EXAMINATION OF SUBSTITUTION BOXES OF SAFER

E. Aras* and M. D. Yücel**
***ASELSAN Inc., MGEO Division, P.O. Box: 30, Etlik 06011, Ankara**
****Dept. of Electrical and Electronics Eng., Middle East Technical University, 06531, Ankara**

Abstract–**Two operations, called "exponentiating box" and "logarithm-taking box" which we call S-boxes of SAFER family of ciphers, are the "only nonlinear layers" of these ciphers and they apply two different "highly nonlinear" transformations, which map 8-bit inputs to 8-bit outputs. Therefore their characteristics have significant effects on the strength of the entire system. The characteristics of S-boxes of SAFER family of ciphers are examined for the criteria of strict avalanche, bit independence, and XOR table distribution. Our experiments show that the "exponentiating" S-box has a weakness for the input difference of 128 (=10000000$_2$) and the "logarithm-taking" S-box has a weakness for the input difference of 253 (=11111101$_2$).**

I. INTRODUCTION

Secure And Fast Encryption Routine with a Key of length 64 bits [1] (SAFER K-64) is a symmetric (one-key) block cipher, which was designed by J. L. Massey. SAFER K-64 is a byte-oriented block-enciphering algorithm. The block length is 8 bytes (64 bits) for plaintext and ciphertext; the user-selected key is also 8 bytes (64 bits) in length. SAFER K-64 is the first designed cipher of the SAFER family of ciphers consisting of SAFER K-64, SAFER K-128, SAFER SK-64, SAFER SK-128, and SAFER SK-40. The block size of all the ciphers in the SAFER family is 64 bits, while the key length is 40 or 64 or 128 bits as indicated in the name of the cipher. The other ciphers in the SAFER family differ from SAFER K-64 only in their key schedules and in the number of rounds used. The encryption round structure of SAFER K-64 is shown in Figure 1. The operations labeled "$45^{(.)}$" and "log$_{45}$" in Figure 1 are the "only nonlinear layers" of the cipher and they apply two different "highly nonlinear" transformations to their inputs. Therefore their characteristics have significant effects on the strength of the entire system. These two operations are called "exponentiating box" and "logarithm-taking box" which we call S-boxes of SAFER family of ciphers. They are used both in the encryption and decryption, but in different locations of the round structures, since the encryption and decryption are slightly different. The S-boxes of SAFER family of ciphers are built on a mathematical structure, in that;

- the operation labeled "$45^{(.)}$" in Figure 1, which notation is to suggest that if the byte input is the integer j the byte output is 45^j modulo 257 (except that this output is taken to be 0 if the modular result is 256, which occurs for j = 128), and
- the operation labeled "log$_{45}$" in Figure 1, which notation is to suggest that if the byte is the integer j then the byte output is log$_{45}$(j) (except that this output is taken to be 128 if the input bit is j = 0).

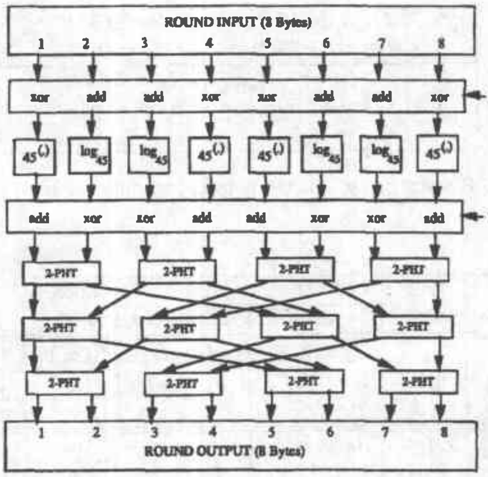

Figure 1: Encryption Round Structure of SAFER K-64

II. DEFINITIONS

1. Completeness: The idea of completeness was introduced by Kam and Davida [2]. If a cryptographic transformation is complete, then each ciphertext bit must depend on all of the plaintext bits. Thus, if it were possible to find the simplest Boolean expression for each ciphertext bit in terms of the plaintext bits, each of those expressions would have to contain all of the plaintext bits if the function was complete. Alternatively, if there is at least one pair of n-bit plaintext vectors P and P_i that differ only in bit i (P_i = P ⊕ e_i, and e_i is the n-bit unit vector with a 1 in position i), and f(P) and f(P_i) differ at least in bit j for all { (i, j) | $1 \le i, j \le n$ }, then the function f must be complete.

2. Avalanche Criterion: The idea of avalanche was introduced by Feistel [3]. For a given transformation to exhibit the avalanche effect, an average of one half of the output bits should change whenever a single input bit is complemented. In order to determine whether a given n x n (*n input bits and n output bits*) function f satisfies this requirement, the 2^n plaintext pairs, P and P_i, such that P and P_i differ only in bit i ($P_i = P \oplus e_i$, and e_i is the n-bit unit vector with a 1 in position i) are used to calculate the 2^n exclusive-or sums, $C_d = f(P) \oplus f(P_i)$. These exclusive-or sums will be referred to as avalanche vectors each of which contains n bits, or avalanche variables. If this procedure is repeated for all i such that $1 \le i \le n$, and one half of the avalanche variables are equal to 1 for each i, then the function f has good avalanche effect.

3. Strict Avalanche Criterion: The concepts of the completeness and the avalanche effect were combined by Webster and Tavares [4] to define the strict avalanche criterion (SAC). If a cryptographic function is to satisfy the strict avalanche criterion, then each output bit should change with a probability of one half whenever a single input bit is complemented. Consider P and P_i, two n-bit, binary plaintext vectors, such that P and P_i differ only in bit i ($P_i = P \oplus e_i$, and e_i is the n-bit unit vector with a 1 in position i), $1 \le i \le n$.

Let $C_d = f(P) \oplus f(P_i)$ and f is the cryptographic transformation, under consideration. If f is to meet the strict avalanche criterion, the probability that each bit in the avalanche vector C_d is equal to 1 should be one half over the set of all possible plaintext vectors P and P_i. This should be true for all values of i. Therefore; completeness and avalanche effect are necessary conditions if the strict avalanche criterion is to be met.

In addition, f is said to satisfy maximum order SAC (MOSAC) if for all j such that $1 \le j \le n$, flipping any combination of one or more input bits changes output bit j with probability one half. SAC is a subset of MOSAC.

Normalized distance to SAC and normalized distance to MOSAC are the measures of the closeness of the cipher function f to SAC and MOSAC respectively. We define normalized distance to SAC for the j^{th} avalanche variable as follows;

$$\{ D_{SAC}[j] \mid P_d = e_i \} = \frac{1}{2^{n-1}} \left| 2^{n-1} - \sum_{all(P,\ P_i)} C_d[j] \right| \quad (1)$$

$$D^{max}_{SAC}[j] = \max_{P_d \in \{1,2,4,\ldots,2^{n-1}\}} \{ D_{SAC}[j] \mid P_d = e_i \} \quad (2)$$

where n is the number of input/output bits of f, e_i is the n-bit unit vector with a 1 in position i, P_d is the input difference between input pairs (P, P_i), $C_d[j]$ is the j^{th} avalanche variable of the avalanche vector, C_d ($= f(P) \oplus f(P \oplus e_i)$), and $1 \le i, j \le n$. If SAC is satisfied, then D^{max} is 0, which is the ideal case. In the worst case, D^{max} equals to 1.

And we define normalized distance to MOSAC for the j^{th} avalanche variable as follows;

$$\{ D_{MOSAC}[j] \mid P_d = \delta \} = \frac{1}{2^{n-1}} \left| 2^{n-1} - \sum_{all(P,\ P^l)} C_d[j] \right| \quad (3)$$

$$D^{max}_{MOSAC}[j] = \max_{1 \le P_d \le 2^n - 1} \{ D_{MOSAC}[j] \mid P_d = \delta \} \quad (4)$$

where n is the number of input/output bits of f, δ is the n-bit binary representation of any integer in the interval [1, 2^n-1], P_d is the input difference between input pairs (P, P^l), $C_d[j]$ is the j^{th} avalanche variable of the avalanche vector, C_d ($= S(P) \oplus S(P \oplus \delta)$), and $1 \le j \le n$. If MOSAC is satisfied, then D^{max} is 0, which is the ideal case. In the worst case, D^{max} equals to 1.

4. Bit Independence Criterion: The idea of bit independence criterion (BIC) was introduced by Webster and Tavares [4]. For a given set of avalanche vectors generated by the complementing of a single plaintext bit, all the avalanche variables should be pairwise independent. Alternatively, consider P and P_i, two n-bit, binary plaintext vectors, such that P and P_i differ only in bit i ($P_i = P \oplus e_i$, and e_i is the n-bit unit vector with a 1 in position i), $1 \le i \le n$. Let $C_d = f(P) \oplus f(P_i)$ and f is the cryptographic transformation, under consideration. If f is to meet the bit independence criterion, the bits j and k in C_d change independently for all i, j, k ($1 \le j$, $k \le n$ with $j \ne k$).

In order to measure the degree of independence between a pair of avalanche variables, their correlation coefficient ρ is calculated. For two variables j and k;

$$\rho\{j, k\} = \frac{\text{cov}\{j, k\}}{\sigma\{j\}\sigma\{k\}} \quad (5)$$

where

$\rho\{j, k\}$ = correlation coefficient of j and k

cov$\{j, k\}$ = covariance of j and k = $E\{jk\} - E\{j\}E\{k\}$

$\sigma^2\{j\} = E\{j^2\} - (E\{j\})^2$

$E\{j\}$ = expected value of j

For the case of binary variables, a correlation coefficient of 0 means that the variables are independent. In addition, the variables will always be identical if the correlation coefficient equals 1, and a value of –1 means that they will always be complements of one another.

In addition, f is said to satisfy maximum order BIC (MOBIC) if the same output bit independence holds whenever flipping any combination of one or more input bits. BIC is a subset of MOBIC.

For the criteria of BIC, if correlation coefficient is calculated for every pair of avalanche variables, a correlation matrix and a maximum correlation matrix of sizes n x n are defined with elements:

$$B_{BIC}(j,\ k \mid P_d = e_i) = \rho\{ C_d[j]\ ,\ C_d[k] \} \quad (6)$$

$$B^{max}_{BIC}(j,\ k) = \max_{P_d \in \{1,2,4,\ldots,2^{n-1}\}} \{ B_{BIC}(j,\ k \mid P_d = e_i) \} \quad (7)$$

where n is the number of input/output bits of f, e_i is the n-bit unit vector with a 1 in position i, P_d is the input difference between input pairs (P, P$_i$), C$_d[j]$ is the j^{th} avalanche variable of the avalanche vector, C$_d$ ($= f(\mathrm{P}) \oplus f(\mathrm{P} \oplus e_i)$), and $1 \leq i, j, k \leq$ n.

Similarly, for the criteria of MOBIC, if correlation coefficient is calculated for every pair of avalanche variables, a correlation matrix and a maximum correlation matrix of sizes n x n are defined with elements:

$$B_{MOBIC}(j,\ k \mid P_d = \delta) = \rho\{\ C_d[j]\ ,\ C_d[k]\ \} \quad (8)$$

$$B_{MOBIC}^{\max}(j,\ k) = \max_{1 \leq P_d \leq 2^n - 1}\{\ B_{MOBIC}(j,\ k \mid P_d = \delta)\ \} \quad (9)$$

where n is the number of input/output bits of f, δ is the n-bit binary representation of any integer in the interval [1, 2^n-1], P$_d$ is the input difference between input pairs (P, P^i), C$_d[j]$ is the j^{th} avalanche variable of the avalanche vector, C$_d$ (= S(P) ⊕ S(P⊕δ)), and $1 \leq j, k \leq$ n.

5. XOR Table Distribution: Differential cryptanalysis [5], a powerful cryptanalytic attack, requires knowledge of the XOR tables of substitution boxes (S-boxes). For an n x n S-box, the XOR table has a size of 2^n x 2^n with its rows and columns indexed by 0, 1, 2, … , 2^n-1. Position $[i, j]$ in the XOR table contains the value;

$$\left|\ \left\{X \in \{0,\ 1\}^n : S(X) \oplus S(X \oplus \eta_i) = \eta_j\right\}\ \right| \quad (10)$$

such that $0 \leq i,\ j \leq 2^n-1$, and η_i and η_j are n-bit binary representations of indices i and j. The pair (i, j) is called an input/output XOR pair. Differential cryptanalysis exploits such XOR pairs with large XOR table entries. A cipher can be secured against differential cryptanalysis by selecting S-boxes with low XOR table entries, ideally 0 or 2 (the one exception is the entry (0, 0) which has the value of 2^n). The sum of the XOR table entries on the each row is equal to 2^n, which is the total number of input vector pairs (X, X ⊕ η$_i$).

III. RESULTS

Exponentiating S-Box

a) SAC and MOSAC:

For the exponentiating S-box, $\{D_{SAC}[j] \mid P_d = e_i\}$ curves, given by (1), are depicted in Figure 2. In the figure there are eight curves each obtained for one of the eight 8-bit unit vector input differences, e_1, …, e_8. Those curves are merged into a single one, $D_{SAC}^{\max}[j]$ curve, by (2) in that it takes the maximum of normalized distance to SAC values for each avalanche variable. But, the $D_{SAC}^{\max}[j]$ curve is not depicted since it is almost the same as the curve $\{D_{SAC}[j] \mid P_d = e_8(=128_{10})\}$ and (1) gives more valuable information than (2) will give. The normalized distance to MOSAC values for all possible 255 input differences are calculated by (3).

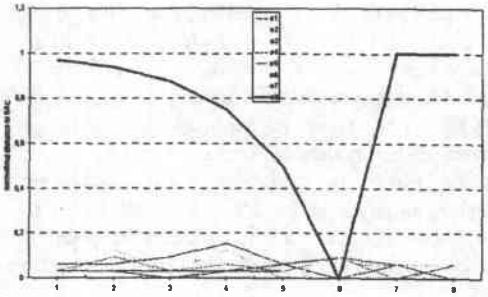

Figure 2: $\{D_{SAC}[j] \mid P_d = e_i\}$ versus j curves for the exponentiating S-box

Instead of drawing all these curves in the same figure the $D_{MOSAC}^{\max}[j]$ curve, given by (4), is depicted in Figure 3. Actually, this curve is also nearly the same as the curve $\{D_{SAC}[j] \mid P_d = 128_{10}\}$ in Figure 2. The only difference is at the 6th avalanche variable and occurs for the input difference of 137 (=10001001$_2$). At all other avalanche variables, the maxima occur for the input difference of 128 (=10000000$_2$). It is observed that normalized distance to (MO)SAC for all avalanche variables other than the 6th are considerably high and the strict avalanche criteria completely fails at the 7th and 8th avalanche variables where $D_{(MO)SAC}^{\max}$ is 1.

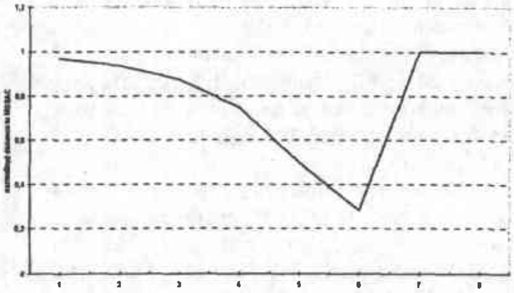

Figure 3: $D_{MOSAC}^{\max}[j]$ versus j curve for the exponentiating S -box

b) BIC and MOBIC:

The $B_{BIC}^{\max}$ and $B_{MOBIC}^{\max}$ matrices are calculated by (7) and (9) respectively as follows;

$$B_{BIC}^{\max} = \begin{bmatrix} 1{,}00 & 0{,}70 & 0{,}48 & 0{,}33 & -0{,}59 & 0{,}15 & 1{,}00 & 1{,}00 \\ 0{,}70 & 1{,}00 & 0{,}69 & 0{,}47 & 0{,}31 & -0{,}30 & 1{,}00 & 1{,}00 \\ 0{,}48 & 0{,}69 & 1{,}00 & 0{,}68 & 0{,}44 & 0{,}25 & 1{,}00 & 1{,}00 \\ 0{,}33 & 0{,}47 & 0{,}68 & 1{,}00 & 0{,}65 & 0{,}37 & 1{,}00 & 1{,}00 \\ -0{,}59 & 0{,}31 & 0{,}44 & 0{,}65 & 1{,}00 & 0{,}57 & 1{,}00 & 1{,}00 \\ 0{,}15 & -0{,}30 & 0{,}25 & 0{,}37 & 0{,}57 & 1{,}00 & 1{,}00 & 1{,}00 \\ 1{,}00 & 1{,}00 & 1{,}00 & 1{,}00 & 1{,}00 & 1{,}00 & 1{,}00 & 1{,}00 \\ 1{,}00 & 1{,}00 & 1{,}00 & 1{,}00 & 1{,}00 & 1{,}00 & 1{,}00 & 1{,}00 \end{bmatrix}$$

$$B_{MOBIC}^{\max} = \begin{bmatrix} 1.00 & 0.70 & 0.48 & 0.33 & -0.59 & -0.25 & 1.00 & 1.00 \\ 0.70 & 1.00 & 0.69 & 0.47 & 0.31 & -0.30 & 1.00 & 1.00 \\ 0.48 & 0.69 & 1.00 & 0.68 & 0.44 & -0.31 & 1.00 & 1.00 \\ 0.33 & 0.47 & 0.68 & 1.00 & 0.65 & 0.37 & 1.00 & 1.00 \\ -0.59 & 0.31 & 0.44 & 0.65 & 1.00 & 0.57 & 1.00 & 1.00 \\ -0.25 & -0.30 & -0.31 & 0.37 & 0.57 & 1.00 & 1.00 & 1.00 \\ 1.00 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 \\ 1.00 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 \end{bmatrix}$$

As seen from the matrices, the correlation coefficient between the 7th and any other avalanche variable, and between the 8th and any other avalanche variable is *1,00* (actually "undefined" since the variance of the avalanche variable is 0 for the 7th and 8th bit positions of the avalanche vector). A search over all correlation matrices defined by (8) shows that these undefined rows correspond to an input difference of 128. Other values in the $B_{MOBIC}^{\max}$ matrix are also quite close to *1*, which means that the avalanche variables are highly correlated.

c) XOR Table Distribution:

The XOR table is a matrix of 256 x 256, whose entries are calculated by (10). If it is divided into 8 pieces, so that each piece is 32 x 256, the maximum entry for each piece is as follows;

1st piece:max. entry = 12 for (i, j) = (21, 184)
2nd piece:max. entry = 16 for (i, j) = (53, 68)
3rd piece:max. entry = 22 for (i, j) = (64, 60)
4th piece:max. entry = 12 for (i, j) = (112, 101)
5th piece:max. entry = 128 for (i, j) = (128, 253)
6th piece:max. entry = 16 for (i, j) = (181, 185)
7th piece:max. entry = 22 for (i, j) = (192, 120)
8th piece:max. entry = 16 for (i, j) = (237, 120)

The maximum entry is 128 for the whole XOR table and occurs for the position [128, 253], which means that when $P_d=128_{10}$, the avalanche vector $C_d=253_{10}$ occurs for 50% of the overall input pairs since the highest possible value is $2^8=256$. The XOR table distribution test also verifies the previous tests in that the maximum table entry occurs for the input difference of 128.

Logarithm-taking S-Box

a) SAC and MOSAC:

For the logarithm-taking S-box, $\{D_{SAC}[j] \mid P_d = e_i\}$ curves, given by (1), are depicted in Figure 4. In the figure there are eight curves each obtained for one of the eight 8-bit unit vector input differences, e_1, ..., e_8. It is seen from the Figure 4 that normalized distances to SAC for all avalanche variables are below *0,25*, which is quite good. Those curves are merged into a single one, $D_{SAC}^{\max}[j]$ curve, by (2) in that it takes the maximum of normalized distance to SAC values for each avalanche variable. The $D_{SAC}^{\max}[j]$ curve is not depicted since (1) gives more valuable information than (2) will give. The normalized distance to MOSAC values for all possible 255 input differences are calculated by (3). Instead of drawing all these curves in the same figure the $D_{MOSAC}^{\max}[j]$ curve, given by (4), is depicted in Figure 5. In Figure 5 the maxima, which are about *0,5*, occur for the input difference of 253 ($=11111101_2$) at all avalanche variables; hence, we obtain the same curve if $\{D_{SAC}[j] \mid P_d = 253_{10}\}$ is depicted.

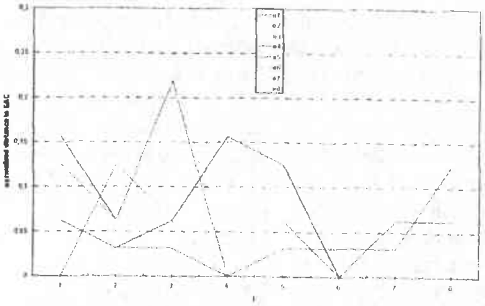

Figure 4: $\{D_{SAC}[j] \mid P_d = e_i\}$ versus *j* curves for the logarithm-taking S-box

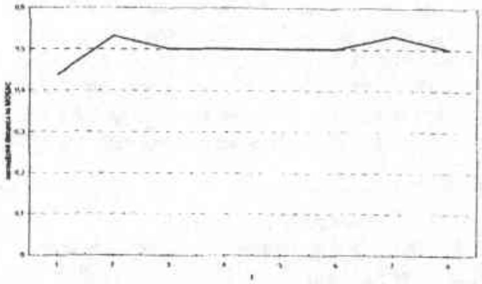

Figure 5: $D_{MOSAC}^{\max}[j]$ versus *j* curve for the logarithm-taking S-box

b) BIC and MOBIC:

The $B_{BIC}^{\max}$ and $B_{MOBIC}^{\max}$ matrices are calculated by (7) and (9) respectively as follows:

$$B_{BIC}^{\max} = \begin{bmatrix} 1.00 & 0.18 & -0.18 & 0.21 & 0.16 & -0.12 & -0.13 & -0.25 \\ 0.18 & 1.00 & 0.09 & 0.06 & -0.22 & 0.18 & 0.15 & -0.19 \\ -0.18 & 0.09 & 1.00 & -0.25 & 0.11 & 0.15 & 0.24 & -0.12 \\ 0.21 & 0.06 & -0.25 & 1.00 & 0.26 & 0.09 & -0.16 & -0.31 \\ 0.16 & -0.22 & 0.11 & 0.26 & 1.00 & 0.06 & -0.12 & 0.19 \\ -0.12 & 0.18 & 0.15 & 0.09 & 0.06 & 1.00 & -0.15 & 0.07 \\ -0.13 & 0.15 & 0.24 & -0.16 & -0.12 & -0.15 & 1.00 & -0.19 \\ -0.25 & -0.19 & -0.12 & -0.31 & 0.19 & 0.07 & -0.19 & 1.00 \end{bmatrix}$$

$$B_{MOBIC}^{\max} = \begin{bmatrix} 1,00 & -0,28 & -0,44 & 0,35 & -0,48 & 0,34 & 0,33 & -0,40 \\ -0,28 & 1,00 & -0,34 & 0,36 & -0,28 & 0,40 & 0,34 & 0,31 \\ -0,44 & -0,34 & 1,00 & 0,33 & -0,50 & 0,37 & 0,27 & 0,40 \\ 0,35 & 0,36 & 0,33 & 1,00 & 0,33 & 0,25 & 0,31 & -0,31 \\ -0,48 & -0,28 & -0,50 & 0,33 & 1,00 & 0,29 & -0,30 & -0,37 \\ 0,34 & 0,40 & 0,37 & 0,25 & 0,29 & 1,00 & 0,27 & 0,37 \\ 0,33 & 0,34 & 0,27 & 0,31 & -0,30 & 0,27 & 1,00 & -0,53 \\ -0,40 & 0,31 & 0,40 & -0,31 & -0,37 & 0,37 & -0,53 & 1,00 \end{bmatrix}$$

c) XOR Table Distribution:

The XOR table is a matrix of 256 x 256, whose entries are calculated by (10). If it is divided into 8 pieces, so that each piece is 32 x 256, the maximum entry for each piece is as follows;

1st piece:max. entry = 12	for (i, j) = (13, 64)
2nd piece:max. entry = 22	for (i, j) = (60, 64)
3rd piece:max. entry = 16	for (i, j) = (68, 53)
4th piece:max. entry = 22	for (i, j) = (120, 192)
5th piece:max. entry = 12	for (i, j) = (133, 109)
6th piece:max. entry = 16	for (i, j) = (185, 181)
7th piece:max. entry = 16	for (i, j) = (193, 192)
8th piece:max. entry = 128	for (i, j) = (253, 128)

The maximum entry is 128 for the whole XOR table and occurs for the position [253, 128], which means that when P_d=253_{10}, the avalanche vector C_d=128_{10} occurs for 50% of the overall input pairs since the highest possible value is 2^8=256. The XOR table distribution test also verifies the SAC test in that the maximum table entry occurs for the input difference of 253, where SAC test has its maxima.

IV. CONCLUSIONS

The exponentiating S-box has a weakness for the input difference of 128 (=10000000_2). In order to compare the exponentiating S-box and the logarithm-taking S-box better in terms of SAC, the $D_{MOSAC}^{\max}[j]$ curves, given by (4), are depicted in Figure 6. As seen from the solid curve in Figure 6, none of the avalanche variables obey the SAC; moreover, it is observed from Figure 2 and Figure 6 that $D_{MOSAC}^{\max}[j]$ = $\{D_{SAC}[j] \mid P_d = 128_{10}\}$ for all j, except j = 6.

For the same input difference of 128, the resulting outputs (C, C^1) always have the same bit values in their 7th bit and the complement bit values in their 8th bit positions, which make the last two rows of $B_{BIC}^{\max}$ and $B_{MOBIC}^{\max}$ matrices undefined. Other values in $B_{MOBIC}^{\max}$ are also quite close to *1*, which means that the avalanche variables are highly correlated, and many of them are the same as the elements of $B_{BIC}(j,\ k \mid P_d = 128_{10})$, given by (6). The XOR table distribution test also verifies SAC and BIC tests in that the maximum table entry occurs for the input difference of 128.

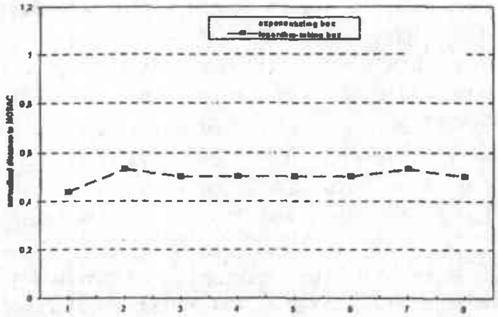

Figure 6: $D_{MOSAC}^{\max}[j]$ versus *j* curves for the exponentiating and logarithm-taking S-boxes

The logarithm-taking S-box has a weakness for the input difference of 253 (=11111101_2). The dashed curve in Figure 6 corresponds to $D_{MOSAC}^{\max}[j]$ and is equal to $\{D_{SAC}[j] \mid P_d = 253_{10}\}$ for all j. However, normalized distance to MOSAC for all j is about *0,5*, which is better than the case of exponentiating S-box.

Many elements of $B_{MOBIC}^{\max}$ are the same as the elements of $B_{BIC}(j,\ k \mid P_d = 253_{10})$, and the maximum entry of the $B_{MOBIC}^{\max}$ is *–0,53.*

The maximum entry of the XOR table also occurs for the input difference of 253, where SAC test has its maxima.

Finally, although the logarithm-taking S-box seems to be more resistive to the attacks than the exponentiating S-box, yet it has a weakness for the input difference of 253 and the exponentiating S-box has a weakness for the input difference of 128.

REFERENCES

1. *J.L. Massey, SAFER K-64: A Byte-Oriented Block-Ciphering Algorithm. Fast Software Encryption – Proc. Cambridge Security Workshop, Cambridge, U.K., LNCS 809, pages 1-17. Springer Verlag, 1994.*

2. *J.B. Kam and G.I. Davida, Structured design of substitution-permutation encryption networks, IEEE Transactions on Computers, Vol. C-28, No. 10, pp. 747-753, October 1979.*

3. *H. Feistel, Cryptography and computer privacy, Scientific American, Vol. 228, No. 5, pp. 15-23, May 1973.*

4. *A.F. Webster and S.E. Tavares, On the Design of S-Boxes, Advances in Cryptology: Proceedings of CRYPTO'85, Springer Verlag, New York, 1986, pp.523-534.*

5. *E. Biham and A. Shamir, Differential cryptanalysis of DES-like cryptosystems, Journal of Cryptology, Vol. 4, No. 1, pp. 3-72, 1991.*