n)Y(n) - k_2Z(n) - k_3X(n) \\ Y(n+1) = k_4X(n) - k_5Y(n) \\ Z(n+1) = k_6Y(n) - k_7Z(n) \end{cases}$$

where

$$k_1 = 1 - 10^{-6}, k_2 = 1 + 10^{-6}, k_3 = 2 \times 10^{-6},$$
$$k_4 = 1 + 10^{-6}, k_5 = 3 \times 10^{-6}, k_6 = 1 - 10^{-6}, k_7 = 10^{-6}.$$

The Lyapunov exponents of the NCM are $[\lambda_1, \lambda_2, \lambda_3] = [+0.0824, 0, -0.0824]$. If select an initial condition $[X_0, Y_0, Z_0] = [0.5\ 0.5\ -1]$, the numerical simulations of the orbits of the NCM display are given in Fig. 1. Observe that the dynamic patterns are similar to those of the 3D Lorenz map[15].

Let

$$K_n = \sqrt{3}X(n) + \sqrt{5}Y(n) + \sqrt{2}Z(n), n = 1, 2, \cdots, N;$$
$$Min(K) = \min_{1\le n\le N} K_n, Max(K) = \max_{1\le n\le N} K_n.$$

Define a transformation T by

$$T(K_n) = mod\left(round\left(\frac{255\sqrt{2}\times 10^5(K_n - Min(K))}{Max(K) - Min(K)}\right), 256\right), n = 1, 2, \cdots, N.$$

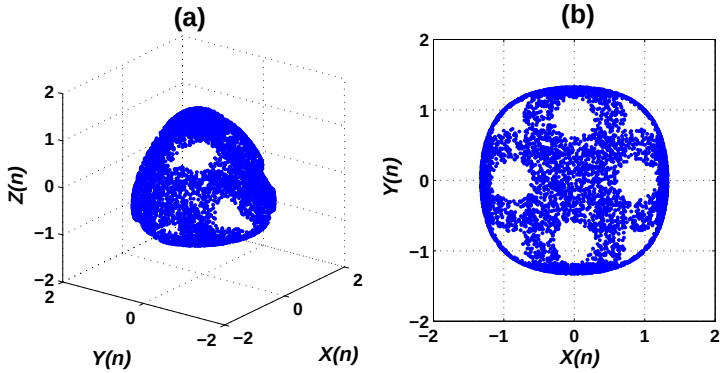

Fig. 1. Orbits of the first 5000 iterations: (a) $X(n), Y(n), Z(n)$, and (b) $X(n)$ and $Y(n)$.

Transferring $T(K_n)$ into binary codes, we obtain a binary sequence

$$s(k) = binary(T(K_n)), n = 1, 2, \cdots, N. \tag{2}$$

Hence, we construct a chaos-based pseudorandom number generator (CPNG).

4 FIPS 140-2 test

The RC4 was designed by Ron Rivest of RSA Security in 1987, and widely used in popular protocols such as Secure Sockets. Now we use the FIPS 140-2 test to test the 1000 keystreams randomly generated by the RC4, and the 1000 keystreams generated by the CPNG with an initial condition $[X(0), Y(0), Z(0)]$ = [0.5, 0.5, -1] perturbed randomly in the range $|\epsilon| \in [10^{-16}, 10^{-4}]$. The results are shown in Table 2. It follows that the statistical properties of the randomness of the sequences generated via the CPNG and the RC4 do not have significant differences.

Matlab commands for implement the RC4 algorithms are listed as follows.

```
L=8; K=randint(1,2^L,[0 2^L-1]);S=[0:2^L-1]; j=0;
for i=1:2^L
j=mod(j+S(i)+K(i),2^L);
Sk=S(j+1); S(j+1)=S(i); S(i)=Sk;
end
l=1; C=zeros(1,20000/8+10); j=0;i=0; k=1;
```

```
    for l=1:20000/8+10; i=mod(i+1,2^L); j=mod(j+S(i+1),2^L);
Sk=S(j+1); S(j+1)=S(i+1); S(i+1)=Sk;
C(k)=S(mod(S(j+1)+S(i+1),2^L)+1);
k=k+1;
end
```

Table 2. The confident intervals of the FIPS 140-2 tested values of 1000 key streams generated by the RC4 and the CPNG respectively. The significant level. $\alpha = 0.00001$

Test	bits	Golomb's	RC4	CPNG
item	$\{0,1\}$	Postulates	Confident Interval	Confident Interval
MT	0	10000	9992.2 ∼ 10012	9990.1 ∼ 10010
	1	10000	9988 ∼ 10008	9989.6 ∼ 10009
PT	–	16.01	14.408 ∼ 15.899	13.373 ∼ 13.914
LT	0	< 26	13.443 ∼ 13.971	13.405 ∼ 13.913
	1	< 26	13.340∼13.872	13.328∼ 13.823
LR			Run Test	
1	0	2500	2493.6 ∼ 2506.9	2492.0 ∼ 2504.9
	1	2500	2493.7 ∼ 2506.6	2489.9 ∼ 2503.3
2	0	1250	1244.9 ∼ 1253.8	1244.7∼ 1253.9
	1	1250	1242.6 ∼ 1251.3	1243.6∼ 1252.2
3	0	625	621.46 ∼ 628	622.10 ∼ 628.60
	1	625	622.44 ∼ 629.25	622.96 ∼ 629.31
4	0	313	310.09 ∼ 314.68	309.92∼ 314.56
	1	313	311.27 ∼ 315.74	310.29∼ 314.83
5	0	156	154.8 ∼ 158.21	154.18∼157.44
	1	156	154.79 ∼ 158.2	154.66∼ 158.14
6^+	0	156	154.29 ∼ 157.64	155.32∼ 158.56
	1	156	154.54 ∼ 157.93	155.28 ∼158.67

5 Concluding Remarks

Based on Golomb's postulates for the randomness of pure pseudorandom sequences, this paper analyzes the required intervals of the statistic quantities of three tests given in the FIPS 140-2. The results show that the required intervals for different tests do not have the same significant levels.

This study introduces a perturbed 3D Lorenze discrete map. The Lyapunov exponents and the dynamic orbits of the map are both similar to those of the 3D Lorenz map.

This paper constructs a chaos-based PRNG which has 7 key parameters. This feature of the PRNG may make it have large key space. Comparing the results of the FIPS 140-2 test for the RC4 PRNG and the chaos-based PRNG shows that statistical properties of the randomness of the sequences generated via the PRNG and the RC4 PRNG do not have significant differences.

The results confirm once again that suitable designed chaos-based PRNGs may generate sound random sequences, in particular for a replacement for the one-time pad system[9]. Further research along this line is promising.

Acknowledgements

L. Min would like to thank Professor Leon O. Chua at the UB Berkeley for directing him to study the fascinating chaos field. This work is jointly supported by the NNSF of China (Nos. 61074192, 61170037).

References

1.K. Binder, and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics: An Introduction* (4th edition). 2002. Springer.
2.N. Ferguson, B. Schneier, and T. Kohno, *Cryptography Engineering: Design Principles and Practical Applications*, 2010. Wiley Publishing.
3.Wegenkittl S. Gambling tests for pseudorandom number generator, *Mathematics and Computers in Simulation*, 55: 281-288, 2001.
4.NIST. *FIPS PUB 140-2, security requirements for cryptographic modules*. 2001.
5.R. Rukhin, J. Soto, J. Nechvatal et al., *A statistical test suite for random and pseudorandom number generator for cryptographic applications*, NIST Special Publication, 2001.
6.G. Marsaglia, http://www.stat.fsu.edu/pub/diehard/, 1996 [2012-03-30].
7.E. N. Lorenz, Deterministic nonperiodic flow, *J. of Atmospheric Sciences*, 20(2): 2130–148, 1963.
8.T. Y. Li and J. A. York, Period three implies chaos,*American Mathematical Monthly*, 82(10): 481–485, 1975.
9.R. A. J. Maatthews, On the derivation of a chaotic encryption algorithm, *Cryptologia*, XIII(1): 29–42, 1989.
10.T. Stojanovski and L. Kocarev01, Choas-based random number generators-part I: analysis, *IEEE Transaction on Circuitts and Systems-I: Fundamental Theory and Applications*, 48(3): 281–288, 2001.
11.L. Gámez-Guzmán, C. Cruz-Hernández, R.M. Lérrez, and E.E. Garacía-Guerrero, Synchronization of Chuas circuits with multi-scroll attractors Application to communication, *Commum Nonlinear Sci Numer Simulat*, 14: 2765–2775, 2009.
12.C. Li, S. Li, G. Alvarez et al., Cryptanalysis of a chaotic block cipher with external key and its improved version, *Chaos Solitons & Fractals*, 37: 299–307, 2008.
13.X. Yu, L. Min, and T. Chen. Chaos criterion on some quadric polynomial maps and design for chaotic pseudorandom number generator.*In Proc. of the 2011 Sewventh Int. Conf. on Nutural Computation*(26-28 July 2011, Shaghai, China), Vol.3: 1399–1402, 2011.
14.S. W. Golomb. *Shift Register Sequences*. Revised edition, CA: Aegean Park, 1982. Laguna Hills.
15.J. C. Sprott, *Chaos and Time-Sries Analysis*, page 427, Oxford. 2003. Oxford University Press.