Properties of binary functions

Sheelagh Lloyd Hewlett Packard Laboratories

1. INTRODUCTION

In this paper, we shall investigate the connections between three properties of a binary function : the Strict Avalanche Criterion, balance and correlation immunity. The strict avalanche criterion was introduced by Webster and Tavares [7] in order to combine the ideas of completeness and the avalanche effect. A cryptographic transformation is said to be complete if each output bit depends on each input bit, and it exhibits the avalanche effect if an average of one half of the output bits change whenever a single input bit is changed. Forré [1] extended this notion by defining higher order Strict Avalanche Criteria. A function is balanced if, when all input vectors are equally likely, then all output vectors are equally likely. This is an important property for many types of cryptographic functions. The idea of correlation immunity is also extremely important, especially in the field of stream ciphers, where combining functions which are not correlation immune are vulnerable to ciphertext only attacks (see, for example [4]). The concept of mth order correlation immunity was introduced by Siegenthaler [5] as a measure of resistance against such an attack.

In a previous paper [2], we found conditions under which a function satisfying the highest possible order Strict Avalanche Criterion was also balanced and/or correlation immune. Here we shall look at functions satisfying the next highest order Strict Avalanche Criterion. We shall also investigate higher orders of correlation immunity.

In Section 2, we establish some notation, define the properties to be examined and state characterisations of functions with the various properties. Section 3 is devoted to some preliminary calculations which will enable us to identify conditions the functions must satisfy. We present results on balance in Section 4, on correlation immunity in Section 5, and on simultaneous balance and correlation immunity in Section 6. In each of sections 4, 5 and 6, we shall produce necessary and sufficient conditions for a function to satisfy the criteria.

2. NOTATION AND DEFINITIONS

Although we are really dealing with functions of binary vectors of length n which take values in $\{-1,1\}$, we shall find it convenient to identify a binary vector with its support, that is the set of positions in which it has a 1. We shall, therefore, deal instead with functions from subsets of $\{1, 2, ..., n\}$ to $\{-1, 1\}$.

Let S be the set $\{1, 2, .., n\}$, and let \mathcal{B}_S denote the set of functions which takes subsets of S to $\{-1, 1\}$. We formulate all the definitions and characterisations in terms of such functions.

2.1 Balance

This is the simplest of the three properties, and ensures that the number of 1's produced by f is the same as the number of -1's produced.

Definition 2.1.1 Let $f \in \mathcal{B}_S$. Then f is balanced if and only if

$$\sum_{V\subseteq \mathcal{S}}f(V)=0$$

2.2 Correlation Immunity

Definition 2.2.1 $f \in B_S$ is said to be first order correlation immune if, for any $i \in S$, the probability that $i \in V$, given that V satisfies f(V) = 1, is equal to $\frac{1}{2}$.

The definition is extended to higher orders as follows.

Definition 2.2.2 Let *m* be an integer with $1 \le m \le n$. Then $f \in \mathcal{B}_S$ is said to be *m*th order correlation immune if for any $J \subseteq S$ with |J| = m and any $Y \subseteq J$, the probability that $V \cap J = Y$, given that f(V) = 1, is equal to $\frac{1}{2^m}$.

Note that, for any m with $2 \le m \le n$, mth order correlation immunity implies (m-1)th order correlation immunity.

In order to characterise correlation immune functions, we need to define the Hadamard-Walsh transform.

Definition 2.2.3 The Hadamard-Walsh transform of $f \in B_S$ is defined by

$$H(U) = \sum_{V \subseteq S} f(V)(-1)^{|U \cap V|}.$$

There is a well known formula for inverting the Hadamard-Walsh transform, which we give below.

$$f(W) = \frac{1}{2^n} \sum_{U \subseteq S} H(U)(-1)^{|U \cap W|} \quad \text{for all } W \subseteq S.$$

Xiao and Massey [6] have proved the following theorem characterising correlation immune functions in terms of the values of their Hadamard-Walsh transforms.

Theorem 2.2.4 The function $f \in \mathcal{B}_S$ is mth order correlation immune if and only if H(U) = 0 for all $U \subseteq S$ with $1 \leq |U| \leq m$.

Let us define the integer valued function X by

$$X(W) = \sum_{V \subseteq W} f(V) \text{ for } W \subseteq S.$$

We will find it more convenient to express the characterisation of correlation immunity in terms of the function X. In order to do so, we need the following result.

Lemma 2.2.5 If X and H are defined as above, then

$$X(W) = \frac{1}{2^{n-|W|}} \sum_{\substack{U \subseteq S \\ U \cap W = \emptyset}} H(U) \text{ for all } W \subseteq S.$$

Proof

Since H is the Hadamard-Walsh transform of f, we know that

$$f(W) = \frac{1}{2^n} \sum_{U \subseteq S} H(U)(-1)^{|U \cap W|} \text{ for all } W \subseteq S.$$

Substituting this into the definition of X, we obtain

$$X(W) = \sum_{V \subseteq W} \frac{1}{2^n} \sum_{U \subseteq S} H(U) (-1)^{|U \cap V|}$$

= $\frac{1}{2^n} \sum_{U \subseteq S} H(U) \sum_{V \subseteq W} (-1)^{|U \cap V|}.$

For any $V \subseteq W$, we can write $V = A \cup B$ with $A \subseteq (W \cap U)$ and $B \subseteq (W \setminus U)$. So

$$\sum_{V\subseteq W} (-1)^{|U\cap V|} = \sum_{B\subseteq (W\setminus U)} \sum_{A\subseteq (W\cap U)} (-1)^{|A|}$$

If $W \cap U \neq \emptyset$, there are as many subsets of odd size as of even size, so the sum is 0. If $W \cap U = \emptyset$, then the sum is just $2^{|W|}$. Hence

$$X(W) = \frac{1}{2^{n-|W|}} \sum_{\substack{U \subseteq S \\ U \cap \overline{W} = \emptyset}} H(U) \text{ for all } W \subseteq S.$$

Note that $X(\mathcal{S}) = H(\emptyset)$.

We shall now use this to produce a formulation of mth order correlation immunity in terms of X.

Lemma 2.2.6 If H and X are defined as above, then the following three conditions are equivalent:

- (i) f is mth order correlation immune
- (ii) H(U) = 0 for all $U \subseteq S$ with $1 \le |U| \le m$.
- (iii) $X(W) = 2^{|W|-n}X(S)$ for all $W \subseteq S$ with $(n-m) \leq |W| \leq (n-1)$.

Proof

The equivalence of (i) and (ii) is given by Theorem 2.2.4. We shall now show the equivalence of (ii) and (iii), using Lemma 2.2.5.

Suppose that (ii) holds. Let $W \subseteq S$ be such that $(n-m) \leq |W| \leq (n-1)$, and let $U \subseteq S$ be such that $W \cap U = \emptyset$. Then $0 \leq |U| \leq (n-|W|) \leq m$, so either $U = \emptyset$, or H(U) = 0. So

$$X(W) = \frac{1}{2^{n-|W|}} \sum_{\substack{U \subseteq S\\ U \cap W = \emptyset}} H(U) = \frac{1}{2^{n-|W|}} H(\emptyset).$$

Now $X(S) = H(\emptyset)$, so we have (iii).

Now suppose that (iii) holds. We shall prove (ii) by induction on the size of U. Suppose first that |U| = 1. Let $W = S \setminus U$. Then $V \cap W = \emptyset$ if and only if either $V = \emptyset$ or V = U, so

$$X(W) = \frac{1}{2}(H(\emptyset) + H(U)).$$

But since |W| = (n-1), we know also that $X(W) = \frac{1}{2}H(\emptyset)$. Hence H(U) = 0. Now suppose that $2 \le |U| \le m$ and that H(V) = 0 for all V with $1 \le |V| < |U|$. Let $W = S \setminus U$, then $V \cap W = \emptyset$ if and only if $V \subseteq U$, so

$$X(W) = \frac{1}{2^{n-|W|}} (H(\emptyset) + H(U) + \sum_{V \in U, V \neq \emptyset} H(V)).$$

Now, for any $V \subset U$, $V \neq \emptyset$, we see that $1 \leq |V| < |U|$, so H(V) = 0. Since $(n-m) \leq |W| \leq (n-1)$, we know also that $X(W) = 2^{|W|-n}H(\emptyset)$. Thus we may conclude that H(U) = 0 as required.

Note that we may write the condition that f is balanced as X(S) = 0.

2.3 The Strict Avalanche Criterion

Definition 2.3.1 Let $f \in \mathcal{B}_S$. Then f satisfies the strict avalanche criterion (SAC) if and only if

$$\sum_{V \subseteq (\mathcal{S} \setminus \{j\})} f(V) f(V \cup \{j\}) = 0 \quad \text{for all } j, 1 \le j \le n.$$

We now define the higher order SAC. The SAC defined above is deemed to be the SAC of order 0, and the SAC of order m for $1 \le m \le n-2$ is defined as follows.

Definition 2.3.2 [1] A function $f \in \mathcal{B}_S$ satisfies the SAC of order m, where $1 \leq m \leq (n-2)$ if and only if given any subset \mathcal{T} of \mathcal{S} with $|\mathcal{T}| = n - m$ and any subset P of $\mathcal{S} \setminus \mathcal{T}$, the function $g \in \mathcal{B}_{\mathcal{T}}$ obtained from f by setting $g(V) = f(V \cup P)$ for each $V \subseteq \mathcal{T}$ satisfies the SAC.

Let \overline{f} denote the algebraic normal form of f (so \overline{f} also takes subsets of S to $\{-1,1\}$). We shall sometimes find it convenient to write F for the function from S to $\{1,-1\}$ such that $F(x) = \overline{f}(\{x\})$. To reduce confusion between sets and elements of those sets, we shall use capital letters to denote subsets of S and small letters to denote elements of S. In [3], we proved the following result characterising functions satisfying the SAC of order (n-3). Note that, since we are dealing exclusively with functions satisfying the SAC of order (n-3), we insist throughout that $n \ge 3$.

Theorem 2.3.3 [3] Suppose that $f \in \mathcal{B}_S$. Then f satisfies the SAC of order (n-3) if and only if

$$f(V) = \prod_{U \subseteq V, |U| < 3} \overline{f}(U) \quad ext{for all } V \subseteq \mathcal{S}$$

and for each $x \in S$, there is at most one $y \in S$ for which $\overline{f}(\{x, y\}) = 1$.

Suppose that f satisfies the SAC of order (n-3). Then Theorem 2.3.3 tells us that for any $x \in S$, either $\overline{f}(\{x, y\}) = -1$ for all $y \in S$, or there is exactly one $y \in S$ for which $\overline{f}(\{x, y\}) = 1$. Given any $W \subseteq S$, and any $x \in W$, there can therefore be at most one $y \in W$ for which $\overline{f}(\{x, y\}) = 1$. Suppose there are exactly m pairs (x, y) in W with $\overline{f}(\{x, y\}) = 1$. Let us write these as $(x_1, y_1), \dots, (x_m, y_m)$ (where $0 \le m \le n/2$), and let us denote the remaining elements of W (if any) by $x_{2m+1}, \dots, x_{|W|}$. Then we have $W = \{x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_m, x_{2m+1}, \dots, x_{|W|}\}$ where $\overline{f}(\{x_j, y_j\}) = +1$ for all $1 \le j \le m$ and $\overline{f}(\{a, b\}) = -1$ otherwise. Note that, although the elements may be numbered in various ways, the value of m is determined uniquely by W (and f). We shall find the following definition useful.

Definition 2.3.4 We shall write $A_W(n,m)$ $(W \subseteq S, 0 \leq 2m \leq |W|)$ for the set of functions $f \in \mathcal{B}_S$ satisfying the following conditions:

f satisfies the SAC of order (n-3) and there exist $x_1, x_2, ..., x_m, y_1, y_2, ..., y_m, x_{2m+1}, ..., x_{|W|}$ such that

$$W = \{x_1, x_2, ..., x_m, y_1, y_2, ..., y_m, x_{2m+1}, ..., x_{|W|}\},\$$

 and

$$\overline{f}(\{x_j, y_j\}) = +1, \quad 1 \le j \le m$$

$$\overline{f}(\{a, b\}) = -1 \quad \text{otherwise.}$$

In what follows, we shall want to distinguish the cases where there exists a pair (x, y), such that $\overline{f}(\{x, y\}) = 1$ and F(x) = -F(y), from those where no such pair exists. It turns out that the case where such pairs exist is much simpler than the other case. In order to be able to state some subsequent results concisely in the cases where no such pair exists, we introduce the following notation.

Definition 2.3.6 We shall write $C_W(n, m, r, t, q)$ $(W \subseteq S, 0 \leq 2m \leq |W|, 0 \leq r \leq m, 0 \leq t \leq |W| - 2m, q = 2r + 2t + m - |W|)$ for the set of functions $f \in \mathcal{B}_S$ satisfying the following conditions:

f belongs to $A_W(n,m)$

and

$$F(x_j) = F(y_j) = +1, \quad 1 \le j \le r$$

$$F(x_j) = F(y_j) = -1, \quad r+1 \le j \le m$$

$$F(x_j) = +1, \quad 2m+1 \le j \le 2m+t$$

$$F(x_j) = -1, \quad 2m+t+1 \le j \le |W|$$

For ease of notation, we shall write simply C(n,m,r,t,q) for $C_S(n,m,r,t,q)$. So we see that, if $f \in \mathcal{B}_S$ satisfies the SAC of order (n-3), then either there exists a pair (x,y), such that $\overline{f}(\{x,y\}) = 1$, and F(x) = -F(y), or f belongs to C(n,m,r,t,q) for some values of m, r, t and q.

3. PRELIMINARY CALCULATIONS

We want to express $f(V \cup \{x\})$ in terms of f(V). We know that, given x and V, if f satisfies the SAC of order (n-3), then there is at most one z in V with $\overline{f}(\{x,z\}) = 1$. We first deal with the case where no such z exists.

Proposition 3.1 Suppose that $f \in \mathcal{B}_S$ satisfies the SAC of order (n-3). Suppose further that $x \notin V$, and that $\overline{f}(\{x,y\}) = -1$ for all $y \in V$. Then

$$f(V \cup \{x\}) = (-1)^{|V|} f(V) F(x).$$

Proof

Straightforward application of Theorem 2.3.3.

We turn now to the case where there is a unique element z in V with $\overline{f}(\{x,z\}) = 1$.

Proposition 3.2 Suppose that $f \in \mathcal{B}_S$ satisfies the SAC of order (n-3). Suppose further that $x \notin V$, and that $\overline{f}(\{x,z\}) = 1$ (so $\overline{f}(\{x,y\}) = -1$ for all $y \in V$, $y \neq z$). Then

$$f(V \cup \{x\}) = (-1)^{|V|-1} f(V) F(x).$$

Proof

Straightforward application of Theorem 2.3.3.

We are now able to produce an expression for $X(W) = \sum_{V \subseteq W} f(V)$ in terms of the values of \overline{f} . In order to prove this, we need also to produce the corresponding expression for $X^{-}(W) = \sum_{V \subseteq W} (-1)^{|V|} f(V)$ as well.

Lemma 3.3 Suppose that f satisfies the SAC of order (n-3). Let $W \subseteq S$, $x \in W$ and $U = W \setminus \{x\}$. Suppose that $\overline{f}(\{x, y\}) = -1$ for all $y \in U$. Then

$$X(W) = X(U) + F(x)X^{-}(U)$$

and

$$X^{-}(W) = X^{-}(U) - F(x)X(U)$$

Proof

$$X(W) = \sum_{V \subseteq W} f(V) = \sum_{V \subseteq U} f(V) + \sum_{V \subseteq U} f(V \cup \{x\}) = X(U) + \sum_{V \subseteq U} f(V \cup \{x\}).$$

To calculate the second sum, since $\overline{f}(\{x, y\}) = -1$ for all $y \in U$, we may use Proposition 3.1 to obtain

$$\sum_{V \subseteq U} f(V \cup \{x\}) = \sum_{V \subseteq U} (-1)^{|V|} f(V) F(x)$$
$$= F(x) \sum_{V \subseteq U} (-1)^{|V|} f(V)$$
$$= F(x) X^{-}(U)$$

as required. Similarly

$$X^{-}(W) = \sum_{V \subseteq U} (-1)^{|V|} f(V) - \sum_{V \subseteq U} (-1)^{|V|} f(V \cup \{x\}) = X^{-}(U) - \sum_{V \subseteq U} (-1)^{|V|} f(V \cup \{x\})$$

Using Proposition 3.1 again, we have

$$\sum_{V \subseteq U} (-1)^{|V|} f(V \cup \{x\}) = F(x) X(U).$$

Hence $X^{-}(W) = X^{-}(U) - F(x)X(U)$.

Lemma 3.4 Suppose that f satisfies the SAC of order (n-3). Let $W \subseteq S$, and suppose that $x, y \in W$ are such that $\overline{f}(\{x, y\}) = 1$. Let $U = W \setminus \{x, y\}$, then

$$X(W) = X(U) + F(y)X^{-}(U) + F(x)X^{-}(U) + F(x)F(y)X(U)$$

and

$$X^{-}(W) = X^{-}(U) - F(y)X(U) - F(x)X(U) + F(x)F(y)X^{-}(U).$$

Proof

As before, we have

$$\begin{split} X(W) &= \sum_{V \subseteq W} f(V) = \sum_{V \subseteq (U \cup \{y\})} f(V) + \sum_{V \subseteq (U \cup \{y\})} f(V \cup \{x\}) \\ &= X(U \cup \{y\}) + \sum_{V \subseteq (U \cup \{y\})} f(V \cup \{x\}). \end{split}$$

By Propositions 3.1 and 3.2, we know that

$$f(V \cup \{x\}) = \begin{cases} (-1)^{|V|} f(V) F(x) & \text{if } y \notin V \\ (-1)^{|V|-1} f(V) F(x) & \text{if } y \in V \end{cases}.$$

So we have

$$\sum_{\substack{V \subseteq (U \cup \{y\})\\ y \notin V}} f(V \cup \{x\}) = \sum_{\substack{V \subseteq (U \cup \{y\})\\ y \notin V}} (-1)^{|V|} f(V) F(x) + \sum_{\substack{V \subseteq (U \cup \{y\})\\ y \in V}} (-1)^{|V|-1} f(V) F(x)$$
$$= \sum_{\substack{V \subseteq U\\ V \subseteq U}} (-1)^{|V|} f(V) F(x) + \sum_{\substack{V \subseteq U\\ V \subseteq U}} (-1)^{|V|} f(V \cup \{y\}) F(x)$$
$$= F(x) X^{-}(U) + F(x) \sum_{\substack{V \subseteq U\\ V \subseteq U}} (-1)^{|V|} f(V \cup \{y\})$$

Applying Proposition 3.1 again, since $\overline{f}(\{y,z\}) = -1$ for all $z \in U$, we have

$$\sum_{V \subseteq U} (-1)^{|V|} f(V \cup \{y\}) = \sum_{V \subseteq U} f(V) F(y) = F(y) X(U)$$

So

$$\sum_{V \subseteq (U \cup \{y\})} f(V \cup \{x\}) = F(x)X^{-}(U) + F(x)F(y)X(U).$$

Now by Lemma 3.3, $X(U \cup \{y\}) = X(U) + F(y)X^{-}(U)$, so putting these two together, we obtain the desired result.

We may now do exactly the same with $X^{-}(W)$ as follows.

$$\begin{aligned} X^{-}(W) &= \sum_{V \subseteq (U \cup \{y\})} (-1)^{|V|} f(V) + \sum_{V \subseteq (U \cup \{y\})} (-1)^{|V|+1} f(V \cup \{x\}) \\ &= X^{-}(U \cup \{y\}) - \sum_{V \subseteq (U \cup \{y\})} (-1)^{|V|} f(V \cup \{x\}). \end{aligned}$$

 and

$$\sum_{\substack{V \subseteq (U \cup \{y\}) \\ y \notin V}} (-1)^{|V|} f(V \cup \{x\}) = \sum_{\substack{V \subseteq (U \cup \{y\}) \\ y \notin V}} f(V)F(x) - \sum_{\substack{V \subseteq (U \cup \{y\}) \\ y \notin V}} f(V)F(x)$$
$$= F(x)X(U) - \sum_{\substack{V \subseteq U \\ V \subseteq U}} f(V \cup \{y\})F(x)$$
$$= F(x)X(U) - F(x)\sum_{\substack{V \subseteq U \\ V \subseteq U}} f(V \cup \{y\})$$

Then

$$\sum_{V \subseteq U} f(V \cup \{y\}) = \sum_{V \subseteq U} (-1)^{|V|} f(V) F(y) = F(y) X^{-}(U)$$

so

$$\sum_{V \subseteq (U \cup \{y\})} (-1)^{|V|} f(V \cup \{x\}) = F(x)X(U) - F(x)F(y)X^{-}(U).$$

Now by Lemma 3.3, $X^-(U \cup \{y\}) = X^-(U) - F(y)X(U)$, so putting these two together, we obtain the desired result.

We are now able to prove our main result in this section.

Theorem 3.5 Suppose that $W \subseteq S$ and that f belongs to $A_W(n,m)$. Let i denote the square root of -1, and let

$$G_W = f(\emptyset) \prod_{j=1}^m (1 + F(x_j)F(y_j) + i(F(x_j) + F(y_j))) \prod_{j=2m+1}^{|W|} (1 + iF(x_j)),$$

then

$$X(W) = \Re(G_W) + \Im(G_W) \text{ and}$$
$$X^-(W) = \Re(G_W) - \Im(G_W)$$

where $\Re(x)$ and $\Im(x)$ denote the real and imaginary parts of x respectively.

Proof

The proof is by induction on the size of W. Firstly, we assume that |W| = 0, so that $W = \emptyset$. Then

$$G_W = f(\emptyset)$$
 and $X(W) = f(\emptyset)$ and $X^-(W) = f(\emptyset)$.

Now suppose the result true for all W with $|W| \leq K$, and let W be such that |W| = K + 1. Choose $x \in W$. We shall split the proof into two cases. Since f satisfies the SAC of order (n-3), either $\overline{f}(\{x,y\}) = -1$ for all $y \in (W \setminus \{x\})$, or there exists a unique $y \in (W \setminus \{x\})$ for which $\overline{f}(\{x,y\}) = 1$.

Suppose that the first case holds, and let $U = W \setminus \{x\}$. Now, by Lemma 3.3,

$$X(W) = X(U) + F(x)X^{-}(U).$$

By the inductive hypothesis, since |U| = K, we have $X(U) = \Re(G_U) + \Im(G_U)$ and $X^-(U) = \Re(G_U) - \Im(G_U)$. We deduce, therefore, that

$$X(W) = \Re(G_U) + \Im(G_U) + F(x)(\Re(G_U) - \Im(G_U)) = \Re(G_W) + \Im(G_W).$$

since, in this case, $G_W = (1 + iF(x))G_U$.

We now turn to the second case, and let $U = W \setminus \{x, y\}$. By Lemma 3.4, we have

$$X(W) = X(U) + F(y)X^{-}(U) + F(x)X^{-}(U) + F(x)F(y)X(U).$$

By the inductive hypothesis, since |U| = K, we have

$$X(U) = \Re(G_U) + \Im(G_U)$$
 and $X^-(U) = \Re(G_U) - \Im(G_U)$

so we have

$$\begin{aligned} X(W) &= \Re(G_U) + \Im(G_U) + F(y)(\Re(G_U) - \Im(G_U)) \\ &+ F(x)(\Re(G_U) - \Im(G_U)) + F(x)F(y)(\Re(G_U) + \Im(G_U)) \\ &= \Re(G_W) + \Im(G_W) \end{aligned}$$

since, in this case, $G_W = (1 + F(x)F(y) + i(F(x) + F(y)))G_U$.

Proof

Suppose, without loss of generality, that $F(x_1) = -F(y_1)$. Then by Theorem 3.5, $\sum_{V \subset W} f(V) = \Re(G) + \Im(G)$, where

$$G = f(\emptyset) \prod_{j=1}^{m} (1 + F(x_j)F(y_j) + i(F(x_j) + F(y_j))) \prod_{j=2m+1}^{|W|} (1 + iF(x_j)).$$

But $1 + F(x_1)F(y_1) + i(F(x_1) + F(y_1)) = 0$, since $F(x_1) = -F(y_1)$, so G = 0. Hence $\sum_{V \subseteq W} f(V) = 0$.

Corollary 3.7 Suppose that f belongs to $C_W(n, m, r, t, q)$. Write k for |W|; then

$$X(W) = \begin{cases} f(\emptyset)(-1)^{\frac{1}{4}q}2^{\frac{1}{2}(m+k)} & q \equiv 0 \pmod{4} \\ f(\emptyset)(-1)^{\frac{1}{4}(q-1)}2^{\frac{1}{2}(m+k+1)} & q \equiv 1 \pmod{4} \\ f(\emptyset)(-1)^{\frac{1}{4}(q-2)}2^{\frac{1}{2}(m+k)} & q \equiv 2 \pmod{4} \\ 0 & q \equiv 3 \pmod{4} \end{cases}$$

Proof

Let G be defined as above; we shall examine each term in turn. Now if $1 \le j \le m$, then

$$(1 + F(x_j)F(y_j) + iF(x_j) + iF(y_j)) = \begin{cases} 2(1+i) & \text{if } F(x_j) = 1\\ 2(1-i) & \text{if } F(x_j) = -1 \end{cases}$$

and if $2m+1 \leq j \leq n$, then

$$(1 + iF(x_j)) = \begin{cases} 1 + i & \text{if } F(x_j) = 1\\ 1 - i & \text{if } F(x_j) = -1 \end{cases}$$

So

$$G = f(\emptyset)2^{r}(1+i)^{r}2^{m-r}(1-i)^{m-r}(1+i)^{t}(1-i)^{k-2m-t}$$

= $f(\emptyset)2^{m}(1+i)^{r+t}(1-i)^{k-m-r-t}$
= $f(\emptyset)2^{k-r-t}(1+i)^{2r+2t+m-k}$

Now if $0 \le b \le 3$, then

$$\Re(1+i)^{4a+b} + \Im(1+i)^{4a+b} = \begin{cases} (-4)^a & \text{if } b = 0\\ 2(-4)^a & \text{if } b = 1\\ 2(-4)^a & \text{if } b = 2\\ 0 & \text{if } b = 3 \end{cases}$$

Now 2r + 2t + m - k = q, and so

$$X(W) = \begin{cases} f(\emptyset)2^{k-r-t}(-4)^{\frac{1}{4}q} & \text{if } q \equiv 0 \pmod{4} \\ f(\emptyset)2^{k-r-t+1}(-4)^{\frac{1}{4}(q-1)} & \text{if } q \equiv 1 \pmod{4} \\ f(\emptyset)2^{k-r-t+1}(-4)^{\frac{1}{4}(q-2)} & \text{if } q \equiv 2 \pmod{4} \\ 0 & \text{if } q \equiv 3 \pmod{4} \end{cases}$$
$$= \begin{cases} f(\emptyset)(-1)^{\frac{1}{4}q}2^{\frac{1}{2}(m+k)} & \text{if } q \equiv 0 \pmod{4} \\ f(\emptyset)(-1)^{\frac{1}{4}(q-1)}2^{\frac{1}{2}(m+k+1)} & \text{if } q \equiv 1 \pmod{4} \\ f(\emptyset)(-1)^{\frac{1}{4}(q-2)}2^{\frac{1}{2}(m+k)} & \text{if } q \equiv 2 \pmod{4} \\ 0 & \text{if } q \equiv 3 \pmod{4} \end{cases}$$

4. BALANCE

We shall use the results of the preceding section to obtain necessary and sufficient conditions for a function satisfying the SAC of order (n-3) to be balanced.

Theorem 4.1 Suppose that $f \in \mathcal{B}_S$ satisfies the SAC of order (n-3). Then f is balanced if and only if either

(i) there exist x and y with $\overline{f}(\{x,y\}) = 1$ and F(x) = -F(y) or

(ii) f belongs to C(n,m,r,t,q) and $q \equiv 3 \pmod{4}$.

Proof

Since f satisfies the SAC of order (n-3), we know that either (i) holds, or there exist m, r, t and q such that f belongs to C(n, m, r, t, q). We recall that f is balanced if and only if X(S) = 0.

If (i) holds, then by Corollary 3.6, we know that X(S) = 0.

If (ii) holds, then by Corollary 3.7, we have

$$X(S) = \begin{cases} f(\emptyset)(-1)^{\frac{1}{4}q} 2^{\frac{1}{2}(m+n)} & \text{if } q \equiv 0 \pmod{4} \\ f(\emptyset)(-1)^{\frac{1}{4}(q-1)} 2^{\frac{1}{2}(m+n+1)} & \text{if } q \equiv 1 \pmod{4} \\ f(\emptyset)(-1)^{\frac{1}{4}(q-2)} 2^{\frac{1}{2}(m+n)} & \text{if } q \equiv 2 \pmod{4} \\ 0 & \text{if } q \equiv 3 \pmod{4} \end{cases}$$

So in this case, f is balanced if and only if $q \equiv 3 \pmod{4}$, since $f(\emptyset) = +1$ or -1.

5. CORRELATION IMMUNITY

We shall now obtain necessary and sufficient conditions for a function satisfying the SAC of order (n-3) to be correlation immune.

Proposition 5.1 Suppose $f \in \mathcal{B}_S$ satisfies the SAC of order (n-3). Suppose there are exactly p pairs (x_j, y_j) such that $\overline{f}(\{x_j, y_j\}) = 1$ and $F(x_j) = -F(y_j)$. Then f is exactly (p-1)th order correlation immune.

Proof

Let us write $S = \{x_1, y_1, ..., x_p, y_p, x_{2p+1}, ..., x_n\}$ where, as usual, $\overline{f}(\{x_j, y_j\}) = 1$, and $\overline{f}(\{u, v\}) = -1$ otherwise. By Corollary 3.6, X(W) = 0 whenever there exists j, $1 \le j \le p$, with $x_j, y_j \in W$. Any W with |W| > n - p must contain at least one such pair, so X(W) = 0 for any such W (including S). By Lemma 2.2.6, therefore, f is at least (p-1)th order correlation immune.

Let $U = S \setminus \{x_1, y_1, y_2, ..., y_p\}$. Then U contains no pairs (x_j, y_j) , and so $G_U \neq 0$. Let us write x for x_1 and y for y_1 , and let $U_x = U \cup \{x\}$, and $U_y = U \cup \{y\}$. Since F(x) = -F(y), we may assume without loss of generality that F(x) = 1 and F(y) = -1. Now

$$G_{U_x} = (1 + iF(x))G_U = (1 + i)G_U,$$

SO

$$X(U_x) = \Re(G_U) + \Im(G_U) + \Re(G_U) - \Im(G_U) = 2\Re(G_U).$$

On the other hand,

$$G_{U_y} = (1 + iF(y))G_U = (1 - i)G_U,$$

so

$$X(U_{\boldsymbol{y}}) = \Re(G_U) + \Im(G_U) - \Re(G_U) + \Im(G_U) = 2\Im(G_U).$$

If f is pth order correlation immune, then $X(U_x) = X(U_y) = 0$. But this forces $G_U = 0$, which is not true. Hence f is not pth order correlation immune.

We shall now prove some results on the values of X(W). In the four lemmas which follow, we assume that f belongs to $C_{W_j}(n, m_j, r_j, t_j, q_j)$ for j = 1, 2, and that $|W_j| = k_j$ for j = 1, 2. We shall calculate the relationship between $X(W_1)$ and $X(W_2)$ for various values of W_1 and W_2 . The proofs of these results are straightforward applications of Corollary 3.7, and are omitted for brevity.

Lemma 5.2 Suppose that $W_1 \subseteq S$ and that $x, y \in W_1$ are such that $\overline{f}(\{x, y\}) = +1$. Let $W_2 = W_1 \setminus \{x\}$, then $X(W_2) = \frac{1}{2}X(W_1)$.

Lemma 5.3 Suppose that $W_1 \subseteq S$ and that $x \in W_1$ is such that $\overline{f}(\{x, y\}) = -1$ for all $y \in W_1$, and that F(x) = +1. Let $W_2 = W_1 \setminus \{x\}$. Then

$$X(W_1)/X(W_2) = \begin{cases} \infty & q_1 \equiv 0 \pmod{4} \\ 2 & q_1 \equiv 1 \pmod{4} \\ 1 & q_1 \equiv 2 \pmod{4} \\ 0 & q_1 \equiv 3 \pmod{4} \end{cases}$$

Lemma 5.4 Suppose that $W_1 \subseteq S$ and that $x \in W_1$ is such that $\overline{f}(\{x, y\}) = -1$ for all $y \in W_1$, and that F(x) = -1. Let $W_2 = W_1 \setminus \{x\}$. Then

$$X(W_1)/X(W_2) = \begin{cases} 1 & q_1 \equiv 0 \pmod{4} \\ 2 & q_1 \equiv 1 \pmod{4} \\ \infty & q_1 \equiv 2 \pmod{4} \\ 0 & q_1 \equiv 3 \pmod{4} \end{cases}$$

Corollary 5.5 Suppose that $W_1 \subseteq S$ and that $x \in W_1$ is such that $\overline{f}(\{x, y\}) = -1$ for all $y \in W_1$. Let $W_2 = W_1 \setminus \{x\}$. Then

$$X(W_2) = \frac{1}{2}X(W_1)$$
 if and only if $q_1 \equiv 1 \pmod{4}$

Lemma 5.6 Suppose that $W_1 \subseteq S$ and that $x, y \in W_1$ are such that $\overline{f}(\{x, y\}) = +1$, and F(x) = F(y) = +1. Let $W_2 = W_1 \setminus \{x, y\}$, then

$$X(W_1)/X(W_2) = \begin{cases} \infty & q_1 \equiv 0 \pmod{4} \\ 2^2 & q_1 \equiv 1 \pmod{4} \\ 2 & q_1 \equiv 2 \pmod{4} \\ 0 & q_1 \equiv 3 \pmod{4} \end{cases}$$

Lemma 5.7 Suppose that $W_1 \subseteq S$ and that $x, y \in W_1$ are such that $\overline{f}(\{x, y\}) = +1$, and F(x) = F(y) = -1. Let $W_2 = W_1 \setminus \{x, y\}$, then

$$X(W_1)/X(W_2) = \begin{cases} 2 & q_1 \equiv 0 \pmod{4} \\ 2^2 & q_1 \equiv 1 \pmod{4} \\ \infty & q_1 \equiv 2 \pmod{4} \\ 0 & q_1 \equiv 3 \pmod{4} \end{cases}$$

Corollary 5.8 Suppose that $W_1 \subseteq S$ and that $x, y \in W_1$ are such that $\overline{f}(\{x, y\}) = +1$. Let $W_2 = W_1 \setminus \{x, y\}$, then

$$X(W_2) = rac{1}{2^2}X(W_1)$$
 if and only if $q_1 \equiv 1 \pmod{4}$

Proposition 5.9 Suppose that f belongs to C(n, m, r, t, q). If 2m < n, and $q \neq 1 \pmod{4}$, then f is not correlation immune.

Proof

We must find W with |W| = n - 1, and $X(W) \neq \frac{1}{2}X(S)$. Let $W = S \setminus \{x_n\}$. Since 2m < n, we may apply Corollary 5.5, with $W_1 = S$. Since $q \not\equiv 1 \pmod{4}$, we deduce that $X(W) \neq \frac{1}{2}X(S)$. So f is not correlation immune.

Proposition 5.10 Suppose that f belongs to C(2m, m, r, 0, q). If $q \not\equiv 1 \pmod{4}$, then f is exactly 1st order correlation immune.

Proof

Let us write $S = \{x_1, y_1, .., x_m, y_m\}$ where, as usual, $\overline{f}(\{x_j, y_j\}) = 1$, and $\overline{f}(\{u, v\}) = -1$ otherwise. We show first that if |W| = n - 1, then $X(W) = \frac{1}{2}X(S)$. Let W be such that |W| = n - 1. Then either $W = S \setminus \{x_j\}$ for some j or $W = S \setminus \{y_j\}$ for some j. By Lemma 5.2, therefore, with $W_1 = S$, $X(W) = \frac{1}{2}X(S)$.

So we have shown that f is at least 1st order correlation immune. We now need to find W with |W| = n - 2, and $X(W) \neq \frac{1}{2^2}X(S)$. We take $W = S \setminus \{x_1, y_1\}$. Then we may use Corollary 5.8, with $W_1 = S$. Since $q \neq 1 \pmod{4}$, $X(W) \neq \frac{1}{2^2}X(S)$. So f is not 2nd order correlation immune.

We turn now to the case where $q \equiv 1 \pmod{4}$.

Lemma 5.11 Suppose that f belongs to C(n, m, r, t, q) and that $q \equiv 1 \pmod{4}$. Then f is 1st order correlation immune.

Proof

Let us write $S = \{x_1, y_1, ..., x_m, y_m, x_{2m+1}, ..., x_n\}$ where, as usual, $\overline{f}(\{x_j, y_j\}) = 1$, and $\overline{f}(\{u, v\}) = -1$ otherwise. We must show that $X(W) = \frac{1}{2}X(S)$ for any W with |W| = n - 1. Choose any such W. Then we have the following possibilities for W:

$$W = S \setminus \{x_j\} \text{ for some } j, 1 \le j \le m \text{ or}$$
$$W = S \setminus \{y_j\} \text{ for some } j, 1 \le j \le m \text{ or}$$
$$W = S \setminus \{x_j\} \text{ for some } j, 2m + 1 \le j \le n$$

In either of the first two cases, we may apply Lemma 5.2, to obtain $X(W) = \frac{1}{2}X(S)$, while in the third case we may apply Corollary 5.5 to obtain $X(W) = \frac{1}{2}X(S)$. Hence f is 1st order correlation immune.

Lemma 5.12 Suppose that f belongs to C(n, m, r, t, q) and that $q \equiv 1 \pmod{4}$. Then f is 2nd order correlation immune if and only if $2m \geq n-1$.

Proof

Let us write $S = \{x_1, y_1, ..., x_m, y_m, x_{2m+1}, ..., x_n\}$ where, as usual, $\overline{f}(\{x_j, y_j\}) = 1$, and $\overline{f}(\{u, v\}) = -1$ otherwise. We already know that f is 1st order correlation immune. We must show that $X(W) = \frac{1}{2^2}X(S)$ for any W with |W| = n - 2. Choose any such W. Then we have the following possibilities for W:

$$W = S \setminus \{x_j, y_k\}, j \neq k, 1 \leq j, k \leq m$$

$$W = S \setminus \{x_j, x_k\}, j \neq k, 1 \leq j, k \leq m$$

$$W = S \setminus \{y_j, y_k\}, j \neq k, 1 \leq j, k \leq m$$

$$W = S \setminus \{x_j, y_j\}, 1 \leq j \leq m$$

$$W = S \setminus \{x_j, x_k\}, 1 \leq j \leq m, 2m + 1 \leq k \leq n$$

$$W = S \setminus \{y_j, x_k\}, 1 \leq j \leq m, 2m + 1 \leq k \leq n$$

$$W = S \setminus \{y_j, x_k\}, 1 \leq j \leq m, 2m + 1 \leq k \leq n$$

$$W = S \setminus \{x_j, x_k\}, j \neq k, 2m + 1 \leq j, k \leq n$$

In the first case, we may first apply Lemma 5.2 with $W_1 = S$ and $W_2 = S \setminus \{x_j\}$, and then apply Lemma 5.2 again with $W_1 = S \setminus \{x_j\}$ and $W_2 = W$ to obtain $X(W) = \frac{1}{2}X(S \setminus \{x_j\}) = \frac{1}{2^2}X(S)$ as required. This may also be done in the second and third cases. In the fourth case, we may apply Corollary 5.8, with $W_1 = S$ to obtain $X(W) = \frac{1}{2^2}X(S)$, as required. In the fifth and sixth cases, we may proceed in a similar manner as in the first case, applying Lemma 5.2, and then Corollary 5.5 to obtain the result (noting that q is unchanged after applying Lemma 5.2). When we come to the seventh case, however, we see that if we apply Corollary 5.5 with $W_1 = S$, and $W_2 = S \setminus \{x_j\}$, we obtain $X(W_2) = \frac{1}{2}X(W_1)$, but when we come to apply Corollary 5.5 again with $W_1 = S \setminus \{x_j\}$ and $W_2 = W$, we now have $q_1 \equiv 0$ (mod 4), or $q_1 \equiv 2 \pmod{4}$, according as $j \leq 2m_1 + t_1$ or $j > 2m_1 + t_1$, and so $X(W) \neq \frac{1}{2^2}X(S)$ in this case. This case can only occur when 2m + 1 < n, so f is 2nd order correlation immune if and only if $2m \geq n-1$.

Lemma 5.13 Suppose that f belongs to C(n, m, r, t, q) and that $q \equiv 1 \pmod{4}$. Then f is 3rd order correlation immune if and only if 2m = n.

Proof

Let us write $S = \{x_1, y_1, .., x_m, y_m, x_{2m+1}, .., x_n\}$ where, as usual, $\overline{f}(\{x_j, y_j\}) = 1$, and $\overline{f}(\{u, v\}) = -1$ otherwise. Suppose first that 2m = n. We therefore know that f is 2nd order correlation immune, since $2m \ge n-1$. We must show that $X(W) = \frac{1}{2^3}X(S)$ for any W with |W| = n-3. Let W be such that |W| = n-3. If $W = S \setminus \{x_j, x_k, x_l\}$, with j, k and l all different, and $1 \le j, k, l \le m$, then we may apply Lemma 5.2 three times to obtain the result. The same method will also work in the cases $W = S \setminus \{x_j, x_k, y_l\}, W = S \setminus \{x_j, y_k, y_l\}$ and $W = S \setminus \{y_j, y_k, y_l\}$. The cases $W = S \setminus \{x_j, y_j, x_k\}$ and $W = S \setminus \{x_j, y_j, y_k\}$, where $1 \le j, k \le m$ may each be dealt with using first Corollary 5.8, and then Lemma 5.2. This means that when 2m = n, f is 3rd order correlation immune. When, however, 2m < n, we must consider the case $W = S \setminus \{x_1, y_1, x_n\}$. We apply Corollary 5.8 with $W_1 = S$, and $W_2 = S \setminus \{x_1, y_1\}$, and then apply Corollary 5.5 with $W_1 = S \setminus \{x_1, y_1\}$, and $W_2 = W$. But this time either $q_1 \equiv 0 \pmod{4}$ or $q_1 \equiv 2 \pmod{4}$ according as r > 0 or r = 0. So in this case, f is not 3rd order correlation immune.

Lemma 5.14 Suppose that f belongs to C(n, m, r, t, q) and that $q \equiv 1 \pmod{4}$. Then f is not 4th order correlation immune.

Proof

We shall produce W with |W| = n - 4 but $X(W) \neq \frac{1}{24}X(S)$.

In order for f to be fourth order correlation immune, it must certainly be third order correlation immune. So, by Lemma 5.13, we must have 2m = n. Let us write $S = \{x_1, y_1, ..., x_m, y_m\}$ where, as usual, $\overline{f}(\{x_j, y_j\}) = 1$, and $\overline{f}(\{u, v\}) = -1$ otherwise. We take $W = S \setminus \{x_1, y_1, x_2, y_2\}$. (Note that this is possible since n is even and at least 3). Let us also denote $S \setminus \{x_1, y_1\}$ by U. Then by Corollary 5.8, we see that $X(U) = \frac{1}{2^2}X(S)$, since $q_1 \equiv 1 \pmod{4}$. We now apply Corollary 5.8 with $W_1 = U$. This time, however, we have $q_1 \equiv 0 \pmod{4}$ or $q_1 \equiv 2 \pmod{4}$, (according as $r \geq 1$ or not) so $X(W) \neq \frac{1}{2^2}X(U)$, and therefore $X(W) \neq \frac{1}{2^4}X(S)$. Hence f is not 4th order correlation immune.

We thus have, combining the preceding four lemmas.

Corollary 5.15 Suppose that f belongs to C(n, m, r, t, q) and that $q \equiv 1 \pmod{4}$. Then

- (i) if 2m < n-1, then f is exactly 1st order correlation immune and
- (ii) if 2m = n 1, then f is exactly 2nd order correlation immune and

(iii) if 2m = n, then f is exactly 3rd order correlation immune.

Combining all the results of this section, we have the following theorems and corollaries.

Theorem 5.16 If $f \in \mathcal{B}_S$ satisfies the SAC of order (n-3), then f is not correlation immune if and only if either

- (i) there is exactly one pair (x, y) with $\overline{f}(\{x, y\}) = 1$ and F(x) = -F(y) or
- (ii) f belongs to C(n, m, r, t, q) and 2m < n and $q \not\equiv 1 \pmod{4}$.

Theorem 5.17 If $f \in \mathcal{B}_S$ satisfies the SAC of order (n-3), then f is exactly 1st order correlation immune if and only if one of the following holds

- (i) there are exactly two pairs (x, y) with $\overline{f}(\{x, y\}) = 1$ and F(x) = -F(y) or
- (ii) f belongs to C(n, m, r, t, q) and 2m = n and $q \not\equiv 1 \pmod{4}$ or
- (iii) f belongs to C(n, m, r, t, q) and 2m < n-1 and $q \equiv 1 \pmod{4}$.

Theorem 5.18 If $f \in \mathcal{B}_S$ satisfies the SAC of order (n-3), then f is exactly 2nd order correlation immune if and only if either

(i) there are exactly three pairs (x, y) with $\overline{f}(\{x, y\}) = 1$ and F(x) = -F(y) or

(ii) f belongs to C(n, m, r, t, q) and 2m = n - 1 and $q \equiv 1 \pmod{4}$.

Theorem 5.19 If $f \in \mathcal{B}_S$ satisfies the SAC of order (n-3), then f is exactly 3rd order correlation immune if and only if either

(i) there are exactly four pairs (x, y) with $\overline{f}(\{x, y\}) = 1$ and F(x) = -F(y) or

(ii) f belongs to C(n, m, r, t, q) and 2m = n and $q \equiv 1 \pmod{4}$.

Theorem 5.20 If $f \in \mathcal{B}_S$ satisfies the SAC of order (n-3), then f is pth order correlation immune (p > 3) if and only if there are exactly (p+1) pairs (x, y) with $\overline{f}(\{x, y\}) = 1$ and F(x) = -F(y).

6. BALANCE AND CORRELATION IMMUNITY

Combining the results in sections 4 and 5, we have the following result.

Theorem 6.1 If $f \in \mathcal{B}_S$ satisfies the SAC of order (n-3), then f is both balanced and correlation immune if and only if either

- (i) there exist at least two pairs (x, y) such that $\overline{f}(\{x, y\}) = 1$ and F(x) = -F(y) or
- (ii) f belongs to C(n, m, r, t, q) and n = 2m and $q \equiv 3 \pmod{4}$.

7. REFERENCES

- Forré, R., "The Strict Avalanche Criterion : Spectral Properties of Boolean Functions and an Extended Definition", Abstracts CRYPTO88, 1988
- [2] Lloyd, S.A, "Balance, uncorrelatedness and the Strict Avalanche Criterion", Hewlett-Packard Research Laboratories, Bristol, Technical Memo no. HPL-ISC-TM-89-012, 1989 (also submitted to Discrete Applied Mathematics)
- [3] Lloyd, S.A, "Characterising and counting functions satisfying the Strict Avalanche Criterion of order (n-3)", to appear in *Proceedings of the Second IMA Conference on Cryptography and Coding*, 1989
- [4] Siegenthaler, T., "Decrypting a class of stream ciphers using ciphertext only", IEEE Transactions on Computers, vol. C-34, (1985), pp.81-85
- [5] Siegenthaler, T., "Correlation immunity of nonlinear combining functions for cryptographic applications", *IEEE Transactions on Information Theory*, vol. IT-30, (1984), pp776-780
- [6] Xiao, G-Z, and Massey, J.L., "A Spectral Characterization of Correlation-Immune Combining Functions", *IEEE Transactions on Information Theory*, Vol. 34, No. 3 (1988), pp.569-571

[7] Webster, A.F. and Tavares, S.E., "On the design of S-boxes", Advances in Cryptology, Proceedings CRYPT085, Springer Verlag, Heidelberg, 1986, pp. 523-534