Improving the strict avalanche characteristics of cryptographic functions

Jennifer Seberry *, Xian-Mo Zhang **, Yuliang Zheng ***

The Centre for Computer Security Research, Department of Computer Science, The University of Wollongong, Wollongong, NSW 2522, Australia

(Communicated by S.G. Aki; received 23 May 1993; revised 7 September 1993)

Abstract

This letter presents a simple yet effective method for transforming Boolean functions that do not satisfy the strict avalanche criterion (SAC) into ones that satisfy the criterion. Such a method has a wide range of applications in designing cryptographically strong functions, including substitution boxes (S-boxes) employed by common key block encryption algorithms.

Key words: Boolean functions; Cryptography; Security in digital systems; Strict avalanche criterion (SAC); Substitution boxes (S-boxes)

1. The strict avalanche criterion

A (Boolean) function on $V_{2^n}$, where $V_{2^n}$ denotes the vector space of $n$-tuples of elements from $GF(2)$, is said to satisfy the strict avalanche criterion (SAC) if complementing a single bit in its input results in the output of the function being complemented half the time over all the input vectors. The SAC is a very important requirement for cryptographic functions. The formal definition for the SAC seems to appear first in the open literature in 1985 [16,17].

Definition 1. Let $f$ be a function on $V_{2^n}$. $f$ is said to satisfy the SAC if $f(x) \oplus f(x \oplus a)$ assumes the values $0$ and $1$ an equal number of times, or simply, $f(x) \oplus f(x \oplus a)$ is balanced, for every $a \in V_{2^n}$ with $W(a) = 1$, where $x = (x_1, \ldots, x_n)$ and $W(a)$ denotes the number of ones in (or the Hamming weight of) the vector $a$.

A closely related concept is propagation criterion [1,12,11]:

Definition 2. Let $f$ be a function on $V_{2^n}$. We say that $f$ satisfies

1) the propagation criterion with respect to a non-zero vector $a$ in $V_{2^n}$ if $f(x) \oplus f(x \oplus a)$ is a balanced function,

2) the propagation criterion of degree $k$ if it satisfies the propagation criterion with respect to all $a \in V_{2^n}$ with $W(a) = k$.

As the SAC is equivalent to the propagation criterion of degree 1, the latter can be viewed as a generalization of the former. In another direc-
tion, the SAC has been generalized to higher order SAC. This work is represented by [5]. In this letter we shall not pursue further the developments in these two directions. Instead we shall focus our attention on how to transform functions which do not satisfy the SAC into ones that satisfy the criterion.

2. Single functions

First we introduce the following basic theorem.

**Theorem 3.** Let $f$ be a function on $V_n$, and $A$ be a nondegenerate matrix of order $n$ whose entries are from GF(2). Suppose that $f(x) \oplus f(x \oplus y)$ is balanced for each row $y$ of $A$, where $i = 1, \ldots, n$ and $x = (x_1, \ldots, x_n)$. Namely $f$ satisfies the propagation criterion with respect to all rows of $A$. Then $\phi(x) = f(\alpha x)$ satisfies the SAC.

**Proof.** Let $\delta_i$ be a vector in $V_n$ whose entries, except the ith, are all zero. Note that $W(\delta_i) = 1$ and $\delta_i \cdot A = y_i$, $i = 1, \ldots, n$. Then we have $\phi(x) \oplus \phi(x \oplus \delta_i) = f(\alpha x) \oplus f(\alpha x \oplus \delta_i) = f(\alpha x \oplus f(\alpha x \oplus y_i)$, where $u = \alpha x$. Since $A$ is nondegenerate, $u$ runs through $V_n$ while $x$ does. By assumption, $f(u) \oplus f(u \oplus y_i)$ runs through the values zero and one an equal number of times while $u$ runs through $V_n$. Consequently $\phi(x) \oplus \phi(x \oplus \delta_i)$ runs through the values zero and one an equal number of times while $x$ runs through $V_n$. That is, $\phi(x)$ satisfies the SAC.

Note that the algebraic degree, the nonlinearity, and the balancedness of a function is unchanged under a nondegenerate linear transformation of coordinates [6,13]. In addition the number of nonzero vectors with respect to which the function satisfies the propagation criterion is also invariant under the transformation [13]. In the case of S-boxes (tuples of functions), the profile of its difference distribution table, which measures the strength against the differential cryptanalysis [3,4], also remains invariant under such a transformation [15]. Thus Theorem 3 provides us with a very useful tool to improve the strict avalanche characteristics of cryptographic functions. In the following we consider two applications of the theorem.

**Application 4.** Our first application shows that a SAC-satisfying function on a higher dimensional space can be easily obtained from a SAC-satisfying function on a lower dimensional space.

Let $g(y_1, \ldots, y_s)$ be a function on $V_s$ that satisfies the SAC. Adding $t$ dummy-coordinates $x_{i+1}, \ldots, x_t$ into $y$, we obtain a function $f$ on $V_{s+t}$, namely,

$$f(y_1, \ldots, y_s, x_{i+1}, \ldots, x_t) = g(y_1, \ldots, y_s).$$

The $t$ newly added coordinates have no influence on the output of $f$. Hence $f$ does not satisfy the SAC.

Let $A$ be a nondegenerate matrix of order $s + t$. Assume that each row $y_i$ of $A$ can be written as $y_i = (y_{i1}, a_i)$, whose $W(\beta_i) = 1$, $y_{i2} \in V_s$, and $a_i \in V_s$. Let $x = (x_1, \ldots, x_t)$, $y = (y_1, \ldots, y_s, x_{i+1}, \ldots, x_t)$, and $z = (y, x)$. Then we have $f(z) = f(z \oplus y_i) = g(y) \oplus f(y \oplus \beta_i)$. This shows that $f(z) \oplus f(z \oplus y_i)$ is balanced for $y_i = 1$, $t + 1, \ldots, s + t$. By Theorem 3, $\phi(z) = f(\alpha z)$ satisfies the SAC.

An example of the matrices that satisfy the requirements is as follows,

$$A = \begin{pmatrix} \beta_1 & 0_{s \times t} \\ 0_{s \times s} & I_t \end{pmatrix},$$

where $I$ denotes the identity matrix, $0$ denotes the zero matrix, and $Q$ is a matrix that contains precisely a one in each of its rows.

$\phi$ and $f$ have the same nonlinearity, algebraic degree, and balancedness as $f(z)$ does. The two functions also have the same number of nonzero vectors with respect to which $\phi$ satisfies the propagation criterion. The net gain of $\phi$ over $f$ is the SAC. However, it should be pointed out that for this particular example, the resulting function $\phi$ does not satisfy the propagation criterion with respect to vectors whose entries are zeros except in the first and the ($s + j$)th, where $1 \leq j \leq t$. This property might be undesirable in certain applications. We can get around the problem by selecting a nondegenerate matrix $A$ that introduces a function.
more inter-dependencies among the coordinates. Here is such a matrix,

\[
A = \begin{bmatrix}
I_r & 0_{r,s-1} & A_{1} & B_{s-1} \\
0_{r,s-1} & I_r & 0_{s-1} & 0_{s-1} \\
A_{1} & 0_{s-1} & I_r & 0_{r,s-1} \\
B_{s-1} & 0_{s-1} & 0_{r,s-1} & I_r
\end{bmatrix}
\]

(2)

where \( B \) is an arbitrary matrix whose entries are taken from GF(2).

Application 5. Let \( g_0 \) and \( g_1 \) be functions on \( V_r \). Then

\[
f(y_1, x_1, \ldots, x_s) = (1 \oplus y_1) g_0(x_1, \ldots, x_s) \oplus y_1 g_1(x_1, \ldots, x_s)
\]

is a function on \( V_r \). The truth table of \( f \) can be obtained by concatenating the truth tables of \( g_0 \) and \( g_1 \). For this reason, we say that \( f \) is the concatenation of \( g_0 \) and \( g_1 \). Similarly, we can define the concatenation of 2\(^s\) functions on \( V_r \). The result is a function on \( V_r \). To simplify the representation of the concatenation of 2\(^s\) functions, we introduce the following notation.

For each vector \( \delta = (i_1, \ldots, i_s) \in \{0, 1\}^s \), we define a function \( D_\delta \) on \( V_r \) by

\[
D_\delta(x) = (y_1 \oplus i_1) \ldots (y_s \oplus i_s)
\]

where \( y = (y_1, \ldots, y_s) \) and \( i \) denotes the binary complement of \( i \), namely \( i = 1 \oplus i \). For instance, when \( s = 2 \) we have \( D_{00}(y_1, y_2) = (y_1 \oplus 1)(y_2 \oplus 1) \), and when \( s = 3 \) we have \( D_{000}(y_1, y_2, y_3) = (y_1 \oplus 1)(y_2 \oplus 1)(y_3 \oplus 1) \). Note that \( D_{000}(x) = 1 \) if and only if \( y = 0 \).

Using this notation, the concatenation of 2\(^s\) functions on \( V_r \), \( s_0, \ldots, s_t \), is defined as

\[
f(y, x) = \bigoplus_{\delta \in \{0, 1\}^s} D_\delta(x) g_\delta(x)
\]

(3)

where \( x = (x_1, \ldots, x_s) \). Note that each \( g_\delta \) is a function on \( V_r \) and is indexed by a vector in \( \{0, 1\}^s \).

Of particular interest is the concatenation of linear functions on \( V_r \). In Theorems 4 and 5 of [14], the following result is proved:

Lemma 6. When \( t > s \) and all \( g_\delta \), \( \delta \in \{0, 1\}^s \), are distinct nonzero linear functions on \( V_r \), the function \( f \) constructed by (3) is highly nonlinear and balanced. In addition, \( f \) satisfies the propagation criterion with respect to all \( y = (b_1, \ldots, b_s) \), where \( b_1 \neq b_s \), is a nonzero vector in \( V_r \) and \( a \) is an arbitrary vector in \( V_r \).

Let \( A \) be a nondegenerate matrix of order \( r + s \). Suppose that the \( r \)th row \( y_r \) of \( A \) can be written as \( y_r = (b_1, a) \) with \( b_1 \neq 0 \), where \( b_1 \neq b_s \), and \( a \in V_r \). From Lemma 6 we know that \( f \) satisfies the propagation criterion with respect to all rows of \( A \). By Theorem 3, \( \phi(x) = f(xA) \) satisfies the SAC. Note that the matrix \( A \) defined by (1) or (2) satisfies the requirements.

These discussions hold also for the more general case where \( f \) is defined by

\[
f(y, x) = \bigoplus_{\delta \in \{0, 1\}^s} D_\delta(x) g_\delta(x) \oplus r(y)
\]

where \( r \) is an arbitrary function on \( V_r \).

3. A set of functions

In computer security practice, such as the design of S-boxes, we often consider a set of functions. It is desirable that all component functions in a set simultaneously satisfy the SAC. From Theorem 3 we can see that given a set of functions \( f_1, \ldots, f_m \) on \( V_r \), if \( A \) is a nondegenerate matrix of order \( n \) such that \( f_i(x) \oplus f_j(x \oplus y) \) is balanced for every function \( f_j \) and every row \( y \) in \( A \), then \( g(x) = f(xA) \) satisfies the SAC. The following theorem gives a sufficient condition for the existence of such a nondegenerate matrix.

Theorem 7. Let \( f_1, \ldots, f_m \) be functions on \( V_r \). Denote by \( B \) the set of all functions \( g \) on \( V_r \) such that \( f_j(x) \oplus f_j(x \oplus y) \) is not balanced for some \( 1 \leq j \leq m \), and by \( |B| \) the number of vectors in \( B \). If \( |B| < 2^{n-1} \), then there exists a nondegenerate matrix \( A \) of order \( n \) with entries from GF(2) such that each \( \phi(x) = f_j(xA) \) satisfies the SAC.

Proof. We show how to construct a nondegenerate matrix \( A \) of order \( n \), under the condition that \( |B| < 2^{n-1} \). Denote by \( S_{m-1, m-1} \) the set of vec-
tors consisting of all the linear combinations of vectors $\alpha_1, \ldots, \alpha_k$. The first row of $A$, $y_1$, is selected from $V_r$ excluding those in $R$ and the zero vector, i.e., from the vector set $V_r - B - S_r$. There are $2^n - |B| - 2^h$ different choices for $y_1$. The second row of $A$, $y_2$, is selected from the vector set $V_r - B - S_r$. This guarantees that $y_2$ is linearly independent of $y_1$. We have $2^n - |B| - 2^h$ different choices for $y_2$.

In general, once the first $k-1$ linearly independent rows $y_1, \ldots, y_{k-1}$ of $A$ are selected, the $k$th row $y_k$, $k \leq n$, will be selected from the vector set $V_r - B - S_r - y_1, \ldots, y_{k-1}$. This process ensures that $y_1, \ldots, y_k$ are all linearly independent.

The number of choices for the last row $y_k$ is $2^n - |B| - 2^n - 1 = 2^n - 1 - |B| > 0$. Therefore, we can always find a nondegenerate matrix $A$ such that $f(x) \equiv f(x \oplus y)$ is balanced for every $1 \leq i \leq m$ and $1 \leq j \leq n$. By Theorem 3, $\phi(x) = f_1(x), \ldots, f_n(x) = f_n(x)$ all satisfy the SAC.

As is discussed in Section 2, the transformation technique does not affect the nonlinearity, the algebraic degree and the balancedness of the component functions of an S-box. The profile of the difference distribution table of the S-box, and the number of nonzero vectors with respect to which the component functions satisfy the propagation criterion are not altered either. This technique has been successfully applied in [15] to design S-boxes that possess many desirable cryptographic properties, which include the high nonlinearity, the SAC, the balancedness and the robustness against differential cryptanalysis. As is shown below, the technique can also be applied to other approaches to the construction of S-boxes.

**Application 8.** S-boxes based on permutation polynomials are studied in [2, 7-10]. In general, these permutations do not satisfy the SAC. Employing the transformation technique discussed above, the strict avalanche characteristics of these permutations can be improved. In particular, with the permutations constructed by the "cubing" method [8-10], each component function $f_j$ satisfies the propagation criterion with respect to all but one nonzero vectors in $V_r$, where $n > 3$ is odd. Note that $|B| \leq n$. A component function fails to satisfy the SAC if the Hamming weight of the nonzero vector with respect to which the propagation criterion is not satisfied is one. If this is the case, by Theorem 7 we can use a nondegenerate matrix to transform the component functions of such a permutation so that they all satisfy the SAC.

### 4. A final remark

In [13], we have constructed highly nonlinear balanced functions on $\mathbb{F}_{2^n}$ that satisfy the propagation criterion of degree $2k$, and highly nonlinear balanced functions on $\mathbb{F}_{2^n}$ that satisfy the propagation criterion of degree $\leq k$. A transformation technique similar to that presented in this letter has played an important role in the constructions.

### 5. Acknowledgment

The authors would like to thank the anonymous referee for his or her helpful comments. The first author was supported in part by the Australian Research Council under the reference numbers A49130102, A4930136, A4917885 and A49223172, the second author by A49180102, and the third author by A49232172.

### 6. References


