International Journal of Cryptology Research 3(1): 1-14 (2011)

Spectral Test via Discrete Tchebichef Transform for Randomness

[1]Nur Azman Abu and [1,2]Shahrin Sahib

[1]Faculty of Information and Communication Technology,
Universiti Teknikal Malaysia Melaka (UTeM),
Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia

Email: [1]nura@utem.edu.my and [2]shahrinsahib@utem.edu.my

ABSTRACT

Random key plays essential roles in cryptography. NIST statistical test suite for randomness is the most comprehensive set of random tests. It has been popular and used as a benchmark test for randomness. One of the random tests is spectral test. There has been some serious problem in spectral test as pointed out by few researchers. In this paper, an alternative test shall be proposed to replace the spectral test. The distribution of discrete orthonormal Tchebichef transform has been obtained based on computational observation being made on random noise. A recommendation on the new random test setting for short cryptographic keys shall also be made.

INTRODUCTION

Spectral test is one of the random test of random NIST test suite. As recommended in Rukhin *et al.* (2001), it is only suitable test for long binary sequence. At the same time, the first author found it so difficult to produce a sample input that gives fail result on spectral test for short practical cryptographic keys. The authors are interested in using the random tests for short cryptographic keys. A recommendation in the NIST statistical test suite is only for 1024-bit and above for spectral test based on discrete Fourier transform. Since the practical cryptographic keys are 128, 256, 1024-bit keys and so on, closer attention has been made on the spectral test (Kim *et al.* (2004)).

The focus of spectral test is to detect periodic features within a binary sequence that would indicate a deviation from the assumption of randomness. The test spectral test is developed based on the discrete Fourier transform. The discrete Fourier transform coefficients have been used to detect periodic features in the bit series that would indicate a deviation from the assumption of randomness. Unfortunately, this spectral test is a bit computationally complex and it is not suitable for short cryptographic keys of size 128 and 256-bit. It has been observed that spectral test may easily give a wrong result for short linear binary sequence (Abu *et al*. (2010)). At the same time, discrete Fourier transform is computationally extensive involving complex numbers.

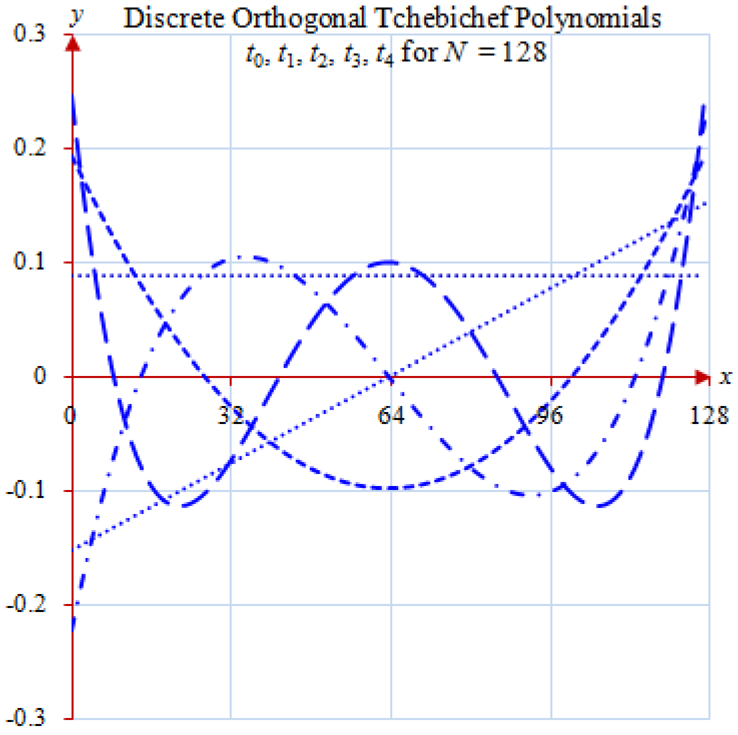

Figure 1: The graphs of the first few Tchebichef polynomials.

This paper has been written to overcome such problem and propose an alternative to random spectral test based on discrete orthonormal Tchebichef transform. Section 2 will briefly go through the theoretical setting of Discrete Orthonormal Tchebichef Polynomials. Section 3 will touch on computational techniques of recurrence relations of discrete Tchebichef transform for large *N*. Section 4 will discuss about an experiment on discrete Tchebichef transform for random binary sequence. In Section 5, a new random spectral statistical test based on discrete Tchebichef transform will be proposed. Section 6 discusses on the consistent result of the new spectral test before further concluded in Section 7.

DISCRETE ORTHONORMAL TCHEBICHEF POLYNOMIALS

Let a one-dimensional discrete orthonormal Tchebichef functions (Mukundan *et al.* (2006)) be given by the following recurrence relations in polynomials $t_n(x)$ of degree n defined on a discrete domain x = 0, 1,…, N−1 and range [−1, +1].

$$t_0(x) = \frac{1}{\sqrt{N}}, \tag{1}$$

$$t_1(x) = (2x+1-N)\sqrt{\frac{3}{N(N^2-1)}} \tag{2}$$

and

$$t_n(x) = (\alpha_1 x + \alpha_2)t_{n-1}(x) - \alpha_3 t_{n-2}(x) \text{ for } n = 2, \ldots, N-1, \tag{3}$$

where

$$\alpha_1 = \frac{2}{n}\sqrt{\frac{4n^2-1}{N^2-n^2}},$$

$$\alpha_2 = \frac{(1-N)}{n}\sqrt{\frac{4n^2-1}{N^2-n^2}} \quad \text{and}$$

$$\alpha_3 = \frac{(n-1)}{n}\sqrt{\frac{2n+1}{2n-3}}\sqrt{\frac{N^2-(n-1)^2}{N^2-n^2}}.$$

This set of discrete orthonormal Tchebichef polynomials satisfies the following properties of orthogonality and completeness.

$$\sum_{x=0}^{N-1} t_m(x)t_n(x) = \delta(m-n) \tag{4}$$

and

$$\sum_{n=0}^{N-1} t_n(x)t_n(y) = \delta(x-y). \tag{5}$$

The graphs of the first few polynomials are given in Figure 1 above. The discrete Tchebichef Moment Transform (TMT) is a two-dimensional transform method using Tchebichef polynomials. It has been shown to provide compact support and perform better than discrete Cosine transform in image compression (Lang *et al.* (2009); Abu *et al.* (2010)).

NUMERICAL INSTABILITY OF RECURRENCE RELATIONS

Numerical instability can easily occur in evaluating the polynomials above if the recurrence relations are being used directly for large *N*. A closer look at recurrence relations above will tell us than none of the coefficients α_1, α_2, α_3 is large and the only factor that contributes to large magnitude variations is the repeated multiplication by *x*. We will use this recurrence relation only to compute the polynomial values for and then fill each row of Table 1 using the recursion given in (Mukundan (2004)).

TABLE 1: Set of orthonormal polynomials is represented by row vector of polynomials, with dotted arrows denoting *n* recursion and *x* recursion respectively.

$t_n(x)$	$x = 0$	1	2	…	$N-1$
$n = 0$	$t_0(0)$	$t_0(1)$	$t_0(2)$	…	$t_0(N-1)$
1	$t_1(0)$	$t_1(1)$	$t_1(2)$	…	$t_1(N-1)$
2	$t_2(0)$	$t_2(1)$	→	…	
□	↓	↓	□		
$N-1$	$t_{N-1}(0)$	$t_{N-1}(1)$	→	…	

Let us start with the scale factor. The above equations (1)-(5) use the following scale factor for the polynomial of degree *n*, $\beta(n,N) = N^n$.

This scale factor was introduced to counteract the asymptotic growth in magnitude of the discrete orthonormal Tchebichef polynomial values with the degree *n*. Then initial value shall be computed as,

$$t_0(x) = \frac{1}{\sqrt{N}} \tag{6}$$

and

$$t_1(x) = (2x+1-N)\sqrt{\frac{3}{N(N^2-1)}} \quad \text{for } x = 0, 1, 2, \ldots, N-1. \tag{7}$$

At the same time, we need also the initial value for domain *x*. Next, the first and second columns in Table 1 may be computed as,

$$t_n(0) = -\sqrt{\frac{N-n}{N+n}}\sqrt{\frac{2n+1}{2n-1}}t_{n-1}(0),$$

(8)

$$t_n(1) = \left[1 + \frac{n(1+n)}{1-N}\right]t_n(0) \text{ for } n = 1, 2, \ldots, N-1. \qquad (9)$$

Then the cell value shall be computed and filled up row-wise from left to right according recurrence relation below.

$$t_n(x) = \gamma_1 t_n(x-1) + \gamma_2 t_n(x-2)$$

for $n = 1, 2, \ldots, N-1$ and $x = 2, 3, \ldots, \frac{N}{2}$ (10)

where

$$\gamma_1 = \frac{-n(n+1)-(2x-1)(x-N-1)-x}{x(N-x)} \text{ and } \gamma_2 = \frac{(x-1)(x-N-1)}{x(N-x)}.$$

Also note that in (10), the recursion can be terminated at $x = \frac{N}{2}$, since the symmetry condition, $t_n(N-1-x) = (-1)^n t_n(x)$ can be utilised to evaluate the polynomial values where *x* is in the range $[N/2, N-1]$. This greatly reduces the computational time and the amount of accumulated computational errors.

Then the binary sequence shall be expressed as a linear combination of the discrete orthonormal Tchebichef polynomials.

Let

$$\varepsilon = c_0 \cdot t_0(x) + c_1 \cdot t_1(x) + \cdots + c_{N-1} \cdot t_{N-1}(x)$$

$$\begin{bmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \vdots \\ \varepsilon_{N-1} \end{bmatrix} = c_0 \cdot \begin{bmatrix} t_0(0) \\ t_0(1) \\ \vdots \\ t_0(N-1) \end{bmatrix} + c_1 \cdot \begin{bmatrix} t_1(0) \\ t_1(1) \\ \vdots \\ t_1(N-1) \end{bmatrix} + \cdots + c_{N-1} \begin{bmatrix} t_{N-1}(0) \\ t_{N-1}(1) \\ \vdots \\ t_{N-1}(N-1) \end{bmatrix} \text{ where } \varepsilon_i = \begin{cases} -1 \\ +1 \end{cases} \text{ for}$$

$i = 0, 1, \ldots, N-1.$

So that the expression may be simplified as matrix equation,

$$\varepsilon = A\, c$$

where

$$\varepsilon = \begin{bmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \vdots \\ \varepsilon_{N-1} \end{bmatrix}, \; A = \begin{bmatrix} t_0(0) & t_1(0) & \cdots & t_{N-1}(0) \\ t_0(1) & t_1(1) & \cdots & t_{N-1}(1) \\ \vdots & \vdots & \ddots & \vdots \\ t_0(N-1) & t_1(N-1) & \cdots & t_{N-1}(N-1) \end{bmatrix} \text{ and } c = \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{N-1} \end{bmatrix}.$$

Thus, the discrete Tchebichef transform may be computed as

$$c = A^{-1}\, \varepsilon$$

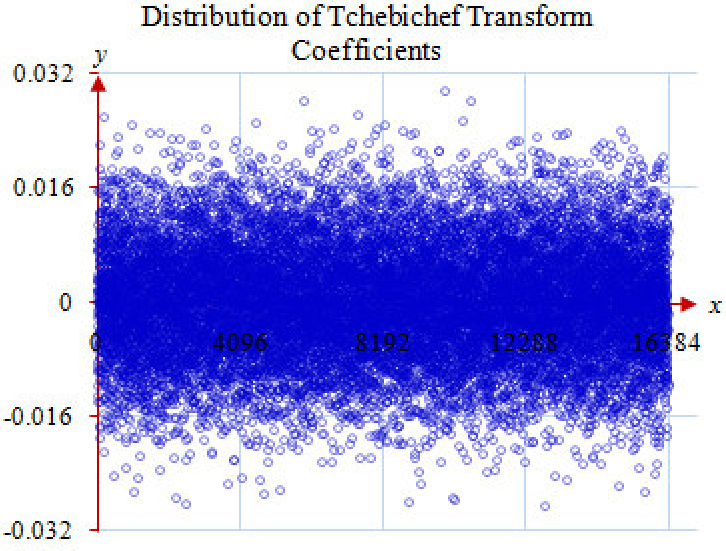

Figure 2: The coefficients of the discrete orthonormal Tchebichef transform for $N = 128$.

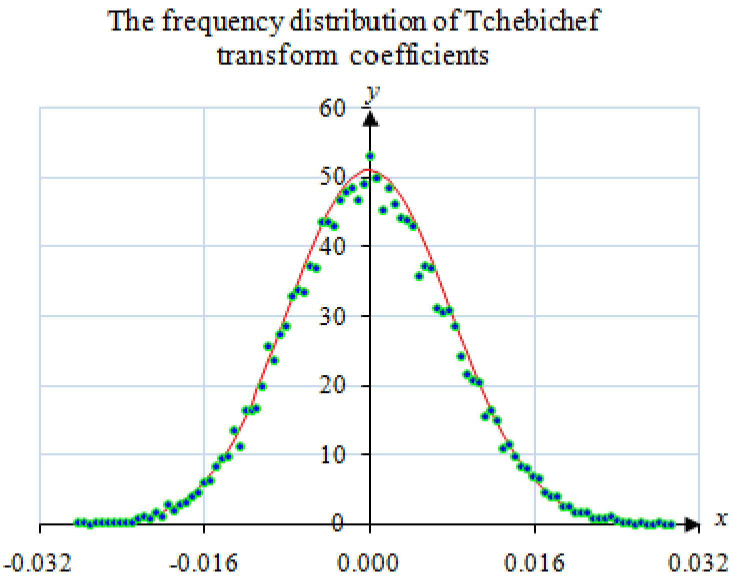

Figure 3a: The frequency distribution of the discrete orthonormal Tchebichef transform coefficients with normal distribution for $N = 128$.

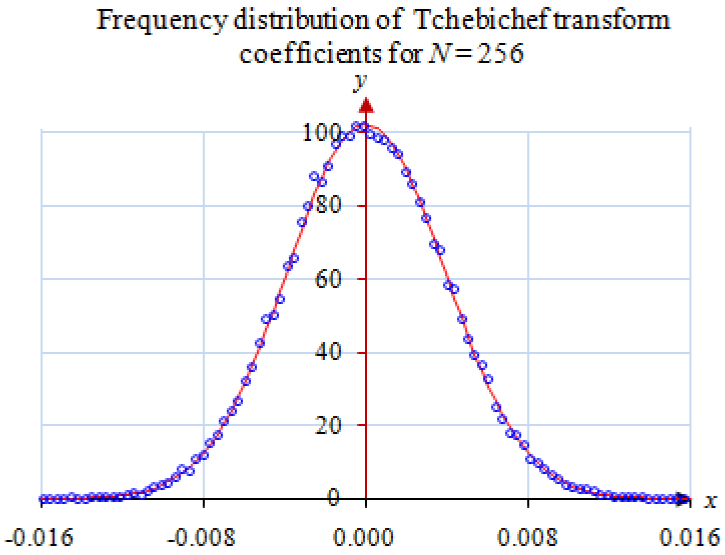

Figure 3b: The frequency distribution of the discrete orthonormal Tchebichef transform coefficients with normal distribution for $N = 256$.

AN EXPERIMENT ON DISCRETE TCHEBICHEF TRANSFORM ON BINARY SEQUENCE

An experiment has been done earlier for N = 128, 256 and 1024-bit binary random ambience sequence (Abu and Sahib (2010)). For M = 1, 128, 256, 1024 blocks, the coefficients of the discrete orthonormal Tchebichef transform have been computed. For N = 128, M = 128, the discrete orthonormal Tchebichef transform coefficients have been plot in Figure 2. The distribution of the coefficients has been plotted with the normal distribution in Figure 3a for N = 128-bit sequence of M = 128 blocks and Figure 3b for N = 256-bit sequence of M = 256 blocks.

TABLE 2: The mean and standard deviation of discrete orthonormal Tchebichef transform coefficients based on random ambience for M blocks of N-bit binary sequence.

N	M	**Mean** $\bar{c}$	**Standard deviation** s_c	**1/N**
128	1	-0.0004130784	0.0078322264	0.0078125000
128	128	-0.0000225058	0.0078127060	0.0078125000
128	256	-0.0000021171	0.0078126189	0.0078125000
128	1024	-0.0000301589	0.0078124716	0.0078125000
256	1	-0.0001751444	0.0039099657	0.0039062500
256	128	-0.0000043217	0.0039063072	0.0039062500
256	256	-0.0000086032	0.0039062703	0.0039062500
256	1024	-0.0000051527	0.0039062541	0.0039062500
1024	1	-0.0000313541	0.0009765360	0.0009765625
1024	128	-0.0000058426	0.0009765487	0.0009765625
1024	256	-0.0000031651	0.0009765592	0.0009765625
1024	1024	-0.0000007596	0.0009765627	0.0009765625

TABLE 3: The sample mean and standard deviation of discrete orthonormal Tchebichef transform coefficients based on *e* for *M* blocks of *N*-bit binary sequence.

N	M	Mean $\overline{c}$	Standard deviation s_c	$1/N$
128	1	-0.0004221496	0.0078317389	0.0078125000
128	128	-0.0000222284	0.0078127068	0.0078125000
128	256	-0.0000834204	0.0078121738	0.0078125000
128	1024	-0.0000118703	0.0078125208	0.0078125000
256	1	-0.0001633110	0.0039104798	0.0039062500
256	128	-0.0000286113	0.0039062048	0.0039062500
256	256	-0.0000105411	0.0039062656	0.0039062500
256	1024	0.0000054469	0.0039062537	0.0039062500
1024	1	0.0000094509	0.0009769939	0.0009765625
1024	128	-0.0000018404	0.0009765645	0.0009765625
1024	256	-0.0000006696	0.0009765641	0.0009765625
1024	1024	-0.0000005249	0.0009765628	0.0009765625

In this experiment, it is found that the random variable of discrete orthonormal Tchebichef transform coefficients follows normal distribution with mean μ zero and standard deviation $\sigma = \frac{1}{N}$. The sample mean moves toward zero and the standard deviation moves towards $\frac{1}{N}$ as the number of block sample gets large as shown in Table 2. The same experiment has been repeated using the first 2^{20}-bit binary sequence from exponent *e* after decimal. The experiment gives similar results as shown in Table 3. This effort has been done to make sure that the normal distribution comes from the randomness of the binary sequence on the discrete Tchebichef transform coefficients rather than from the random ambience audio signals.

DISCRETE TCHEBICHEF TRANSFORM RANDOM TEST PROCEDURE

This spectral test is suitable for short cryptographic keys of size 128, 256, 1024-bit and so on. This test shall perform a chi-square test of the null hypothesis the binary sequence input comes from a normal distribution with specified variance against the alternative that it comes from a normal distribution with a different variance. The step by step procedure is as follows;

(i) Let us define the random variables in hypothesis testing. Let $\boldsymbol{\varepsilon} = \varepsilon_0, \varepsilon_2, \ldots, \varepsilon_{N-1}$ the binary sequence to be tested for randomness. Convert the input binary sequence $\boldsymbol{\varepsilon}$ into values of –1 and +1 to create the sequence $Y = y_1, y_2, \ldots, y_{N-1}$, where $y_i = 2\varepsilon_i - 1$.

(ii) Let $p_1 = \mathrm{P}(\ M < L\ )$ and $p_2 = \mathrm{P}(\ M > R\ )$ then set the hypothesis testing as $H_0 : p_1 = 0.05$ versus $H_1 : p_1 \neq 0.05$ and $H_0 : p_2 = 0.05$ versus $\mathrm{H}_2 : p_2 \neq 0.05$, or equivalently let a random variable y follows Bernoulli distribution with probability p, then $\sum_{i-1}^{n} y_1$ shall follow Binomial distribution with mean $\mu = np$ and variance $\sigma^2 = np(1-p)$. Thus $H_0 : \mu_1 = N_0$ versus $H_1 : \mu_1 \neq N_0$ and $H_0 : \mu_2 = N_0$ versus $H_2 : \mu_2 \neq N_0$.

(iii) Set the significant level $\alpha = 0.05$. Count N_1 = the actual observed percentage of number of coefficients in C that are less than L and N_2 = the actual observed percentage of number of coefficients in C that are larger than R. The lower L and upper R 0.05 tails are displayed in Table 4 below.

(iv) Apply a discrete Tchebichef transform (DTT) on Y to produce: C = DTT(X). The coefficient sequence C represents periodic components of the sequence of bits at different frequencies.

(v) Count N_1 = the actual observed percentage of number of coefficients in M that are less than L and N_2 = the actual observed percentage of number of coefficients in M that are larger than R.

(vi) Compute $N_0 = 0.05n$. N_0 is the expected theoretical (5%) number of lows and peaks (under the assumption of randomness) that are less than L and of larger than R respectively.

$$d_1 = \frac{N_1 - N_0}{\sqrt{n \cdot (0.95)(0.05)}} \text{ and } d_2 = \frac{N_2 - N_0}{\sqrt{n \cdot (0.95)(0.05)}}$$

(vii) Compute

P□value1 = 2[1 □ $\phi(|d_1|)$] and P□value2 = 2[1 □ $\phi(|d_2|)$].

(viii) The null hypothesis shall be rejected if any one of the two P □values < 0.05.

TABLE 4: The lower and upper 0.05 tails for respective n-bit cryptographic key

n	N_0	L	R
128	6.40	-0.01285042	0.01285042
256	12.80	-0.00642521	0.00642521
1024	51.20	-0.00160630	0.00160630

Let e_N be the first N-bit decimal. For instance, take $N = 1024$, then e_N =

```
B7 E1 51 62 8A ED 2A 6A BF 71 58 80 9C F4 F3 C7
62 E7 16 0F 38 B4 DA 56 A7 84 D9 04 51 90 CF EF
32 4E 77 38 92 6C FB E5 F4 BF 8D 8D 8C 31 D7 63
DA 06 C8 0A BB 11 85 EB 4F 7C 7B 57 57 F5 95 84
90 CF D4 7D 7C 19 BB 42 15 8D 95 54 F7 B4 6B CD
D5 5C 4D 79 FD 5F 24 D6 61 3C 31 C3 83 9A 2D DF
8A 9A 27 6B CF BF A1 C8 77 C5 62 84 DA B7 9C D4
C2 B3 29 3D 20 E9 E5 EA F0 2A C6 0A CC 93 ED 87
```

written in hexadecimal.

Let a counter example ε_N be an increment byte counter from zero to $\frac{N}{8}-1$. For instance, for $N = 1024$, then ε_{1024} =

```
00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F
10 11 12 13 14 15 16 17 18 19 1A 1B 1C 1D 1E 1F
20 21 22 23 24 25 26 27 28 29 2A 2B 2C 2D 2E 2F
30 31 32 33 34 35 36 37 38 39 3A 3B 3C 3D 3E 3F
```

```
40 41 42 43 44 45 46 47 48 49 4A 4B 4C 4D 4E 4F
50 51 52 53 54 55 56 57 58 59 5A 5B 5C 5D 5E 5F
60 61 62 63 64 65 66 67 68 69 6A 6B 6C 6D 6E 6F
70 71 72 73 74 75 76 77 78 79 7A 7B 7C 7D 7E 7F
```

written in hexadecimals.

DISCUSSION

It has been shown in (Abu *et al.* (2010)) that the traditional spectral random test using FFT gave the wrong conclusion which was unable to reject the increment byte counter ε_N above for N=128 and N=256 as random binary sequence. The authors proposed a spectral test based on both lower left-hand-side and the upper right-hand-side 0.05 tails. Similarly, here the new spectral test should make full use of the symmetric normal distribution of the DTT coefficients.

TABLE 5: The upper 0.05 tails give consistent passing result on randomness for the first N-bit exponent e_N.

n	N_0	R	N_2	d_2	P-value2
128	6.4	0.01285042	9	1.054439	0.291682
256	12.8	0.00642521	16	0.917663	0.358795
1024	51.2	0.00160630	52	0.114708	0.908677

TABLE 6: The upper right 0.05 tail gives consistent failing result on randomness for the first N-bit byte counter sequence ε_N.

n	N_0	R	N_2	d_2	P-value2
128	6.4	0.01285042	1	-2.189989	0.028525
256	12.8	0.00642521	2	-3.097112	0.001954
1024	51.2	0.00160630	30	-3.039758	0.002368

The upper right-hand-side 0.05 tail is similar to the original test via Fast Fourier Transform. The result in Table 5 above shows that this new test consistently produces passing result on randomness for short n-bit from exponent e_N. Whereas the result in Table 6 above shows that this test consistently give correct failing results on the randomness of short *n*-bit byte counter sequence for as short as 128-bit numbers.

CONCLUSION

Two-dimensional Tchebichef discrete orthonormal transform is well-known to provide a compact support for image analysis especially compression via sub-block reconstruction. In this paper, a clear statistical distribution on one-dimensional DTT coefficients has been explored. A recommendation on a new spectral random test setting for short cryptographic keys has also been made. The random variable of DTT coefficients follows normal distribution.

This paper has proposed a more accurate and simpler spectral test based on DTT coefficients than original proposal in NIST test suite based on FFT. This test consistently gives failing result on randomness for short *n*-bit byte counter sequence for as short as 128-bit numbers. Tchebichef discrete orthonormal transform provides more efficient and accurate spectral statistical test for randomness.

REFERENCES

Andrew Rukhin, Juan Soto, James Nechvatal, Miles Smid, Elaine Barker, Stefan Leigh, Mark Levenson, Mark Vangel, David Banks, Alan Heckert, James Dray, and San Vo. (2001). A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications. *NIST Special Publication*. 800-22, 15 May 2001.

Nur Azman Abu, Nanna Suryana Herman and Shahrin Sahib. (2010). An Enhancement of the Spectral Statistical Test for Randomness. *International Conference on Computational and Statistical Science (ICCSS 2010)*: 521-525.

Nur Azman Abu, Wong Siaw Lang Nanna Suryana and Ramakrishnan Mukundan .(2010). An Efficient Compact Tchebichef moment for Image Compression. *10th International Conference on Information Science, Signal Processing and their applications (ISSPA2010)*: 448-451.

Nur Azman Abu and Shahrin Sahib. (2010). Random Ambience Key Generation Live on Demand. *Proceedings 2nd International Conference on Signal Processing Systems (ICSPS 2010)*. **1**: 110-114.

Ramakrishnan Mukundan. (2004). Some Computational Aspects Of Discrete Orthonormal Moments. *IEEE Transactions On Image Processing*. **13**(8): 1055-1059, August 2004.

Ramakrishnan Mukundan. (2006). Transform Coding Using Discrete Tchebichef Polynomials. *Proceedings IASTED International Conference on Visualization Imaging and Image Processing VIIP 2006*: 270-275.

Song-Ju Kim, Ken Umeno, and Akio Hasegawa. (2004). Corrections of the NIST Statistical Test Suite for Randomness, Cryptology ePrint Archive: Report 2004/018, 26 Jan 2004, *http://eprint.iacr.org/2004/018.pdf.*

Wong Siaw Lang, Nur Azman Abu, Hidayah Rahmalan. (2009). Fast 4x4 Tchebichef Moment Image Compression. *Proceedings International Conference of Soft Computing and Pattern Recognition (SoCPaR) 2009*: 295-300.