The Comparative Study of Randomness Analysis between Modified Version of LBlock Block Cipher and its Original Design

Nik Azura Nik Abdullah*, Liyana Chew Nizam Chew, Abdul Alif Zakaria
Cryptography Development Department
CyberSecurity Malaysia
Selangor, Malaysia
*Email: azura [AT] cybersecurity.my

Kamaruzzaman Seman, Norita Md Norwawi
Faculty of Science and Technology
University Sains Islam Malaysia (USIM)
Negeri Sembilan, Malaysia

Abstract— **In this research paper, we present and compare the randomness analysis conducted towards LBlock block cipher and its modified version namely Modified LBlock. Among the important requirement when designing a block cipher algorithm is that the algorithm can act as a random number generator. Therefore, the aim of performing modification towards LBlock algorithm is to enhance its randomness results. Modification were made by replacing the eight 4 X 4 S-boxes with four different 16 X 16 S-boxes which has the same security strength as S-box of AES. During experimentation, this research project considers a full rounds of LBlock and Modified LBlock algorithms which both algorithms accepts a 64-bit plaintext, utilizes an 80-bit key, executes in 32 rounds and produces a 64-bit ciphertext. Nine different data categories were used to generate inputs (plaintext and key), with each having 100 samples. Blocks of ciphertext were generated from these algorithms and were concatenated to construct a binary sequence. NIST Statistical Test Suite consisting of sixteen tests was used to conduct testing and analysis, and the significance level was set to 1%. From the comparative analysis done, it is concluded that the randomness analysis results for Modified LBlock block cipher has 42.96% improvement compared to its original design when tested under 1% significance level and using the same samples.**

Keywords-LBlock Block Cipher, Data Categories, NIST, significance level, randomness

I. INTRODUCTION

LBlock block cipher used in Lightweight Cryptography was developed by Wenling Wu and Lei Zhang in 2011. This algorithm was first published at ACNS2011. This algorithm can be implemented efficiently not only in hardware environment but also in software platforms and can achieve competitive performances compared with other known ultra-lightweight block ciphers. Wu and Zhang claimed that their security analysis showed that the full-round LBlock encryption can provide enough security margins against known cryptanalytic techniques, such as differential cryptanalysis, linear cryptanalysis, impossible differential cryptanalysis, related-key attack and etc [11].

Over the past couple of years, there has been some work conducted in order to evaluate the security of LBlock block cipher against various cryptanalysis attacks. In 2013, statistical analysis has been performed towards the full round LBlock block cipher. The purpose of conducting this analysis was to ensure that this algorithm is suitable to act as a random number generator, as it is an important requirement when designing an encryption algorithm [5]. This is to ensure that ciphertext produce from LBlock block cipher is uniformly distributed should all of the plaintext blocks used during encryption are different. Results from this analysis concluded that this block cipher algorithm is not random based on 1% significance level only [6]. Therefore, to enhance its randomness results, a few modification designs has been conducted towards LBlock block cipher. Design with the best randomness analysis result is presented in this paper.

This paper illustrates the comparative statistical analysis results on full round LBlock and Modified-LBlock block ciphers using NIST Statistical Test Suite and is organized as follows:- A short description of LBlock block cipher and the modification made to this algorithm are described in **Section II**. **Section III** firstly provides detail explanations of nine different data type used to generate plaintext and keys, which act as inputs to LBlock and Modifed LBlock block ciphers. This section also provides the information on NIST Statistical Test Suite. Finally, this section demonstrates the research methodology and experimentation setup conducted in analyzing the ciphertext produced from both algorithms. The comparative randomness results for LBlock and Modified LBlock block ciphers are discuss in **Section IV**. Summary and conclusion of this research are demonstrated in **Section V**.

II. MODIFIED LBLOCK BLOCK CIPHER

A. Description of LBlock Block Cipher

LBlock Block Cipher is based on a modified Fiestel structure with block size of 64-bit. This algorithm accepts an 80-bit key and executes in 32 rounds. It also employs S-box layers in both encryption algorithm and key scheduling,

confusion function that act as a non-linear layer and diffusion function that act as a linear layer.

In the encryption algorithm of LBlock block cipher, the 64-bit plaintext, M is processed by first dividing it into two separate sequences; X_1 is the left 32-bit half and X_0 is the right 32-bit half. This is denoted as:-

$$M\ (64\text{-bit}) = X_1\ (32\text{-bit}) \,||\, X_0\ (32\text{-bit}) \quad (1)$$

The encryption process continues in 32 rounds for $2 \le i \le 33$ as follows:-

$$X_i = F\ (X_{i-1}\ ,\ K_{i-1})\ \oplus (X_{i-2} <<< 8) \quad (2)$$

where F, and K_i are denoted as the Round Function and the 32-bit round subkey respectively, meanwhile $||$, $\oplus$ and $<<<$ are denoted as concatenation of two binary strings, exclusive-OR operation and a left cyclic shift operation respectively. Figure 1 shows the Round Function that is used in each round.

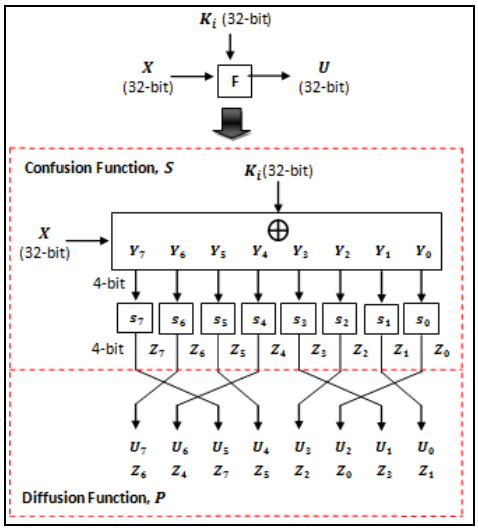

Figure 1. Round Functiion, F which includes Confusion Function,S and Diffusion Function, P.

The produced ciphertext, C from this encryption process is a concatenation of X_{32} and X_{33} and is presented as:-

$$C = X_{32} \,||\, X_{33} \quad (3)$$

1) *Round Function, F:* Constructed from two other functions; Confusion function, S and Diffusion Function, P. This is defined as follows:-

$$U = F(X, K_i) = P\ (\ S\ (\ X \oplus K_i)\ )\ \quad (4)$$

2) *Confusion Function, S:* The non-linear layer of the Round Function, F that consists of eight 4 X 4 S-boxes namely s_0, s_1, s_2, s_3, s_4, s_5, s_6, and s_7 (shown in Table I). This function is defined as follows:-

$$Y=Y_7||Y_6||Y_5||Y_4||Y_3||Y_2||Y_1||Y_0$$
$$\rightarrow \quad Z=Z_7||Z_6||Z_5||Z_4||Z_3||Z_2||Z_1||Z_0 \quad (5)$$

where $Z_7 = s_7(Y_7)$, $Z_6 = s_6(Y_6)$, $Z_5 = s_5(Y_5)$, $Z_4 = s_4(Y_4)$, $Z_3 = s_3(Y_3)$, $Z_2 = s_2(Y_2)$, $Z_1 = s_1(Y_1)$, $Z_0 = s_0(Y_0)$.

TABLE I. CONTENT OF S-BOXES S_0, S_1, S_2, S_3, S_4, S_5, S_6,AND S_7.

S-boxes	Content
s_0	14, 9, 15, 0, 13, 4, 10, 11, 1, 2, 8, 3, 7, 6, 12, 5
s_1	4, 11, 14, 9, 15, 13, 0, 10, 7, 12, 5, 6, 2, 8, 1, 3
s_2	1, 14, 7, 12, 15, 13, 0, 6, 11, 5, 9, 3, 2, 4, 8, 10
s_3	7, 6, 8, 11, 0, 15, 3, 14, 9, 10, 12, 13, 5, 2, 4, 1
s_4	14, 5, 15, 0, 7, 2, 12, 13, 1, 8, 4, 9, 11, 10, 6, 3
s_5	2, 13, 11, 12, 15, 14, 0, 9, 7, 10, 6, 3, 1, 8, 4, 5
s_6	11, 9, 4, 14, 0, 15, 10, 13, 6, 12, 5, 7, 3, 8, 1, 2
s_7	13, 10, 15, 0, 14, 4, 9, 11, 2, 1, 8, 3, 7, 5, 12, 6

3) *Diffusion Function, P:* The linear layer of the Round Function, F that permutes inputs Z to produce outputs U. This function is defined as follows:-

$$Z=Z_7||Z_6||Z_5||Z_4||Z_3||Z_2||Z_1||Z_0$$
$$\rightarrow \quad U=U_7||U_6||U_5||U_4||U_3||U_2||U_1||U_0 \quad (6)$$

where $U_7 = Z_6$, $U_6 = Z_4$, $U_5 = Z_7$, $U_4 = Z_5$, $U_3 = Z_2$, $U_2 = Z_0$, $U_1 = Z_3$, $U_0 = Z_1$

The Key Schedule of LBlock block cipher accepts an 80-bit master key, $K = k_{79}, k_{78}, k_{77}, \ldots, k_1, k_0$ which is stored in a key register. After 32 rounds of execution, this algorithm will produce 32 round subkeys denoted as K_i for $1 \le i \le 31$.

1) *Step 1:* Round subkey, K_1 is the leftmost 32 bits of the master key, K.
2) *Step 2:* For $1 \le i \le 31$, the key register is updated as follows:-
   a) Contents in key register are shifted 29-bits to the left with a cyclic shift operation.
   b) Contents in key register at position 79, 78, 77, 76 and 75, 74, 73, 72 are replaced with contents from two 4 X 4 S-boxes s_9 and s_8 (shown in Table 2) respectively.
   c) Contents in key register at position 50, 49, 48, 47 and 46 are XORed with binary form of current round.

$$[k_{50},k_{49},k_{48},k_{47} \text{ and } k_{46}] \oplus [i]_2 \quad (7)$$

   d) Round subkey, K_{i+1} is the leftmost 32 bits of the current key register.

B. Modification of LBlock Block Cipher

In LBlock block cipher, the Confusion Function, *S* contains eight 4 X 4 S-boxes in parallel namely s_0, s_1, s_2, s_3, s_4, s_5, s_6, and s_7. For modification, four different 16 X 16 S-boxes which has same security strength as AES S-box are used to replace the eight 4 X 4 S-boxes. The four S-boxes used in this design are denoted as s_0, s_1, s_2, and s_3. Figure 2 shows the Round Function that is used in each round of Modified LBlock block cipher.

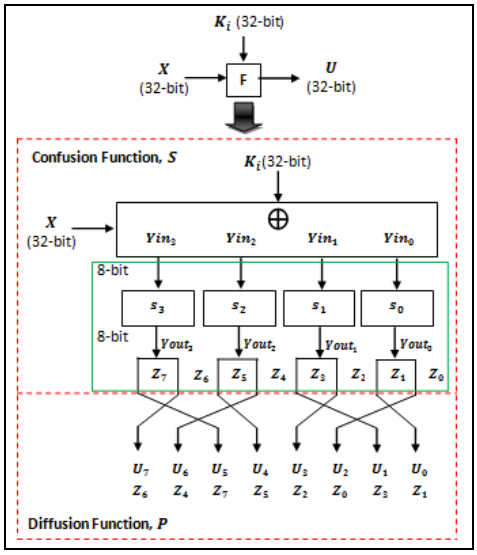

Figure 2. Round Function, *F* of Modified LBlock block cipher.

The Confusion Function, *S* of Modified LBlock block cipher is clearly defined as follows:-

$$Yin = Yin_3 \,||\, Yin_2 \,||\, Yin_1 \,||\, Yin_0$$
$$\rightarrow \quad Yout = Yout_3 \,||\, Yout_2 \,||\, Yout_1 \,||\, Yout_0$$
$$\rightarrow \quad Z = Z_7 \,||\, Z_6 \,||\, Z_5 \,||\, Z_4 \,||\, Z_3 \,||\, Z_2 \,||\, Z_1 \,||\, Z_0 \qquad (8)$$

where $Yout_3 = s_3(Yin_3) = Z_7 \,||\, Z_6, Yout_2 = s_2(Yin_2) = Z_5 \,||\, Z_4,$
$Yout_1 = s_1(Yin_1) = Z_3 \,||\, Z_2, Yout_0 = s_0(Yin_0) = Z_1 \,||\, Z_0.$

The four S-boxes used in this design are constructed using the same method as constructing the S-box for AES; taking multiplicative inverse of a matrix element in $GF(2^8)$ and applying an affine (over $GF(2)$) transformation as follows:-

For s_0,

$$\begin{bmatrix} 1&1&0&1&1&0&1&0\\ 0&1&1&0&1&1&0&1\\ 1&0&1&1&0&1&1&0\\ 0&1&0&1&1&0&1&1\\ 1&0&1&0&1&1&0&1\\ 1&1&0&1&0&1&1&0\\ 0&1&1&0&1&0&1&1\\ 1&0&1&1&0&1&0&1 \end{bmatrix} \begin{bmatrix} x_7\\ x_6\\ x_5\\ x_4\\ x_3\\ x_2\\ x_1\\ x_0 \end{bmatrix} + \begin{bmatrix} 0\\ 1\\ 0\\ 1\\ 1\\ 0\\ 1\\ 1 \end{bmatrix}$$

For s_1,

$$\begin{bmatrix} 0&1&1&0&1&1&0&1\\ 1&0&1&1&0&1&1&0\\ 0&1&0&1&1&0&1&1\\ 1&0&1&0&1&1&0&1\\ 1&1&0&1&0&1&1&0\\ 0&1&1&0&1&0&1&1\\ 1&0&1&1&0&1&0&1\\ 1&1&0&1&1&0&1&0 \end{bmatrix} \begin{bmatrix} x_7\\ x_6\\ x_5\\ x_4\\ x_3\\ x_2\\ x_1\\ x_0 \end{bmatrix} + \begin{bmatrix} 0\\ 1\\ 0\\ 1\\ 1\\ 0\\ 1\\ 1 \end{bmatrix}$$

For s_2,

$$\begin{bmatrix} 1&1&0&0&0&1&1&1\\ 1&1&1&0&0&0&1&1\\ 1&1&1&1&0&0&0&1\\ 1&1&1&1&1&0&0&0\\ 0&1&1&1&1&1&0&0\\ 0&0&1&1&1&1&1&0\\ 0&0&0&1&1&1&1&1\\ 1&0&0&0&1&1&1&1 \end{bmatrix} \begin{bmatrix} x_7\\ x_6\\ x_5\\ x_4\\ x_3\\ x_2\\ x_1\\ x_0 \end{bmatrix} + \begin{bmatrix} 0\\ 1\\ 1\\ 0\\ 0\\ 0\\ 1\\ 1 \end{bmatrix}$$

For s_3,

$$\begin{bmatrix} 1&1&1&0&0&0&1&1\\ 1&1&1&1&0&0&0&1\\ 1&1&1&1&1&0&0&0\\ 0&1&1&1&1&1&0&0\\ 0&0&1&1&1&1&1&0\\ 0&0&0&1&1&1&1&1\\ 1&0&0&0&1&1&1&1\\ 1&1&0&0&0&1&1&1 \end{bmatrix} \begin{bmatrix} x_7\\ x_6\\ x_5\\ x_4\\ x_3\\ x_2\\ x_1\\ x_0 \end{bmatrix} + \begin{bmatrix} 0\\ 1\\ 1\\ 0\\ 0\\ 0\\ 1\\ 1 \end{bmatrix}$$

The constructed S-boxes are as follows:-

```
s0 = {{5d,06,1c,23,11,88,62,c8,97,d1,b7,26,2e,16,7b,a6},
      {38,43,f7,bc,c4,01,e0,64,e4,d9,14,91,a2,8b,4c,98},
      {ef,e1,52,A3,08,67,41,a7,7d,31,9f,de,83,e6,2d,90},
      {81,bd,f3,57,15,02,3b,3d,4e,b5,36,ba,d5,19,bf,69},
      {04,f9,03,60,da,d8,ce,61,1b,8d,ac,05,53,00,cc,dc},
      {a1,89,6b,4a,3c,5b,70,74,32,82,80,f8,65,fa,bb,68},
      {33,3f,c1,7e,0a,40,58,66,95,0b,1e,c9,6e,a8,6d,b4},
      {d4,ae,c5,c6,e8,18,42,34,f5,55,93,86,c0,94,ab,46},
      {1d,54,0f,f0,9e,b8,2f,e5,72,44,73,12,78,eb,af,f6},
      {92,8c,35,0c,49,9c,9d,e3,5a,85,1f,2a,79,24,71,29},
      {cf,e2,37,cb,aa,f4,d6,51,ed,fc,5e,e7,27,cd,25,56},
      {ea,4d,b2,17,b3,63,8f,0e,ad,f1,8e,6f,c2,99,2b,be},
      {6a,ec,6c,76,ff,fe,20,07,1a,b6,d3,30,df,4d,2c,96},
      {39,dd,9a,0d,10,a4,fb,21,28,c7,4b,a5,a9,22,45,d7},
      {75,b1,48,7a,fd,5c,7c,77,87,8a,13,a0,d2,3e,e9,e3},
      {09,db,59,ee,3a,50,b0,47,7f,f2,b9,9b,ca,84,d0,5f}}.

s1 = {{5d,eb,df,a1,c5,f6,23,76,c8,44,88,ab,bb,cb,11,aa},
      {97,61,08,9e,6e,e5,26,2f,2e,54,cf,c4,a2,f0,7d,d6},
      {38,24,43,a0,f7,29,65,a8,1d,85,d8,5a,e0,2a,bd,c6},
      {e4,9c,00,49,cd,e3,91,9d,7b,8c,8b,92,4c,d5,98,35},
      {ef,14,e1,27,52,56,7a,25,d1,fc,be,ed,41,e7,7e,5e},
      {a4,f4,31,73,9f,51,07,0f,83,e2,e6,16,2d,12,90,37},
      {81,99,64,1b,f3,67,57,2b,cc,f1,db,74,3b,b6,3d,8e},
      {4e,ba,6c,6a,36,d7,63,8f,0c,4d,c0,ea,66,ce,b0,6b},
      {dd,4f,f9,06,da,96,b9,2c,03,6f,01,c3,17,30,b8,0a},
      {c2,fe,8d,ff,75,de,dc,20,53,ec,d9,b3,15,af,05,b5},
      {78,22,89,70,b2,0e,4a,45,3c,1e,5b,28,a9,7c,ad,4b},
      {32,7d,82,c9,80,21,f8,fb,bc,04,fa,39,62,d4,b1,9a},
      {33,3e,3f,0b,18,1a,a7,e9,d3,8a,40,87,58,79,bf,ca},
      {95,5c,d2,fd,c7,ae,10,a5,b7,68,71,ac,b4,a3,6d,48},
      {0d,84,77,13,1c,5f,1f,09,e8,f2,c1,a6,42,9b,34,60},
      {f5,50,55,3a,93,47,86,69,19,02,94,d0,72,ee,46,59}}.

s2 = {{63,80,e1,60,51,62,e2,b7,09,2f,e3,6a,d0,f5,7a,c1},
      {56,5f,36,a0,50,24,e7,11,ba,95,5b,fa,9c,9b,41,17},
      {f9,b0,7d,cb,c9,e8,f1,96,89,bb,b3,31,21,14,29,ad},
      {8f,f7,6b,77,0c,dd,af,5c,ef,4d,1f,53,72,eb,59,a1},
      {2e,7f,8a,4c,6c,c2,44,1b,45,68,d5,79,2a,73,ea,9f},
      {65,35,0f,b2,0b,90,39,64,42,76,d8,28,46,86,04,f6},
      {15,f2,5a,70,67,a6,69,bf,a7,30,4f,e0,05,88,fc,1a},
      {25,7b,07,fe,5d,bc,08,b1,98,d9,54,2b,0d,f0,71,55},
      {b6,8e,6d,92,e4,fd,87,ed,97,fb,c0,a8,83,a4,2c,61},
      {03,3f,e6,94,4b,4a,1d,1e,c7,d2,18,8d,d4,c4,6e,74},
      {13,49,48,4e,26,cf,8b,84,57,75,9a,43,3d,bd,93,20},
      {f3,16,e9,a2,be,b5,c6,3a,82,c5,91,52,a3,40,da,0e},
      {58,00,ab,ca,8c,db,99,d7,12,b4,81,ec,66,b8,7e,5e},
      {01,c8,b9,c3,06,6f,d1,ce,23,a9,e5,38,df,37,ac,dc},
      {33,10,1c,2d,22,34,de,9d,7c,cc,ff,32,d6,a5,0a,f4},
      {9e,3b,3e,ae,f8,d3,47,02,27,3c,aa,ee,19,85,78,cd}}.

s3 = {{63,a4,66,65,07,61,60,ca,b7,fb,62,71,04,4e,51,26},
      {09,1b,c9,e4,05,ed,6a,87,d0,8e,13,50,9c,92,27,8b},
      {56,c4,5f,32,36,74,46,88,b6,d2,c2,c7,e7,8d,f7,fe},
      {ba,4a,73,4b,bd,1e,fa,1d,7a,3f,9b,03,41,72,17,e6},
      {f9,5b,b0,3d,7d,20,2d,93,2f,75,0e,57,f1,43,70,9a},
      {6f,cf,bb,c0,b3,84,d7,6d,21,49,14,f5,29,a8,ad,48},
      {8f,40,11,45,6b,e8,77,da,ea,c5,3b,64,af,b4,5c,91},
      {ef,53,ab,58,1f,dc,b5,c6,94,16,0d,f3,bf,44,47,0f},
      {c8,b8,7f,80,6c,5e,aa,7e,8a,52,24,f4,a2,ec,fd,67},
      {a3,db,68,8c,33,31,9f,99,2a,00,95,be,0c,2c,79,4d},
      {83,37,35,39,e9,3a,b2,ac,0b,4f,90,23,df,de,82,e5},
      {42,89,76,e0,d8,ce,28,d1,a0,2e,86,01,e2,25,10,b9},
      {15,a5,f2,30,bc,12,96,0a,81,cc,a6,7c,69,d4,59,19},
      {a7,34,d6,22,a9,7b,06,38,e3,f6,6e,d5,1a,cb,fc,1c},
      {c3,85,9d,ff,e1,cd,18,9e,5d,3c,5a,c1,08,ee,b1,4c},
      {98,d3,d9,f8,54,02,2b,a1,eb,dd,f0,78,97,ae,55,3e}}.
```

Strength and weakness of S-boxes can be evaluated using two approaches. The first evaluation focuses on the properties of S-box which covers the area of Nonlinearity, Algebraic Degree and Differential Uniformity properties [10]. The other testing that has been implemented to evaluate S-boxes is based on new techniques for Strict Avalanche Criterion (SAC) that was proposed by Phu Phu Mar and Knin Maung Latt [9]. This method highlighted the avalanche effect, completeness and strong S-box properties.

Nonlinearity features makes an S-box design to be resistance to linear cryptanalysis. For S-box in $GF(2^n)$, the upper bound of nonlinearity is defined as $N(f) = 2^{n-1} - (2^{-1} * 2^{n/2})$. AES S-box is in $GF(2^8)$, therefore the upper bound value for N is 120. The smallest value of nonlinearity parameter, NL must be between $100 < NL < 120$, or else the S-box will be susceptible to linear cryptanalysis. The Nonlinearity NL of the AES S-box equals to 112 [10], thus S-box of AES is not susceptible to linear cryptanalysis attack. This satisfies that AES S-box meets the characteristic of good nonlinearity property.

A good S-box requires for the Algebraic Degree (AD) to be high. In order to resist higher order differential cryptanalysis the measurement of AD is suggested to be $AD \geq 4$. AD of an S-box is calculated using $AD = r - 1$ where r is the degree of the S-box. Therefore AD of AES S-box is $AD = 8 - 1 = 7$ [10]. This satisfies that AES S-box meets the characteristic of good algebraic degree property.

Differential Uniformity (DU) property of an S-box provides information regarding the susceptibility of the block cipher against differential cryptanalysis attack. To examine the DU properties of an S-box, difference pairs of input and output of the S-box are recorded in a difference uniformity table, where each element from this table represents the number of occurrences of output difference value corresponding to input difference value which DU indicates the highest. A good S-box should have DU value in the range of $2 \leq DU \leq 6$. AES S-box has $DU = 4$ [10], thus S-box of AES is not susceptible to differential cryptanalysis attack. This satisfies that AES S-box meets the characteristic of good differential uniformity property.

To determine the avalanche effect property of any S-box, frequency of various hamming weight is analyzed when using various random inputs to the S-box. The avalanche effect property is satisfied if frequencies of various hamming weight is normally distributed [9].

The completeness property of an S-box is established by analyzing the frequency of various differential value of outputs when using various random inputs to the S-box. If the frequencies of differential output vary, the S-box tested is poor, whereas if the frequencies is uniformly distributed, then the S-box tested is considered as a good S-box [9].

To determine the strong S-box property of an S-box, analysis of hamming weights according to the bit position when using various random inputs to the S-box needs to be conducted. If the frequencies of hamming weight vary, the S-box tested is poor, whereas if the frequencies is uniformly distributed, then the S-box tested is considered as a good S-box [9].

Security strength of s_0, s_1, s_2, and s_3 are tested using the method described above. Results obtained shows that all four S-boxes has similar security strength as compared to AES S-box; Nonlinearity, $NL = 112$ (which satisfies $100 < NL < 120$), Algebraic Degree, $AD = 7$ (which satisfies $AD \geq 4$), Differential Uniformity, $DU = 4$ (which satisfies $2 \leq DU \leq 6$), normal distribution graph for avalanche effect property, uniform distribution graph for completeness property and also uniform distribution graph for strong S-box property. Therefore, it is concluded that s_0, s_1, s_2, and s_3 has good characteristics of a strong S-Box.

III. Experimental Setup

A. Data Categories for LBlock and Modified LBlock

The randomness testing activity was conducted on a full round of LBlock and Modified LBlock block ciphers using NIST Statistical Test Suite. Output (ciphertext) produced by these two algorithms are analyzed. Inputs (plaintext and key) for these algorithms were generated using nine different categories of data. These data categories are described in Appendix A [3], [4]:- Strict Key Avalanche (SKA)[A], Strict Plaintext Avalanche (SVA)[B], Plaintext / Ciphertext Correlation (PCC)[C]. Cipher Block Chaining Mode (CBCM)[D], Random Plaintext / Random Key (RPRK)[E], Low Density Key (LDK)[F], High Density Key (HDK)[G], Low Density Plaintext (LDP)[H] and High Density Plaintext (HDP)[I]. In this experiment, 100 samples are generated using each data category for each algorithm. The number of blocks produced for every sample is depending on the block size or key size of LBlock and Modified LBlock block ciphers [11]. To construct a large sequence of bits to be inserted into the NIST Statistical Test Suite, derived blocks produced from each data categories are concatenated. The number of derived blocks and derived bits per sample produces from each data categories are listed in Table II [6].

TABLE II. Number of derived block and derived bits per sample for all nine data categories.

Data Categories	Derived blocks	Derived bits
SKA, SPA	15,680	1,003,520
PCC, CBCM, RPRK	15,625	1,000,000
LDK, HDK	3,241	207,424
LDP, HDP	3,241	133,184

B. NIST Statistical Test Suite

After reviewing many possible randomness testing tools available, NIST Statistical Test Suite is contemplated to be a reliable tool to perform tests in this research study. This statistical package focuses on a variety of different types of non-randomness that could exist in a sequence [2]. The NIST Statistical Test Suite consists of sixteen tests which are separated into two categories. Nine tests categorizes as Non-

Parameterized Test Selection do not require users to enter any parameter in obtaining the p-values for each test, whereas the other seven tests categorized as the Parameterized Test Selection require users to define one or two parameter value(s). All sixteen tests are listed according to their category as in Table III. This table also presents the minimum number of bits required for each test as recommended by NIST.

TABLE III. MINIMUM NUMBER OF BITS RECOMMENDED BY NIST FOR ALL SIXTEEN TESTS.

	Statistical Test	Derived blocks
Non – Parameterized Test Selection	Frequency	$n \geq 100$
	Runs	$n \geq 100$
	Longest Runs of Ones	$n \geq 750{,}000$
	DFT	$n \geq 1{,}000$
	Lempel-Ziv Complexity	$n \geq 10^6$
	Cumulative Sums	$n \geq 100$
	Random Excursion Variant	$n \geq 10^6$
	Random Excursion	$n \geq 10^6$
	Binary Matrix Rank	$n \geq 38{,}912$
Parameterized Test Selection	Block Frequency	$n \geq 100$
	Non-Overlapping Templates	Not specified
	Overlapping Template	$n \geq 10^6$
	Maurer's Universal	$n \geq 387{,}480$
	Linear Complexity	Not specified
	Serial	Not specified
	Approximate Entropy	Not specified

The Parameterized Test Selection requires users to define one or two parameter value(s). The requirements to be considered in choosing the parameter(s) are listed as in Appendix B [2].

C. Research Methodology and Experimental Setup

The statistical analysis activity is conducted on full round LBlock and Modified LBlock block ciphers (which accept a 64-bit plaintext, utilizes an 80-bit key, executes in 32 rounds and produces a 64-bit ciphertext) using NIST Statistical Test Suite (consisting of sixteen tests separated into two categories; Non-Parameterized Test Selection and Parameterized Test Selection). Section II. *B.* concluded that, in order to determine the randomness of ciphertext produced by any algorithm, NIST Statistical Test Suite requires a large sequence of bit. To achieve this requirement, nine different data categories as specified in Section II. *A.* are used to generate inputs (plaintext and key) for LBlock and Modified LBlock block ciphers, to produce ciphertexts that will be concatenated in a certain way. This will results in having a large sequence of bit, which will be suitable to be inserted into the NIST Statistical Test Suite. Figure 3 illustrates this process [6], [7], [8].

Eleven out of sixteen tests in the NIST Statistical Test Suite reported only one p-value, whereas two tests (Cumulative Sums and Serial tests) reported two p-values and the other three tests reported eight (random Excursion test) , eighteen Random Excursion Variant test) and 148 p-values (Non-Overlapping test) [4]. These give a total of 189 P-values when analyzing a single binary sequence. Due to the limitation of minimum number of bits required, only five categories of data were able to undergo all sixteen tests; SKA, SPA, PCC, CBCM and RPRK, whereas the other four categories; LDK, LDP, HDK and HDP, were able to undergo only eleven tests. This can be seen by comparing Table II and Table III. Therefore, there are 1,585 P-values reported per sample during experimentation. As there are two algorithms with 100 samples per algorithm tested in this research study, there are in total 317,000 P-values evaluated. During experimentation, a total of 1,800 binary sequences (i.e 2 algorithms X 100 samples X 9 data categories) were evaluated. Five categories of data which evaluated all sixteen test constructed 16,000 test values (i.e 2 algorithms X 100 samples X 5 data categories X 16 tests), whereas the other four categories of data which evaluated only eleven test constructed 8,800 test values (ie 2 algorithms X 100 samples X 4 data categories X 11 tests). Therefore, there were 24,800 test values constructed altogether.

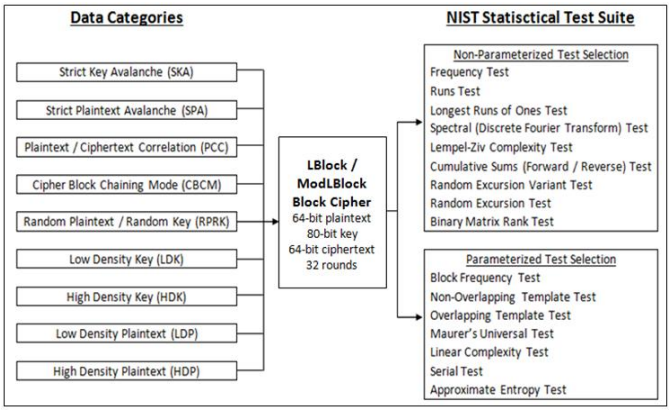

Figure 3. The process for conducting statistical analysis towards LBlock and Modified LBlock block ciphers.

In this experiment, 100 samples are used by all nine data categories for both block ciphers to produce ciphertexts. The level of significance, α has been set to 0.01 (1%). If the P-value $\geq \alpha$, then the ciphertext tested is accepted to be random with a confidence level of 99%. Whereas if the P-value $< \alpha$, then the ciphertext tested is rejected or in other word appears to be non-random with a confidence level of 99%.

The maximum number of binary sequences that is allowable to be rejected at the chosen significance level is computed using formula (9) [4]:-

$$s\left(\alpha + 3\left(\left(\alpha\left(1-\alpha\right)/s\right)^{1/2}\right)\right) \qquad (9)$$

where α is the significance level that was set at 0.01, and s is the number of samples used in the experiment.

IV. Empirical Results and Analysis

Table IV and Table V present and compare the randomness analysis result gathered between LBlock and Modified LBlock block ciphers. Cells highlighted in green ( ) shows the number of maximum rejection allowable for each test. Tests which exceeded the number of allowable rejection rate are highlighted in yellow ( ). Grey cells ( ) indicate that tests are not applicable for that particular data category.

For LBlock block cipher algorithm, at 1 % significance level, four P-values generated from LDK data category were rejected when tested using the Spectral DFT test. The number of rejected P-values has exceeded the maximum number of P-values that is allowable to be rejected, which is three. The same scenario occurs for P-values generated from LDP, RPRK and SKA data categories which were tested using Cumulative Sums - Forward, Binary Matrix Rank and Maurer's Universal tests respectively. Six P-values generated from CBCM data category were rejected when tested using Maurer's Universal test, also exceeded the maximum allowable rejection rate of three sequences. For Non-Overlapping test, the number of P-values that is allowable to be rejected for all data categories is 184. The actual result shows that P-values generated from LDP and HDP data categories exceeded these maximum rejection which are 196 and 187 P-values respectively. For Random Excursion Variant test, the number of P-values that is allowable to be rejected for RPRK data category is 19. The actual result shows that 20 P-values generated from this data category has failed this test and has exceeded this maximum number. For Modified LBlock block cipher algorithm, it is shown that there are no test which exceeded the maximum number of allowable rejection rate at 1 % significance level. In other words, all 100 samples used are accepted to be random with a confidence level of 99%.

V. Summary and Conclusion

We have presented the statistical analysis on full round LBlock and Modified LBlock block cipher using NIST Statistical Test Suite. To determine whether these two algorithms are random or non-random, the significance level was set at 1%. Overall result shows that, Modified LBlock block cipher is 42.96% better than LBlock block cipher. Therefore, modification made to the original LBlock block cipher algorithm, which is replacing the eight 4 X 4 S-boxes with four different 16 X 16 S-boxes which has the same security strength as S-box of AES, has enhanced the randomness results.

TABLE IV. Number of samples rejected in full round testing of LBlock and Modified LBlock block ciphers for Non-Parameterized Test Selection of NIST Statistical Test Suite.

		Non-Parameterized Test Selection									
		Frequency		Runs		Longest Runs of Ones		Spectral DFT		Lempel-Ziv	
# of max rejection		**3**		**3**		**3**		**3**		**3**	
		LBlock	Modified LBlock	LBlock	Modified LBlock	LBlock	Modified LBlock	LBlock	Modified LBlock	LBlock	Modified LBlock
Data Category	SKA	1	1	0	0	2	1	1	0	0	0
	SPA	3	1	0	0	3	2	2	2	0	0
	PCC	3	2	0	1	0	0	1	0	1	3
	CBCM	2	0	1	1	2	2	1	2	2	2
	RPRK	1	0	1	0	0	1	3	1	0	1
	LDK	0	0	0	0	1	2	4	3		
	HDK	0	0	2	0	2	0	2	2		
	LDP	3	3	0	1	1	0	3	2		
	HDP	0	1	2	1	2	0	2	1		

		Binary Matrix Rank		Cumulative Sums				Random Excursion Variant				Random Excursion			
				Forward		Reverse									
# of max rejection		**3**		**3**		**3**									
		LBlock	Modified LBlock	LBlock	Modified LBlock	LBlock	Modified LBlock	LBlock		Modified LBlock		LBlock		Modified LBlock	
Data Category	SKA	1	1	1	1	1	1	**22**	9	**20**	18	**12**	6	**11**	9
	SPA	2	0	2	1	3	1	**21**	10	**20**	7	**12**	5	**11**	5
	PCC	0	1	0	2	2	2	**21**	4	**21**	12	**11**	2	**12**	6
	CBCM	0	0	2	0	2	1	**20**	12	**22**	10	**11**	7	**12**	5
	RPRK	4	1	1	0	1	1	**19**	20	**20**	10	**10**	9	**11**	7
	LDK	1	1	0	0	0	0								
	HDK	1	0	0	0	0	0								
	LDP	0	2	4	2	2	3								
	HDP	2	0	0	0	0	1								

TABLE V. NUMBER OF SAMPLES REJECTED IN FULL ROUND TESTING OF LBLOCK AND MODIFIED LBLOCK BLOCK CIPHERS FOR PARAMETERIZED TEST SELECTION OF NIST STATISTICAL TEST SUITE.

		Parameterized Test Selection							
		Block Frequency		Non-Overlapping		Overlapping		Maurer's Universal	
# of max rejection		**3**		**184**		**3**		**3**	
		LBlock	Modified LBlock	LBlock	Modified LBlock	LBlock	Modified LBlock	LBlock	Modified LBlock
Data Category	SKA	0	0	145	162	1	1	4	1
	SPA	0	2	154	141	0	1	1	1
	PCC	2	1	139	154	0	2	1	2
	CBCM	0	1	140	159	1	1	6	0
	RPRK	3	2	152	143	0	0	3	2
	LDK	0	0	153	168				
	HDK	0	0	166	158				
	LDP	2	0	196	175				
	HDP	0	0	187	182				

		Linear Complexity		Serial				Approximate Entropy	
				P-value 1		P-value 2			
# of max rejection									
		LBlock	Modified LBlock	LBlock	Modified LBlock	LBlock	Modified LBlock	LBlock	Modified LBlock
Data Category	SKA	2	2	0	1	0	0	1	1
	SPA	2	1	2	1	0	0	3	1
	PCC	3	1	1	3	0	1	1	0
	CBCM	1	0	2	0	1	1	2	1
	RPRK	0	2	2	0	1	0	2	1
	LDK	1	1	0	0	0	0	0	1
	HDK	1	2	1	0	2	0	0	0
	LDP	2	0	2	2	0	1	1	2
	HDP	1	1	2	0	2	1	2	1

APPENDIX A: CATEGORIES OF DATA FOR LBLOCK AND MODIFIED LBLOCK BLOCK CIPHERS

Strict Key Avalanche (SKA)[A] & Strict Plaintext Avalanche (SPA)[B]:-The Strict Key {Plaintext} Avalanche data category is used to examine the sensitivity of LBlock and Modified LBlock block ciphers to changes in the 80-bit key {plaintext}. Each sample for this data category utilizes plaintext {key} that is set to all zero, and 196 blocks of random 80-bit base-keys {245 blocks of random 64-bit base-plaintext}. Each base-key {base-plaintext} is flipped at the i^{th} bit, for $1 \leq i \leq 80$ {$1 \leq i \leq 64$} giving a total of 15,680 perturbed-keys {perturbed-plaintexts}. The all-zero plaintext {key} is then encrypted using each perturbed-key {perturbed-plaintexts}. All resultant ciphertexts are XORed with the ciphertext resulting from the encryption using its corresponding base-key {base-plaintext}. Output product of the XOR operation is called a derived block.

Plaintext / Ciphertext Correlation (PCC)[C]:- The Plaintext / Ciphertext Correlation data category is used to examine the correlation between plaintext / ciphertext pairs using ECB mode of operation. Each sample for this data category utilizes 15,625 blocks of random 64-bit plaintext and one random 80-bit key. Each plaintext block is encrypted using the random 80-bit key. The resultant ciphertext is XORed with its corresponding plaintext. Output product of the XOR operation is called a derived block.

Ciphertext Block Chaining Mode (CBCM)[D]:- The Ciphertext Block Chaining Mode data category is computed using CBC mode of operation. Each sample for this data category utilizes plaintext that is set to all zero (*PT*), a random 80-bit key (*k*), and an all-zeroes initialization vector (*IV*). Derived blocks for this data category are ciphertext blocks computed in CBC mode of operation. The first ciphertext block (CT_1) is define as $CT_1 = E_k\ (IV \oplus PT_0)$, whereas subsequent ciphertext blocks is define as $CT_{i+1} = E_k\ (CT_i \oplus PT_i)$ for $1 \leq i \leq 15{,}625$. To construct a large sequence of bits to be inserted into the NIST Statistical Test Suite, derived blocks are concatenated.

Random Plaintext / Random Key (RPRK)[E]:- The Random Plaintext / Random Key data category is used to examine the randomness of ciphertext based on random plaintext and random key. Each sample for this data category utilizes 15,625

blocks of random 64-bit plaintext and one random 80-bit key. Each plaintext block is encrypted using the random 80-bit key. Derived blocks for this data category are ciphertext blocks computed in ECB mode of operation.

Low Density Keys (LDK)[F] & High Density Keys (HDK)[G]:- The Low {High} Density Keys data category is formed based on low-density {high-density} 80-bit keys. Each sample for this data category utilizes 3,241 blocks of random 64-bit plaintext. The first plaintext block is encrypted using an all-zeroes {all-ones} 80-bit key. Plaintext blocks 2 – 81 are encrypted using an 80-bit key with a single '1' {'0'} in each of the 80-bit position of the key and all other key bits are set to '0' {'1'}. Plaintext blocks 82 – 3,241 are encrypted using an 80-bit key with two '1's {'0's} in each combination of two bit positions of the key and all other key bits are set to '0' {'1'}. Derived blocks for this data category are ciphertext blocks computed in ECB mode of operation.

Low Density Plaintext (LDP)[H] & High Density Plaintext (HDP)[I]:- The Low {High} Density Plaintext data category is formed based on low-density {high-density} 64-bit plaintext blocks. Each sample for this data category utilizes 2,081 blocks of random 80-bit keys. The first plaintext block is encrypted using an all zeroes {all-ones} 64-bit plaintext. Plaintext blocks 2 - 65 consist of one bit of '1' {'0'} and all other bits are '0' {'1'}. Plaintext blocks 66 - 2,081 consist of two bits of '1' {'0'} and all other bits are '0' {'1'}. The '1's {'0's} appear in each combination of two bit positions of the plaintext block. Derived blocks for this data category are ciphertext blocks computed in ECB mode of operation.

APPENDIX B: REQUIREMENT IN CHOOSING PARAMETER(S) FOR PARAMETERIZED TEST SELECTION

Abbreviation: block length (M or L), length of bit string (n), non-overlapping blocks ($N = n/M$), template length (m), number of degrees of freedom (K), theoretical probabilities (π_i) and number of block in the initialization sequence (Q).

- Block Frequency test: M is selected such that $n \geq MN$, $M \geq 20$, $M > 0.01n$ and $N < 100$. Therefore, M chosen is 20,000 for all categories of data.
- Non-Overlapping Template test: $N = 8$ has been specified, m is recommended that $m = 9$ or $m = 10$, $N \leq 100$ and $M > 0.01n$ where $N = [n/M]$. Therefore, m chosen is 10 for all categories of data.
- Overlapping Template test: K, M and N have been fixed in the test code such that $n \geq 10^6$, m is recommended that $m = 9$ or $m = 10$, $n \geq MN$, N (min π_i), $\lambda = (M–m+1) / 2^m \approx 2$, $m \approx log_2 M$ and $K \approx 2\lambda$. Therefore, m chosen is 10. Note that this test is only suitable for categories of data which produce sequences larger than 10^6 bits.
- Maurer's Universal test: the value of L must be in the range of $6 \leq L \leq 16$, $Q = 10.^{2L}$ and $n \geq (Q+K)L$ where $K = [n/L]-Q \approx 1000.^{2L}$. Therefore, L and Q chosen are 7 and 1,280 respectively. Note that this test is only suitable for categories of data which produce sequences larger than 904,960 bits.
- Linear Complexity test: the value of M must be in the range of $500 \leq M \leq 5000$ and $N \geq 200$. Therefore, M chosen is 2,000 for categories of data with $n \geq 10^6$ and 600 for categories of data with $n < 10^6$.
- Serial and Approximate Entropy tests: m and n selected such that $m < [log_2 n] – 2$. Therefore, m chosen is 2.

ACKNOWLEDGMENT

Firstly, I would like to express my sincere gratitude to my supervisor, Prof. Dr. Kamaruzzaman bin Seman and my co-supervisor, Prof. Dr. Norita binti Md Norwawi for their support, guidance motivation, and immense knowledge. Their guidance has helped me in all the time of research and writing of this thesis. A special thanks also goes to my fellow colleague in Cryptography Development Department of CyberSecurity Malaysia for the advice, recommendation and knowledge shared which helps me a lot during the analysis of this research.

REFERENCES

[1] A.A. Zakaria, N.A.M. Yusof, W.Z.W. Omar, N.A.N. Abdullah and H.A. Rani, "Analysis of Steganography Substitute System Methods Using New Testing Techniques Proposed for Strict Avalanche Criterion", International Journal of Cryptology Research, No 5 Issue 1, pp. 61-76, 2015.

[2] A. Rukhin, J. Soto, J. Nechvatal, M. Smid, E. Barker, S. Leigh, M. Levenson, M. Vangel, D. Banks, A. Heckert, J. Dray and S. Vo, "A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications", in NIST Special Publication 800-2, 2001.

[3] J. Soto, "Randomness Testing of the AES Candidate Algorithms", in NIST Interagenc Reports, NISTIR 6390, 1999.

[4] J. Soto and L. Bassham, "Randomness Testing of the Advanced Encryption Standard Finalist Candidates", in NIST Interagency Reports, NISTIR 6483.

[5] Lynn Hathaway, "National Policy on the Use of the Advanced Encryption Standard (AES) to Protect National Security Systems and National Security Information", in *The Committee on National Security Systems (CNSS)*, June 2003.
<http://csrc.nist.gov/groups/ST/toolkit/documents/aes/CNSS15FS.pdf>

[6] N.A.N. Abdullah, K. Seman and N.M. Norwawi, "Statistical Analysis on LBlock Block Cipher", in International Conference on Mathematical Sciences and Statistics 2013, Selected Papers, pp. 233 - 245, 2013.

[7] N.A.N. Abdullah, N.H. Lot and H.A. Rani, "Analysis on Lightweight Block Cipher, KTANTAN", in Proceedings of Information Assurance and Security (IAS), 2011 7th International Conference, pp. 46 – 51, 2011.

[8] N.H. Lot, N.A.N. Abdullah & H.A. Rani, "Statistical Analysis on KATAN Block Cipher", in Proceedings of Research and Innovation in Information Systems (ICRIIS), 2011.

[9] P.P Mar and K.M Latt. "New Analysis Method on Strict Avalanche Criterion of SBoxes", World Academy of Science, Engineering and Technology 24, pp. 150-154, 2008.

[10] S. Kavut and M.D. Yucel. "On Some Cryptographic Properties of Rifndael", Information Assurance in Computer Networks, LNCS, vol. 2052/2011, pp. 300-311, 2001.

[11] W. Wu and L. Zhan, "LBlock: A Lightweight Block Cipher", Lecture Notes in Computer Science: Applied Cryptography and Network Security, vol. 6715, pp. 327 – 344, 2011.