Transform Domain Analysis of DES

Guang Gong and Solomon W. Golomb
Communication Sciences Institute
University of Southern California
Electrical Engineering-Systems, EEB # 500
Los Angeles, California 90089-2565
Tels. 213 740-7332 and 213 740-7333
Fax. 213 740-8729
E-mails. guanggon@milly.usc.edu and milly@mizar.usc.edu

Abstract

DES can be regarded as a nonlinear feedback shift register (NLFSR) with input. From this point of view, the tools for pseudo-random sequence analysis are applied to the S-boxes in DES. The properties of the S-boxes of DES under Fourier transform, Hadamard transform, extended Hadamard transform and Avalanche transform are investigated. Two important results about the S-boxes of DES are found. The first result is that nearly two-thirds of the total 32 functions from $GF(2^6)$ to $GF(2)$ which are associated with the 8 S-boxes of DES have the maximal linear span 63, and the other one-third have linear span greater than or equal to 57. The second result is that for all S-boxes, the distances of the S-boxes approximated by monomial functions has the same distribution as for the S-boxes approximated by linear functions. Some new criteria for the design of permutation functions for use in block cipher algorithms are discussed.

Index Terms

DES, nonlinear feedback shift register, transform domain analysis, block cipher.

1 Introduction

The Data Encryption Standard (DES) is a block cipher involving 64-bit data encryption with a 56-bit key, which was adopted by the U.S. National Institute of Standards and Technology in 1976. DES has been widely used in bank activities and Internet communications. The security of DES has been extensively investigated by many researchers [2, 15, 3, 4, 26, 6, 23, 9, 5, 20]. DES can be implemented in hardware as well as software.

In this paper, we will consider DES as a nonlinear feedback shift register (NLFSR) with input. From this point of view, we will apply the tools for pseudo-random sequence analysis to the S-boxes in DES, i.e., the feedback function when DES is regarded as a NLFSR with input. We will exhibit several new properties of the S-boxes of DES and discuss some new criteria for design of block cipher algorithms.

Concerning the analysis of shift register sequences, we have three kinds of known transforms: the Fourier transform, the Hadamard transform and the Avalanche transform. Pieprzyk [19], Nyberg [18], and Webster and Tavares [26] discussed the nonlinearity and the Strict Avalanche Criterion (SAC) of a function from Z_2^n to Z_2^m. For a function from Z_2^n to Z_2^m, the Fourier transform represents its linear span property, the Hadamard transform reflects its nonlinearity, and the Avalanche transform shows whether it satisfies SAC. In analysis of shift register sequences [7], we consider that all m-sequences are equivalent under the decimation operation for elements in a sequence. We apply this idea to approximate the S-boxes in DES, i.e., we use monomial functions instead of linear functions to approximate S-boxes. We call this transform an *extended Hadamard transform* whose definition will be given in Section 4. We found that distributions of the extended Hadamard transform spectra of all S-boxes are the same as the distributions of the Hadamard transform spectra of the S-boxes.

This paper is organized as follows. In Section 2, we present some basic concepts that will be used thoughtout this paper. In Section 3, we show the Fourier transform spectrum of the S-boxes of DES, which gives that nearly two-thirds of the total of 32 functions from $GF(2^6)$ to $GF(2)$, which are associated with the 8 S-boxes of DES, have the maximal linear span value 63. In Section 4, first we introduce the extended Hadamard transform for a function from $GF(2^n)$ to $GF(2)$, then we discuss new criteria for design of permutation functions for use in block cipher algorithms, and we present Hadamard transform spectra, extended Hadamard transform spectra and Avalanche transform spectra of the S-boxes of DES.

Remark 1 RC5, which was invented by Rivest [21] in 1994, is also a block cipher with parameters that can be easily switched into a mode of 64-bit, or 128-bit, or 256-bit data encryption. RC5 is widely used in Internet communications [24]. The security of RC5 was discussed at recent Crypto conferences [10, 12]. Until now, RC5 has been implemented in software. The approach developed in this paper can be applied to RC5, since RC5 has the same NLFSR structure as DES, only differing in their feedback functions. So, RC5 can be easily implemented in hardware in terms of its NLFSR architecture. For RC5, the feedback function is a function from a ring Z_{32} to Z_{32} (here we assume that it is in a 64-bit mode). But it can be transformed into a function from Z_2^{32} to Z_2^{32}, or equivalently, a function from $GF(2^{32})$ to $GF(2^{32})$. Thus an analysis of transform domain properties of the feedback function of RC5 can be partially done by computation. We will discuss this in a separate paper. Other block ciphers widely used in

Internet communications [22], such as IDEA [13] and SAFER K-64 [17], are different modes. They directly use a permutation function from Z_2^n toZ_2^n instead of feedback shift resister structures. But the transform spectrum analysis techniques used for DES also can be applied to them.

2 Preliminaries

In this section, we will adopt some tools from pseudo-random sequences for the analysis of functions used in the design of conventional cryptosystems.

Let $p = 2^n - 1$ and $f(x)$ be a map from a finite field $GF(2^n)$ to $GF(2)$, i.e., $f(x) : GF(2^n) \to GF(2)$. Let α be a primitive element in $GF(2^n)$. A positive integer r is a coset leader modulo p means that r is the smallest integer in the set $\{r2^i (\bmod p) | i = 0, \ldots, n-1\}$. The trace function from $GF(2^n)$ to $GF(2)$ is defined by

$$Tr_1^n(x) = x + x^2 + \cdots + x^{2^{n-1}}. \tag{1}$$

When n is clear from the context we may also write $Tr(x)$ for short. We now list the following basic facts about $f(x)$.

Fact 1 *$f(x)$ can be represented as*

$$f(x) = c + \sum_{r \in H} Tr_1^{n_r}(\gamma_r x^r), x \in GF(2^n), c \in GF(2) \tag{2}$$

where $n_r | n$, $\gamma_r \in GF(2^{n_r})$, and H is a set of coset leaders modulo p.

Definition 1 *The expression for $f(x)$ given by the formula (2) is called the Fourier transform [1] of $f(x)$. The elements of the γ_r's, together with their conjugates, $\gamma_r^{2^i}$, $i = 0, \ldots, n_r - 1$, are called the spectrum of the Fourier transform of the function $f(x)$.*

Let

$$LS(f) = | \{r2^i (\mathrm{mod} p) | \gamma_r \neq 0, i = 0, \ldots, n_r - 1\} | . \tag{3}$$

In the literature [16, 11] on the analysis of pseudo-random sequences, the value given by (3) is called the linear span of the binary sequence $\{f(\alpha^i)\}_{i \geq 0}$. We will adopt this concept for the function $f(x)$. I.e., we will also call $LS(f)$, defined by (3), the linear span of $f(x)$.

Definition 2 *The Hadamard transform of $f(x)$ is defined by [25, 1]*

$$\hat{f}(\lambda) = \sum_{x \in GF(2^n)} (-1)^{f(x) + Tr(\lambda x)}, \lambda \in GF(2^n). \tag{4}$$

Definition 3 *The Avalanche spectrum of $f(x)$ is given by*

$$D(f, \lambda) = \sum_{x \in GF(2^n)} (-1)^{f(x) + f(x+\lambda)}, \lambda \neq 0, \lambda \in GF(2^n). \tag{5}$$

The nonlinearity and the Strict Avalanche Criterion (SAC) of a function from Z_2^n to Z_2 were discussed in [18, 26, 19]. In the following, we will show the relationship between functions from Z_2^n to Z_2 and from $GF(2^n)$ to $GF(2)$.

Let $g(x) = (g_0(x), \ldots, g_{m-1}(x))$ be a function from Z_2^n to Z_2^m and

$$x = x_0 + x_1 2 + \cdots + x_{n-1}2^{n-1}, x_i \in Z_2. \tag{6}$$

In terms of the following 1-1 correspondence

$$x_0 + x_1 2 + \cdots + x_{n-1}2^{n-1} \longleftrightarrow x_0 + x_1\alpha + \cdots + x_{n-1}\alpha^{n-1} \tag{7}$$

between Z_2^n and $GF(2^n)$, each $g_i(x), i = 0, \ldots, m-1$, can be regarded as a function from $GF(2^n)$ to $GF(2)$. In this paper, we will use this map to investigate the properties of the S-boxes in DES.

Remark 2 The Hadamard transform spectrum of $f(x)$, given by definition 2, exhibits the nonlinearity of $f(x)$. The absolute value of $\hat{f}(\lambda)$ reflects the difference between agreements and disagreements of $f(x)$ and the linear function $Tr(\lambda x)$. It is difficult to find a function $f(x)$ with a constant spectrum of its Hadamard transform. (Note. Only Bent functions [18, 25] have a constant spectrum of their Hadamard transform, but Bent functions are not permutations of $GF(2^n)$ to $GF(2)$, which means that the appearance of zeros and ones is not equally likely, so they are not suitable for use in the design of cryptosystems.) Therefore, we propose that using the Hadamard transform spectrum of a function is a useful way to exhibit the nonlinearity in the behavior of the function $f(x)$.

Remark 3 The Hadamard transform spectrum and the Avalanche transform spectrum have a strong relationship [25]. The absolute value of $D(f, \lambda)$ is the difference between agreements and disagreements between $f(x)$ and $f(x + \lambda)$. The latter denotes the output under the operation of complementation of several input variables. A function $f(x)$ satisfies the Strict Avalanche Criterion if the Avalanche transform spectrum of $f(x)$ is identically zero.

3 Fourier Transform Spectrum of the S-boxes of DES

The full description of DES is given in [14]. According to the definition of a nonlinear feedback shift register (NLFSR) [7], DES can be viewed as a NLFSR with input which is a round key in Fig. 1.

For Fig. 1, (a_0, a_1) is an initial state of F, and each component, i.e., each register, holds a 32-bit number. The feedback function F is given by

$$F(x_0, x_1) = x_0 \oplus f(x_1, k_t) \tag{8}$$

where k_t is a round key in DES, $\oplus$ is bit-wise addition in $GF(2)$, and

$$f(x_1, k_t) = S(E(x_1)) \oplus k_t \tag{9}$$

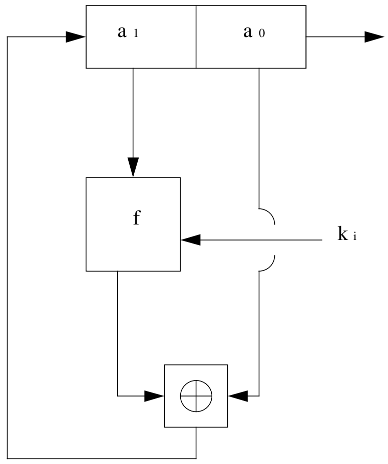

Figure 1: DES Viewed as a NLFSR with Input

where E is the E bit selection map and S is the function given by the eight S-boxes in DES. Let $\{a_i\}$ be a sequence over $GF(2^{32})$ generated by the NLFSR in Fig. 1. For a fixed DES key K, which is a 56-bit number, we have a key sequence $\{k_i\}$ where k_i is a 48-bit number and

$$k_{16i+j} = k_j, i = 0, 1, \ldots; j = 0, 1, \ldots, 15. \tag{10}$$

Therefore we have the following recursion formula:

$$a_{i+2} = a_i \oplus f(a_{i+1}, k_i), i = 0, 1, \ldots. \tag{11}$$

For a pair of a plaintext (m_0, m_1) and a ciphertext (c_0, c_1), where each of m_i and c_i is a 32-bit number, since the initial permutation IP and its inverse IP^{-1} are known, we can eliminate them. Under this assumption, each pair of a plaintext and a ciphertext can be regarded as two states of the nonlinear feedback shift register given by Fig.1; or we may say they are on the same cycle of the vector diagram of the feedback function F [7]. Namely, if (a_0, a_1) is a message, then a ciphertext corresponding to it is (a_{16}, a_{17}). Note that DES has "round 16" which corresponds to shifting 16 times in Fig. 1. We summarize the above discussion with the following proposition.

Proposition 1 *For a fixed key in DES, we have all of the following states as pairs of messages and ciphers in DES:*

$$(a_i, a_{i+1}) \longrightarrow (a_{i+16}, a_{i+17}), i = 0, 1, \ldots. \tag{12}$$

When we consider DES as a NLFSR with input, the data flow in DES for each "round" is very clear and simple. This is all shown in Fig.1.

In the following, we begin to derive the Fourier transform of the S-boxes of DES. From the construction of S-boxes in DES, we can write $f(z, k)$, given by (9), as follows:

$$f(z, k) = g(y) = (f_1(x), f_2(x), \ldots, f_8(x)), \tag{13}$$

where z and k are 48-bit numbers, $y = z \oplus k$, x is a 6-bit number, and f_i is a function from Z_2^6 to Z_2^4 which is given by the ith S-box. From Section 2, for each i, f_i can be decomposed as follows:

$$f_i(x) = (f_{i1}(x), f_{i2}(x), f_{i3}(x), f_{i4}(x)), i = 1, 2, \ldots, 8 \tag{14}$$

where f_{ij} is a function from Z_2^6 to Z_2, $i = 1, \ldots, 8$ and $j = 1, 2, 3, 4$. From the mapping given by (7), we can now consider f_{ij} as a function from $GF(2^6)$ to $GF(2)$.

Let α be a primitive element of $GF(2^6)$ with the minimal polynomial $h(x) = x^6 + x + 1$, i.e., $h(\alpha) = 0$. There are 12 coset leaders modulo 63 which are $\{1, 3, 5, 7, 9, 11, 13, 15, 21, 23, 27, 31\}$. In the following theorem, we only write out the complete forms for the Fourier transforms of S-box 1. For all other $f_{ij}, i > 1$, we only list them in the tables 2-8. Each table should be read as follows: the first row represents the exponents of x, the rows 2-5 represent the exponents of the primitive element α where these powers of α are the coefficients corresponding to the indicated terms. Notice that the trace function is Tr_1^6 for all coset leaders modulo 63 except for 9, 27 and 21. For coset leaders 9 and 27, the trace function is Tr_1^3. For coset leader 21, the trace function is Tr_1^2. The symbol ∞ means that the coefficient of the corresponding term is zero. The second column is the constant term in the Fourier transform which takes values in $\{0, 1\}$.

Theorem 1 *The algebraic forms of the feedback functions of DES, or equivalently, the Fourier transforms of $f_{ij}(x)$, are given by*

S-box 1:

$$\begin{aligned} f_{11}(x) &= Tr_1^6(\alpha^5 x) + Tr_1^6(\alpha^{22}x^3) + Tr_1^6(\alpha^6 x^5) + Tr_1^6(\alpha^{31}x^7) + Tr_1^6(\alpha^{19}x^{11}) + \\ &\quad Tr_1^6(\alpha^2 x^{13}) + Tr_1^6(\alpha^{32}x^{15}) + Tr_1^6(\alpha^{28}x^{23}) + Tr_1^6(\alpha^{18}x^{31}) + \\ &\quad Tr_1^3(\alpha^{36}x^9) + Tr_1^3(\alpha^{36}x^{27}) + Tr_1^2(\alpha^{42}x^{21}). \end{aligned}$$

$$\begin{aligned} f_{12}(x) &= 1 + Tr_1^6(\alpha^{57}x) + Tr_1^6(\alpha^5 x^3) + Tr_1^6(\alpha^3 x^5) + Tr_1^6(\alpha^{55}x^7) + \\ &\quad Tr_1^6(\alpha^{19}x^{13}) + Tr_1^6(\alpha^{13}x^{15}) + Tr_1^6(\alpha^{53}x^{23}) + Tr_1^6(\alpha^{41}x^{31}) + \\ &\quad Tr_1^3(\alpha^{36}x^9) + Tr_1^3(\alpha^{36}x^{27}) + Tr_1^2(\alpha^{42}x^{21}). \end{aligned}$$

$$\begin{aligned} f_{13(x)} &= 1 + Tr_1^6(\alpha^{24}x) + Tr_1^6(\alpha^5 x^3) + Tr_1^6(\alpha^{53}x^5) + Tr_1^6(\alpha^{58}x^7) + Tr_1^6(\alpha^{50}x^{11}) + \\ &\quad Tr_1^6(\alpha^{44}x^{13}) + Tr_1^6(\alpha^7 x^{15}) + Tr_1^6(\alpha^{24}x^{23}) + Tr_1^6(\alpha^{41}x^{31}) + \\ &\quad Tr_1^3(\alpha^{36}x^9) + Tr_1^3(\alpha^{36}x^{27}) + Tr_1^2(x^{21}). \end{aligned}$$

$$\begin{aligned} f_{14}(x) &= 1 + Tr_1^6(\alpha^3 x) + Tr_1^6(\alpha^{35}x^3) + Tr_1^6(\alpha^{16}x^5) + Tr_1^6(\alpha^{37}x^7) + Tr_1^6(\alpha^{61}x^{11}) + \\ &\quad Tr_1^6(\alpha^{44}x^{13}) + Tr_1^6(\alpha^2 x^{15}) + Tr_1^6(\alpha^{35}x^{31}) + \\ &\quad Tr_1^3(\alpha^{27}x^9) + Tr_1^3(\alpha^{45}x^{27}) + Tr_1^2(x^{21}). \end{aligned}$$

S-box 1:

	Const.	1	3	5	7	11	13	15	23	31	9	27	21
f_{11}	0	5	22	6	39	19	2	32	28	18	36	36	42
f_{12}	1	57	5	3	55	∞	19	13	53	41	36	36	42
f_{13}	1	24	5	53	58	50	44	7	24	41	36	36	0
f_{14}	1	3	35	16	37	61	44	2	64	35	27	45	0

S-box 2:

	Const.	1	3	5	7	11	13	15	23	31	9	27	21
f_{21}	1	21	22	1	54	24	56	56	60	1	0	0	21
f_{22}	1	2	18	8	45	21	13	4	6	4	27	64	42
f_{23}	1	48	30	10	16	0	2	48	12	4	45	18	0
f_{24}	1	21	19	56	56	23	29	41	21	33	45	45	42

S-box 3:

	Const.	1	3	5	7	11	13	15	23	31	9	27	21
f_{31}	1	2	14	6	58	17	61	48	24	32	18	54	21
f_{32}	0	7	0	5	7	39	58	37	33	56	18	36	42
f_{33}	1	14	20	23	9	55	5	27	13	7	18	36	42
f_{34}	0	47	38	52	51	7	4	30	10	36	45	18	42

S-box 4:

	Const.	1	3	5	7	11	13	15	23	31	9	27	21
f_{41}	1	7	33	31	15	8	58	33	36	7	45	27	42
f_{42}	1	45	11	14	39	48	53	56	13	7	18	54	21
f_{43}	1	50	54	24	20	21	56	54	8	7	54	∞	21
f_{44}	0	42	5	5	57	55	4	49	4	7	9	45	42

S-box 5:

	Const.	1	3	5	7	11	13	15	23	31	9	27	21
f_{51}	0	37	52	21	49	7	6	25	62	7	27	27	∞
f_{52}	1	61	24	33	61	16	4	21	20	41	9	54	21
f_{53}	0	31	18	19	59	47	20	16	62	32	0	∞	42
f_{54}	0	20	51	24	20	4	42	58	23	33	36	9	42

S-box 6:

	Const.	1	3	5	7	11	13	15	23	31	9	27	21
f_{61}	0	42	52	31	32	8	25	20	3	36	∞	18	42
f_{62}	0	51	2	60	53	56	3	11	29	4	0	9	∞
f_{63}	1	57	9	35	44	13	20	14	9	35	27	45	0
f_{64}	1	26	1	23	40	60	13	11	24	45	0	18	∞

S-box 7:

	Const.	1	3	5	7	11	13	15	23	31	9	27	21
f_{71}	0	0	43	38	57	15	7	54	17	2	27	27	42
f_{72}	0	7	21	40	13	3	12	33	6	18	0	27	21
f_{73}	1	23	26	29	8	52	1	30	23	33	∞	54	∞
f_{74}	0	50	43	56	61	58	61	48	7	48	9	54	21

S-box 8:

	Const.	1	3	5	7	11	13	15	23	31	9	27	21
f_{81}	1	30	60	0	53	8	43	39	22	35	27	∞	0
f_{82}	0	35	59	2	24	19	5	16	49	4	0	0	∞
f_{83}	1	44	38	13	2	17	21	2	2	32	∞	54	0
f_{84}	1	3	62	13	59	26	61	24	32	45	36	18	0

Corollary 1 *All S-box functions have linear span 63, which is the maximal value for the field* $GF(2^6)$, *except for* $LS(f_{12}) = 57$, $LS(f_{43}) = 60$, $LS(f_{51}) = 61$, $LS(f_{53}) = 60$, $LS(f_{61}) = 60$, $LS(f_{62}) = 61$, $LS(f_{64}) = 61$, $LS(f_{73}) = 58$, $LS(f_{81}) = 60$, $LS(f_{82}) = 61$, *and* $LS(f_{83}) = 60$.

4 Further Transform Domain Analysis for the S-Boxes in DES

In this section, first we introduce the extended Hadamard transform for a function from $GF(2^n)$ to $GF(2)$. Then we discuss new criteria for design of permutation functions for use in block cipher algorithms, and present the Hadamard transform spectrum, extended Hadamard transform spectrum and Avalanche transform spectrum of the S-boxes of DES.

In the analysis of shift register sequences [7], we consider that all m-sequences are equivalent under the decimation operation on elements in a sequence. We apply this idea to approximate the S-boxes in DES, i.e., we use monomial functions instead of linear functions to approximate S-boxes. We give a definition of this transform as follows.

Definition 4 *Let $f(x)$ be a function from* $GF(2^n)$ *to* $GF(2)$. *Let*

$$\hat{f}(\lambda, t) = \sum_{x \in GF(2^n)} (-1)^{f(x)+Tr(\lambda x^t)} \tag{15}$$

where $\lambda \in GF(2^n)$ and t is a coset leader modulo $2^n - 1$ coprime to $2^n - 1$. Then we call $\hat{f}(\lambda, t)$ an extended Hadamard transform of the function f.

Notice that the Hadamard transform of f, defined by (4), is $\hat{f}(\lambda, 1)$.

In the DES case, we have 6 coset leaders modulo 63 coprime to 63, which are $P = \{1, 5, 11, 13, 23, 31\}$. The numerical results show that the distribution of $\hat{f}(\lambda, t)$ in λ is invariant for each $t \in P$. This result leads to a new criterion for the design of permutation functions used in conventional cryptosystems. The reason is as follows. For example, the monomial function x^d of $GF(2^n)$, when $gcd(d, 2^n - 1) = 1$, is a permutation function on $GF(2^n)$. Notice that

$$x^d = Tr(\eta_0 x^d) + Tr(\eta_1 x^d)\alpha + \cdots + Tr(\eta_{n-1} x^d)\alpha^{n-1} \tag{16}$$

where $\{\eta_0, \eta_1, \ldots, \eta_{n-1}\}$ is the dual basis of the basis $\{1, \alpha, \ldots, \alpha^{n-1}\}$. For some choice of d [8, 19], the function x^d satisfies the Strict Avalanche Criterion. Namely, all Avalanche transform spectra of each component function $Tr(\eta_j x^d)$ are zero. At the same time, the Hadamard transform spectrum of each component function $Tr(\eta_j x^d)$ is bounded by $2^{n/2}$, which means x^d has very good nonlinearity [18]. But there exists some λ such that the extended Hadamard transform of $Tr(\eta_j x^d)$ is $\hat{Tr}(\eta_j x^d)(\lambda, t) = 2^n$. So, in this sense, x^d has a very poor approximative property when we use a monomial x^d to replace linear functions. The reason is that the functions $Tr(\eta_j x^d)$ and $Tr(\lambda x)$ have the same linear span. From the point of view of m-sequences, both of the sequences $\{Tr(\eta \alpha^{id})\}_{i \geq 0}$ and $\{Tr(\lambda \alpha^i)\}_{i \geq 0}$ are m- sequences of period $2^n - 1$. The former can be obtained from the latter by decimation d.

Let $h(x)$ be a permutation function from $GF(2^n)$ to $GF(2^n)$. Then $h(x) = \sum h_j(x)\alpha^j$, where the $h_j(x)$'s are functions from $GF(2^n)$ to $GF(2)$. If $h(x)$ is selected to use in block cipher mode, then component functions $h_j(x)$ should satisfy the following two fundamental requirements:

a For each j, $0 \leq n < n$, h_j has lower Hadamard transform spectrum, which corresponds to high nonlinearity.

b For each j, $0 \leq n < n$, h_j has lower Avalanche transform spectrum (The ideal case is for h_j to satisfy SAC, i.e., the Avalanche transform spectrum of h_j is equal to zero.)

Regarding the above consideration, we develop the following two criteria for block cipher design, which together with the two fundamental requirements listed above serve as a set of criteria for the design of "good" block cipher algorithms.

c For each j, $0 \leq n < n$, h_j has a large linear span.

d For a fixed j, $0 \leq n < n$, the spectrum distribution of the extended Hadamard transform $\hat{h}_j(\lambda, t)$ of $h_j(x)$ is the same as for the Hadamard transform $\hat{h}_j(\lambda)$ of $h_j(x)$ for all coset leaders $t \neq 1$ modulo $2^n - 1$ satisfying $(t, 2^n - 1) = 1$. Combined with Criterion a, this requires that h_j has high distances from monomial functions $Tr(\lambda x^t)$, $x \in GF(2^n)$ as well as from linear functions.

Next we will show the distribution properties of the Hadamard transform spectrum and the extended Hadamard transform spectrum in the following theorem. The possible values for $\hat{f}_{ij}(\lambda, t)$ are

$$S = \{0, 4, -4, 8, -8, 12, -12, 16, -16, 20, -20, 24, -24, 28, -28\}.$$

We now write S as $S = \{s_k | k = 0, 1, \ldots, 14\}$. Let

$$| \{\lambda \in GF(2^6) | \hat{f}_{ij}(\lambda) = s_k\} | = v_k, k = 0, 1, \ldots, 14. \tag{17}$$

Theorem 2 *For each component function f_{ij} of S-box i, the distribution of $\hat{f}_{ij}(\lambda, t)$ in λ does not depend on t. I.e., for a fixed pair i and j,*

$$| \{\lambda \in GF(2^6) | \hat{f}_{ij}(\lambda, t) = s_k\} | = v_k, k = 0, 1, \ldots, 14, \tag{18}$$

for each $t \in P = \{1, 5, 11, 13, 23, 31\}$,

For the complete data on these distributions, see Appendix 1.

Theorem 3 *For the Avalanche transform spectrum of f_{4j}, we have the following new patterns. For all $\lambda \in GF(2^6)^*$, the cyclic group of $GF(2^n)$,*

$$\begin{aligned} D(f_{41}, \lambda) &= D(f_{44}, \lambda), \\ D(f_{42}, \lambda) &= D(f_{43}, \lambda), \\ |D(f_{41}, \lambda)| &= |D(f_{42}, \lambda)|. \end{aligned}$$

The complete data for the distributions of the Avalanche transform spectra of S-boxes is given in Appendix 2.

5 Conclusions

DES can be regarded as a nonlinear feedback shift register (NLFSR) with input. From this point of view, the tools for pseudorandom sequence analysis are applied to the S-boxes in DES. We introduced a new transform, called an extended Hadamard transform, for a function from $GF(2^n)$ to $GF(2)$. We investigated the properties of the S-boxes of DES under Fourier transform, Hadamard transform, extended Hadamard transform and Avalanche transform. There were two important results. One is that for the total of 32 functions from $GF(2^6)$ to $GF(2)$, which are associated with the 8 S-boxes of DES, nearly two-thirds of them have the maximal linear span 63 and the other one-third have linear spans greater than or equal to 57. The other is that for each of these 32 functions, the distributions of the extended Hadamard transform spectra are the same as for the Hadamard transform spectra of that function, which does not depend on the specific monomial function associated with the extended Hadamard transform.

Considering the decimation property of m-sequences, we propose that two new criteria should be considered for the design of block cipher algorithms: larger linear span for each component function, and the same spectral distribution for all extended Hadamard transforms as for the Hadamard transforms. Together with the requirement of high nonlinearity, the latter corresponds to each component function having high distance whenever it is approximated by monomial functions or linear functions. The S-boxes of DES satisfy these two criteria. For some d, the power function x^d over $GF(2^n)$ is a counter-example in that it has high nonlinearity and lower distance when it is approximated by a monomial function.

Appendix 1. Distribution of the Hadamard Transform Spectra of S-boxes

For each table, the first row lists all possible values of $\hat{f}_{ij}(\lambda), j = 1, 2, 3, 4$. The following four rows are frequencies for four components in each S-box. For example, in table 1 of S-box 1, the second row should be read as follows: $\mid \{\lambda \in GF(2^6)^* | \hat{f}_{11}(\lambda) = 0\} \mid = 13$; $\mid \{\lambda \in GF(2^6)^* | \hat{f}_{11}(\lambda) = 4\} \mid = 15$; $\mid \{\lambda \in GF(2^6)^* | \hat{f}_{11}(\lambda) = -4\} \mid = 10$; etc..

S-box 1:

0	4	-4	8	-8	12	-12	16	-16	20	-20	-24	-28
13	15	10	10	6	4	0	2	0	0	2	0	1
16	9	15	4	7	4	2	1	2	1	1	1	0
13	13	11	4	8	4	3	2	4	0	1	0	0
11	10	15	11	8	1	4	0	0	1	0	1	1

S-box 2:

0	4	-4	8	-8	12	-12	16	-16	20	-20	-24	-28
19	8	14	2	6	1	6	3	1	2	1	0	0
16	13	11	4	7	1	5	1	2	2	0	1	0
21	12	12	4	4	3	2	1	1	0	1	0	2
21	11	10	2	6	4	4	2	0	0	2	0	1

S-box 3:

0	4	-4	8	-8	12	-12	16	-16	20	-20	24	-24	28	-28
19	11	12	4	4	4	3	4	0	1	0	0	0	0	1
15	11	11	6	6	3	5	2	2	0	2	0	0	0	0
17	16	10	5	5	2	2	1	1	2	0	1	1	0	0
13	9	18	5	10	1	2	1	1	1	0	0	1	1	0

S-box 4:

0	4	-4	8	-8	16	-16	20	-20
11	11	17	10	6	1	3	1	3
11	11	17	6	10	3	1	1	3
11	11	17	6	10	3	1	1	3
11	17	11	6	10	3	1	3	1

S-box 5:

0	4	-4	8	-8	12	-12	16	-16	20	-20	24	-24	28
11	14	8	10	6	7	2	1	3	0	1	0	0	0
16	7	15	6	5	4	5	2	1	0	1	0	1	0
21	6	13	4	4	9	2	1	1	0	1	0	0	1
15	9	13	10	5	5	2	0	0	1	2	1	0	0

S-box 6:

0	4	-4	8	-8	12	-12	16	-16	20	-20	24	-24
16	14	10	8	3	4	2	3	0	0	2	0	1
19	12	10	5	2	6	3	2	2	1	0	0	1
12	13	13	6	9	2	2	1	2	1	1	0	1
17	11	15	3	7	1	3	0	2	2	0	1	1

S-box 7:

0	4	-4	8	-8	12	-12	16	-16	20	24	-24	28
19	13	10	5	5	5	3	0	0	0	1	1	1
17	12	12	4	4	4	2	3	3	2	0	0	0
15	18	11	6	6	0	1	0	4	1	0	0	0
20	13	9	4	2	7	3	1	2	0	1	1	0

S-box 8:

0	4	-4	8	-8	12	-12	16	-16	20	-20	24	-24
14	9	15	5	7	3	3	3	2	0	2	0	0
18	12	10	5	6	5	2	0	1	2	1	1	0
18	12	10	6	5	3	4	1	0	0	3	0	1
15	6	16	8	4	3	5	1	3	1	1	0	0

Appendix 2. Distribution of the Avalanche Transform Spectra of S-boxes

For each table, the first row lists all possible values of $D(f_{ij}, \lambda), j = 1, 2, 3, 4$. The following four rows are frequencies for four components in each S-box. For example, in table 1 of S-box 1, the second row should be read as follows $|\{\lambda \in GF(2^6)^* | D(f_{11}, \lambda) = 0\}| = 13$; $|\{\lambda \in GF(2^6)^* | D(f_{11}, \lambda) = 8\}| = 11$; $|\{\lambda \in GF(2^6)^* | D(f_{11}, \lambda) = -8\}| = 12$; etc..

S-box 1:

0	8	-8	16	-16	24	-24	32	-32	40
13	11	12	5	7	5	6	2	2	0
19	11	13	4	7	1	4	2	1	1
17	12	16	6	8	2	2	0	0	0
16	8	9	6	8	7	4	1	4	0

S-box 2:

0	8	-8	16	-16	24	-24	32	-32	40	-40
19	9	10	9	12	1	1	1	0	0	1
11	13	18	7	4	4	3	0	3	0	0
4	16	16	4	4	6	5	2	3	1	1
12	12	11	8	8	3	3	2	3	0	1

S-box 3:

0	8	-8	16	-16	24	-24	32	-32	-40	-48
13	10	11	11	11	1	0	2	3	0	1
19	15	16	4	2	2	3	0	2	0	0
20	12	11	2	3	5	4	2	3	0	1
17	11	10	4	6	3	4	4	2	2	0

S-box 4:

0	8	-8	16	-16	24	-24
23	8	4	6	12	5	5
23	8	4	6	12	5	5
23	8	4	6	12	5	5
23	8	4	6	12	5	5

S-box 5:

0	8	-8	16	-16	24	-24	32	-32	-40
19	13	15	7	7	0	2	0	0	0
19	9	8	10	11	2	3	0	1	0
16	8	10	8	8	6	5	0	1	1
17	7	12	9	9	3	4	1	1	0

S-box 6:

0	8	-8	16	-16	24	-24	32	-32	-40	48
19	10	14	6	5	3	2	1	2	1	0
13	12	12	10	10	2	3	0	0	1	0
18	15	11	5	4	2	4	1	3	0	0
20	10	13	2	3	5	4	1	4	0	1

S-box 7:

0	8	-8	16	-16	24	-24	32	-32	40	-40	-48
9	16	11	4	8	2	3	5	4	0	0	1
21	11	13	5	5	2	4	1	1	0	0	0
0	10	6	8	12	6	8	6	4	1	1	1
9	15	11	9	10	2	4	1	2	0	0	0

S-box 8:

0	8	-8	16	-16	24	-24	32	-32	-40
23	11	10	6	8	2	2	0	0	1
11	14	11	10	10	1	3	1	1	1
9	17	14	7	7	2	4	1	1	1
22	15	11	1	9	1	1	2	1	0

References

[1] R.E. Blahut, *Theory and Practice of Error Control Codes*, Addison-Wesley Publishing Company, 1983.

[2] E. Biham and A. Shamir, Differential cryptanalysis of the full 16-round DES, *Advances in Cryptology, Proceedings of Crypt'92*, pp. 487-496, Lecture Notes in Computer Science, Vol. 740, Springer-Verlag, Berlin, 1992.

[3] E. Biham, On Matsui's linear cryptanalysis, *Advances in Cryptology, Proceedings of EuroCrypt'94*, pp. 341-355, Lecture Notes in Computer Science, Vol. 950, Springer-Verlag, Berlin, 1994.

[4] D. Davies and S. Murphy, Pairs and triples of DES S-boxes, *Journal of Cryptology*, vol. 8, No. 1, pp.1-25, 1995.

[5] Y. Desmedt, Jeam-Jacques Quisquater and M. Davio, Dependence of putput on input in DES: small Acalaance characteritics, *Advances in Cryptology, Proceedings of Crypt'84*, pp. 359-375, Lecture Notes in Computer Science, Vol. 196, Springer-Verlag, 1984.

[6] W. Diffie and M.E. Hellman, Exhaustive cryptanalysis of the NBS data encryption standard, *Computer*, Vol. 10, No. 6, pp. 74-84, June 1977.

[7] S.W. Golomb, *Shift Register Sequences*, Revised Edition, Aegean Park Press, 1982.

[8] G. Gong and Z.D. Dai, Construction of SAC permutations, *Systems Science and Mathematical Sciences*, Vol. 10, No.2, pp. 120-128, April 1997.

[9] J.A. Gorden and H. Retkin, Are big S-boxes best? *Advances in Cryptology, Proceedings of Crypt'82.*

[10] B.S. Kaliski and Y.L. Yin, On differential and linear cryptanalysis of RC5 encryption algorithm, *Advances in Cryptology, Proceedings of Crypt'95*, pp. 171-183, Lecture Notes in Computer Science, Springer-Verlag, 1995.

[11] E.L. Key, An analysis of the structure and complexity of nonlinear binary sequence generators, *IEEE Trans. on Inform. Theory*, Vol. IT-22, No.6, pp.732-736, November 1976.

[12] L.R. Knudsen and W. Meier, Improved differential attacks on RC5, *Advances in Cryptology, Proceedings of Crypt'96*, pp. 216-228, Lecture Notes in Computer Science, Vol. 1109, Springer-Verlag, Berlin, 1996.

[13] X. Lai and J. Massey, A proposal for a new block encryption standard, *Advances in Cryptology, Proceedings of EuroCrypto'90.*

[14] National Bureau of Standards, Data Encryption Standard, FIPS Publication 46, U.S. Department of Commerce, 1977.

[15] M. Matsui, Linear cryptanalysis method for DES cipher, *Advances in Cryptology, Proceedings of EuroCrypto'93*, pp. 386-397, Lecture Notes in Computer Science, Vol. 765, Springer-Verlag, Berlin, 1993.

[16] J.L. Massey, Shift-register synthesis and BCH decoding, *IEEE Trans. on Inform. Theory*, Vol. IT-15, January 1969.

[17] J.L.Massey, "SAFER K-64: A byte-oriented block-ciphering algorithm", *Proceedings of Fast Software Encryption* (Ed. R. Anderson), Lecture Notes in Computer Science No. 809, pp.1-17. New York Springer, 1994

[18] K. Nyberg, Differentially uniform mappings for cryptography, *Advances in Cryptology, Proceedings of EuroCrypt'93*, Lecture Notes in Computer Science, 1993.

[19] J.P. Pieprzyk, Nonlinearity of exponent permutations, *Advances in Cryptology, Proceedings of EuroCrypt'89*, pp. 80-92, Lecture Notes in Computer Science, 1989.

[20] I.A. Reeds and J.L. Manferdelli, DES has no per round factors, *Advances in Cryptology, Proceedings of Crypt'84*, pp. 377-389, Lecture Notes in Computer Science, Vol. 196, Springer-Verlag, 1984.

[21] R.L. Rivest, The RC5 encryption algorithm, *Proceedings of the Workshop on Cryptographic Algorithms*, K.U. Leuven, December 1994.

[22] W. Stallings, *Network and Internetwork Security*, IEEE Press, Inc., New York, 1995.

[23] I. Schaumüller-Bichl, Cryptanalysis of the Data Encryption Standard by the method of formal coding, *Advances in Cryptology, Proceedings of Crypt'82*, pp. 235-255.

[24] B. Schneier, *Applied Cryptography*, John Wiley, revised version, 1996.

[25] M.K. Simon, J.K. Omura, R.A. Scholtz and B.K. Levitt, *Spread Spectrum Communications Handbook*, McGraw-Hill, Inc. , Revised version, 1994, Computer Science Press, 1985.

[26] A.F. Webster and S.E. Tavares, On the design of S-boxes, *Advances in Cryptology, Proceedings of Crypt'85*, pp. 523-534, Lecture Notes in Computer Science, Vol. 218, Springer-Verlag, 1986.