Analysis of the Use of Whirlpool's *S-box*, S1 and S2 SEED's *S-box* in AES Algorithm with SAC Test

Novita Angraini, Bety Hayat Susanti, Magfirawaty
Sekolah Tinggi Sandi Negara

Keywords:

Advanced Encryption Standard(AES)
Substitution Box (*S-box*)
Strict Avalanche Criterion (SAC)
Bit Independence Criterion (BIC)
Nonlinearity

ABSTRACT

Since 2001, Advanced Encryption Standard (AES) block cipher has become a standard algorithm and widely used in cryptographic applications. The heart of AES is a nonlinear substitution box (*S-box*) that generated using Affine transformation. In this paper, we change the original *S-box* with Whirlpool's *S-box*, S1 and S2 Seed's *S-box*. After that, we analyze the effect of their usage in AES and compared them to the original. We decide to use these *S-box*es due to the similarity dimension with AES's *S-box*. The cryptographic properties such as strict avalanche criterion (SAC), bit independence criterion (BIC), XOR table distribution and nonlinearity of these*S-box*es are analyzed in details. According to our experiments, S1 and S2 Seed's *S-box* have an error value close to AES's *S-box* error value for all tests but not so with Whirlpool's *S-box*. We also test the original and modified AES using SAC. The test results showed that they satisfy SAC.

Corresponding Author:

Bety Hayat Susanti,
Sekolah Tinggi Sandi Negara,
Jalan H.Usa Putat Nutug, Ciseeng, Bogor, 16330.
Email: bety.hayat@lemsaneg.go.id

1. INTRODUCTION

The Rijndael block cipher algorithm was published as FIPS 197 on November 2001 and established by NIST as the new advanced encryption standard (AES) [1]. AES uses key sizes of either 128, 192, or 256 bits, which perform 10, 12, and 14 rounds respectively. Each round consists of four transformations (SubBytes, ShiftRows, MixColumns, and AddRoundKey) except the final round which omits the MixColumns stage [1]. The SubBytes transformations are non-linear substitutions. They replace each of the 16 bytes in the input block with another apparently random byte and contain a permutation for each of the 256 byte values. This substitution table has 8×8 dimensions and it's called the S-box. The AES's S-box test results satisfy the properties of AC, SAC, and BIC with error values close to 0 and has an evenly nonlinearity value.

The block ciphers: Skipjack, Camellia, and SEED as well as Whirlpool hash function are the algorithms that also use an 8×8 S-Box like AES. To the best our knowledge, S-Box testing has not been done to the SEED's S-box and Whirlpools's S-box. Therefore, we carry out S-box testing of the Whirlpool Hash Function, S1 and S2 SEED's S-boxes using SAC, BIC, XOR-table distribution, and the nonlinearity tests. After that, we analyze the effect of their usage in AES and compared them to the original using SAC test.

2. THEORETICAL BACKGROUND

a. Substitution Box (S-box)

In general, S-box takes some number of input bits, m, and mappings them into some number of output bits, n: an $m \times n$ S-Box can be implemented as a lookup table with 2^m words of n bits each [2]. Static tables are normally applied as in the Data Encryption Standard (DES), but in some ciphers the tables are generated dynamically from the key; e.g. the Blowfish and the Twofish encryption algorithms. Another kind of S-box used in the International Data Encryption Algorithm (IDEA) modular multiplication step as a key-dependent S-Box [3].

Definition 1. [4] : An *n x n* S-box is a mapping function $f:\{0,1\}^n \rightarrow \{0,1\}^n$, which maps *n*-bit input strings, $X = \{x_1, x_2, \dots, x_n\}$ to *n*-bit output strings, $Y = \{y_1, y_2, \dots, y_n\}$ where $Y = f(x)$.

Mister and Adams [5] explained an *n x m* S-box *S* is a mapping $S:\{0,1\}^n \rightarrow \{0,1\}^m$. *S* can be represented as 2^n *m*-bit numbers, denoted $r_0, \dots, r_{2^n-1}$ in which case $S(x) = r_x$, $0 \le x < 2n$ and the r_i are the rows of the S-box. $S(x) = [c_{m-1}(x) c_{m-2}(x) \dots c_0(x)]$, where the c_i are fixed Boolean functions $c_i : \{0,1\}^n \rightarrow \{0,1\} \; \forall i$; these are the columns of the S-box. Finally, *S* can be represented by a 2^n *x m* binary matrix *M* with the *i, j* entry being bit *j* of row *i*.

b. Strict Avalanche Criterion (SAC)

According to [6], Strict Avalanche Criterion (SAC) is the combination between the concepts of completeness and avalanche effect. If a cryptographic function $f:\{0,1\}^n \rightarrow \{0,1\}^n$ satisfies SAC for all $i, j \in \{1, 2, \dots, n\}$, then each output bit should change with a probability of one half whenever a single input bit is complemented, formulated as follows [6]:

$$\frac{1}{2^n} W\left(a_j^{e_i}\right) = \frac{1}{2} \text{ for all } i, j \qquad \text{...............(1)}$$

We can modify equation (1) to determine the parameter of SAC, $k_{SAC}(i,j)$ as follows:

$$k_{SAC}(i,j) = \frac{1}{2^n} W\left(a_j^{e_i}\right) = \frac{1}{2} \qquad \text{...............(2)}$$

$k_{SAC}(i,j)$ in the range of [0,1] and can be interpreted as probability of a change in the j-th bit output when the i-th bit input change. If $k_{SAC}(i,j)$ is not equal to ½ for every pair of (*i,j*), then it is not satisfying SAC.
Relative error of SAC results can be obtained by the formula:

$$\epsilon = \max_{\substack{1 \le i \le n \\ 1 \le j \le n}} \left|2 k_{SAC}(i,j) - 1\right| \qquad \text{................(3)}$$

c. Bit Independence Criterion (BIC)

A function $f:\{0,1\}^n \rightarrow \{0,1\}^n$ is to satisfies BIC if $\forall i, j, k \in \{1, 2, \dots, n\}$, with $j \neq k$, inverting input bit i causes output bits j and k to change independently [6].

To measure the bit independence concept, one needs the correlation coefficient between the j'th and k'th components of the output difference string, which is called the avalanche vector A^{e_i}. Bit independence criteria corresponding to the effect of the i'th input bit change on the j'th and k'th bits of A^{e_i} is defined as:

$$BIC(a_j, a_k) = \max_{1 \le i \le n} \left|corr\left(a_j^{e_i}, a_k^{e_i}\right)\right| \qquad \text{..............(4)}$$

In the process of criteria analysis of BIC, the $BIC(f)$ value will be the relative error ϵ_B. Thus, for an nxn S-box, the maximum value of $\epsilon_B = BIC(f)$ is said the maximum value of relative error of BIC results, denoted by ϵ_{BIC},

$$\epsilon_{BIC} = \max_{\text{over all s-box}} \{\epsilon_B\} \qquad \text{.................(5)}$$

d. XOR Table Distribution

For an *n* x *n* S-box,an XOR-Table of the S-box is a matrix that has the given row and column indexed by 0,1,2, ..., 2^n – 1, and entries in the table indexed by (*δ*, *b*), where *δ* indicates the number of input vectors P is modified by *δ*, and *b* shows the change in output [7]

$b = f(P) \oplus f(P \oplus \delta)$

The XOR table formula is given by : $XORf(\delta, b) = \#\{P \mid f(P) \oplus f(P \oplus \delta)\}$ ....(6)

Where $\delta \in \{0,1\}^n$ and $b = \{0,1\}^m$.

Number of entries in the XOR table are always even and the sum of all values in each row is always 2^n. Ideally, an entry in the XOR Table S-boxes all zero or two with the exception of the entry (0,0) has been always worth 2^n.

e. Nonlinearity

According to [3], the nonlinearity of the function $f:\{0,1\}^n \rightarrow \{0,1\}^m$ is defined as the minimum Hamming distance between the set of Affine functions and every nonzero linear combination of the output coordinates of *f*, i.e.

$$NL_f = \min_{a,b,w} \#\{x \in \{0,1\}^n \mid a.f(x) \neq w.x \oplus b\} \qquad \text{..................(7)}$$

where $w \in \{0,1\}^n$, $a \in \{0,1\}^m \backslash \{0\}$, $b \in \{0,1\}$, and $w \cdot x$ denotes the dot product between *w* and *x* over $\{0,1\}$,

$$c.f(x) = \bigoplus_{i=1}^{m} c_i f_i(x), \qquad ....(8)$$

where $c = \{c_1, c_2, \ldots, c_m\} \in \{0,1\}^m$.

For a cryptosystem not to be susceptible to linear cryptanalysis, NLN_f is required to be as close as possible to its maximum value (perfect nonlinearity). The maximum nonlinearity value (perfect nonlinearity) of the Boolean function given by $N_f \leq 2^{n-1} - 2^{\frac{n}{2}-1}$ [6].

f. Advanced Encryption Standard (AES)

Rijndael is a symmetric block cipher algorithm, created by Joan Daemen and Vincent Rijmen and established by NIST as the Advanced Encryption Standard (AES) [1]. It encrypts data in 128-bit input block into 128-bit output block. The key length varies from $128, 192,$ and 256 bits. The number of rounds depends on the key length used in the algorithm.

The algorithm begins with an AddRoundKey transformation followed by 9 rounds of four transformations and a tenth round of three transformations. This applies for both encryption and decryption with the exception that each transformation of a round, the decryption algorithm is the inverse of the encryption algorithm. The four transformations are namely: SubBytes, ShiftRows, MixColumns, and AddRoundKey. SubBytes is a simple lookup table using a 16×16 matrix of byte values called an s-box. ShiftRows is a simple permutation. MixColumns is a matrix multiplication in GF(2^8). AddRoundKey are bitwise XORed between the 128 bits of state and the 128 bits of the round key.

An S-box used in the AES algorithm generated by two steps. Firstly, we use the multiplicative inverse in the field GF (2^8). Secondly, we apply Affine transformation as follows:

$$b_i' = b_i \oplus b_{(i+4) \bmod 8} \oplus b_{(i+5) \bmod 8} \oplus b_{(i+6) \bmod 8} \oplus b_{(i+7) \bmod 8} \oplus c_i$$

for $0 \leq i \leq 8$, which b_i is i –th bit of the byte, and c_i is i –th bit of *byte c* with value (63) or (01100011).

g. S-box S1 and S2 SEED Block Cipher

S-box used in the SEED block cipher algorithm is defined as follows [8]:

$$S_i : Z_{2^8} \rightarrow Z_{2^8}, S_i(x) = A^{(i)} . x^{n_i} \oplus b_i$$

where $n_1 = 247, n_2 = 251, b_1 = 169, b_2 = 56$.

This S-boxisan Affine transformation of x^{n_i}. The matrix used in the S-box are:

$$A^{(1)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}, \qquad A^{(2)} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}$$

h. S-box Whirlpool Hash Function

Whirlpool S-box is composed of three "mini-box" sized 4×4 (E, E^{-1}, and R). The three mini-box can be seen on [9]. The structure of Whirlpool's S-box can be seen in Figure 1.

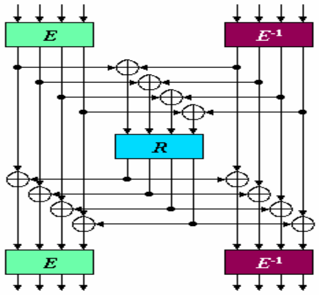


Figure 1. Structure of Whirlpool's S-box

3. RESEARCH METHOD

We perform the SAC test on AES algorithm as a whole while the tests performed on SEED's S-box and Whirlpool's S-box are SAC, BIC, XOR Table Distribution and Nonlinearity tests. SAC test is applied to measure the level of confusion and diffusion in the AES algorithm. AES algorithm used in this study has 128-bits plaintext input and key input.Therefore, the total population of plaintexts and keys is 2^{128} each.

SAC testing on the modified AES algorithm is done in two phase. The first phase is to generate as many as 20000 samples of independent variables that have been defined. When the plaintexts are treated as an independent variables then the key as the control variables are held constant with a value of zero. Similarly, when the keys are treated as independent variables then the plaintext as control variables are held constant with a value of zero. We use constant zero value on the control variables in order to eliminate the influence of the control variables. The output of this process is the value of the dependent variable (ciphertext). The second phase is the testing of the samples that have been generated using the SAC test calculations. SAC test is done every round in order to determine the round position that satisfies the SAC properly.

In this study, the S-box is stated to fulfill all the criteria of the test if the test results from each S-box approach to the test results of the AES's S-box. While the AES algorithm that apply different S-boxes is said to fulfill the SAC test if the test result is close to 50% or close to the test results of the original AES algorithm.

Table 1.The Research Variables of S-box and AES Algorithm

No.	Test	Object of the test	Variable		
			Input		Output
			Independent	Control	Dependent
!.	SAC	AES algorithm withdifferent S-boxes	Key	Plaintext	Ciphertext
			Plaintext	Key	Ciphertext
		Whirlpool's S-box, S1 and S2 SEED'sS-box	Input S-box	-	Output S-box
2.	BIC	Whirlpool's S-box, S1 and S2 SEED'sS-box	Input S-box	-	Output S-box
3.	XOR Table	Whirlpool's S-box, S1 and S2 SEED'sS-box	Input S-box	-	Output S-box
4.	Nonlinearity	Whirlpool's S-box, S1 and S2 SEED'sS-box	Input S-box	-	Output S-box

4. RESULTS AND ANALYSIS

a. Test Results of S-boxes

Based on the test results, it can be determined that the results of the S1 and S2 SEED's S-box close to the test results of AES's S-box but not so with Whirlpool's S-box. Comparison of test results of each S-box is presented in Table 2.

Based on Table 2, on the SAC test, it can be seen that the S-box which has a minimum error value is AES's S-box and S1 SEED's S-box with an error value of 0.125. While the S-box which has a maximum error value is the Whirlpool's S-box with error value of 0.203125. At the BIC test, AES's S-box has the smallest BIC value of 0.13412,while Whirlpool's S-box has the largest BIC value of 0.30693.

Table 2. Comparison of test results of each *S-box*

No.	*S-box*	ε_S	ε_{BIC}	Max XOR	$NLM_{(f)}$ min
1.	Whirlpool	0.203125	0.30693	10	98
2.	S1 SEED	0.125	0.13867	6	111
3.	S2 SEED	0.14063	0.13767	6	111
4.	AES	0.125	0.13412	4	112

At the XOR-Table test, it can be seen that the largest entry is the Whirlpool's S-box with 10 entries as much as 1 piece. This value indicates that there are 10 specific output difference value of 256 possible of output difference. The maximum probability of this S-box is $10/256$. Whirlpool S-box has 10 entries with total 1, means that there is only one pair of input and output difference that produces a certain maximum output difference as many as 10 of 256 possibilities. Therefore, it is vulnerable to differential cryptanalysis.

At the nonlinearity test, AES's S-box which has become a standard, its minimum nonlinearity value is 112, and it is proved that the AES S-box close to perfect nonlinearity value (for 8×8 S-box, the perfect nonlinearity value close to $2^{n-1} - 2^{\frac{n}{2}-1} = 2^7 - 2^3 = 120$). The minimum nonlinearity value of S1 and S2 SEED's S-boxes is 111 close to the value of AES's S-box. However, the Whirlpool's S-box has minimum nonlinearity value of 98, far from perfect nonlinearity value. Moreover, on Whirlpool's S-box, the number of vectors that contained in minimum nonlinearity value only 6 vectors, then it is likely to be susceptible to linear cryptanalysis.

The differences of test results of the S-boxes is influenced by the structure of its construction. S1 and S2 SEED's S-boxes have test results which is almost close to the test results of the AES's S-box because of the structure of the construction used is similar. S1 and S2 SEED's S-boxes are generated using the Affine transformation matrix. It is the same as the AES’s S-box that generated using Affine transformation which also uses a matrix. However, the matrix and Affine transformation used in each S-box are different.

Whirlpool’s S-box generated by the recursive method using three "mini-box" sized 4×4 (E, E^{-1}, and R). To see the randomness of the three mini-box, we conducted SAC and BIC tests. But the test results of the three mini box have an extreme error value. It also led to the Whirlpool's S-box generated is not good, so the test results are away from the value of the AES's S -box standard. Table 3 is the error value of the mini box E, E^{-1}, and R.

Table 3. Error value of Mini Box E, E^{-1}, and R

No.	Mini Box	ε_I	ε_{SAC}
1.	Mini Box E	0,5	0,87738
2.	Mini Box E^{-1}	0,5	0,87738
3.	Mini Box R	0,5	0,87738

b. Test Results of Algorithms with SAC Test

Based on SAC test results on the AES algorithm with different S-box, it can be concluded that all of these algorithms passed the test (see Table 4 and Table 5).The SAC test results of the AES algorithm with different S-box when the keys are treated as independent variables showed that all of them fulfilled confusion property as shown in Table 4. The original AES algorithm achieves good confusion properties after the second round with the minimum value of 48,635% and the maximum value of 51,56%. On the AES algorithm that using S1 and S2 SEED's S-boxes also achieve good confusion properties after the second round with both minimum value of 48,57% and the maximum value of 51,52% and 51,665%, respectively. However, the AES algorithm that using Whirlpool's S-box, after the second round still has SAC maximum value 52,07% and can satisfy the test in the third round with a minimum value of 48,5% and a maximum value of 51, 81%.

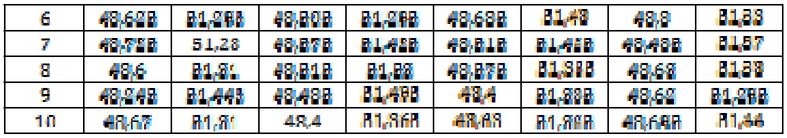
Table 4. SAC test results with key as independent variable

Round	SAC test result of each algorithm (%)							
	AES		AES S1 SEED		AES S2 SEED		AES Whirlpool	
	Min	Max	Min	Max	Min	Max	Min	Max
1	0	100	0	100	0	100	0	100
2	48,635	51,56	48,57	51,52	48,57	51,665	48,235	52,07
3	48,71	51,38	48,81	51,825	48,48	51,46	48,5	51,81
4	48,55	51,35	48,55	51,36	48,695	51,42	48,59	51,555
5	48,785	51,34	48,725	51,305	48,84	51,58	48,78	51,445
6	48,695	51,305	48,37	51,365	48,68	51,505	48,435	51,31
7	48,555	51,555	48,605	51,435	48,76	51,57	48,635	51,495
8	48,68	51,555	48,675	51,355	48,87	51,415	48,725	51,27
9	48,74	51,35	48,635	51,485	48,635	51,545	48,785	51,28
10	48,645	51,36	48,67	51,825	48,58	51,07	48,615	51,37

Based on SAC test results of the AES algorithm with different S-box when the plaintexts are treated as independent variables showed that all of them fulfilled diffusion property as shown in Table 5. The original AES algorithm achieves good diffusion properties after the second round with the minimum value of 48,61% and the maximum value of 51,665%. On the AES algorithm that using S1 and S2 SEED's S-boxes also achieve good confusion properties after the second round with the minimum value of 48,67% and 48,79% and the maximum value of 51,735% and 51,835%, respectively. However, the AES algorithm that using Whirlpool's S-box, after the second round still has SAC maximum value 52,13% and fulfill good diffusion properties in the third round with a minimum value of 48,55% and a maximum value of 51,875%.

Table 5. SAC test results with plaintext as independent variable

Round	SAC test result of each algorithm (%)							
	AES		AES S1 SEED		AES S2 SEED		AES Whirlpool	
	Min	Max	Min	Max	Min	Max	Min	Max
1	0	57,075	0	57,725	0	56,565	0	61,38
2	48,61	51,665	48,67	51,735	48,79	51,835	48,17	52,13
3	48,575	51,4	48,62	51,48	48,58	51,53	48,55	51,875
4	48,58	51,37	48,555	51,48	48,68	51,555	48,645	51,365
5	48,805	51,355	48,55	51,355	48,625	51,44	48,715	51,245



6	48,625	51,265	48,805	51,265	48,685	51,48	48,8	51,33
7	48,725	51,28	48,875	51,425	48,815	51,455	48,495	51,57
8	48,6	51,31	48,815	51,33	48,875	51,315	48,68	51,37
9	48,245	51,445	48,485	51,495	48,4	51,335	48,62	51,285
10	48,67	51,31	48,4	51,365	48,49	51,365	48,685	51,44

5. CONCLUSION

In this study, we analyze Whirlpool's S-box and S1 and S2 SEED's S-boxes using SAC, BIC, XOR Table Distribution, and nonlinearity test. After that, the three S-boxes are applied to the AES. By applying the three S-boxes, the modified AES algorithm still has the good confusion and diffusion properties. The difference lies only in the round position of the algorithm that satisfies good confusion and diffusion properties.It means that AES has a good design structure, even when applied by an S-box that not good, it still achieves good confusion and diffusion for a full round of the algorithm.In future, further research needs to be done to the other components in the AES algorithm to determine the level of confusion and diffusion of the algorithm if the modification done to it.

REFERENCES

[1] NIST,**“**Federal Information Processing Standards Publication (FIPS) 197”,*Springfield : National Institute of Standards and Technology (NIST)*, 2001.

[2] B. Schneier,“Applied Cryptography: Protocols, Algorithms, and Source Code in C, Second Edition”,New York: John Wiley & Sons, Inc, 1996.

[3] A.M. Youssef, “Analysis and Design of Block Cipher”, PhD Thesis, Queen’s University, Canada, 1997.

[4] M.D.Yucel and I. Vergili, “Avalanche and Bit Independence Properties for the Ensembles of Randomly Chosen nxn S-boxes”, EE Department of METU, Turkey, 2001.

[5] S. Mister and Carlisle Adams, “Practical S-box Design”, Canada, 1996.

[6] A.F. Webster and S.E.Tavares,“On the design of S-boxes”, Department of Electrical Engineering. Queen’s University, 1989.

[7] S. Kavut and M.D. Yucel, “On Some Cryptographic Properties of Rijndael”, Middle East Technical University, 2000.

[8] Korea Information Security Agency, “SEED Algorithm Specification”.

[9] P. Barreto and V. Rijmen,”The Whirlpool Hashing Function”, 2003.

BIBLIOGRAPHY OF AUTHORS

	Novita Angraini is a student at Sekolah Tinggi Sandi Negara, Bogor. Her interests are in mathematics, cryptography, and its related.
	Bety Hayat Susanti is a lecturer at Sekolah Tinggi Sandi Negara, Bogor. She obtained undergraduate degree in Math and master’s degree in Planning and Public Policy at University of Indonesia. Her interests are in mathematics, cryptography, and its related.
	Magfirawaty is a lecturer at Sekolah Tinggi Sandi Negara, Bogor. She obtained both undergraduate and master’s degree in Physics at University of Indonesia. Her interests are in physics, cryptography, and its related.