Turk J Elec Engin, VOL.9, NO.2 2001, © TÜBİTAK

Performance Evaluation of Safer K-64 and S-Boxes of the Safer Family

Ekrem ARAS, Melek D. YÜCEL
Electrical Engineering Department of Middle East Technical University,
06531 Ankara-TURKEY

Abstract

If the characteristics of s-boxes of the SAFER family of ciphers are examined for the criteria of strict avalanche, bit independence, and XOR table distribution, experiments show that the "exponentiating" s-box has a weakness for an input difference of 128 (=10000000_2) and the "logarithm-taking" s-box has a weakness for an input difference of 253 (=11111101_2). However, since these experiments are performed by isolating the s-boxes from the general structure, they do not necessarily indicate a weakness in the overall algorithm. We propose a quick and rough test method, called the avalanche weight distribution criterion, to evaluate the overall performance of block ciphers. We then apply this novel criterion and the conventional strict avalanche criterion to SAFER K-64, and show that the algorithm passes both tests successfully despite the specific weaknesses of its isolated s-boxes.

1. Introduction

Secure **A**nd **F**ast **E**ncryption **R**outine with a **K**ey of length **64** bits [1] (SAFER K-64) is a symmetric (one-key) block cipher, designed by James L. Massey. SAFER K-64 is the first designed cipher of the SAFER family of ciphers, which differ only in their key schedules and in the number of rounds used.

The encryption and decryption blocks of the SAFER family of ciphers contain two nonlinear operations, called the "exponentiating box" and "logarithm-taking box", which may have significant effects on the strength of the entire system. If the "exponentiating" and "logarithm-taking" boxes of the SAFER family of ciphers are examined with respect to the criteria of completeness, avalanche, strict avalanche (SAC) and bit independence (BIC), by considering them to be isolated from the general structure of the cipher [2,3,4], experiments show that the "exponentiating" s-box has a weakness for an input difference of 128 (=10000000_2) and the "logarithm-taking" s-box has a weakness for an input difference of 253 (corresponding to 11111101_2). However, since these experiments are performed by isolating the boxes from the general structure, these results do not indicate an overall weakness in the SAFER algorithms.

We propose a quick and rough test method, called the Avalanche Weight Distribution (AWD) criterion, to evaluate the overall performance of block ciphers [4,5,6]. We then apply this novel criterion and the conventional strict avalanche criterion to SAFER K-64, and show that the cryptographic algorithm passes both tests successfully.

In Section II, we introduce the **A**valanche **W**eight **D**istribution (*AWD*) criterion after briefly discussing the conventional criteria of completeness, avalanche, strict avalanche, bit independence and *X-OR* distribution. SAFER K-64 is described and the corresponding s-box results are summarized in Section III. The overall round by round performance of SAFER K-64 is presented in terms of *AWD* curves in Section IV, and with reference to the conventional strict avalanche criterion in Section V. Conclusions are discussed in Section VI.

2. Cryptographic Test Criteria

Ideally, a block cipher should be hard to break, easy to implement, and fast to encrypt. In 1949, Shannon gave us the fundamental theory of symmetric (secret-key) cryptosystems [7], in which he presented the principles of *diffusion* and *confusion*. Since then, these principles have become the essential properties that block ciphers must have and methods of achieving good diffusion and confusion have been at the heart of block cipher design. Actually both principles are defined in quite similar forms and try to achieve the same goal: *hiding the statistical features of the plaintext*.

To design a cipher according to the principle of diffusion means that one designs it to ensure that "*the statistical structure of plaintext which leads to its redundancy is dissipated into long term statistics*" [7]. Lai [8] states the principle of diffusion as follows: "*for virtually every key, the encryption function should be such that there is no statistical dependence between simple structures in the plaintext and simple structures in the ciphertext and that there is no simple relation between different encryption functions*". In other words, every bit of the ciphertext should depend on every bit of the plaintext and every bit of the key, i.e., the effect of changing one bit in the plaintext or in the key should be sensed on all ciphertext bits. This ensures that the statistics of the plaintext are dissipated within the ciphertext so that an attacker cannot predict the plaintext that corresponds to a particular ciphertext, even after observing a number of similar plaintexts and their corresponding ciphertexts.

To design a cipher according to the principle of confusion means that one designs it so as "*to make the relation between the simple statistics of ciphertext and the simple description of key a very complex and involved one*" [7]. The principle of confusion is also stated [8] as follows: "*the dependence of the key on the plaintext and ciphertext should be so complex that cryptanalysis is useless*".

We first describe the conventional test criteria which can be applied either to the overall transformation describing a block cipher or to the isolated s-boxes of substitution permutation networks. We then define our novel criterion of **A**valanche **W**eight **D**istribution (*AWD*), which measures the overall performance roughly but quickly.

2.1. Conventional Test Methods

1) Completeness

The idea of completeness was introduced by Kam and Davida [9]. 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.

2) Avalanche Criterion

The idea of avalanche was introduced by Feistel [10]. 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 function $f: Z_2^n \rightarrow Z_2^n$ (from the n dimensional binary vector space into itself) satisfies this requirement, 2^n plaintext pairs, $\mathbf{P}$ and $\mathbf{P}_i$, such that $\mathbf{P}$ and $\mathbf{P}_i$ differ only in bit i ($\mathbf{P}_i = \mathbf{P} \oplus \mathbf{e}_i$, and $\mathbf{e}_i$ is the n-bit unit vector with a one in position i) are used to calculate the 2^n exclusive-or sums, $\mathbf{C}_d = f(\mathbf{P}) \oplus f(\mathbf{P}_i)$. These output difference vectors $\mathbf{C}_d$ are referred to as avalanche vectors, and their elements are called avalanche variables. If one half of the avalanche variables are equal to 1 for each i in $1 \leq i \leq n$, then the function f has a good avalanche effect.

3) Strict Avalanche Criterion

The concepts of the completeness and the avalanche effect were combined by Webster and Tavares [11] to define the **S**trict **A**valanche **C**riterion (*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 two input vectors which differ only in bit i, with the corresponding avalanche vector $\mathbf{C}_d$. f meets the strict avalanche criterion, if the probability that each bit in the avalanche vector $\mathbf{C}_d$ is equal to 1 is one half over the set of all possible input vectors $\mathbf{P}$ and $\mathbf{P}_i$, for all values of i. Therefore, completeness and the avalanche effect are necessary conditions if the strict avalanche criterion is to be met.

In addition, f is said to satisfy the **M**aximum **O**rder **S**trict **A**valanche **C**riterion (*MOSAC*) if for all j such that $1 \leq j \leq n$, flipping any combination of one or more input bits changes the output bit j with probability one half [12].

Distance to *SAC* and 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:

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

$$D_{SAC}^{\max}[j] = \max_{1 \leq i \leq n} \{D_{SAC}[j] \mid P_d = e_i\} \tag{2}$$

where $\mathbf{e}_i$ is the n-bit unit vector with a 1 in position i, $\mathbf{P}_d$ is the input difference, $\mathbf{C}_d\ [_j]$ is the j^{th} avalanche variable of the avalanche vector $\mathbf{C}_d$. If *SAC* is satisfied perfectly, then $D_{SAC}^{\max}[j]$ is 0 for all output bits, and in the worst case normalized distance $D_{SAC}^{\max}[j]$ is equal to 1.

Similarly, we define normalized distance to *MOSAC* for the j^{th} avalanche variable as

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

$$D_{MOSAC}^{\max}[j] = \max_{1 \leq \delta \leq 2^n - 1} \{D_{MOSAC}[j] \mid P_d = \delta\} \tag{4}$$

where δ is the n-bit binary representation of any integer in $[1, 2^n - 1]$, and the avalanche vector $\mathbf{C}_d = f(\mathbf{P}) \oplus f(\mathbf{P} \oplus \delta)$. If *MOSAC* is satisfied, then $D_{MOSAC}^{\max}[j]$ is 0, which is the ideal case. In the worst case, the normalized distance $D_{MOSAC}^{\max}[j]$ is equal to 1.

4) Bit Independence Criterion

The idea of **B**it **I**ndependence **C**riterion (*BIC*) was introduced by Webster and Tavares [11]. For a given set of avalanche vectors generated by complementing a single plaintext bit, all avalanche variables should be pairwise independent. Alternatively, consider two n-bit input vectors which differ only in bit i, with the corresponding avalanche vector $\mathbf{C}_d$. If f meets the bit independence criterion, the bits j and k in $\mathbf{C}_d$ change independently for all $i,\ j,\ k$ $(1 \leq j,\ k \leq n$ with $j \neq k)$. In order to measure the degree of independence between a pair of avalanche variables, their correlation coefficient $\rho\ (\mathbf{C}_d\ [_j],\ \mathbf{C}_d\ [k])$ is calculated. The cryptographic function f is said to satisfy the **M**aximum **O**rder **B**it **I**ndependence **C**riterion (*MOBIC*) if the same output bit independence holds whenever any combination of one or more input bits are flipped.

The correlation matrix $\mathbf{B}_{BIC}$ and the maximum correlation matrix $B_{BIC}^{\max}$ of size $n \times n$ are defined using the correlation coefficient $\rho\ (\mathbf{C}_d\ [_j],\ \mathbf{C}_d\ [k])$:

$$B_{BIC}\,(j,k|\,P_d = e_i) = \rho(C_d[j], C_d[k]) \tag{5}$$

$$B_{BIC}^{\max}\,(j,k) = \underset{1 \leq i \leq n}{max}\,\{B_{BIC}\,(j,k|\,P_d = e_i)\,\} \tag{6}$$

Similarly, for the criteria of *MOBIC*, the correlation coefficient is calculated for every pair of avalanche variables, and correlation and maximum correlation matrices of size $n \times n$ are defined with elements

$$B_{MOBIC}\,(j,k\mid P_d = \delta) = \rho(C_d\,[j]\,, C_d\,[k]) \tag{7}$$

$$B_{MOBIC}^{\max}\,(j,k) = \underset{1 \leq \delta \leq 2^n - 1}{\max}\{B_{MOBIC}\,(j,k|\,P_d = \delta)\,\} \tag{8}$$

where δ is the n-bit binary representation of any integer in the interval [1, $2^n - 1$], and $\mathbf{C}_d\ [_j]$ is the j^{th} avalanche variable of the avalanche vector $\mathbf{C}_d = f(\mathbf{P}) \oplus f(\mathbf{P} \oplus \delta)$.

5) X-OR Table Distribution

Differential cryptanalysis [13], which is a powerful cryptanalytic attack, requires knowledge of the $D_{MOSAC}^{\max}[j]$ tables of substitution boxes (s-boxes). For an $n \times n$ s-box, the *X-OR* table has a size of $2^n \times 2^n$, with its rows and columns indexed by 0, 1, 2, ..., $2^n - 1$. Position $[i,\ j]$ in the *X-OR* table contains the number of input vectors:

$$|\{P \in \{0,1\}^n : f\,(P) \oplus f\,(P \oplus \eta_i) = \eta_j\}| \tag{9}$$

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. $\mathbf{P}$ is the input vector, $f(.)$ corresponds to the cryptographic function of the s-box, and the pair $(i,\ j)$ is called an input/output *X-OR* pair. Differential cryptanalysis exploits such *X-OR* pairs with large *X-OR* table entries. A cipher can be secured against differential cryptanalysis by selecting s-boxes with low *X-OR* table entries, ideally 0 or 2 (the only exception is the entry (0, 0) which has the value of 2^n). The sum of the *X-OR* table entries on each row is equal to 2^n, which is the total number of input vector pairs ($\mathbf{P}$, $\mathbf{P} \oplus \eta_i$).

2.2. Avalanche Weight Distribution Criterion

We define the *"Avalanche Weight Distribution (AWD) curve"* as the *"histogram of the Hamming weight of the ciphertext difference vector"* [4,5,6]. We then propose *AWD* as a simple criterion for fast and rough

analysis of the diffusion and confusion properties mentioned by Shannon [7].

If the Avalanche Weight Distribution (*AWD*) criterion measures the diffusion properties of block ciphers [4], then we state the criterion as follows: Even for quite similar plaintext pairs ($\mathbf{P}_1$, $\mathbf{P}_2$), histograms of the Hamming weight of the avalanche vectors should be completely random. Hence, *AWD* curves corresponding to all possible pairs of similar inputs should be binomially distributed around $n/2$ for a well diffused block cipher of blocklength n.

If the *AWD* criterion reveals the confusion properties of block ciphers [4], then we state the criterion as follows: For all pairs of similar secret keys ($\mathbf{K}_1$, $\mathbf{K}_2$), histograms of the Hamming weight of the differences of corresponding ciphertext pairs ($\mathbf{C}_1$, $\mathbf{C}_2$) should be binomially distributed around $n/2$ for a well confused block cipher of blocklength n.

3. Test Results for Safer S-Boxes

Secure **A**nd **F**ast **E**ncryption **R**outine with a **K**ey of length 64 bits [1] (SAFER K-64) is a symmetric (one-key) block cipher, 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 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 "$45^{(\cdot)}$" 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. These two operations are called the "exponentiating box" and "logarithm-taking box", which can be considered as the s-boxes of the 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 the SAFER family of ciphers use the two operations defined by:

- the operation labeled "$45^{(\cdot)}$" in Figure 1, which maps the byte input j to the byte output 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 maps the byte input j to the byte output, is $\log_{45}(\mathrm{j})$ (except that this output is taken to be 128 if the input bit is $j = 0$).

The encryption algorithm consists of r rounds of identical transformations that are applied in sequence to the plaintext, followed by an output transformation to produce the final ciphertext. $r = 6$ is recommended but larger values of r can be used, if desired, for even greater security. Each round is controlled by two 8-byte subkeys and the output transformation is controlled by one 8-byte subkey.

We apply the conventional criteria discussed in Section II to SAFER s-boxes and summarize the test results [2,3,4] in Section III.1.

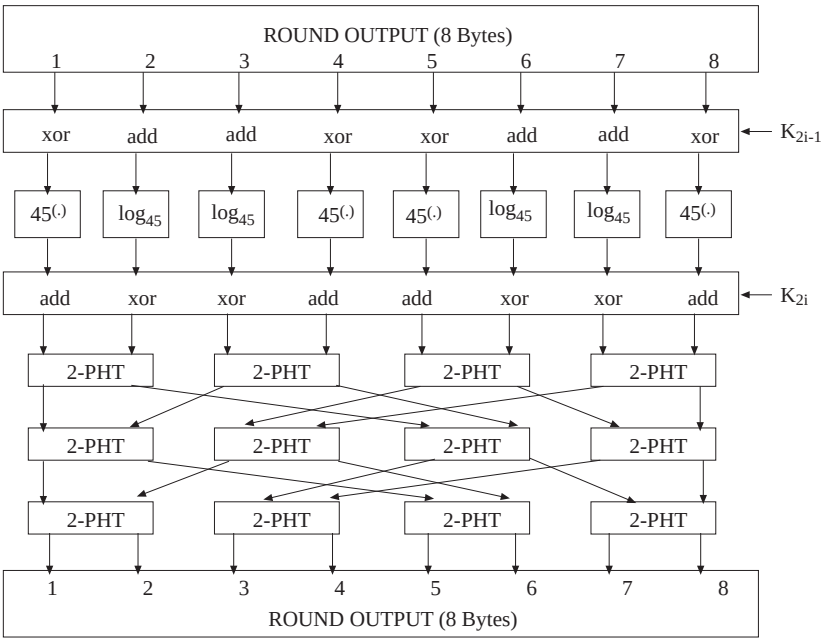

Figure 1. Encryption Round Structure of SAFER K-64

3.1. Exponentiating S-Box

1) SAC and MOSAC

For the exponentiating s-box, $\{D_{SAC}[j]|P_d = e_i\}$ curves, given by (1), are depicted in Figure 2. In the figure there are eight curves corresponding to input differences, $\mathbf{e}_1, \ldots, \mathbf{e}_8$. If those curves are merged into a single one by (2), the maximum of normalized distance to SAC for each avalanche variable, $D^{\max}_{SAC}[j]$, is found to be almost the same as the curve $\{D_{SAC}[j]|P_d = e_8(= 128_{10})\}$.

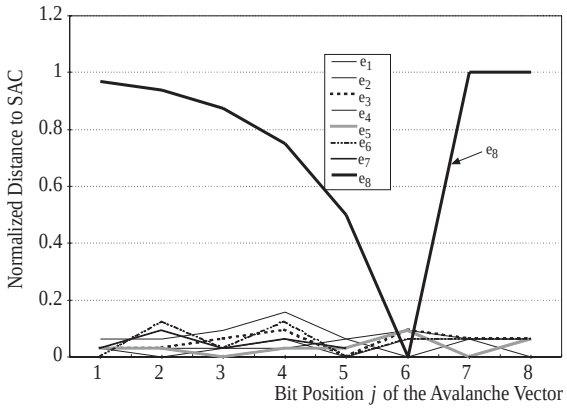

Figure 2. $\{D_{SAC}[j]|P_d = e_i\}$ versus j Curves for the Exponentiating S-box

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, only the $D^{\max}_{MOSAC}[j]$ curve, given by (4), is depicted in Figure 3. Actually, this curve is also nearly the same as the curve $\{D_{SAC}[j]|P_d = 128_{10}\}$ in Figure 2. The only difference is at the 6^{th} avalanche variable and it occurs for an input difference of 137 (=10001001_2).

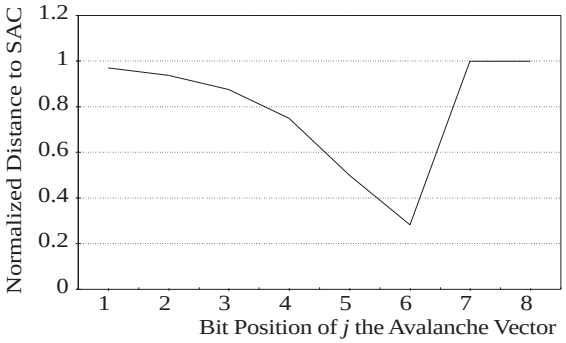

Figure 3. $D^{\max}_{MOSAC}[j]$ versus j Curve for the Exponentiating S–box

At all other avalanche variables, the maxima occur for an input difference of 128 (=10000000_2). It is observed that normalized distance to $MOSAC$ for all avalanche variables other than the 6^{th} are considerably high and SAC completely fails at the 7^{th} and 8^{th} avalanche variables where $D^{\max}_{(MO)SAC}$ is equal to 1.

2) BIC and MOBIC

The $B^{\max}_{BIC}$ and $B^{\max}_{MOBIC}$ matrices are calculated by (6) and (8) respectively as follows:

$$B^{\max}_{BIC} = \begin{bmatrix} +1.00 & +0.70 & +0.48 & +0.33 & -0.59 & +0.15 & \infty & \infty \\ +0.70 & +1.00 & +0.69 & +0.47 & +0.31 & -0.30 & \infty & \infty \\ +0.48 & +0.69 & +1.00 & +0.68 & +0.44 & +0.25 & \infty & \infty \\ +0.33 & +0.47 & +0.68 & +1.00 & +0.65 & +0.37 & \infty & \infty \\ -0.59 & +0.31 & +044 & +0.65 & +1.00 & +0.57 & \infty & \infty \\ +0.15 & -0.30 & +0.25 & +0.37 & +0.57 & +1.00 & \infty & \infty \\ \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty \\ \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty \end{bmatrix}$$

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

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

3) X-OR Table Distribution

The X-OR table is a matrix of size 256x256, whose entries are calculated by (9). If it is divided into 8 pieces, so that each piece is 32x256, the maximum entry for each piece is as follows:

1^{st} piece: max. entry = 12 for (i, j) = (21, 184)

2^{nd} piece: max. entry = 16 for (i, j) = (53, 68)

3^{rd} piece: max. entry = 22 for (i, j) = (64, 60)

4^{th} piece: max. entry = 12 for (i, j) = (112, 101)

5^{th} piece: max. entry = 128 for (i, j) = (128, 253)

6^{th} piece: max. entry = 16 for (i, j) = (181, 185)

7^{th} piece: max. entry = 22 for (i, j) = (192, 120)

8^{th} piece: max. entry = 16 for (i, j) = (237, 120)

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

3.2. Logarithm-Taking S-Box

1) SAC and MOSAC:

For the logarithm-taking box, $\{D_{SAC}[j]|P_d = e_i\}$ 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, $\mathbf{e}_1$, ..., $\mathbf{e}_8$. It is seen from 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 $D^{\max}_{SAC}[j]$ curve by (2), which takes the maximum of normalized distance to *SAC* values for each avalanche variable. However, the $D^{\max}_{SAC}[j]$ curve is not depicted since (1) gives more valuable information than (2) does.

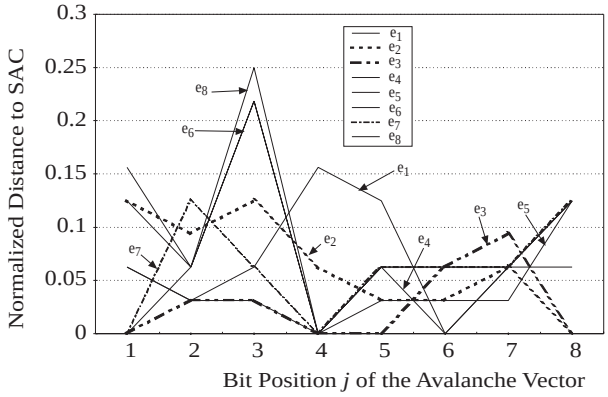

Figure 4. $\{D_{SAC}[j]|P_d = e_i\}$ versus j Curves for the Logarithm-Taking S-box

The normalized distance to *MOSAC* values for all possible 255 input differences are calculated by (3). Because of the difficulty in drawing all these curves, only the $D^{\max}_{MOSAC}[j]$ curve given by (4) is depicted in Figure 5. In this figure, the maxima, which are about 0.5, all occur for an input difference of 253 (=11111101_2) at all avalanche variables; hence, we obtain the same curve if $\{D_{MOSAC}[j]|P_d = 253_{10}\}$ is depicted instead.

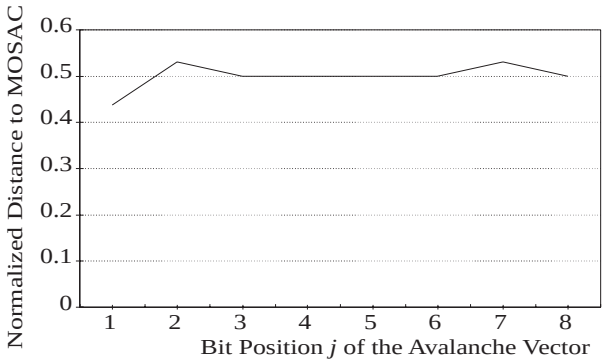

Figure 5. $D^{\max}_{MOSAC}[j]$ versus j Curve for the Logarithm-Taking S-box

2) BIC and MOBIC

$B^{\max}_{BIC}$ and $B^{\max}_{MOBIC}$ matrices are calculated by (6) and (8) respectively as follows:

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

$$B^{\max}_{MOBIC} = \begin{bmatrix} +1.00 & -0.28 & -0.44 & +0.35 & -0.48 & +0.34 & +0.33 & -0.40 \\ -0.28 & +1.00 & -0.34 & +0.36 & -0.28 & +0.40 & +0.34 & +0.31 \\ -0.44 & -0.34 & +1.00 & +0.33 & -0.50 & +0.37 & +0.27 & +0.40 \\ +0.35 & +0.36 & +0.33 & +1.00 & +0.33 & +0.25 & +0.31 & -0.31 \\ -0.48 & -0.28 & -0.50 & +0.33 & +1.00 & +0.29 & -0.30 & -0.37 \\ +0.34 & +0.40 & +0.37 & +0.25 & +0.29 & +1.00 & +0.27 & +0.37 \\ +0.33 & +0.34 & +0.27 & +0.31 & -0.30 & +0.27 & +1.00 & -0.53 \\ -0.40 & +0.31 & +0.40 & -0.31 & -0.37 & +0.37 & -0.53 & +1.00 \end{bmatrix}$$

3) X-OR Table Distribution

The *X-OR* table is a matrix of size 256 x 256, whose entries are calculated by (9). If it is divided into 8 pieces, so that each piece is 32 x 256, the maximum entry for each piece is as follows:

1^{st} piece: max. entry = 12 for (i, j) = (13, 64)

2^{nd} piece: max. entry = 22 for (i, j) = (60, 64)

3^{rd} piece: max. entry = 16 for (i, j) = (68, 53)

4^{th} piece: max. entry = 22 for (i, j) = (120, 192)

5^{th} piece: max. entry = 12 for (i, j) = (133, 109)

6^{th} piece: max. entry = 16 for (i, j) = (185, 181)

7^{th} piece: max. entry = 16 for (i, j) = (193, 192)

8^{th} piece: max. entry = 128 for (i, j) = (253, 128)

The maximum entry is 128 for the whole *X-OR* table and occurs for the position [253, 128], which means that when $\mathbf{P}_d = 253_{10}$, the avalanche vector $\mathbf{C}_d = 128_{10}$ occurs for 50% of the overall input pairs since the highest possible value is 2^8=256. The *X-OR* table distribution test also verifies the *SAC* test in that the maximum table entry occurs for an input difference of 253, where *SAC* test has its maxima.

3.3. Comparison of Exponentiating and Logarithm-Taking Boxes

The exponentiating s-box has a weakness for an 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 $D^{\max}_{MOSAC}[j]$ curves, given by Figure 3 and Figure 5, are sketched together in Figure 6.

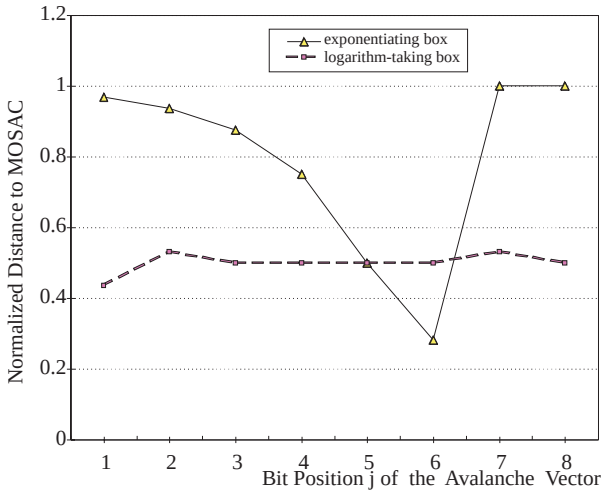

Figure 6. $D^{\max}_{MOSAC}[j]$ Curves for the Exponentiating and Logarithm-Taking S-boxes

As seen from the solid curve in Figure 6, none of the avalanche variables obey the *SAC*; moreover, it is observed by comparing Figure 6 with Figure 2 that

$D^{\max}_{MOSAC}[j] = \{D_{SAC}[j] | P_d = 128_{10}\}$ for all j value, except $j = 6$. For an input difference of 128, many of the avalanche variables have large distances to *SAC*, and $D_{SAC}[\ _j]$ becomes equal to 1 for $j = 7$ and $j = 8$; since the outputs of the exponentiating box always have the same bit values in their 7^{th}, and the complement bit values in their 8^{th} bit positions.

For the same reason, the last two rows of $B^{\max}_{BIC}$ and $B^{\max}_{MOBIC}$ matrices are undefined. Other values in $B^{\max}_{MOBIC}$ 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 $B_{BIC}(j, k | P_d = 128_{10})$ given by (5). The *X-OR* table distribution test also verifies *SAC* and *BIC* tests in that the maximum table entry occurs for an input difference of 128.

The logarithm-taking s-box has a weakness for an input difference of 253 (=11111101_2). The dashed curve in Figure 6 corresponds to $D^{\max}_{MOSAC}[j]$ and is equal to $\{D_{MOSAC}[j] | P_d = 253_{10}\}$ for all j. However, normalized distance to *MOSAC* for all j values is about 0.5, which is better than the case of the exponentiating s-box. Many elements of $B^{\max}_{MOBIC}$ are the same as the elements of $B_{BIC}(j, k | P_d = 253_{10})$, and the maximum entry of the $B^{\max}_{MOBIC}$ is –0.53. The maximum entry of the *X-OR* table also occurs for an input difference of 253, where the *SAC* test has its maxima for all avalanche variables.

Therefore, the logarithm-taking s-box seems to be have a weakness for an input difference of 253 and the exponentiating s-box has a weakness for an input difference of 128.

Finally, we should mention the fact that since the results presented in this section are obtained by isolating the substitution boxes from the general structure, they donot indicate an overall weakness in the SAFER algorithm. The overall performance of a specific algorithm from the SAFER family is investigated by two different criteria in the following two sections.

4. AWD Curves of Safer K-64

We use the following test procedure [4] to examine the overall round by round diffusion properties of SAFER K-64 by the criterion of avalanche weight distribution (*AWD*):

- A plaintext $\mathbf{P}$ is chosen at random and the plaintext $\mathbf{P}_i$ is calculated so that the difference between $\mathbf{P}$ and $\mathbf{P}_i$ is $\mathbf{e}_i$, i.e., $\mathbf{P}_i = \mathbf{P} \oplus \mathbf{e}_i$ and $\mathbf{P}$ and $\mathbf{P}_i$ differ only in bit i, where $\mathbf{e}_i$ is a 64-bit unit vector with a 1 in position i, and $i \in \{1, 2, \ldots, 64\}$,
- $\mathbf{P}$ and $\mathbf{P}_i$ are submitted to r-rounds of SAFER K-64 for encryption under a random key,
- From the resultant ciphertexts $\mathbf{C}$ and $\mathbf{C}_i$, the Hamming weight of the avalanche vector $wt(\mathbf{C}_d) = wt(\mathbf{C} \oplus \mathbf{C}_i) = j$ is calculated, where $j \in \{0, 1, 2, \ldots, 64\}$,
- The value of the j^{th} element of an avalanche weight distribution array of size 65 is incremented by 1, i.e., $AWD\_\text{array}\,[j] = AWD\_\text{array}\,[j] + 1$,
- The steps above are repeated 10000 times and the values in the avalanche weight distribution array are sketched versus its index.

With the help of the test procedure explained above, 64 figures, each corresponding to a different input difference $\mathbf{P}_d = \mathbf{e}_i$ where i=1,..., 64, are obtained for 1-round encryption of SAFER K-64. Some of these curves seem to obey the binomial distribution around a mean value of 32 and this is what we expect due to the *AWD* criterion. However, some curves seem to be far from being binomially distributed. Such histograms show that after a single round of encryption the avalanche vectors may have very small Hamming weights such as 1, 2, 4, 5 and 6. We concentrate on the worst diffusion curves of the first round and the second round. The resultant curves after 2 rounds of encryption are depicted in Figures 7, 8 and 9 together with the histograms obtained after a single round of encryption.

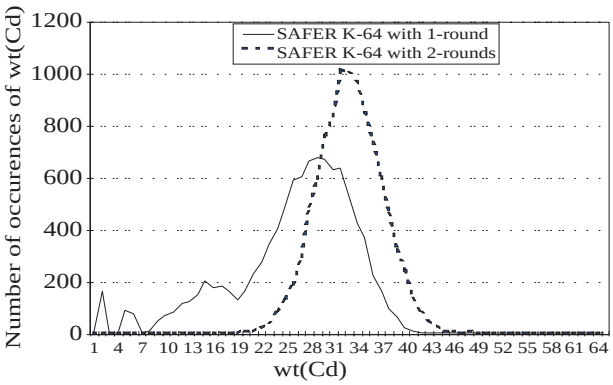

Figure 7. *AWD* Curves of SAFER K-64 for $\mathbf{e}_{57}$

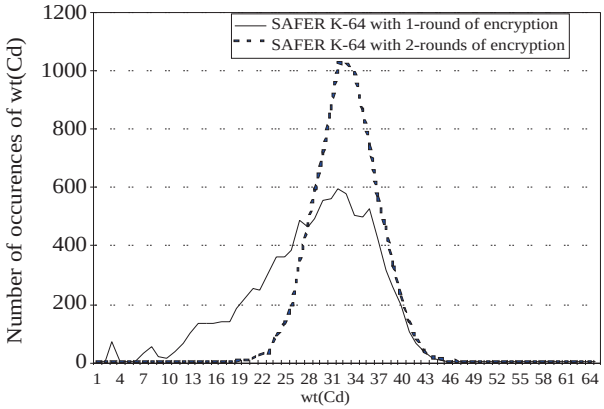

Figure 8. *AWD* Curves of SAFER K-64 for $\mathbf{e}_{41}$

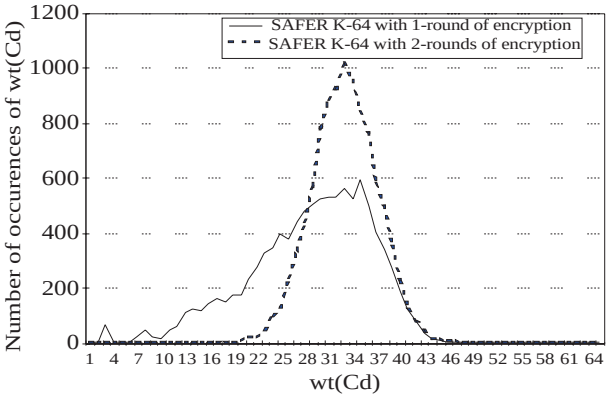

Figure 9. *AWD* Curves of SAFER K-64 for $\mathbf{e}_{49}$

It is observed from Figures 7, 8, and 9 that when the number of encryption rounds is increased to 2, the desired diffusion is achieved even for the worst *AWD* curves of the first round. Moreover, for more than 2 rounds of encryption we did not observe any distortion in the binomial shape of the *AWD* curves .

Hence, we conclude that the SAFER K-64 achieves the desired diffusion after at most the second of its six rounds.

5. Avalanche Curves of Safer K-64

We now use the following test procedure to examine the overall round by round diffusion properties of SAFER K-64, by the conventional criterion of strict avalanche;

- A plaintext $\mathbf{P}$ is chosen at random and the plaintext $\mathbf{P}_i$ is calculated so that the difference between $\mathbf{P}$ and $\mathbf{P}_i$ is $\mathbf{e}_i$, i.e., $\mathbf{P}_i = \mathbf{P} \oplus \mathbf{e}_i$ and $\mathbf{P}$ and $\mathbf{P}_i$ differ only in bit i, where $\mathbf{e}_i$ is a 64-bit unit vector with a 1 in position i, and $i \in \{1, 2, \ldots, 64\}$,
- $\mathbf{P}$ and $\mathbf{P}_i$ are submitted to r-rounds of SAFER K-64 for encryption under a random key,
- From the resultant ciphertexts $\mathbf{C}$ and $\mathbf{C}_i$, the avalanche vector $\mathbf{C}_d = (\mathbf{C} \oplus \mathbf{C}_i)$ is calculated,
- The 64 bit avalanche vector is summed up to an avalanche sum array, i.e.,
  avalanche_sum_array = avalanche_sum_array + $\mathbf{C}_d$,

- The steps above are repeated 10000 times and the values in the avalanche sum array are sketched versus its index.

With the help of the test procedure explained above, 64 avalanche curves, each corresponding to an input difference e_i where $i \in \{1,2,..., 64\}$, are obtained for r-round encryption of SAFER K-64. In Figures 10 and 11 we give the worst case examples, each showing 3 curves corresponding to 1-round, 2-round and 6-round encryptions of SAFER K-64.

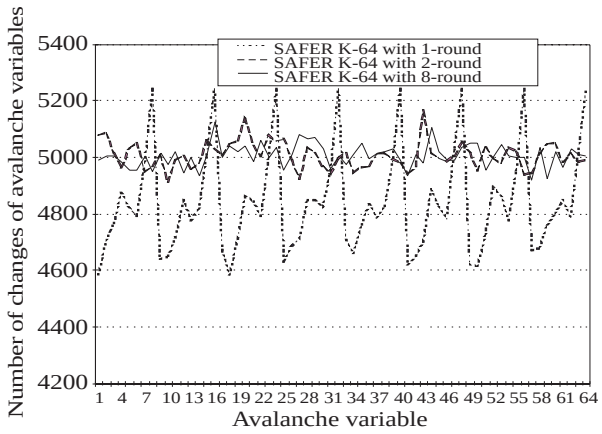

Figure 10. Avalanche Curves of SAFER K-64 for $\mathbf{e}_2$

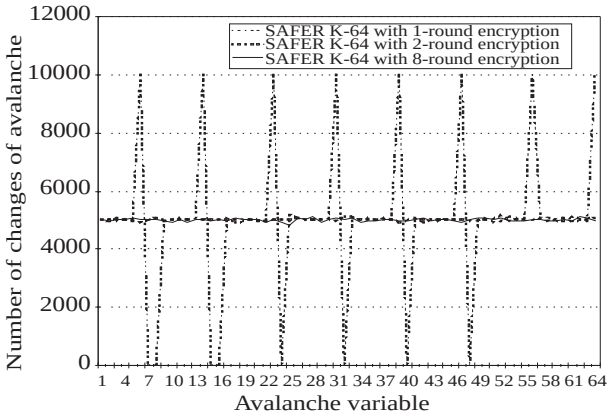

Figure 11. Avalanche Curves of SAFER K-64 for $\mathbf{e}_{32}$

It is expected, due to the strict avalanche criterion, that an average of one half of the output bits should change whenever a single input bit is complemented, i.e., avalanche_sum_array [i] $\approx$ 5000 for all i values since the number of trials is 10000.

However, it is observed from Figures 10 and 11 that the number of changes is very different from 5000 for most of the avalanche variables and the strict avalanche criterion completely fails at some avalanche variables for 1-round encryption of SAFER K-64. It is also observed that when the number of encryptions is increased to 2, the desired diffusion is achieved, even for the worst avalanche variables. Moreover, for 6-rounds of encryption, the resultant avalanche curves are very similar to the curves obtained for 2-rounds of encryption.

Hence, we conclude that the SAFER K-64 achieves the desired diffusion after at most the second of its six rounds and note that this result is also verified by the tests applied for the criterion of *AWD* in Section IV.

6. Conclusions

The characteristics of s-boxes of the SAFER family of ciphers in terms of *SAC*, *MOSAC*, *BIC*, *MOBIC* and *X-OR* distributions, and the overall diffusion properties of SAFER K-64 in terms of *AWD* and *SAC* are examined. Experiments show that although the logarithm-taking s-box seems to be more resistive to differential attacks than the exponentiating s-box, it has a weakness for an input difference of 253 ($=11111101_2$) and the exponentiating s-box has a weakness for an input difference of 128 ($=10000000_2$).

However, as far as the two different criteria (the rough test method of *AWD* and a more sensitive measure of strict avalanche curves) indicate, the overall performance of SAFER K-64 does not suffer from the weaknesses of its s-boxes and achieves the desired diffusion after at most the second step of its six rounds. It is necessary to direct further work on this subject towards differential cryptanalysis of SAFER K-64 by making use of the sensitive inputs of its exponentiating and logarithm-taking boxes.

References

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

[2] E. Aras and M. D. Yücel, "Examination of Substitution Boxes of SAFER", *Proc. ELECO'99 Int. Conf. on Electrical & Electronics Engineering*, Bursa, Türkiye, pp.418-422, December 1999.

[3] E. Aras and M. D. Yücel, "Some Cryptographic Properties of Exponentiating and Logarithm-Taking Boxes", *Proc. 20th Biennial Symp. on Communications,* Queen's Univ., Kingston, Ontario, Canada, pp. 69-73, May 2000.

[4] E. Aras, "Analysis of Security Criteria for Block Ciphers", *M.Sc. Thesis,* Electrical & Electronics Eng. Dept. of Middle East Technical University, Ankara, Türkiye, September 1999.

[5] R. C. Acar, "Cryptographic Test Methods for Block Ciphers", *M.Sc. Thesis,* Electrical & Electronics Eng. Dept. of Middle East Technical University, Ankara, Türkiye, December 1999.

[6] M. D. Yücel and R. C. Acar, "Comparison of the Block Ciphers DES and IDEA", *Proc. ELECO'99 Int. Conf. on Electrical & Electronics Engineering*, Bursa, Türkiye, pp. 281-285, December 1999.

[7] C. E. Shannon, "Communication Theory of Secrecy Systems", *Bell Systems Technical Journal,* Vol. 28, pp. 656-715, 1949.

[8] X. Lai, "On the Design and Security of Block Ciphers", *ETH Series in Information Processing,* Vol. 1, Hartung-Gorre Verlag Konstanz, 1992.

[9] 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.

[10] H. Feistel, "Cryptography and Computer Privacy", *Scientific American*, Vol. 228, No. 5, pp. 15-23, May 1973.

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

[12] S. Mister and C. M. Adams, "Practical S-Box Design", *SAC'96 – Third Annual Workshop on Selected Areas in Cryptography*, Queen's Univ., Kingston, Ontario, Canada, pp. 61-76, August 1996.

[13] E. Biham and A. Shamir, "Differential Cryptanalysis of DES-like Cryptosystems", *Journal of Cryptology,* Vol. 4, No. 1, pp. 3-72, 1991.