Difference Distribution Table of a Regular Substitution Box

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: yzheng@fcit.monash.edu.au

March 26, 1997

This short paper reports an interesting property of the difference distribution table of an S-box or substitution box, which has been discovered by the authors while studying relationships between differential and other cryptographic characteristics of an S-box. Namely, an $n \times m$ S-box is regular if and only if the sum of the entries in a column in the difference distribution table of the S-box is 2^{2n-m}.

Denote by V_n the vector space of n tuples of elements from $GF(2)$. An $n \times m$ S-box is 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 from V_n to $GF(2)$ (or on V_n for short).

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.

In cryptography we are mainly concerned with *regular* S-boxes. An S-box $F = (f_1, \ldots, f_m)$ 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. It is well-known that a regular S-box can be characterized by the balance of the linear combinations of its component functions. The following is a re-statement of Corollary 7.39 of [1]:

Lemma 1 *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 $f_1, \ldots, f_m$, $f(x) = \bigoplus_{j=1}^m c_j f_j(x)$, is balanced.*

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

Definition 1 *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 introduce three matrices in the following:

Definition 2 For $F = (f_1, \ldots, f_m)$, 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.

In designing a strong S-box, many cryptographic criteria should be examined not only against component functions, but also against their linear combinations. Such criteria include those related to nonlinearity, propagation characteristics and difference distribution tables. The matrix K characterizes the differential characteristics of an S-box. The matrix D indicates the auto-correlation of all linear combinations of the component functions. While the matrix P represents the inner product between the sequence of each linear combination of the component functions and each linear sequence. P is helpful in studying the correlation immunity, as well as the nonlinearity, of each linear combination of the component functions (see [2]).

As one immediately expects, the three matrices K, D and P are closely related. In particular the following result has been proven in [4]:

Theorem 1 *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) $D = KH_m$,

(ii) $P = H_n D$,

(iii) $P = H_n K H_m$.

Using Theorem 1, we now show that a regular S-box can be completely characterized by its difference distribution table.

Corollary 1 *Let $F = (f_1, \dots, f_m)$ 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 is a function on V_n. Then F is regular if and only if the sum of a column in the difference distribution table is 2^{2n-m}, i.e., $\sum_{\alpha \in V_n} k_i(\alpha) = 2^{2n-m}$, $i = 0, 1, \dots, 2^m - 1$.*

Proof. Compare the first rows in both sides of the formula in (iii) of Theorem 1,

$$(\sum_{\alpha \in V_n} k_0(\alpha), \sum_{\alpha \in V_n} k_1(\alpha), \dots, \sum_{\alpha \in V_n} k_{2^m-1}(\alpha))H_m = (\langle \eta_0, \ell_0 \rangle^2, \langle \eta_1, \ell_0 \rangle^2, \dots, \langle \eta_{2^m-1}, \ell_0 \rangle^2). \tag{1}$$

Obviously, if $\sum_{\alpha \in V_n} k_i(\alpha) = 2^{2n-m}$, $i = 0, 1, \dots, 2^m - 1$. then $\langle \eta_1, \ell_0 \rangle^2 = \dots = \langle \eta_{2^m-1}, \ell_0 \rangle^2 = 0$. Note that ℓ_0 is the all-one sequence of length 2^n. Hence $g_1, \dots, g_{2^m-1}$ are balanced, where $g_1, \dots, g_{2^m-1}$ are defined in Theorem 1. By Lemma 1, F is regular.

Conversely, suppose F is regular. By Lemma 1, $g_1, \dots, g_{2^m-1}$ are balanced. Hence $\langle \eta_1, \ell_0 \rangle^2 = \dots = \langle \eta_{2^m-1}, \ell_0 \rangle^2 = 0$. Note that $\langle \eta_0, \ell_0 \rangle^2 = 2^{2n}$. Rewrite (1) as

$$2^m(\sum_{\alpha \in V_n} k_0(\alpha), \sum_{\alpha \in V_n} k_1(\alpha), \cdots, \sum_{\alpha \in V_n} k_{2^m-1}(\alpha)) = (2^{2n}, 0, \dots, 0)H_m.$$

This proves that $\sum_{\alpha \in V_n} k_i(\alpha) = 2^{2n-m}$, $i = 0, 1, \dots, 2^m - 1$. □

Corollary 1 has also been obtained independently by Tapia-Recillas, Daltabuit and Vega [3].

The following corollary shows the uniqueness of the first column of the difference distribution table of a regular mapping.

Corollary 2 *Let $F = (f_1, \dots, f_m)$ 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 is a function on V_n. Then F is regular if and only if the sum of the leftmost column is 2^{2n-m}, i.e., $\sum_{\alpha \in V_n} k_0(\alpha) = 2^{2n-m}$.*

Proof. Multiply both sides of the equality in (iii) of Theorem 1 by e^T where, e denotes the all-one sequence of length 2^m. Hence we have

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

and hence

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

Compare the two sides of equality (2), obtaining

$$2^m \sum_{i=0}^{2^n-1} k_0(\alpha_i) = \sum_{j=0}^{2^m-1} \langle \eta_j, \ell_0 \rangle^2. \tag{3}$$

Since g_0 is the constant zero, η_0 is the all-one sequence of length 2^n and hence $\langle \eta_0, \ell_0 \rangle^2 = 2^{2n}$.

If $\sum_{\alpha \in V_n} k_0(\alpha) = 2^{2n-m}$, hen from (3), $\langle \eta_1, \ell_0 \rangle^2 = \cdots = \langle \eta_{2^m-1}, \ell_0 \rangle^2 = 0$. Note that ℓ_0 is the all-one sequence of length 2^n. Hence $g_1, \ldots, g_{2^m-1}$ are balanced, where $g_1, \ldots, g_{2^m-1}$ are defined in Theorem 1. By Lemma 1, F is regular.

Conversely, if F is regular, then by Corollary 1 $\sum_{\alpha \in V_n} k_0(\alpha) = 2^{2n-m}$. □

From Corollaries 1 and 2, we conclude that (1) an S-box is regular, (2) the sum of the first column in its difference distribution table is 2^{2n-m}, and (3) the sum of each column in the difference distribution table is 2^{2n-m}, are all equivalent statements.

References

[1] Lidl, R., and Niederreiter, H. *Finite Fields, Encyclopedia of Mathematics and Its Applications.* Cambridge University Press, 1983.

[2] Seberry, J., Zhang, X. M., and Zheng, Y. On constructions and nonlinearity of correlation immune functions. In *Advances in Cryptology - EUROCRYPT'93* (1994), vol. 765, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York, pp. 181–199.

[3] Tapia-Recillas, H., Daltabuit, E., and Vega, G. Some results on regular mappings, 1996. (preprint).

[4] Zhang, X. M., and Zheng, Y. Relationships between differential and other cryptographic characteristics of an S-box, 1996. (submitted for publication).