International Journal of Cryptology Research x(x): 1 - 26 (20xx)

Recent Advances on the Theory of S-box Design Strategies

Nur Fasihah Mohd Esa[*1], **Shekh Faisal Abdul-Latip** [1], and **Zaheera Zainal Abidin**[1]

[1]*INSFORNET - Crypto Research Group (CRYPTREG) Centre of Advanced Computing Technology (C-ACT), Universiti Teknikal Malaysia Melaka 76100 Durian Tunggal, Melaka, Malaysia*

E-mail: fasihah.esa@gmail.com

Abstract

S-box is a nonlinear transformation that essential in cryptography which providing an excellent security in the symmetric key primitives. The important part of the design strategies of S-box is to consider the cryptographic properties of S-box which contribute to the confusion property during encryption and decryption process. The methods discussed such Algebraic construction, chaotic maps, and heuristics techniques which have their own principle and their own significant. The analysis result of performance experiment show that the method based on the design strategies has good cryptographic properties and future direction of the design of S-box is proposed at the end on this paper.

Keywords: S-box, Confusion, Bijective, Non-linearity, Strict Avalanche Criterion (SAC)

1 INTRODUCTION

S-box provides a nonlinear transformation for the iterated symmetric key primitives such stream ciphers and block ciphers(Carlet, 2010a,b). According to

Shannon's philosophy (Shannon, 1949), the security cipher itself is depending on the strength the S-box which provide *confusion* property during encryption process. An $n \times m$ S-box has mapping of *n* input bits and *m* output bits which expecting the contribution of the resistance against differential and linear cryptanalysis if the S-box is in larger size (Xiao and Heys, 2005). However, large dimension *n* leads to larger lookup table which required a costing in terms of time. The size of the lookup table decides the size of the program memory (Canright, 2005).

Therefore, the small S-box is required for the hardware with less program memory and large S-box can be used with hardware having more program memory which will perform slower speed (Canright, 2005). For example, AES uses 16×16 S-box. This is implemented in a suite of hardware platforms: 8051 based micro-controllers, PIC processor, ARM processors, FPGA based processors, ASIC, etc. It is possible to implement 256×256 S-box in high end processors (Daemen and Rijmen, 2002).

Another practical consideration is that the larger the S-box, the more difficult it is to design it properly (Webster and Tavares, 1985). S-box is required for both encryption and decryption. An $n \times m$ S-box typically consists of 2^n rows of *m* bits each. The *n* bits of input select one of the rows of the S-box, and the *m* bits in that row are the output. For example, in an 8×32 S-box, if the input is 00001001, the output consists of the 32 bits in row 9 (the first row is labeled row 0).

According to (Shannon, 1949), the purpose of *confusion* property is to make the relationship between the ciphertext and the secret key to be complex as possible which means each bit of the ciphertext should depend on all bits of the key. In contrast, diffusion refers to the influence of single bit plaintext flipping highly effect to the changing of the ciphertext bits.

1.1 Outline

First, we briefly discuss on the preliminaries of Boolean function followed by the concept of Vectorial Boolean function in Section 2 and Section 3 respectively. In Section 4, we review the significant of S-box properties which

contribute to the robustness of S-box. For instance, *Bijective*, *Non-linearity*, *Strict Avalanche Criterion (SAC)*, *Bit Independence Criterion (BIC)*, *Algebraic Degree*, *Autocorrelation*, *Differential Approximation Probabilities (DAP)* and *Linear Approximation Probabilities (LAC)*. Section 5 contains the selected recent design theory of S-box which been used by many researchers namely, *Algebraic Construction*, *Chaotic Map Technique* and *Heuristic Technique*. We present the comparison of S-box performance against the selected S-box properties in Section 6. Finally, we conclude the paper and review future work direction in Section 7.

1.2 Contribution

1. We present the state of the art of the S-box design strategies.
2. We stated the open problem for the S-box design strategies.
3. A future direction of the design strategies of S-box is proposed in the conclusion

2 PRELIMENARIES OF BOOLEAN FUNCTION

Definition 2.1. *A Boolean function, $\mathbb{B}_2$ is a set value of function f on the map of $\mathbb{B} : \{0,1\}^n \mapsto \{0,1\}^m$ for simplicity, $m = 1$ is assumed. Then the function can be represented as a binary vector $\overrightarrow{f}$ of length 2^n where $\overrightarrow{f}$ is the rightmost column of the truth table describing the function.*

A *truth table* is one of the significant representation for Boolean function in form of binary or polarity output vector of $\mathbb{B}_2 \in 2^n$ which denote as $f(x)$ and $\widehat{f}(x) = -1^{f(x)}$. To represent the algebraic degree, *deg(f)* (see 4.3) of Boolean function expression, we can use *algebraic normal form (ANF)* for the representation. ANF consist of the operation of a unique XOR sum of AND products and described as :

$$f(x) = a_0 \oplus a_1 x_1 \oplus a_{1,2} x_1 x_2 \oplus \cdots \oplus a_{1,2,\ldots,n} x_1 x_2 \cdots x_n$$

where the coefficient $a \in [0, 1]$, form the elements of the ANF truth table, $\mathbb{B}_2$. *deg(f)* is the variable numbers of the largest product term of the function's ANF having a non-zero coefficient.

Two *n*-variable of Boolean functions , $f(x)$ and $g(x)$ are affine equivalent if and only if some invertible $n \times n$ bimary matrix A are exist, such $g(x) = f(Ax \oplus b) \oplus < c, x > \oplus d$ where vectors $b, c \in \mathbb{B}^n$ and a scalar $d \in \mathbb{B}$. The *Walsh Hadamard transform (WHT)* of an *n*-variable Boolean function in polarity table form $\widehat{f}(x)$, denoted as $\widehat{F}(x)$ is defined by,

$$\begin{aligned}\widehat{F}(x) &= \sum_{x \in \mathbb{B}_2} (-1)^{f(x)} (-1)^{l\omega(x)} \\ &= \sum_{x \in \mathbb{B}_2} \widehat{f}(x)\, \widehat{l}_\omega(x)\end{aligned}$$

where $\widehat{l}_\omega(x) \in \{1, -1\}$ represent the polarity form of linear function denote for $\forall \omega \in \mathbb{B}_2^n$ and $\widehat{l}_\omega(x)$ is the *signed* function of the linear Boolean function $l_\omega(x) = \sum_{i=1}^{n} \omega_i x_i$.

3 VECTORIAL BOOLEAN FUNCTION

Definition 3.1. *Let multiple output of Boolean functions from $\mathbb{B}_2^n \times \mathbb{B}_2^m$, where m and n be two positive integers.*

Vectorial Boolean function is the extended of Boolean function which have the same conceptual in the transformation from the single output to the multiple output. However, there are difference in properties manner in S-box and Boolean function respectively. While considering the cryptographic properties, it is important to consider the linear combination of Boolean functions instead of only considering their coordinate.

To resist trivial attacks, a good cryptographic S-box should be balanced. If the S-box is imbalance, some output would appear more often than others when the input to the S-box is randomly chosen and the bias can be exploited

by the cryptanalyst. An $(n \times m)$ with $n \geq m$ is regular if and only if all non-zero its component Boolean function is balanced.

Matsui (Matsui, 1993) introduced a known-plaintext attack, *linear cryptanalysis* analyse the approximation of the relationship between plaintext, ciphertext as well as key bits. Matsui proposed the construction of a linear expression and then the probability *P* is evaluated. The best linear or affine approximation is chosen by those highest or lowest probability among all possible expression. Therefore, the cipher is resisted to the linear and affine approximation if and only if all the possibilities P are approximately close to $\frac{1}{2}$.

4 CRYPTOGRAPHIC PROPERTIES OF S-BOX

There are many features that can determine the strength of the S-Box. Jakimoski and Kocarev (Jakimoski et al., 2001) demonstrated the strategies of the properties selection which can measure or prove the strength of an $n \times n$ S-Box.

4.1 Bijective

Boolean function, f_i such that $wt\,(\sum_{i=1}^{n} a_i f_i) = 2^{n-1} \mod 2$ where $a_i \in \{0, 1\}$ and *wt* is the hamming weight, which the linear combination that possible input vector mapped into unique output vector (Adams and Tavares, 1990). Bijection property requires a one-to-one and onto mapping from input vectors to output vectors if the S-box is *m* by *n* bit.

4.2 Non-linearity of S-boxes

Nonlinearity of S-box S, is defined as $NL(S) = min\ NL(f)$, $f \in LC$, where LC represent the set of all linear combinations of the columns S. We note that nonlinearity is equal to minimum Hamming distance, H_d between all the

component output Boolean function of S-box and all affine function on n-variables. Thus, we can closely relate this generalization with linear attack due to it's role which tend to find the opportunity that can easily approximate the S-box structure by the set of linear equation. Besides, the non-linearity of S-box will never change if we add the function, f an affine function and the nonlinearity is clearly shows a left and right invariant. According to Seberry in (Seberry et al., 1993), the nonlinearity of each of the component Boolean function should be high as possible.

4.3 Algebraic Degree of S-boxes

The algebraic degree of the S-box (and similarly Boolean function) is need to be as high as possible in purpose to resist a cryptanalytic attack which is known as *low order approximation* (Golić, 1996, Millan et al., 1999). We can measure the (minimal) algebraic degree of an $(n \times m)$ S-box, denoted by $deg(S_{n,m})$ as shown in definition below.

Definition 4.1. *(Adane, 2013) Let $S = (f_1, f_2, f_3, \ldots, f_m)$ be an $n \times m$ S-box where $f_i(i = 1, 2, \ldots, m)$ are n-variable of non-trivial Boolean functions. Let g_j be the set of linear combinations of Boolean functions. Then, the algebraic degree of S, is defined as*

$$deg(S_{n,m}) = min_g\{deg(g_j)\}(j = 1, 2, \ldots 2^m - 1) \tag{1}$$

4.4 Autocorrelation of S-box

The autocorrelation of S-box is the maximum autocorrelation of all linear combination of each Boolean function of an S-box.

Definition 4.2. *(Adane, 2013) Let $S = (f_1, f_2, \ldots, f_m)$ be an $n \times m$ S-box where $f_i(i = 1, 2, \ldots m)$. Let A_i be the set of linear combination of each Boolean function of an S-box. Then the autocorrelation of an S-box denoted by $AC(S_{n,m})$ is defined as*

$$AC(S_{n,m}) = max_A r(A_i)(i = 1, 2, \ldots, 2^m - 1) \tag{2}$$

4.5 Strict Avalanche Criterion (SAC) of S-box

Webster and Teveres (Webster and Tavares, 1985) introduced strict avalanche criterion combining the property of avalanche and completeness of Boolean function. SAC is achieved if and only if there have any slight changes in the input vector, there will be a significant change in the output vector. If one input bit *i* is changed, each output bit will change with probability of one half. To achieve this effect, we will need a function that has an approximately 50% dependency on each of its *n* input bits.

$$\begin{aligned} wt(x \oplus e) \oplus f(x) &= \sum_{k=0}^{2^n-1} [f(x^k \oplus e) \oplus f(x^k)] \\ &= 2^{n-1} \end{aligned} \tag{3}$$

where affine function $e \in \mathbb{B}_2^n$ with hamming weight $wt(e) = 1$ and $\oplus$ represent the XOR operation.

4.6 Bit Independence Criterion (BIC)

Bit Independence Criterion of S-box requires a Boolean function, $f : \{0,1\}^n \rightarrow \{0,1\}^m$ satisfies the BIC if for all $i, j, k\varepsilon\{1, 2, ..., n\}$ with $j \neq k$, the invertible input i the output bits j and k act independently from each other's in order to avoid any statistical pattern or statistical dependencies between output bits from the output vectors. The Bit independence criterion (BIC) analyzes all the avalanche variables and determined the extent of their pair-wise independence in reference to a given set of avalanche vectors. A bit independence parameter corresponding to the effect of the ith input change on the jth and kth bits of avalanche vector,$A^e i$ defined as

$$BIC(a_j, a_k) = max_{1 \leq i \leq n} \, |corr(a_j^{ei}, a_k^{ei})| \tag{4}$$

Therefore, the Bit Independence Criterion (BIC) parameter for the S-box function f is then found as:

$$BIC(f) = max_{1 \leq j,k \leq n, j \neq k} \, BIC(a_j, a_k) \tag{5}$$

The equation 5 indicate the function (f) is closely satisfying the BIC which take value $[0, 1]$. Literally, 0 is an ideal case and if the worst case, BIC is equal to 1.

4.7 Linear Approximate Probability (LAP)

The Linear approximation probability is defines as the maximum value of the imbalance of an event. The uniformity of the input bits is selected by mask Γ_x equivalence to the uniformity of the output bits selected by Γ_y. This property was originally define by Matsui (Matsui, 1993) or known as probability of bias. LAP of the given S-box is given as,

$$LAP = max_{\Gamma_x \Gamma_y \neq 0} \left| \frac{\#\{x \in X / x \bullet \Gamma_x = S(x) \bullet \Gamma_y\}}{2^n} - \frac{1}{2} \right| \tag{6}$$

where Γ_x as the input and Γ_y as the output masks, X represents the set of all possible inputs; and 2^n is the number of elements. Therefore, the immunity of an $(n \times m)$ S-box against linear cryptanalysis is improved if the LAP of all the entries magnitudes are as small as possible (Matsui, 1993).

4.8 Differential Approximate Probability (DAP)

Differential approximation probability (DAP) is a measurement for differential uniformity. In every nonlinear transformation of S-box, there have differential uniformity. An input differential Δx_i is supposed to be mapped uniquely with an output differential y_i, hence ensuring ta uniform mapping for each i. The DAP is defined as:

$$DAP(\Delta x \rightarrow \Delta y) = \left[\frac{\#\{x \in X / S(x) \oplus S(x \oplus \Delta x) = \Delta y\}}{2^n} \right] \tag{7}$$

where X is the set of all possible input values and 2^n represent the number of its element. Therefore, the immunity of an $(n \times m)$ S-box against differential cryptanalysis is improved if the DAP of all the entries magnitudes are as small as possible (Biham and Shamir, 1991).

5 DESIGN STRATEGIES OF S-BOX DESIGN

To construct a strong cryptographic S-box, many researchers proposed a variant of techniques to fulfill the design criteria namely Algebraic Construction, Chaotic Map technique, Heuristic technique and Pseudo-random S-box generation. In this section, we review some of relevant work presented by researchers in the literature. At the end of this section, we do a summarizes of the whole techniques of S-box construction in Table 1.

5.1 Algebraic Construction

Algebraic construction is based on mathematical principles such using formula or any transformation to contribute the confusion S-box generation. Most of the ciphers are employing the algebraic construction to provide a complex computation which intends to enhance the development of stronger S-boxes. In previous work, many algorithms consist of algebraically complex and strong cryptographic properties e.g. AES, Skipjack, Gray and Residue Prime S-boxes (Abuelyman and Alsehibani, 2008, Daemen and Rijmen, 2002, Knudsen and Wagner, 2001, Tran et al., 2008).

Daemen and Rijmen (Daemen and Rijmen, 2002) designed Advance Encryption Standard, AES by using multiplicative inverse over the Galois Field $GF(2^8)$ with irreducible polynomial $m(x) = x^8 + x^4 + x^3 + x^+1$. Daemen and Rijmen (Daemen and Rijmen, 2002) carefully selecting this irreducible polynomial due the first list out of 30 irreducible polynomials. Besides that, Gangadari, Bhoopal and Ahamed (Gangadari and Ahamed, 2015) analysed algebraic construction of AES S-box in their papers. Gangadari et al found that affine transformation is necessary in AES design in order to increase the complexity of the encryption process and it will become more complicated expression (Gangadari and Ahamed, 2015). As the result shown in (Daemen and Rijmen, 2002), (8×8) S-box of AES has high algebraic degree of 7, provide high nonlinearity of 112, low autocorrelation is 32 and low differential uniformity of 4. Thus, AES S-box become a preferred S-box due to their excellent cryptographic properties. Although AES S-box have good criteria in S-box design, the algebraic construction of AES is simple yet has possibility for future

vulnerability to algebraic attack. In addition, the number of S-boxes is small and all of them are affine equivalent (Courtois and Pieprzyk, 2002). Furthermore, according to Ivanov et al (Ivanov et al., 2016), algebraic construction is not good enough to generate larger size of S-box . Maximum Distance Separable (MDS) and Maximun Distance Binary Linear (MDBL) are the choices of many block ciphers for their diffusion layer(Aslan and Sakallı, 2014).

Many researchers are concerning to the improvement of S-box design quality using algebraic technique. Hussain *et al*., (Hussain et al., 2011), used $\frac{17z+13}{29z+41}$ fractional linear transformation which is the one of an action of projective general linear group, PGL$(2, GF(2^8))$ applied on Galois field $GF(2^8)$, where $(z \in GF(2^8)$. However, they still need an improvement of non-linearity average of 104, eventhough the authors managed to propose a good criteria of S-boxes. In a few years, Hussein et al (Hussain et al., 2013) took an effort by adopting the same algebraic method as in (Hussain et al., 2011) and the only difference is the fractional linear transformation. According to the proposed algorithm, The linear fractional transformation used in the construction of S-boxes $f(z) = \frac{35z+15}{9z+5}$ is also formed with the action of projective linear group $PGL(2, GF(2^8))$ on $GF(2^8)$. The calculated values of $f(z)$ are replaced with their binary value equivalent, represented as some power of w, where w is the root of the primitive irreducible polynomial, $P(x) = x^8 + x^4 + x^3 + x^2 + x$: The resulting values from $GF(2^8)$ are then solved to determine the eight-bit binary value to be used in S-box. However, the average non-linearity of the proposed S-box is 104 and 106 respectively.

Recently Safraz, Hussein and Ali (Sarfraz et al., 2016) presented a technique of strong S-box construction based on Mobius transformation of $\frac{(230t+33)}{(93t+204)}$ and then, the authors modified the scheme depending to the invertible function $h(t) = mt + n$. The main purpose of this paper is to increase the rate of confusion. The authors managed to increase the nonlinearity of S-boxes, however the non-linearity did not achieving an optimum value in (Sarfraz et al., 2016) or the same value as the AES nonlinearity value. The authors presented two results based on transform and modified S-box as shown in Table 6.(Farwa et al., 2016) applied fractional linear transformation to produce a highly nonlinear S-box. For the irreducible polynomial, they selected $x^8 + x^6 + x^5 + x^4 + 1$ as the generator. Therefore, algebraic construction contribute to the strong properties S-box but it is highly vulnerable to algebraic attack. Table 6 shows the perfor-

mance of cryptographic properties of S-box in literature e.g nonlinearity, strict avalanche criterion (SAC), Bit independent criterion (BIC), Linear approximation probability(LAP) and differential approximation probability (DAP).

5.2 Chaotic Map S-box

Chaotic map is a non-linear function which involves the mixing and random-like behaviour that satisfy the confusion properties. Chaotic map is has variant of map function e.g logistic map, Baker map and Piecewise-Linear chaotic map. The properties of chaos such as stochastically, similarity to random behavior, and high sensitivity to the initial conditions are highly useful in communication networks security based on chaotic design. The randomness in chaotic system is not in a stochastic nature (Garg and Upadhyay, 2013). Jakimoski and Kocarev (Jakimoski et al., 2001) were originally introduced the design of S-boxes based on chaotic map are in 2001 with four-step including the selection of chaotic map, discretizing the chaotic map chosen, key independent with logistic and followed by exponential chaotic maps to strengthen the security. However, the chaos S-box not achieving the performance of security as good as in AES (Wang et al., 2012).

Wang *et al*, (Wang et al., 2012) improved the chaotic S-box by employing genetic algorithm which only considering nonlinear properties. Genetic algorithm is taking role as an optimization technique. This method implies generating the initial S-box by iterating S-box, while genetic algorithm is useful to search for high performance of S-box. Although this paper has good result in nonlinearity properties, but Wang *et al* did not consider other design criteria of S-boxes for instance avalanche property, BIC, Linear and Differential Approximation probability. However, this paper did not mentioned regarding the efficiency of design based on chaos and evolutionary algorithm. In 2014, Lambic (Lambić, 2014)proposed a method of S-box construction based on chaotic and composition method. Logistic map is the one of important chaotic maps to generate the S-boxes which can be written as (Rehman et al., 2016):

$$X_{n+1} = rX_n(1 - X_n) \tag{8}$$

where r is the parameter control with $r \in (0, 4)$ and X_n represent the output of random sequence. In 2015, Luma *et al*., (Luma et al., 2015) generated dy-

namical key dependent S-box based on the 2D logistic map and 2D cross map which is sensitive to the secret keys, has larger key space and contribute to the best image encryption. Rehman et al. (Rehman et al., 2016) presented a new logistic map technique with dynamic S-boxes generation which specifically for image encryption. They used generalized chaotic 2D Burgers map to reduce the autocorrelation between adjacent neighbourhood pixels by permuting the column- and row-wise.

$$x(n+1) = PWLCM(x,p) = \begin{cases} \dfrac{x_n}{p} & \text{for} \quad x_n \in (0,p] \\ \dfrac{1-x_n}{1-p} & \text{for} \quad x_n \in (p,1) \end{cases} \tag{9}$$

where $x \in (0,1)$ for all variable $n \geq 0$ with p is the control parameter.

Ahmad et al., (Ahmad et al., 2016) designed an efficient S-boxes based on the piece-wise linear chaotic map (PWLCM) (see function 9) which satisfies the boolean cryptographic criteria in order to solve the traveling salesman problem applications. The traveling salesman problem (TSP) intent to find the set of all the cities in ordered for the salesman to visit in terms cost minimization.

Baker's Map is the iteration of 8-bit sequence of binary random variable in chaotic logistic map introduced by Tang and Liao (Tang and Liao, 2005). The Chaotic Baker's map transformation is described as follows (Gondal et al., 2014) in form of Geometrical coordinates $p, q \in [0,1) \times [0,1)$ as illustrate as :

$$\begin{aligned} p' &= p + \frac{2q}{2} \\ q' &= 2q - |2q| \end{aligned} \tag{10}$$

where p represent as the area compression and q is the area stretching. The only nonlinear part in this transformation is when the unit square is divided into half and pile up the halves on top of each other.

$$x_(n+1) = \begin{cases} \lambda_a x_n & \text{if} \;\; y_n < \alpha \\ (1-\lambda_b) + \lambda_b x_n & \text{if} \;\; y_n > \alpha \end{cases} \tag{11}$$

$$y_(n+1) = \begin{cases} \dfrac{y}{\alpha} & \text{if} \;\; y_n < \alpha \\ \dfrac{y_n - \alpha}{\beta} & \text{if} \;\; y_n > \alpha \end{cases} \tag{12}$$

Gondal *et al* (Gondal et al., 2014) (see function 10) presented a scheme of S-box construction based on Baker's chaotic map. Gondal's Method also applied for different dimension $m \times n$ S-box effectively. The results shows satisfied with the design criteria. For instance, although the nonlinearity is achieved more than 100, yet the nonlinearity of S-box still need an improvement. In other research paper, *Özkaynak et al.,* (Özkaynak et al.) constructed 8-bit lookup table based on Chen's fractional-order Chaotic system given in (Li and Peng, 2004). The main reason of *Özkaynak et al*., chose Frictional-order Chen's system as their main source randomness because the function is convenient based on chaos and the S-box construction processing is much simpler than the other previous chaotic system.

Wang et al., (Wang et al., 2015) proposed the algorithm to enhance the nonlinearity of Chaotic map based and optimization concept since the nonlinearity of recent Chaotic S-box is not high sufficiently. In this paper, Wang *et al* only consider one criterion e.g nonlinearity optimization. Thus, more criterion need to be fulfilled to achieve a dynamic S-box construction. Also, Dragan Lambic (Lambić) proposed a simple algorithm or S-box construction using discrete chaotic map based on the composition of permutation equation of t he set with arbitrary number of element m [see eq (4) in (Lambić)]. On the top of that, the chaotic map represents in fully digital approach. Furthermore, this approach is suitable for large value n of $n \times n$ S-boxes. Table 6 shows the comparison of S-boxes performance according to the design criteria properties. As we can see, all of the S-box reviewed fulfilled the performance criterion. However, we still need to improve the performance in term of time complexity to achieve fast implementation instead of the strong development of S-boxes.

5.3 Heuristic Technique

Heuristic technique involve choosing a value in random and then experimenting it against a set of criteria to determine if it is well-suited for inclusion in the s-box. The main goal of heuristic technique is for s-boxes generation that meet those design criteria such the strict avalanche criteria (SAC), high non-linear properties, and have a high degree of resistance to differential cryptanalysis (Lineham and Gulliver, 2008). Heuristic techniques are able to construct bigger set than algebraic construction but the cryptographic properties of S-box is

not good as the S-box generation using algebraic construction which is stated in (Ivanov et al., 2015). To overcome the weakness of this technique, many recent researchers invented new improved heuristic technique such hill climbing method, simulated annealing method and genetic algorithm method.

Hill Climbing S-box is the method involves the application of small modifications of one or more distinct elements in order iteratively to improve one or more cryptographic properties. The highest nonlinearity achieved by this method is 100 (Millan, 1998). Whereas, simulated annealing method provides an extension to the hill climbing technique in which the search process is able to move out of a local optimum in order to continue. For the case of 8×8 S-boxes by using this method S-boxes possessing nonlinearity 102 are generated (Clark et al., 2005).

In 2015, Inovov et al. (Ivanov et al., 2015) proposed an extended heuristics algorithm with a good combination of big sets ($n \times n$) called immune algorithm. The main goal of this proposed algorithm is to achieve good cryptographic properties including low differential uniformly. This algorithm involves five steps including initialization (definition of variables), initial selection (start the modified hill climbing method), somatic hypermutation (apply the mutation function twice), selection and stopping the criterion. The experimental result shows good result in non-linearity and differential uniform (104,6). Their future work is to apply some changes in the number of the mutation functions and in the functions themselves aiming at producing S-boxes with $N > 104$ and $N = 4$ that are different from the finite field inversion-based ones.

Other than that Ivanov *et al* (Ivanov et al., 2016), also proposed a modified genetic algorithm of heuristic design known as reverse genetic algorithm. Ivanov *et al*, (Ivanov et al., 2016) worked on the large number of good bijective by S-box by proposing a new methodology using genetic algorithms in a reverse way which give a better result with the non-linearity at most at $108-112$. The motivation of their research is to enhance performance of the security of 8×8 or even larger S-boxes. Ivanov *et al* considered linear redundancy need to be zero and have complex algebraic structure due to achieve an optimal result. Ivanov *et al* proposed two genetic algorithm (GA1 and GA2). Genetic Algorithm 1 consist of three main functions; the breeding function, mutation

function and the fitness function. While in genetic algorithm 2, there have slightly modification of GA1 with the application of additional cost function together with the fitness function necessarily to determine whether the respective child will survive to the next generation or not. The cost function is define by WHT spectrum from the family of function in (Clark et al., 2005). However, the additional cost function effect the time complexity which become slower.

Isa *et al* (Isa et al., 2016) developed an alternative heuristic method called *the bee waggle dance* algorithm that involving the application of trinomial power function as their initial S-box. The main concept of this bee waggle dance algorithm is mainly based four types of information distance, direction, quality and quantity of food found. Isa *et al* considered for elements of parameters e.g; direction of dance, wriggle of dance, distance of dance, and loop. Experimental result shows that this method contribute good result of cryptographic properties specifically in nonlinearity and differential uniform in Table 6. In further studies, Isa *et al* suggested in the paper to investigate further regarding the different parameters of the algorithm.

5.4 Pseudo-random Generation

The a pseudo-random technique of S-box generation consists of the construction of S-box from a table of random numbers followed by the compliance test. However, Inavov *et al*. (Ivanov et al., 2015) was critically reviewed that this technique approach was unsuccessful at the very beginning since most cryptographic required criteria often in opposite way with each other, which greatly reduces the number of S-boxes are good with respect to all criteria and reduce the probability of good S-box selected. For instance, paper in Inovov *et al* attempt to review the highest value of non-linearity for pseudo-random method is at least $98 - 100$ for 8×8 S-box case, and these with nonlinearity 100 found were only four out of 50 million S-boxes generated.

Table 1 gives a summary of the selected techniques for S-box design strategies namely Algebraic construction, Chaotic Map S-box, Heuristics techniques and Pseudorandom technique. We summarized briefly the concept of design, type of techniques used, advantages of the design, as well as disadvantages of each techniques.

Selected Techniques	Design Concept	Type of techniques	Advantages	Disadvantages
Algebraic Construction	Based on mathematical principle.	Multiplicative inverse, fractional linear transformation, and Mobius Transformation	Have high complexity in the expression, high nonlinear property and have random behaviour.	Lack performance for larger dimension of S-box generation, vulnerable to algebraic attack
Chaotic Map	Non-linear function involves mixing and random-like behaviour.	Logistic Map, Baker's Map, PLWCM	Provide random behaviour, have high sensitivity to the initial condition of S-box	Low amount of nonlinearity
Heuristics Techniques	Experimenting a chosen random value against a set of criteria for well-suited S-box's inclusion.	Hill climbing, Simulated annealing, Immune algorithm, Bee Waggle Dance Algorithm	High amount of nonlinearity, able to construct bigger size of S-box , provide random behaviour	Cryptographic properties is not good as Algebraic construction
Pseudo-random	Construction of S-box from a table of random variable followed by the compliance test.	Pseudorandom S-box generation	The process can be easily replicated	low amount of nonlinearity 100, and only 4 S-box found from 1000 trials

Table 1: Summary of Selected Techniques of S-box Design Strategies

6 ANALYSIS OF S-BOX PERFORMANCE

In this section, we compare the result performance from the literature against significant cryptographic properties. As in Table 2, we present the comparison of S-box performance with respect to the nonlinearity, Bit Independence Criterion (BIC), Strict Avalanche Criterion (SAC), Linear Approximation Probability (LAP) and Differential Approximation Probability (DAP). From Table 1 and Table 6, graphs in Figure 1, 2, 3, 4 and Figure 5 are generated in evaluating the performance of S-box in terms of Nonlinearity, Strict Avalanche Criterion (SAC), Bit Independence Criterion (BIC), Linear Approximation Probabilities (LAP) and Differential Approximation Probabilities (DAP). Figure 1

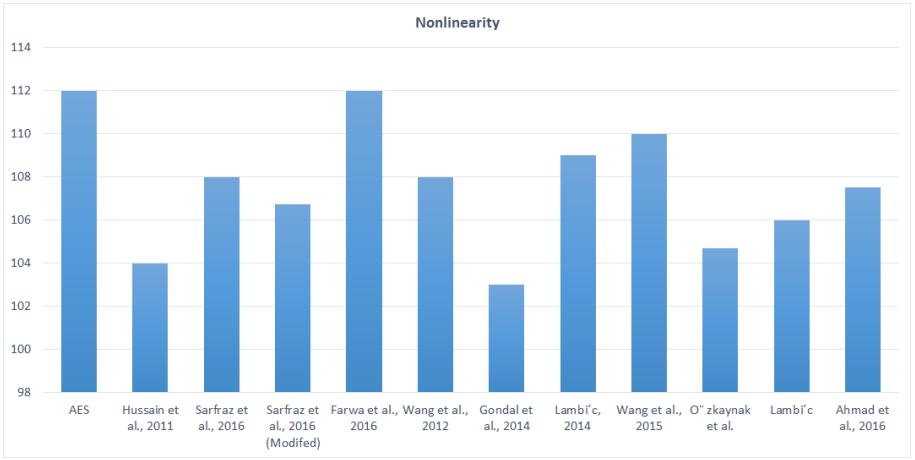

Figure 1: Comparison of Nonlinearity Performance

shows the comparison of average value of nonlinearity performance between algebraic construction and chaotic map technique. According to (Isa et al., 2016), an S-box that would categorized as cryptographically strong if and only if $NL > 100$. Therefore, the nonlinearity performance of the selected S-box in the literature are cryptographically strong as shown in Figure 1. Nonlinear properties are significant to avoid the differential and linear cryptanalysis which to ensure that the S-box is not linear mapping from input to the output vectors. (Carlet and Ding, 2007, Garg and Upadhyay, 2013). The range values of nonlinearity are between 103 to 112. AES (Daemen and Rijmen, 2002) obtained the highest value of nonlinearity, 112 due to the multiplicative inverse

S-box	Nonlin-earity	BIC	BIC of SAC	SAC	LAP	DAP	Max XOR
AES	112	112	0.504	0.504	0.062	0.0156	4
(Hussain et al., 2011)	104	103	0.46	0.493	0.125	0.125	NR
(Sarfraz et al., 2016)	108	106.571	0.501	0.498	0.125	0.0625	NR
(Sarfraz et al., 2016) (Modified)	106.75	105.571	0.497	0.502	0.125	0.0625	NR
(Farwa et al., 2016)	112	112	NR	0.51025	0.0625	0.01563	NR
(Wang et al., 2012)	108	103.36	0.5017	0.5068	NR	NR	10
(Gondal et al., 2014)	103	NR	NR	NR	0.04688	0.1484	NR
(Lambić, 2014)	109	104	NR	0.5012	NR	0.03516	NR
(Luma et al., 2015)	NR	NR	NR	0.51342	0.07601	0.04688	4
(Wang et al., 2015)	110	103.86	0.5033	0.4937	0.125	0.03906	NR
(Özkaynak et al.)	104.7	103.1	0.4942	0.5781	NR	NR	10
(Lambić)	106	100	NR	0.5034	0.07056	NA	10
(Ahmad et al., 2016)	107.5	NR	NR	0.5036	NR	0.039063	NR

Table 2: Performance comparison of the S-boxes Design Strategies

which contribute to the complex algebraic construction (Daemen and Rijmen, 2002). Meanwhile, nonlinearity value in scheme (Gondal et al., 2014) indicate the lowest amount that is 103.

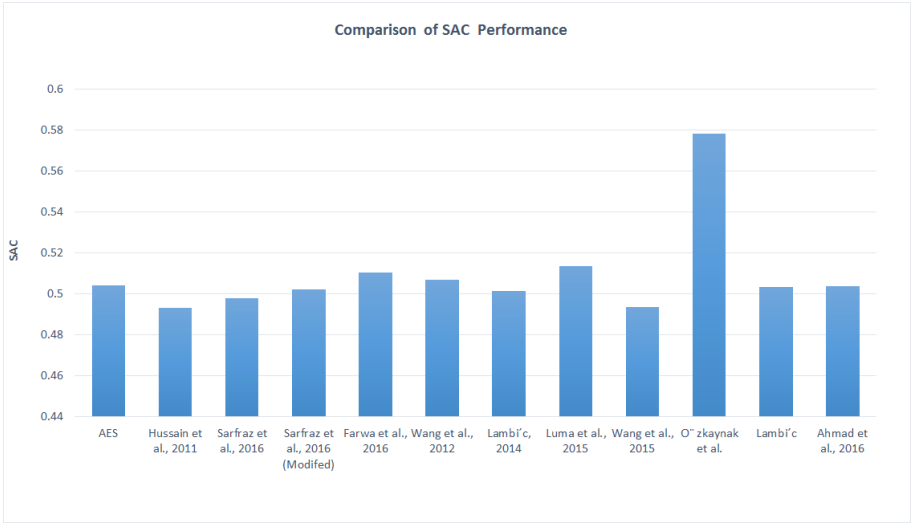

Figure 2: Comparison of Strict Avalanche Criterion (SAC) Performance

Figure 2 indicate the performance of selected S-box in the literature with respect to the strict avalanche criterion properties (SAC). The result in figure 2 show all of the selected S-box satisfies the SAC optimum value which required approximately close to 0.5 value. (Özkaynak et al.) shows the highest value of SAC, 0.5781 and (Hussain et al., 2011) have the lowest value, which can reach only 0.46. SAC is easily to analysed using these three method. Marr and Latt (Mar and Latt, 2008) did proposed easier and economical method for SAC analysis by considering the frequency of hamming weight, frequency of differential value and the analysis of hamming weights according to the bit position.

Figure 3 shows the results of the analysis of BIC analysis of the selected 8×8 S-box which contribute to the strength of encryption. Detombe and Teveres (Detombe and Tavares, 1992) was the first introduced the output of BIC properties. According to the definition of BIC stated in the previous section, by complement only a single bit of plaintext, all the avalanche variable

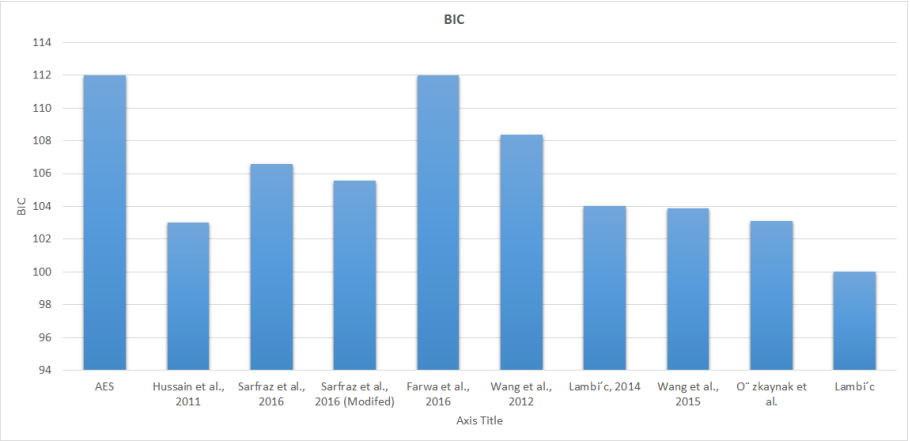

Figure 3: Comparison of Bit Independence Criterion (BIC) Performance

should be pair-wise for the given set of generated avalanche vectors.

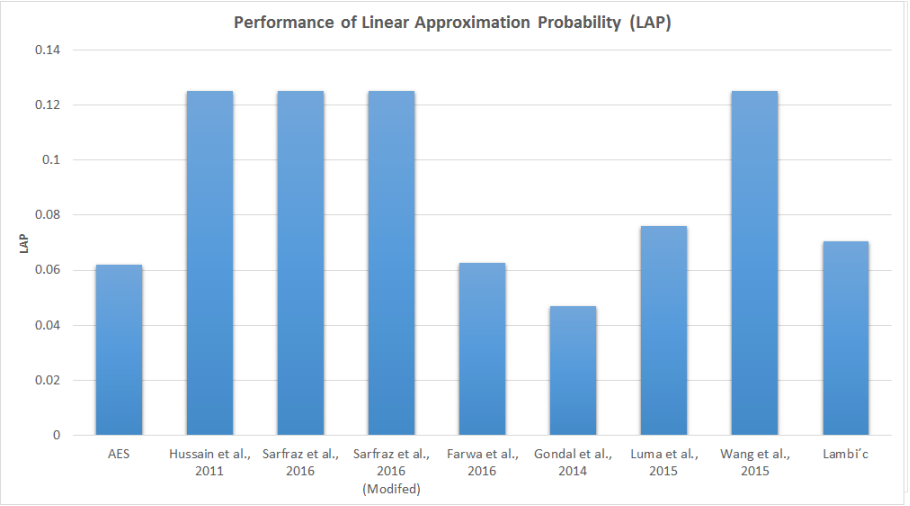

Figure 4: Comparison of Linear Approximation Probability (LAP) Performance

Figure 4 and 5 indicate the comparison of Linear and Differential Approximation respectively. These approximation is significant for the resistance of

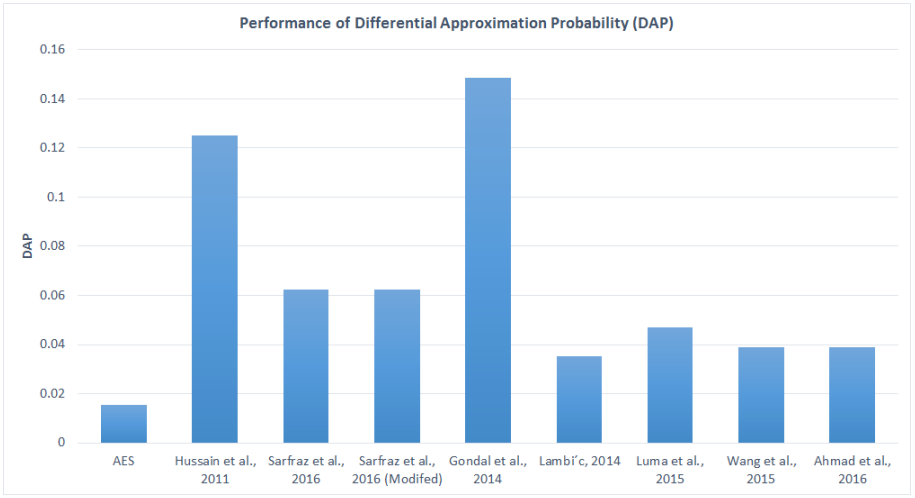

Figure 5: Comparison of Differential Approximation Probability (DAP) Performance

the most powerful attack namely Differential and Linear cryptanalysis.

7 CONCLUSION

The S-box designed using Algebraic construction have high nonlinearity value. Algebraic construction contributes to the complex computation of S-box. However, algebraic construction is vulnerable to algebraic attack. For chaotic system, the performance of S-box contribute to good cryptographic properties such has sensitivity towards the initial randomness and ergodicity which highly demand in cryptosystem. The S-box provide immunities against various of attack specifically linear and differential attack. To the best of our knowledge, there are still many work can be done to improve the security of S-box. For instance, we propose a possibility of hybrid computation between algebraic construction and chaotic map to obtain more complex computation in the S-box generation.

ACKNOWLEDGEMENT

We would like to thank the reviewers of International Journal of Cryptology Research for giving valuable suggestions and corrections in the paper. This research paper is supported by Fundamental Research Grant Scheme FRGS/1/2015/ICT05/FTMK/02/F00293 funded by the Ministry of Higher Education, Malaysia.

REFERENCES

Abuelyman, E. S. and Alsehibani, A.-A. S. (2008). An optimized implementation of the s-box using residue of prime numbers. *International Journal of Computer Science and Network Security*, 8(4):304–309.

Adams, C. M. and Tavares, S. E. (1990). Good s-boxes are easy to find. In *Proceedings of the 9th Annual International Cryptology Conference on Advances in Cryptology*, CRYPTO '89, pages 612–615, London, UK, UK. Springer-Verlag.

Adane, H. (2013). *Towards the Development of Stronger S-box*. PhD thesis, Addis Ababa University.

Ahmad, M., Mittal, N., Garg, P., and Khan, M. M. (2016). Efficient cryptographic substitution box design using travelling salesman problem and chaos. *Perspectives in Science*, 8:465–468.

Aslan, B. and Sakallı, M. T. (2014). Algebraic construction of cryptographically good binary linear transformations. *Security and Communication Networks*, 7(1):53–63.

Biham, E. and Shamir, A. (1991). Differential cryptanalysis of des-like cryptosystems. *Journal of CRYPTOLOGY*, 4(1):3–72.

Canright, D. (2005). A very compact s-box for aes. In *International Workshop on Cryptographic Hardware and Embedded Systems*, pages 441–455. Springer.

Carlet, C. (2010a). Boolean functions for cryptography and error correcting codes. *Boolean models and methods in mathematics, computer science, and engineering*, 2:257.

Carlet, C. (2010b). Vectorial boolean functions for cryptography. *Boolean models and methods in mathematics, computer science, and engineering*, 134:398–469.

Carlet, C. and Ding, C. (2007). Nonlinearities of s-boxes. *Finite Fields and Their Applications*, 13(1):121–135.

Clark, J. A., Jacob, J. L., and Stepney, S. (2005). The design of s-boxes by simulated annealing. *New Generation Computing*, 23(3):219–231.

Courtois, N. T. and Pieprzyk, J. (2002). Cryptanalysis of block ciphers with overdefined systems of equations. In *International Conference on the Theory and Application of Cryptology and Information Security*, pages 267–287. Springer.

Daemen, J. and Rijmen, V. (2002). *The Design of Rijndael*. Springer-Verlag New York, Inc., Secaucus, NJ, USA.

Detombe, J. and Tavares, S. (1992). Constructing large cryptographically strong s-boxes. In *International Workshop on the Theory and Application of Cryptographic Techniques*, pages 165–181. Springer.

Farwa, S., Shah, T., and Idrees, L. (2016). A highly nonlinear s-box based on a fractional linear transformation. *SpringerPlus*, 5(1):1658.

Gangadari, B. R. and Ahamed, S. R. (2015). Analysis and algebraic construction of s-box for aes algorithm using irreducible polynomials. In *Contemporary Computing (IC3), 2015 Eighth International Conference on*, pages 526–530. IEEE.

Garg, S. and Upadhyay, D. (2013). S-box design approaches: Critical analysis and future directions. *International Journal of Advanced Research in Computer Science and Electronics Engineering (IJARCSEE)*, 2(4):pp–426.

Golić, J. D. (1996). Fast low order approximation of cryptographic functions. In *International Conference on the Theory and Applications of Cryptographic Techniques*, pages 268–282. Springer.

Gondal, M. A., Raheem, A., and Hussain, I. (2014). A scheme for obtaining secure s-boxes based on chaotic baker's map. *3D Research*, 5(3):1–8.

Hussain, I., Shah, T., Gondal, M. A., and Khan, W. A. (2011). Construction of cryptographically strong 8× 8 s-boxes. *World Applied Sciences Journal*, 13(11):2389–2395.

Hussain, I., Shah, T., Mahmood, H., and Gondal, M. A. (2013). A projective general linear group based algorithm for the construction of substitution box for block ciphers. *Neural Computing and Applications*, 22(6):1085–1093.

Isa, H., Jamil, N., and Zaba, M. R. (2016). Construction of cryptographically strong s-boxes inspired by bee waggle dance. *New Generation Computing*, 34(3):221–238.

Ivanov, G., Nikolov, N., and Nikova, S. (2015). Cryptographically strong s-boxes generated by modified immune algorithm. In *International Conference on Cryptography and Information Security in the Balkans*, pages 31–42. Springer.

Ivanov, G., Nikolov, N., and Nikova, S. (2016). Reversed genetic algorithms for generation of bijective s-boxes with good cryptographic properties. *Cryptography and Communications*, 8(2):247–276.

Jakimoski, G., Kocarev, L., et al. (2001). Chaos and cryptography: block encryption ciphers based on chaotic maps. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 48(2):163–169.

Knudsen, L. and Wagner, D. (2001). On the structure of skipjack. *Discrete Applied Mathematics*, 111(1):103–116.

Lambić, D. A novel method of s-box design based on discrete chaotic map. *Nonlinear Dynamics*, pages 1–7.

Lambić, D. (2014). A novel method of s-box design based on chaotic map and composition method. *Chaos, Solitons & Fractals*, 58:16–21.

Li, C. and Peng, G. (2004). Chaos in chen's system with a fractional order. *Chaos, Solitons & Fractals*, 22(2):443–450.

Lineham, A. and Gulliver, T. A. (2008). Heuristic s-box design. *Contemporary Engineering Sciences*, 1(4):147–168.

Luma, F., Hilal, H., and Ekhlas, A. (2015). New dynamical key dependent s-box based on chaotic maps. *IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN*, 17(4):91–101.

Mar, P. P. and Latt, K. M. (2008). New analysis methods on strict avalanche criterion of s-boxes. *World Academy of Science, Engineering and Technology*, 48(150-154):25.

Matsui, M. (1993). Linear cryptanalysis method for des cipher. In *Workshop on the Theory and Application of of Cryptographic Techniques*, pages 386–397. Springer.

Millan, W. (1998). How to improve the nonlinearity of bijective s-boxes. In *Australasian Conference on Information Security and Privacy*, pages 181–192. Springer.

Millan, W., Burnett, L., Carter, G., Clark, A., and Dawson, E. (1999). Evolutionary heuristics for finding cryptographically strong s-boxes. In *International Conference on Information and Communications Security*, pages 263–274. Springer.

Özkaynak, F., Çelik, V., and Özer, A. B. A new s-box construction method based on the fractional-order chaotic chen system. *Signal, Image and Video Processing*, pages 1–6.

Rehman, A. U., Khan, J. S., Ahmad, J., and Hwang, S. O. (2016). A new image encryption scheme based on dynamic s-boxes and chaotic maps. *3D Research*, 7(1):1–8.

Sarfraz, M., Hussain, I., and Ali, F. (2016). Construction of s-box based on mobius transformation and increasing its confusion creating ability through invertible function. *International Journal of Computer Science and Information Security*, 14(2):187.

Seberry, J., Zhang, X.-M., and Zheng, Y. (1993). Systematic generation of cryptographically robust s-boxes. In *Proceedings of the 1st ACM Conference on Computer and Communications Security*, pages 171–182. ACM.

Shannon, C. E. (1949). Communication theory of secrecy systems. *Bell system technical journal*, 28(4):656–715.

Tang, G. and Liao, X. (2005). A method for designing dynamical s-boxes based on discretized chaotic map. *Chaos, Solitons & Fractals*, 23(5):1901–1909.

Tran, M. T., Bui, D. K., and Duong, A. D. (2008). Gray s-box for advanced encryption standard. In *Computational Intelligence and Security, 2008. CIS'08. International Conference on*, volume 1, pages 253–258. IEEE.

Wang, Y., Lei, P., and Wong, K.-W. (2015). A method for constructing bijective s-box with high nonlinearity based on chaos and optimization. *International Journal of Bifurcation and Chaos*, 25(10):1550127.

Wang, Y., Wong, K.-W., Li, C., and Li, Y. (2012). A novel method to design s-box based on chaotic map and genetic algorithm. *Physics Letters A*, 376(6):827–833.

Webster, A. and Tavares, S. E. (1985). On the design of s-boxes. In *Conference on the Theory and Application of Cryptographic Techniques*, pages 523–534. Springer.

Xiao, L. and Heys, H. M. (2005). Software performance characterisation of block cipher structures using s-boxes and linear mappings. *IEE Proceedings-Communications*, 152(5):567–579.