Journal of Mathematical Sciences: Advances and Applications
Volume 49, 2018, Pages 29-50
Available at http://scientificadvances.co.in
DOI: http://dx.doi.org/10.18642/jmsaa_7100121860

DES-LIKE CIPHERS, DIFFERENTIAL ATTACKS AND APN FUNCTIONS

MOISÉS DELGADO[1], ROBERTO REYES[2] and CARLOS AGRINSONI[1]

[1]Department of Mathematics
University of Puerto Rico
Cayey Campus
San Juan PR
USA
e-mail: moises.delgado@upr.edu
carlos.agrinsoni@upr.edu

[2]Department of Mathematics
University of Puerto Rico
Ponce Campus
San Juan PR
USA
e-mail: roberto.reyes@upr.edu

Abstract

Special types of high nonlinear functions (APN functions) defined over finite fields of characteristic 2 have important applications in cryptography. The design of block ciphers by using APN functions provides high resistance against differential attacks. The goal of this paper is to supply a comprehensive review of the most important facts connecting differential attacks, DES-like ciphers

2010 Mathematics Subject Classification: 11T06, 11T55, 11T71.
Keywords and phrases: DES-like cipher, *S*-box, *s*-round characteristic, *s*-round differential, differential attack, almost perfect nonlinear, Markov cipher.
Received July 24, 2017

and APN functions. In this paper, we show how differential attacks work against DES-like ciphers and how APN functions work against differential attacks.

1. Introduction

Cryptographic algorithms are sequences of processes used to encipher and decipher messages. These algorithms allow two parties to communicate while preventing unauthorized third parties from intersecting the message. Encryption transforms human readable data (plaintext) into unreadable data, known as ciphertext.

Symmetric-key ciphers consist of encryption methods in which both parties share the same key or related keys. Symmetric-key ciphers are implemented as either block ciphers or stream ciphers. A block cipher enciphers the data as blocks, while a stream cipher do it as individual characters. A block cipher consists of encryption and decryption algorithms, the decryption algorithm is defined to be the inverse process of the encryption. Examples of symmetric ciphers include Twofish, Serpent, DES, 3DES, AES, DES-like, etc. We will focus our study on DES-like ciphers, which are block ciphers with a similar structure of DES.

Data encryption standard (DES) algorithm is a symmetric-key cipher selected in 1977 by the National Bureau of Standards as an official Federal Information Processing Standard for the USA and used world wide. *S*-Boxes are basic components of block ciphers, they are typically used to obscure the relationship between the key and the inputs (plaintext). They can take an n-bit block input and produce an n-bit block output (ciphertext). For more information about DES and its properties, see [4].

Nowadays DES is considered insecure by several applications because of its small 56-bit key size and many attempts to increase the security have failed. DES has been extensively analyzed in order to

capture its property of weakness. Special attention has been focused on the nonlinear properties of the round function F, which is composed of permutations and substitutions transformations (S-boxes). As observed by Nyberg and Knudsen [23, 24], the security of the cipher can be increased by replacing the round function F by a function with high nonlinear properties, a function that provides resistance against differential cryptanalysis.

During the application of differential cryptanalysis, the attacker selects inputs and examines outputs in an attempt to derive the key. The attacker will select pairs of inputs x, x', satisfying $x - x' = \Delta x$, knowing that for this value Δx, a particular value $y - y' = \Delta y$ will occurs with high probability. As higher this probability is, higher is the possibility to reveal the secret key.

Almost perfect nonlinear (APN) functions provide high resistance against differential cryptanalysis when are used as alternatives S-boxes of DES-like ciphers. APN functions can be chosen in such a way that for any difference Δx, the probability of the occurrence of Δy is as small as possible.

In this paper, we explain how DES-like ciphers are vulnerable to differential attacks, and how APN functions provide high resistance against differential attacks.

This paper is a compilation of the main results obtained by Biham, Shamir, Lay, Massey, Murphy, Nyberg, Knudsen, Budaghyan, Dobbertin, among others, whose works relates DES-like ciphers, differential cryptanalysis and almost perfect nonlinear functions. Our motivation for this paper is to show attractively, in just one document, one of the most important applications of high nonlinear functions defined over finite fields in cryptography, in order to be understandable for more diverse readers, others without the expertise in this area.

2. DES-like Ciphers

An r-round iterated cipher is a cryptographic algorithm based on iterating r times a function F (each iteration is called a round). The function F (called the round function) is applied to a plaintext x and a round subkey k. The round function is such that, for every pair x and k, F establishes a one to one correspondence between the round input x and the round output X.

DES-like ciphers are iterated ciphers. The process of encryption of a DES-like cipher is essentially as follows. The input x in each round is divided into two halves (x_L, x_R). F applies to the right half x_R and a round key k derived from a key schedule algorithm. The output $F(x_R, k)$ is added modulo 2 to the left half x_L and the two halves are swapped. This process, for a DES-like block cipher of r rounds, block size $2n$ and round function F, can be described as follows:

Let $(\mathbb{F}_2)^n$ be the finite field of 2^n elements. For $m \geq n$, let

$$f : (\mathbb{F}_2)^m \to (\mathbb{F}_2)^n,$$

$$E : (\mathbb{F}_2)^n \to (\mathbb{F}_2)^m,$$

and let $K = (k_1, k_2, \ldots, k_r)$, $k_i \in (\mathbb{F}_2)^m$, be the r-round key. The round function F,

$$F : (\mathbb{F}_2)^n \times (\mathbb{F}_2)^m \to (\mathbb{F}_2)^n,$$

is defined, in the i-th round, by

$$F(x, k_i) = f(E(x) + k_i), \tag{1}$$

where + is the bitwise addition modulo 2 (key mixing operation), E is an affine mapping (extension affine mapping from n-bit inputs to m-bit outputs) and f is a substitution mapping (8 parallel nonlinear substitutions called S-boxes).

The Figure 1 illustrates the operation of the round function F of DES, with a block of size 64.

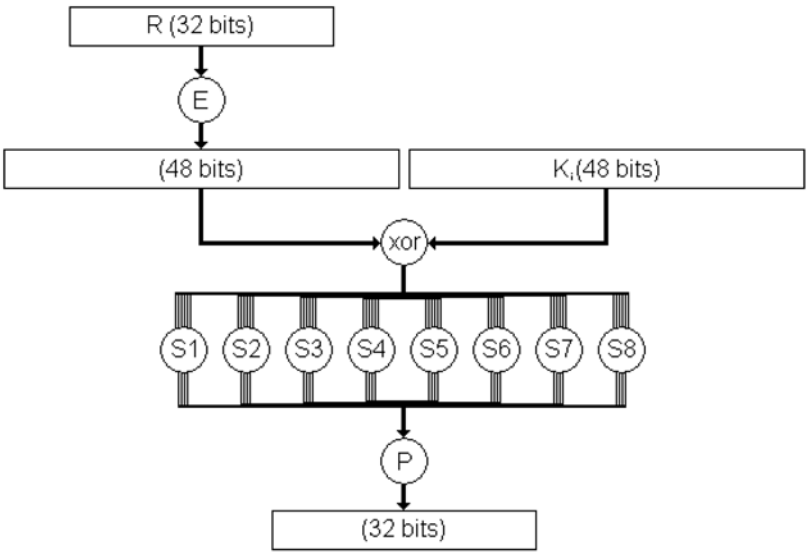

Figure 1. The round function F in DES [27].

Given a plaintext $x = (x_L, x_R)$ and a key $k = (k_1, k_2, \ldots, k_r)$, the ciphertext $X = (X_L, X_R)$ is computed in r rounds as follows:

Set $x_L(0) = x_L$ and $x_R(0) = x_R$. Compute for $i = 1, 2, \ldots, r$:

$$x_L(i) = x_R(i-1),$$

$$x_R(i) = F(x_R(i-1), k_i) + x_L(i-1),$$

$$x(i) = (x_L(i), x_R(i)).$$

Set $X_L = x_R(r)$, $X_R = x_L(r)$.

The Figure 2 illustrates the encryption process for one round. In the next chapter, we will show that, in a difference sense, DES-like ciphers are related with homogeneous Markov chains, and consequently they are vulnerable to differential cryptanalysis.

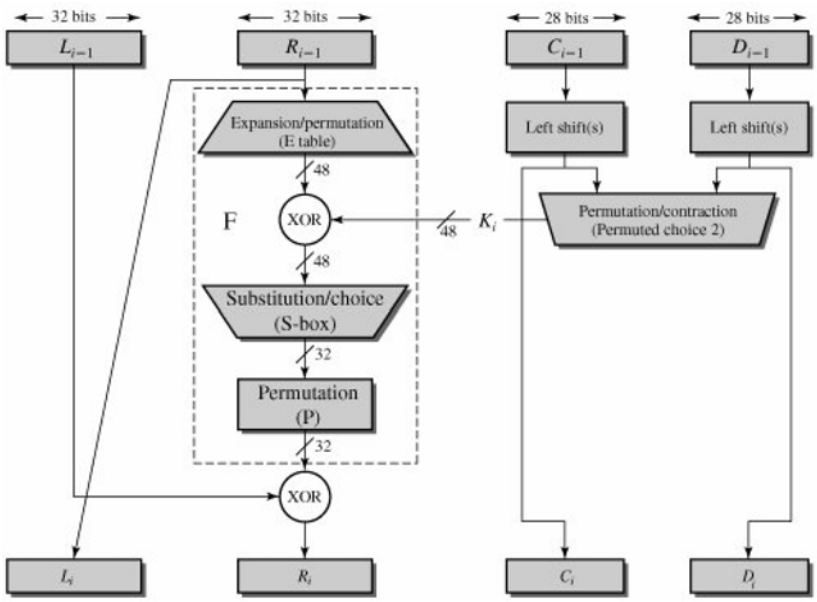

Figure 2. One round in DES [27].

3. Differential Cryptanalysis

3.1. Markov ciphers

A sequence of discrete random variables $a_0, a_1, a_2, \ldots$ is a Markov chain if, for any i, the probability of the occurrence of the state a_{i+1} is given by:

$$P(a_{i+1} = b_{i+1} | a_i = b_i, a_{i-1} = b_{i-1}, \ldots, a_0 = b_0) = P(a_{i+1} = b_{i+1} | a_i = b_i).$$

As can be seen, the probability of the next state depends only on the present state and not on the previous. A Markov chain is called homogeneous if $P(a_{i+1} = b_{i+1} | a_i = b_i)$ is independent of i, for all b_{i+1}, b_i.

Suppose that, in a DES-like cipher, the plaintext x is independent of the subkeys k_i. The following definition and theorem, given in [22], are important for introducing differential attacks on DES-like ciphers in the next subsection. (Let us use Δ to denote the difference of two variables, let us say $\Delta a = a' - a''$.)

Definition 1. An iterative cipher with round function $F = f(x, k)$ is a Markov cipher if there exist a group operation such that, for all choices of α and $\beta (\alpha \neq 0$ and $\beta \neq 0)$, the probability $P(\Delta y = \beta | \Delta x = \alpha, x = \gamma)$ is independent of γ when the round subkey k is uniformly random, i.e.;

$$P(\Delta y = \beta | \Delta x = \alpha, x = \gamma) = P(\Delta y = \beta | \Delta x = \alpha),$$

for all choices of γ when k is uniformly random.

Theorem 1. *If an r-round iterated cipher is a Markov cipher and the r round keys are independent and uniformly random, then the sequence of differences*

$$\Delta y(0), \ldots, \Delta y(r)$$

is an homogeneous Markov chain.

In [1], Biham and Shamir showed that DES-like ciphers are Markov ciphers with the addition modulo 2 bitwise operation.

3.2. *S*-round differentials and differential attacks

To start defining differential cryptanalysis on an r-round DES-like cipher, let us consider for encryption a pair of distinct plaintext x, x' of n bits, and its nonzero difference Δx defined as

$$\Delta x = x - x'.$$

For $i = 0, 1, \ldots, r,$ let

$$y(i) = F(x, k_i),$$

$$y'(i) = F(x', k_i),$$

be the outputs after i-rounds, which are also inputs to the $(i + 1)$-round.

From the above formulas, we obtain the sequence of differences

$$\Delta y(0), \ldots, \Delta y(r),$$

where $\Delta y(i) = y(i) - y'(i)$ and $y(0) = x, y'(0) = x'$.

In [1], the authors introduced differential cryptanalysis of DES-like ciphers in terms of "s-round characteristics".

An s-round characteristic is a $(s+1)$-tuple $(\beta(0), \beta(1), \ldots, \beta(s))$ considered as the possible value of $(\Delta x, \Delta y(1), \ldots, \Delta y(s))$. The probability of an s-round characteristic is defined as

$$P(\Delta y(1) = \beta(1), \Delta y(2) = \beta(2), \ldots, \Delta y(s) = \beta(s) | \Delta x = \beta(0)),$$

where the plaintext x and the subkeys k_i are independent and uniformly random. Differential cryptanalysis seeks to exploit the fact that, given a particular input difference Δx, a particular output difference $\Delta y(s)$ occurs with a very high probability (much greater than $1/2^{n-1}$), computed as a probability of Markov chains as in Equation (4), in order to reveal the secret key. If this fact is feasible for a few number of Δx, Δy, the round function F is said to be cryptographically weak.

In [22], Lai et al. introduced the notion of "s-round differentials" for a general iterated cipher instead of s-round characteristics. They observed that for the success of differential cryptanalysis it may not be necessary to fix the values of differences for the intermediate rounds in a characteristic, because only the values of $\Delta y(s-1)$ are used to determine the possible values of the subkeys in the last round. An s-round differential is a pair $(\beta(0), \beta(s))$ considered as the possible value of $(\Delta x, \Delta y(s))$. The probability of an s-round differential $(\beta(0), \beta(s))$, denoted as

$$P(\Delta y(s) = \beta(s) | \Delta x(0) = \beta(0)),$$

is the conditional probability that, after s rounds, $\beta(s)$ equals the difference $\Delta y(s)$ given that the plaintext difference Δx has difference $\beta(0)$, when the plaintext x and the subkeys k_i are independent and uniformly random.

Then, a differential attack will be succesfull if the round function of an iterated cipher is cryptographically weak and the attacker counts on s-round differentials with high probability. This is the case for DES-like ciphers.

3.3. How differential cryptanalysis works against DES-like ciphers

Differential cryptanalysis is the first method which reduce the complexity of attacking DES in half of an exhaustive search [1].

Assume we have two inputs x, x' for the same s-box S and we know only their difference Δx. Question: What do we know about the difference of their images $\Delta S(x) = S(x) - S(x')$?. A difference distribution table for this s-box is a distribution table for the input-output $(\Delta x, \Delta S(x))$ for all possible inputs x, x' such that $x + x' = \Delta x$ (see [1] for the distribution table of all s-boxes of DES). A number n_s in the position (α, β) means that for n_s pairs with input difference α, the output difference is β. The minimum value of n_s is 0, the maximum value is 64 (when $\alpha = 0, \beta = 0$). When $n_s > 0$ occurs in the position (α, β) means that the probability that α may cause β is given by $p_s = \frac{n_s}{64}$. The probability of a 1-round differential (α, β) is the product of the probabilities p_s of the 8 s-boxes. This probability equals the fraction of the key that can be extracted from that differential [1]. The probability of an s-round differential is the sum of the probabilities of s-round characteristics as in Equation (4). The process to extract the key of an r-round DES-like cipher, by using differential attacks, can be summarized as follows:

(1) Find a $(r-1)$-differential (α, β) with maximal probability or close to it.

(2) Choose pairs of plaintexts x, x' such that $x + x' = \alpha$ and submit them to encryption under the actual key K to obtain ciphertexts $y(r)$, $y'(r)$.

(3) Using these pairs $y(r)$, $y'(r)$, find the possible values of the subkeys in the last round (if any) that correspond with the previous $y(r-1)$, $y'(r-1)$ with difference $\Delta y(r-1) = \beta$.

(4) Count the number of appearances of the subkeys in Step (3).

(5) Repeat Steps (3) and (4) until a subkey (or a set of subkeys) occurs more likely than the others and take it as the key of the DES.

Thus, to prove resistance against differential cryptanalysis we must ensure that there is not differentials with probability high enough to enable successful attacks. It must be ensured a low probability for any differential. Next, an upper bound to the probability of an s-round differential is provided.

3.4. An upper bound

Assuming that the chosen plaintext x, x' (chosen by the cryptanalyst) and the subkeys k_i are independent and uniformly random, the rate of success of a 1-round differential $(\beta(0), \beta(1))$, taken over the distribution of x and k_i, is

$$P(\Delta y(1) = \beta(1) | \Delta y(0) = \beta(0)), \tag{2}$$

which by the property of Markov ciphers is equal to

$$P(\Delta y(1) = \beta(1) | \Delta y(0) = \beta(0),\ x = \gamma), \tag{3}$$

for all values of γ, if the round keys k_i are uniformly distributed. Hence the probability of a 1-round differential is independent of the distribution of x and is taken only over the distribution of k_i. Then, for independent and uniformly random subkeys k_i, the probability of an s-round characteristic is equal to the product of the probabilities of the individuals rounds (by the Chapman-Kolmogorov equation of Markov chains [22]), i.e.,

$$\prod_{i=1}^{s} P(\Delta y(i) = \beta(i) | \Delta y(i-1) = \beta(i-1)).$$

Then, the probability of the s-round differential $(\beta(0), \beta(s))$ is the sum of the probabilities of all s-round characteristics with input difference $\beta(0)$ and output difference $\beta(s)$, i.e.;

$$P(\Delta y(s) = \beta(s) | \Delta x = \beta(0))$$

$$= \sum_{\beta(1)} \sum_{\beta(2)} \cdots \sum_{\beta(s-1)} \prod_{i=1}^{s} P(\Delta y(i) = \beta(i) | \Delta y(i-1) = \beta(i-1)). \quad (4)$$

Let us denote by $p_{\max}$ the highest probability for a 1-round differential, i.e.;

$$p_{\max} = \max_{\beta} \max_{\alpha_R} P(\Delta y(1) = \beta | \Delta x = \alpha), \quad (5)$$

where $\alpha_R \neq 0$ is the right half of α (to avoid a trivial probability).

In [23], Nyberg and Knudsen proved the following theorem:

Theorem 2. *It is assumed that in a DES-like cipher with* $f : GF(2)^m \to GF(2)^n$ *the round keys are independent and uniformly random. Then the probability of an s-round differential,* $s \geq 4$, *is less than or equal to* $2p_{\max}^2$.

Nyberg proved that the bound $2p_{\max}^2$ *is also reached for less rounds differentials* $(s \geq 3)$ *when f is a permutation* [23].

In the next section special functions which minimize this bound, and hence provide high resistance against differential attacks when used as alternative S-boxes, are defined.

4. APN Functions

Let $L = \mathbb{F}_q$ be the field with q elements, $q = p^n$ for some prime number p and some positive integer n. Consider a function $f : L \to L$.

For $a, b \in L$, $a \neq 0$, let

$$N_f(a, b) = \{x \in L : f(x + a) - f(x) = b\}.$$

We say that f is nonlinear if $\Delta_f = \max\{|N_f(a, b)| : a, b \in L, a \neq 0\}$ is smaller than q. Obviously, if f is linear, $\Delta_f = q$.

For $p = 2$, necessarily $\Delta_f \geq 2$. A function f for which $\Delta_f = 2$ is called almost perfect nonlinear (APN).

Definition 2. Let $L = \mathbb{F}_q$, with $q = 2^n$ for some positive number n. A function $f : L \to L$ is said to be *almost perfect nonlinear* (APN) on L if for all $a, b \in L$, $a \neq 0$, the equation

$$f(x + a) + f(x) = b, \tag{6}$$

has at most 2 solutions.

Because L has characteristic 2, it is easy to see that if x is a solution of the equation (6), $x + a$ is also a solution, then an equivalent definition is: f is an APN function if the "derivative" set:

$$D_a(f) = \{f(x + a) + f(x) : x \in L\}, \tag{7}$$

has size at least 2^{n-1} for each $a \in L^*$. As we will see later, these functions have good resistance against differential attacks when used as S-boxes of DES-like algorithms.

Some examples of APN functions are monomials of the form $f(x) = x^{2^t+1}$ (the family of Gold functions) defined over $\mathbb{F}_{2^n}$. This class was shown to be APN for all n, provided $(n, t) = 1$, by Janwa et al. [19], Janwa and Wilson [20], as well as by Nyberg [25].

The class of monomials $f(x) = x^{4^t - 2^t + 1}$ (called the Kasami-Welch functions) defined over $\mathbb{F}_{2^n}$ are also APN when $(t, n) = 1$. These functions are maximally nonlinear (and hence APN) for odd n as proved by Kasami [21], and for even n as proved by Dobbertin [13].

It is obvious that a constant function is not APN. Neither the identity function $f(x) = x$, since for all $x \in F_{2^n}$, $f(x + a) + f(x) = a$; then the number of solutions of the Equation (6) have $|\mathbb{F}_{2^n}| = 2^n$ solutions, by selecting $b = a$. The same applies for $f(x) = x^2$. In general, a monomial function $f(x) = cx^d$, where c is a constant and d is a power of 2, is not APN.

The APN property is invariant under some transformations of functions. A function $f : L \to L$ is called *affine* if

$$f(x) = a + \sum_{i=0}^{n-1} a_i x^{2^i}, \quad a, a_i \in L.$$

Two functions f and g are Carlet, Charpin, Zinoviev equivalent (CCZ-equivalent) if the graph of f, $\{x, f(x)\}$, can be obtained from the graph of g, $\{x, g(x)\}$, by an affine permutation. Two CCZ-equivalent functions preserves the APN property. Mostly, the CCZ-equivalence is very hard to prove (for more details, see [3]).

Until 2006, the list of known inequivalent APN functions on $L = \mathbb{F}_{2^n}$ was rather short as it is shown in Table 1 [17, 21, 13, 14, 15, 25].

Table 1. Monomial APN functions

$f(x) = x^d$	Exponent d	Constraints
Gold	$2^r + 1$	$(r, n) = 1$
Kasami-Welch	$2^{2r} - 2^r + 1$	$(r, n) = 1$
Welch	$2^r + 3$	$n = 2r + 1$
Niho	$2^r + 2^{r/2} - 1$ $2^r + 2^{(3r+1)/2} - 1$	$n = 2r + 1$, r even $n = 2r + 1$, r odd
Inverse	$2^{2r} - 1$	$n = 2r + 1$
Dobbertin	$2^{4r} + 2^{3r} + 2^{2r} + 2^r - 1$	$n = 5r$

It was conjectured that these monomial functions were the only APN functions, up to equivalence. In 2006, Edel et al. [16] showed that the function

$$x^3 + ux^{36} \in \mathbb{F}_{2^{10}}[x]$$

is APN on $\mathbb{F}_{2^{10}}$ for a special seleccion of u. They also showed that this function is not CCZ-equivalent with any of the functions listed in Table 1. Since the emergence of this first example, there exist now several infinite families of non-monomial APN functions (see [18] and references therein). The Table 2 list all the known APN polynomials until now.

Table 2. Nonmonomial APN functions

$f(x)$	Constraints
$x^{2^s+1} + a^{2^t-1}x^{2^{it}+2^{rt+s}}$	$n = 3t$, $(t, 3) = (s, 3t) = 1$, $t \geq 3$ $i \equiv st \mod 3$, $r = 3 - i$, a is primitive in L
$x^{2^s+1} + a^{2^t-1}x^{2^{it}+2^{rt+s}}$	$n = 4t$, $(t, 2) = (s, 2t) = 1$, $t \geq 3$ $i \equiv st \mod 4$, $r = 4 - i$, a is primitive in L
$ax^{2^s+1} + a^{2^m}x^{2^{m+s}+2^m} + bx^{2^m+1} + \sum_{j=1}^{m-1} c_j x^{2^{m+i}+2^i}$	$n = 2m$, m odd $c_j \in \mathbb{F}_{2^m}$, $(s, m) = 1$, s odd a, b are primitive in L
$ax^{2^{n-t}+2^{t+s}} + a^{2^t}x^{2^{s+1}} + bx^{2^{n-t}+1}$	$n = 3t$, $(s, 3t) = 1$, $(3, t) = 1$, $3 \mid (t + s)$ a is primitive in L, $b \in \mathbb{F}_{2^t}$
$a^{2^t}x^{2^{n-t}+2^{t+s}} + ax^{2^{s+1}} + bx^{2^{n-t}+1}$	$n = 3t$, $(s, 3t) = 1$, $(3, t) = 1$, $3 \mid (t + s)$ a is primitive in L, $b \in \mathbb{F}_{2^t}$
$a^{2^t}x^{2^{n-t}+2^{t+s}} + ax^{2^{s+1}} + bx^{2^{n-t}+1} + ca^{2^t+1}x^{2^{t+s}+2^s}$	$n = 3t$, $(s, 3t) = 1$, $(3, t) = 1$, $3 \mid (t + s)$ a is primitive in L, $b, c \in \mathbb{F}_{2^t}$, $bc \neq 1$
$x^{2^{2k}+2^k}bx^{q+1} + cx^{q(2^{2k}+2^k)}$	$n = 2m$, m odd, c a power of $(q - s)$ but not a power of $(q-1)(2^i+1)$, $cb^q + b \neq 0$
$x^3 + \mathrm{tr}_1^n(x^9)$	
$x^{2^k+1} + \mathrm{tr}_m^n(x)^{2^k+1}$	$n = 2m = 4t$, $(n, k) = 1$

4.1. The nonlinear property

We now display the nonlinear property of an APN function f by computing the number of elements in its derivative set $D_a(f)$.

Let $\mathbb{F}_{2^n}$ be the finite field with irreducible polynomial $p(x) \in \mathbb{F}_2[x]$ of degree n, i.e., $\mathbb{F}_{2^n} = \mathbb{F}_2(\alpha)$, where α is a root of $p(x)$. $\mathbb{F}_{2^n}$ can be identified as the n-dimensional vector space $(\mathbb{F}_2)^n$ over $\mathbb{F}_2$. Then, any

univariate polynomial $f : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ can be represented as a multivariate polynomial $f : (\mathbb{F}_2)^n \to (\mathbb{F}_2)^n$, called "the algebraic normal form" of f [12],

$$f(x_1, \ldots, x_n) = \sum_{u \in F_2^n} c(u)\left(\prod_{i=1}^{n} x_i^{u_i}\right), \quad c(u) \in (\mathbb{F}_2)^n, \tag{8}$$

whose degree with respect to each variable is at most one.

The following example show the nonlinearity of the particular APN function $f(x) = x^5$ over $\mathbb{F}_8$.

We can represent $\mathbb{F}_8$ as $\mathbb{F}_2(\alpha)$:

$$\mathbb{F}_8 = \{0, 1, \alpha, \alpha^2, \alpha + 1, \alpha^2 + \alpha, \alpha^2 + \alpha + 1, \alpha^2 + 1\},$$

where α is a root of $p(x) = x^3 + x + 1$. Then we get the permutation of elements in $\mathbb{F}_8$:

x	$f(x) = x^5$
0	0
1	1
α	$\alpha^2 + \alpha + 1$
α^2	$\alpha + 1$
$\alpha + 1$	α
$\alpha^2 + \alpha$	$\alpha^2 + 1$
$\alpha^2 + \alpha + 1$	$\alpha^2 + \alpha$
$\alpha^2 + 1$	α^2

Using the identification as a 3-dimensional vector space $(\mathbb{F}_2)^3$:

$a = (a_1, a_2, a_3)$	$f(a)$
(0, 0, 0)	(0, 0, 0)
(0, 0, 1)	(0, 0, 1)
(0, 1, 0)	(1, 1, 1)
(0, 1, 1)	(0, 1, 0)
(1, 0, 0)	(0, 1, 1)
(1, 0, 1)	(1, 0, 0)
(1, 1, 0)	(1, 0, 1)
(1, 1, 1)	(1, 1, 0)

The algebraic normal form of f is:

$$f(x_1, x_2, x_3) = (x_2 + x_1x_3 + x_2x_3, x_1 + x_2 + x_1x_3, x_1 + x_2 + x_3 + x_1x_2).$$

Computing $|D_a(f)|$ for each $a \in (\mathbb{F}_2)^3$, $a \neq 0$, we get:

$a = (a_1, a_2, a_3)$	$D_a(f)$	$\lvert D_a(f)\rvert$
(0, 0, 1)	{(0, 0, 1), (0, 1, 1), (1, 0, 1), (1, 1, 1)}	4
(0, 1, 0)	{(0, 1, 0), (0, 1, 1), (1, 1, 0), (1, 1, 1)}	4
(0, 1, 1)	{(0, 0, 1), (0, 1, 0), (1, 0, 1), (1, 1, 0)}	4
(1, 0, 0)	{(0, 1, 0), (0, 1, 1), (1, 0, 0), (1, 0, 1)}	4
(1, 0, 1)	{(0, 0, 1), (0, 1, 0), (1, 0, 0), (1, 1, 1)}	4
(1, 1, 0)	{(1, 0, 0), (1, 0, 1), (1, 1, 0), (1, 1, 1)}	4
(1, 1, 1)	{(0, 0, 1), (0, 1, 1), (1, 0, 0), (1, 1, 0)}	4

As can be seen, for each $a \in (\mathbb{F}_2)^3$, $a \neq 0$, the size of the derivative $D_a(f)$ is $|D_a(f)| = 4 \geq 2^{3-1}$, implying that f is APN.

As we already discussed at the beginning of this section, the function $f(x) = x^5$ defined over $\mathbb{F}_{2^3}$ belongs to the family of Gold functions (see Table 1), which is APN because of the relatively prime condition.

4.2. Resistance against differential attacks

Now we will show in detail, as it was briefly did it in [23], that the round function of a DES-like cipher can be chosen such that $p_{\max}$ has a small value, i.e., that it is possible to choose the round function so that no single differential is useful. For simplicity let us use (α, β) for a 1-round differential. Let f be the function as defined in a DES-like cipher in Section 2.

Let us denote

$$P_f = \max_b \max_{a \neq 0} P(f(X + a) + f(X) = b). \tag{9}$$

From $\Delta y(0) = \alpha$, $\Delta y(1) = \beta$, we have:

$$y'(0) - y(0) = \alpha, \tag{10}$$

$$y'(1) - y(1) = \beta. \tag{11}$$

Then, dividing y, y' in two halves we get

$$(y'_L(0) - y_L(0),\ y'_R(0) - y_R(0)) = (\alpha_L, \alpha_R), \tag{12}$$

$$(y'_R(0) - y_R(0),\ f(E(y'_R(0)) + K) + y'_L(0) - f(E(y_R(0)) + K) - y_L(0)) = (\beta_L, \beta_R). \tag{13}$$

From (12), we get $y'_L(0) - y_L(0) = \alpha_L$, $y'_R(0) = y_R(0) + \alpha_R$. Replacing this in (13) and equating components, we get

$$y'_R(0) - y_R(0) = \beta_L, \tag{14}$$

$$\alpha_L + f(E(y_R(0) + \alpha_R) + K) + f(E(y_R(0)) + K) = \beta_R. \tag{15}$$

Thus, from (5),

$$\begin{aligned} p_{\max} &= \max_\beta \max_{\alpha_R} P(\Delta y(1) = \beta | \Delta y(0) = \alpha) \\ &= \max_\beta \max_{\alpha_R} P(\alpha_L + f(E(y_R(0) + \alpha_R) + K) \\ &\quad + f(E(y_R(0)) + K) = \beta_R) \\ &= \max_\beta \max_{\alpha_R} P(\alpha_L + f(E(y'_R(0)) + K) + f(E(y_R(0)) + K) = \beta_R). \end{aligned}$$

Let $\widetilde{x} = y_R(0)$

$$= \max_{\beta} \max_{\alpha_R} P(f(E(\widetilde{x}) + K) + f(E(\widetilde{x} + \alpha_R) + K) = \alpha_L + \beta_R)$$

$$= \max_{\beta} \max_{\alpha_R} P(f(X) + f(X + a) = b),$$

where we have assumed that E is affine, and denoted $E(\widetilde{x}) + K$ by X, $E(\alpha_R)$ by a and $\alpha_L + \beta_R$ by b.

This means that, if we choose a function f in (1) such that for any nonzero difference $X + X' = a$, the occurrence of the difference $f(X) + f(X') = b$ has very low probability, i.e., an APN function, then the success of any 1-round differential (and as a consequence, any s-round differential) is hardly probable. Therefore, APN functions provides good resistance against differential attacks when used as the round function of DES-like ciphers.

4.3. The major APN open problem

As showed in Section 2, the design of a cryptosystem consist of algorithms of encryption and decryption. Then, an important condition for the round function is to be a bijective function (a permutation). Also, as showed in the same section, the process of encryption of DES-like ciphers start by dividing the plaintexts into two halves, i.e., many ciphers applies as functions of even number of variables, it means, functions defined over $\mathbb{F}_{2^n}$ for an even number n.

It is well known that for $n = 2, 4$ does not exist APN permutations over $\mathbb{F}_{2^n}$. For $n = 6$, there exist (up to now) only one APN permutation (The Dillon function [8]). For $n > 6$ is unknown the existence or non existence of APN permutations. Then, the following is considered a major open question about APN functions.

Does there exist APN permutations over $\mathbb{F}_{2^n}$ for n an even number greater than 6?

References

[1] E. Biham and A. Shamir, Differential Cryptanalysis of DES-like Cryptosystems, Lect. Notes Comp. Sci. Springer, 537 (1990), 2-21.

[2] E. Biham and A. Shamir, Differential cryptanalysis of the full 16-round DES, Differential cryptanalysis of the Data Encryption Standard, pp. 79-88. Springer, New York.

[3] C. Carlet, P. Charpin and V. Zinoviev, Codes, bent functions and permutations suitable for DES-like cryptosystems, Designs, Codes and Cryptography 15(2) (1998), 125-156.

DOI: https://doi.org/10.1023/A:1008344232130

[4] D. Coppersmith, The Data Encryption Standard and its strength against attacks, IBM Journal of Research and Development 38(3) (1994), 243-250.

DOI: https://doi.org/10.1147/rd.383.0243

[5] C. Blondeau and K. Nyberg, Perfect nonlinear functions and cryptography, Finite Fields and Their Applications 32 (2015), 120-147.

DOI: https://doi.org/10.1016/j.ffa.2014.10.007

[6] C. Bracken, E. Byrne, N. Markin and G. McGuire, New families of quadratic almost perfect nonlinear trinomials and multinomials, Finite Fields and Their Applications 14(3) (2008) 703-714.

DOI: https://doi.org/10.1016/j.ffa.2007.11.002

[7] M. Brinkman and G. Leander, On the classification of APN functions up to dimension five, International Workshop on Coding and Cryptography (WCC), Versailles, France, 2007.

[8] K. A. Browning, J. F. Dillon, M. T. McQuistan and A. J. Wolfe, An APN permutation in dimension six, Finite Fields: Theory and Applications 518 (2010), 33-42.

[9] L. Budaghyan, C. Carlet, P. Felke and G. Leander, An infinite class of quadratic APN functions which are not equivalent to power mappings, Proceedings of ISIT 2006, Seattle, USA, 2006.

[10] L. Budaghyan, C. Carlet and G. Leander, Another class of quadratic APN binomials over $\mathbb{F}_{2^n}$: the case n divisible by 4, Proceedings of the Workshop on Coding and Cryptography (WCC07) (2007), 49-58.

[11] L. Budaghyan, C. Carlet and G. Leander, Constructing new APN functions from known ones, Finite Fields and Their Applications 15(2) (2009), 150-159.

DOI: https://doi.org/10.1016/j.ffa.2008.10.001

[12] C. Carlet, Vectorial Boolean functions for cryptography, Boolean Models and Methods in Mathematics, Computer Science and Engineering 134 (2010), 398-469.

[13] H. Dobbertin, Almost perfect nonlinear power functions on GF(2^n): The Welch case, IEEE Transact. Inform. Th. 45(4) (1999), 1271-1275.

DOI: https://doi.org/10.1109/18.761283

[14] H. Dobbertin, Almost perfect nonlinear power functions on GF(2^n): The Niho case, Information and Computation 151(1-2) (1999), 57-72.

DOI: https://doi.org/10.1006/inco.1998.2764

[15] H. Dobbertin, Almost perfect nonlinear power functions on GF(2^n): A new case for n divisible by 5, In: D. Jungnickel and H. Niederreiter, Editors; Proceedings of the Conference on Finite Fields and Applications, Augsburg 1999, Springer-Verlag, Berlin (2001), 113-121.

[16] Y. Edel, G. Kyureghyan and A. Pott, A new APN function which is not equivalent to a power mapping, IEEE Trans. Inform. Th. 52(2) (2006), 744-747.

DOI: https://doi.org/10.1109/TIT.2005.862128

[17] R. Gold, Maximal recursive sequences with 3-valued recursive cross-correlation functions, IEEE Trans. Inform. Th. 14(1) (1968), 154-156.

DOI: https://doi.org/10.1109/TIT.1968.1054106

[18] F. Gologlu, APN Trinomials and Hexanomials, arXiv: 1411.2981v1, (2014).

[19] H. Janwa, G. McGuire and R. M. Wilson, Double-error-correcting cyclic codes and absolutely irreducible polynomials over GF(2), Journal of Algebra 178(2) (1995), 665-676.

DOI: https://doi.org/10.1006/jabr.1995.1372

[20] H. Janwa and M. Wilson, Hyperplane sections of Fermat varieties in $\mathbb{P}^3$ in chapter 2 and some applications to cyclic codes, Applied Algebra, Algebraic Algorithms and Error-Correcting Codes, Proceedings AAECC-10 (G. Cohen, T. Mora and O. Moreno Editors), Lecture Notes in Computer Science, Springer-Verlag, New York/Berlin 673 (1993), 180-194.

[21] T. Kasami, The weight enumerators for several classes of subcodes of the second order binary Reed-Muller codes, Information and Control 18(4) (1971), 369-394.

DOI: https://doi.org/10.1016/S0019-9958(71)90473-6

[22] X. Lai, J. L. Massey and S. Murphy, Markov ciphers and differential cryptanalysis, Advances in Cryptology-Eurocrypt'91, Lecture Notes in Computer Science, Springer-Verlag 547 (1991), 17-38.

[23] K. Nyberg and L. R. Knudsen, Provable Security Against Differential Cryptanalysis, Proceedings of the 12th Annual International Cryptology Conference on Advances in Cryptology, Springer-Verlag (1992), 566-574.

[24] K. Nyberg and L. R. Knudsen, Provable security against a differential attack, Journal of Cryptology 8(1) (1995), 27-37.

DOI: https://doi.org/10.1007/BF00204800

[25] K. Nyberg, Differentially uniform mappings for cryptography, Advances in Cryptology-Eurocrypt'93, Springer-Verlag (1994), 55-64.

[26] F. Rodier, Bornes sur le degre des polynomes presque parfaitement non-lineaires, Contemporary Math., AMS, Providence (RI) USA 487 (2009), 169-181.

[27] W. Stallings, Cryptography and Network Security Principles and Practice, Fifth Edition, New York, Pearson, 2010.

■