Study of S-box Properties in Block Cipher

Conference Paper · September 2014

DOI: 10.1109/I4CT.2014.6914206

Study of S-box Properties in Block Cipher

Kamsiah Mohamed, Mohd Nazran Mohammed Pauzi
Department of Computer Science
Universiti Selangor, Bestari Jaya
Selangor, Malaysia
kamsh@unisel.edu.my, nazran@unisel.edu.my

Fakariah Hani Hj Mohd Ali, Suriyani Ariffin, Nurul Huda Nik Zulkipli
Faculty of Computer and Mathematical Sciences
Universiti Teknologi MARA, Shah Alam
Selangor, Malaysia
fakariah@tmsk.uitm.edu.my, suriyani@tmsk.uitm.edu.my

Abstract—**In the field of cryptography, the substitution box (S-box) becomes the most widely used ciphers. The process of creating new and powerful S-boxes never end. Various methods are proposed to make the S-box becomes strongest and hard to attack. The strength or weakness of S-box will be determined through the analysis of S-box properties. However, the analysis of the properties of the S-box in block ciphers is still lacking because there is no specific guidelines and technique based on S-box properties. Hence, the cipher is easier to attack by an adversary if the S-box properties are not robust. The purpose of this paper is to describe and review of the S-box properties in block ciphers. As a result, for future work, a new model for analysis S-box properties will be proposed. The model can be used to analysis the properties to determine the strength and weakness of any S-boxes.**

Index Terms—**Cryptography, S-box, cryptanalysis.**

I. Introduction

Cryptography is becoming necessary when sensitive data is being transacted over any untrusted medium. It becomes extremely useful, especially in many applications, for example, identification and authentication, secure communication systems, online billing, secure log in, emails, etc. The encryption algorithm is used to generate and encrypt a key. The strength of encryption depends on the ability of S-box in distorting the data; hence, the processes of discovering new and powerful S-boxes are of great interest in the field of cryptography [1]. Many researchers were emulated to redesign, reconstruct or renew the design and implementation of the S-box in order to make their S-box is strong and secure. They used many methods to construct their S-box to resist cryptanalysis attack. There are two types of S-box in block cipher which are static and dynamic S-box. For example, in 2001, a Latin square S-box approach was proposed to build up dynamic change S-boxes [2]. The secret key of the length 128 bits is used to generate new Latin square S-box. This approach will solve the problem of the static structure S-boxes and consequently will increase the security level of the block cipher system. However, the dynamic S-box is needed to make cryptanalysis is difficult to discover the key in a block cipher. This is because the problem of a static S-box is easy to to identify a weak key. Thus, the data is less secure to achieve the highest security. And so, a dynamic S-boxes are designed using Latin Square doubly stochastic matrix was proposed by Wu and Noonan [3]. Then, a new approach for generating dynamically S-boxes using spatiotemporal chaotic system was presented by Peng [4]. Within the algorithm, the key is mapped to system parameters and the hyper-chaotic sequences are generated to construct an S-box. However, to determine the strength and weakness of their S-box the properties in S-box should be analyzed. It must fulfill several cryptographic properties such as high nonlinearity, low differential uniformity and complex algebraic expression resist against linear, differential and interpolation attacks [5]. With good cryptographic properties, it is possible to design a dynamic S-box in a block cipher [6]. This paper will discuss the security of S-box in a block cipher based on the properties.

This paper is organized as follows: In Section II, an overview of S-box is discussed. In Section III, S-box properties are presented. In Section IV a study of S-box properties is discussed. In Section V, a conclusion and possible further works are given.

II. Overview of S-box

The security of data relies on the substitution process. Substitution is a nonlinear transformation which performs confusion of bits. It provides the cryptosystem with the confusion property described by Shannon [7]. He suggested that strong ciphers could be built by combining substitutions with transposition repeatedly. The earliest block ciphers were simple networks that combined substitution and permutation circuits, and called substitution permutation networks (SPN). In modern encryption algorithm a nonlinear transformation is essential and is proved to be a strong cryptographic primitive against linear and differential cryptanalysis [8]. An example of a nonlinear transformation algorithm is Advanced Encryption Standard (AES). This standard specifies the Rijndael algorithm. It is widely used in cryptographic applications approved by the National Institute of Standards and Technology (NIST) in 2001. It was designed to handle additional block sizes and key lengths 128, 192 and 256 bits. However, AES limits the block length to 128 bits [9]. In the Rijndael algorithm, S-box is the most important part because of the encryption algorithm. Encryption algorithm means that it requires the key to be the same length as the message to be encoded. However, it causes the most delay of the encryption algorithm. Table I shows nonlinear transformations of S-box in AES. The S-box component used in AES is fixed and not changeable [8]. The static S-box will use the same S-box in each round while for key-dependent or dynamic S-Box it will change in round of S-box depends on the key and number of rounds. The dynamic or dependent key algorithm should be generated to increase the cryptographic strength of the AES cipher system. S-box depend on key values are slower, but more secure than independent ones [10]. When design and analysis AES, the cryptographic properties of S-box must be considered, especially the avalanche effect [11]. For the moment the AES hasn't been broken, but the cryptanalysis of Rijndael (AES) has not stopped [12].

TABLE I. S-BOX USED IN AES

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

An example of S-box mapping is shown in Figure 1. The S-box mapping represent with input X and output Y [13]. In order to detect the linear cryptanalysis of the S-box, the linear approximation should be examined to determine the probability bias of S-box. The basic idea is to approximate the operation of a portion of the cipher with an expression that is linear where the linearity refers to a mod-2 bit-wise operation (i.e., exclusive-OR denoted by " $\oplus$ ").

For example, consider the linear equation is

$$X_2 \oplus X_3 \oplus Y_1 \oplus Y_3 \oplus Y_4 = 0 \quad (1)$$

Assume 16 possible input values for X was applied and examining the corresponding output values Y, it may be noted that for exactly 12 out the 16 cases, the equation 1 holds true. Hence, the probability bias is 12/16−1/2 = 0.25. Nevertheless, the success of the approach is based on the minimum and maximum bias. The cryptanalyst has to take sufficiently many plaintext/ciphertext pairs to examine the value is correct with minimum bias.

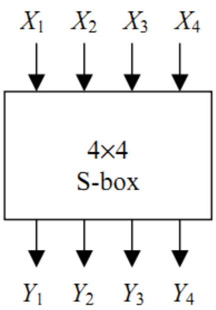

FIGURE I. S-BOX MAPPING

III. S-BOX PROPERTIES

The properties S-boxes have been widely used as a base of new encryption strategies such as nonlinearity, differential uniformity, and strict avalanche criterion [14].

A (x,y)-S-box is a map, $S:\{0,1\}^x \rightarrow \{0,1\}^y$.
It comprises of n-variable component Boolean functions: $(f_1\ (x_1, ..., x_n), f_2\ (x_1, ..., x_n), ..., f_n(x_1, ..., x_n))$ each of which need to satisfy S-box properties.

The following are the list of several properties in S-box.

i) Robustness
Let $F = (f_1, f_2, ..., f_n)$ be an $n \times n$ S-box, where f_i is a component function of S-box mapping
$f_i : \{0, 1\}n \rightarrow \{0, 1\}$

$$R = \left(1 - \frac{N}{2^n}\right)\left(1 - \frac{L}{2^n}\right)$$

F must be Robust to against differential cryptanalysis [15].

ii) Balancing
$S:\{0,1\}^n \rightarrow \{0,1\}^m$ balanced, if $HW(f) = 2^{n-1}$. The significance of the balance property is based on the higher the magnitude of function imbalance, a high probability linear approximation being obtained.

iii) Strict Avalanche Criterion (SAC)
A change in one bit of input bits of S-box should produce a change in half of output bits of S-box. It is harder to perform an analysis of ciphertext, when trying to come up with an attack.
A cryptographic function which satisfies the above condition is said to be satisfied Strict Avalanche Criteria.

iv) Nonlinearity
$S:\{0,1\}^x \rightarrow \{0,1\}^y$ is defined as the least value of nonlinearity of all nonzero linear combinations of x boolean functions $f_i : \{0,1\} \rightarrow \{0,1\}$, $i = x-1, ..., 1, 0$.
The nonlinearity of an S-box must be high to resist against linear cryptanalysis.

v) Differential Uniformity
The smaller is the Differential Uniformity, the better is the S-box's resistance against differential cryptanalysis.

vi) Linear Approximation
The lower is the Linear Approximation value, the better is the S-box's resistance against linear cryptanalysis.

vii) Algebraic Complexity
The Algebraic Complexity is important to resist against interpolation attack and other concerning algebraic attacks.

viii) Fixed (Fp) and Opposite Fixed Points (OFp)
The number of these Fp and OFp should be kept as low as possible to avoid leakage in any statistic cryptanalysis.

ix) Bit independence criterion
The bit independence is a highly desirable property as with increasing independence between bits, it becomes more difficult to understand and predict the design of the system.

Most of researchers always construct their S-box to resist against linear cryptanalysis and differential cryptanalysis attacks. Every new cipher should be tested in the case of the weak keys [16].

In Table II, there are several static S-boxes have been broken by the linear and the differential cryptanalysis attacks. It attacks the S-box properties based on the linear approximation and avalanche effect. For example, the linear cryptanalysis attack is successful broken the number of rounds in linear approximation property at DES S-box. In DES all the S-boxes in a round are different, while all rounds use the same set of S-boxes. The approach in linear cryptanalysis is to determine expressions of the form above which have a high or low probability of occurrence.

Hence, each S-box properties are important to analyze in order to avoid cryptanalysis attacks. The strong and secure S-box is needed to protect the number of rounds, a key, confusion and round function in S-boxes. As a result, the key should be difficult to discover in order to show that the more secure of the S-boxes mechanism.

There are four common types of cryptanalysis attacks which are ciphertext only, known plaintext, chosen plaintext and chosen ciphertext [17]. The aim of classifying attacks is to clear about the types of attacks are applicable when certain information is accessible to an assailant. The ciphertext only means that, only ciphertext was needed to attack the S-box by the cryptanalyst. Various methods can be employed to attack ciphertext only such as a brute force method, statistical method or pattern attack. For example, when the most commonly used character of the ciphertext and the plaintext was identified with the cryptanalyst the statistical method is applied. After identifying the pairs between ciphertext and plaintext characteristics, the analyst can find the key and apply it to decrypt the message. In order to prevent this type of attack, the S-box in block cipher should hide the characteristics of the language. For the known plaintext attack, pairs of plaintext or ciphertext is collected earlier. The previous pairs are used to analyze the current ciphertext. In many applications and scenarios, it is reasonable to assume that the attacker has knowledge of a random set of plaintexts and the corresponding ciphertexts. This approach is easier to enforce because the analyst have more information to break the key based on the previous message. Examples of known plaintext attack are linear cryptanalysis, interpolation attacks and slide attacks. The chosen plaintext attack is similar to the known plaintext attack. It can be created to go with the only known plaintext. However the plaintext/ciphertext pairs have been chosen by attacker herself. Chosen plaintext attacks consist of differential cryptanalysis, square attacks and boomerang attacks.

TABLE II. ANALYSIS OF S-BOXES

Static S-Box	Technique	Type of Attack	S-Box Property	Crypt-analysis	Weakness
DES [18]	Linear Cryptanalysis	Known Plain text	Linear Approximation	Successful break 8 rounds, 12 rounds and 16 rounds DES cipher.	Ignore the initial permutation - compression of the key scheduling.
AES [19]	Linear Cryptanalysis	Chosen Plain text	Linear Approximation	Successful break 1 and 2 AES cipher	Not work well on more than 3 rounds
PRE-SENT [20]	Linear Cryptanalysis	Known Plain text	Low Avalanche effect	Can attack up to 24 rounds.	The 32 % key is weak. The evaluation is only one input-output mask.
AES and Camellia [21]	Related S-Box attacks (the round function and key schedule)	Un-known components	None	Successful reverse engineering of two different ciphers for the first round only.	Applicable for 8 bits S-boxes not for 4 bit s-boxes.
CTC2 DES Serpent [22]	Differential-linear attacks (the numbers of attack rounds)	Chosen Plain Text	Linear Approximation and Time complexity	CTC2- 10 rounds DES – 13 rounds Serpent – 12 rounds	At present, these attack techniques appear to be hard to apply to obtain good cryptanalytic results in practice.

IV. A STUDY OF S-BOXES PROPERTIES

A good S-box should satisfy a lot of criteria, for example nonlinear properties to determine the performance of the whole block cipher. In order to classify the S-box are strong and secure the properties of S-box should be analyzed as shown in Table III. However, some of them did not analyze their S-box. This is because such properties of S-boxes are difficult to analyze in the context of a single S-box [23]. In order to solve the problem of S-box properties an S-box simulator were created by Niemiec and Machowski [16]. However, not all properties can be analyzed by using the simulator. It can be used to analyze certain properties which are balanced, SAC, completeness, diffusion order, low XOR table and nonlinearity. Moreover, the result of S-box only can be referred by their method. In this paper, several studies related to S-box properties is compared in order to determine the sufficient properties to assess S-box.

As a result, this study found that most of the S-box properties are important to analyze in order to resist against linear cryptanalysis, differential cryptanalysis, interpolation attack and statistical cryptanalysis. All studies were analyzing a nonlinearity property to protect their S-box from the linear cryptanalysis attack. If the analysis indicates that, the high nonlinearity, it provides stronger resistance to linear cryptanalysis.

Based on the previous studies, there is no specific guideline exist to analyze the S-box properties. For instance, one of the most important characteristics of an S-box is an avalanche criterion that is a bit change in the input byte of an S-box must result in a change in the output byte at least by 50% of bits studied by Chandrasekharappa [24]. While another study said that two of the important characteristics which decides the strength of an S-box are robustness and SAC, both of which are derived from the Difference Distribution Table (DDT) [25]. Besides that, to be considered as cryptographically strong, an S-box needs to satisfy balancing, high nonlinearity, low differential uniformity, high algebraic degree, low linear approximation, high algebraic complexity, and low/no fixed and opposite fixed points [5]. Consequently, there are no algebraic procedures that can give the preferred and the complete set of properties for an S-box [26]. As a result, most of the researches assess their S-box properties based on their perceptions and commonly used. Therefore, a specific guideline will assist researchers about appropriate S-box properties to protect from cryptanalysis attack.

TABLE III. COMPETITIVE STUDY OF S-BOX PROPERTIES

S-Box Properties	**Study by Researchers**				
	Hussain, et al. [1]	**Isa, Jamil, Zaaba [5]**	**Mazumdar [15]**	**Niemiec and Machowski [16]**	**Radhakrishnan and S. Subramanian [25]**
Robustness	√		√		
Randomness	√				√
Balancing			√		√
Strict Avalanche Criterion	√		√	√	√
High Nonlinearity	√	√	√	√	√
Low Differential Uniformity		√		√	
High Algebraic Degree (AD)		√	√		
Low Linear Approximation		√	√	√	
Low Fixed (Fp) and Opposite Fixed Points (OFp)		√			
High Algebraic Complexity		√			
Completeness					√
Diffusion Order					√
Low XOR table					√
Bit Dependence Criterion				√	
Low Transparency Order			√		
Majority Logic Criterion				√	
Correlation			√		
Propagation Characteristic			√		

V. CONCLUSION AND FUTURE WORKS

Most of researchers aware that S-box properties are very important to make sure that their S-box is secure [1, 5, 15,16 , 25]. However, most of them analyze their S-box properties based on their perceptions without any proper guideline. In addition, cryptanalysis attempt to break the S-box properties with all kinds of method.

As a future work, a new model that can be analyzed all the S-box properties in any S-box ciphers will be proposed. The model will be given a broad insight into S-box properties and guidelines for proper use and implementation. The most important aspect is the model can be used to protect the S-box from any cryptanalysis attacks.

Acknowledgment

This research was supported by the Research Management Institute, Universiti Teknologi MARA and registered under the Research Acculturation Grant Scheme (RAGS) #600-RMI/RAGS 5/3(204/2013) by Ministry of Education of Malaysia.

References

[1] I. Hussain, et al., "A projective general linear group based algorithm for the construction of substitution box for block ciphers," *Neural Computig. Appliction*, Vol. 22, no. 6, pp. 1085–1093, Feb. 2012.

[2] S. H. El-Ramly, T. El-Garf, and A.H. Soliman. "Dynamic Generation of S-Boxes In Block Cipher Systems". Eighteenth National Radio Science Conference. Mansoura University, Egypt. 2001. pp.389-397 vol 2.

[3] Y. Wu and J. P. Noonan, "Dynamic and Implicit Latin Square Doubly Stochastic S-Boxes with Reversibility," pp. 3358–3364, 2011.

[4] J. Peng et al., "Construction and Analysis of Dynamic S-boxes Based on Spatiotemporal Chaos," vol. 1. Proc. 11th IEEE Int. Conf. on Cognitive Informatics & Cognitive Computing ©2012 IEEE.

[5] H. Isa, N. Jamil, and M. R. Z'aba, "S-box construction from non-permutation power functions," *Proc. 6th Int. Conf. Secur. Inf. Networks - SIN '13*, pp. 46–53, 2013.

[6] M. Wang, "How to Search Linear Approximation for Large for Non-Surjective S-box."ACM. 2011

[7] Shannon,C.E "Communication Theory of Secrecy Systems", Bell Sysit. Tech 5.28, PP.656-715, 1949.

[8] Hosseinkhani and Javadi. "Using Cipher Key to Generate Dynamic S-Box in AES Cipher Syst. International Journal of Computer Science and Security (IJCSS), Volume (6): Issue (1): 2012.

[9] Shivkumar and Umamaheswari. "Performance Comparison of Advanced Encryption Standard (AES) and AES key dependent S-Box –Simulation Using MATLAB". 2011. IEEE.

[10] Schneier. "Applied Cryptography: Protocols, Algorithms, and Source Code in C", New York: Wiley. 1996.

[11] H. Shi and Y. Deng, "Analysis of the avalanche effect of the AES S box," *2011 2nd Int. Conf. Artif. Intell. Manag. Sci. Electron. Commer.*, pp. 5425–5428, Aug. 2011.

[12] Arrag, et. al. "Implementation of Stronger AES by using dynamic S-Box dependent of master key". Journal of Theoretical and Applied Information Technology. Vol. 53., July 2013.

[13] H. M. Heys, "A Tutorial on Linear and Differential Cryptanalysis," *Cryptologia*, vol. 26, no. 3, pp. 189–221, Jul. 2002.

[14] Y. Zhang and D. Xiao, "Cryptanalysis of S-box-only chaotic image ciphers against chosen plaintext attack," pp. 751–756, 2013.

[15] B. Mazumdar, D. Mukhopadhyay, and I. Sengupta, "Design for Security of Block Cipher S-Boxes to Resist Differential Power Attacks," pp. 0–5, 2012.

[16] Niemiec and Machowski. "A new symmetric block cipher based on key-dependent S- boxes". IV International Congress on Ultra Modern Telecommunications and Control System 2012.

[17] A. Forouzan, "Cryptography and Network Security Forouzan", Tata McGraw Hill Publishing Company Limited, 2008.

[18] Matsui, "Linear Cryptanalysis method for DES cipher." Springer-Verlag.1994.

[19] K. Sakamura, W. Xiao, and H. Ishikawa, "A Study on the Linear Cryptanalysis of AES Cipher," vol. 2, no. 1, pp. 19–26, 2004.

[20] K. Ohkuma, "Weak Keys of Reduced-Round PRESENT for Linear Cryptanalysis," vol. 100, pp. 249–265, 2009.

[21] M. Macchetti, "Cryptanalysis of AES and Camellia with Related S-boxes," pp. 208–221, 2013.

[22] J. Lu, "A methodology for differential-linear cryptanalysis and its applications," *Des. Codes Cryptogr.*, Jun. 2014.

[23] M. Saarinen. "Cryptographic Analysis of All 4 × 4 -Bit S-Boxes". Cryptography, Springer, 2012.

[24] T G S Chandrasekharappa, "Enhancement of confidentiality and integrity using cryptographic techniques", 2012. Manipal Institute of Technology http://hdl.handle.net/10603/5051

[25] S. V. Radhakrishnan and S. Subramanian, "An Analytical Approach to S-box Generation," pp. 1–5, 2012. IEEE.

[26] S. Picek and M. Golub, "On Using Genetic Algorithms for Intrinsic Side-Channel Resistance□: The Case of AES S-Box Categories and Subject Descriptors."Copyright 2014 ACM.