RESEARCH ARTICLE OPEN ACCESS

New DNA Based Dynamical S-Box for Block Cipher

Chng Chern Wei[1], Sharifah Md. Yasin[2], Mohd. Taufik Abdullah[3] And Nur Izura UDZIR[4]

[1-4]Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, 43400, Serdang,Selangor.
Corresponding Author: Chng Chern Wei

ABSTRACT
DNA based technique has grown rapidly among researchers in introducing the latest network security algorithms, which can enhance the strength of the current cryptosystem. DNA-based techniques are able to provide a high degree of cryptography algorithm. This article discusses a new DNA based Dynamical S-Box for the symmetric keys block ciphers. The DNA based Dynamical S-Box is proposed using a polynomial calculation in producing an unknown DNA sequences of {A,C,G,T} in mapping the S-Box table in the form of metric [16x16]. The National Institute of Technology and Technology (NIST) 15 Test is used to verify the cipher for this DNA based Dynamical S-Box. However, S-Box testing criteria have been used to verify the crypto security vulnerabilities by the new DNA based Dynamical S-Box. The simulated end result shows that the proposed DNA based Dynamical S-Box can provide a good level of safety after the NIST Randomness Test shows the value is random.

Keywords - Block Cipher, AES, Biology Chemistry, DNA, S-Box, NIST

Date of Submission: 16-07-2018 Date of acceptance: 30-07-2018

I. INTRODUCTION

A. Stream Cipher vs Block Cipher

Stream cipher is a technique to conclude a symmetric key with plaintext coupled with a stream of digit pseudorandom cipher (keystream). In the stream cipher, each plaintext bit is encrypted one bit by one bit with bits corresponding to the main stream, to provide a bit of ciphertext flow[24][24].

Block Cipher is an plaintext/ciphertext split in a block that uses a tough symmetric key that operates on a fixed set of bits; this is named as a block, with unchanged any of transformations. A plaintext is split into a fixed-length block with the same block sizes of each block split [23][24]. Encrypt a block of plaintext to produce the same sized of the block ciphertext. Typically the block sizes are 64, 128, 192 and 256 bits, as mention by Hamdan et al, [25].

However, larger plaintexts require an operation mode for the perfect encryptions.

They are five modes are Electronics codebook mode (ECB), Cipher-block chaining (CBC), Counter mode (CTR), Output Feedback modes (OFB) and Cipher Feedback (CFB). The Electronic Codebook (ECB) is the easiest encryption mode with the ECB mode that provides encryption parallelizable, decryption parallelizable both in 128 bits and Memory pointer support [29]. This mode will divide the message into more slick blocks and each block will run the encryption process separately [29].

The ideal of block cipher is allows a maximum numbers of possible encryption mappings from the plaintext block. Examples of Block Ciphers are DES, Triple-DES, Blowfish Algorithm and Advanced Encryption Systems (AES) [1][2] [25].

B. DNA

DNA (Deoxyribonucleic Acid) consists of two Polynucleotide Strands (nucleotide polymer), which looks like a ladder. The Nitrogen base in DNA is keeping the instructions to create a polypeptide chain, essentially coding is for every feature of an organism.

Both polynucleotide strands will run 'antiparallel' with each other, with the projected Nitrogenous Base. The term 'antiparallel' here means the strand runs in the opposite direction, parallel to each other. The antiparallel twist in the complete DNA structure will form the Double Helix [7][8][14].

This strand is jointly held by the Hydrogen Bond between the Nitrogenous Bases which is contrary to each other. The bases joined together are called 'paired', and are specific to their Base who will join together. A Purine will only pair with Pyrimidine [12]. However, Adenine Purine will only be paired with Thymine Pyrimidine (A-T), and Guanine Purine will only be paired with Cytosine

Pyrimidine (G-C). This base pair is known as Complementary Base Pairings [6].

In the first section of this paper, the authors' discuss about the related work. The methodology discussed in section II and section III discuss about the proposed S-box that be dependent on DNA sequences of {A,C,G,T}. Finally, we will discuss about the results in the section IV and conclusions discussed in section V.

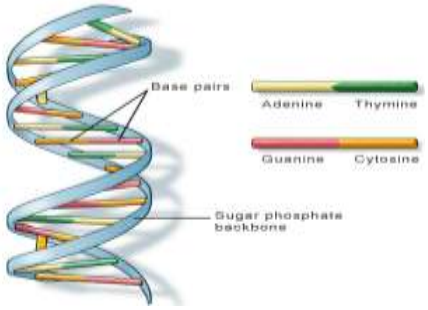

Fig.1. Dna Double Helix [6]

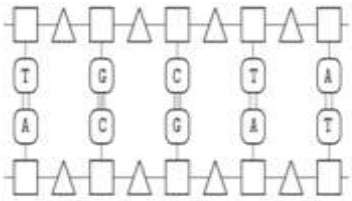

Fig.2. Complementary Base Paring [6]

II. RELATED WORK

Block cipher is rely on S-boxes to provide a more secure and safe cryptographic result. The S-Box has no connection with the secret key [27]. The Non-linear components of S-Box criteria for AES that can be provide important aspects in order to strengthen the cryptography resources [27]. This is the discussion about the related work done by passed researchers.

Kazlauskas & Kazlauskas (2009) introduced key-dependent S-box generation algorithm in AES block cipher system. The advantage of this approach is that a large amount of S-box can be generated by simply changing the secret key. This approach will lead to the creation of strong block ciphers and able to solve the problem of fixed S-box structure, and will increase the security level of the block cipher cryptosystem [27].

In the same year of 2009, Mohammad al. et., proposed the new S-Box development by A key-dependent S-box of AES algorithm using a variable mapping technique. This approach highlighted that, the number of rounds for VMS-AES is 7 laps versus 10 laps for AES. Although the number of encryption rounds has been reduced, this will not affect the security of algorithm and will increase the QoS for VoIP transmission [32].

In 2010, Lai, XueJia et al. [30] proposed a DNA sequences as asymmetric encryption and digital signature methods with DNA matrix generated for encrypt an image. The DNA sequence matrix is divided by into the small block and performs an operation of the additions between the blocks. In this work, the original image will be contrasted by addition and complementary operations. Therefore, this technique will provide a large secretive and high sensitive secret key to the secret key of the encryption algorithm. This algorithm provides defense strengths against statistical attacks and differential attacks.

Hamdy et al. (2011) in his paper proposed that a customized version of the AES block cipher by using key-dependent S-box generation methodology. This implementation generated by RC4 algorithm and used a different primitive polynomial for Key Expansion and Mix Column Transformations. Resultance had shown the positive on the NIST test. The linear and differential cryptanalysis was proved that strengths of the cryptosystem [32].

El-Sheikh et al. (2012), introduced an approach for designing a key-dependent S-box by using GF (2^4) in AES algorithm. Sheikh implemented a small s-box by using a small value of irreducible polynomials with GF (2^4). With this implementation, the number of rounds for encryption and decryption will be reduced and the strength of the cryptosystem has not affected.

Juremi et al. (2012) highlighted that the S-box rotation method is able to produce a new AES-like design for key-dependent in AES. In this paper discuss that the small modification on the key-dependent for the S-Box on the original AES will enhanced the strength of the AES algorithms to become a new Block Cipher cryptosystem. The author verifies the strength of the proposed algorithm by using NIST 15 Test and Cryptanalysis attack [31].

Mahmoud et al. (2013) in his paper discuss that an implementation of dynamic AES-128 cryptosystem with a key-dependent S-box. In this paper, highlight that, this technique is suitable for substituting the secret keys in the insecure media transmissions in order to achieve the secure transmission objectives. This technique provided high degree of avalanche effect and balance degree of correction factor [25]. Das et al. (2013) in his research introduced that a new S-box generation by using various modulus and additive constant polynomial.

Das et al., emphasizes that researchers can use a variety of polynomials to produce an S-Box that is in its own choice from a large set of

polynomials, and this is to prevent attacks from linear and different cryptanalysis [3].

However, Al-Wattar et al. (2015), in his paper introduced a new DNA – Based S-Box. In his paper discussed that the S-Box is performed by [8 x 8] matric from with DNA – Based component. However, this technique will cause high usage of memory once performed an S-Box for the cryptosystem [2].

This paper framework that the authors' study into the design of a new DNA method with a modulus and additive constant polynomial, which no researchers investigated before.

The modeling hypotheses indicate that the proposed algorithm has a satisfactory cryptographic strength, with an algorithmic property that able to withstand the linear and differential cryptanalysis, against known S – boxes.

III. METHODOLOGY

A. NIST STATISTICAL TEST SUITE

NIST Statistical Test Suite is an important verification tools for randomness analysis. The NIST test suite consists of 15 empirical test sets, mainly to test the continuous of binary sequence (Bitstreams).

The NIST Test Suite been used to evaluate the randomness of the cryptosystem algorithm for the block cipher in the properties of confusion and diffusion. Test results are convincing if there are 15 sets of NIST Randomness p-values of the empirical tests to produce randomly.

B. NIST 15 Empirical Test Suite includes [1] [2] [10][11]:

Test 1: Frequency

This test is intended to test the binary bit sequence of values 0s and 1s is completely random in a sequence. The sequence in this test environment is randomized if the binary bit rate 0s and 1 is half the total proportion.

Test 2: Frequency within a Block

This test is to determine whether the frequency block size remains around (block size) / 2, as expected in the environment. Therefore, these are randomly estimated.

Test 3: Runs

This will take into account that the number of complete runs in that sequence. Runs are a continuous sequence with the same bit. The purpose of the test is to solve if the number of run for zero and one bit is of different size as guessable in a random order.

Test 4: Longest run of ones in a Block

The test emphasizes the longest run in a long bits length of block with fixed bits. The objective of this test is to determine whether the longest run term of the test sequence is homogeneous with a predicted random sequence.

Test 5: Binary Matrix Rank

This test is concerned with the rank of disjoint sub-matrices of the entire sequence. The objective of this test is to confirm the linear reliance among fixed size substrings of the genuine sequence.

Test 6: Spectral

It discusses the height of the Fourier Discrete Transformation for a sequence. The objective of this test is to know about the periodic characteristics of a revised sequence that will determine the deviation in terms of random assumptions. However, the ultimate goal is to determine whether the peak number exceeds the limit of 95% wider than 5%.

Test 7: Non-overlapping Template Matching

This is in relation to the number of predefined target of series occurrences. The purpose of the test is to expose the generator that creates too many instances of the non-episodes model set.

Test 8: Overlapping Template Matching

It is similar to Non-Overlapping Template Matching Test; one of the differences between Non-overlapping Template Matching with Overlapping Template Matching test is when the pattern is detected, the block is glides slightly and appears to be continuous.

Test 9: Maurer's Universal Statistical

This discusses the number of bits in between (matching) for traces the same pattern. The objective of the test is to determine whether the sequence can be substantially compressed without involving the loss of information. The arrangement that can show compression is believed to be non-random sequence.

Test 10: Linear complexity

This is concern about the sizes of the Linear Feedback Shift Register (LFSR). The objective of this test is to determine whether the order is complex enough to be considered random and has the necessary complexity.

Test 11: Serial

This emphasizes the frequency of all k-bit patterns that are likely to overlap in a sequence. The purpose of this test is to determine whether the total event of 2kk-bit overlapping pattern is the same as it should in a random order.

Test 12: Approximate Entropy

This takes note of the frequency of all overlapping k-bit patterns in a sequence. The objective of this test is the comparison between the overlapping blocks of frequencies for two consecutive / adjacent lengths (k & k + 1) as opposed to expected results in random order.

Test 13: Cumulative sums (Cusum)

This focuses on the highest excursion (zero) random step, defined by the amount of accrued accruals (-1, +1) digits in sequence. The purpose of the test is to determine whether the cumulative number of partial sequences that occur in the sequence being tested is a very large or small size, with respect to the action predicted of the cumulative sum for a random sequence.

Test 14: Random Excursions

This takes into consideration the number of cycles that have the exact K-ride in cumulative random numbers cumulatively. The purpose of the test is to check whether the number of visits to a particular situation in the cycle deviates from what is expected for random order.

Test 15: Random Excursions Variant

This discusses the total number of situations that occur in cumulative random accumulations. This test is intended to test the difference in the number of visits expected to various states in a random excursion. This test should be carried out in a series of eighteen test reactions.

IV. PROPOSED METHOD

In this paper, we proposed a S-Box based on DNA is considered reliable after satisfy the S-Box Test Criteria. S-Box Test Criteria must able to performed an invert-ability, complexity in the algebraic production in the Galois Field of GF(2^8), able to form in non-linearity, have an fair value of avalanche effect, constancy and reliability to differential crypto analysis[2][17].

However, the S-Box is considered as good [2] if satisfy Strict Avalanche (SAC) as follow:
Minimize the correlation of the merger between linear input bytes and the linear combination of output bytes. A function of S : $\{0,1\}^m \rightarrow \{0,1\}^m$ for all input, **u** and output, **v** ∈ (1,2,….m), is consider to satisfy the SAC. The S-Box able to provide a great security in protection if this criterion is practiced [2]:

For all input, **u** and output, **v** forms an equation:

$$Strict\ Avalanche\ Effect = \frac{1}{2^m} = \left[\oplus\left(u_{output.v}^{input.Cu}\right)\right] = \frac{1}{2} \quad (1)$$

According to Al-Wattar al el., highlighted that if completing one input, **u** bit, it will also cause an alter of output, **v** bit with an occasion of 50 percent [2].

Experiments measure the new DNA based Dynamical S-Box by using the S-Box Test Criteria. Experiments measure the level of security of the cryptosystem algorithm that applies proposed S-Boxes by using the NIST empirical Randomness 15 test suite. Experiments measure the key sensitivity. Experiment to analysis the information entropy.

The NIST empirical Randomness 15 Test Suite 128 bits block of plaintext and 128 bit key was generated for the experiments by using Intel Pentium i7 microprocessor with 4GB of RAM, running in Ubuntu 12 platform. Experiment conduced with breaking the plaintext into few blocks. Each block will encrypt differently using the same key. The block cipher is running in the mode of ECB only. The experimental conducted in text file, image file and video file of plaintext. The minimum sequences of each file must generated 1,000,000 bits in plaintext with 128 sequences by using the NIST Randomness test. To conduct this experiment, 1,064,000 bits which size file was 133KB simulated in 128 sequences.

The NIST Randomness test of this experiment conducted in 3 rounds for the proposed algorithm and each round will produce a p-value. Each p-value interpreted of a randomness statistical test value for particular block cipher. According to Rukhin & Soto [15], the block cipher is a combination of the binary bits sequence.

As proposed by the NIST, the significant level, α is 0.01. Al-Wattar al et., in his research highlighted that the p-value can be interpreted as follow [2]:

i. p-value = 0, indicate that the block cipher sequences are shown perfectly not random.
ii. p-value ≥ 0.01, indicate that the block cipher sequences are shown random.
iii. p-value < 0.01, indicate that the block cipher sequences are shown not random.
iv. p-value = 1, indicate that the block cipher sequences are shown perfectly random.

The p-values passed the conditions must be above the value of the proportional calculation as described in the following equation [2]:

$$p_\alpha = (1 - \alpha) - (3)\sqrt{\frac{\alpha \times (1 - \alpha)}{n}} \quad (2)$$

Where:

p_α - Proportion Value
α - Significant value
n – Numbers of testing sequences

V. RESULTS AND DISCUSSION

A. NIST Randomness Test

This section discusses the analysis and assessment of NIST Randomness test from the proposed S-Box.

The proposed S-Boxed is from an irreversible {TCTCAC} polynomial and {TTTC} from an additional constant by converting to the DNA of the Code of the book {A, G, C, T}.

The NIST Randomness 15 test Results shown in Figure 5. From the Figure 5, NIST Randomness 15 test results show random.

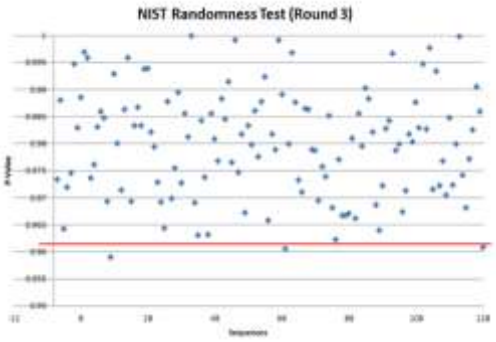

Fig. 5. NIST Randomness Test for Round 3.

The proportion of the sequences that exceeded a particular statistical test must be greater than the proportion value, p_α.The proportion of experiment executed has shown that the value is greater than the statistical test of proportion value, p_α. The proportional equation as defined in equation (2) and the proportion value of this experiment as:

$$p_\alpha = (1 - 0.01) - (3)\sqrt{\frac{0.01 \times (1 - 0.01)}{128}}$$

p_α = 96.36%

Where, n is the sample size of 128 and the Significant value of α is 0.01.

Figure 5 demonstrate the randomness test for 15 empirical tests for Round 3 of AES block cipher that used the proposed S-Box. From this diagram, at the end of the 3rd round, 98% of the 41 empirical tests fall over 96.36%. This provided an evident that the output from the algorithm is completely random.

B. Linear and Differential Cryptanalysis

The process of generating the proposed S-Box does used the mathematical operations that were employed in generating the AES S-Box, including the inverse multiplication and all operations related to it.

Also the use of DNA segments makes the attempts to attack the S-Box and cipher look like a hard job or infeasible for the cryptanalyses, since it is difficult for him to recognize or anticipate the DNA sequence and their construction and structures.

The number of modulus and additive constant polynomials are able to produce the secured S-Box with high degree of randomization of block cipher, which highly preventing the linear and differential attacks.

VI. CONCLUSSION

This paper proposes and discusses a new method for producing a new DNA Dynamical S-Box by applying the DNA techniques and concepts.
The proposed method applies the characteristics of DNA in producing the latest S-Box and satisfying the S-Box criteria with polynomial mathematical computation.

The proposed S-Box is suitable for use in block ciphers such as the AES algorithm.
The type of data used to test in the experiment against the strength of the proposed algorithm is like Text, Video and Image, and as well as random generator, BBS. Such data types may be considered as one of the most important data types in terms of Encryption and Decryption.

This proposed DNA based Dynamical S-Box is tested by using the NIST Suite and S-Box test criteria. Experimental results and test results have shown that this proposed DNA S-Box has strong security endurance.

For future work, some modifications can be made to the proposed S-Box to produce a hybrid-dynamic DNA S-Box with DNA techniques.

Finally, this work is open for continue research to create the latest algorithm with various features in DNA methodology and adapting it in cryptographic fields.

REFERENCES

[1]. A. M. Alabaichi, A Dynamic 3D S-Box based on Cylindrical Coordinate System for Blowfish Algorithm, Indian Journal of Science and Technology, vol. 8, no. 30, pp. 1-17, 2015.
[2]. A. H. Al-Wattar, R. Mahmod, Z. A. Zukarnain and N. I. Udzir, A New DNA-Based S-Box, International Journal of Engineering & Technology, vol. 15, no. 4, pp. 1-9, 2015.
[3]. S. Das, J. K. M. S. Uz Zaman and R. Ghosh, Generation of AES S-Boxes with Various Modules and Additive Constant Polynomials and Testing

Their Randomization, Procedia Technology, vol. 10, pp. 957-962, 2013.
[4]. R. Hosseinkhani and H. H. S. Javadi, Using Cipher Key to Generate Dynamic S-Box in AES Cipher System, International Journal of Computer Science and Security, vol. 6(1), 2012.
[5]. A. Jain, P. Agarwal, R. Jain and V. Singh, Chaotic Image Encryption Technique using S-Box based on DNA Approach, International Journal of Computer Applications, vol. 92, no. 13, pp. 30-34, 2014.
[6]. C.W.Chng, DNA approach for password conversion generator, IEEE, 2014.
[7]. J.Shipra & B. Vishal, Anology of Various DNA Based Security Algorithms Using Cryptography and Steganography, IEEE, 2014.
[8]. J.Shipra & B. Vishal, A Novel DNA Sequence Dictionary method/or Securing Data in DNA using Spiral Approach and Framework 0/ DNA Cryptography, IEEE, 2014.
[9]. D. Canright, A very compact s-box for AES, ACM: Springer-Verlag Berlin, 2005.
[10]. J.K.M. Sadique & G.Ranjan, Review on fifteen Stitistical Tests Proposed by NIST, Journal of Theoretical Physics & Cryptography: Vol.1, Nov. 2012.
[11]. L. E. Bassham III, et al., SP 800-22 Rev. 1a. A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications, NIST: 2010.
[12]. Y.P. Zhang & B.C. Fu, Research on DNA Cryptography, Applied Cryptography and Network Security, Dr. Jaydip Sen (Ed.), 2012, ISBN: 978-953-51-0218-2
[13]. Beenish Anam et al., Review on the Advancements of DNA Cryptography, eprint arXiv:1010.0186, 10/2010
[14]. Leier A et al., Cryptography with DNA binary strands [J]. Biosystems Vol.57(1), 2000.
[15]. Rukhin A., Soto J., et al., A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications, NIST, Technology Administration, U.S. Department of Commerce, 2008.
[16]. A. Doganaksoy, B. Ege, O. Kocak & F. Sulak, Cryptographic Randomness Testing of Block Ciphers and Hash Functions, IACR, 2010.
[17]. J. Cui, L. Huang, H. Zhong, C. Chang & W. Yang, An Improved AES S-Box and Its Performance Analysis, International Journal of Innovative Computing, Information and Control, Vol.7(5), 2011.
[18]. A. Leier, C. Richter, W. Banzhaf & H. Rauhe, Cryptography with DNA Binary Strands, BioSystem, Vol.57(1), 2000.
[19]. I. Das, R. Sanjoy, N. Subhrapratim & M. Subhash, Random S0Box Generation in ASE by Changing Irreducible Polynomial, Meghnad Saha Institute of Technology, 2013.
[20]. C. Easttom, Modern Cryptography, New York City, New York: McGraw Hill, 2015.
[21]. S.Sinha & C. Arya, Algebraic Construction and Cryptographic Properties of Rijndael Substitution Box, Defence Science Journal, vol. 62(1), 2012.
[22]. O.A.Hamdan, B.B.Zaidan, A.A.Zaidan, A.J.Hamid, M.Shabbir & Y. Al-Nabhani, New Comparative Study Between DES, 3DES and AES within Nine Factors, Journal of Computing, vol. 2, no. 3, March 2010.
[23]. M.A. Daniyal, H.H.Syed & S.T.Mohamed, Stream Ciphers: A Comparative Study of Attacks and Structures, International Journal of Computer Application, Vol.83 (1), December 2013.
[24]. C. Paar, J. Pelzl, Understanding Cryptography, Springer-Verlag Berlin Heidelberg, 2010.
[25]. E.M.Mahmoud, A.A.Hafez & A.Talaat, Dynamic AES-128 with Key-Dependent S-Box, International Journal of Engineering Research and Applications, Vol. 3(1), 2013.
[26]. L.R Knudsen & M.J.B.Robshaw, The Block Cipher Companion, Springer, 2011.
[27]. K.Kazlauskas, G.Vaicekauskas & R.Smaluikas, An Algorithm for Key-Dependent S-Box Generation in Block Cipher System, INFORMATICA, Vol. 26 (1), 2015.
[28]. C.P.Ruisanchez, A New Algorithm to Construct S-Boxes with High Diffusion, International Journal of Soft Computing, Mathematics and Contorl, Vol. 4 (3), August 2015.
[29]. K.Huang, J.Chiu and S.Shen, A Novel Structure with Dynamic Operation Mode for Symmetric-Key Block Cipers, International Journal of Network Security & Its Applications (IJNSA). Vol.5 (1), 2013.
[30]. Lai, XueJia, Asymmetric encryption and signature method with DNA technology, Science China Information Sciences 53.3, page 506-514, 2010.
[31]. J.Juremi, M.Ramlan, S.Sulaiman & J.Ramli, Enhancing Advanced Encryption Standard S-Box Generation Based on Round Key, International Journal of Cyber-Security and Digital Forensics, Vol.1(3), 2012.
[32]. N.Hamdy, K.Shehata & H.Eldemerdash, Design and Implementation of Encryption Unit Based on Customized AES Algorithm, International Journal of Video & Image Processing and Network Security IJVIPNS-IJENS Vol: 11(1), 2011.

Chng Chern Wei "New DNA Based Dynamical S-Box for Block Cipher "International Journal of Engineering Research and Applications (IJERA), vol. 8, no.7, 2018, pp.64-69