Study on the Pass Rate of NIST SP800-22 Statistical Test Suite

Dong Lihua
National ISN Key Laboratory
Xidian University, Xi'an
Shanxi 710071, P.R.China
Lih_dong@mail.xidian.edu.cn

Zeng Yong
School of Computer Science and Technology
Xidian University, Xi'an
Shanxi 710071, P.R. China
yzeng@mail.xidian.edu.cn

Ji Ligang, Han Xucang
CEC Huada Electronic Design Co., Ltd
Beijing,100102, P.R.China

***Abstract*—NIST SP800-22 is a statistical test suite for determining whether given sequences are random or not for each statistical test. The statistical test suite has been used widely. However, it was not mentioned in the statistical test suite how many the ratio of passing all the 15 kinds of tests should be for the target generator to be regarded as the ideally true random number generator. In this paper, the ratio of passing all the 15 kinds of tests for the ideally true random number sequences is derived by the theoretical analysis. To verify the theoretical deduction, the statistical tests have been performed on several well known random number generators, such as AES, 3DES, SHA1, stream ciphers wined in eStream project, and some perfect pseudo-random number generator recommended by NIST. The result of the numerical simulations is accord with the theoretical analysis.**

Keywords- NIST SP800-22, statistical test, random number, randomnes

I. INTRODUCTION *(HEADING 1)*

Random number generators (RNGs) are one of basic cryptographic primitives, which are used in - but are not limited to - the generation of cryptographic keys, initialization vectors, challenges, nonces and padding values. For these applications, the quality of the generated random numbers is of crucial importance, which is checked by statistical tests. The basic statistical tests include NIST (National Institute of Standards and Technology) Special Publication 800-22[1], Diehard [2] and some statistical tests given by D.E.Knuth [3].

In this paper we consider the statistical test suite SP 800-22 which was firstly published in 2001 by the U.S. National Institute of Standard and Technology (NIST). However, some deviations on some of the statistical tests were founded [4-6]. After modification in 2010, SP 800-22 now consists of 15 p-value based tests that were developed to test the randomness of (arbitrarily long) binary sequences. These tests focus on a variety of different types of non-randomness that could exist in a sequence. Thus, it can not indicate the sequence is random just by pass one test. However, it was not mentioned in the statistical test suite how many the ratio of passing all the 15 kinds of tests should be for the target generator to be regarded as the ideally true random number generator. In 2010, the probability of passing all tests was analyzed by Okutomi et al [7] firstly. Then in [8], the authors augured that the analysis given in [7] was insufficient to obtain the probability of passing all tests and given an improvement. However, their improvement is derived under the assumption that there is no correlation among the 15 kinds of tests. So the probability of passing all tests that is obtained in [8] is so low that just about 50%. But the statistical tests that we performed on several perfect pseudo-random number generator recommended by NIST showed that the probability of passing all tests were all about 60%-70%. Thus, in this paper, a further analysis of the problem is made and the numerical simulations are given.

II. NIST SP800-22

NIST SP800-22 is a widely used statistical test suite. For every test, the basic testing process is the same and the data are analyzed using P-value- test firstly [9]. Thereafter, the distribution of P-values is examined to ensure uniformity.

The null hypothesis under P-value-test is that the sequence being tested is random. Associated with this null hypothesis is the alternative hypothesis (H1), which is that the sequence is not random.

For every test, the basic testing process is as follows.

- Firstly, a relevant randomness statistic must be chosen and a theoretical reference distribution of this statistic under the null hypothesis is determined by mathematical methods.
- Secondly, partition the sequence of length n to be tested into N subsequences and discard any unused bits.
- Then for every subsequence, a test statistic value is computed and this test statistic value is compared to the significance level (α). If the test statistic value exceeds the significance level (α),

978-1-4799-7434-4/14 $31.00 © 2014 IEEE
DOI 10.1109/CIS.2014.120

the null hypothesis for randomness is rejected. Otherwise, the null hypothesis (the randomness hypothesis) is not rejected (i.e., the hypothesis is accepted).

After the test, we have obtained N P-value. That is, for every sequence of length n to be tested, we can make N decisions on every test. According to the properties of the normal distribution, the range of acceptable proportions can be determined using 3□ rules. If the proportion falls outside of this interval, then there is evidence that the data is non-random. Otherwise if the sequence passes the P-value- test and the obtained P-values are uniform also, then the sequences are random, otherwise not random.

III. PASSING RATE OF TESTS OF NIST SP800-22

In NIST SP800-22, some of the statistical tests have several sub-tests. And the total number of sub-tests, that is the actual number of tests, is 188. From the definition of the significance level (α) in the P-value- test, we know that the probability that the test will indicate that the tested sequence is not random when it really is random is 0.01. That is, the probability that one ideally true random sequence passed the P-value- test is 1-0.01=0.99.

For every subtest of the NIST SP800-22, when there are more than 1052 subsequences passed the P-value- test, we think that the sequence tested passed the test. By the 3□ rules, the probability that there are more than 1052 subsequences passed the test in 1073 subsequences is 0.997.

In the following, we denote Ai as the event that all the 1073 ideally true random number subsequences can pass the i'th subtest. When there is no correlation among these subtests, the probability that all the 1073 ideally true random number subsequences can pass all the 188 subtests is

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2)\cdots P(A_n)$$
$$=0.997^{188}=56.84\%$$

But correlations have been found among some of the subtests [10-12]. At this time, the probability can be computed as

$$P(A_1A_2\cdots A_n)=P(A_n\mid A_1A_2\cdots A_{n-1})P(A_{n-1}\mid A_1A_2\cdots A_{n-2})\cdots P(A_2\mid A_1)P(A_1)$$

When some subtest, such as i'th subtest, has some positive correlation with some other subtests, we have

$$P(A_i)<P(A_i\mid A_1A_2\cdots A_{i-1})<1;$$

When some subtest, such as j'th subtest, has no positive correlation with other subtests, we have

$$P(A_j\mid A_1A_2\cdots A_{j-1})=P(A_j)$$

Thereafter when there are k subtests which have no positive correlations with the subtests before it, then the probability that all the 1073 ideally true random number subsequences can pass these subtests is

$$P(A_1A_2\cdots A_n)=P(A_{j_1})P(A_{j_2})\cdots P(A_{j_k})$$

We have performed practical correlation-test, the test process is that

- Extract the P-values for every subtest (188 in all) , and save them in a TXT File which named 'Freq.txt', 'BlockFreq.txt', 'NonOverLap_data1.txt' and so on;
- Compute the correlation between any two subtests according to the correlation definition

$$c=\frac{E(A-E(A))\cdot E(B-E(B))}{\sqrt{Var(A)\cdot Var(B)}}$$

By analyzing the result of the correlation-test, there are about 70 subtests whose correlation coefficients with others are less than 0.1. So we have

$$0.997^{70}=81\%.$$

For the uniform, the pass rate is determined by its significance level $\alpha_T=0.0001$, that is, $1-\alpha_T=0.9999$.

Thus, the upper bound that all the 1073 ideally true random number subsequences can pass all the subtests of NIST SP800-22 is

$$0.81\times0.9999=0.8099.$$

IV. SIMULATION OF THE NIST SP800-22

We have performed the statistical tests on several stream ciphers wined in eStream[13], some perfect pseudo-random number generator recommended by NIST, and some well known algorithms, such as AES, 3DES, SHA1. In our numerical simulation, the contents of the 15 tests and tests parameters are shown in Tables 1 and 2.

TABLE I. CONTENTS OF THE NIST SP800-22 STATISTICAL TEST

No	Test Name
1	The Frequency (Monobit) Test
2	Frequency Test within a Block
3	The Serial Test
4	The Runs Test
5	Tests for the Longest-Run-of-Ones in a Block
6	The Binary Matrix Rank Test
7	The Discrete Fourier Transform (Spectral) Test
8	The Non-overlapping Template Matching Test
9	The Overlapping Template Matching Test
10	Maurer's \Universal Statistical" Test
11	The Cumulative Sums (Cusums) Test
12	The Approximate Entropy Test
13	The Linear Complexity Test
14	The Random Excursions Test
15	The Random Excursions Variant Test

TABLE II. PARAMETERS USED IN OUR SIMULATIONS

No	Term
1	Sequence Length
2	Sample Size
3	Level of significance

The results of numerical simulation are shown in Table 3, which showed that the probability of passing all the 188 subtests is all about 60%-70%. The obtained pass rates were close to the theoretical analysis.

TABLE III. PASS RATE OF SOME KNOWN PSEUDO-RANDOM NUMBER GENERATOR

pseudo-random number generator	Pass rate	Note
HMAC_DRBG	61.7%	SP800-90A NIST recommended algorithm
Hash_DRBG	71%	SP800-90A NIST recommended algorithm
CTR_DRBG	61%	SP800-90A NIST recommended algorithm
CTR_TDES	67.5%	ANSI X9.31 recommended algorithm
AES	61%	
grain	62%	Wined algorithm in eStream
SHA1	59%	
3DES	61%	
Hc-128	67%	Wined algorithm in eStream
Rabbit	60%	Wined algorithm in eStream
Trivium	66%	Wined algorithm in eStream
mickey	69%	Wined algorithm in eStream

V. CONCLUSIONS

The upper bound in theory that an ideally true random sequence passes all the 188 subtests of the NIST SP800-22 is 80.99%. And the actual pass rates of the numerical simulation are about 60%-70%. Thus the result of the actual numerical simulation closes to the upper bound of the theoretical deviation, which confirmed the correctness of the theoretical deviation.

ACKNOWLEDGMENT

Project supported by the the Young Scientists Fund of the National Natural Science Foundation of China (Grant No. 61100235) ; the 111 Project (No. B08038) ; the Fundamental Research Funds for the Central Universities

REFERENCES

[1] Andrew Rukhin, Juan Soto, et al. A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications, NIST Special Publication 800-22 Revision 1, 2008.

[2] Marsaglia G. The Marsaglia Random Number CDROMincluding the Diehard Battery of Tests of Randomness. (1995). http : //stat.fsu.edu/geo/ diehard.html

[3] Knuth D E. The Art of Computer Programming, Volume2: Semi-numerical algorithms : Addison-Wesley, 1981. http://product.china-pub.com/7471

[4] Song-Ju Kim, Ken Umeno, and Akio Hasegawa. Corrections of the NIST Statistical Test Suite for Randomness, arXiv preprint nlin/0401040, 2004-arxiv. org.

[5] SHI Hongsong, ZHANG Chongbin, YANG Yongsheng and GAO Jinping. On the randomness test and its incompleteness, Journal of Tsing hua University (Science & Technology), 2011, 51(10), 1269-1273.

[6] Kenji Hamano, Toshinobu Kaneko. Correction of Overlapping Template Matching Test Included in NIST Randomness Test Suite, IEICE transactions on Fundamentals of Electronics, Communications and Computer Sciences, E90-A(9), 1788-1792.

[7] H. Okutomi and K. Nakamura. A study on rational judgment method of randomness property using NIST randomness test (NIST SP. 800-22) (in Japanese), IEICE Trans. A, J93-A(2010), 11-22.

[8] Akihiro Yamaguchi, Takaaki Seo and Keisuke Yoshikawa. On the pass rate of NIST statistical test suite for randomness, Japan Society for Industrial and Applied Mathematics Letters, 2010, Vol.2, pp.123-126.

[9] Sheng zhou, Xie shiqian, Pan chengyi. Probability and Mathematical Statistics (the 4th edition), Higher Education Press, 2010.10.

[10] Fan Limin, Feng Denggu and Chen Hua. On the Relativity of Binary Derivation and Autocorrelation Randomness Test, Journal of Computer Research and Development, 2009, 46(6), 956-961.

[11] Fan Limin, Feng Denggu and Chen Hua. Study on the Correlation Between Randomness Tests Based on Entropy, Journal of Software, 2009, 20(7), 1967-1976.

[12] HUANG Jia-lin, LAI Xue-jia. Eliminating Ability and Correlation of Random Statistical Tests, Information Security and Communications Privacy, 2009, 10, 43-46.

[13] http://www.ecrypt.eu.org/stream/endofphase3.html