BOUNDS ON THE NUMBER OF FUNCTIONS SATISFYING THE STRICT AVALANCHE CRITERION

Thomas W. Cusick and Pantelimon Stănică

Department of Mathematics, SUNY at Buffalo, 106 Diefendorf Hall
Buffalo, New York 14214-3093

1. Introduction

The Strict Avalanche Criterion (SAC) was introduced by Webster and Tavares [3] in a study of design criteria for certain cryptographic functions. For $\mathbb{Z}_2 = \{0, 1\}$, a Boolean function $f : \mathbb{Z}_2^n \to \mathbb{Z}_2$ is said to satisfy the SAC if complementing a single input bit results in changing the output bit with probability exactly one half. Thus f satisfies the SAC if and only if

$$\sum_{\boldsymbol{x} \in \mathbb{Z}_2^n} \hat{f}(\boldsymbol{x}) \hat{f}(\boldsymbol{x} + \boldsymbol{e}_i) = 0, \ i = 1, ..., n \tag{1}$$

where $\boldsymbol{e}_i$ is the i-th standard basis vector and $\hat{f}(\boldsymbol{x}) = (-1)^{f(\boldsymbol{x})}$.

As is the case with any criterion of cryptographic significance, it is of interest to count the functions which satisfy the criterion. Two recent papers [1, 2] have been concerned with counting functions that satisfy the SAC. In this paper we continue

Key words and phrases. Cryptography, Strict Avalanche Criterion, Boolean functions, enumeration, combinatorial problems.

the work in [1]. We improve an estimate given in [1] and we prove some conjectures stated there.

2. Main results

Throughout the paper we assume that $n \geq 4$. We define two sets as follows:

$$\begin{aligned} T_1 = \{&A = 0,0,1,1;\ \bar{A} = 1,1,0,0;\ B = 0,1,0,1;\ \bar{B} = 1,0,1,0;\\ &C = 0,1,1,0;\ \bar{C} = 1,0,0,1;\ D = 0,0,0,0;\ \bar{D} = 1,1,1,1\} \end{aligned} \tag{2}$$

$$\begin{aligned} T_2 = \{&U = 1,0,0,0;\ \bar{U} = 0,1,1,1;\ V = 0,0,0,1;\ \bar{V} = 1,1,1,0;\\ &X = 0,1,0,0;\ \bar{X} = 1,0,1,1;\ Y = 0,0,1,0;\ \bar{Y} = 1,1,0,1\}. \end{aligned} \tag{3}$$

We will use these sets to construct a large family of functions satisfying the SAC. First we state a very helpful lemma, namely

Lemma 1. *If the Boolean function f is given as a vector $(v_1, v_2, \ldots, v_{2^n})$, $v_i \in \{0,1\}$, then f satisfies the SAC if and only if*

$$\begin{aligned} &(w_1 w_{2^{i-1}+1} + w_2 w_{2^{i-1}+2} + \cdots + w_{2^{i-1}} w_{2^i}) + \\ &(w_{2^i+1} w_{2^i+2^{i-1}+1} + \cdots + w_{2^i+2^{i-1}} w_{2^{i+1}}) + \cdots + \\ &(w_{2^n-2^i+1} w_{2^n-2^{i-1}+1} + \cdots + w_{2^n-2^{i-1}} w_{2^n}) = 0 \end{aligned} \tag{4}$$

for each $i = 1, 2, \ldots, n$, where $w_i = (-1)^{v_i}$.

Proof. The lemma follows directly from (1). *qed.*

Now we state a result that enables us to construct a family of functions satisfying the SAC. We say that a string of elements of $\mathbb{Z}_2$ is *balanced* if it has an equal number of 0's and 1's. When we write $f = (v_1, \ldots, v_{2^n})$ we assume that the values of $f(\boldsymbol{x})$ are given as $\boldsymbol{x}$ runs through the vectors in $\mathbb{Z}_2^n$ given in the lexicographical order.

Theorem 1. *If $f = (X_i)_{i=1,2,\ldots,2^{n-2}}$, where either all $X_i \in T_1$, or all $X_i \in T_2$ is*

such that all of the following sums of vectors in $\mathbb{Z}_2^4$

$$X_1 + X_{2^{i-3}+1}, X_2 + X_{2^{i-3}+2}, \ldots, X_{2^{i-3}} + X_{2^{i-2}},$$

$$X_{2^{i-2}+1} + X_{2^{i-2}+2^{i-3}+1}, \ldots, X_{2^{i-2}+2^{i-3}} + X_{2^{i-1}}, \tag{5}$$

$$X_{2^{n-2}-2^{i-2}+1} + X_{2^{n-2}-2^{i-3}+1}, \ldots, X_{2^{n-2}-2^{i-3}} + X_{2^{n-2}},$$

for each $i = 3, 4, \ldots, n$, *are balanced, then* f *satisfies the SAC.*

Proof. We use Lemma 1 and formula (4). It will suffice to prove that in (4) each parenthesis is zero. In fact we shall show that the sum of each four consecutive terms in each parenthesis is zero, if $i \geq 3$. Any function obtained by concatenation of blocks from our sets T_1 or T_2 satisfies (4) for $i = 1$. For $i = 2$, the equation (4) is a simple calculation. We restrict our attention to the case of $i \geq 3$. In each parenthesis in (4) we have 2^{i-1} terms and since $i \geq 3$ the number of terms is divisible by 4, which means each string either has all or none of its elements in any parenthesis. From the relations (5) we know that the sum of each two strings X_i is balanced and so in (4) we have the same number of 1's and -1's. We take for instance the first four terms $w_1 w_{2^{i-1}+1} + w_2 w_{2^{i-1}+2} + w_3 w_{2^{i-1}+3} + w_4 w_{2^{i-1}+4}$ which is zero because it corresponds to the concatenation of two strings with balanced sum.

qed.

We let S_n denote the number of Boolean functions f in n variables which satisfy the SAC. We define

$$\exp_2(x) = 2^x.$$

There are $\exp_2(2^n)$ Boolean functions $f : \mathbb{Z}_2^n \to \mathbb{Z}_2$. One of the simplest questions we can ask about the number S_n concerns the size of the ratio L_n defined by

$$L_n = 2^{-n} \log_2 \ S_n.$$

Conjecture 1 of [1] stated that

$$L = \lim_{n\to\infty} L_n \text{ exists and } 1 > L \geq \frac{1}{2},$$

but in [1] it was only proved that $L_n \geq \frac{1}{4}$, for all n. Our next theorem improves this to $L_n \geq \frac{1}{2}$.

The same inequality was proved independently by Youssef and Tavares [4], using a different method. More recently, Daniel Biss has given a much more complicated argument that shows $L = 1$, thereby disproving Conjecture 1 of [1].

Theorem 2. *We can construct* $\exp_2(2^{n-1})$ *Boolean functions of* n *variables which satisfy the SAC.*

Proof. First we introduce a function defined on a pair of blocks, that will turn out to be very useful in what follows, namely $h_n(P, Q)$ which denotes the strings $PQ,\ PQQP,\ PQQPQPPQ, \ldots$, for $n = 4,\ 5,\ 6$ (the string for $n = k+1$ is the string for $n = k$ followed by the string for $n = k$ with P and Q interchanged). Next we define the Boolean functions expressed by the vector

$$(M_{2^{n-3}}, h_n(P, \bar{P})), \tag{6}$$

where either both $M, P \in T_1$ or $M, P \in T_2$. It is very easy to show (for example, by using Theorem 1) that these functions are SAC-functions. Now we shall apply a procedure to (6) to produce functions that satisfy the SAC. If in the vector (6) we interchange the j-th block M counted from the left ($1 \leq j \leq 2^{n-3}$) in the first string $M_{2^{n-3}}$ with the j-th block P or $\bar{P}$ in the second string $h_n(P, \bar{P})$, j counted from the left, we get SAC-functions. Furthermore, applying any subset of the 2^{n-3}

possible interchanges we still obtain SAC-functions. So we obtain $\exp_2(2^{n-2})$ functions satisfying the SAC.

In order to prove the general result we shall introduce a map denoted $\odot$ from the Cartesian product $\{A, \bar{A}, \dots, X, \bar{X}, \dots\} \times \{A, \bar{A}, \dots, X, \bar{X}, \dots\}$ to $\{-4, -2, 0, 2, 4\}$ defined by

$$M \odot N = \#0's - \#1's \text{ in the sum } M \oplus N, \tag{7}$$

where $\oplus$ denotes component-wise addition of strings over $\mathbb{Z}_2$. Then the relation (4) reads as follows, for an arbitrary $i \geq 3$:

$$(X_1 \odot X_{2^{i-3}+1} + X_2 \odot X_{2^{i-3}+2} + \cdots + X_{2^{i-3}} \odot X_{2^{i-2}}) + \cdots = 0. \tag{8}$$

We shall use the previous idea of interchanging to produce the necessary number of SAC-functions. Consider the following vector (a special case of (6))

$$(U_{2^{n-3}}, h_n(X, \bar{X})). \tag{9}$$

When we replace the j-th block U in the first string $U_{2^{n-3}}$ by M and the j-th block X or $\bar{X}$ in the second string $h_n(X, \bar{X})$ by N, j counted from the left in both strings, we denote this by (M, N). We replace in (9) the block pair $(U, \bar{X})$ by one of the pairs

$$\begin{aligned} &(U, \bar{X}); (\bar{X}, \bar{U}); (\bar{U}, X); (X, U); (V, \bar{Y}); (\bar{Y}, \bar{V}); (\bar{V}, Y); (Y, V); \\ &(A, B); (B, \bar{A}); (\bar{A}, \bar{B}); (\bar{B}, A); (\bar{C}, \bar{D}); (\bar{D}, C); (C, D); (D, \bar{C}); \end{aligned} \tag{10}$$

or the block pair (U, X) by one of the pairs

$$\begin{aligned} &(U, X); (X, \bar{U}); (\bar{U}, \bar{X}), (\bar{X}, U); (V, Y); (Y, \bar{V}); (\bar{V}, \bar{Y}); (\bar{Y}, V); \\ &(A, \bar{B}); (\bar{B}, \bar{A}); (\bar{A}, B); (B, A); (C, \bar{D}); (\bar{D}, \bar{C}); (\bar{C}, D); (D, C). \end{aligned} \tag{11}$$

Performing replacements on each of the 2^{n-3} possible blocks U we get $\exp_{16}(2^{n-3})$ $= \exp_2(2^{n-1})$ functions. To complete the proof, we show that all of these functions satisfy the SAC. This is easy to see if we observe that for any two pairs (M, N) from (10) and (P, Q) from (11) we have

$$\begin{aligned} M \odot P + N \odot Q = 0, \\ M \odot N = 0,\ P \odot Q = 0. \end{aligned} \tag{12}$$

Since for any pair (M, N) from either of the two lists (10) and (11) we have $M \odot N = 0$, the equation (8) is satisfied for $i = n$. For $i \leq n-1$ the equation (8) is obviously satisfied because in the equation (8), for any product $M \odot N$ in the first half of this equation we have an antidote in the second half, according to (12). *qed.*

Our next theorem proves Conjecture 4 of [1].

Theorem 3. *If the Boolean function f has values $(v_1, \ldots, v_{2^n})$ and we are given any fixed choice of the values v_i, $1 \leq i \leq 2^{n-1}$, then there exists a choice of v_i, $2^{n-1}+1 \leq i \leq 2^n$, such that the resulting function f satisfies the SAC.*

Proof. We assume that we are given the first 2^{n-1} elements in a 2^n vector and we try to find the second half of that vector such that the function represented by it is a SAC-function. Since $n \geq 4$ we must have the number of elements of the first half divisible by 4 and thus we have the first half split up into 2^{n-3} blocks of length 4. These blocks must be from the two sets T_1 and T_2 (no other possibilities exist). In the corresponding positions in the second half we put the blocks given by either the list (10) or the list (11), according to the following rule: If the pair of blocks, one from each half, would be in the same position as a pair $(U, \bar{X})$ in (9) we choose from (10), and if it would be in the same position as a pair (U, X) in (9) we choose from (11). We can do this since we

can find any of the blocks (2) or (3) in the first position in some pair in either (10) or (11), so all possibilities are covered. The proof of Theorem 2 shows that the condition (8) is satisfied. Thus our function satisfies the SAC and our Theorem 3 is proved.

qed.

The construction of the SAC functions in the proof of Theorem 3 uses the lists (10) and (11). It is natural to ask if a similar construction could be carried out with different lists. We can prove that there are exactly 9 other possibilities for the lists (10) and (11), as follows: In (9) we cannot replace the j-th block by a pair with one component in (2) and the other one in (3). This is clear from applying (8) for $i = n$. So if the first component is in (2) so is the second component. First we restrict our attention to the case of replacing $(U, \bar{X})$. Suppose that we try to replace it by (A, M), M to be determined later. From (12) we get $A \odot M = 0$, which implies that $M \in \{B, \bar{B}, C, \bar{C}, D, \bar{D}\}$. Furthermore, from (12) with $i = n - 1$ we get $A \odot U + M \odot \bar{X} = 0$, that is $M \odot \bar{X} = 2$ and thus we obtain $M \in \{\bar{B}, \bar{C}, \bar{D}\}$. For $i \leq n - 2$ the equation (8) is fulfilled for these values of M. Performing a similar analysis with A replaced by one of the other possible blocks, we come up with the

following list of possible replacements for $(U, \bar{X})$:

$$(A,B);(A,C);(\bar{A},\bar{D});(\bar{A},\bar{B});(\bar{A},\bar{C});(A,D);(B,C);(B,\bar{A});(\bar{B},\bar{D});$$

$$(\bar{B},A);(B,D);(\bar{B},\bar{C});(C,B);(C,\bar{A});(\bar{C},\bar{D});(\bar{C},A);(C,D);(\bar{C},\bar{B});$$

$$(D,\bar{B});(D,\bar{C});(D,A);(\bar{D},\bar{A});(\bar{D},B);(\bar{D},C);(\bar{U},X);(X,U);(X,\bar{U});$$

$$(X,V);(X,\bar{V});(X,Y);(X,\bar{Y});(U,\bar{X});(\bar{X},U);(\bar{X},\bar{U});(\bar{X},V);(\bar{X},\bar{V});$$

$$(\bar{X},Y);(\bar{X},\bar{Y});(V,U);(V,\bar{U});(V,Y);(V,\bar{Y});(\bar{V},U);(\bar{V},\bar{U});(\bar{V},Y);$$

$$(\bar{V},\bar{Y});(Y,U);(Y,\bar{U});(Y,V);(Y,\bar{V});(\bar{Y},U);(\bar{Y},\bar{U});(\bar{Y},V);(\bar{Y},\bar{V});$$

and possible replacements for (U, X) :

$$(\bar{A},D);(A,\bar{B});(A,\bar{C});(\bar{A},B);(\bar{A},C);(A,\bar{D});(B,A);(\bar{B},D);(B,\bar{C});$$

$$(\bar{B},C);(\bar{B},\bar{A});(B,\bar{D});(C,A);(\bar{C},D);(C,\bar{B});(\bar{C},B);(\bar{C},\bar{A});(C,\bar{D});$$

$$(D,\bar{A});(D,B);(D,C);(\bar{D},\bar{B});(\bar{D},\bar{C});(\bar{D},A);(\bar{U},\bar{X});(V,U);(V,\bar{U});$$

$$(V,Y);(V,\bar{Y});(\bar{V},U);(\bar{V},\bar{U});(\bar{V},Y);(\bar{V},\bar{Y});(Y,U);(Y,\bar{U});(Y,V);$$

$$(Y,\bar{V});(\bar{Y},U);(\bar{Y},\bar{U});(\bar{Y},V);(\bar{Y},\bar{V});(X,U);(X,\bar{U});(X,V);(X,\bar{V});$$

$$(U,X);(X,Y);(X,\bar{Y});(\bar{X},U);(\bar{X},\bar{U});(\bar{X},V);(\bar{X},\bar{V});(\bar{X},Y);(\bar{X},\bar{Y}).$$

Replacing two blocks $(U, \bar{X})$ or (U, X) and looking at the incompatibilities, that is pairs that do not satisfy the equations (12), we see that we can get ten sets each of which could be used to do the job. One set is the one displayed in (10) and (11), which we used above. In order to display the rest of the nine sets we shall uncover another side of our construction. We shall use an idea of [4] and our Theorem 3. We want to write f as

$$f = [f_{n-1}\ [f_{n-1} \oplus g_{n-1}]]. \tag{13}$$

Let H be any block in the function f_{n-1} which is the first half of f. The block from the second half of f is $H \oplus G$, where the block G is defined as follows: We choose a pair of blocks, one each from the first and second halves of the function in (9), such that both blocks have the same position in their respective halves as the block H does in f_{n-1}. If this block pair is U, X, we define G to be $H \oplus P$, where P is the second member of the block pair in a list similar to (11), whose first member is H. If the block pair is $U, \bar{X}$, we define G to be $H \oplus Q$, where Q is the second member of the block pair in a list similar to (10), whose first member is H. This can be done because the union of the first members of any pair in (10) (or (11)) is $\mathbb{Z}_2^4$:

$$D, V, Y, A, X, B, C, \bar{U}, U, \bar{C}, \bar{B}, \bar{X}, \bar{A}, \bar{Y}, \bar{V}, \bar{D}. \tag{14}$$

For our ten lists, the inverse process can also be performed: given g_{n-1} we can write the two lists similar to (10) and (11), using the equations (12). For any function g_{n-1}, which gives (13) as a function which satisfies the SAC (obtained starting from a function of the form (6)), we can complete the two sets of replacements if and only if the union of the first members in any of the two obtained lists is $\mathbb{Z}_2^4$. In what follows we shall display the ten functions which give SAC-functions starting from (9) and using our method. They are (starting with the function g_{n-1} corresponding to (10)

and (11) and using the ordering (14) for the blocks):

$$
\begin{array}{ll}
(C\bar{A}A\bar{C}\bar{A}C\bar{C}A \mid A\bar{C}C\bar{A}\bar{C}A\bar{A}C) & (\bar{A}AB\bar{B}B\bar{B}\bar{A}A \mid A\bar{A}\bar{B}B\bar{B}BA\bar{A}) \\
(BA\bar{B}\bar{A}\bar{B}\bar{A}BA \mid AB\bar{A}\bar{B}\bar{A}\bar{B}AB) & (C\bar{C}A\bar{A}\bar{C}C\bar{A}A \mid A\bar{A}C\bar{C}\bar{A}A\bar{C}C) \\
(\bar{A}CA\bar{C}C\bar{A}\bar{C}A \mid A\bar{C}\bar{A}C\bar{C}AC\bar{A}) & (\bar{A}\bar{A}\bar{A}\bar{A}A\bar{A}\bar{A}A \mid A\bar{A}\bar{A}A\bar{A}\bar{A}\bar{A}\bar{A}) \\
(\bar{A}AA\bar{A}A\bar{A}\bar{A}A \mid A\bar{A}\bar{A}A\bar{A}AA\bar{A}) & (BA\bar{A}\bar{B}\bar{A}\bar{B}BA \mid AB\bar{B}\bar{A}\bar{B}\bar{A}AB) \\
(C\bar{A}B\bar{C}\bar{B}C\bar{C}A \mid A\bar{C}C\bar{B}\bar{C}B\bar{A}C) & (C\bar{A}\bar{B}\bar{C}BB\bar{B}A \mid A\bar{B}BB\bar{C}\bar{B}\bar{A}C).
\end{array}
\tag{15}
$$

References

[1] T. W. Cusick, *Bounds on the number of functions satisfying the Strict Avalanche Criterion*, *Inform. Process. Lett.* **57** (1996), pp. 261-263.

[2] L. O'Connor, *An upper bound on the number of functions satisfying the Strict Avalanche Criterion, Inform. Process. Lett.* **52** (1994) 325-327.

[3] A. F. Webster and S. E. Tavares, *On the design of S-boxes, in: H. C. Williams, ed., Advances in Cryptology-Crypto '85*, Lectures Notes in Computer Science **218** (Springer, Berlin, 1986) 523-534.

[4] A.M. Youssef and S.E. Tavares, *Comment on "Bounds on the number of functions satisfying the Strict Avalanche Criterion"*, *Inform. Process. Lett.*, to appear.