Crowds, Anonymous Web Transactions

Avi Rubin
AT&T Labs Research, Florham Park, NJ

Abstract

In this talk we introduce a system called Crowds for protecting users' anonymity on the world-wide-web. Crowds, named for the notion of "blending into a crowd", operates by grouping users into a large and geographically diverse group (crowd) that collectively issues requests on behalf of its members. Web servers are unable to learn the true source of a request because it is equally likely to have originated from any member of the crowd, and indeed collaborating crowd members cannot distinguish the originator of a request from a member who is merely forwarding the request on behalf of another. We describe the design, implementation, security, performance, and scalability of our system. Our security analysis introduces degrees of anonymity as an important tool for describing and proving anonymity properties. (Joint work with Mike Reiter.)

See `http://www.cs.nyu.edu/~rubin/crowds.ps.gz` for a copy of the technical report or visit the Crowds home page at `http://www.research.att.com/projects/crowds/`.

DESING OF SUBSTITUTION BLOCKS SATISFYING STRICT AVALANCHE CRITERION

MARTIN STANEK AND DANIEL OLEJÁR

Comenius University
Bratislava
{ stanek, olejar }@dcs.fmph.uniba.sk

ABSTRACT. The strength of various cryptosystems depends substantially on the properties of elements they consist of. Substitution blocks are often used as nonlinear transformations in both stream ciphers, block ciphers and one-way hash functions. One of the most important properties of substitution blocks is their ability to transform small changes of input into large output changes, which is measured by the Strict Avalanche Criterion, the SAC. We present a flexible method for construction of large substitution blocks from smaller ones, preserving the SAC-property and discuss the other cryptographically important properties of resulting substitution blocks as regularity, nonlinearity and the robustness against the linear cryptanalysis.

1. Introduction

The strength of various cryptosystems depends substantially on the properties of elements they consist of. Substitution blocks are often used as nonlinear transformations in stream ciphers, block ciphers and one-way hash functions. There are sveral methods how to construct cryptographically strong S-boxes. Small, regular $(n \times m)$ S-boxes $(n \geq m)$ with $n < 6$ can be generated by exhaustive search. Larger, SAC satisfying S-boxes can be constructed by means of methods presented in [SZZ93a] or by expanding smaller S-boxes according to [KMI90]. Another approach, introduced in [P91, N93] yields cryptographically strong S-boxes which do not satisfy SAC. But the resulting S-boxes can be modified by transforming their inputs by a suitable linear transformation [SZZ93b] into SAC satisfying S-boxes. Both previous constructions yield large but complex cryptographically strong S-boxes. Youssef and Tavares [YT96] presented simplier S-blocks, so called Substitution-Permutation Networks (SPN) with cryptographically good properties. The SPN's constructed by their method have very good avalanche characteristics, but do not satisfy the ideal – SAC.

Using the idea of constructing large S-blocks from smaller ones by proper choise of basic S-boxes and tranforming their outputs by some controlled tranformations, we introduce a method of design of regular substitution blocks with high nonlinearity, satisfying SAC and discuss their resistance against linear cryptanalysis, too.

Key words and phrases. SAC, Strict Avalanche Criterion, S-boxes.

2. Preliminaries

The vector space of dimension n over $GF(2)$ is denoted by V_n. The elements from V_n will be denoted by small letters of greek aplhabet $(\alpha, \beta, \dots)$. Since there is a very natural correspondence between vectors from V_n and numbers from the set $\{0, 1, \dots, 2^n - 1\}$, we will often treat the vectors from V_n as the integers. To avoid the missunderstandings, we denote by the symbol $\sigma(n, j)$ the vector $(\sigma_{j,1}, \dots, \sigma_{j,n}) \in V_n$, such that

$$j = \sum_{k=1}^{n} 2^{n-k} \sigma_{j,k}$$

The n-ary Boolean function f is mapping from V_n to $GF(2)$. The *truth table* of a Boolean function f is defined as the vector $tt(f) = (f(0), f(1), \dots, f(2^n - 1))$. Let $\alpha = (a_1, a_2, \dots, a_m)$ and $\beta = (b_1, b_2, \dots, b_m)$ be two binary vectors of length m. The operation $\alpha \oplus \beta$ denotes bitwise XOR of α and β. The symbol $\oplus$ itself denotes the sum modulo 2. The *Hamming weight* of a vector α, denoted by $wt(\alpha)$, is the number of ones in α. The n-ary Boolean function f is *balanced*, if $wt(tt(f)) = 2^{n-1}$; that is, the truth table of f contains the same number of ones and zeroes.

Let α, β be two vectors in V_m. Then $d(\alpha, \beta) = wt(\alpha, \beta)$ is called the *Hamming distance* between α and β. Analogically, distance between functions f, g (both depending on the same number of variables) is defined as $d(f, g) = d(tt(f), tt(g))$.

The Boolean function $h : V_n \to \{0, 1\}$ of the form $h(x) = a_0 \oplus a_1 x_1 \oplus a_2 x_2 \oplus \cdots \oplus a_n x_n$, where $x = (x_1, \dots, x_n)$ and $a_0, \dots, a_n \in \{0, 1\}$, is called *affine*. In particular, h is *linear*, if $a_0 = 0$.

Since linear cryptosystems can be easily broken, one of the most important cryptographic properties of Boolean functions is *nonlinearity*. The nonlinearity of a Boolean function is defined as the (minimal) distance of the function and the set of all affine functions (we denote the nonlinearity of f by $N(f)$):

$$N(f) = \min_h \{d(f, h) \mid h(x) = a_0 \oplus \bigoplus_{i=1}^{n} a_i x_i, \quad a_0, \dots, a_n \in \{0, 1\}\}.$$

The other important criterion (Strict Avalanche Criterion) was introduced by Webster and Tavares in [WT86]:

Definition 2.1. The Boolean function $f : V_n \to \{0, 1\}$ is said to satisfy SAC, if the function $f(x) \oplus f(x \oplus \alpha)$ is balanced, for all $\alpha \in V_n$ such that $wt(\alpha) = 1$.

The SAC is a global criterion. To describe "local" avalanche properties of Boolean functions, we said, the function $f : V_n \to \{0, 1\}$ satisfies *propagation criterion* with respect to a vector $\gamma \in V_n$, if the function $f(x) \oplus f(x \oplus \gamma)$ is balanced.

The substitution box S $n \times m$ $(n \geq m)$ is an m-tuple of Boolean functions $(f_1, \dots, f_m)$, $f_i : V_n \to \{0, 1\}$ $\forall\, i = 1, \dots, m$. S-box $S = (f_1, \dots, f_m)$ satisfies the SAC, if every f_i $(i = 1, \dots, m)$ satisfies SAC. The S-box S $(S : V_n \to V_m)$, is said to be *regular*, if

$$|\{x \in V_n \mid S(x) = y\}| = 2^{n-m}$$

for all $y \in V_m$. The following lemma was proved in [SZZ94] and will be an useful tool in the next sections.

Lemma 2.2. *Let $(f_1, \ldots, f_m)$, where each f_i is Boolean function on V_n, is mapping from V_n to V_m. This mapping is regular if and only if all nonzero linear combinations of $f_1, \ldots, f_m$ are balanced.*

Let $f_0, f_1, \ldots, f_{2^k-1}$ be Boolean functions on V_n. Then *concatenation* of these functions is a function $g : V_{n+k} \to \{0,1\}$ defined in the following way:

$$g(x,y) = \bigoplus_{\sigma \in V_k} x^\sigma f_\sigma(y),$$

where $x = (x_1, \ldots, x_k)$, $y = (y_1, \ldots, y_n)$, $x^\sigma = x_1^{\sigma_1} x_2^{\sigma_2} \cdot \ldots \cdot x_k^{\sigma_k}$ $(\sigma = (\sigma_1, \ldots, \sigma_k))$ and

$$x_i^{\sigma_i} = \begin{cases} \overline{x_i}, \text{ if } \sigma_i = 0; \\ x_i, \text{ if } \sigma_i = 1. \end{cases}$$

The truth table of the function g can be obtained by concatenation of truth tables of functions $f_0, f_1, \ldots, f_{2^k-1}$: $tt(g) = (tt(f_0), tt(f_1), \ldots, tt(f_{2^k-1}))$.

The following lemma provide us with a useful lower bound on the nonlinearity of concatenation of Boolean functions.

Lemma 2.3. *Let $f_0, f_1, \ldots, f_{2^k-1}$ be Boolean functions defined on V_n. Let $N^* = \min\{N(f_j) \mid j = 0, 1, \ldots, 2^k - 1\}$. Let g be concatenation of these functions. Then $N(g) \geq 2^k N^*$.*

Proof. (by contradiction)

Let $N(g) < 2^k N^*$. Then there exists an affine function $h(x,y) = a_1 x_1 \oplus \cdots \oplus a_k x_k \oplus b_1 y_1 \oplus \cdots \oplus b_n y_n \oplus c$, where $x = (x_1, \ldots, x_k)$, $y = (y_1, \ldots, y_n)$, such that $d(g,h) < 2^k N^*$. Let us split function h into 2^k parts $h_0, h_1, \ldots, h_{2^k-1} : V_n \to \{0,1\}$, such that

$$h_j(y) = \bigoplus_{i=1}^{n} b_i y_i \oplus d_j,$$

where $d_j = c \oplus a_1 \sigma_{j,1} \oplus a_2 \sigma_{j,2} \oplus \cdots \oplus a_k \sigma_{j,k}$ $((\sigma_{j,1}, \ldots, \sigma_{j,k}) = \sigma(j,k))$. Then there must exist an integer l $(0 \leq l \leq 2^k - 1)$, such that $d(f_l, h_l) < N^*$. Since h_l is affine function, we have $N(f_l) < N^*$ – a contradiction. □

Linear cryptanalysis ([M93]) is a powerful method to attack various kinds of cryptosystems. It is based on approximation of outputs of S-boxes by linear (or affine) functions. To achieve immunity against linear cryptanalysis it is sufficient to use (construct) S-boxes with high nonlinearity of every nonzero linear combination of their functions – see [SZZ93a].

3. General construction

The main idea of our construction is to split the input vector into two parts – one of them is used as input variables of an S-box S; while the other controls a transformation T, modifying the output of the S-box. The elementary building block (EBB) is depicted in figure 1.

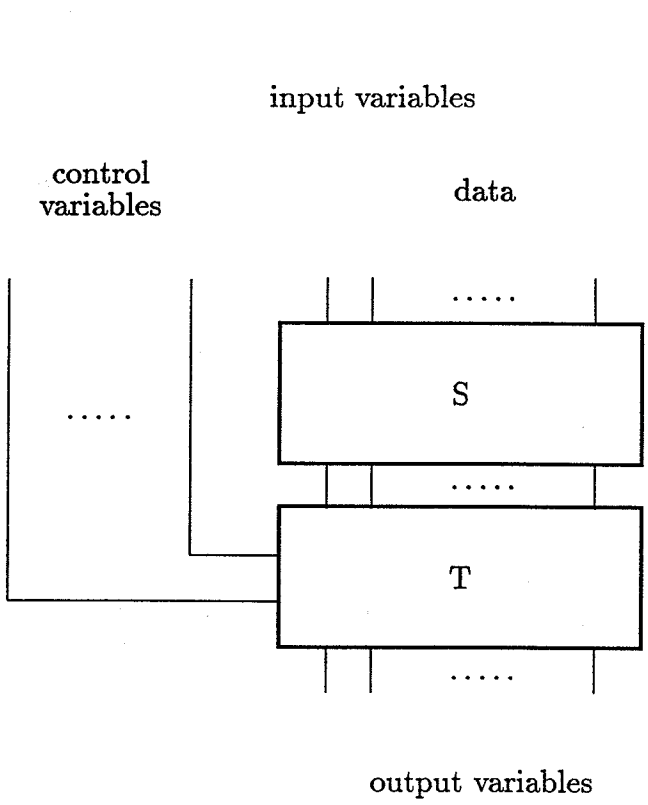

Figure 1: Elementary building block

There are many ways how to interconnect these elementary blocks into a network. Since the cryptographic properties of resulting network are substantially influenced by the depicted transformation, we concentrate our attention on transformations, the use of which results in cryptographically strong networks, namely permutations and inversions (inversion of a Boolean function $f(x)$ is the Boolean function $f(x) \oplus 1$).

4. Permutations

The input variables of the EBB B are $(x_1, \ldots, x_k)$ – control variables and $(y_1, \ldots, y_n)$ – data. B (cf. figure 1) consists of regular $n \times m$ S-box S satisfying SAC. Let us suppose $S = (g_1, \ldots, g_m)$ such that $g_i : V_n \to \{0,1\}$ for $i = 1, \ldots, m$. To abbreviate the notation we denote the input vectors $(x_1, \ldots, x_k)$ and $(y_1, \ldots, y_n)$ by the symbols x, y respectively. The output of S-box S, a vector $(g_1(y), \ldots, g_m(y))$, is permuted by a permutation P. The permutation P depends on control variables x (and therefore will be denoted by P_x). P_x is an m-bit permutation, transforming the vector $(g_1(y), \ldots, g_m(y))$. The m-ary output of P_x is a vector $(G_1(x,y), \ldots, G_m(x,y))$ where $G_1, \ldots, G_m$ are mappings from V_{n+k} to $\{0,1\}$:

$$(G_1(x,y), \ldots, G_m(x,y)) = P_x(g_1(y), \ldots, g_m(y)).$$

Let permutation P be fixed, that is – P does not depend on $x_1, \ldots, x_k$. Without loss of generality we can assume that P is identity. Then $G_i(x,y) = g_i(y)$, for all $i = 1, \ldots, m$. Since g_i satisfies SAC, $G_i(x,y)$ satisfies propagation criterion with

respect to all vectors $\gamma = (0, \alpha)$, where $0 \in V_k,\ \alpha \in V_n \wedge wt(\alpha) = 1$. A problem appears for vectors $\gamma = (\beta, 0)$, where $0 \in V_n,\ \beta \in V_k \wedge wt(\beta) = 1$:

$$G_i(x, y) \oplus G_i(x \oplus \beta, y \oplus 0) = g_i(y) \oplus g_i(y) = 0,$$

what is the worst case from the "avalanche" point of view. Let us look at the following example:

Example 4.1. Let g_1, g_2, g_3 are functions of 3×3 regular S-box satisfying SAC. Let the truth table of a function G be obtained by concatenation of these functions:

$$G = \langle g_1, g_2, g_3, g_1, g_2, g_3, g_1, g_2 \rangle.$$

Function G depends on 6 variables. It is not hard to verify, that G satisfies SAC, too. There are 6 vectors of Hamming weight 1. In 3 cases (for $\gamma \in \{(000001), (000010), (000100)\}$) G satisfies the Strict Avalanche Criterion with respect to vector γ (this follows directly from the fact that S-box satisfies SAC). The remaining vectors are analyzed bellow:

a) Let $\gamma = (001000)$. Then

$$G(x) \oplus G(x \oplus (001000)) =$$
$$\langle g_1, g_2, g_3, g_1, g_2, g_3, g_1, g_2 \rangle \oplus \langle g_2, g_1, g_1, g_3, g_3, g_2, g_2, g_1 \rangle =$$
$$\langle g_1 \oplus g_2, g_2 \oplus g_1, g_3 \oplus g_1, g_1 \oplus g_3, g_2 \oplus g_3, g_3 \oplus g_2, g_1 \oplus g_2, g_2 \oplus g_1 \rangle$$

b) Let $\gamma = (010000)$. Then

$$G(x) \oplus G(x \oplus (010000)) =$$
$$\langle g_1 \oplus g_3, g_2 \oplus g_1, g_3 \oplus g_1, g_1 \oplus g_2, g_2 \oplus g_1, g_3 \oplus g_2, g_1 \oplus g_2, g_2 \oplus g_3 \rangle$$

c) Let $\gamma = (100000)$. Then

$$G(x) \oplus G(x \oplus (100000)) =$$
$$\langle g_1 \oplus g_2, g_2 \oplus g_3, g_3 \oplus g_1, g_1 \oplus g_2, g_2 \oplus g_1, g_3 \oplus g_2, g_1 \oplus g_3, g_2 \oplus g_1 \rangle$$

Since the S-box is regular, every function $g_i \oplus g_j$ $(\forall\ 1 \leq i < j \leq 3)$ is balanced (Lemma 2.2) and therefore the functions $G(x) \oplus G(x \oplus \gamma)$ are balanced, too.

Our elementary building block (with controlled permutation – cf. figure 1) exactly concatenation of functions of S-box S. We are interesting in such concatenations, which preserve SAC (just like in the example above).

Definition 4.2. Let $\pi : V_k \to \{1, 2, \ldots, m\}$. This mapping is **SAC-preserving**, if and only if $\forall\ \alpha \in V_k\ \forall\ x \in V_k$:

$$wt(\alpha) = 1 \implies \pi(x) \neq \pi(x \oplus \alpha).$$

Following lemma extends previous example.

Lemma 4.3. *Let $S = (g_1, \ldots, g_m)$ be a regular $n \times m$ S-box satisfying SAC (that is, $\forall\ i = 1, \ldots, m : g_i$ satisfies SAC). Let $\pi : V_k \to \{1, 2, \ldots, m\}$ be SAC-preserving mapping. Then the function G defined by*

$$G(x, y) = \bigoplus_{\sigma \in V_k} x^{\sigma}\ g_{\pi(\sigma)}(y)$$

satisfies SAC.

Proof. Let us consider two cases:

a) Let $\gamma = (0, \alpha), \quad \alpha \in V_n \wedge wt(\alpha) = 1,\ 0 \in V_k$. Then

$$\begin{aligned} G(x, y) \oplus G((x, y) \oplus \gamma) &= G(x, y) \oplus G(x \oplus 0, y \oplus \alpha) = \\ &= \bigoplus_{\sigma \in V_k} (x^{\sigma}\ g_{\pi(\sigma)}(y) \oplus x^{\sigma}\ g_{\pi(\sigma)}(y \oplus \alpha)) = \\ &= \bigoplus_{\sigma \in V_k} x^{\sigma}(g_{\pi(\sigma)}(y) \oplus\ g_{\pi(\sigma)}(y \oplus \alpha)). \end{aligned} \tag{4.1}$$

Since g_i $(i = 1, \ldots, m)$ satisfies SAC, Boolean function (4.1) is concatenation of 2^k balanced Boolean functions. Therefore G satisfies propagation criterion with respect to all above vectors γ.

b) Let $\gamma = (\beta, 0), \quad \beta \in V_k \wedge wt(\beta) = 1,\ 0 \in V_n$. Then

$$\begin{aligned} G(x, y) \oplus G((x, y) \oplus \gamma) &= G(x, y) \oplus G(x \oplus \beta, y \oplus 0) = \\ &= \bigoplus_{\sigma \in V_k} (x^{\sigma}\ g_{\pi(\sigma)}(y) \oplus (x \oplus \beta)^{\sigma}\ g_{\pi(\sigma)}(y)) = \\ &= \bigoplus_{\sigma \in V_k} (x^{\sigma}\ g_{\pi(\sigma)}(y) \oplus x^{\sigma}\ g_{\pi(\sigma \oplus \beta)}(y)) = \\ &= \bigoplus_{\sigma \in V_k} x^{\sigma}(g_{\pi(\sigma)}(y) \oplus\ g_{\pi(\sigma \oplus \beta)}(y)). \end{aligned} \tag{4.2}$$

We already know that S-box S is regular. Hence, all functions $g_i \oplus g_j$ $(\forall\ 1 \leq i < j \leq m)$ are balanced. Since mapping π is SAC-preserving and $wt(\beta) = 1$, $\pi(\sigma) \neq \pi(\sigma \oplus \beta)$. Therefore Boolean function (4.2) is (again) concatenation of 2^k balanced Boolean functions and thus G satisfies SAC. $\square$

Recall our goal: to create controlled permutation P such that resulting substitution block satisfies SAC. Lemma 4.3 did the first step toward the solution. Now we need to transform input functions $g_1, \ldots, g_m$ into output functions $G_1, \ldots, G_m$ in such a way, that we save the regularity and avalanche property of the S-box. So the same SAC-preserving mapping can be used only once. But how to create SAC-preserving functions? The following lemma provide us the instructions.

Lemma 4.4. *Let $\pi : V_k \to \{1, \ldots, m\}$ be SAC-preserving mapping. Let π' be a permutation on the set $\{1, \ldots, m\}$ defined in following way:*

$$\pi'(x) = (\pi(x)\ mod\ m) + 1.$$

Then π' is SAC-preserving mapping.

Proof. Let us consider mapping $h : \{1,\ldots,m\} \to \{1,\ldots,m\}$; $h(x) = (x \bmod m)+1$. Since h is a permutation on the set $\{1,\ldots,m\}$, for each $x, \alpha \in V_k$, such that $\pi(x) \neq \pi(x \oplus \alpha)$, $h(\pi(x)) \neq h(\pi(x \oplus \alpha))$, too. Since $\pi'(x) = h(\pi(x))$, the lemma follows. □

Some serious questions remain open, namely whether SAC-preserving mappings exist for arbitrary number of variables, whether it is possible to generate them efficiently and finally, how many of SAC-preserving functions can be found.

Definition 4.5. Let $\lambda_{k,m} : V_k \to \{1,2,\ldots,m\}$ is defined by

$$\lambda_{k,m}(x) = (x \bmod m) + 1,$$

where $+$ denotes arithmetical addition. Then $\lambda_{k,m}$ is called **natural cyclical mapping.**

Lemma 4.6. *Let $m > 2$ and $m \nmid 2^{k-1}$ (m is not a divisor of 2^{k-1}) and let $\lambda_{k,m}$ be natural cyclical mapping. Then $\lambda_{k,m}$ is SAC-preserving mapping.*

Proof. (by contradiction)
Let $\alpha = (0\ldots010\ldots0) \in V_k$ (1 is i-th element from the left; $i \in \{1,\ldots,k\}$) is vector for which $\exists\, x \in V_k : \lambda_{k,m}(x) = \lambda_{k,m}(x \oplus \alpha)$. Then (since $\lambda_{k,m}$ is cyclical) the following equalities hold:

$$\begin{aligned} \lambda_{k,m}(0) &= \lambda_{k,m}(\alpha); \\ 1 &= (2^{k-i} \bmod m) + 1; \\ 2^{k-i} &\equiv 0 \pmod m. \end{aligned}$$

Hence, $m \mid 2^{k-i}$. But i is an integer from the set $\{1,2,\ldots,k\}$. The obtained contradiction proves the proposition of our Lemma. □

Corollary 4.7. *For $k > 2$ there is at least*

$$\sum_{\substack{2<r\leq m \\ r \nmid 2^{k-1}}} \binom{m}{r} \cdot r!$$

SAC-preserving mappings from V_k to $\{1,2,\ldots,m\}$.

Proof. Let $\lambda_{k,r}$ be a natural cyclical mapping and $r \nmid 2^{k-1}$. Then, by Lemma 4.6, $\lambda_{k,r}$ is also SAC-preserving mapping. We can define following mapping $\lambda_{k,r}^{(i_1,\ldots,i_r)} : V_k \to \{i_1,\ldots,i_r\}$, by:

$$\lambda_{k,r}^{(i_1,\ldots,i_r)}(x) = i_{(x \bmod r)+1},$$

where $i_1,\ldots,i_r$ are distinct elements from the set $\{1,\ldots,m\}$. Clearly, $\lambda_{k,r}^{(i_1,\ldots,i_r)}$ is also SAC-preserving mapping. We can choose $i_1,\ldots,i_r$ in $\binom{m}{r}$ different ways and their ordering in $r!$ ways. Thus, the corollary follows. □

Let us define the binary operation $\boxplus$ over the set $\{1,\ldots,m\}$. Let $a,b \in \{1,\ldots,m\}$. Then $a \boxplus b = (a+b-1) \mod m+1$.

Now, we shall summarize our construction.

The permutation block is based on the SAC-preserving function $\pi : V_k \to \{1,\ldots,m\}$:

$$\begin{aligned} G_1(x,y) &= g_{\pi(x)\boxplus 1}(y) \\ G_2(x,y) &= g_{\pi(x)\boxplus 2}(y) \\ &\ldots \\ G_m(x,y) &= g_{\pi(x)\boxplus m}(y), \end{aligned} \tag{4.3}$$

Remark 4.8 It is possible to choose various different ways, how to implement permutation block to achieve the regularity of the resulting substitution block, too.

SAC and regularity are only two from the set of very important cryptographic criteria. Impacts of described construction on nonlinearity and resistance against linear cryptanalysis are summarized in the following theorem.

Theorem 4.9. *Let S be regular $n \times m$, SAC-satisfying S-box. Let $\pi : V_k \to \{1,\ldots,m\}$ be SAC-preserving mapping. Let us consider substitution block constructed according to (4.3). Let N_g denote minimal nonlinearity from the set of all g_i's $(1 \le i \le m)$:*

$$N_g = min_{1\le i\le m}\{N(g_i)\}.$$

Let N_g^* be minimal nonlinearity of nonzero linear combinations of $g_1, g_2, \ldots, g_m$:

$$N_g^* = min\{N(\bigoplus_{i=1}^{m} a_i g_i) \mid (a_1,\ldots,a_m) \in V_m \setminus \{0\}\}.$$

Then

(i) $\forall\, i = 1,\ldots,m:$ *G_i satisfies SAC;*
(ii) *constructed substitution block is regular;*
(iii) $\forall\, i = 1,\ldots,m:\ N(G_i) \ge 2^k N_g$;
(iv) $\forall\, (a_1,\ldots,a_m) \in V_m \setminus \{0\}:\ N(\bigoplus_{i=1}^{m} a_i G_i) \ge 2^k N_g^*$.

Proof.

(i) Each Boolean function $G_i(x,y) = g_{\pi(x)\boxplus i}(y)$ (for $i = 1,\ldots,m$) can be written in the form $G_i(x,y) = \bigoplus_{\sigma\in V_k} x^\sigma g_{\pi(\sigma)\boxplus i}(y)$. But π is SAC-preserving mapping. Then, by applying $(i-1)$ times Lemma 4.4, the mapping $\pi(\sigma) \boxplus i$ is also SAC-preserving. Therefore, by Lemma 4.3, G_i satisfies SAC.

(ii) According to Lemma 2.2 it is sufficient to prove the balancedness of each nonzero linear combination of G's. Let $(a_1,\ldots,a_m) \in V_m \setminus \{0\}$. Then

$$\begin{aligned} \bigoplus_{i=1}^{m} a_i G_i(x,y) &= \bigoplus_{i=1}^{m} a_i \bigoplus_{\sigma\in V_k} x^\sigma g_{\pi(\sigma)\boxplus i}(y) \\ &= \bigoplus_{\sigma\in V_k} x^\sigma \bigoplus_{i=1}^{m} a_i g_{\pi(\sigma)\boxplus i}(y). \end{aligned}$$

Let us denote $\bigoplus_{i=1}^{m} a_i g_{\pi(\sigma)\boxplus i}(y)$ by $h_\sigma(y)$. The function $h_\sigma(y)$ is nonzero linear combination of functions $g_1, \ldots, g_m$. Since S-box $S = (g_1, \ldots, g_m)$ is regular, according to Lemma 2.2, the function $h_\sigma(y)$ is balanced. Hence, the function $\bigoplus_{i=1}^{m} a_i G_i$ is a concatenation of 2^k balanced functions. Therefore, $\bigoplus_{i=1}^{m} a_i G_i$ is balanced, too.

(iii) We know that G_i is concatenation of 2^k functions:

$$g_{\pi(0)\boxplus i}(y), g_{\pi(1)\boxplus i}(y), \ldots, g_{\pi(2^k-1)\boxplus i}(y).$$

It is obvious, that the minimal nonlinearity among these functions is at least N_g. Applying Lemma 2.3, we can conclude: $N(g_i) \geq 2^k N_g$.

(iv) The proof is analogous to the proof of (iii); $\bigoplus_{i=1}^{m} a_i G_i$ $((a_1, \ldots, a_m) \in V_m \backslash \{0\})$ is concatenation of 2^k functions

$$\bigoplus_{i=1}^{m} a_i g_{\pi(0)\boxplus i}(y), \bigoplus_{i=1}^{m} a_i g_{\pi(1)\boxplus i}(y), \ldots, \bigoplus_{i=1}^{m} a_i g_{\pi(2^k-1)\boxplus i}(y).$$

Again, we use Lemma 2.3 to obtain our proposition. □

5. Inversions

In this section we study another (but similiar) approach to building SAC-satisfying substitution block. The basic idea is to replace controlled permutation (used in previous section) by controlled inversion(s):

Let S be (again) the regular $n \times m$ S-box satisfying SAC. I is the block of controlled inversions – functions G_i $(i = 1, \ldots, m)$ are formed with help of them.

$$G_i(x, y) = g_i(y) \oplus \nu_i(x) \qquad i = 1, \ldots, m, \tag{5.1}$$

where $x = (x_1, \ldots, x_k)$, $y = (y_1, \ldots, y_n)$ and $\nu_i : V_k \to \{0,1\}$. By suitable selection of ν_i's it is possible (like in Section 4) to guarantee desirable cryptographic properties of resulting substitution block.

Theorem 5.1. *Let S be regular $n \times m$, SAC-satisfying S-box. Let $\nu_i : V_k \to \{0,1\}$ $(i = 1, \ldots, m)$ be SAC-satisfying functions. Let us consider substitution block (5.1). Then*

(i) *$\langle G_1, \ldots, G_m \rangle$ is regular $(n+k) \times m$ S-box;*
(ii) *$G_1, \ldots, G_m$ satisfy SAC;*
(iii) *$\forall\, (a_1, \ldots, a_m) \in V_m \setminus \{0\} : N(\bigoplus_{i=1}^{m} a_i G_i) \geq 2^k N(\bigoplus_{i=1}^{m} a_i g_i)$.*

Proof.

(i) According to Lemma 2.2 it is sufficient to prove the balancedness of every nonzero linear combination of $G_1, \ldots, G_m$. Let $(a_1, \ldots, a_m) \in V_m \setminus \{0\}$. Then

$$\begin{aligned}
\bigoplus_{i=1}^{m} a_i G_i(x, y) &= \bigoplus_{i=1}^{m} g_i(y) \oplus \nu_i(x) \\
&= \left(\bigoplus_{i=1}^{m} g_i(y) \right) \oplus \bigoplus_{i=1}^{m} \nu_i(x) \\
&= h(y) \oplus \bigoplus_{i=1}^{m} \nu_i(x),
\end{aligned} \tag{5.2}$$

where $h(y) = \bigoplus_{i=1}^{m} g_i(y)$. From the regularity of S-box S and Lemma 2.2 it follows that the function $h(y)$ is balanced. The function (5.2) is concatenation of 2^k functions $h(y)$ or $h(y) \oplus 1$ (depending on value of $\bigoplus_{i=1}^{m} \nu_i(x)$). Hence, $\bigoplus_{i=1}^{m} a_i G_i(x,y)$ is balanced.

(ii) There are two cases for each $i = 1, \ldots, m$:

a) Let $\gamma = (\alpha, 0) \in V_{n+k},\ \alpha \in V_k,\ wt(\alpha) = 1,\ 0 \in V_n$. Then

$$\begin{aligned} G_i(x,y) \oplus G_i((x,y) \oplus (\alpha,0)) &= g_i(y) \oplus \nu_i(x) \oplus g_i(y) \oplus \nu_i(x \oplus \alpha) \\ &= \nu_i(x) \oplus \nu_i(x \oplus \alpha). \end{aligned}$$

Since ν_i satisfies SAC, G_i satisfies propagation criterion with respect to vector γ.

b) Let $\gamma = (0, \beta) \in V_{n+k},\ \beta \in V_n,\ wt(\beta) = 1,\ 0 \in V_k$. Then

$$\begin{aligned} G_i(x,y) \oplus G_i((x,y) \oplus (0,\beta)) &= g_i(y) \oplus \nu_i(x) \oplus g_i(y \oplus \beta) \oplus \nu_i(x) \\ &= g_i(y) \oplus g_i(y \oplus \beta). \end{aligned}$$

S-box S satisfies SAC, therefore $g_i(y) \oplus g_i(y \oplus \beta)$ is balanced and G_i satisfies propagation criterion with respect to vector γ.

(iii) The function $\bigoplus_{i=1}^{m} a_i G_i(x,y)$ is concatenation of 2^k functions from the set $\{\bigoplus_{i=1}^{m} a_i g_i(y), 1 \oplus \bigoplus_{i=1}^{m} a_i g_i(y)\}$. It depends on the value of $\bigoplus_{i=1}^{m} \nu_i(x)$. Since the nonlinearity of both functions is equal, using Lemma 2.3, we can conclude that $N(\bigoplus_{i=1}^{m} a_i G_i) \geq 2^k N(\bigoplus_{i=1}^{m} a_i g_i)$. □

Remark 5.2. Notice, ν_i's do not need to be balanced.

6. Conclusion

We presented methods for constructing larger substitution blocks from smaller ones. This methods preserved SAC, that is, from SAC-fulfilling substitution blocks (or S-boxes) are constructed again SAC-fulfilling subtitution blocks. We used controlled permutations and inversions, as the basic building elements. Resulting blocks are regular, have high nonlinearity and are immune against linear cryptanalysis. These blocks can be interconnected in many different ways (e.g. the data inputs of one block are control inputs of the other block and vice versa). Such SPN-like networks have interesting properties and are object of further study. Such constructions can be suitable for use in design of stream ciphers, one-way hash functions and similiar cryptographic applications.

References

[KMI90] Kim K., Matsumoto T., Imai H., *On Generating Cryptographically Desirable Substitutions*, The Transactions of the IEICE **E 73** (1990), 1031–1035.

[M93] Matsui, M., *Linear cryptanalysis method for DES cipher*, Advances in Cryptology – EUROCRYPT'93, Springer-Verlag, 1993.

[N93] Nyberg K., *Differentially uniform mappings for cryptography*, Advances in Cryptology – EUROCRYPT'93, Springer-Verlag, 1993.

[P91] Pieprzyk J., *Bent permutations*, Proceedings of the International Conference on Finite Fields, Coding Theory, and Advances in Communications and Computing, Las Vegas, USA, 1991.

[SZZ93a] Seberry J., Zhang X.-M., Zheng Y., *Systematic Generation of Cryptographically Robust S-boxes*, The Proceedings of the First ACM Conference on Computer and Communication Security, Fairfax, Virginia, USA, 1993.

[SZZ93b] Seberry J., Zhang X.-M., Zheng Y., *Improving the Strict Avalanche Characteristics of Cryptographic Functions*, Technical Report 93-9, The Centre for Computer Security Research, University of Wollongong, Australia, 1993.

[SZZ94] Seberry J., Zhang X.-M., Zheng Y., *Relationship among nonlinearity*, Advances in Cryptology – EUROCRYPT '94, Springer-Verlag, 1994.

[WT86] Webster A.F., S.E. Tavares, *On the designs os S-boxes*, Advances in Cryptology – CRYPTO'85, Springer-Verlag, 1986.

[YT96] Youssef A.M., Tavares S.E., *Modelling Avalanche Characteristics of a Class of Substitution-Permutation Networks*, Proceedings of the 1st International Conference on Theory and Applications of Cryptology, PRAGOCRYPT'96, Czech Republic, 1996.

A New Substitution-Permutation Network Cipher Using Key-Dependent S-Boxes*

Liam Keliher †‡ Henk Meijer†

1 Introduction

The use of key-dependent s-boxes in block cipher design has not been widely investigated in the literature. Research into s-box design has focussed on determination of s-box properties which yield cryptographically strong ciphers, with the goal of selecting a small number of "good" s-boxes for inclusion in a block cipher (e.g., DES [5], CAST[1]). Simultaneously, however, a series of combinatorial results have demonstrated that a randomly chosen s-box of sufficient size will possess several of these desirable properties with high probability. This paper outlines the ongoing work of the authors' investigation into the design of a new block cipher incorporating key-dependent, pseudo-randomly generated s-boxes. Other systems using key-dependent s-boxes have been proposed in the past, the most well-known being perhaps Blowfish [17] and Khufu [12]. Each of these two systems, however, uses the cryptosystem itself to generate the s-boxes, which renders analysis difficult — we choose to avoid this approach. Preliminary results indicate that our proposed system has good cryptographic strength, with the added benefit that it is immune to linear and differential cryptanalysis, which require that the s-boxes be known. In addition, the system can easily be extended through the use of larger s-boxes and an increased number of rounds.

2 Substitution-Permutation Networks

Shannon's principles of confusion and diffusion [19, 20] are effectively realised through a substitution-permutation network (SPN) cryptosystem [3]. An SPN with key $\mathbf{K}$ is an invertible mapping $f_{\mathbf{K}} : \{0,1\}^N \rightarrow \{0,1\}^N$, where N is the number of plaintext and ciphertext

*This work was partially supported by NSERC Canada.

†Department of Computing and Information Science, Queen's University, Kingston, Ontario, Canada.
‡Corresponding author, E-mail: keliher@qucis.queensu.ca

bits. An SPN consists of R rounds, each made up of a *substitution* stage and a *permutation* stage. In the substitution stage, the current N-bit string (*block*) is fed into a series of M *substitution boxes* (*s-boxes*). An $n \times m$ s-box is a mapping $S : \{0,1\}^n \rightarrow \{0,1\}^m$, for integers n and m . For our purposes, we consider only the case that $n = m$, $N = nM$, and S is invertible. We can view S as a look-up table with row indices $0, 1, \ldots, 2^n - 1$, such that row $\mathbf{X}$ contains $S(\mathbf{X})$ (with the usual correspondence between $\{0, 1, \ldots, 2^n - 1\}$ and $\{0,1\}^n$). We will use the notation $\mathbf{Z} = Z_1 Z_2 \cdots Z_t$ for $\mathbf{Z} \in \{0,1\}^t$, where Z_i is the i^{th} bit of $\mathbf{Z}$ (numbering from most to least significant). If $S(\mathbf{X}) = \mathbf{Y}$ for some $\mathbf{X}, \mathbf{Y} \in \{0,1\}^n$, S can also be viewed as a vector of n functions, each mapping $\{0,1\}^n \rightarrow \{0,1\}$:

$$S(\mathbf{X}) = [v_1(\mathbf{X}), v_2(\mathbf{X}), \ldots, v_n(\mathbf{X})],$$

where $v_j(\mathbf{X}) = Y_j$. The v_j are called the *columns* of S. As with any function mapping $\{0,1\}^n \rightarrow \{0,1\}$, we will at times view v_j as the 2^n-bit vector,

$$v_j(0)v_j(1)\cdots v_j(2^n - 1).$$

The substitution stage is followed by a permutation of the N bits. Decryption is accomplished by running the SPN "backwards", reversing the order of the rounds, and in each round first performing the inverse permutation followed by application of the inverse s-boxes. A sample SPN with $N = 16$, $n = M = 4$, $R = 3$, and using the permutation of Kam and Davida [8] is given in Figure 1 (key not shown).

The two standard ways to incorporate the key into an SPN are shown in Figure 2 (see [6]). In the first method (a), the input to each s-box is first XOR'd with n bits derived from the key before being fed into the s-box. This may be performed during each round, or only during certain rounds. The second method (b) uses one or more key bits to select among multiple s-boxes for each sub-block of n s-box input bits.

3 S-box and SPN Properties

Since the s-boxes comprise the only nonlinear component of an SPN, they are a crucial source of cryptographic strength. S-box research has focussed largely on determining which properties yield a cryptographically "good" s-box. Some of the important properties are given below. In this section, we use $\mathbf{e_i}$ to denote a unit vector with 1 in position i, and $w(\mathbf{v})$ to mean the Hamming weight of vector $\mathbf{v}$.

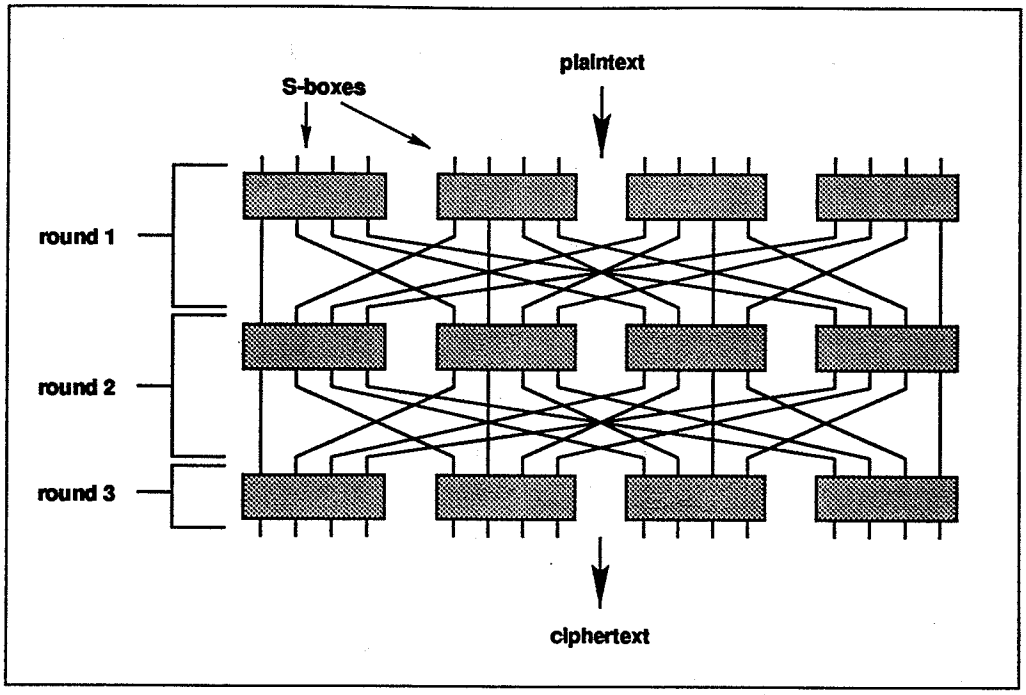

Figure 1: Example SPN with $N = 16$, $n = M = 4$, $R = 3$

3.1 Completeness

In 1979, Kam and Davida [8] defined the property of *completeness* for a bijective s-box $S : \{0,1\}^n \rightarrow \{0,1\}^n$. S is complete if for all $i, j \in \{1, 2, \ldots, n\}$, there exists $\mathbf{X} \in \{0,1\}^n$ such that $S(\mathbf{X})$ and $S(\mathbf{X} \oplus \mathbf{e_i})$ differ in at least bit j. That is to say, every output bit depends upon every input bit. An SPN is complete if it satisfies the above property for every key. Kam and Davida proposed permutations for each round which produce a complete SPN after a minimum number of rounds (given complete s-boxes).

3.2 Avalanche and Strict Avalanche

Feistel et al. defined a property of s-boxes and SPNs known as the *avalanche criterion* (AVAL) [3, 4]. A function $f : \{0,1\}^t \rightarrow \{0,1\}^t$ satisfies AVAL if whenever one input bit is changed, on average half the output bits change. In 1985, Webster and Tavares combined the completeness and avalanche properties into the *strict avalanche criterion* (SAC) [21]. A function $f : \{0,1\}^t \rightarrow \{0,1\}^t$ satisfies SAC if for all $i, j \in \{1, 2, \ldots, t\}$, flipping input bit i changes output bit j with probability exactly one half. It is easy to demonstrate that a function f which satisfies SAC is complete, and satisfies AVAL. In addition, f is said to satisfy *higher order SAC* (HOSAC) if for all $j \in \{1, 2, \ldots, t\}$, flipping any combination of

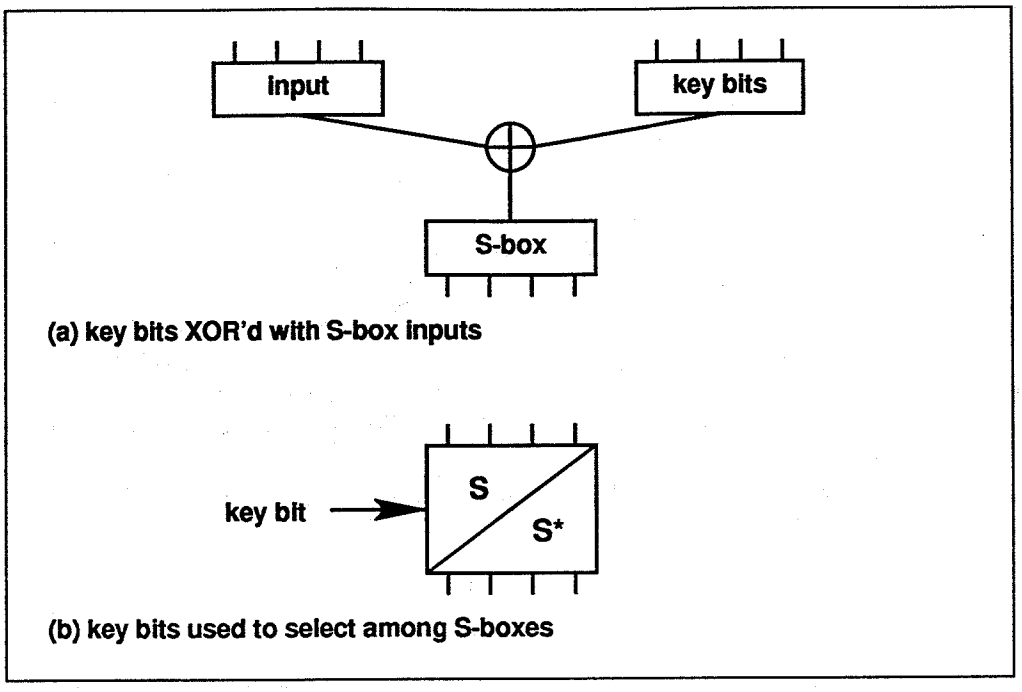

Figure 2: Two methods of incorporating the key into an SPN

one or more input bits changes output bit j with probability one half [13]. The *distance to SAC* (DSAC) and *distance to HOSAC* (DHOSAC) of a function $g : \{0,1\}^t \to \{0,1\}$ are defined as follows [13]:

$$\begin{aligned} \mathrm{DSAC}(g) &= \max_{i=1,2,\ldots,t} \frac{1}{2} \left| 2^{n-1} - \sum_{\mathbf{X}=0}^{2^n-1} g(\mathbf{X}) \oplus g(\mathbf{X} \oplus \mathbf{e_i}) \right| \\ \mathrm{DHOSAC}(g) &= \max_{\delta \in \{0,1\}^t} \frac{1}{2} \left| 2^{n-1} - \sum_{\mathbf{X}=0}^{2^n-1} g(\mathbf{X}) \oplus g(\mathbf{X} \oplus \delta) \right|. \end{aligned}$$

If $f : \{0,1\}^t \to \{0,1\}^t$ and the columns of f are $v_1, v_2, \ldots, v_t$, we define

$$\begin{aligned} \mathrm{DSAC}(f) &= \max_{j=1,2,\ldots,t} \mathrm{DSAC}(v_i) \\ \mathrm{DHOSAC}(f) &= \max_{j=1,2,\ldots,t} \mathrm{DHOSAC}(v_i). \end{aligned}$$

Note that if f satisfies SAC (HOSAC), then $\mathrm{DSAC}(f) = 0$ $(\mathrm{DHOSAC}(f) = 0)$.

3.3 Bit Independence

Webster and Tavares, in the paper in which they introduced SAC [21], also defined a property called the *bit independence criterion* (BIC). A function $f : \{0,1\}^t \to \{0,1\}^t$ satisfies BIC if for all $i, j, k \in \{1, 2, \ldots, t\}$, with $j \neq k$, inverting input bit i causes output bits j and k

to change independently. Symbolically, if $v_1, v_2, \ldots, v_t$ are the columns of f, and $j \neq k$, we have

$$\mathrm{BIC}(v_j, v_k) = \max_{i=1,2,\ldots,t} \left| \frac{\frac{1}{2^t} \sum_{\mathbf{X} \in \{0,1\}^t} [v_j(\mathbf{X}) \oplus v_j(\mathbf{X} \oplus \mathbf{e_i})]\,[v_k(\mathbf{X}) \oplus v_k(\mathbf{X} \oplus \mathbf{e_i})] - \overline{v_j}\,\overline{v_k}}{(\overline{v_j} - \overline{v_j}^2)(\overline{v_k} - \overline{v_k}^2)} \right|, \quad (1)$$

where

$$\overline{v_j} = \frac{1}{2^t} \sum_{\mathbf{X} \in \{0,1\}^t} v_j(\mathbf{X}) \oplus v_j(\mathbf{X} \oplus \mathbf{e_i}).$$

$\mathrm{BIC}(v_j, v_k)$ gives the bit independence correlation coefficient of columns j and k. The *higher order BIC* correlation coefficient of of columns j and k, $\mathrm{HOBIC}(v_j, v_k)$, is defined analogously to (1), except that the maximum is taken over all input changes consisting of one or more bit flips. We then define

$$\begin{aligned} \mathrm{BIC}(f) &= \max_{j \neq k} \mathrm{BIC}(v_j, v_k) \\ \mathrm{HOBIC}(f) &= \max_{j \neq k} \mathrm{HOBIC}(v_j, v_k) \end{aligned}$$

as a measure of how close f is to satisfying BIC (HOBIC). If f satisfies BIC (HOBIC) exactly, then $\mathrm{BIC}(f) = 0$ $(\mathrm{HOBIC}(f) = 0)$.

3.4 Nonlinearity

A function $f : \{0,1\}^t \to \{0,1\}$ is called *affine* if there exist constants $a_i \in \{0,1\}$, for $i = 0, 1, \ldots, t$, such that for all $\mathbf{X} \in \{0,1\}^t$,

$$f(\mathbf{X}) = a_0 \oplus a_1 X_1 \oplus a_2 X_2 \oplus \cdots \oplus a_t X_t.$$

An affine function is called *linear* if $a_0 = 0$. S-boxes with "high nonlinearity" are needed to make an SPN immune to *linear cryptanalysis* [10]. Let A_t be the set of all affine functions $g : \{0,1\}^t \to \{0,1\}$. For $f : \{0,1\}^t \to \{0,1\}$, we define the *nonlinearity* of f as

$$\mathrm{nl}(f) = \min_{g \in A_t} w(f \oplus g)$$

(in this expression, we view f and g as 2^t-bit vectors). If S is an s-box, let $\mathcal{C}$ be the set of all nonzero linear combinations of the columns of S. Then the nonlinearity of S is

$$\mathrm{nl}(S) = \min_{c \in \mathcal{C}} \mathrm{nl}(c).$$

3.5 XOR Table Distribution

In 1991, Biham and Shamir introduced a powerful cryptanalytic technique known as *differential cryptanalysis* [2]. They have successfully applied their attack to a variety of SPNs, including DES. Differential cryptanalysis requires knowledge of the *XOR tables* of the s-boxes. For an $n \times n$ s-box, S, the XOR table has rows and columns indexed by $0, 1, 2, \ldots, 2^n - 1$, and the table entries are defined as follows. If $i, j \in \{0, 1, 2, \ldots, 2^n - 1\}$, position (i, j) in the XOR table contains the value

$$|\{\mathbf{X} \in \{0,1\}^n : S(\mathbf{X}) \oplus S(\mathbf{X} \oplus i) = j\}| \tag{2}$$

(in (2) we are treating i and j as their equivalent n-bit strings). Note that (2) always evaluates to an even number. The pair (i, j) is called an input/output XOR pair. Differential cryptanalysis exploits such XOR pairs with large XOR table entries. An SPN can be secured against differential cryptanalysis by selecting s-boxes with low XOR table entries, ideally all 0 or 2 (the one exception is entry $(0, 0)$ which has value 2^n). Even if the XOR table is not directly calculated, resistance to differential cryptanalysis can be achieved by assuring that the s-boxes have good *diffusive* properties, i.e., they reasonably satisfy AVAL or SAC [16].

4 Properties of Random S-boxes

Since the object of this proposal is a new SPN cipher using key-dependent s-boxes, it will be useful to investigate the average properties of random invertible $n \times n$ s-boxes. Since the introduction of DES [5], a number of such results have appeared in the literature.

4.1 Completeness

O'Connor proved in [14] that a randomly chosen $n \times n$ invertible s-box has a high probability of being complete for sufficiently large n. In fact, he showed that the probability that such an s-box is *not* complete is

$$o\left(\frac{\sqrt{2^n}}{2^{2^{n-1}+n-1}}\right).$$

For an exact formula, see [15].

4.2 Avalanche and Strict Avalanche

The authors have not found any results giving the probability that a random invertible $n \times n$ s-box satisfies AVAL or SAC (although there are bounds on the probability that a random

function $f : \{0,1\}^t \rightarrow \{0,1\}$ satisfies SAC). On the other hand, a number of theoretical and experimental results exist concerning the AVAL property for SPNs. Heys and Tavares developed a probabilistic model for the AVAL property of an SPN [7]. Their results for $N = 64$ and $n = M = 8$, using randomly selected s-boxes and the fixed permutation of Kam and Davida [8], indicate that AVAL is reasonably satisfied after 5 or more rounds. In fact, if E_R is the expected number of output bit changes after R rounds when one input bit is flipped, and we define $\epsilon = |1 - E_R/(N/2)|$, then $\epsilon \leq 10^{-5}$ for $R \geq 7$.

4.3 Nonlinearity

For an invertible $n \times n$ s-box S and an integer $2L$ $(0 \leq L \leq 2^{n-2})$, Youssef and Tavares [22] prove that

$$\textbf{prob}\left[\text{nl}(S) \leq \left(2^{n-1} - 2L\right)\right] < \frac{2(2^{n-1}!)^2(2^n - 1)^2}{2^n!} \sum_{l=L}^{2^{n-2}} \binom{2^{n-1}}{2^{n-2}+l}^2. \tag{3}$$

If $n = 8$, we have, for example, $\textbf{prob}\left[\text{nl}(S) \leq 64\right] < 1.4 \times 10^{-11}$ and $\textbf{prob}\left[\text{nl}(S) \leq 80\right] < 4.6 \times 10^{-5}$. Experimental results support the theoretical result of (3). For example, Heys [6] generated 200 random invertible 8×8 s-boxes and found that each satisfied $86 \leq \text{nl}(S) \leq 98$.

4.4 XOR Table Distribution

If S is a randomly chosen invertible $n \times n$ s-box, and $0 \leq A \leq 2^n$ is an *even* integer, a formula of Youssef and Tavares gives the probability that the maximum XOR table entry of S (denoted maxXOR(S)) is $\geq A$ [22]. For example, if $n = 8$ and $A = 16$, we have $\textbf{prob}[\text{maxXOR}(S) \geq 16] < 0.0042$.

4.5 Cyclic Properties

There is some indication that the cyclic properties (cycle length, number of cycles) of an s-box are related to other cryptographic properties. Youssef et al. give experimental results which show that, on average, s-boxes with fewer fixed points have higher nonlinearity and lower maximum XOR table entries [23]. They prove that the expected number of fixed points for a random invertible s-box is 1, with a variance of 1. They also state that the expected value and variance of the number of cycles is approximately $\log_e 2^n \approx 0.69n$; and that the expected cycle length is $2^{n-1} + 1/2$, where the expected cycle length is defined as the value of the length of the cycle to which a randomly chosen element belongs.

5 The Proposed Cryptosystem

5.1 Design and Rationale

The cryptosystem we are investigating is a 64-bit SPN with 8×8 key-dependent s-boxes. The SPN design has the advantages of being simple, and having been subjected to extensive cryptanalysis [6]. In each round we use the permutation of Kam and Davida [8], which connects output bit i of s-box j in round r, to input bit j of s-box i in round $r+1$ ($1 \leq i, j \leq 8$, $1 \leq r < R$). The s-boxes are changed from round to round, so the total number of s-boxes generated is $M \cdot R$. The number of rounds will be determined from the results of the testing described in Section 5.3.

The fact that the s-boxes are unknown to the cryptanalyst is one of the principal strengths of our system, since both linear and differential cryptanalysis require known s-boxes. It is not apparent that the pseudo-random nature of the s-boxes introduces any exploitable weakness into the system. The results of Section 4 indicate that if the s-boxes are generated from the key in a sufficiently random fashion, each s-box has a high probability of being complete, possessing fairly high nonlinearity, and having its largest XOR table entry < 16.

5.2 Random S-Box Generation Process

Figure 3 depicts the conceptual layout of our s-box generation process. The key, $\mathbf{K}$, is used

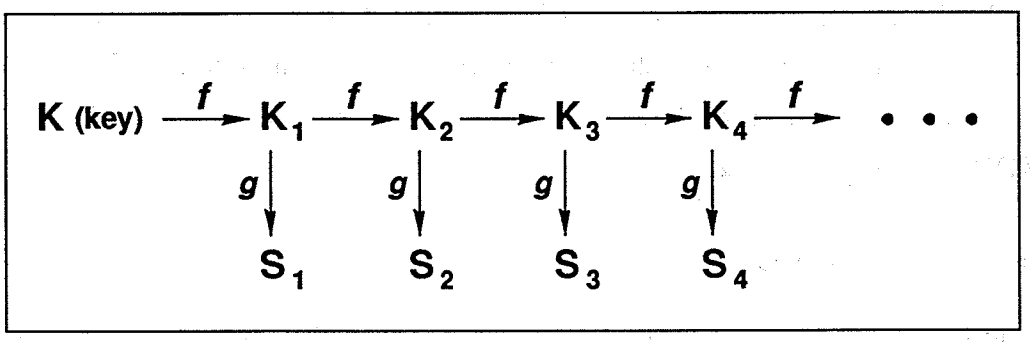

Figure 3: Conceptual approach to random s-box generation

to generate a series of *subkeys* $K_1, K_2, \ldots, K_{MR}$, by application of a function f. A second function g generates the i^{th} s-box, S_i, from K_i. The functions f and g must meet certain requirements. First of all, they must produce s-boxes which are satisfactorily random, in order to achieve the results of Section 4. Secondly, even though the s-boxes are in principle secret, we want the generation process to be such that if a cryptanalyst determines one of

the s-boxes, this does not yield any information about any other s-box. There are a number of ways to achieve this security. One approach is to make f cryptographically secure; g can then be any simple function which generates pseudo-random s-boxes from the K_i. Another method is to make g one-way, for example a secure hash function, or a many-to-one function. Then it is only required that f be sufficiently random.

We are currently pursuing the first approach—we have chosen the RC4 stream cipher (given in [18]) for f. RC4 is a simple and widely used cipher with a variable-length key (up to 2048 bits). Software implementation of RC4 is extremely short, requiring about ten lines of C code. (Although technically proprietary, the RC4 cipher is publicly known and has undergone cryptanalysis [18].) For g we use the following simple algorithm.

```
for i = 1 to 256
  sbox[i] := i;       /* initialise to identity s-box */

for i = 1 to 256
  j := RandomInRange (i, 256);
  swap (sbox[i], sbox[j]);
end for
```

where `RandomInRange` uses RC4 to generate a random integer in the interval $[i, 256]$. This algorithm chooses an s-box uniformly from the set of all invertible 8×8 s-boxes. If RC4 is assumed to be cryptographically secure, it is clear from Figure 3 that knowledge of S_i will not give any information about S_j, for $i \neq j$.

We are also considering a number of other options for f and g. As this work progresses, test results for these different methods will enable us to determine the approach which yields the most secure SPN. Note that we plan to keep the s-box generation process separate from the SPN itself. Two other well-known cryptosystems, Blowfish [17] and Khufu [12], also use key-dependent s-boxes, but do not make this separation. Each uses the cryptosystem itself in some initial state to pseudo-randomly generate s-boxes; these are then used for the actual encryption and decryption. By avoiding this self-referential approach, we hope to simplify the analysis of our system.

Our system will require a small amount of startup time to generate the s-boxes, making it suitable for applications such as cellular phones, which can tolerate a short startup time. For example, a non-optimised software version of the SPN using 16 rounds required 0.07 sec. of startup time on a SUN Ultra 1 (140 MHz UltraSPARC CPU).

5.3 Testing Approaches

We are in the process of subjecting our SPN design to a variety of analytical and statistical tests.

1. Verification that the s-box generation process is sufficiently random. The s-boxes produced must be consistent with the theoretical results of Section 4.

2. Randomness tests of sample ciphertext (see [9, 11]).

3. Statistical testing to determine minimum number of rounds for SPN to satisfy properties such as AVAL and completeness.

4. Investigation of concept of *correlation between s-boxes.* We want to ensure that disclosure of one s-box does not yield any statistical information about a second s-box. Let S_i and S_j be two s-boxes generated for our cryptosystem. We are considering (among others) correlations between:

   (a) the columns of S_i and S_j (possibly cyclically shifted up or down)

   (b) the XOR table entries of S_i and S_j

   (c) the nonlinearities of S_i and S_j

5.4 Preliminary Results

Using the method of Section 5.2 involving RC4, we generated 500 random s-boxes and tested them for the properties of Section 4 (a 128-bit key was selected at random). We summarise the results in Figure 4. Section 4.1 states that each s-box has a high probability of being complete. It is easy to see that an $n \times n$ s-box, S, is complete if and only if $\mathrm{DSAC}(S) \neq 2^{n-2}$. Therefore, since the maximum DSAC value is 22, all 500 s-boxes are complete.

Figure 5 compares the distribution of s-box nonlinearities as given by the theoretical result of Youssef and Tavares [22] (Section 4.4), with that obtained by our generation of 500 invertible s-boxes. The middle row gives upper bounds on the probability that the nonlinearity of a randomly chosen invertible s-box is less than or equal to the NL value given in the first row. The last row gives the fraction of our s-boxes with nonlinearity less than or equal to the given value NL. The similarity of theoretical and experimental values suggests that our s-box generation process behaves sufficiently randomly.

	MEAN	STD DEV	MIN	MAX
nl(S)	92.76	2.13	86	96
maxXOR(S)	11.25	1.17	10	14
DSAC(S)	14.64	2.37	10	22
DHOSAC(S)	20.37	1.79	16	28
BIC(S)	0.032	0.265	0.188	0.414
HOBIC(S)	0.066	0.341	0.293	0.443
num. fixpt	0.95	1.07	0	6
num. cycles	6.14	2.05	1	12
exp. cycle length	125.07	50.77	38.02	256.00
column corr.	0.0033	0.2578	0.2188	0.3594
lin. comb. corr.	0.0241	0.2760	0.2344	0.3438

Figure 4: Results of random s-box generation using RC4

NL	84	86	88	90	92	94	96	98
theoret. **prob**[nl(S) <= NL]	0.0037	0.015	0.058	0.21	0.70	> 1	> 1	> 1
exper. **prob**[nl(S) <= NL]	0.0	0.012	0.058	0.18	0.49	0.88	1.00	1.00

Figure 5: Theoretical and experimental distribution of nonlinearities

In Figure 6 we present a similar probability distribution comparison, this time for the values maxXOR(S). The middle row gives a theoretical upper bound that **prob**[maxXOR(S) $\geq$ A], while the third row gives the fraction of the 500 s-boxes we randomly generated for which maxXOR(S) $\geq A$.

The s-boxes generated exhibit predictable cyclic properties (Section 4.5). The mean and variance (square of standard deviation) of the number of fixed points are 0.95 and 1.14, respectively, both close to the theoretical value of 1. The mean and variance of the number of cycles, 6.14 and 4.20, roughly approximate the theoretical value $\log_e 2^8 \approx 5.55$. And the mean of the expected cycle length is 125.07, with a theoretical value of 128.5.

A	10	12	14	16
theor. **prob**[maxXOR(S) $\geq A$]	> 1	0.94	0.067	0.0042
exper. **prob**[maxXOR(S) $\geq A$]	1.00	0.57	0.054	0.0

Figure 6: Theoretical and experimental distribution of maximum XOR table entries

The last two rows of Figure 4 contain values resulting from our investigation of correlation between s-boxes. If the 500 randomly generated 8×8 s-boxes are $S_1, S_2, \ldots, S_{500}$, for a given S_i and $s \in \{0, 1, \ldots, 255\}$ define the s-box $S_i^s(\mathbf{X}) = S_i((\mathbf{X} - s) \bmod 256)$, i.e., S_i^s is S_i cyclically shifted down s rows. For all i, j, k, s ($1 \leq i \leq 499$, $1 \leq j, k \leq 8$, $0 \leq s \leq 255$), we compute the correlation coefficient (see Section 3.3) between column j of S_i and column k of S_{i+1}^s. The row of Figure 4 labelled "column corr." reports the results of these computations. The last row of Figure 4 is obtained by calculating each correlation coefficient between an element of $\mathcal{C}_i$ and an element of $\mathcal{C}_{i+1}$, where $\mathcal{C}_i$ is the set of all nontrivial linear combinations of the columns of S_i. The last two rows of Figure 4 do not hold any extreme values, suggesting that the correlations investigated do not yield useful cryptanalytic information.

We performed another statistical test to verify the lack of correlation between columns of consecutive s-boxes. We generated 30,000 pairs of consecutive s-boxes using RC4 (with a 128-bit key): (S_1, T_1), $(S_2, T_2), \ldots, (S_{30000}, T_{30000})$. For each $j, k \in \{1, 2, \ldots, 8\}$, we computed the distribution of the correlation coefficient between column j of S_i and column k of T_i for $1 \leq i \leq 30,000$. Each such distribution was compared to the theoretical distribution of correlation values obtained when two balanced 256-bit vectors are chosen at random (this is easy to calculate). Using a chi-square test, we determined that the 64 distributions obtained behaved as expected, appearing to be drawn from the theoretical distribution.

6 Conclusion

The research direction presented in this proposal holds promise because few SPN cryptosystems exist which make use of key-dependent s-boxes. This research has the potential of resulting in a new cryptosystem that appears to be secure. The key-dependent nature of the s-boxes makes our proposed system immune to linear and differential cryptanalysis, and the use of (relatively large) 8×8 s-boxes guarantees with high probability that the s-boxes which are generated will possess good cryptographic properties.

References

[1] C.M. Adams, *Constructing Symmetric Ciphers Using the CAST Design Procedure,* to appear in: Designs, Codes and Cryptography.

[2] E. Biham and A. Shamir, *Differential cryptanalysis of DES-like cryptosystems,* Journal of Cryptology, Vol. 4, No. 1, pp. 3–72, 1991.

[3] H. Feistel, *Cryptography and computer privacy,* Scientific American, Vol. 228, No. 5, pp. 15–23, May 1973.

[4] H. Feistel, W.A. Notz, and J.L. Smith, *Some cryptographic techniques for machine to machine data communications,* Proceedings of the IEEE, Vol. 63, No. 11, pp. 1545–1554, November 1975.

[5] *Data Encryption Standard (DES),* National Bureau of Standards FIPS Publication 46, 1977.

[6] H.M. Heys, *The design of substitution-permutation network ciphers resistant to cryptanalysis,* Ph.D. Thesis, Queen's University, Canada, 1994.

[7] H.M. Heys and S.E. Tavares, *Avalanche characteristics of substitution-permutation encryption networks,* IEEE Transactions on Computers, Vol. 44, No. 9, September 1995.

[8] J.B. Kam and G.I. Davida, *Structured design of substitution-permutation encryption networks,* IEEE Transactions on Computers, Vol. C-28, No. 10, October 1979.

[9] J.C. Lagarias, *Pseudo-random number generators in cryptography and number theory,* Proceedings of Symposia in Applied Mathematics, Vol. 42, 1990.

[10] M. Matsui, *Linear cryptanalysis method for DES cipher,* Advances in Crytology: Proceedings of EUROCRYPT'93, Springer-Verlag, Berlin, pp. 386–397, 1994.

[11] A.J. Menezes, P.C. van Oorschot, and S.A. Vanstone, *Handbook of Applied Cryptography,* CRC Press, 1997.

[12] R. Merkle, *Fast software encryption functions,* Advances in Cryptology: Proceedings of CRYPTO'90, Springer-Verlag, Berlin, pp. 476–501, 1991.

[13] S. Mister and C.M. Adams, *Practical S-Box Design,* SAC'96 — Third Annual Workshop on Selected Areas in Cryptography, Queen's University, Kingston, Ontario, pp. 61–76.

[14] L. O'Connor, *Enumerating nondegenerate permutations,* Advances in Cryptology: EUROCRYPT'91, Springer-Verlag, pp. 368–377, 1992.

[15] L. O'Connor, *Enumerating nondegenerate permutations,* Technical Report 2527, University of Waterloo Waterloo, Ontario, Canada, 1991.

[16] L. O'Connor, *On the distribution of characteristics in bijective mappings,* Advances in Cryptology: Proceedings of EUROCRYPT'93, Springer-Verlag, Berlin, pp. 360–370, 1994.

[17] B. Schneier, *Description of a new variable-length, 64-bit block cipher (Blowfish),* Fast Software Encryption, pp. 191–204.

[18] B. Schneier, *Applied Cryptography, Second Ed.,* John Wiley and Sons, 1996.

[19] C.E. Shannon, *A mathematical theory of communication,* Bell System Technical Journal, Vol. 27, No. 4, pp. 379–423, 623–656, 1948.

[20] C.E. Shannon, *Communication theory of secrecy systems,* Bell System Technical Journal, Vol. 28, no. 4, pp. 656–715, 1949.

[21] A.F. Webster and S.E. Tavares, *On the design of S-boxes,* Advances in Cryptology: Proceedings of CRYPTO'85, Springer-Verlag, Berlin, pp. 523–534, 1986.

[22] A.M. Youssef and S.E. Tavares, *Resistance of balanced s-boxes to linear and differential cryptanalysis,* Information Processing Letters, Vol. 56, pp. 249–252, 1995.

[23] A.M. Youssef, S.E. Tavares, and H.M. Heys, *A new class of substitution-permutation networks,* SAC '96 — Third Annual Workshop on Selected Areas in Cryptography, Queen's University, Kingston, Ontario, pp. 132–147.

Non-existence of Certain Quadratic S-boxes and Two Bounds on Nonlinear Characteristics of General S-boxes

Xian-Mo Zhang
Department of Computer Science
The University of Wollongong
Wollongong, NSW 2522, AUSTRALIA
E-mail: xianmo@cs.uow.edu.au

Yuliang Zheng
School of Computing and Information Technology
Monash University
Melbourne, VIC 3199, AUSTRALIA
E-mail: yuliang@mars.fcit.monash.edu.au
URL: http://www-pscit.fcit.monash.edu.au/~yuliang/

Hideki Imai
Institute of Industrial Science
The University of Tokyo
7-22-1 Roppongi, Minato-ku, Tokyo 106, JAPAN
Email: imai@iis.u-tokyo.ac.jp

July 17, 1997

Abstract

S-boxes with a uniformly half-occupied difference distribution table are considered useful in designing a block cipher secure against differential attacks. Researchers, however, have conjectured that for all $n > m$, there exist no $n \times m$ S-boxes with a uniformly half-occupied difference distribution table. Prior to this work, the best known result that supports the conjecture is that there exist no *quadratic* S-boxes with a uniformly half-occupied difference distribution table *if n or m is even*. In this paper we provide further evidence to support the conjecture. In particular, we show that there exists no *quadratic* S-box with a uniformly half-occupied difference distribution table *if* $n \geqq 2m - 1$. The other two contributions of this work are concerned about two of most important nonlinear characteristics of (general) S-boxes, namely differential uniformity and nonlinearity. In particular, we derive a non-trivial and tight lower bound on the differential uniformity of an S-box, and then reveal a relationship between the nonlinearity and differential characteristics of an S-box.

1 Introduction

This paper deals with $n \times m$ S-boxes with $n > m$. Success of the notable differential cryptanalysis on various block ciphers [3, 4] has motivated researchers to search for S-boxes

whose difference distribution tables are relatively flat. As S-boxes with a *completely flat* difference distribution table have been known to be weak in resisting against differential attacks, naturally one of the research focuses has been on designing S-boxes with a uniformly half-occupied difference distribution table (UHODDT), i.e., S-boxes whose differential distribution tables contain an equal number of zero and identical non-zero entries in each of their rows (not taking into account the first row). Previous works directly or indirectly related to this line of research include, but not limited to, [1, 2, 12, 13, 14, 15, 16].

Defying efforts by a number of researchers, no $n \times m$ S-box with a UHODDT has emerged. This has led to a conjecture which states that

for all $n > m$, there exists no $n \times m$ S-box with a UHODDT.

Some progress in proving the conjecture was made in [22] where it was shown that when n or m is even, there exists no *quadratic* $n \times m$ S-box with a UHODDT (see Theorem 1 of [22]). This paper reports further progress in proving the conjecture. In particular, we show that *when* $n \geqq 2m - 1$, there exists no *quadratic* $n \times m$ S-box with a UHODDT. We hope that this new piece of evidence can be of some contribution to the eventual success in proving the conjecture.

The second issue addressed in this paper is on the lower bound of differential uniformity. The differential uniformity of an S-box is defined as the largest value in the differential distribution table of the S-box. For an $n \times m$ S-box, it is easy to see that its differential uniformity is at least 2^{n-m}. As another contribution of this paper, we will show a new tight lower bound that considerably improves the "trivial" bound of 2^{n-m}.

The final issue addressed in this work relates more specifically the nonlinearity of an S-box to its difference distribution table. In particular, it shows an upper bound on the nonlinearity of the S-box expressed in terms of three parameters: the number of input bits, the number of output bits and the values in the leftmost column of its difference distribution table. We also compare the new upper bound with previous works in the same area.

The remainder of this paper is organized as follows: Section 2 introduces formal notations and definitions used in this paper. Section 3 represents a sample of known results on S-boxes that are relevant to this paper. This is followed by Section 4 where it is proved that for $n \geqq 2m - 1$, there exists no quadratic $n \times m$ S-box with a UHODDT. A general tight lower bound on the differential uniformity of an S-box is presented in Section 5, and then a relationship between the nonlinearity of an S-box and its difference distribution table is proved in Section 6. Finally the paper is closed with some remarks in Section 7.

2 Basic Notations and Definitions

This section is intended as a summary of the minimum amount of mathematical knowledge required in rigorously treating issues on S-boxes to be discussed in this paper.

The vector space of n tuples of elements from $GF(2)$ is denoted by V_n. These vectors, in ascending alphabetical order, are denoted by α_0, α_1, ..., α_{2^n-1}. As vectors in V_n and integers in $[0, 2^n - 1]$ have a natural one-to-one correspondence, it allows us to switch from a vector in V_n to its corresponding integer in $[0, 2^n - 1]$, and vice versa.

Let f be a function from V_n to $GF(2)$ (or simply, a function on V_n). The *sequence* of f is defined as $((-1)^{f(\alpha_0)}, (-1)^{f(\alpha_1)}, \ldots, (-1)^{f(\alpha_{2^n-1})})$, while the *truth table* of f is defined as $(f(\alpha_0), f(\alpha_1), \ldots, f(\alpha_{2^n-1}))$. f is said to be *balanced* if its truth table assumes an equal number of zeros and ones. We call $h(x) = a_1x_1 \oplus \cdots \oplus a_nx_n \oplus c$ an *affine function*, where

$x = (x_1, \ldots, x_n)$ and $a_j, c \in GF(2)$. In particular, h will be called a *linear function* if $c = 0$. The sequence of an affine (linear) function will be called an *affine (linear) sequence.*

The *Hamming weight* of a vector v, denoted by $W(v)$, is the number of ones in v. Let f and g be functions on V_n. Then $d(f,g) = \sum_{f(x) \neq g(x)} 1$, where the addition is over the reals, is called the *Hamming distance* between f and g. Let $\varphi_0, \ldots, \varphi_{2^{n+1}-1}$ be the affine functions on V_n. Then $N_f = \min_{i=0,\ldots,2^{n+1}-1} d(f, \varphi_i)$ is called the *nonlinearity* of f. It is well-known that the nonlinearity of f on V_n satisfies $N_f \leqq 2^{n-1} - 2^{\frac{1}{2}n-1}$. The equality holds if and only if f is bent (see P. 426 of [10]).

Given two sequences $a = (a_1, \ldots, a_m)$ and $b = (b_1, \ldots, b_m)$, their component-wise product is denoted by $a*b$, while the scalar product (sum of component-wise products) is denoted by $\langle a, b \rangle$.

Definition 1 Let f be a function on V_n. For a vector $\alpha \in V_n$, denote by $\xi(\alpha)$ the sequence of $f(x \oplus \alpha)$. Thus $\xi(0)$ is the sequence of f itself and $\xi(0) * \xi(\alpha)$ is the sequence of $f(x) \oplus f(x \oplus \alpha)$. Define the *auto-correlation* of f with a shift α by

$$\Delta(\alpha) = \langle \xi(0), \xi(\alpha) \rangle.$$

The *Sylvester-Hadamard matrix (or Walsh-Hadamard matrix)* of order 2^n, denoted by H_n, is generated by the recursive relation

$$H_n = \begin{bmatrix} H_{n-1} & H_{n-1} \\ H_{n-1} & -H_{n-1} \end{bmatrix}, \; n = 1, 2, \ldots, \; H_0 = 1.$$

Each row (column) of H_n is a linear sequence of length 2^n.

The following two lemmas can be found in [20].

Lemma 1 *Let ξ be the sequence of a function f on V_n. Then the nonlinearity of f, N_f can be calculated by*

$$N_f = 2^{n-1} - \frac{1}{2} max\{|\langle \xi, \ell_i \rangle|, 0 \leqq i \leqq 2^n - 1\}$$

where ℓ_i is the ith row of H_n, $i = 0, 1, \ldots, 2^n - 1$.

3 Some Known Results on S-boxes

An $n \times m$ S-box or substitution box is actually a mapping from V_n to V_m, i.e., $F = (f_1, \ldots, f_m)$, where n and m are integers with $n \geqq m \geqq 1$ and each component function f_j is a function on V_n. In this paper, we use the terms of mapping and S-box interchangeably.

As can be seen from the design of many practical block ciphers, researchers are mainly concerned with *regular* S-boxes only. A mapping $F = (f_1, \ldots, f_m)$ is said to be regular if $F(x)$ runs through each vector in V_m 2^{n-m} times while x runs through V_n once. One can easily see that $n \times m$ regular S-boxes exist only for $n \geqq m$.

The following lemma states a useful result on the regularity of an S-box. This result has appeared in many different forms in the literature. Our description follows the binary case of Corollary 7.39 of [9].

Lemma 2 *Let $F = (f_1, \ldots, f_m)$ be a mapping from V_n to V_m, where n and m are integers with $n \geqq m \geqq 1$ and each $f_j(x)$ is a function on V_n. Then F is regular if and only if every non-zero linear combination of $f_1, \ldots, f_m$ is balanced.*

The concept of nonlinearity can be extended to the case of an S-box. The standard definition of the nonlinearity of $F = (f_1, \ldots, f_m)$ is

$$N_F = min_g\{N_g | g = \bigoplus_{j=1}^{m} c_j f_j,\ c_j \in GF(2), g \neq 0\}.$$

Now we introduce three more notations associated with $F = (f_1, \ldots, f_m)$. Namely, $k_j(\alpha)$, $\Delta_j(\alpha)$ and η_j.

Definition 2 *Let $F = (f_1, \ldots, f_m)$ be an $n \times m$ S-box, $\alpha \in V_n$, $j = 0, 1, \ldots, 2^m - 1$ and $\beta_j = (b_1, \ldots, b_m)$ be the vector in V_m that corresponds to the binary representation of j. In addition, set $g_j = \bigoplus_{u=1}^{m} b_u f_u$ be the jth linear combination of the component functions of F. Then we define*

1. *$k_j(\alpha)$ as the number of times $F(x) \oplus F(x \oplus \alpha)$ runs through $\beta_j \in V_m$ while x runs through V_n once,*
2. *$\Delta_j(\alpha)$ as the auto-correlation of g_j with shift α,*
3. *η_j as the sequence of g_j.*

Using the three notations, we formally define three tables/matrices related to $F = (f_1, \ldots, f_m)$.

Definition 3 For an S-box $F = (f_1, \ldots, f_m)$, set

$$K = \begin{bmatrix} k_0(\alpha_0) & k_1(\alpha_0) & \ldots & k_{2^m-1}(\alpha_0) \\ k_0(\alpha_1) & k_1(\alpha_1) & \ldots & k_{2^m-1}(\alpha_1) \\ & \vdots & & \\ k_0(\alpha_{2^n-1}) & k_1(\alpha_{2^n-1}) & \ldots & k_{2^m-1}(\alpha_{2^n-1}) \end{bmatrix}$$

$$D = \begin{bmatrix} \Delta_0(\alpha_0) & \Delta_1(\alpha_0) & \ldots & \Delta_{2^m-1}(\alpha_0) \\ \Delta_0(\alpha_1) & \Delta_1(\alpha_1) & \ldots & \Delta_{2^m-1}(\alpha_1) \\ & \vdots & & \\ \Delta_0(\alpha_{2^n-1}) & \Delta_1(\alpha_{2^n-1}) & \ldots & \Delta_{2^m-1}(\alpha_{2^n-1}) \end{bmatrix}$$

and

$$P = \begin{bmatrix} \langle \eta_0, \ell_0 \rangle^2 & \langle \eta_1, \ell_0 \rangle^2 & \cdots & \langle \eta_{2^m-1}, \ell_0 \rangle^2 \\ \langle \eta_0, \ell_1 \rangle^2 & \langle \eta_1, \ell_1 \rangle^2 & \cdots & \langle \eta_{2^m-1}, \ell_1 \rangle^2 \\ & \vdots & & \\ \langle \eta_0, \ell_{2^n-1} \rangle^2 & \langle \eta_1, \ell_{2^n-1} \rangle^2 & \cdots & \langle \eta_{2^m-1}, \ell_{2^n-1} \rangle^2 \end{bmatrix}$$

where ℓ_i is the ith row of H_n, $i = 0, 1, \ldots, 2^n - 1$. The three $2^n \times 2^m$ matrices K, D and P are called *difference distribution table, auto-correlation distribution table and correlation immunity distribution table* of the S-box F respectively.

The following lemma, first appeared in [21], shows an intimate relationship between the three tables K, D and P defined above. It turns out that the lemma is very useful in examining cryptographic properties of an S-box. In particular, part (ii) of the lemma will be used in proving one of the main results in this paper.

Lemma 3 *Let $F = (f_1, \ldots, f_m)$ be a mapping from V_n to V_m, where n and m are integers with $n \geqq m \geqq 1$ and each $f_j(x)$ is a function on V_n. Set $g_j = \bigoplus_{u=1}^{m} c_u f_u$ where $(c_1, \ldots, c_m)$ is the binary representation of integer j, $j = 0, 1, \ldots, 2^m - 1$. Then*

(i)

$$(k_0(\alpha_i), k_1(\alpha_i), \ldots, k_{2^m-1}(\alpha_i))H_m = (\Delta_0(\alpha_i), \Delta_1(\alpha_i), \ldots, \Delta_{2^m-1}(\alpha_i))$$

where α_i is the binary representation of integer i,

(ii) $D = KH_m$,

(iii) $P = H_n D$,

(iv) $P = H_n K H_m$.

Now we consider an S-box in terms of its usefulness in designing a block cipher secure against differential cryptanalysis [3, 4]. The essence of a differential attack is that it exploits particular entries in the difference distribution tables of S-boxes employed by a block cipher. The difference distribution table of an $n \times m$ S-box is a $2^n \times 2^m$ matrix. The rows of the matrix, indexed by the vectors in V_n, represent the changes in the inputs, while the columns, indexed by the vectors in V_m, represent the change in the output of the S-box. An entry in the table indexed by (α, β) indicates the number of input vectors which, when changed by α (in the sense of bit-wise XOR), result in a change in the output by β (also in the sense of bit-wise XOR).

Note that an entry in a difference distribution table can only take an even value, the sum of the values in a row is always 2^n, and the first row is always $(2^n, 0, \ldots, 0)$. As entries with higher values in the table are particularly useful to differential cryptanalysis, a necessary condition for an S-box to be immune to differential cryptanalysis is that it does not have large values in its differential distribution table (not counting the first entry in the first row).

In measuring the strength of an S-box (in terms of the security of a block cipher that employs the S-box) against differential attacks, a useful indicator commonly used is *differential uniformity* whose formal definition follows [14].

Definition 4 *Let F be an $n \times m$ S-box, where $n \geqq m$. Let δ be the largest value in the differential distribution table of the S-box (not counting the first entry in the first row), namely,*

$$\delta = \max_{\alpha \in V_n, \alpha \neq 0} \max_{\beta \in V_s} |\{x | F(x) \oplus F(x \oplus \alpha) = \beta\}|$$

or equivalently

$$\delta = \max\{k_j(\alpha) | j = 0, 1, \ldots, 2^m - 1,\ \alpha \in V_n,\ \alpha \neq 0\}.$$

Then F is said to be differentially δ-uniform, *and accordingly, δ is called the differential uniformity of F.*

Obviously the differential uniformity δ of an $n \times m$ S-box is constrained by $2^{n-m} \leqq \delta \leqq 2^n$. Extensive research has been carried out to construct differentially δ-uniform S-boxes with low δ [1, 2, 12, 13, 14, 15, 16]. Some constructions, in particular those based on permutation polynomials on finite fields, are simple and elegant. However, caution must be

taken with Definition 4. In particular, it should be noted that low differential uniformity (a small δ) is only a *necessary*, but not a *sufficient* condition for immunity to differential attacks. This is shown by the fact that for $n \geqq m$, $n \times m$ S-boxes constructed in [1, 12], which have a flat difference distribution table, are extremely weak against differential attacks, despite the fact that they achieve the lowest possible differential uniformity $\delta = 2^{n-m}$ [4, 5, 17].

We are particularly interested in $n \times m$ S-boxes that have the following property: for each nonzero vector $\alpha \in V_n$, $F(x) \oplus F(x \oplus \alpha)$ runs through 2^{m-t}, $1 \leqq t \leqq m$, of the vectors in V_m, each 2^{n-m+t} times, but not through the other $2^m - 2^{m-t}$ vectors in V_m. With each row in the difference distribution table of such an S-box, 2^{m-t} of its entries contain a value 2^{n-m+t} while the remaining entries contain a value zero. For simplicity, we say such a difference distribution table to be *uniformly 2^{m-t}-occupied*.

For n odd, $n = m$ (i.e., permutation S-boxes) and $t = 1$, there has been a large body of research (see for instance [2, 12, 13, 14, 15, 16]). One of the properties of these permutations is that their differential distribution tables are all 2-uniform, namely, half of the entries in a row contain a value zero while the other half contain a value 2. For this reason, it has been believed that these permutations achieve the highest possible robustness against the (first order) differential attack.

As an extension of the above observation to a general $n \times m$ S-box with $n \geqq m$, one would expect that the S-box would be highly useful in resisting differential attacks if its difference distribution table is *uniformly 2^{m-1}-occupied*, i.e., each row in the difference distribution table contains an equal number of zero and non-zero entries with all the non-zero values being identical to 2^{n-m+1}. For simplicity, we say that such an $n \times m$ S-box has a *uniformly half-occupied difference distribution table (UHODDT)*.

Intuitively, an $n \times m$ S-box with a UHODDT is expected to be useful as it seems to sit nicely in the middle of two undesirable extremes: S-boxes whose differential distribution tables contain too few non-zero entries and S-boxes whose differential distribution tables contain too many non-zero entries. At one extreme, the differential distribution tables contain high-valued entries which may be exploited by differential attacks, while at the other extreme, the differential distribution tables may be so close to a flat one that the S-box is again exploitable by differential attacks.

As we mentioned earlier, despite efforts by a number of researchers around the world, we have not witnessed the appearance of an $n \times m$ S-box with a UHODDT, except for the case of $n = m$ with n odd. This has led us to a conjecture:

Conjecture 1 *For all $n > m$, there exists no $n \times m$ S-box with a UHODDT.*

The first major step towards proving the conjecture was made in Theorem 1 of [22] for a special class of S-boxes called *quadratic S-boxes* whose algebraic degrees are two. In particular, it has been proved in [22] that *for $n \geqq 4$, there exists no quadratic $n \times m$ S-box with a UHODDT if n or m is even.*

In the next section, we provide further evidence to support the correctness of the conjecture. Before proceeding to the discussions, we would like to stress that a UHODDT is not a sufficient condition for the cryptographic usefulness of an S-box. A really "good" S-box should satisfy a set of conditions associated with all currently known cryptanalytic attacks. The size of the set seems to be expanding with advances in cryptanalytic attacks. For instance, recent progress in "high-order" differential attacks [8, 18] shows that it is desirable for an S-box to have a high algebraic degree.

4 Non-existence of Certain Quadratic S-boxes

There are a few directions one can follow to improve the result in [22]. These directions may include (1) proving the conjecture for higher-degree (say cubic) S-boxes, (2) proving the conjecture for quadratic S-boxes, but with different parameters. In what follows we report our progress in the second direction.

Theorem 1 *There exists no quadratic $n \times m$ S-box with a UHODDT when $n \geqq 2m-1$.*

Proof. Assume for contradiction that there exists a quadratic $n \times m$ S-box with a UHODDT, say $F = (f_1, \ldots, f_m)$, for $n \geqq 2m-1$. Write all the nonzero linear combination of $f_1, \ldots, f_m$ as $g_1, \ldots, g_{2^m-1}$. From the proof of Theorem 1 of [22], each nonzero vector in V_n is a linear structure of a unique g_j, i.e., there is a unique g_j such that $g_j(x) \oplus g_j(x \oplus \alpha)$ is a constant. It is easy to verify that for each $j = 1, \ldots, 2^{m-1}$, the nonzero linear structures of g_j, together with the zero vector, form a t_j-dimensional subspace of V_n for an integer t_j. We denote the subspace by W_j.

Note that

$$V_n = W_1 \cup \cdots \cup W_{2^m-1} \tag{1}$$

where

$$W_j \cap W_i = \{0\} \text{ if } j \neq i. \tag{2}$$

Thus $2^{t_1} + \cdots + 2^{t_{2^m-1}} = 2^n + 2^m - 2$ and thus there is a j_0, $1 \leqq j_0 \leqq 2^m - 1$, such that

$$2^{t_{j_0}} \geqq \frac{2^n + 2^m - 2}{2^m - 1} \geqq 2^{n-m} + 1$$

Since $2^{t_{j_0}}$ must be an integer power of 2, we conclude

$$2^{t_{j_0}} \geqq 2^{n-m+1}.$$

Now consider W_{j_0}. From linear algebra, V_n can be expressed as a partition

$$V_n = U_0 \cup U_1 \cup \cdots \cup U_{2^{n-t_{j_0}}} \tag{3}$$

satisfying

(i) $U_0 = W_{j_0}$,

(ii) $|U_j| = 2^{t_{j_0}}$,

(iii) $U_j \cap U_i = \phi$ where ϕ denotes the empty set,

(iv) two vectors α', α'' belong the same class U_j (also called a coset) if and only if $\alpha' \oplus \alpha'' \in U_0$.

Now we focus on U_1. Since $U_1 \cap U_0 = \phi$, from (1), we have

$$U_1 \subseteq (W_1 \cup \cdots \cup W_{j_0-1} \cup W_{j_0+1} \cup \cdots \cup W_{2^m-1}). \tag{4}$$

Note that $|U_1| = 2^{t_{j_0}} \geqq 2^{n-m+1}$. By the assumption, we have $n - m + 1 \geqq m$. Thus $|U_1| > 2^m - 1$. (4) implies that there is i_0, $i_0 \in \{1, \cdots, j_0 - 1, j_0 + 1, \cdots, 2^m - 1\}$, such that

$|W_{i_0} \cap U_1| \geqq 2$. Let $\alpha', \alpha'' \in W_{i_0} \cap U_1$. Since $\alpha', \alpha'' \in U_1$, from the above property (iv), we have $\alpha' \oplus \alpha'' \in U_0 = W_{j_0}$. On the other hand, since $\alpha', \alpha'' \in W_{i_0}$ and W_{i_0} is a subspace, we must have $\alpha' \oplus \alpha'' \in W_{i_0}$. This proves that

$$\alpha' \oplus \alpha'' \in W_{i_0} \cap W_{j_0}. \tag{5}$$

Since $i_0 \in \{1, \cdots, j_0 - 1, j_0 + 1, \cdots, 2^m - 1\}$, we have $i_0 \neq j_0$. This contradicts (2). □

We note that both Theorem 1 in this paper and Theorem 1 in [22] can be extended to S-boxes with partially bent component functions introduced in [6].

5 A Lower Bound on Differential Uniformity

Recall that the differential uniformity, denoted by δ, of an $n \times m$ S-box is defined as the largest value in the differential distribution table of the S-box (not counting the first entry in the first row), namely,

$$\delta = \max_{\alpha \in V_n, \alpha \neq 0} \max_{\beta \in V_s} |\{x | F(x) \oplus F(x \oplus \alpha) = \beta\}|$$

(See Definition 4). As discussed earlier, δ is bounded by $2^{n-m} \leqq \delta \leqq 2^n$, and generally speaking S-boxes with a smaller δ are desirable in designing a block cipher secure against differential attacks. This motivates us to improve the "trivial" lower bound 2^{n-m} on the differential uniformity δ.

The following lemma will be used in our discussions. It is identical to Lemma 2 of [19].

Lemma 4 *Let real valued sequences $a_0, \ldots, a_{2^n-1}$ and $b_0, \ldots, b_{2^n-1}$ satisfy*

$$(a_0, \ldots, a_{2^n-1}) H_n = (b_0, \ldots, b_{2^n-1}).$$

For any integer p and q, $p + q = n$, $1 \leqq p, q \leqq n - 1$, set $\sigma_j = \sum_{s=0}^{2^q-1} b_{j2^q+s}$, where $j = 0, 1, \ldots, 2^p - 1$. Then

$$2^q (a_0, a_{2^q}, a_{2 \cdot 2^q}, \ldots, a_{(2^p-1)2^q}) H_p = (\sigma_0, \sigma_1, \ldots, \sigma_{2^p-1}). \tag{6}$$

Now we prove the second main result of this paper.

Theorem 2 *Let $F = (f_1, \ldots, f_m)$ be an $n \times m$ S-box, where n and m are integers with $n \geqq m \geqq 1$ and each $f_j(x)$ is a function on V_n. Set $g_j = \bigoplus_{u=1}^m c_u f_u$ where $(c_1, \ldots, c_m)$ is the binary representation of integer j, $j = 0, 1, \ldots, 2^m - 1$. Denote by $\Delta_j(\alpha)$ the auto-correlation of g_j with a shift α, and set $\Delta_M = \max\{|\Delta_j(\alpha)| \mid j = 1, \ldots, 2^m - 1, \alpha \in V_n, \alpha \neq 0\}$. Then we have*

$$\delta \geqq 2^{n-m} + 2^{-m} \Delta_M.$$

Proof. Let $\Delta_{j'}(\alpha_{i'}) = \Delta_M$. By Part (ii) of Lemma 3, we have

$$2^{-m} (\Delta_0(\alpha_{i'}), \Delta_1(\alpha_{i'}), \ldots, \Delta_{2^m-1}(\alpha_{i'})) H_m = (k_0(\alpha_i'), k_1(\alpha_i'), \ldots, k_{2^m-1}(\alpha_i')) \tag{7}$$

Applying Lemma 4 to (7), we get

$$2^{m-1} 2^{-m} (\Delta_0(\alpha_{i'}), \Delta_{2^{m-1}}(\alpha_{i'})) H_1 = (\sigma_0, \sigma_1)$$

where $\sigma_j = \sum_{s=0}^{2^{m-1}-1} k_{j2^{m-1}+s}$, $j = 0, 1$. Hence

$$2^{-1}(\Delta_0(\alpha_{i'}) + \Delta_{2^{m-1}}(\alpha_{i'})) = \sigma_0$$

and

$$2^{-1}(\Delta_0(\alpha_{i'}) - \Delta_{2^{m-1}}(\alpha_{i'})) = \sigma_1$$

Thus there is a $j_0 2^q + s_0$ for $0 \leqq s_0 \leqq 2^{m-1} - 1$ and $j_0 = 0$ or 1, such that

$$k_{j_0 2^q + s_0} \geqq 2^{-m}(\Delta_0(\alpha_{i'}) + \Delta_{2^{m-1}}(\alpha_{i'})).$$

Recall that $\Delta_0(\alpha) = 2^n$ for all $\alpha \in V_n$. So we have

$$k_{j_0 2^q + s_0} \geqq 2^{-m}(2^n + \Delta_{2^{m-1}}(\alpha_{i'})).$$

According to Section 5.3 of [17], the differential uniformity of F is invariant under a nonsingular linear transformation on the variables of F. Thus by choosing an appropriate nonsingular linear transformation on the variables of F, we have

$$k_{j_0 2^q + s_0} \geqq 2^{n-m} + 2^{-m}\Delta_M$$

and hence

$$\delta \geqq 2^{n-m} + 2^{-m}\Delta_M.$$

□

When $\Delta_M = 0$, every nonzero linear combination of the components of F is a bent function. (Such S-boxes do exist [1, 12], but are not regular.) In this case we have $\delta = 2^{n-m}$. This indicates that the bound in Theorem 2 is tight for $n \times m$ S-boxes with $n \geqq m$.

6 Relating nonlinearity of S-boxes to their differential characteristics

After the discovery of differential attacks in [4], an equally notable cryptanalysis method, the linear cryptanalytic attack, was subsequently introduced in [11]. Identifying relationships between these two types of attacks has been an interesting research area, both from the view point of cryptanalysis and the design of secure ciphers. We will show in this section a relationship between the differential characteristics of an S-box and a upper bound on the nonlinearity of the S-box. The usefulness of such an explicit relationship is obvious: the nonlinearity of an S-box represents a key indicator for the strength of a block cipher that employs the S-box. We also compare our result on the relationship with a related theorem in [7].

We begin with examining the leftmost column of the difference distribution table of an S-box (not necessarily regular).

Lemma 5 *Let F be a mapping from V_n to V_m and K is the difference distribution table of F. Then the leftmost column of K is determined by a 2^m-partition of V_n, say $V_n = \Omega_0 \cup \cdots \cup \Omega_{2^m-1}$, that satisfies the condition that $\Omega_j \cap \Omega_i = \phi$ for all $j \neq i$.*

Proof. For each $\beta \in V_m$, define $\Omega_\beta = \{\alpha \in V_n | F(\alpha) = \beta\}$. Note that we use an integer in $[0, \ldots, 2^m - 1]$ and a vector in V_m interchangeably. Clearly

$$V_n = \cup_{\beta \in V_m} \Omega_\beta \tag{8}$$

and $\Omega_{\beta'} \cap \Omega_{\beta''} = \phi$ if $\beta' \neq \beta''$. Note that $F(x) \oplus F(x \oplus \alpha) = 0$ if and only if both x and $x \oplus \alpha$ belong to the same class, say Ω_β.

Now we modify the mapping F into F' by applying an arbitrary permutation on V_m to the output of F. Clearly the partition in (8) remains unchanged, and $F'(x) \oplus F'(x \oplus \alpha) = 0$ if and only if both x and $x \oplus \alpha$ belong to the same class in (8). This proves that the leftmost columns of the difference distribution tables of F and F' are the same. □

To study a $n \times m$ S-box, the two parameters n and m alone are not adequate in finding out detailed information on the S-box. On the other hand, it will be too complex to take into account all the $k_j(\alpha)$, $\Delta_j(\alpha)$, or $\langle \eta_j, \ell_i \rangle^2$, for $j = 0, 1, \ldots, 2^m - 1$, $i = 0, 1, \ldots, 2^n - 1$ and $\alpha \in V_n$. The following theorem can be viewed as a compromise between the two approaches. It relates the nonlinearity of a regular $n \times m$ S-box to three of its parameters, namely n, m and the leftmost column of its difference distribution table K.

Theorem 3 *For any regular $n \times m$ S-box F, its nonlinearity satisfies*

$$N_F \leqq 2^{n-1} - \frac{1}{2}\sqrt[4]{\frac{2^{3n+2m} - 2^{4n} + 2^{2m+n}\sum_{i=1}^{2^n-1} k_0^2(\alpha_i)}{(2^n - 1)(2^m - 1)^2}}.$$

Proof. Multiply both sides of the equality in (iii) of Theorem 3 by e^T, where e denotes the all-one sequence of length 2^m,

$$\begin{bmatrix} \sum_{j=0}^{2^m-1}\langle \eta_j, \ell_0 \rangle^2 \\ \sum_{j=0}^{2^m-1}\langle \eta_j, \ell_1 \rangle^2 \\ \vdots \\ \sum_{j=0}^{2^m-1}\langle \eta_j, \ell_{2^n-1} \rangle^2 \end{bmatrix} = H_n \begin{bmatrix} k_0(\alpha_0) & k_1(\alpha_0) & \ldots & k_{2^m-1}(\alpha_0) \\ k_0(\alpha_1) & k_1(\alpha_1) & \ldots & k_{2^m-1}(\alpha_1) \\ & \vdots & & \\ k_0(\alpha_{2^n-1}) & k_1(\alpha_{2^n-1}) & \ldots & k_{2^m-1}(\alpha_{2^n-1}) \end{bmatrix} \begin{bmatrix} 2^m \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{9}$$

Hence

$$\begin{bmatrix} \sum_{j=0}^{2^m-1}\langle \eta_j, \ell_0 \rangle^2 \\ \sum_{j=0}^{2^m-1}\langle \eta_j, \ell_1 \rangle^2 \\ \vdots \\ \sum_{j=0}^{2^m-1}\langle \eta_j, \ell_{2^n-1} \rangle^2 \end{bmatrix} = 2^m H_n \begin{bmatrix} k_0(\alpha_0) \\ k_0(\alpha_1) \\ \vdots \\ k_0(\alpha_{2^n-1}) \end{bmatrix} \tag{10}$$

Multiply the transposes of the two sides in (10), we have

$$(\sum_{j=0}^{2^m-1}\langle \eta_j, \ell_0 \rangle^2)^2 + (\sum_{j=0}^{2^m-1}\langle \eta_j, \ell_1 \rangle^2)^2 + \cdots + (\sum_{j=0}^{2^m-1}\langle \eta_j, \ell_{2^n-1} \rangle^2)^2 = 2^{2m+n}\sum_{i=0}^{2^n-1} k_0^2(\alpha_i) \tag{11}$$

Since both η_0 and ℓ_0 are an all-one sequence, we have $\langle \eta_0, \ell_0 \rangle = 2^n$.

Recall that F is regular. By Lemma 2, each nonzero linear combination of the component functions of F is balanced. Thus for $j = 1, \ldots, 2^m - 1$, η_j is $(1, -1)$ balanced and we have $\langle \eta_j, \ell_0 \rangle = 0$. Also since ℓ_j is $(1, -1)$ balanced for $j > 0$, we have $\langle \eta_0, \ell_j \rangle = 0$ for $j = 1, \ldots, 2^n - 1$.

Note that $k_0(\alpha_0) = 2^n$. So (11) can be specialized as

$$(\sum_{j=1}^{2^m-1} \langle \eta_j, \ell_1 \rangle^2)^2 + \cdots + (\sum_{j=1}^{2^m-1} \langle \eta_j, \ell_{2^n-1} \rangle^2)^2 = 2^{2m+3n} - 2^{4n} + 2^{2m+n} \sum_{i=1}^{2^n-1} k_0^2(\alpha_i) \qquad (12)$$

Thus there is a i_0, $1 \leqq i_0 \leqq 2^n - 1$, such that

$$\sum_{j=1}^{2^m-1} \langle \eta_j, \ell_{i_0} \rangle^2 \geqq \sqrt{\frac{2^{3n+2m} - 2^{4n} + 2^{2m+n} \sum_{i=1}^{2^n-1} k_0^2(\alpha_i)}{2^n - 1}}$$

Furthermore there is a j_0, $1 \leqq j_0 \leqq 2^m - 1$, such that

$$|\langle \eta_{j_0}, \ell_{i_0} \rangle| \geqq \sqrt[4]{\frac{2^{3n+2m} - 2^{4n} + 2^{2m+n} \sum_{i=1}^{2^n-1} k_0^2(\alpha_i)}{(2^n - 1)(2^m - 1)^2}}.$$

Now the theorem follows immediately from Lemma 1. □

The significance of Theorem 3 lies in its generality: it applies to all regular S-boxes that have more input bits than output bits.

Before closing this section, we note that a paper by Chabaud and Vaudenay [7] is a prior work most relevant to this research. The main result in [7] is their Theorem 4 which is equivalent to stating that for every mapping from V_n to V_m, say F, the nonlinearity of F, N_F, satisfies

$$N_F \leqq 2^{n-1} - \frac{1}{2}(3 \cdot 2^n - 2 - \frac{2(2^n - 1)(2^{n-1} - 1)}{2^m - 1})^{\frac{1}{2}}.$$

Examining the expression under the square root in the above bound, one can see that it is negative if $m \leqq n - 2$. Therefore, a condition for the validity of the theorem, which has not been spelled out in their paper, is that $m \geqq n - 1$. The same un-spelled condition of $m \geqq n - 1$ is also implied in Lemma 4 in the same paper, and hence its proof presented in the paper should be corrected.

7 Concluding Remarks

We have proved that there exists no quadratic $n \times m$ S-box with a UHODDT when $n \geqq 2m - 1$, which acts as evidence that further supports the conjecture on the non-existence of an $n \times m$ S-box with a UHODDT for all $n > m$. We have also proved a tight lower bound on the differential uniformity of an S-box, and an upper bound on the nonlinearity of an S-box that serves as an bridge between the nonlinearity and differential characteristics of the S-box.

The technique used in proving the non-existence result is essentially similar to that used in [22]. This technique, however, seems to have its limitation in that it may not be applicable to a research topic that deserves immediate attention in light of the progress made in this work, namely proving the non-existence of higher-degree S-boxes.

Acknowledgment

The first author was supported by a Queen Elizabeth II Research Fellowship (223 23 1001). Part of the second author's work was completed while on sabbatical leave at the University of Tokyo. Thanks also go to referees for SAC'97 for helpful comments.

References

[1] ADAMS, C. M. On immunity against Biham and Shamir's "differential cryptanalysis". *Information Processing Letters 41* (1992), 77–80.

[2] BETH, T., AND DING, C. On permutations against differential cryptanalysis. In *Advances in Cryptology - EUROCRYPT'93* (1994), vol. 765, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York, pp. 65–76.

[3] BIHAM, E., AND SHAMIR, A. Differential cryptanalysis of DES-like cryptosystems. *Journal of Cryptology Vol. 4, No. 1* (1991), 3–72.

[4] BIHAM, E., AND SHAMIR, A. *Differential Cryptanalysis of the Data Encryption Standard.* Springer-Verlag, New York, Heidelberg, Tokyo, 1993.

[5] BROWN, L., KWAN, M., PIEPRZYK, J., AND SEBERRY, J. Improving resistance to differential cryptanalysis and the redesign of LOKI. In *Advances in Cryptology - ASIACRYPT'91* (1993), vol. 739, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York, pp. 36–50.

[6] CARLET, C. Partially-bent functions. *Designs, Codes and Cryptography 3* (1993), 135–145.

[7] CHABAUD, F., AND VAUDENAY, S. Links between differential and linear cryptanlysis. In *Advances in Cryptology - EUROCRYPT'94* (1995), vol. 950, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York, pp. 256–265.

[8] JAKOBSEN, T., AND KNUDSEN, L. The interpolation attack on block ciphers. In *Fast Software Encryption* (Berlin, New York, Tokyo, 1997), Lecture Notes in Computer Science, Springer-Verlag.

[9] LIDL, R., AND NIEDERREITER, H. *Finite Fields, Encyclopedia of Mathematics and Its Applications.* Cambridge University Press, 1983.

[10] MACWILLIAMS, F. J., AND SLOANE, N. J. A. *The Theory of Error-Correcting Codes.* North-Holland, Amsterdam, New York, Oxford, 1978.

[11] MATSUI, M. Linear cryptanalysis method for DES cipher. In *Advances in Cryptology - EUROCRYPT'93* (1994), vol. 765, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York, pp. 386–397.

[12] NYBERG, K. Perfect nonlinear S-boxes. In *Advances in Cryptology - EUROCRYPT'91* (1991), vol. 547, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York, pp. 378–386.

[13] NYBERG, K. On the construction of highly nonlinear permutations. In *Advances in Cryptology - EUROCRYPT'92* (1993), vol. 658, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York, pp. 92–98.

[14] NYBERG, K. Differentially uniform mappings for cryptography. In *Advances in Cryptology - EUROCRYPT'93* (1994), vol. 765, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York, pp. 55–65.

[15] NYBERG, K., AND KNUDSEN, L. R. Provable security against differential cryptanalysis. In *Advances in Cryptology - CRYPTO'92* (1993), vol. 740, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York, pp. 566–574.

[16] PIEPRZYK, J. Bent permutations. In *Proceeding of the International Conference on Finite Fields, Coding Theory, and Advances in Communications and Computing* (Las Vegas, 1991).

[17] SEBERRY, J., ZHANG, X. M., AND ZHENG, Y. Systematic generation of cryptographically robust S-boxes. In *Proceedings of the first ACM Conference on Computer and Communications Security* (1993), The Association for Computing Machinery, New York, pp. 172 – 182.

[18] SHIMOYAMA, T., MORIAI, S., AND KANEKO, T. Cryptanalysis of the cipher KN, May 1997. (presented at the rump session of Eurocrypt'97).

[19] ZHANG, X. M., AND ZHENG, Y. Auto-correlations and new bounds on the nonlinearity of boolean functions. In *Advances in Cryptology - EUROCRYPT'96* (1996), vol. 1070, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York, pp. 294–306.

[20] ZHANG, X. M., AND ZHENG, Y. Characterizing the structures of cryptographic functions satisfying the propagation criterion for almost all vectors. *Design, Codes and Cryptography 7*, 1/2 (1996), 111–134. special issue dedicated to Gus Simmons.

[21] ZHANG, X. M., AND ZHENG, Y. Difference distribution table of a regular substitution box ,. In *Proceedings of the Third Annual Workshop on Selected Area in Cryptography (SAC'96)* (1996), Queen's University at Kingston, Ontario, pp. 57–60.

[22] ZHANG, X. M., AND ZHENG, Y. On the difficulty of constructing cryptographically strong substitution boxes. *Journal of Universal Computer Science 2*, 3 (1996), 147–162. (available at `http://hgiicm.tu-graz.ac.at/`).

On the Design of Linear Transformations for Substitution Permutation Encryption Networks

A.M. Youssef, S. Mister and S.E. Tavares

Department Of Electrical and Computer Engineering

Queen's University, Kingston, Ontario, Canada, K7L 3N6

E-mail: {amr_y, misters and tavares}@ee.queensu.ca

http://adonis.ee.queensu.ca:8000

Abstract— In this paper we study the security of Substitution Permutation Encryption Networks (SPNs) with randomly selected bijective substitution boxes and a randomly selected invertible linear transformation layer. In particular, our results show that for such a *64*–bit SPN using 8×8 s-boxes, the number of s-boxes involved in any 2 rounds of a linear approximation or a differential characteristic is equal to 8 with probability exceeding 0.8. For these SPNs the number of plaintext/ciphertext pairs that are required for the basic linear and differential cryptanalysis exceeds 2^{64} within 6 rounds. We also provide two construction methods for involution linear transformations based on Maximum Distance Separable Codes.

1 Introduction

Heys and Tavares [3][4][5] showed that replacing the permutation layer of Substitution Permutation encryption Networks (SPNs) with a diffusive linear transformation improves the avalanche characteristics of the cipher and increases the cipher's resistance to differential and linear cryptanalysis . Linear [8] and differential [1] cryptanalysis are two of the most powerful attacks on block ciphers. In particular it was shown [3][4] that with such a linear transformation we can develop upper bounds on the differential characteristic probability [1] and on the probability of a linear approximation [9] as a function of the number of rounds of substitution. These bounds are achieved by choosing the linear transformation in such a way that we can have a lower bound on the number of s-boxes involved in any 2 rounds of a differential characteristic or linear approximation expression. Letting N represent the block size of an SPN consisting of R rounds of $n \times n$ s-boxes (M per round), a simple example of an SPN with $N = 16,\ n = 4, M = \frac{N}{n} = 4$, and $R = 3$ is illustrated in Figure 1.

An interesting class of linear transformations is the one based on Maximum Distance Separable (MDS) codes [7]. The use of such linear transformations was first proposed by Vaudenay in [13] and then utilized in the cipher SHARK [12] and later in the cipher SQUARE [2]. This class of linear transformations has the advantage that the number of s-boxes involved in any 2 rounds of a linear approximation or in any 2 rounds of a differential characteristic is equal to $M + 1$ which is the maximum theoretically possible number.

In this paper we study the security of SPNs with randomly selected n-bit bijective substitution boxes and a randomly selected linear transformation layer over $GF(2^n)$. We also provide two construction methods for involution linear transformations based on Maximum Distance

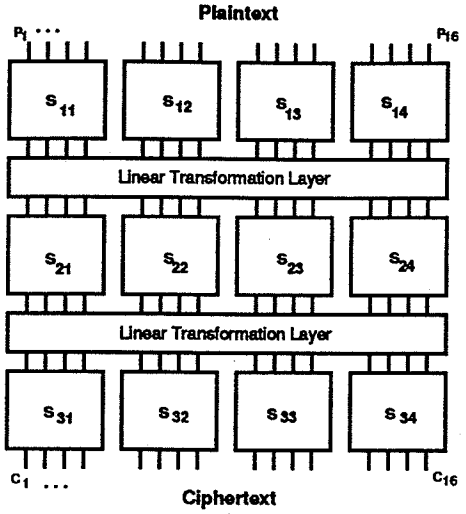

Figure 1 SPN with $N = 16$, $n = 4$, and $R = 3$.

Separable Codes. Involution linear transformations have the advantage that the resulting network can be used to perform both the encryption and the decryption operations [16].

Rijmen *et al* [12] noted that the framework of linear codes over $GF(2^n)$ provides an elegant way to construct the linear transformation layer. More details about the theory of error correcting codes can be found in [7].

Let C be a $(2M, M, d)$ code over $GF(2^n)$. Let $G = [I|A]$ be the generator matrix in echelon form where A is a nonsingular $M \times M$ matrix and I is the $M \times M$ identity matrix. Then A defines an invertible linear mapping

$$GF(2^n)^M \rightarrow GF(2^n)^M : X \rightarrow Y = AX. \tag{1}$$

If the matrix A is used in the implementation of the linear transformation of the SPN, then it is easy to see that the number of s-boxes involved in any 2 rounds of a differential characteristic or linear approximation expression is lower bounded by d, the minimum distance of the code [12]. The minimum distance of the code is equal to the minimum number of linearly dependent columns in its null matrix (also known as the parity-check matrix). For an MDS code with parameters $(2M, M, d)$, the minimum distance d is equal to $M + 1$. Throughout this paper we assume that M is an even number.

2 Randomly Selected Linear Transformations

Lemma 1

Let $G = [I|A]$ be the generator matrix of a code in echelon form where A is a randomly selected $M \times M$ nonsingular matrix and I is the $M \times M$ identity matrix with elements over $GF(q), q = 2^n$. Then the probability that this code has a minimum distance $d \geq r$, $2 \leq r \leq M + 1$, is lower

bounded by

$$\frac{1}{\Psi(M,q)}\prod_{i=1}^{M}\left(q^M-\sum_{j=0}^{r-2}\binom{M+i-1}{j}(q-1)^j-\sum_{j=r-1}^{i-1}\binom{i-1}{j}(q-1)^j\right), \tag{2}$$

where

$$\Psi(M,q)=\prod_{i=0}^{M-1}\left(q^M-q^i\right) \tag{3}$$

is the number of nonsingular $M \times M$ matrices over $GF(q)$.

Proof: If $G=[I|A]$ then the null matrix H is given by

$$H=\left[-A^T|I\right]=\left[A^T|I\right] \tag{4}$$

since we are working over $GF(2^n)$. It is clear that as A varies over all possible nonsingular matrices, A^T varies over the same set. We construct the matrix A^T column by column to meet our criterion.

The columns of A^T must not equal any linear combination of up to $r-2$ of the other columns of H, and, for A^T to be invertible, no column of A^T should be a linear combination of the other columns of A^T.

Suppose we have already assigned $i-1$ columns of A^T. We may choose any of the q^M possibilities for column i except the

$$\sum_{j=0}^{r-2}\binom{M+i-1}{j}(q-1)^j \tag{5}$$

linear combinations of up to $r-2$ of the $M+i-1$ assigned columns of H and the

$$\sum_{j=r-1}^{i-1}\binom{i-1}{j}(q-1)^j \tag{6}$$

linear combinations of known columns of A^T not counted in (5).

Note that the combinations counted in (5) and (6) may not be distinct. Thus, the number of choices available for column i is at least

$$q^M-\sum_{j=0}^{r-2}\binom{M+i-1}{j}(q-1)^j-\sum_{j=r-1}^{i-1}\binom{i-1}{j}(q-1)^j \tag{7}$$

and hence the number of choices of A is at least

$$\prod_{i=1}^{M}\left(q^M-\sum_{j=0}^{r-2}\binom{M+i-1}{j}(q-1)^j-\sum_{j=r-1}^{i-1}\binom{i-1}{j}(q-1)^j\right). \tag{8}$$

The lemma follows by dividing the expression above by the total number of nonsingular $M \times M$ matrices over $GF(q)$. □

O'Connor [11], and Youssef and Tavares [15], [14] studied the XOR distribution table and the Linear Approximation Table (LAT) properties of randomly selected bijective s-boxes. From the analysis in [11], [15] and [14] the expected value of the maximum XOR table entry of an 8×8 randomly selected bijective mapping Δ is less than or equal to 12 and the expected nonlinearity $\mathcal{NL}$ is greater than 92.

Using an approach similar to the analysis in [4], it is possible to establish upper bounds on the most likely differential characteristic and linear approximation expression using a randomly selected SPN for which the number of s-boxes involved in any 2 rounds of a differential characteristic is greater than or equal to d. The results are obtained by assuming that all the round keys are independent.

The number of chosen plaintext/ciphertext pairs required for differential cryptanalysis of an R round SPN (based on the best *characteristic* and not the best *differential* [10], [6]) may be approximated by [1], [4]

$$N_D \geq \frac{1}{(P_\delta)^\alpha}, \tag{9}$$

where $P_\delta = \frac{\Delta}{2^n}$ and

$$\alpha \geq d\left(\frac{R}{2} - 1\right) + 1. \tag{10}$$

Similarly, the number of known plaintexts required for the *basic* linear cryptanalysis (algorithm 1 in [9]) may be approximated by [4]

$$N_L \geq \frac{1}{|2^{\alpha-1} P_\epsilon^\alpha|^2} \tag{11}$$

where

$$P_\epsilon = \frac{2^{n-1} - \mathcal{NL}}{2^n}, \tag{12}$$

and

$$\alpha \geq \frac{dR}{2}. \tag{13}$$

Letting R_L and R_D denote the minimum even number of rounds required so that N_L and N_D are greater than 2^{64}, Table 1 shows R_L and R_D as a function of d for $n = 8$, $\Delta = 12$ and $\mathcal{NL} = 92$.

d	4	5	6	7	8
R_L	10	10	8	8	6
R_D	10	8	8	6	6

Table 1 R_L and R_D as a function of d ($n = 8$, $\Delta = 12$ and $\mathcal{NL} = 92$)

Table 2 shows the theoretical lower bound (equation (2)) as well as the experimental result (sample size $= 10^5$) for the probability of picking a random invertible linear transformation, with $n = M = 8$, for which d is lower bounded by r, $4 \leq r \leq 8$.

r	4	5	6	7	8
Theoretical bound (eqn. 2)	$1 - 1.58 \times 10^{-12}$	$1 - 1.51 \times 10^{-9}$	$1 - 9.78 \times 10^{-7}$	$1 - 4.66 \times 10^{-4}$	0.839
Experimental (Random)	1.0	1.0	1.0	$1 - 4.6 \times 10^{-4}$	0.844
Experimental (Involution)	1.0	1.0	1.0	$1 - 1.18 \times 10^{-3}$	0.922

Table 2 Lower Bounds for $P(d \geq r)$ for a Randomly Chosen Linear Transformation ($n = M = 8$)

3 Involution Linear Transformations based on MDS codes

In general, SPNs need two different modules for the encryption and the decryption operations. In an SPN, decryption is performed by running the data backwards through the inverse network (i.e., applying the key scheduling algorithm in reverse and using the inverse s-boxes and the inverse linear transformation layer). In [16] the authors proposed a special class of SPNs that has the advantage that the same network can be used to perform both the encryption and the decryption operations. The basic idea is to use involution substitution layers and involution linear transformations. In this section we study two construction methods for involution linear transformations based on MDS codes.

For a linear (n, k, d) code over any field, $d \leq n - k + 1$. Codes with $d = n - k + 1$ are called Maximum Distance Separable Codes, or MDS codes for short [7].

Lemma 2[7]:

An (n, k, d) code with generator matrix $G = [I|A]$, where A is a $k \times (n - k)$ matrix, is MDS if and only if every square submatrix (formed from any i rows and any i columns, for any $i = 1, 2, \cdots, min\{k, n - k\}$) of A is nonsingular.

3.1 Random Construction

One way to obtain an involution matrix A which satisfies the above constraint is to pick a random involution matrix and test it for the above constraint.

Let

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \tag{14}$$

be an $M \times M$ random matrix where A_{11}, A_{12}, A_{21} and A_{22} are nonsingular $\frac{M}{2} \times \frac{M}{2}$ matrices. An involution matrix is one which satisfies $A^2 = I$, and thus A is an involution iff

$$A_{11}A_{12} \oplus A_{12}A_{22} = 0, \tag{15}$$

$$A_{11}^2 \oplus A_{12}A_{21} = I, \tag{16}$$

$$A_{21}A_{11} \oplus A_{22}A_{21} = 0, \tag{17}$$

$$A_{21}A_{12} \oplus A_{22}^2 = I. \tag{18}$$

If we let $A_{22} = A_{11}$ then equation (15) is satisfied iff A_{11} and A_{12} commute with each other. To achieve this we let $A_{12} = A_{11}^{-1}$. For these choices of A_{12} and A_{22}, equations (16), (17) and (18) are linearly dependent with the solution $A_{21} = A_{11}^3 \oplus A_{11}$.

Thus the $M \times M$ matrix

$$A = \begin{bmatrix} A_{11} & A_{11}^{-1} \\ A_{11}^3 \oplus A_{11} & A_{11} \end{bmatrix}, \tag{19}$$

where A_{11} is a random nonsingular $\frac{M}{2} \times \frac{M}{2}$ matrix, is an involution over $GF(2^n)$.

For $n = 8$, a random search for a matrix A, with the structure in equation (19), that satisfies the condition in lemma 2, terminates within a few seconds for even values of M, $M \leq 6$. For $M = 8$ we were unable to obtain any matrix that satisfies the conditions in lemma 2 by random search. Table 2 shows the experimental results for 10^5 randomly chosen involution linear transformations in the form of equation (19) for $M = n = 8$.

3.2 Algebraic Construction

In this section we show how to obtain an involution matrix satisfying lemma 2 by a simple algebraic construction.

Lemma 3[7]:

Given $\mathrm{x}_0, \cdots, \mathrm{x}_{n-1}$, and $\mathrm{y}_0, \cdots, \mathrm{y}_{n-1}$ the matrix $A = [\mathrm{a}_{ij}], 0 \leq i, j \leq n-1$ where $\mathrm{a}_{ij} = \frac{1}{\mathrm{x}_i + \mathrm{y}_j}$ is called a Cauchy matrix. It is known that

$$det(A) = \frac{\prod\limits_{0 \leq i < j \leq n-1} (\mathrm{x}_j - \mathrm{x}_i)(\mathrm{y}_j - \mathrm{y}_i)}{\prod\limits_{0 \leq i,j \leq n-1} (\mathrm{x}_i + \mathrm{y}_j)}. \tag{20}$$

Hence, provided the $\mathbf{x}_i$ are distinct, the $\mathbf{y}_i$ are distinct, and $\mathbf{x}_i + \mathbf{y}_j \neq \mathbf{0}$ for all i, j, it follows that any square submatrix of a Cauchy matrix is nonsingular over any field.

Let

$$\begin{aligned} \mathbf{x}_i &= \mathbf{i}, \\ \mathbf{y}_i &= \mathbf{i} \oplus \mathbf{r}, \end{aligned} \tag{21}$$

where

$$\mathbf{i} = (00\cdots 0 i_\tau \cdots i_1 i_0) \in GF(2^n), \sum_{l=0}^{\tau} 2^l i_l = i, \tau = \lceil log_{2M} \rceil - 1, \tag{22}$$

and the least significant $log_2(M)$ bits of $\mathbf{r} \neq \mathbf{0}$ are zeros.

For $A^2 = H = [\mathbf{h}_{ij}]$ we have

$$\mathbf{h}_{ij} = \bigoplus_{k=0}^{M-1} \frac{1}{(\mathbf{i} \oplus \mathbf{k} \oplus \mathbf{r})(\mathbf{j} \oplus \mathbf{k} \oplus \mathbf{r})} = \begin{cases} \bigoplus_{k=0}^{M-1} \frac{1}{(\mathbf{k} \oplus \mathbf{r})^2}, & i = j \\ \mathbf{0}, & i \neq j, \end{cases} \tag{23}$$

where i, j and k are evaluated as in equation (22). Thus the matrix A will satisfy $A^2 = \mathbf{c}^2 I$, $\mathbf{c} = \bigoplus_{i=1}^{n} \mathbf{a}_{1i}^2$ over $GF(2^n)$. Dividing (division over $GF(2^n)$) each element of A by

$$\sqrt{\mathbf{c}} = \bigoplus_{k=0}^{M-1} \frac{1}{(\mathbf{k} \oplus \mathbf{r})} = \bigoplus_{i=1}^{n} \mathbf{a}_{1i}, \tag{24}$$

we obtain an involution matrix for which every square submatrix is nonsingular over $GF(2^n)$. Figure 2 shows an example for $M = n = 8$, using the irreducible polynomial 11d†.

93	13	57	*da*	58	47	*c*	1*f*
13	93	*da*	57	47	58	1*f*	*c*
57	*da*	93	13	*c*	1*f*	58	47
da	57	13	93	1*f*	*c*	47	58
58	47	*c*	1*f*	93	13	57	*da*
47	58	1*f*	*c*	13	93	*da*	57
c	1*f*	58	47	57	*da*	93	13
1*f*	*c*	47	58	*da*	57	13	93

Figure 2 Involution Linear Transformation Based on MDS Codes ($M = n = 8$, Irreducible Polynomial = 11d)

† All numbers are in hexadecimal format

Conclusions

In this paper we studied SPNs with randomly selected s-boxes and a randomly selected invertible linear transformation layer. The results of our analysis show that SPNs with good cryptographic properties can be obtained using this random construction approach. Although this random construction can be used to implement an actual cipher, the analysis in this paper was aimed to prove the robustness of the SPN model.

We also provided two construction methods for involution linear transformations based on MDS codes. Involution linear transformations have the advantage that the resulting network can be used to perform both the encryption and the decryption operations, which enhances the practical aspects of this the class of SPN ciphers.

References

[1] E. Biham and A. Shamir. Differential cryptanalysis of DES-like cryptosystems. *Journal of Cryptology, vol. 4, no. 1, pp. 3-72*, 1991.

[2] J. Daemen, L. Knudsen, and V. Rijmen. The block cipher SQUARE. *Proc. of Fast Software Encryption (4), LNCS , Springer-Verlag*, 1997.

[3] H.M. Heys and S.E. Tavares. The design of substitution-permutation networks resistant to differential and linear cryptanalysis. *Proceedings of 2nd ACM Conference on Computer and Communications Security, Fairfax, Virginia, pp. 148-155*, 1994.

[4] H.M. Heys and S.E. Tavares. The design of product ciphers resistant to differential and linear cryptanalysis. *Journal of Cryptology, Vol. 9, no. 1, pp. 1-19*, 1996.

[5] H.M. Heys and S.E. Tavares. Avalanche characteristics of substitution-permutation encryption networks. *IEEE Trans. Comp., Vol. 44, pp.1131-1139*, Sept. 1995.

[6] X. Lai, J.L. Massey, and S. Murphy. Markov ciphers and differential cryptanalysis. *Advances in Crytology: Proc. of EUROCRYPT '91, Springer-Verlag, pp.17-38*, 1992.

[7] F.J. MacWilliams and N.J.A. Sloane. *The theory of error correcting codes.* North-Holland Publishing Company, 1977.

[8] M. Matsui. The first experimental cryptanalysis of the Data Encryption Standard. *Advances in Cryptology: Proc. of CRYPTO '94, Springer-Verlag, Berlin, pp. 1-11*, 1994.

[9] M. Matsui. Linear cryptanalysis method for DES cipher. *Advances in Cryptology: Proc. of EUROCRYPT '93, Springer-Verlag, Berlin, pp. 386-397*, 1994.

[10]K. Nyberg and L.R. Knudsen. Provable security against differential cryptanalysis. *Advances in Crytology: Proc. of CRYPTO '92, Springer-Verlag, pp. 566-574*, 1993.

[11]L.J. O'Connor. On the distribution of characteristics in bijective mappings. *Advances in Cryptology: Proc. of EUROCRYPT '93, Springer-Verlag, Berlin, pp. 360-370*, 1994.

[12]V. Rijmen, J. Daemen, B. Preneel, A. Bosselaers, and E. De Win. The cipher SHARK. *Fast Software Encryption, LNCS 1039, D. Gollmann, Ed., Springer-Verlag, pp. 99-112*, 1996.

[13]S. Vaudenay. On the need for multipermutations: Cryptanalysis of MD4 and SAFER. *Proc. of Fast Software Encryption (2), LNCS 1008, Springer-Verlag, pp. 286-297*, 1995.

[14]A.M. Youssef. Ph.D. thesis, under preparation.

[15]A.M. Youssef and S.E. Tavares. Resistance of balanced s-boxes to linear and differential cryptanalysis. *Information Processing Letters, 56(1995), pp. 249-252*, 1995.

[16]A.M. Youssef, S.E. Tavares, and H.M. Heys. A new class of substitution-permutation networks. *Workshop on Selected Areas in Cryptography, SAC '96, Workshop Record, pp.132-147*, 1996.