

A Complete Formulation of Generalized Affine Equivalence

Marco Macchetti, Mario Caironi, Luca Breveglieri, and Alessandra Cherubini

Politecnico di Milano, Milan, Italy

{macchett, caironi, brevegli}@elet.polimi.it, aleche@mate.polimi.it

Abstract. In this paper we present an extension of the generalized linear equivalence relation, proposed in [7]. This mathematical tool can be helpful for the classification of non-linear functions $f : F_p^m \rightarrow F_p^n$ based on their cryptographic properties. It thus can have relevance in the design criteria for substitution boxes (S-boxes), the latter being commonly used to achieve non-linearity in most symmetric key algorithms. First, we introduce a simple but effective representation of the cryptographic properties of S-box functions when the characteristic of the underlying finite field is odd; following this line, we adapt the linear cryptanalysis technique, providing a generalization of Matsui's lemma. This is done in order to complete the proof of Theorem 2 in [7], also by considering the broader class of generalized affine transformations. We believe that the present work can be a step towards the extension of known cryptanalytic techniques and concepts to finite fields with odd characteristic.

Keywords: Boolean functions, generalized linear equivalence, linear cryptanalysis, S-boxes.

1 Introduction

Symmetric key cryptographic algorithms play a crucial role in today's secure communication protocols and secure storage applications, due to their high efficiency and key-agility. The class of block ciphers has recently known a flourishing of proposals, also due to the Advanced Encryption Standard establishment process.

Block ciphers are usually characterized by an iterative nature: a constant set of transformations, called *round* or more generically *step*, is applied several times to the plaintext block in order to obtain the corresponding ciphertext. The round transformation must necessarily possess several properties in order to enforce the robustness of the whole algorithm, and to maximize efficiency: it must be key-dependent, it should be highly non-linear and it should guarantee a high level of diffusion of information.

These constraints must be satisfied regardless of the block cipher structural scheme, e.g. they are valid for Feistel networks [1], Lai-Massey [13] and Substitution-permutation networks [8]. A common and well-studied method to implement the non-linear step is to use bricklayer functions composed by S-boxes. These are usually defined over the binary domain, i.e. $S : F_2^m \rightarrow F_2^n$ and they are chosen such that their cryptographic characteristics are optimal.

Several works have been focused on the characterization of the non-linear properties of S-boxes, some examples being [3],[16],[17],[18], and on the possibility of partitioning them into equivalence classes [14],[11],[12]; recent papers propose efficient algorithms that can be used to decide if two S-boxes [6] or two Boolean functions [10] are linearly equivalent. This research activity is motivated by the relevant link between the properties of S-boxes and the security and efficiency of block ciphers.

Two attacks that are particularly relevant for block ciphers are linear cryptanalysis [15],[4] and differential cryptanalysis [5], thus most of the efforts have been directed to the problem of finding S-boxes with optimal differential and linear characteristics.

In [7], Breveglieri, Cherubini and Macchetti proposed an extension of the criterion of functional linear equivalence called *generalized linear equivalence*; it has been shown that generalized equivalence classes result from merging of classical equivalence classes and that linear and differential characteristics are indeed invariant under this broader class of transformation.

The aim of this paper is double fold. First, we elaborate more on the theoretical basis of the original formulation of generalized equivalence; in fact, although it has been proved that the linear characteristics of S-boxes are invariant in the context of these transformations, the proof, as it is, is rigorously valid only for fields of even characteristic. No formalization of the linear cryptanalysis technique over fields with odd characteristic can be found in the scientific literature. This is provided in this paper, along with the corresponding generalization of Matsui's lemma. The original proof of invariance is then completely and coherently derived.

A second contribution is the introduction of generalized *affine* transformations; these are indeed the natural extension of classical affine transformations and complete the results of [7]. The proof of invariance for S-box cryptographic robustness is thus obtained within the largest possible context.

This paper is organized as follows: in Section 2 we define and analyze the linear characteristics of S-boxes defined over finite fields of odd characteristic. In Section 3 we extend the linear cryptanalysis technique, providing a generalization of Matsui's lemma. In Section 4 we give a complete proof for the invariance of cryptographic robustness of S-boxes, also considering generalized affine transformations. Section 5 concludes the paper.

2 Linear Biases over F_p^m

Broadly speaking, the goal of symmetric-key cryptanalysis is to distinguish a block cipher from a set of random permutations, and to get information on key material faster than it could be done via a trivial brute-force attack.

In the case of differential cryptanalysis¹, the suggested distinguisher is the probability of finding certain differences in the ciphertext blocks, given certain

¹ We assume here that the reader has some basic familiarity with the ideas beyond differential and linear cryptanalysis.

differences in the corresponding plaintexts. If this probability deviates significantly from what would be expected from a random permutation, the attack is successful and information about the key can be found.

Differential characteristics for N rounds of a block cipher can be constructed starting from differential characteristics of the single S-boxes; these in fact are usually the only non-linear component of a symmetric key algorithm. Conventional algorithms are defined over the finite field F_2 for evident reasons of efficiency. However, it is easy to extend the differential cryptanalysis technique to finite fields of odd characteristic; this extension is of theoretical interest as it may be used to test the cryptographic robustness of basic arithmetic operations, such as multiplication, inversion, and power functions defined over a field F_{p^m} . Recent work has been done by Dobbertin [9] regarding the problem of finding power monomials in such fields with optimal differential characteristics.

For a given S-box function $f : F_p^m \rightarrow F_p^n$ with p prime and $m, n > 1$ the Difference Distribution Table (DDT) is built by computing the number $\delta_f(a, b)$ of solutions x of the equation

$$f(x \oplus a) \ominus f(x) = b \quad a \in F_p^m, b \in F_p^n \quad (1)$$

where \oplus and \ominus respectively indicate sum over F_{p^m} , the finite field associated with the vector space F_p^m , and difference over F_p^n . The lower the value of the maximum entry in the table, $\Delta_f = \max_{a \neq 0, b}(\delta_f(a, b))$, the more robust function f is versus differential cryptanalysis, since the differential characteristics $\delta_f(a, b)$ of S-boxes located in different rounds of the cipher are, roughly speaking, connected to form a multi-round characteristic. The amount of plaintext-ciphertext block pairs needed to highlight a hypothetical differential bias is inversely proportional to its magnitude.

The linear cryptanalysis technique is built upon a very similar concept, namely that of linear distinguisher. The objective of linear cryptanalysis is to build linear (over F_2) equations involving plaintext, ciphertext and key bits that hold with a probability significantly different from 50%. If this is possible for a high number of rounds, then the attack may reveal information about the key bits faster than brute-force attacks (this indeed is the case for the DES cipher). Again, linear characteristics for a full cipher are built starting from those of the single S-boxes.

More formally, the Linear Approximation Table (LAT) of an S-box function $f : F_2^m \rightarrow F_2^n$ is built by counting the number $\lambda_f(a, b)$ of solutions x of the equation

$$a \bullet x = b \bullet f(x) \quad a \in F_2^m, b \in F_2^n \quad (2)$$

where the inner product over F_2^m and F_2^n is indicated with \bullet and gives a value in F_2 . The robustness to linear cryptanalysis is measured with the maximum value $\Lambda_f = \max_{a, b \neq 0}(|\lambda_f(a, b) - 2^{m-1}|)$. In fact, the event when the number of solutions of (2) is always very near to 2^{m-1} is the best case for the designer and the worst case for the attacker; this is because a random Boolean function is expected to be equal to any linear Boolean function for roughly half of the points

of its domain, i.e. 2^{m-1} times in the specific context². The attacker can then infer nothing about the function, apart from the fact that it behaves randomly, a thing which does not help in distinguishing the block cipher from a random permutation.

Several extensions of linear cryptanalysis are known, see for instance [2] which contains also a very good survey. An extension of linear cryptanalysis to finite fields of odd characteristic may be beneficial for the same motivations outlined above for the differential case.

We start by defining the affine biases for an S-box over F_p . Let $f : F_p^m \rightarrow F_p^n$ be an S-box function, then we introduce the Affine Approximation Table (AAT) of f , which is built by counting the number $\lambda_f(a, b, c)$ of solutions x of the equation

$$a \bullet x \oplus b \bullet f(x) = c \quad a \in F_p^m, b \in F_p^n, c \in F_p \tag{3}$$

The constant c in (3) is introduced to take into account the fact that the inner product now gives a value in F_p , and for this reason it is not sufficient to compare f to all the linear functions, and in fact all affine functions must be taken into consideration. Another way of looking at (3) is to say that we count the number of times that function $b \bullet f(x)$ is equal to the affine function $(-a) \bullet x \oplus c$, hence the name AAT.

Each cell of the AAT of f is indexed by the triplet $\{a, b, c\}$ and the AAT is indeed a three-dimensional array rather than a table, still we keep the old terminology for clearness. The value of interest to the cryptanalyst becomes in this case $\lambda_f(a, b, c) - p^{m-1}$; the reason is that a random function over the range F_p is expected to be equal to any affine function in roughly one case out of p , and since the cardinality of the domain of f is p^m a simple division gives the expected result p^{m-1} . The overall robustness of the function can then be characterized with the parameter $\Lambda_f = \max_{a,b \neq 0,c} (|\lambda_f(a, b, c) - p^{m-1}|)$.

In the binary case there is no need to consider affine functions, because the number of solutions of equation

$$a \bullet x \oplus b \bullet f(x) = c \quad a \in F_2^m, b \in F_2^n, c \in F_2 \tag{4}$$

is equal to the number of solutions for the pair $\{a, -b\}$ in (2) when $c = 0$ and is equal to the difference between 2^m and the previous number when $c = 1$. In a sense, in fields of even characteristic, affine distinguishers are totally redundant and give no advantage to the attacker in addition to linear distinguishers.

Even in the odd characteristic case a partial redundancy is present, because it must hold that

$$\sum_{c=0}^{p-1} \lambda_f(a, b, c) = p^m \tag{5}$$

and this implies that one value out of p in the AAT is redundant. At this point it is useful to give a visual representation of the AAT, since this will also be

² The functions for which $\Lambda_f = 2^{\frac{m}{2}-1}$ are called Bent, and exist only under additional hypotheses on the number of input/output variables.

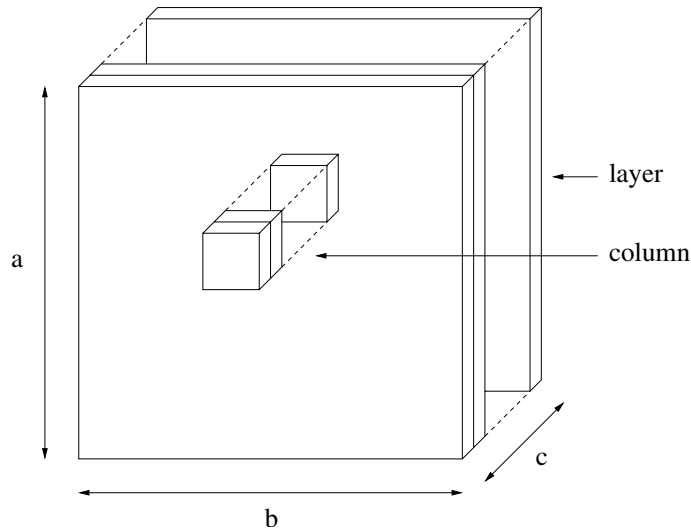


Fig. 1. A graphical representation of the Affine Approximation Table

useful in the following Sections to understand what is the concrete effect of generalized linear and affine transformations. Figure 1 depicts the AAT of a generic S-box function f ; the values of the three parameters a, b, c vary along the three different axes of the table. We call *layers* the set of cells with constant index c , and *columns* the set of cells with constant a, b indexes; these structures are highlighted in the Figure.

Elaborating on these definitions, we can say that one out of the p layers is then redundant, because essentially all the information is already contained in the remaining $p - 1$ ones. The choice of the redundant layer is arbitrary, thus in the following discussion we will assume that all the layers are maintained. It is clear that the number of possible biases for S-boxes defined over F_p is higher than that of S-boxes defined over the field F_2 ; the latter only have two layers in the AAT. This means that more computational effort is in general required to calculate the AAT versus the LAT, but also that the cryptanalyst may have more freedom in the choice of the biases to be used in an *affine* attack.

In the next Section we will see how the AATs can be used in such a generalization of the linear cryptanalysis attack.

3 Extending Linear Cryptanalysis

The next step towards a complete formulation of linear cryptanalysis on fields with odd characteristic is the extension of Matsui's Piling-Up Lemma [15]; this is classically used to obtain the bias of a sum of linearly biased variables, even if it is rigorously valid only if the variables are strictly uncorrelated.

An extension of Matsui's Lemma has been proposed in [2], in the context of a specific variant of linear cryptanalysis: the input/output Boolean sums are

substituted with linear (and non-linear) projections over F_2^l . The formulation is indeed quite complex; our goal here is rather to obtain a simple formula, involving only the affine biases, which is the direct extension of that of Matsui.

Let X_1 be a variable with values over F_p ; the affine bias ϵ_1^i is defined via the following equation:

$$\Pr(X_1 = i) = \frac{1}{p} + \epsilon_1^i \quad i \in F_p \tag{6}$$

Thus a vector of affine biases $\Xi_1 = \langle \epsilon_1^0, \dots, \epsilon_1^{p-1} \rangle$ is associated with X_1 . Now, we want to be able to compute the affine bias vector of a sum of two such variables starting from the two single bias vectors; this is done in the following calculations, where all the sums performed over F_p are indicated with \oplus .

$$\begin{aligned} \Pr(X_1 \oplus X_2 = k) &= \sum_{i \oplus j = k} \left(\frac{1}{p} + \epsilon_1^i\right) \left(\frac{1}{p} + \epsilon_2^j\right) = \\ &= \sum_{i \oplus j = k} \frac{1}{p^2} + \frac{1}{p} \sum_{i \oplus j = k} \epsilon_1^i + \frac{1}{p} \sum_{i \oplus j = k} \epsilon_2^j + \sum_{i \oplus j = k} \epsilon_1^i \epsilon_2^j = \\ &= \frac{1}{p} + \sum_{i \oplus j = k} \epsilon_1^i \epsilon_2^j \end{aligned} \tag{7}$$

By re-writing the last passage using only the affine biases we obtain:

$$\epsilon_{1,2}^k = \sum_{i \oplus j = k} \epsilon_1^i \epsilon_2^j \tag{8}$$

which is a direct generalization of Matsui’s formula for the sum of two variables; moreover, the link between the affine bias vectors is given by:

$$\Xi_{1,2} = \Xi_1 \star \Xi_2 \tag{9}$$

where \star stands for a variant of the discrete convolution operation where the sum and differences of vector indexes are computed over F_p . We note that both (8) and (9) could be easily extended to a number $n > 2$ of variables, and that they reduce to the well-known formula for linear cryptanalysis if $p = 2$.

In the proposed extension of linear cryptanalysis the variables X_i are indeed approximations of the non-linear components of the algorithm under consideration, typically the active S-boxes, i.e. they will have the form of (3). Thus, the affine bias vector Ξ_i is indeed nothing but a column of the AAT of f , indexed by the specific values of a, b . We think that the name *affine cryptanalysis* could be effectively used to identify the extension of linear cryptanalysis to fields of odd characteristic.

An outline of Matsui’s algorithm 2 targeting a cryptographic algorithm which operates on the base field F_p is roughly as follows:

1. The attacker chooses an affine characteristic over $N - 1$ rounds of the cipher; the affine distinguishers of the active S-boxes are calculated and stored in the AATs.

2. The attacker chooses the affine approximations for all active S-boxes, i.e. the values of the parameters a, b are selected for each active S-box.
3. Equation (9) can then be used to calculate the affine bias vector of the global characteristic, Ξ_T .
4. Last-round decryptions of the ciphertexts will eventually reveal the bias of the affine approximation under the correct key hypothesis.

We underline a difference with regard to the case of even characteristic. The maximum affine bias inside Ξ_T must always be searched for and identified; this happens because the position of the ϵ_T^i with maximum (minimum) value depends on the key material which is added along the affine trail. This has the effect of changing the i in a key-dependent way, and roughly increases the complexity of the attack by a factor of p compared to standard linear cryptanalysis.

An anonymous referee has pointed out that in the case of composite fields $F_{p^{mn}}$, the field F_{p^m} can be taken as a base field in place of F_p and all equations could be re-written properly to obtain affine approximations at a higher level. The constant c would in this case belong to F_{p^m} , and the inner product in (3) would be modified accordingly. This leads to an interesting formulation of the affine biases that may have practical applications in the case of composite fields with even characteristic.

4 A Complete Formulation of Generalized Equivalence

4.1 An Extended Proof

Given the previous background, it is now possible to give a coherent proof of Theorem 2 in [7]. The need for a formal extension basically derives from the fact the original formulation of generalized linear equivalence does not take into account the differences between linear cryptanalysis and affine cryptanalysis. The part about differential characteristics remains unchanged and will not be repeated here (it will be expanded when generalized affine transformation are considered).

We summarize here the basics of the approach outlined in [7]. It is possible to associate a particular geometric representation to any completely specified function $f : F_p^m \rightarrow F_p^n$. Let S be a linear space of dimension $k = m + n$, where the vector components are defined over F_p ; consider the set \mathcal{F} of p^m vectors, belonging to S , formed by the rows of the truth-table of f (each vector belonging to \mathcal{F} is the concatenation of an input vector of f and its corresponding output vector). We refer to \mathcal{F} as the implicit embedding of f in the linear space S .

If an invertible linear transformation of coordinates is applied to S , the essential information contained in \mathcal{F} is not changed. Every such invertible linear transformation is governed by a non-singular $(m + n) \times (m + n)$ matrix over F_p . The non-singularity of this matrix, while providing the possibility to invert the transformation, also assures that we do not lose information while transforming the coordinates. The extended Theorem follows.

Theorem 1. *Given two functions $f, g : F_p^m \rightarrow F_p^n$ and a non-singular $(m + n) \times (m + n)$ matrix T over F_p , if $g = T(f)$ then the distributions of values in the AATs of f and g are equal.*

Proof. A cell of the AAT table of f indexed by $\{a, b, c\}$ contains the number of input vectors x such that $a^T \bullet x \oplus b^T \bullet f(x) = c$, where, for sake of clearness, the transposed of vector v is indicated with v^T .

Thus, if we consider the geometric representation for function f we have that the cell contains the number of vectors w belonging to the implicit embedding of f such that $k^T \bullet w = c$ where $k = (a)|(b)$ (the concatenation of vectors a and b); note that $a \in F_p^m$, $b \in F_p^n$ and $k \in F_p^{m+n}$. The merged index k is unique for every column in the AAT of the two functions.

These vectors will be transformed by the change of basis into other vectors w' belonging to the implicit embedding of function g such that $w' = Tw$. We can rewrite the equation as:

$$k^T \bullet Tw = c \iff (T^T k)^T \bullet w = c \iff (k')^T \bullet w = c$$

Since matrix T is non-singular, there is a bijection between the values of k and those of $k' = T^T k$, i.e. the cells of the AAT of g are just a (linear) rearrangement of the cells of the AAT of f . □

Note that the cells belonging to a given layer cannot be shifted to different layers; indeed the cells of all the layers are reordered in a uniform way, given only by matrix T . The question if there are even more general transformations can thus arise, and we positively answer in the following Section.

4.2 Generalized Affine Transformations

We introduce the following functional equivalence relation.

Definition 1. *Two functions $f, g : F_p^m \rightarrow F_p^n$ are called generally affine equivalent if and only if the implicit embedding of g can be obtained from the implicit embedding of f as*

$$\mathcal{G} = T(\mathcal{F}) \oplus e \tag{10}$$

where T is a non-singular $(m + n) \times (m + n)$ matrix over F_p and e is a vector belonging to F_p^{m+n} .

This is the natural and most elegant definition of generalized equivalence. A proof of invariance for the cryptographic characteristics of functions f, g is given in the following theorem.

Theorem 2. *Given two functions $f, g : F_p^m \rightarrow F_p^n$, a non-singular $(m + n) \times (m + n)$ matrix T and a constant vector $e \in F_p^{m+n}$, if $g = T(f) \oplus e$ then the distributions of values in the AATs and DDTs of f and g are equal.*

Proof. We first prove the relation regarding the DDTs of f and g .

A cell of the DDT of f located in the i -th row and in the j -th column contains the number of the input vector pairs (x, y) such that $y = x \oplus i$ and $f(y) = f(x) \oplus j$.

Thus, if we consider the geometric representation for function f we have that the cell contains the number of vector pairs (w, z) belonging to the implicit embedding of f such that $w = z \oplus k$ where $k = (i)|(j)$; note that $i \in F_p^m$, $j \in F_p^n$ and $k \in F_p^{m+n}$. These pairs will be transformed by the change of basis into other pairs (w', z') belonging to the implicit embedding of function g such that $w' = Tw \oplus e$, $z' = Tz \oplus e$. If we define k' equal to $w' - z'$, we obtain that $k' = Tk$ and the relation $w' = z' \oplus k'$ holds. Since matrix T is non-singular, there is a bijection between the values of k and those of k' : this means that the cells of the DDT of g are just a rearrangement of the cells of the DDT of f .

Now we prove the relation between the AATs.

A cell of the AAT of f located in the column indexed by a, b and in the c -th layer contains the number of input vectors x such that $a^T \bullet x \oplus b^T \bullet f(x) = c$. Thus, if we consider the geometric representation of function f we have that the cell contains the number of vectors w belonging to the implicit embedding of f such that $k^T \bullet w = c$ where $k = (i)|(j)$; note that $a \in F_p^m$, $b \in F_p^n$, $k \in F_p^{m+n}$ and $c \in F_p$. These vectors will be transformed by the change of basis into other vectors w' belonging to the implicit embedding of function g such that $w' = Tw \oplus e$. We can rewrite the equation as:

$$k^T \bullet (Tw \oplus e) = c \iff (T^T k)^T \bullet w = c \ominus k^T \bullet e \iff (k')^T \bullet w = c' \quad (11)$$

Given the non-singularity of matrix T , we have a bijection between the values of k and k' ; moreover, for fixed k the values of c and those of c' are bound by a permutation. Thus (11) states that the cells of the AAT of g are an (affine) rearrangement of the cells of the AAT of f . \square

The reordering of the cells in the AAT is now more complex than in the preceding case, because here the cells can also migrate among the different layers due to the presence of e ; actually the columns of the AAT of f are permuted, and the cells inside each column are also re-ordered in a column-specific way.

The nature of the transformation defined in (10) is quite general; an open question is if this is indeed the most general instance of affine functional equivalence relation.

5 Conclusions

In this paper we have given an extension of the generalized equivalence relation between functions defined over finite fields; the generalized affine transformations are the most general instance of equivalence relations proposed so far in the scientific literature.

As a side result, we have derived an extension of the linear cryptanalysis technique that is applicable to finite fields of odd characteristic; this may be practically useful to test the cryptographic robustness of arithmetic operations (and cryptographic algorithms) defined over such fields. The extension has been named affine cryptanalysis.

References

1. Announcing the Standard for DATA ENCRYPTION STANDARD (DES). FIPS Publication 46-2, NIST, 1993.
2. Baignères, T., Junod, P., Vaudenay, S.: How Far Can We Go Beyond Linear Cryptanalysis? Proceedings of ASIACRYPT 2004, 432-450, 2004.
3. Beth, T., Ding, C.: On Almost Perfect Nonlinear Permutations. Proceedings of EUROCRYPT '93, 65-76, 1994.
4. Biham, E.: On Matsui's Linear Cryptanalysis. Proceedings of EUROCRYPT '94, 341-355, 1994.
5. Biham, E., Shamir, A.: Differential Cryptanalysis of DES-like Cryptosystems. Journal of Cryptology, 4(1):3-72, 1991.
6. Biryukov, A., De Canniere, C., Braeken, A., Preneel, B.: A Toolbox for Cryptanalysis: Linear and Affine Equivalence Algorithms. Proceedings of EUROCRYPT 2003, 33-50, 2003.
7. Breveglieri, L., Cherubini, A., Macchetti, M.: On the Generalized Linear Equivalence of Functions over Finite Fields. Proceedings of ASIACRYPT 2004, 79-91, 2004.
8. Daemen, J., Rijmen, V.: The Design of Rijndael: AES-The Advanced Encryption Standard. Springer-Verlag, 2002.
9. Dobbertin, H., Mills, D., Muller, E.N., Pott, A., Willems, W.: APN functions in odd characteristic. Discrete Mathematics, 267(1-3):95-112, 2003.
10. Fuller, J., Millan, W.: Linear Redundancy in S-Boxes. Proceedings of FSE 2003, 74-86, 2003.
11. Harrison, M.A.: The Number of Classes of Invertible Boolean Functions. Journal of ACM, 10:25-28, 1963.
12. Harrison, M.A.: On Asymptotic Estimates in Switching and Automata Theory. Journal of ACM, 13(1):151-157, 1966.
13. Junod, P., Vaudenay, S.: FOX : A New Family of Block Ciphers. Proceedings of SAC 2004, 114-129, 2004.
14. Lorens, C.S.: Invertible Boolean Functions. IEEE Transactions on Electronic Computers, EC-13:529-541, 1964.
15. Matsui, M.: Linear Cryptanalysis method for DES cipher. Proceedings of EUROCRYPT '93, 386-397, 1994.
16. Nyberg, K.: Differentially Uniform Mappings for Cryptography. Proceedings of EUROCRYPT '93, 55-64, 1994.
17. Nyberg, K.: Perfect Nonlinear S-Boxes. Proceedings of EUROCRYPT '91, 378-386, 1991.
18. Nyberg, K., Knudsen, L. R.: Provable security against differential cryptanalysis. Proceedings of CRYPTO '92, 566-574, 1992.