J. Korea Soc. Math. Educ. Ser. B: Pure Appl. Math.
Volume 8, Number 2 (November 2001), Pages 163–174

ISSN 1226-0657

# GLOBAL AVALANCHE CRITERION FOR THE $S$-BOXES OF DES

WANSOON KIM, YANG-SU KIM, AND MIN SURP RHEE

ABSTRACT. In this paper we modify two indicators related to the global avalanche criterion (GAC) and discuss their properties. Also, we apply the modified indicators to measure the GAC of $S$-boxes of DES.

## 1. INTRODUCTION

A symmetric key block cipher system requires strong substitution boxes ($S$-boxes) satisfying a number of critical properties. Among these properties are balancedness, nonlinearity, propagation characteristics and correlation immunity. The propagation characteristics can be investigated by the SAC (Strict Avalanche Criterion) and PC (Propagation Criterion), in a local manner. The SAC was introduced by Webster [7] and Webster-Tavares [8] in 1985 and measures the balancedness of output bits when a single bit of input bits is complemented. Independently the propagation criterion was introduced by Preneel et al [3] and measures the balancedness of output bits when input bits are complemented with respect to a vector $\alpha \in GF(2)^n$, where $GF(2)$ is the Galois field. But all these concepts have their limitations in capturing some of the vital characteristics required by a cryptographically strong function.

A new criterion called GAC (Global Avalanche Characteristic) to measure the propagation characteristic of substitution boxes of cryptographic system in a global manner was introduced by Zhang-Zheng [10] in 1997. Two indicators the sum-of-square indicators $\sigma_f$ and the absolute indicator $\Delta_f$ for the $S$-boxes were introduced. The smaller $\sigma_f$ and $\Delta_f$ are, the better the GAC of a $S$-box is.

Received by the editors September 7, 2001.
2000 *Mathematics Subject Classification.* 94A60.
*Key words and phrases.* A Boolean function, bent, nonlinearity, strict avalanche criterion, propagation criterion, global avalanche criterion.

In this paper we modify the concepts of two indicators so that we are able to compare the GACs of different $S$-boxes. Applying the modified indicators to $S$-boxes of DES(Data Encryption Standard) we find the $S_4$ box is strong and the $S_7$ box is weak as far as GAC concerns.

## 2. Basic notations and definitions

In this section, we introduce notations, definitions and well known properties for cryptographic Boolean functions.

Let $GF(2)^n$ be an $n$-dimensional vector space over the Galois field $GF(2)$. A function $f$ from $GF(2)^n$ to $GF(2)$ is called a Boolean function on $GF(2)^n$. A Boolean function of the form

$$f(x) = a_1x_1 + \cdots + a_nx_n + c, \text{ where } a_j,\ c \in GF(2),\ \ i = 1, 2, \cdots, n$$

is called affine. Furthermore $f$ is called linear if $c = 0$.

The Hamming weight $W(x)$ of $x \in GF(2)^n$ is the number of ones in $x$. The Hamming distance $d(f, g)$ between two functions $f$ and $g$ is the number of function values in which they differ. A Boolean function $f$ is *balanced* if $d(f, o) = \frac{1}{2^{n-1}}$, where $o$ is the Boolean function on $GF(2)^n$ all of whose function values are 0. This implies that when input coordinates of a Boolean function are selected independently, at random, the output of the function must behave as a uniformly distributed random variable.

The minimum Hamming distance between $f$ and all the affine functions on $GF(2)^n$ is called the *nonlinearity* of a Boolean function $f$ and it is denoted by $N_f$. Then the nonlinearity of an affine function is 0.

Strong substitution boxes also require that complementing a single bit results in the output of the function being complemented with a probability of a half. The following definition was introduced by Webster [7] and Webster-Tavares [8].

**Definition 2.1.** A Boolean function $f$ on $GF(2)^n$ is said to *satisfy the strict avalanche criterion* (SAC) if

$$|\{x \in GF(2)^n |\ f(x + \alpha) = f(x)\}| = 2^{n-1},$$

for any $\alpha \in GF(2)^n$ with $W(\alpha) = 1$.

It is obvious that $f$ on $GF(2)^n$ satisfies the SAC if and only if $f(x) + f(x + \alpha)$ is balanced for any $\alpha \in GF(2)^n$ with $W(\alpha) = 1$.

**Example 2.2.** Let $f$ be a Boolean function on $GF(2)^3$ defined by $f(x_1, x_2, x_3) = x_1x_2 + x_2x_3$. Then for all $\alpha$'s with $W(\alpha) = 1$ we have

$$\left|\{x \in GF(2)^3 |\, f(x+\alpha) = f(x)\}\right| = 4$$

Thus $f$ satisfies the strict avalanche criterion.

The notion of SAC has been generalized to the propagation criterion by Adams and Tavares [1] and independently by Preneel et al [4].

**Definition 2.3.** A Boolean function $f$ is said to *satisfy the propagation criterion* (PC) *with respect to a vector* $\alpha \in GF(2)^n$ if $|\{x \in GF(2)^n |\, f(x+\alpha) = f(x)\}| = 2^{n-1}$, or equivalently $f(x+\alpha) + f(x)$ is balanced. Furthermore, a Boolean function $f$ is said to *satisfy the propagation criterion of degree* $k$ if it satisfies the propagation criterion with respect to all nonzero vectors whose Hamming weight is at most $k$.

**Example 2.4.** Let $f$ be a Boolean function on $GF(2)^4$ defined by $f(x_1, x_2, x_3) = x_1x_2 + x_2x_3 + x_1x_3$. Then $f$ satisfies the PC of degree 2.

Note that a function satisfying the PC of degree 1 satisfies the SAC. Generally speaking, $f$ satisfies the PC of degree $k$ if complementing $k$ or less bits results in the out bit of $f$ being complemented by a probability of half. As we mentioned above, the concept of PC measure the avalanche characteristic of a function. A complementary concept to the PC is the linear structure which measures the smoothness of a function.

**Definition 2.5.** A Boolean function $f$ on $GF(2)^n$ is said to *have a linear structure* if

$$f(x) + f(x+\alpha) = c, \quad \text{where } x \in GF(2)^n \text{ and } c \in GF(2),$$

for some nonzero vector $\alpha \in GF(2)^n$.

From above definition every affine function has a linear structure. The following example is not an affine function, but it has a linear structure.

**Example 2.6.** Let $f$ be a Boolean function on $x \in GF(2)^4$ defined by

$$f(x_1, x_2, x_3, x_4) = x_1 + x_1x_2 + x_1x_3 + x_4.$$

Then for $\alpha = (0, 1, 1, 0)$ we have

$$f(x) + f(x+\alpha) = f(x_1, x_2, x_3, x_4) + f(x_1, x_2+1, x_3+1, x_4) = 0. \ .$$

Thus $f$ has a linear structure with respect to $\alpha = (0, 1, 1, 0)$.

Let $f$ be a Boolean function, and let

$$\alpha_0 = (0, \cdots, 0, 0),\ \alpha_1 = (0, \cdots, 0, 1),\ \cdots,\ \alpha_{2^n-1} = (1, \cdots, 1, 1)$$

be all vectors of $GF(2)^n$. Then the sequence $(f(\alpha_0), f(\alpha_1), \cdots, f(\alpha_{2^n-1}))$ is a $(0,1)$-sequence of $f$ and the sequence $((-1)^{f(\alpha_0)}, (-1)^{f(\alpha_1)}, \cdots, (-1)^{f(\alpha_{2^n-1})})$ is a $(1,-1)$-sequence of $f$.

Let $\bar{a} = (a_1, \cdots, a_m)$ and be $\bar{b} = (b_1, \cdots, b_m)$ two vectors (or sequences), the scalar product of $\bar{a}$ and $\bar{b}$, denoted by $\langle \bar{a}, \bar{b} \rangle$, is defined as the sum of the component-wise multiplications. In particular, when $\bar{a}$ and $\bar{b}$ are $(0,1)$-sequences, $\langle \bar{a}, \bar{b} \rangle = a_1b_1 + \cdots + a_mb_m$, where the addition and multiplications are over $GF(2)$, and when $\bar{a}$ and $\bar{b}$ are $(1,-1)$-sequences, $\langle \bar{a}, \bar{b} \rangle = \sum_{i=1}^{m} a_ib_j$, where the addition and multiplication are over the reals.

**Definition 2.7.** Let $f$ be a Boolean function on $GF(2)^n$. The *Walsh-Hadamard transform* of $f$ is defined as

$$\hat{f}(\alpha) = 2^{-\frac{n}{2}} \sum_{x \in GF(2)^n} (-1)^{f(x) + \langle \alpha, x \rangle}$$

where $\alpha = (a_1, \cdots, a_n) \in GF(2)^n$, $x = (x_1, \cdots, x_n)$ and $f(x) + \langle \alpha, x \rangle$ is regarded as a real-valued function.

**Definition 2.8.** A Boolean function $f$ on $GF(2)^n$ is called a *bent function* if its Walsh-Hadamard transform satisfies

$$\hat{f}(\alpha) = \pm 1 \quad \text{for all } \alpha \in GF(2)^n.$$

The following result can be found in Adams-Tavares [1], Dillon [2], Seberry-Zhang [5] and Yarlagadda-Hershey [9].

**Proposition 2.9.** *Let $f : GF(2)^n \to GF(2)$ be a Boolean function. Then the following statements are equivalent:*

(1) *$f$ is bent.*

(2) *$\langle \xi, l \rangle = \pm 2^{\frac{1}{2}n}$ for any affine sequence $l$ of length $2^n$, where $\xi$ is the sequence of $f$.*

(3) *$f(x) + f(x + \alpha)$ is balanced for any non-zero $\alpha \in GF(2)^n$.*

The following properties of bent functions are well known (cf. Zhang-Zheng [10]).

**Proposition 2.10.** *Let $f$ be a bent function. Then*

(1) *$f$ satisfies PC of order of degree $k$ for all $1 \le k \le n$,*

(2) *$f$ satisfies SAC,*
(3) *$f$ has maximum nonlinearity,*
(4) *$f$ has no linear structure, and*
(5) *$f$ is not balanced.*

Even though bent functions satisfy strong properties such as SAC, PC and have maximum nonlinearity, they are not balanced. Also we observe that functions satisfying SAC can have a large number of vectors of Hamming weight larger than one as its linear structure. Therefore they can not be directly used in a cipher system where balancedness is needed.

## 3. The Criterion for Global Avalanche Characteristics

In this section we modify two indicators the sum-of-square indicator and the absolute indicator to measure the global avalanche characterstics (GACs) of Boolean functions.

Given a Boolean function $f$ on $GF(2)^n$ and a vector $\alpha$ in $GF(2)^n$, we denoted by $\xi(\alpha)$ the sequence of $f(x+\alpha)$. Set $\Delta_f(\alpha) = \langle \xi(0), \xi(\alpha) \rangle$. To further simplify our discussion, $\Delta_f(\alpha)$ will be written as $\delta(\alpha)$ if the function under consideration is clear.

The overall avalanche characteristic of a function $f$ can be measured by examining $|\Delta(\alpha)|$ for all nonzero vectors $\alpha$. We can say that a function has a good GAC if for most nonzero $\alpha$, $|\Delta(\alpha)|$ is zero or very close to zero. This observation leads us to the following definition (cf. Zhang-Zheng [10]).

**Definition 3.1.** Let $f : GF(2)^n \to GF(2)$ be a Boolean function. Then the *sum-of-square indicator for the avalanche characteristic* of $f$ is defined by

$$\sigma_f = \sum_{\alpha \in GF(2)^n} \Delta^2(\alpha)$$

and the *absolute indicator for the characteristic* of $f$ is defined by

$$\Delta_f = \max_{\alpha \in GF(2)^n, \alpha \neq 0} |\Delta(\alpha)|.$$

The smaller $\sigma_f$ and $\Delta_f$ are, the better the GAC of a function $f$ is. The following theorem gives us the upper bounds and lower bounds on $\sigma_f$.

**Theorm 3.2.** *Let $f$ be a Boolean function on $GF(2)^n$. Then*

(1) $2^{2n} \leq \sigma_f \leq 2^{3n}$,
(2) $\sigma_f = 2^{2n}$ *if and only if f is a bent function, and*
(3) $\sigma_f = 2^{3n}$ *if and only if f is an affine function.*

*Proof.* The proof can be found in Zhang-Zheng [10]. □

Now we modify two indicators in order to compare the GACs of two different functions.

**Definition 3.3.** Let $f$ be a Boolean function on $GF(2)^n$. Then the value

$$R_{\sigma_f} = \frac{\sigma_f - 2^{2n}}{2^{3n} - 2^{2n}} \quad \text{and} \quad R_{\Delta_f} = \frac{\Delta_f}{2^n} \quad \text{(resp.)}$$

are called the *sum-of-square measure* and the *absolute measure* (resp.) of $f$.

**Proposition 3.4.** *Let f be a Boolean function on* $GF(2)^n$. *Then*

(1) $0 \leq R_{\sigma_f} \leq 1$,
(2) $f$ *is bent if and only if* $R_{\sigma_f} = 0$, *and*
(3) $f$ *is affine if and only if* $R_{\sigma_f} = 1$.

*Proof.* (1) Since $2^{2n} \leq \sigma_f \leq 2^{3n}$, it is clear by Definition 3.3.
(2) By Theorem 3.2, $f$ is bent if and only if $\sigma_f = 2^{2n}$ if and only if $R_{\sigma_f} = 0$ by Definition 3.3.
(3) By Theorem 3.2, $f$ is an affine function if and only if $\sigma_f = 2^{3n}$ if and only if $R_{\sigma_f} = 1$ by Definition 3.3. □

**Proposition 3.5.** *Let f be a Boolean function on* $GF(2)^n$. *Then*

(1) $0 \leq R_{\Delta_f} \leq 1$,
(2) $f$ *is bent if and only if* $R_{\Delta_f} = 0$, *and*
(3) $f$ *a linear structure if and only if* $R_{\Delta_f} = 1$.

*Proof.* (1) Note that $\Delta_f$ is defined as the maximum among all $\Delta(\alpha)$, $\alpha \neq 0$. Therefore $\Delta(\alpha) = \pm 2^n$ if and only if $\alpha$ is a linear structure of $f$. Hence we get $0 \leq \Delta_f \leq 2^n$.
(2) It is clear that $\Delta_f = 0$ if and only if $f$ is bent if and only if $R_{\Delta_f} = 0$.
(3) $\Delta_f = 2^n$ if and only if $f$ has a linear structure. □

**Proposition 3.6.** *Let f be a non-bent cubic function on* $GF(2)^n$. *Then*

$$R_{\Delta_f} \geq 2^{-\frac{n-1}{2}}.$$

*Proof.* It follows from Seberry-Zhang [5, Theorem 3] that $\Delta_f \geq 2^{\frac{n+1}{2}}$. Hence

$$R_{\Delta_f} = \frac{\Delta_f}{2^n} \geq 2^{-\frac{n-1}{2}}. \qquad \square$$

**Example 3.7.** Let $w$ be a permutation on the set $V_k - \{0\} = \{a_1, \cdots, a_{2^k-1}\}$. Let $f$ be a Boolean function on $GF(2)^{2k}$ defined by

$$f(z) = f(y,x) = \begin{cases} \langle \alpha_{j_0}, x \rangle & \text{if } y = 0 \\ \langle W(y), x \rangle & \text{if } y \neq 0 \end{cases}$$

where $z = (y, x)$, $y \in GF(2)^k$, $x \in GF(2)^k$, $y = (y_1, \cdots, y_k)$ and $\alpha_{j_0}$ is a fixed nonzero vector in $GF(2)^k$. Then we have the following Table 3.1.

TABLE 3.1. The sum-of-square measure and the absolute measure of $f$

| $k$ | $\alpha_{j_0}$ | $\sigma_f$ | $\Delta$ | $R_{\sigma_f}$ | $R_{\Delta_f}$ |
|---|---|---|---|---|---|
| 3 | (1,0,1) | 7168 | 16 | 0.012 | 0.25 |
| 4 | (1,0,0,1) | 90112 | 32 | 0.001 | 0.125 |
| 5 | (1,0,0,0,1) | 1245184 | 64 | 0.0002 | 0.0625 |
| 6 | (1,0,0,0,0,1) | 18350080 | 128 | 0.00002 | 0.03125 |

We observe that as $k$ increase, both $R_{\sigma_f}$ and $R_{\Delta_f}$ decrease very rapidly.

**Example 3.8.** Let $f$ be a function on $GF(2)^{2k}$ defined by Example 3.7. Then

$$R_{\sigma_f} = \frac{2^{3k+3} - 2^{3k+1}}{2^{6k} - 2^{4k}} \quad \text{and} \quad R_{\Delta_f} \leq 2^{-k+1}.$$

*Remark.* It follows from Seberry-Zhang [5, Theorem 4] that $\sigma_f = 2^{4k} + 2^{3k+3} - 2^{3k+1}$ and $\Delta_f \leq 2^{k+1}$. Hence we have

$$R_{\sigma_f} = \frac{\sigma_f - 2^{4k}}{2^{6k} - 2^{4k}} = \frac{2^{3k+3} - 2^{3k+1}}{2^{6k} - 2^{4k}} \quad \text{and} \quad R_{\Delta_f} = \frac{\Delta_f}{2^{2k}} \leq 2^{-k+1}.$$

## 4. The GAC for the $S$-boxes of DES

In this section we present the values of the sum-of-square measure and the absolute measure for the $S$-boxes of DES (cf. Stinson [6]).

The DES(Data Encryption Standard) is the most well-known symmetric key block cipher, which was developed at IBM. DES was first published in the Federal

TABLE 4.1. The $S$-boxes of DES

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $S_1$ | 14 | 4 | 13 | 1 | 2 | 15 | 11 | 8 | 3 | 10 | 6 | 12 | 5 | 9 | 0 | 7 |
| | 0 | 15 | 7 | 4 | 14 | 2 | 13 | 1 | 10 | 6 | 12 | 11 | 9 | 5 | 3 | 8 |
| | 4 | 1 | 14 | 8 | 13 | 6 | 2 | 11 | 15 | 12 | 9 | 7 | 3 | 10 | 5 | 0 |
| | 15 | 12 | 8 | 2 | 4 | 9 | 1 | 7 | 5 | 11 | 3 | 14 | 10 | 0 | 6 | 13 |
| $S_2$ | 15 | 1 | 8 | 14 | 6 | 11 | 3 | 4 | 9 | 7 | 2 | 13 | 12 | 0 | 5 | 10 |
| | 3 | 13 | 4 | 7 | 15 | 2 | 8 | 14 | 12 | 0 | 1 | 10 | 6 | 9 | 11 | 5 |
| | 0 | 14 | 7 | 11 | 10 | 4 | 13 | 1 | 5 | 8 | 12 | 6 | 9 | 3 | 2 | 15 |
| | 13 | 8 | 10 | 1 | 3 | 15 | 4 | 2 | 11 | 6 | 7 | 12 | 0 | 5 | 14 | 9 |
| $S_3$ | 10 | 0 | 9 | 14 | 6 | 3 | 15 | 5 | 1 | 13 | 12 | 7 | 11 | 4 | 2 | 8 |
| | 13 | 7 | 0 | 9 | 3 | 4 | 6 | 10 | 2 | 8 | 5 | 14 | 12 | 11 | 15 | 1 |
| | 13 | 6 | 4 | 9 | 8 | 15 | 3 | 0 | 11 | 1 | 2 | 12 | 5 | 10 | 14 | 7 |
| | 1 | 10 | 13 | 0 | 6 | 9 | 8 | 7 | 4 | 15 | 14 | 3 | 11 | 5 | 2 | 12 |
| $S_4$ | 7 | 13 | 14 | 3 | 0 | 6 | 9 | 10 | 1 | 2 | 8 | 5 | 11 | 12 | 4 | 15 |
| | 13 | 8 | 11 | 5 | 6 | 15 | 0 | 3 | 4 | 7 | 2 | 12 | 1 | 10 | 14 | 9 |
| | 10 | 6 | 9 | 0 | 12 | 11 | 7 | 13 | 15 | 1 | 3 | 14 | 5 | 2 | 8 | 4 |
| | 3 | 15 | 0 | 6 | 10 | 1 | 13 | 8 | 9 | 4 | 5 | 11 | 12 | 7 | 2 | 14 |
| $S_5$ | 2 | 12 | 4 | 1 | 7 | 10 | 11 | 6 | 8 | 5 | 3 | 15 | 13 | 0 | 14 | 9 |
| | 14 | 11 | 2 | 12 | 4 | 7 | 13 | 1 | 5 | 0 | 15 | 10 | 3 | 9 | 8 | 6 |
| | 4 | 2 | 1 | 11 | 10 | 13 | 7 | 8 | 15 | 9 | 12 | 5 | 6 | 3 | 0 | 14 |
| | 11 | 8 | 12 | 7 | 1 | 14 | 2 | 13 | 6 | 15 | 0 | 9 | 10 | 4 | 5 | 3 |
| $S_6$ | 12 | 1 | 10 | 15 | 9 | 2 | 6 | 8 | 0 | 13 | 3 | 4 | 14 | 7 | 5 | 11 |
| | 10 | 15 | 4 | 2 | 7 | 12 | 9 | 5 | 6 | 1 | 13 | 14 | 0 | 11 | 3 | 8 |
| | 9 | 14 | 15 | 5 | 2 | 8 | 12 | 3 | 7 | 0 | 4 | 10 | 1 | 13 | 11 | 6 |
| | 4 | 4 | 2 | 12 | 9 | 5 | 15 | 10 | 11 | 14 | 1 | 7 | 6 | 0 | 8 | 13 |
| $S_7$ | 4 | 11 | 2 | 14 | 15 | 0 | 8 | 13 | 3 | 12 | 9 | 7 | 5 | 10 | 6 | 1 |
| | 13 | 0 | 11 | 7 | 4 | 9 | 1 | 10 | 14 | 3 | 5 | 12 | 2 | 15 | 8 | 6 |
| | 1 | 4 | 11 | 13 | 12 | 3 | 7 | 14 | 10 | 15 | 6 | 8 | 0 | 5 | 9 | 2 |
| | 6 | 11 | 13 | 8 | 1 | 4 | 10 | 7 | 9 | 5 | 0 | 15 | 14 | 2 | 3 | 12 |
| $S_8$ | 13 | 2 | 8 | 4 | 6 | 15 | 11 | 1 | 10 | 9 | 3 | 14 | 5 | 0 | 12 | 7 |
| | 1 | 15 | 13 | 8 | 10 | 3 | 7 | 4 | 12 | 5 | 6 | 11 | 0 | 14 | 9 | 2 |
| | 7 | 11 | 4 | 1 | 9 | 12 | 14 | 2 | 0 | 6 | 10 | 13 | 15 | 3 | 5 | 8 |
| | 2 | 1 | 14 | 7 | 4 | 10 | 8 | 13 | 15 | 12 | 9 | 0 | 3 | 5 | 6 | 11 |

Register of March 17, 1975. After adopting as a standard for unclassified applications on January 1977 it has been renewed until 1998. One of important factors in security is $S$-boxes. In the encryption process of DES there are 8 $S$-boxes, given in Table 4.1, which are functions from $GF(2)^6$ to $GF(2)^4$. We denote component functions of an $S$-boxes by $f_1(x), f_2(x), f_3(x)$ and $f_4(x)$.

TABLE 4.2. Measure values for component functions of DES

| | $S_1$-box | | | | $S_2$-box | | | |
|---|---|---|---|---|---|---|---|---|
| $i$ | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| $\sigma_{f_i}$ | 20224 | 11776 | 16000 | 19072 | 19840 | 25984 | 16000 | 14464 |
| $R_{\sigma,i}$ | 0.063 | 0.030 | 0.046 | 0.058 | 0.061 | 0.085 | 0.046 | 0.040 |
| $\Delta_{f_i}$ | 32 | 24 | 40 | 32 | 40 | 56 | 32 | 40 |
| $R_{\Delta,i}$ | 0.5 | 0.38 | 0.63 | 0.5 | 0.63 | 0.88 | 0.5 | 0.63 |
| | $S_3$-box | | | | $S_4$-box | | | |
| $i$ | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| $\sigma_{f_i}$ | 21376 | 19456 | 12544 | 19072 | 15232 | 15232 | 15232 | 15232 |
| $R_{\sigma,i}$ | 0.067 | 0.060 | 0.033 | 0.058 | 0.043 | 0.043 | 0.043 | 0.043 |
| $\Delta_{f_i}$ | 40 | 48 | 32 | 48 | 24 | 24 | 24 | 24 |
| $R_{\Delta,i}$ | 0.63 | 0.75 | 0.5 | 0.75 | 0.38 | 0.38 | 0.38 | 0.38 |
| | $S_5$-box | | | | $S_6$-box | | | |
| $i$ | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| $\sigma_{f_i}$ | 16000 | 18304 | 14464 | 10624 | 19456 | 15616 | 15232 | 16000 |
| $R_{\sigma,i}$ | 0.046 | 0.055 | 0.040 | 0.025 | 0.060 | 0.045 | 0.043 | 0.046 |
| $\Delta_{f_i}$ | 32 | 40 | 32 | 24 | 48 | 32 | 40 | 40 |
| $R_{\Delta,i}$ | 0.5 | 0.63 | 0.5 | 0.38 | 0.75 | 0.5 | 0.63 | 0.63 |
| | $S_7$-box | | | | $S_8$-box | | | |
| $i$ | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| $\sigma_{f_i}$ | 17152 | 34048 | 13696 | 23296 | 12544 | 16768 | 16768 | 12928 |
| $R_{\sigma,i}$ | 0.051 | 0.116 | 0.037 | 0.074 | 0.033 | 0.049 | 0.049 | 0.034 |
| $\Delta_{f_i}$ | 32 | 48 | 32 | 48 | 32 | 40 | 40 | 40 |
| $R_{\Delta,i}$ | 0.5 | 0.75 | 0.5 | 0.75 | 0.5 | 0.63 | 0.63 | 0.63 |

Each $S_i$ is a fixed $4 \times 16$ array whose entries come from the integers 0–15. Given a bit string of length 6, say $B_j = b_1 \cdots b_6$, we compute $S_j(B_j)$ as follows. The two bits $b_1b_6$ determine the binary representation of a row $r(0 \le r \le 3)$ of $S_j$, and the four bits $b_2b_3b_4b_5$ determine the binary representation of a column $c(0 \le c \le 15)$ of $S_j$. Then $S_j(B_j)$ is defined to be the entry $S_j(r,c)$, written in binary as a bit string of length 4.

For our convenience, let $R_{\sigma,i}$ and $R_{\Delta,i}$ be the sum-of-square measure and the absolute measure of $f_i$, respectively. Then we have Tables 4.2 as measure values for component functions of each $S_i$-boxes, $i = 1, \cdots, 8$ of DES.

Let $m_{\sigma_f}(m_{\Delta_f})$ and $s_{\sigma_f}(s_{\Delta_f})$ (resp.) be the mean and the standard deviation (resp.) of

$$\{\sigma_{f_1}, \sigma_{f_2}, \sigma_{f_3}, \sigma_{f_4}\} \text{ and } (\{\Delta_{f_1}, \Delta_{f_2}, \Delta_{f_3}\Delta_{f_4}\}), \text{ respectively.}$$

Similarly, let $m_{R_\sigma}(m_{R_\Delta})$ and $s_{R_\sigma}(s_{R_\Delta})$ (resp.) be the mean and the standard deviation of $\{R_{\sigma,1}, R_{\sigma,2}, R_{\sigma,3}, R_{\sigma,4}\}$ and $(\{R_{\Delta,1}, R_{\Delta,2}, R_{\Delta,3}, R_{\Delta,4}\})$ (resp.). From Table 4.1 of $S_i$-boxes, the following results can be presented (see Table 4.3).

From Table 4.2 $R_{\sigma_f}$ is in between 0.030 and 0.116, which is close to 0. That implies that the nonlinearity of each component function in all $S$-boxes is pretty large. Furthermore, from Table 4.3 the average value $R_{\sigma_f}$ of 4 components is in

TABLE 4.3. The mean and the standard deviation

| | $S_1$-box | | $S_2$-box | | $S_3$-box | | $S_4$-box | |
|---|---|---|---|---|---|---|---|---|
| $i$ | $m_i$ | $s_i$ | $m_i$ | $s_i$ | $m_i$ | $s_i$ | $m_i$ | $s_i$ |
| $\sigma_f$ | 16768 | 3270 | 19072 | 4445 | 18112 | 3331 | 15232 | 0 |
| $R_\sigma$ | 0.049 | 0.013 | 0.058 | 0.017 | 0.055 | 0.013 | 0.043 | 0 |
| $\Delta_f$ | 32 | 5.66 | 42 | 8.72 | 42 | 6.63 | 24 | 0 |
| $R_\Delta$ | 0.5 | 0.088 | 0.66 | 0.145 | 0.66 | 0.103 | 0.38 | 0 |
| | $S_5$-box | | $S_6$-box | | $S_7$-box | | $S_8$-box | |
| $i$ | $m_i$ | $s_i$ | $m_i$ | $s_i$ | $m_i$ | $s_i$ | $m_i$ | $s_i$ |
| $\sigma_f$ | 14848 | 2796 | 16576 | 1685 | 22048 | 7734 | 14752 | 2021 |
| $R_\sigma$ | 0.042 | 0.011 | 0.049 | 0.007 | 0.07 | 0.03 | 0.041 | 0.008 |
| $\Delta_f$ | 32 | 5.66 | 40 | 5.66 | 40 | 8 | 38 | 3.46 |
| $R_\Delta$ | 0.5 | 0.088 | 0.63 | 0.088 | 0.63 | 0.125 | 0.6 | 0.056 |

between 0.041 and the standard deviations is less than 0.03. We compare the $S_4$-box is better than the others since $R_{\sigma_f} = 0.043, R_{\Delta_f} = 0.38$ are the smallest and their standard deviations are 0. On the other hand, as the standard deviations of $S_7$-box are 7734, 0.03, 8, and 0.125 which are rather larger than the other $S$-boxes, $S_7$-box is easy to be attacked.

## References

1. C. M. Adams and S. E. Tavares: Generating and counting binary bent sequences. *IEEE Trans. Inform. Theory* **36** (1990), no. 5, 1170–1173. MR **91e:**11024
2. J. F. Dillon: A survey of bent functions. *NSA Technical J.* **1972**, 191–215.
3. B. Preneel, W. Van Leekwijck, L. Van Linden, R. Govaerts and J. Vandewalle: Propagation characteristics of Boolean functions. In: *Advances in Cryptology-EUROCRYPT '90 (Aarhus, 1990)*, (pp. 161–173). Lecture Notes in Comput. Sci., 473, Springer, Berlin, 1991. CMP **91:11)** 1 102 479
4. M. S. Rhee, H. Y. Shin and Y. B. Youn: Algorithms for generating nonlinear combiners with given conditions. *Korean J. Comput. Appl. Math. Ser. A* **7** (2000), no. 1, 269–278.
5. J. Seberry and X. M. Zhang: Highly nonlinear 0-1 balanced functions satisfying strict avalanche criterion (extended abstract). In: J. Seberry and Y. Zheng (Eds.), *Advances in Cryptology—AUSCRYPT '92* (pp. 145–155). Lecture Notes in Computer Science, 718. Springer-Verlag, Berlin, 1993. MR **95d:**94002
6. D. R. Stinson: *Cryptography. Theory and practice.* CRC Press Series on Discrete Mathematics and its Applications. CRC Press, Boca Raton, FL, 1995. MR **96k:**94015
7. A. F. Webster: *Plaintext/ciphertext bit dependencies in cryptographic system.* Master's Thesis, Department of Electrical Engineering, Queen's University, Canada, 1985.
8. A. F. Webster and S. E. Tavares: On the design of $S$-boxes. In: H. C. Williams (ed.), *Advances in Cryptology—CRYPTO '85* (pp. 523–534). Lecture Notes in Computer Science, 218. *Springer-Verlag, Berlin*, 1986. MR **87d:**94002
9. R. Yarlagadda and J. E. Hershey: Analysis and synthesis of bent sequences. *IEEE Proceedings (Part E)* **136** (1989), 112–123.
10. X. M. Zhang, and Y. Zheng: *GAC-the Criterion for Global Avalanche Characteristics of Cryptographic Functions. J. UCS* **1** (1995), no. 5, 320–337. MR **97k:**94040

(W. KIM) DEPARTMENT OF MATHEMATICS, HOSEO UNIVERSITY, BAEBANG-MYEON, ASAN, CHUNGNAM 337-850, KOREA
*E-mail address*: `kimws@office.hoseo.ac.kr`

(Y. S. KIM) DEPARTMENT OF MATHEMATICS, HOSEO UNIVERSITY, BAEBANG-MYEON, ASAN, CHUNGNAM 337-850, KOREA
*E-mail address*: `yskim@office.hoseo.ac.kr`

(M. S. RHEE) DEPARTMENT OF APPLIED MATHEMATICS, DANKOOK UNIVERSITY, ANSEO-DONG, CHEONAN, CHUNGNAM 330-714, KOREA
*E-mail address*: `msrhee@anseo.dankook.ac.kr`