IEICE TRANS. FUNDAMENTALS, VOL.E90–A, NO.9 SEPTEMBER 2007

PAPER Special Section on Information Theory and Its Applications

Correction of Overlapping Template Matching Test Included in NIST Randomness Test Suite*

Kenji HAMANO[†a)], ***Member and*** **Toshinobu KANEKO**[††b)], ***Fellow***

SUMMARY Accurate values for occurrence probabilities of the template used in the overlapping template matching test included in NIST randomness test suite (NIST SP800-22) have been analyzed. The inaccurate values used in the NIST randomness test suite cause significant difference of pass rate. When the inaccurate values are used and significance level is set to 1%, the experimental mean value of pass rate, which is calculated by use of random number sequences taken from DES (Data Encryption Standard), is about 98.8%. In contrast, our new values derived from a set of recurrence formulas for the NIST randomness test suite give an empirical distribution of pass rate that meets the theoretical binomial distribution. Here, the experimental mean value of pass rate is about 99%, which corresponds to the significance level 1%.

key words: *overlapping template matching test, NIST randomness test suite, SP800-22, pass rate*

1. Introduction

Random number is used in many different cryptographic systems and security of the system depends largely on randomness of the random numbers. Among statistical test suites for randomness, NIST randomness test suite [8], which originally consists of sixteen tests, is widely used in cryptographic field. Several studies on reliability of the NIST randomness test suite have reported that the Discrete Fourier Transform (Spectral) test and the Lempel-Ziv Compression test included in the NIST randomness test suite were needed to be corrected [1], [2], [5]. In addition, it was found that input size recommendations of the Approximate Entropy test in the NIST randomness test suite should be modified [4]. NIST updated some values used in the Discrete Fourier Transform test and removed the Lempel-Ziv Compression test from the NIST randomness test suite in 2004. Therefore, the NIST randomness test suite presently consists of fifteen tests. Recently, Okutomi et al. [7] evaluated randomness of random number sequences taken from DES and SHA-1, which are well-known for good pseudo-random number generators, using the NIST randomness test suite closely. They observed that each empirical distribution of pass rate of the overlapping template matching test and the Maurer's "Universal Statistical" test did not agree with the theoretical binomial distribution function whether DES or SHA-1, however, they have not clarified the reason of this phenomenon.

Takeda et al. [9] examined two template matching tests, namely, the overlapping template matching test and the non-overlapping template matching test in the NIST randomness test suite. They claimed that both template matching tests had a problem that the mean and variance of the occurrence probability of each template were dependent on its pattern. Their claim against the non-overlapping template matching test was a mistake, however, their idea to count runs of ones in [9] was helpful.

To make the NIST randomness test suite reliable, theoretical studies on the cause of the phenomenon found in [7] are required. This work was motivated by [7] and carried out to determine the cause of the phenomenon of the overlapping template matching test and to correct the inaccurate values used in the NIST randomness test suite. Accurate values for the NIST randomness test suite can be derived using a set of recurrence formulas in [9] and we show that the phenomenon of the overlapping template matching test is caused by use of the inaccurate values described in the NIST randomness test suite. In addition, we demonstrate that use of the inaccurate values described in the NIST randomness test suite actually decreases the reliability of results of the NIST randomness test suite.

2. Overlapping Template Matching Test

In this section, the overlapping template matching test included in the NIST randomness test suite is explained.

A random number sequence is partitioned into independent substrings of length M, and the number of occurrences of a template B, which is an m-bit run of ones, in each of the substrings is calculated and then the chi-square statistic is calculated on the basis of the number of occurrences of B and occurrence probabilities π_i of B. If P-value of the chi-square statistic is less than the significance level, the test concludes that the sequence tested appears to be non-random. Otherwise, the test concludes that the sequence tested appears to be random. Pass rate is defined by the proportion of sequences that pass the test.

Manuscript received December 18, 2006.
Manuscript revised April 16, 2007.
Final manuscript received May 29, 2007.

[†]The author is with the Department of Complexity Science and Engineering, the University of Tokyo, Kashiwa-shi, 277-8561 Japan.
[††]The author is with the Department of Electrical Engineering, Tokyo University of Science, Noda-shi, 278-8510 Japan.

*This paper was presented at the 2006 International Symposium on Information Theory and its Applications (ISITA2006), Seoul, South Korea, 29 October - 1 November 2006.

a) E-mail: hamano@it.k.u-tokyo.ac.jp
b) E-mail: kaneko@ee.noda.tus.ac.jp

DOI: 10.1093/ietfec/e90–a.9.1788

3. Accurate Computation of Occurrence Probability of Template

Let B, n and π_i $(0 \le i \le 4)$ be an m-bit run of ones, the length of a random number sequence and the probability of i occurrences of B, respectively. π_5 is the probability of five or more occurrences of B in the sequence. Let $T_i(n)$ denote the number of n-bit sequences with i occurrences of B. Values for π_i $(0 \le i \le 5)$ are calculated using $T_i(n)$ $(0 \le i \le 5)$ as follows:

$$\pi_i = \frac{T_i(n)}{2^n} \text{ (where } 0 \le i \le 4), \ \pi_5 = 1 - \sum_{i=0}^{4} \pi_i. \tag{1}$$

Takeda et al. [9] has proposed a set of recurrence formulas for $T_i(n)$, in which a seemingly editorial mistake has been found.

A set of correct recurrence formulas can be derived as follows:

$$T_0(n) = \begin{cases} 1, n = -1, \\ 1, n = 0, \\ 2T_0(n-1), 1 \le n \le m-1, \\ 2T_0(n-1) - T_0(n-m-1), n \ge m. \end{cases} \tag{2}$$

$$T_1(n) = \begin{cases} 0, \ n \le m-1, \\ 1, \ n = m, \\ 2, \ n = m+1, \\ \sum_{j=-1}^{n-m-1} T_0(j)T_0(n-m-2-j), \ n \ge m+2. \end{cases} \tag{3}$$

$$T_\alpha(n) = \sum_{j=1}^{n-(m+\alpha-2)} T_0(j-2)T_0(n-(j+m+\alpha-1)) + \sum_{k=1}^{\alpha-1} \sum_{j=1}^{n-2m-\alpha+2} T_0(j-2)T_k(n-(j+m+\alpha-k-1)), \tag{4}$$

where $\alpha \ge 2$.

We simplified the recurrence formula for $T_\alpha(n)$ as follows:

$$T_\alpha(n) = T_{\alpha-1}(n-1) + \sum_{j=-1}^{n-2m-\alpha} T_0(j)T_{\alpha-1}(n-m-2-j). \tag{5}$$

Here, the summention in Eq. (5) is defined by zero when $n-2m-\alpha < -1$. Derivation of the set of recurrence formulas is shown in Appendix A.

Equation (5) can be derived from Eq. (4). Replacing n, α of Eq. (4) with $n-1, \alpha-1$ respectively, we obtain the following equation:

$T_{\alpha-1}(n-1)$
= [First term of R.H.S. of Eq. (4)]
+ [Second term of R.H.S. of Eq. (4) except when $k = \alpha - 1$].

Table 1 Values for $T_i(n)$ when $m = 5$.

n	1	2	3	4	5	6	7	8	9
$T_0(n)$	2	4	8	16	31	61	120	236	464
$T_1(n)$	0	0	0	0	1	2	5	12	28
$T_2(n)$	0	0	0	0	0	1	2	5	12
$T_3(n)$	0	0	0	0	0	0	1	2	5
$T_4(n)$	0	0	0	0	0	0	0	1	2
$T_5(n)$	0	0	0	0	0	0	0	0	1
Total	2	4	8	16	32	64	128	256	512

Table 2 Values for π_i.

	Inaccurate values used in the NIST randomness test suite	Our new accurate values
π_0	0.367879	0.364091
π_1	0.183940	0.185659
π_2	0.137955	0.139381
π_3	0.099634	0.100571
π_4	0.069935	0.0704323
π_5	0.140657	0.139865

Therefore,

$T_\alpha(n) - T_{\alpha-1}(n-1)$
= [Second term of R.H.S. of Eq. (4) when $k = \alpha - 1$]
$= \sum_{j=-1}^{n-2m-\alpha} T_0(j)T_{\alpha-1}(n-m-2-j).$

Specifically, the set of recurrence formulas gives the values for $T_i(n)$ shown in Table 1.

Procedure of computation of $T_i(n)$ is as follows:

1. Compute $T_0(j)$ $(j \le n)$ using Eq. (2).
2. Compute $T_1(j)$ $(j \le n)$ using Eq. (4) and the values for T_0.
3. Compute $T_\alpha(j)$ $(j \le n)$ using Eq. (5) and the values for $T_0, T_{\alpha-1} (2 \le \alpha \le 4)$.
4. Compute $T_5(n)$ using Eq. (1) and the values for T_i $(0 \le i \le 4)$.

Time complexity of the procedure is $O(n^2)$.

Since NIST recommends $m = 9$ and M has been set to 1032 in the code of the NIST randomness test suite, values for π_i when $n = 1032$ and $m = 9$ are needed. The above set of recurrence formulas gives the accurate values for π_i when $n = 1032$ and $m = 9$, which are listed in Table 2 compared with the inaccurate values used in the NIST randomness test suite. A Mathematica program for computing the accurate values is shown in Appendix B. Since the values described in the NIST randomness test suite are found from formulas that require an asymptotic approximation [8], the values used in the NIST randomness test suite are inaccurate.

4. Kolmogorov-Smirnov Test

In this section, Kolmogorov-Smirnov (KS) test [6], which is used for measuring the goodness of fit, is explained. The test statistic K_n^+ of the KS test is defined by the square

root of sample size n multiplied by maximum difference between two distributions investigated. The theoretical distribution function of K_n^+ is closely approximated by $1 - \exp(-2x^2)$, $(x \geq 0)$ for a large value of n like $n = 1000$.

5. Experiment and Results

Random number sequences taken from DES (Data Encryption Standard) were tested using the overlapping template matching test. The test was carried out with the conditions shown in Table 3, which were the same as the previous experiment [7]. Figure 1 shows the empirical distribution of pass rate, which was calculated on the basis of 10^3 iterations of experiments, and the theoretical binomial distribution function when inaccurate values for π_i described in the NIST randomness test suite were used. Significant difference between the empirical distribution and the theoretical binomial distribution function was found. The observation was consistent with [7]. The two distributions had the different mean value of pass rate. That is, from Fig. 1, the experimental mean value of pass rate was about 98.8%, which was less than expected since the significance level was set to 1%. The value of K_{1000}^+ for the two distributions was 6.57833, and therefore the probability that the experimental pass rate follows the theoretical binomial distribution function was less than 1%.

Figure 2 shows the empirical distribution of pass rate and the theoretical binomial distribution function when our new values for π_i were used. No significant difference between the two distributions can be found. The value of K_{1000}^+ for the two distributions was 0.388979, and therefore the probability that the experimental pass rate follows the theoretical binomial distribution function was more than 1%. Therefore, our new values should be used for accurate randomness testing instead of inaccurate values described in the NIST randomness test suite.

When the number of sequences was increased to more than 22275, even the sequences taken from a true random number generator are expected to be rejected by the overlapping template matching test with high probability (See Discussion). For example, when the number of sequences was 4×10^4 and other test conditions remained unchanged, the null hypothesis that random number sequences taken from DES appear to be random was more likely to be rejected. Acceptance region based on [8] was determined to be [0.988508, 0.991492]. Table 4 shows a test result when experiments were iterated ten times. As Table 4 shows, only three out of ten experiments concluded that random num-

Table 3 Test conditions.

The length of random number sequence	10^6
The length of a substring (M)	1032
The number of sequences	10^3
The iteration number of experiments	10^3
The significance level	1%

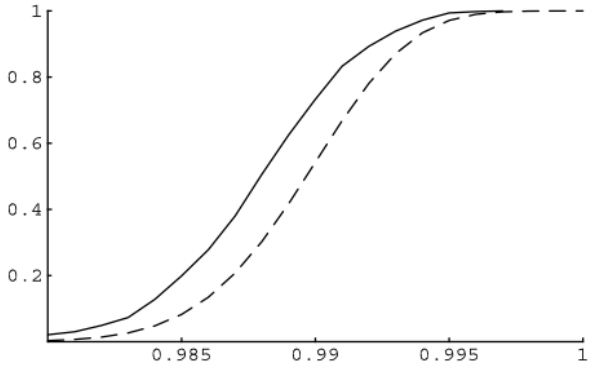

Fig. 1 Empirical distribution (solid line) and theoretical binomial distribution (broken line) of pass rate when inaccurate values for π_i described in the NIST randomness test suite were used.

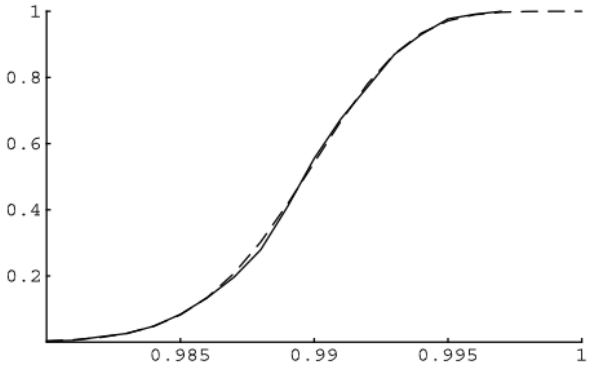

Fig. 2 Empirical distribution (solid line) and theoretical binomial distribution (broken line) of pass rate when our new accurate values for π_i were used.

Table 4 Test result when the number of sequences was increased to 4×10^4 and inaccurate values for π_i described in the NIST randomness test suite were used.

Number	Pass rate	Test result
1	0.987150	non-random
2	0.987950	non-random
3	0.987650	non-random
4	0.989225	random
5	0.987825	non-random
6	0.988425	non-random
7	0.989525	random
8	0.988300	non-random
9	0.988350	non-random
10	0.989175	random

Table 5 Test result when the number of sequences was increased to 4×10^4 and our new values for π_i were used.

Number	Pass rate	Test result
1	0.989025	random
2	0.989850	random
3	0.989775	random
4	0.990675	random
5	0.989300	random
6	0.989575	random
7	0.990675	random
8	0.990450	random
9	0.990075	random
10	0.990650	random

ber sequences taken from DES appear to be random. In contrast, as Table 5 shows, when our new values were used instead of inaccurate values described in the NIST randomness test suite, ten out of ten experiments concluded that random number sequences taken from DES appear to be random.

6. Discussion

On the basis of the results of the experiment, we confirmed that the significant difference shown in Fig. 1 was due to the inaccurate values for π_i used in the NIST randomness test suite. Use of the inaccurate values described in the NIST randomness test suite adversely affects test results of the NIST randomness test suite.

Let s, α be the number of sequences and the significance level, respectively. Let $\hat{p} = 1 - \alpha$ denote the mean value of pass rate. From [8], when an observation of pass rate falls outside of the following range of acceptable pass rates:

$$\left[\hat{p} - 3\sqrt{\frac{\hat{p}(1-\hat{p})}{s}},\ \hat{p} + 3\sqrt{\frac{\hat{p}(1-\hat{p})}{s}}\right],$$

the NIST randomness test suite concludes that the sequences tested are non-random. When the significance level $\alpha = 0.01$ and the number of sequences $s > 22275$, the following inequality holds:

$$0.988 < \hat{p} - 3\sqrt{\frac{\hat{p}(1-\hat{p})}{s}}.$$

Therefore, when the inaccurate values are used and the number of sequences is more than about 22275, even the sequences taken from a true random number generator are expected to be rejected by the overlapping template matching test with high probability.

Since NIST SP800-22 says, "Ideally, many distinct samples should be analyzed.," this problem can not be ignored, where the term "samples" means the number of sequences tested. NIST SP800-22 also says, "Over time, statistical tests are revamped in light of new results. Since many statistical tests are based upon asymptotic approximations, careful work needs to be done to determine how an approximation is."

We therefore strongly suggest that our new values shown in Table 2 should be reflected in the NIST randomness test suite to evaluate randomness of random numbers in cryptographic systems accurately.

7. Conclusions

It has been shown that the overlapping template matching test included in the NIST randomness test suite uses the inaccurate occurrence probabilities π_i of the template. A set of recurrence formulas for $T_i(n)$ has been analyzed, then the accurate occurrence probabilities π_i have been derived. When our new accurate values are used, the empirical distribution of pass rate follows the theoretical binomial distribution function. Thus, we conclude that the significant difference between the empirical distribution of pass rate and the theoretical distribution function, which was previously found in [7], is caused by the inaccurate occurrence probabilities π_i of the template used in the NIST randomness test suite. Since use of the inaccurate values adversely affects the result of the NIST randomness test suite, our new accurate values should be reflected in the NIST SP800-22 document.

References

[1] K. Hamano, F. Satoh, and M. Ishikawa, "Randomness test using discrete Fourier transform," Technical Report 6841, Technical Research and Development Institute, Japan Defense Agency, Sept. 2003.
[2] K. Hamano, "The distribution of the spectrum for the discrete Fourier transform test included in SP800-22," IEICE Trans. Fundamentals, vol.E88-A, no.1, pp.67–73, Jan. 2005.
[3] K. Hamano, "Correction of overlapping template matching test included in NIST randomness test suite," Technical Report 6944, Technical Research and Development Institute, Japan Defense Agency, Oct. 2006.
[4] Kaneko Lab., http://www.ipa.go.jp/security/enc/CRYPTREC/fy16/documents/rep_ID0211_000.pdf, Dec. 2004.
[5] S. Kim, K. Umeno, and A. Hasegawa, "On the NIST statistical test suite for randomness," IEICE Technical Report, ISEC2003-87, Dec. 2003.
[6] D.E. Knuth, The Art of Computer Programming, vol.2, 3rd ed., Addison-Wesley, 1997.
[7] H. Okutomi, M. Kaneda, K. Yamaguchi, and K. Nakamura, "A study on the randomness evaluation method using NIST randomness test," Proc. SCIS 2006, p.10, Jan. 2006.
[8] A. Rukhin, J. Soto, J. Nechvatal, M. Smid, E. Barker, S. Leigh, M. Levenson, M. Vangel, D. Banks, A. Hechert, J. Dray, and S. Vo, "A statistical test suite for random and pseudorandom number generators for cryptographic applications," NIST Special Publication 800-22, (revised May 15, 2001).
[9] Y. Takeda, M. Huzii, T. Kamakura, N. Watanabe, and T. Sugiyama, "The problem of template matching test in the testing randomness by NIST," IEICE Technical Report, ISEC2005-110, Dec. 2005.
[10] Wolfram Research, Inc., Mathematica, Version 4, Champaign, IL, 1999.

Appendix A

A.1 Derivation of Recurrence Formula for $T_0(n)$

Consider a binary sequence of length $n-1$ that does not match an m-bit template B at all. If $n \le m-1$, then the two sequences of length n that are obtained by appending a bit "0" or a bit "1" at the tail of the sequence of length $n-1$, do not match the template B at all. If $n \ge m$ and the sequence of length $n-1$ has an m-bit pattern "$011\cdots11$" at the tail of the sequence, then the sequence of length n that is obtained by appending a bit "1" at the tail of the sequence of length $n-1$, matches the template B once. The number of such sequences is $T_0(n-m-1)$. Therefore, the recurrence formula for $T_0(n)$ is obtained as Eq. (2).

A.2 Derivation of Recurrence Formula for $T_1(n)$

Consider a binary sequence of length n that matches an m-bit template B only once. Such sequences have an $(m+2)$-bit pattern "$011\cdots110$." If the pattern begins from the j-th bit of the sequence, there are $T_0(j)$ patterns before the bit pattern "$011\cdots110$" and $T_0(n-m-2-j)$ patterns after the bit pattern "$011\cdots110$." Therefore, the recurrence formula for $T_1(n)$ is obtained as Eq. (4).

A.3 Derivation of Recurrence Formula for $T_\alpha(n)$

Part of binary sequences of length n with α occurrences of an m-bit template B can be obtained by appending a bit "1" at the first matching run of ones of binary sequence of length $n-1$ that has $\alpha-1$ occurrences of B, and therefore all the sequences obtained match the template B more than once at the first matching run of ones. The number of such sequences is $T_{\alpha-1}(n-1)$.

Other part of binary sequences of length n with α occurrences of an m-bit template B match the template B only once at the first matching run of ones. Consider a binary sequence of length n that has an $(m+2)$-bit pattern "$011\cdots110$" with zero occurrences of B before the bit pattern "$011\cdots110$" and with $\alpha-1$ occurrences of B after the bit pattern "$011\cdots110$." If the pattern begins from j-th bit, then the number of such sequences is $T_0(j)T_{\alpha-1}(n-m-2-j)$.

Therefore, the recurrence formula for $T_\alpha(n)$ is obtained as Eq. (5).

Appendix B

A Mathematica [10] program for computing the accurate values for the occurrence probabilities π_i in Table 2 is shown below.

```
T0=Table[0,{1032}];
T1=Table[0,{1032}];
T2=Table[0,{1032}];
T3=Table[0,{1032}];
T4=Table[0,{1032}];
M=1032;
f0[i_]:=If[i==-1,1,If[i==0,1,T0[[i]]]];
m=9;
T1[[m]]=1;
T1[[m+1]]=2;
makeT0:=Module[{},
For[n=1,n<=m-1,T0[[n]]=2 f0[n-1];n++;];
For[n=m,n<=M,T0[[n]]=2 f0[n-1]-f0[n-m-1];
    n++;];];
makeT0
T0;
makeT1:=Module[{},
For[n=m+2,n<=M,
    T1[[n]]=Sum[f0[j] f0[n-m-2-j],
    {j,-1,n-m-1}];n++;];];
makeT1
T1;
g2[a_,b_]:=
If[a>b,0,Sum[f0[j] T1[[n-m-2-j]],{j,a,b}]];
makeT2:=Module[{},
For[n=2,n<=M,T2[[n]]=T1[[n-1]]
    +g2[-1,n-2 m-2];n++;];];
makeT2
T2;
g3[a_,b_]:=
If[a>b,0,Sum[f0[j] T2[[n-m-2-j]],{j,a,b}]];
makeT3:=Module[{},
For[n=3,n<=M,T3[[n]]=T2[[n-1]]
    +g3[-1,n-2 m-3];n++;];];
makeT3
T3;
g4[a_,b_]:=
If[a>b,0,Sum[f0[j] T3[[n-m-2-j]],{j,a,b}]];
makeT4:=Module[{},
For[n=4,n<=M,T4[[n]]=T3[[n-1]]
    +g4[-1,n-2 m-4];n++;];];
makeT4
T4;
T0[[M]]/2^1032 // N
T1[[M]]/2^1032 // N
T2[[M]]/2^1032 // N
T3[[M]]/2^1032 // N
T4[[M]]/2^1032 // N
0.364091
0.185659
0.139381
0.100571
0.0704323
```

Kenji Hamano received the B.E. and M.E. degrees in Mathematical Engineering from the University of Tokyo in 1999 and 2001, respectively. In 2001, he joined Technical Research and Development Institute, Ministry of Defense. He is currently a doctoral student at the University of Tokyo. His research interests are randomness, statistics and cryptography.

Toshinobu Kaneko received the B.E., M.E., and Ph.D. degrees all in Electrical Engineering, from the University of Tokyo, in 1971, 1973, and 1976, respectively. In 1976, he joined the faculty of Science and Technology, Tokyo University of Science, and since then, as a faculty member, he has been engaged in education and research in the fields of coding theory and information security. Currently, he is a Professor of Department of Electronics Engineering of the university. He is a member of CRPTREC and served as a chairman of Symmetric-Key Cryptography Subcommitte in 2001–2003. He is a member of IEEJ and IPSJ and IEEE.