Characterizations of the Degraded Boolean Function and Cryptanalysis of the SAFER Family

Wentan Yi and Shaozhen Chen

***Abstract*—This paper investigates the degradation properties of Boolean functions from the aspects of the distributions of differences and linear masks, and shows two characterizations of the degraded Boolean function. One is that there exists a linear space of the input differences, where the differentials with the zero output difference have probability 1; Another one is that the input linear masks of the nonzero-correlation linear approximations are included in a linear space. Those two linear spaces are orthogonal spaces. Moreover, the degradation properties are showed about the exponentiation type S-box of the SAFER block ciphers, which are applied to reduce the compute complexity in the zero-correlation linear attacks on 5-round SAFER SK/128, 4(5)-round SAFER+/128(256) and 5(6)-round SAFER++/128(256). In the attacks, some of the linear properties of PHT employed as the linear layer by the SAFER block ciphers are investigated and some zero-correlation approximations for SAFER SK, SAFER+, and SAFER++ are identified, when only the least one or two significant bits are considered. The results show that more rounds of some of the SAFER block ciphers can be attacked, by considering the degradation properties and the zero-correlation linear relations.**

***Index Terms*—Cryptography, Block cipher, Degradation property, Zero-correlation linear cryptanalysis, SAFER**

I. Introduction

S-box (Substitution box) is a basic component of symmetric cryptography algorithms and hash functions. It provides "confusion" and in most cases is the only nonlinear part of the algorithms. S-boxes have been studied widely and some criteria including differential uniformity, non-linearity, algebraic degree, correlation immunity, and the strict avalanche criterion, etc, have been considered, and they are classified according to many criterions such as affine equivalence. The strict avalanche criterion was introduced by Webster and Tavares [1], which is an important properties for an S-box to resist differential cryptanalysis[2] and also the important concepts in designing cryptographic functions [3] and hashing functions[4]. An S-box is said to satisfy the strict avalanche criterion if and only if each of its output bits changes with a probability of one half whenever a single input bit is complemented. On the contrary, an S-box is said to have the degradation properties, if and only if one or some of its output bits never change whenever one or some input bits are complemented. It is obvious that if an S-box has the degradation properties, it must not satisfy the strict avalanche criterion. Form the aspect of cryptanalysis, the degradation properties can help us to find some non-random phenomenons, and then obtain some collision massage or some information of the key in the analysis of hash functions and symmetric cryptography algorithms.

SAFER (Secure And Fast Encryption Routine) block cipher family consists of SAFER K, SAFER SK, SAFER+ and SAFER++. Among them, SAFER K [5] is the first member of the family, SAFER SK [6] is a modified version of SAFER K, SAFER+ [7] is an AES candidate and SAFER++ [8] is used in the custom algorithm of Bluetooth for key derivation and authentication. The S-boxes derived from exponentiation and discrete logarithm functions, the linear layer employing PHT (Pseudo-Hadamard-like mixing transforms) and the methods to perform key-mixing with two-commutative operations are three key features shared by the encryption algorithms of the SAFER block cipher family. Up to now, the security of the SAFER block ciphers has attracted lots of attentions, and there have been several cryptanalytic results for SAFER SK, SAFER+ and SAFER++ by using different approaches, such as non-homomorphic linear cryptanalysis[9][10], integral attack[11][12], boomerang attack[13], impossible differential cryptanalysis[14][15][16][17]. Main cryptanalytic results of the SAFER block ciphers are listed in Table 1.

In this paper, we mainly investigate the degradation properties of Boolean function from the aspects of the distributions of the differences and linear masks. Two characterizations are showed, which can be used to distinguish whether a Boolean function has degradation properties. Moreover, some degradation properties of the exponentiation type S-box of the SAFER block ciphers are discovered. In addition, the degradation properties of the S-boxes and some linear properties of PHT employed as the linear layer are applied to the zero-correlation linear attacks on the SAFER block cipher family. The contributions are two-fold in detail.

1. We show two characterizations of the degraded Boolean function from the aspect of differential and linear. If f is a Boolean function with the degradation properties, then there exists linear spaces V, W, for any difference $\alpha \in V$, $\Pr_x(f(x) \oplus f(x \oplus \alpha) = 0) = 1$, and if $\mathrm{Cor}_x(\beta \cdot x \oplus f(x)) \neq 0$, then $\beta \in W$. For a simple example, for a Boolean function with n-bit variable, if there exist a n-bit value α with weight 1, the outputs of x and $x \oplus \alpha$ are the same, or under the nonzero-correlation condition, there exists a bit-position of the input masks always keeping 0, then the Boolean function has the degradation properties. We study the distributions of the input differences and the input linear masks and show the exponentiation type S-box has some degradation properties, that is, the 7-th bit of the outputs can be obtained by the 2-th to 8-th bits of the inputs, and the XORed values of the 8-th bit of the outputs and the first bit of inputs can be obtained by the 2-th to 8-th bits of the inputs.

W. Yi and S. Chen were with the Department of State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou 450001, China. E-mail: (nlwt8988@gmal.com)

Manuscript received —; revised —-.

TABLE I: Summary of the attacks on the SAFER family

Cipher	Rounds	Attack	Date Complexity	Time Complexity	Source
SAFER SK/128	2.75	ID	2^{39}CPs	2^{64}Enc	[14]
SAFER SK/128	3.75	ID	2^{45}CPs	2^{40}Enc	[16]
SAFER SK/128	4	ID	$2^{61.5}$CPs	2^{80}Enc	[17]
SAFER SK/128	4.75	LNH	2^{64}KPs	2^{90}Enc	[10]
SAFER SK/128	5	ZC	2^{64}KPs	$2^{113.3}$Enc	**This Paper**
SAFER+/128	2.75	ID	2^{64}CPs	2^{58}Enc	[14]
SAFER+/128	3.25	LNH	2^{101}KPs	2^{137}Enc	[10]
SAFER+/128	3.75	ID	2^{78}CPs	2^{73}Enc	[16]
SAFER+/128	4	ID	$2^{122.4}$KPs	2^{121}Enc	[17]
SAFER+/128	4	ZC	2^{128}KPs	2^{125}Enc	**This Paper**
SAFER+/256	3.75	ID	2^{78}CPs	2^{73}Enc	[16]
SAFER+/256	4	ID	$2^{124.4}$KPs	2^{216}Enc	[17]
SAFER+/256	5	ZC	2^{128}KPs	2^{191}Enc	**This Paper**
SAFER++/128	2.75	ID	2^{64}CPs	2^{59}Enc	[14]
SAFER++/128	3.25	LHN	2^{81}KPs	2^{101}Enc	[10]
SAFER++/128	3.75	ID	2^{23}CPs	2^{84}Enc	[15]
SAFER++/128	3.75	ID	2^{78}KPs	2^{64}Enc	[16]
SAFER++/128	4	Integral	2^{64}CPs	2^{120}Enc	[11]
SAFER++/128	4.25	Integral	$--$CPs	$--$Enc	[12]
SAFER++/128	4.5	Multiset	2^{48}CPs	2^{100}Enc	[13]
SAFER++/128	5	ID	2^{124}CPs	2^{118}Enc	[17]
SAFER++/128	5.5	Boomerang	2^{108}CPs	2^{108}Enc	[13]
SAFER++/128	5	ZC	2^{128}KPs	$2^{124.7}$Enc	**This Paper**
SAFER++/256	3.75	LNH	2^{81}KPs	2^{176}Enc	[10]
SAFER++/256	3.75	ID	2^{78}KPs	2^{72}Enc	[16]
SAFER++/256	4	Integral	2^{64}CPs	2^{152}Enc	[11]
SAFER++/256	4.75	Integral	$--$KPs	$--$Enc	[12]
SAFER++/256	5.5	ID	2^{124}CPs	2^{246}Enc	[17]
SAFER++/256	6	ZC	2^{128}KPs	2^{191}Enc	**This Paper**

LHN: Non-homomorphic Linear Attack; ID: Impossible Differential Attack; ZC: Zero-correlation Linear Cryptanalysis; ACC: Adaptive Chosen Ciphertext; CPs: Chosen Plaintexts; KPs: Known Plaintexts; Enc: Encryption.

2. The degradation properties can be applied to the zero-correlation linear attacks on 5-round SAFER SK/128 and 4(5)-round SAFER+/128(256), 5(6)-round SAFER++/128(256). We identify zero-correlation linear approximations for the SAFER block ciphers. As PHT employs the modular 256 addition as basic operation, which makes it difficult for us to understand the mask propagation under the nonzero-correlation condition. Fortunately, if the least significant bit is considered, the PHT can be converted to a 0,1 bit-chosen matrix with the XOR operation. In addition, the principle under the nonzero-correlation condition is used, that is, the most significant nonzero mask bits for the input and the output of modular addition are the same. Zero-correlation linear approximations for 2.75-round SAFER SK and SAFER+, and 3.75-round SAFER++ are constructed. Some of our attacks are the best attack in terms of the number of rounds.

Organization of our paper: In Sec.2, we give some preliminaries. Sec.3 shows two characterizations of the degraded Boolean function, and some degradation properties of the exponentiation type S-box are presented. Sec.4 demonstrates key recovery attacks on SAFER SK, SAFER+ and SAFER++ by means of the zero-correlation linear cryptanalysis. Finally, we summarize our work in Sec.5.

II. Preliminaries

An S-box with n-bit input and output can be represented by a vectorial Boolean function $F : F_2^n \mapsto F_2^n$, and

$$F = (f_1, f_2, ..., f_n),$$

where f_i is a n-bit Boolean function, $i = 1, 2, ..., n$.

Consider a function $F : F_2^n \mapsto F_2^m$ and let the input of the function be $x \in F_2^n$. A differential is given by a pair (δ, Δ) of an input difference δ and an output difference Δ and its probability is defined as

$$\Pr[\delta \xrightarrow{F} \Delta] = 2^{-n}|\{x \in F_2^n | F(x) \oplus F(x \oplus \delta) = \Delta\}|.$$

A linear approximation with the input mask α and the output mask β is the following function:

$$\beta \cdot F(x) \oplus \alpha \cdot x$$

and its correlation is defined as follows

$$\mathrm{Cor}_x\big(\beta \cdot F(x), \alpha \cdot x\big) = 2Pr_x\big(\beta \cdot F(x) \oplus \alpha \cdot x = 0\big) - 1.$$

The link between the differential probabilities and the correlations of linear approximations of vectorial Boolean functions was presented by Chabaud and Vaudenay [18].

Theorem 1: [18] Let $F : F_2^n \mapsto F_2^m$ be a Boolean function and $(\delta, \Delta) \in F_2^n \times F_2^m$, we have

$$\Pr[\delta \xrightarrow{F} \Delta] = 2^{-m} \sum_{u \in F_2^n, v \in F_2^m} (-1)^{u \cdot \delta \oplus v \cdot \Delta} \mathrm{Cor}_x^2(v \cdot F(x) \oplus u \cdot x).$$

For a Boolean function $f : F_2^n \to F_2$, by the theorem, we can obtain that

$$\sum_{u \in F_2^n} \mathrm{Cor}_x^2(f(x) \oplus u \cdot x) = 1.$$

III. Two Characterizations of the Degradation Boolean Function

We show two characterizations of the degraded Boolean function from the aspect of differential and linear. Firstly, we will introduce the definition of the degradation property of Boolean function.

Definition 1: For a Boolean function f with n-bit variable, if there exists a $n \times m$ $(m < n)$ matrix D with the rank $Rank(D) = m$ and a Boolean function $g : F_2^m \to F_2$,

$$f(x) = g(x \cdot D) = g(y),$$

then, f is a $(n - m)$-degradation Boolean function, where $y = x \cdot D$ and for Boolean function g, there exists no Boolean function h and $m \times m'$ $(m' < m)$ matrix D',

$$g(y) = h(y \cdot D') = h(z).$$

For a balance Boolean function f with n-bit variables, we have the following result.

Theorem 2: For two integrals n, m with $m < n$, the three conditions are equivalent.

(1) There exists a linear space V of n-bit values with the dimension $DimV = n - m$, for any n-bit value $\alpha \in V$,

$$\Pr_x\big(f(x) \oplus f(x \oplus \alpha) = 0\big) = 1;$$

(2) There exists a linear space W of n-bit values with the dimension $DimW = m$, for any n-bit value β, if

$$\mathrm{Cor}_x\big(f(x), \beta \cdot x\big) \neq 0,$$

then, we have $\beta \in W$;

(3) Boolean function f has the $(n-m)$-degradation property.

Proof of Theorem 2. It suffices to prove that (1) $\Rightarrow$ (2), (2) $\Rightarrow$ (1), (1) $\Rightarrow$ (3) and (3) $\Rightarrow$ (1).

(1) $\Rightarrow$ (2). Let the space W be an orthogonal linear space of V, that is,

$$W = \{\beta \in F_2^n |\ \beta \cdot \alpha = 0, for\ any\ \alpha\ \in V\},$$

then, we know

$$DimW = n - m.$$

By the definition, the correlation of linear approximation $f(x) \oplus \beta \cdot x$ is nonzero, if and only if

$$\mathrm{Pr}_x(f(x) \oplus \beta \cdot x = 0) \neq 1/2,$$

and for any $\alpha \in V$, we have

$$\begin{aligned}\mathrm{Pr}_x(f(x) \oplus \beta \cdot x = 0) &= \mathrm{Pr}_x(f(x \oplus \alpha) \oplus \beta \cdot (x \oplus \alpha) = 0) \\ &= \mathrm{Pr}_x(f(x) \oplus \beta \cdot x = \beta \cdot \alpha),\end{aligned}$$

then, the nonzero correlation leads to

$$\beta \cdot \alpha = 0,$$

for all $\alpha \in V$, that is,

$$\beta \in W.$$

(2) $\Rightarrow$ (1). Let the space V be an orthogonal linear space of the space W, and for any $\alpha \in V$, denote

$$\mathrm{P}_\alpha = \mathrm{Pr}_x\big(f(x) \oplus f(x \oplus \alpha)\big).$$

By (2), we know that if the linear approximation

$$\beta \cdot x \oplus f(x)$$

has correlation zero, then $\beta \in W$. By Theorem 1, we have

$$\begin{aligned}\mathrm{P}_\alpha &= 2^{-1} \sum_{\xi \in F_2^n, \gamma \in F_2} -1^{(\xi \cdot \alpha)}\ \mathrm{Cor}_x^2\big(\xi \cdot x \oplus \lambda \cdot f(x)\big) \\ &= 2^{-1} \sum_{\xi \in W} \mathrm{Cor}_x^2\big(\xi \cdot x \oplus f(x)\big) + 1/2 \\ &= 1,\end{aligned}$$

which means

$$f(x) = f(x \oplus \alpha),$$

for any $x \in F_2^n$ and $\alpha \in V$.

(1) $\Rightarrow$ (3). We know that V is a subspace of F_2^n. Then, we can find $x_1, x_2, ..., x_{2^m-1}$ and denote $V_0 = 0 \oplus V$ and $V_i = x_i \oplus V$, for $i = 1, 2, ..., 2^m - 1$, satisfying

$$V_0 \cap V_1 \cap V_2, ..., \cap V_{2^m-1} = \varnothing,$$

and

$$V_0 \cup V_1 \cup V_2, ..., \cup V_{2^m-1} = F_2^n.$$

By (1), we know that for any $y, z \in V_i$

$$f(y) = f(z) = f(x_i).$$

Let the space D be an orthogonal linear space of V, and $d_1, d_2, ..., d_m \in D$ be m linear independent n-bit values. Denote the $m \times n$ matrix

$$\bar{D} = \{d_1, d_2, ..., d_m\},$$

then for any $\alpha \in V$,

$$a \cdot \bar{D} = 0.$$

Let g be a Boolean function from F_2^m to F_2, and let

$$y_i = x_i \cdot \bar{D},$$

and

$$f(x_i) = g(y_i),$$

then for any $x \in F_2^n$, there exists a subspace V_i, $x \in V_i$, and

$$f(x) = f(x_i) = g(y_i).$$

(3) $\Rightarrow$ (1). The function f has a $(n-m)$-degradation property, there exist a $n \times m$ ($m < n$) matrix D with $Rank(D) = m$ and a Boolean function g,

$$f(x) = g(x \cdot D) = g(y).$$

Rewrite the matrix $D = (d_1, ..., d_m)$, and let

$$V = \{a \in F_2^n |\ a \cdot d_i = 0, i = 1, 2, ..., m\},$$

then, we know that the space V is an orthogonal linear space of D with

$$DimV = n - m.$$

It can be obtained that for any $\alpha \in V$, we have that

$$\begin{aligned}f(x) \oplus f(x \oplus a) &= g(x \cdot D) \oplus g(x \cdot D \oplus a \cdot D) \\ &= g(x \cdot D) \oplus g(x \cdot D) \\ &= 0.\end{aligned}$$

This, as we have proved, certainly implies the condition (1). Thus, we have proved the theorem.

The two characterizations showed by Theorem 2 can distinguish whether a Boolean function has degradation properties. The following theorem can be seen as a spacial case.

Theorem 3: For two integrals n, m and $m < n$, the following three conditions are equivalent.

(I) There exist m integrals $1 \leq i_1 < i_2 < ... < i_m \leq n$, for all n-bit value α satisfying the conditions that:

$$\alpha_i = \begin{cases} 0, & \text{when } i = i_1, i_2, ..., i_m; \\ 0 \text{ or } 1, & \text{other cases .} \end{cases}$$

we have

$$\mathrm{Pr}_x\big(f(x) \oplus f(x \oplus \alpha) = 0\big) = 1.$$

(II) For a n-bit value β, if

$$\mathrm{Cor}_x\big(f(x), \beta \cdot x\big) \neq 0,$$

then there exists $n - m$ integrals $1 \leq j_1 < j_2 < ... < j_{n-m} \leq n$,

$$\beta_j = 0,\ when\ \ j = j_1, j_2, ..., j_{n-m};$$

(III) Let d_i be n-bit values with $i = 1, 2, ...n$, and $d_1 = (1, 0, ..., 0)$, $d_2 = (0, 1, 0, ..., 0)$,..., $d_n = (0, 0, ..., 0, 1)$, then there exist m integrals $1 \leq r_1 < r_2 < ... < r_m \leq n$ and a Boolean function g,

$$f(x) = g(x \cdot D) = g(y),$$

where $D = (d_{r_1}, d_{r_2}, ..., d_{r_m})$ and $y = x \cdot D$.

Proof of Theorem 3. As (I) $\Rightarrow$ (III) and (III) $\Rightarrow$ (I) are easy to be proved, we need give prove (I) $\Rightarrow$ (II) and (II) $\Rightarrow$ (I).

(I) $\Rightarrow$ (II). By the definition of the correlation, the correlation of linear approximation $f(x) \oplus \beta \cdot x$ is nonzero, if and

only if

$$\mathrm{Pr}_x\big(f(x)\oplus\beta\cdot x=0\big)\neq 1/2,$$

and for any value α of (I), we have

$$\begin{aligned}\mathrm{Pr}_x\big(f(x)\oplus\beta\cdot x=0\big)&=\mathrm{Pr}_x\big(f(x\oplus\alpha)\oplus\beta\cdot(x\oplus\alpha)=0\big)\\&=\mathrm{Pr}_x\big(f(x)\oplus\beta\cdot x=\beta\cdot\alpha\big),\end{aligned}$$

then, the nonzero correlation condition leads to

$$\beta\cdot\alpha=0,$$

that is,

$$\beta_i=\begin{cases}0\ or\ 1, & \text{when } i=i_1,i_2,...,i_m;\\ 0, & \text{other cases .}\end{cases}$$

(II) $\Rightarrow$ (I). For the sake of simplicity, we assume $m=n-1$ and $\beta_1=0$. We claim that the function $f(x)$ is not affected by x_1, that is

$$f(0,x_2,x_3,...,x_n)=f(1,x_2,x_3,...,x_n),$$

for any $(x_2,x_3,...,x_n)\in F_2^{n-1}$.

By (2), we know that for all $(\beta_2,...,\beta_n)\in F_2^{n-1}$, the linear approximation

$$(1,\beta_2,...,\beta_n)\cdot(x_1,x_2,...,x_n)\oplus f(x)$$

has correlation zero.

Let $\alpha=(1,0,...,0)$ and $\mathrm{P}_\alpha=\mathrm{Pr}_x\big(f(x)\oplus f(x\oplus\alpha)\big)$, by Theorem 1, we have

$$\begin{aligned}\mathrm{P}_\alpha&=2^{-1}\sum_{\xi\in F_2^n,\gamma\in F_2}-1^{(\xi\cdot\alpha)}\ \mathrm{Cor}_x^2\big(\xi\cdot x\oplus\lambda\cdot f(x)\big)\\&=2^{-1}\sum_{\xi\in F_2^n}\mathrm{Cor}_x^2\big(\xi\cdot x\oplus f(x)\big)+1/2\\&=1,\end{aligned}$$

which means

$$f(0,x_2,x_3,...,x_n)=f(1,x_2,x_3,...,x_n),$$

for any $(x_2,x_3,...,x_n)\in F_2^{n-1}$. Thus, the theorem has been proved.

For a byte x, define S by

$$S(x)=(45^x\bmod 257)\bmod 256;$$

As 257 is prime and 45 is a primitive element modulo 257, then the vectorial Boolean function S is an invertible function of a byte, which can be denoted

$$S=(S_1,S_2,....,S_8),$$

where $S_i:F_2^8\to F_2$ is a balance function, $i=1,2,...,8$.

Proposition 1: Let S be the vectorial Boolean function defined above, then for the Boolean function S_7, we have

$$\mathrm{Pr}_x\big(S_7(x)\oplus S_7(x\oplus 80_x)=0\big)=1;$$

and for the Boolean function S_8,

$$\mathrm{Pr}_x\big(S_8(x)\oplus S_8(x\oplus 80_x)=1\big)=1.$$

where 0_x is the hexadecimal notation;

Proposition 2: For the linear approximations $\beta\cdot x\oplus S_7(x)$ and $\gamma\cdot x\oplus S_8(x)$, if

$$\mathrm{Cor}_x\big(S_7(x),\beta\cdot x\big)\neq 0,\ \text{and}\ \mathrm{Cor}_x\big(S_8(x),\gamma\cdot x\big)\neq 0,$$

then, $\beta_1=0$ and $\gamma_1=1$.

By either Proposition 1 or Proposition 2, we can deduce the following results from Theorem 3.

Theorem 4: Let S be the vectorial Boolean function defined above, then for the Boolean functions S_7, S_8, there exist two Boolean functions

$$S_7':F_2^7\to F_2,\quad S_8':F_2^7\to F_2,$$

we have

$$S_7(x)=S_7'(x_2,x_3,...,x_8),$$

and

$$S_8(x)=S_8'(x_2,x_3,...,x_8)\oplus x_1.$$

IV. Zero-correlation Linear Cryptanalysis of SAFER Family

A. Notations

$\boxminus$: modulo substraction;
$\boxplus$: modulo addition;
M^T : the transposition of matrix M;
$X\|Y$: the concatenation of X and Y;
$z[i]$: the i-th byte of z, and $'1'$ is the most significant byte;
$z_{[i]}$: the i-th bit of z, and $'1'$ is the most significant bit;
K_i : the i-th subkeys with the upper(lower)key K_i^1 (K_i^2);
I_i : the input of the upper key layer of the i-th round;
X_i : the input of the nolinear layer of the i-th round;
Y_i : the input of the lower key layer of the i-th round;
Z_i : the input of the linear layer of the i-th round;
P,C : the plaintext and the ciphertext.

B. Description of SAFER Block Cipher Family

The paper focuses on SAFER SK, SAFER+ and SAFER++ and we main give the description of SAFER SK. SAFER SK is a block cipher that operates on 64-bit blocks considered as 8 bytes. It consists of a round function iterated r times followed by a final output transformation. After the finial round, an additional key-addition similar to the upper key layer is added. Recommended values of r are 6 for SAFER SK-64 and 10 for SAFER SK-128. The i-th round function is built from four basic operations, see Fig.1.

1. **Upper Key Layer**: Bytes 1, 4, 5, 8 of the round input are XORed with the corresponding bytes of subkey K_i^1. Bytes 2, 3, 6, 7 of the round input are added bytewise (modulo 256) with the corresponding bytes of subkey K_i^1.

2. **Nonlinear Layer**: For a byte x, define X and L by

$$X(x)=(45^x\bmod 257)\bmod 256;\quad L(x)=\mathrm{Log}_{45}x\bmod 256;$$

with the special case that $L(0)=128$.

X is an invertible function of a byte, and L is defined to be its inverse. The X transformation is applied to bytes 1, 4, 5, 8, while the L transformation is applied to bytes 2, 3, 6, 7 with the output of the upper key layer.

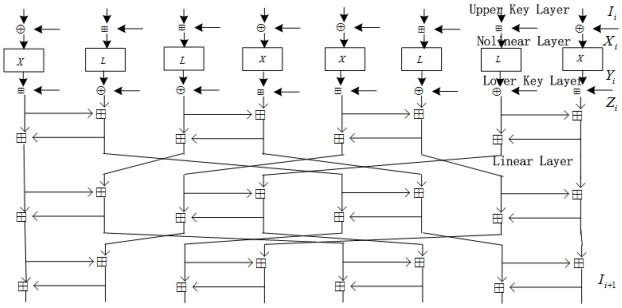


Fig. 1: The round function of SAFER SK

3. **Lower Key Layer**: Bytes 1, 4, 5, 8 of the output of the nonlinear layer are added bytewise (modulo 256) with the corresponding bytes of subkey K_i^2, while bytes 2, 3, 6, 7 of the output of the nonlinear layer are XORed with the corresponding bytes of subkey K_i^2

4. **Linear Layer**: The linear layer denotes a network of twelve 2-PHT boxes, where the latter is defined as

$$2-\text{PHT}(a;b) = (2a \boxplus b; a \boxplus b),$$

where addition is modulo 256. Denoting the input to a PHT layer by $Y = (y_1, y_2, y_3, y_4, y_5, y_6, y_7, y_8)$ and its output by $Z = (z_1, z_2, z_3, z_4, z_5, z_6, z_7, z_8)$, where $y_i, z_i \in F_2^8, 1 \leq i \leq 8$, this transformation can be described by $Z = Y^T \cdot M$, where M is called the PHT matrix:

$$M = \begin{pmatrix} 8 & 4 & 4 & 2 & 4 & 2 & 2 & 1 \\ 4 & 2 & 4 & 2 & 2 & 1 & 2 & 1 \\ 4 & 2 & 2 & 1 & 4 & 2 & 2 & 1 \\ 2 & 1 & 2 & 1 & 2 & 1 & 2 & 1 \\ 4 & 4 & 2 & 2 & 2 & 2 & 1 & 1 \\ 2 & 2 & 2 & 2 & 1 & 1 & 1 & 1 \\ 2 & 2 & 1 & 1 & 2 & 2 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

The final output transformation after r rounds is an application of the upper key layer with the output of the r-th round and the subkey K_r^1. Decryption involves the application of the inverse of each round with reverse order for the subkeys.

For the key schedules of SAFER SK/128, the master key $K = (k_1, k_2, ..., k_{16})$ is split into two parts, the first part including $k_1, k_2, ..., k_8$ and k^{sp1} is used in the upper key layer and the final key addition, while the second part including $k_9, k_{10}, ..., k_{16}$ and k^{sp2} is used in the lower key layer. k^{sp1} and k^{sp2} are computed in the following:

$$k^{sp1} = \bigoplus_{i=1}^{8} k_i; \qquad k^{sp2} = \bigoplus_{i=9}^{16} k_i;$$

The relations between first 6 round subkey bytes and the master key bytes for SAFER SK/128 are shown in Table 2.

SAFER+ and SAFER++ operate on 128-bit blocks considered as 16 bytes. SAFER+ uses four PHT layers composed of four 2-PHT layers with a particular fixed permutation between 2-PHT layers, while the linear layer of SAFER++ constructed with a permutation and two 4-PHT. The linear layers can be expressed by the matrices $M+$ and $M++$, which are shown in Appendix A.

TABLE II: Key schedules for the first 6-round SAFER SK/128.

K_1^1	k_9	k_{10}	k_{11}	k_{12}	k_{13}	k_{14}	k_{15}	k_{16}
K_1^2	k_2	k_3	k_4	k_5	k_6	k_7	k_8	k^{sp1}
K_2^1	k_{11}	k_{12}	k_{13}	k_{14}	k_{15}	k_{16}	k^{sp2}	k_9
K_2^2	k_4	k_5	k_6	k_7	k_8	k^{sp1}	k_1	k_2
K_3^1	k_{13}	k_{14}	k_{15}	k_{16}	k^{sp2}	k_9	k_{10}	k_{11}
K_3^2	k_6	k_7	k_8	k^{sp1}	k_1	k_2	k_3	k_4
K_4^1	k_{15}	k_{16}	k^{sp2}	k_9	k_{10}	k_{11}	k_{12}	k_{13}
K_4^2	k_8	k^{sp1}	k_1	k_2	k_3	k_4	k_5	k_6
K_5^1	k^{sp2}	k_9	k_{10}	k_{11}	k_{12}	k_{13}	k_{14}	k_{15}
K_5^2	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
K_6^1	k_{10}	k_{11}	k_{12}	k_{13}	k_{14}	k_{15}	k_{16}	k^{sp2}

The key size for SAFER+ can be 128 bits, 192 bits and 256 bits, and 128 and 256 bits are allowed for SAFER++. For SAFER+/128 or SAFER++/128, 16 master key bytes together with k^{sp1} will be used, where k^{sp1} is computed in the following:

$$k^{sp1} = \bigoplus_{i=1}^{16} k_i;$$

The 256 bit master keys for SAFER+/256 or SAFER++/256 is $K = (k_1, k_2, ..., k_{32})$. k^{sp1} and k^{sp2} are computed as

$$k^{sp1} = \bigoplus_{i=1}^{16} k_i, \qquad k^{sp2} = \bigoplus_{i=17}^{32} k_i;$$

The relations of the subkeys for SAFER+(++)/128 and SAFER+(++)/256 can be found in[8][19].

C. Basic ideas of zero-correlation linear cryptanalysis

In this subsection, we briefly recall the basic concepts of zero-correlation linear cryptanalysis. Zero-correlation linear cryptanalysis is one of the recent cryptanalytic methods introduced by Bogdanov and Rijmen[20], which is based on linear approximations with zero correlation, which is different from the traditional linear cryptanalysis where linear characteristics with high correlations are used.

Let $(\alpha \rightarrow \beta)$ be a zero-correlation linear distinguisher for the first $r-1$ rounds of an r-round n bit block cipher $E(\cdot, K)$. After partial decryption of the last round, the linear distinguisher to be evaluated becomes:

$$\alpha \cdot P \oplus \beta \cdot E_r^{-1}(C, K),$$

where $E_r^{-1}(\cdot)$ represents a partial decryption of the last round for the k bits of K and C that influence the value of the linear distinguisher. Then, if the guessed K is right, the correlation of the linear approximations must be zero, while the guessed K is wrong, the probability of the correlation of the linear approximations being zero is

$$\text{Pr} = \frac{1}{\sqrt{2\pi}} 2^{\frac{4-n}{2}},$$

which is extremely low probability, when the block size n is a big number.

D. Zero-correlation Linear Approximations of SAFER Family

To construct the zero-correlation linear approximations, one adopts the miss-in-the-middle techniques just like to find impossible differential. Any linear approximations with nonzero correlation is concatenated to any linear approximations with nonzero correlation in the inverse direction, where the intermediate masks states contradict with each other. For the properties of the propagation of linear masks over basic block cipher operations, we have the following lammas.

Lemma 1: [20] For a linear map $h(x) = M \cdot x$, where M is a value-chosen matrix with XOR operation. We have $C(\beta^T \cdot h(x), \alpha^T \cdot x) = 1$, if $\alpha = M^T \cdot \beta$, and $C(\beta^T \cdot h(x), \alpha^T \cdot x) = 0$, if $\alpha \neq M^T \cdot \beta$.

Lemma 2: [20] Let (α, β) be an approximation over an invertible function ϕ, Then $C(\beta^T \cdot h(x), \alpha^T \cdot x) \neq 0$ only if $\alpha = \beta = 0$ or both α and β are nonzero.

For the modular addition of two n-bit inputs x and y, the output z can be computed as:

$$z = (x + y) \ \mathbf{mod}\ 2^n.$$

It can be verified that

$$\big((x+y) \ \mathbf{mod}\ 2^n\big) \oplus \big((x'+y) \ \mathbf{mod}\ 2^n\big) = \big((x' \oplus x)+y\big) \ \mathbf{mod}\ 2^n,$$

can not always be the truth. It means modular addition is not a linear operation for XOR. However, if only the lest significant bit is taken into consideration, the modular addition can be treated as a linear operation.

Lemma 3: [21] Denote by Z (Z') the modular addition (subtraction) operations, for any linear approximations

$$\alpha \cdot x \oplus \beta \cdot y \oplus \gamma \cdot Z(Z'),$$

if the correlation is non-zero, then, the most significant non-zero mask bits for α, β and γ must be the same. Specifically, if $\gamma = 1$ and the correlation is non-zero , then $\alpha = \beta = 1$.

We show a 2.75-round zero-correlation linear approximation for SAFER SK and SAFER+, and a 3.75-round zero-correlation linear approximation for SAFER++.

Theorem 5: For SAFER SK, suppose that the output mask of the nonlinear layer in the i-th round is $(0, 0, 0, 02_x, 0, 0, 0, 0)$, and the output mask of the upper key layer in the $(i + 3)$-th round is $(01_x, 0, 0, 0, 0, 0, 0, 0)$, the correlation of such 2.75-round linear approximation is zero.

Proof of Theorem 5. We just need to show the process to construct the linear approximations adopting the miss-in-the-middle techniques, see Fig. 2.

Along the encryption direction: We consider the linear trail with nonzero correlation. Given the input mask $(0, 0, 0, 02_x, 0, 0, 0, 0)$ for the lower key layer Y_i of the i-th round, the input mask for the nonlinear layer X_{i+1} of the $(i+1)$-th round must have the form $(a_1, a_2, a_3, a_4, a_5, a_6, 0, 0)$, where $a_1, a_2, a_3, a_4, a_5, a_6 \in F_2^8$ are unknown values, because $Y_i[4]_8$ and $X_{i+1}[7, 8]$ are two independence variables.

Along the decryption direction: Given the output mask $(01_x, 0, 0, 0, 0, 0, 0, 0)$ for the upper key layer of the $(i + 3)$-th round, by the 0,1 bit-chosen matrix with XOR operation converted from PHT, the input mask of the PHT is $(0, 0, 0, 0, 0, 0, 0, 01_x)$. Then by Theorem 3 and Lemma 3,

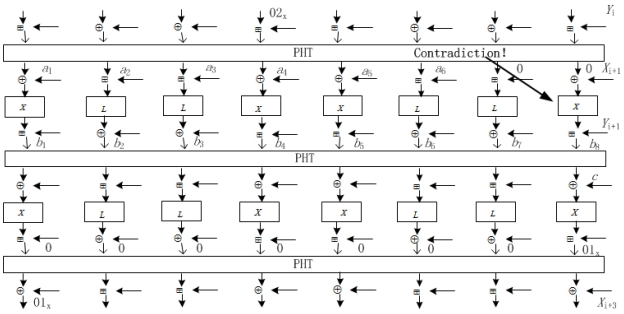


Fig. 2: 2.75-round Distinguisher of SAFER SK

the input mask of the linear layer Z_{i+1} of the $i + 1$-th round must have the form $(b_1, b_2, b_3, b_4, b_5, b_6, b_7, b_8)$, where $b_1, b_2, b_3, b_4, b_5, b_6, b_7, b_8$ are unknown nonzero values.

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

(Bit-chosen Matrix of the PHT (M))

Contradiction: For the 7-th and 8-th S-boxes in the nonlinear layer of the $(i + 1)$-th round, the input masks are zero, but the output masks are nonzero, which is a contradiction. Thus, the linear hull is a zero-correlation linear hull.

Theorem 6: For SAFER SK+, suppose that the output mask of the nonlinear layer in the i-th round is $(0, 0, 0, 0, 0, 0, 0, 0; 02_x, 0, 0, 0, 0, 0, 0, 0,)$ and the output mask of the upper key layer in the $(i + 3)$-th round is $(0, 0, 01_x, 01_x, 0, 0, 01_x, 01_x; 0, 0, 0, 0, 0, 01_x, 0, 0, 0)$, the correlation of such 2.75-round linear approximation is zero.

Theorem 7: For SAFER SK++, suppose that the output mask of the nonlinear layer in the i-th round is $(02_x, 0, 0, 0, 0, 0, 0, 0; 0, 0, 0, 0, 0, 0, 0, 0,)$, and the output mask of the upper key layer in the $(i + 4)$-th round is $(01_x, 01_x, 01_x, 0, 0, 0, 0, 0; 0, 0, 01_x, 01_x, 0, 0, 0, 0)$, the correlation of such 3.75-round linear approximation is zero.

The proofs of Theorem 6 and Theorem 7 are similar to the proof of Theorem 5. The processing of the construction of the linear approximations are showed in Appendix B.

E. Zero-correlation Linear Cryptanalysis of SAFER Family

In this section, we show our attack on 5-round SAFER SK/128 and 4(5)-round SAFER+/128(256), 5(6)-round SAFER++/128(256). The zero-correlation linear approximations and the degradation properties of the S-boxes are used. We extend some rounds forward and backward of the linear approximations, the general attack procedure is to partially encrypt each plaintext and to partially decrypt the corresponding ciphertext with a guess of subkey bits. In the attacks, the partial-sum techniques are used to reduce the compute complexity.

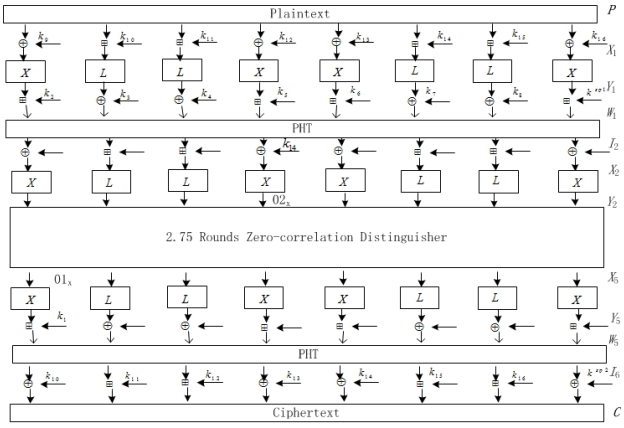


Fig. 3: Attacks on 5-round SAFER SK/128.

1) Key Recovery Attacks on SAFER SK : In this subsection, we will attack 5-round SAFER SK/128 based on 2.75-round zero-correlation linear approximations. We mount the 2.75-round zero-correlation linear approximations from round 1.5 to round 4.25, and extend 1.5 rounds forward and 0.75 rounds backward respectively, see Fig.3.

Attack Process. The key recovery attacks on 5-round SAFER SK/128 are proceeded with the partial-sum technique as follows.

1. Collect all the 2^{64} plaintext-ciphertext pairs (P, C). Allocate 8-bit counters $N_1[y_1]$ for 2^{104} possible values of

$$y_1 = P[1,2,3,4,5,6,7,8] \| C[1,4,5,8] \| M^1,$$

and initialize them to zero, where M^1 is a 8-bit value with

$$M^1 = C[6] \boxplus C[7] \boxminus C[2] \boxminus C[3].$$

For every (P, C) pair, extract the value of y_1 and increase the corresponding counter $N_1[y_1]$.

2. Allocate 8-bit counters $N_2[y_2]$ for 2^{59} possible values of

$$y_2 = P[1,3,4,7,8] \| Y_1[2]_{[3-8]} \| Y_1[5]_{[3-8]} \| Y_1[6]_{[3-8]} \| X_5[1]_{[8]},$$

and initialize them to zero. Guess $k_6^1[1,4,5,8]$ and $k_6^1[6] \boxplus k_6^1[7] \boxminus k_6^1[2] \boxminus k_6^1[3] \boxplus k_5^2[1]$, that is, $k_{10}, k_{13}, k_{14}, k^{sp2}$ and $k_{15} \boxplus k_{16} \boxminus k_{12} \boxminus k_{13} \boxplus k_1$, and partially decrypt y_1 to get the value of y_2, then update the corresponding counter by $N_2[y_2] += N_1[y_1]$.

3. Allocate a counter $N_3[y_3]$ for 2^{51} possible values of

$$y_3 = P[3,4,7,8] \| Y_1[2]_{[3-8]} \| Y_1[5]_{[3-8]} \| Y_1[6]_{[3-8]} \| X_5[1]_{[8]},$$

and initialize them to zero. Guess $k_1^1[1]$, that is, k_9, compute

$$Y_1[5]_{[3-8]} = Y_1[5]_{[3-8]} \boxplus Y_1[1]_{[3-8]},$$

and partially decrypt y_2 to get the value of y_3, then update the corresponding counter by $N_3[y_3] += N_2[y_2]$.

4. Allocate a counter $N_4[y_4]$ for 2^{44} possible values of

$$y_4 = P[3,7,8] \| Y_1[2]_{[3-8]} \| Y_1[5]_{[2-8]} \| Y_1[6]_{[3-8]} \| X_5[1]_{[8]},$$

and initialize them to zero. Guess $k_1^1[4]$, that is, k_{12}, compute

$$Y_1[5]_{[2-8]} = (Y_1[5]_{[3-8]} \| 0) \boxplus Y_1[4]_{[2-8]},$$

and partially decrypt y_3 to get the value of y_4, then update the corresponding counter by $N_4[y_4] += N_3[y_3]$.

5. Allocate a counter $N_5[y_5]$ for 2^{36} possible values of

$$y_5 = P[3,7] \| Y_1[2]_{[3-8]} \| Y_1[5]_{[2-8]} \| Y_1[6]_{[3-8]} \| X_5[1]_{[8]},$$

and initialize them to zero. Guess $k_1^1[8]$, that is, k_{16}, compute

$$Y_1[5]_{[2-8]} = Y_1[5]_{[2-8]} \boxplus Y_1[8]_{[2-8]},$$

and partially decrypt y_4 to get the value of y_5, then update the corresponding counter by $N_5[y_5] += N_4[y_4]$.

6. Allocate a counter $N_6[y_6]$ for 2^{30} possible values of

$$y_6 = P[3,7] \| Y_1[5]_{[2-8]} \| Y_1[6]_{[3-8]} \| X_5[1]_{[8]},$$

and initialize them to zero. Guess $k_1^2[2]_{[3-8]}$, that is, $(k_3)_{[3-8]}$, and get $k_1^1[2](k_{10})$ from the keys guessed before, and compute

$$Y_1[5]_{[2-8]} = Y_1[5]_{[2-8]} \boxplus (Y_1[2]_{[3-8]} \| 0),$$

and partially decrypt y_5 to get the value of y_6, then update the corresponding counter by $N_6[y_6] += N_5[y_5]$.

7. Allocate a counter $N_7[y_7]$ for 2^{24} possible values of

$$y_7 = P[3,7] \| Y_1[5]_{[2-8]} \| X_5[1]_{[8]},$$

and initialize them to zero. Guess $k_1^2[6]_{[3-8]}$, that is, $(k_7)_{[3-8]}$, and get $k_1^1[6]$ (k_{14}) from the keys guessed before, compute

$$Y_1[5]_{[2-8]} = Y_1[5]_{[2-8]} \boxplus (Y_1[6]_{[3-8]} \| 0),$$

and partially decrypt y_6 to get the value of y_7, then update the corresponding counter by $N_7[y_7] += N_6[y_6]$.

8. Allocate a counter $N_8[y_8]$ for 2^{16} possible values of

$$y_8 = P[7] \| Y_1[5]_{[2-8]} \| X_5[1]_{[8]},$$

and initialize them to zero. Guess $k_1^1[3]$ and $k_1^2[3]_{[2-8]}$, that is, k_{11} and $(k_4)_{[2-8]}$, compute

$$Y_1[5]_{[2-8]} = Y_1[5]_{[2-8]} \boxplus Y_1[3]_{[2-8]},$$

and partially decrypt y_7 to get the value of y_8, then update the corresponding counter by $N_8[y_8] += N_7[y_7]$.

9. Allocate a counter $N_9[y_9]$ for 2^8 possible values of

$$y_9 = Y_1[5]_{[2-8]} \| X_5[1]_{[8]},$$

and initialize them to zero. Guess $k_1^2[7]_{[3-8]}$, that is, $(k_8)_{[2-8]}$, and get $k_1^1[7](k_{15})$ from the keys guessed before and the key schedules, compute

$$Y_1[5]_{[2-8]} = Y_1[5]_{[2-8]} \boxplus Y_1[7]_{[2-8]},$$

and partially decrypt y_8 to get the value of y_9, then update the corresponding counter by $N_9[y_9] += N_8[y_8]$.

10. Allocate a counter $N_{10}[y_{10}]$ for 2 possible values of

$$y_{10} = X_5[1]_{[8]},$$

and initialize them to zero. Guess $(8k_1^2[1] \boxplus 2k_1^2[4] \boxplus 4k_1^2[5] \boxplus k_1^2[8] \oplus k_2^4)_{[2-8]}$, that is, $(8k_2 \boxplus 2k_5 \boxplus 4k_6 \boxplus k^{sp1} \oplus k_{14})_{[2-8]}$, compute

$$X_5[1]_{[8]} = X_5[1]_{[8]} \boxplus Y_2[4]_{[7]},$$

and partially decrypt y_9 to get the value of y_{10}, then update the corresponding counter by $N_{10}[y_{10}] += N_9[y_9]$.

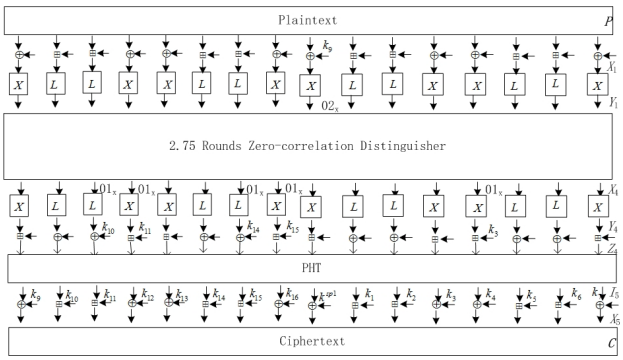


Fig. 4: Attacks on 4-round SAFER+/128.

11. After Step 10, 118 key bits have been guessed. If $N_{10}[0] \neq 2^{63}$, then, discard the guessed keys and guess another subkey until we get the correct subkey. We do exhaustive search for all keys conforming to this possible key candidate.

Complexity of the Attack. There are 118-bit key value guessed during the encryption phase, according to the wrong-key randomization hypothesis, the probability that a wrong subkey candidate can pass the test in Step 11 is approximately

$$\frac{1}{2\sqrt{\pi}} 2^{\frac{4-64}{2}} \approx 2^{-31.3},$$

thus about $2^{96} \times 2^{-31.3} = 2^{64.7}$ subkey candidates will be left.

(1) Step 1 and Step 2 require 2^{64} and $2^{64} \times 2^{40} = 2^{104}$ memory accesses;
(2) Step 3-7 requires $2^{107}+2^{108}+2^{108}+2^{108}+2^{108} \approx 2^{110.3}$ memory accesses;
(3) Step 8-10 requires $2^{115} + 2^{115} + 2^{115} \approx 2^{116.6}$ memory accesses;
(4) Step 11 requires $2^{32} \times 2^{64.7} = 2^{96.7}$ 5-round SAFER SK encryption. Because, there are 32 bit keys being not guessed during the encryption phase.

If we assume that processing each memory accesses is equivalent to 1/2 round encryption, then, the total time complexity is about $2^{116.6} \times 1/2 \times 1/5 \approx 2^{113.3}$ 5-round encryptions. In total, The data complexity of this attack is 2^{64} known plaintexts, the time complexity is about $2^{113.3}$ 5-round encryptions. The memory complexity is primarily owing to storing the vectors N_1, thus, the memory requirement are 2^{102} bytes for counters.

2) Key Recovery Attacks on SAFER+ : In this subsection, we will show our attacks on 4-round SAFER+/128 and 5-round SAFER+/256. For the attack on 4-round SAFER+/128, we mount the 2.75-round zero-correlation linear approximations from round 0.5 to round 3.25, and extend 0.5 round forward and 0.75 round backward, see Fig.4.

Attack Process. The key recovery attacks on 4-round SAFER+/128 are proceeded with the partial-sum technique as follows.

1. Collect all the 2^{128} plaintext-ciphertext pairs (P, C). Allocate 8-bit counters $N_1[y_1]$ for 2^{112} possible values of

$$y_1 = P[9] \| C[1, 4, 5, 8, 9, 12, 13, 16] \| M^1 \| M^2 \| M^3 \| M^4 \| M^5,$$

and initialize them to zero, where M^1, M^2, M^3, M^4, M^5 are 8-bit values with

$$\begin{aligned}
M^1 &= C[3] \boxminus 2C[2] \boxminus 2C[6] \boxplus C[7] \boxminus C[10] \\
&\quad \boxplus C[11] \boxminus 4C[14] \boxminus 4C[15]; \\
M^2 &= 4C[2] \boxminus C[3] \boxplus 4C[6] \boxminus C[7] \boxplus 2C[10] \\
&\quad \boxminus C[11] \boxplus 4C[14] \boxminus 8C[15]; \\
M^3 &= C[3] \boxminus 8C[2] \boxminus C[6] \boxplus C[7] \boxminus 2C[10] \\
&\quad \boxplus 2C[11] \boxminus 2C[14] \boxplus C[15]; \\
M^4 &= 16C[2] \boxminus C[3] \boxplus 2C[6] \boxminus C[7] \boxplus 4C[10] \\
&\quad \boxminus 2C[11] \boxplus 2C[14] \boxminus 2C[15]; \\
M^5 &= 2C[3] \boxminus C[2] \boxminus 8C[6] \boxplus C[7] \boxminus 2C[10] \\
&\quad \boxplus C[11] \boxminus 2C[14] \boxplus C[15];
\end{aligned}$$

For every (P, C) pair, extract the value of y_1 and increase the corresponding counter $N_1[y_1]$.

2. Allocate 8-bit counters $N_2[y_2]$ for 2^{97} possible values of

$$y_2 = C[4, 5, 8, 9, 12, 13, 16] \| M^1 \| M^2 \| M^3 \| M^4 \| M^5 \| Y_1[1]_{[7]},$$

and initialize them to zero. Guess $k_1^9[1]$, that is, k_9 and compute

$$M^1 = M^1 \boxplus (C[1] \oplus k_9); M^2 = M^2 \boxminus 2(C[1] \oplus k_9);$$
$$M^3 = M^3 \boxplus 4(C[1] \oplus k_9); M^4 = M^4 \boxminus 8(C[1] \oplus k_9);$$
$$M^5 = M^5 \boxplus (C[1] \oplus k_9),$$

and partially decrypt y_1 to get the value of y_2, then update the corresponding counter by $N_2[y_2] += N_1[y_1]$.

The following steps in the partial decryption phase are similar to Step 2. We use Table III to show the details of each step of the partial decryption, and in the table, we denote

$$K^{(1)} = -2k_{10} \boxplus k_{11} \boxminus 2k_{14} \boxplus k_{15} \boxminus k_1 \boxplus k_2 \boxminus 4k_5 \boxminus 4k_6 \oplus k_{10};$$
$$K^{(2)} = 4k_{10} \boxminus k_{11} \boxplus 4k_{14} \boxminus k_{15} \boxplus 2k_1 \boxminus k_2 \boxplus 4k_5 \boxminus 8k_6 \boxplus k_{11};$$
$$K^{(3)} = -8k_{10} \boxplus k_{11} \boxminus k_{14} \boxplus k_{15} \boxminus 2k_1 \boxplus 2k_2 \boxminus 2k_5 \boxplus k_6 \oplus k_{14};$$
$$K^{(4)} = 16k_{10} \boxminus k_{11} \boxplus 2k_{14} \boxminus k_{15} \boxplus 4k_1 \boxminus 2k_2 \boxplus 2k_5 \boxminus 2k_6 \boxplus k_{15};$$
$$K^{(5)} = -k_{10} \boxplus 2k_{11} \boxminus 8k_{14} \boxplus k_{15} \boxminus 2k_1 \boxplus k_2 \boxminus 2k_5 \boxplus k_6 \boxplus k_3.$$

15. After Step 14, 104 key bits have been guessed. If $N_{13}[0] \neq 2^{127}$, then, discard the guessed keys and guess another subkey until we get the correct subkey. We do exhaustive search for all keys conforming to this possible key candidate.

TABLE III: Partial decryption of the attack on 4-round SAFER+/128.

Step	Guess Keys	Computed States
3	k_{12}	$y_4 = C[4, 5, 8, 12, 13, 16] \| M^1 \| M^2 \| M^3 \| M^4 \| M^5 \| Y_1[9]_{[7]}$
4	k_{13}	$y_5 = C[5, 8, 12, 13, 16] \| M^1 \| M^2 \| M^3 \| M^4 \| M^5 \| Y_1[9]_{[7]}$
5	k_{16}	$y_6 = C[8, 12, 13, 16] \| M^1 \| M^2 \| M^3 \| M^4 \| M^5 \| Y_1[9]_{[7]}$
6	k^{sp1}	$y_7 = C[12, 13, 16] \| M^1 \| M^2 \| M^3 \| M^4 \| M^5 \| Y_1[9]_{[7]}$
7	k_3	$y_8 = C[13, 16] \| M^1 \| M^2 \| M^3 \| M^4 \| M^5 \| Y_1[9]_{[7]}$
8	k_4	$y_7 = C[16] \| M^1 \| M^2 \| M^3 \| M^4 \| M^5 \| Y_1[9]_{[7]}$
9	k_7	$y_8 = M^1 \| M^2 \| M^3 \| M^4 \| M^5 \| Y_1[9]_{[7]}$
10	$K^{(1)}$	$y_9 = M^2 \| M^3 \| M^4 \| M^5 \| Y_1[9]_{[7]} \oplus X_4[3]_{[8]}$
11	$K^{(2)}$	$y_{10} = M^3 \| M^4 \| M^5 \| Y_1[9]_{[7]} \oplus X_4[3]_{[8]} \oplus X_4[4]_{[8]}$
12	$K^{(3)}$	$y_{11} = M^4 \| M^5 \| Y_1[9]_{[7]} \oplus X_4[3]_{[8]} \oplus X_4[4]_{[8]} \oplus X_4[7]_{[8]}$
13	$K^{(4)}$	$y_{12} = M^5 \| Y_1[9]_{[7]} \oplus (X_4[3] \oplus X_4[4] \oplus X_4[7] \oplus X_4[8])_{[8]}$
14	$K^{(5)}$	$y_{13} = Y_1[9]_{[7]} \oplus (X_4[3] \oplus X_4[4] \oplus X_4[7] \oplus X_4[8] \oplus X_4[8])_{[8]}$

TABLE IV: Partial decryption of the attack on 5-round SAFER++/128.

Step	Guess Keys	Computed States
3	k_{11}	$y_3 = C[4,5,9,12,13,16] \| M^1 \| M^2 \| M^3 \| M^4 \| M^5 \| Y_1[1]_{[7]}$
4	k_{14}	$y_4 = C[5,9,12,13,16] \| M^1 \| M^2 \| M^3 \| M^4 \| M^5 \| Y_1[1]_{[7]}$
5	k_{15}	$y_5 = C[9,12,13,16] \| M^1 \| M^2 \| M^3 \| M^4 \| M^5 \| Y_1[1]_{[7]}$
6	k_2	$y_6 = C[12,13,16] \| M^1 \| M^2 \| M^3 \| M^4 \| M^5 \| Y_1[1]_{[7]}$
7	k_5	$y_7 = C[13,16] \| M^1 \| M^2 \| M^3 \| M^4 \| M^5 \| Y_1[1]_{[7]}$
8	k_6	$y_8 = C[16] \| M^1 \| M^2 \| M^3 \| M^4 \| M^5 \| Y_1[1]_{[7]}$
9	k_9	$y_9 = M^1 \| M^2 \| M^3 \| M^4 \| M^5 \| Y_1[1]_{[7]}$
10	$K^{(1)}$	$y_{10} = M^2 \| M^3 \| M^4 \| M^5 \| Y_1[1]_{[7]} \oplus X_5[1]_{[8]}$
11	$K^{(2)}$	$y_{11} = M^3 \| M^4 \| M^5 \| Y_1[1]_{[7]} \oplus X_5[1]_{[8]} \oplus X_5[2]_{[8]}$
12	$K^{(3)}$	$y_{12} = M^4 \| M^5 \| Y_1[1]_{[7]} \oplus X_5[1]_{[8]} \oplus X_5[2]_{[8]} \oplus X_5[3]_{[8]}$
13	$K^{(4)}$	$y_{13} = M^5 \| Y_1[1]_{[7]} \oplus (X_5[1] \oplus X_5[2] \oplus X_5[3] \oplus X_5[11])_{[8]}$
14	$K^{(5)}$	$y_{14} = Y_1[1]_{[7]} \oplus (X_5[1] \oplus X_5[2] \oplus X_5[3] \oplus X_5[11] \oplus X_5[12])_{[8]}$

Complexity of the Attack. In this attack, there are 104 bit key values guessed during the encryption phase, about $2^{104} \times 2^{-31.3} = 2^{78.7}$ subkey candidates will be left. Step 1 requires 2^{128} memory accesses; Step 2 requires $2^{112} \times 2^8 = 2^{120}$ memory accesses; Step 3-14 requires $11 \times 2^{113} \approx 2^{116.3}$ memory accesses; Step 15 requires $2^{24} \times 2^{78.7}$ 4-round SAFER+ encryption. The total time complexity is about $2^{128} \times 1/2 \times 1/4 = 2^{125}$ 4-round encryptions. The data complexity of this attack is 2^{128} known plaintexts, the time complexity is about 2^{125} 4-round encryptions. The memory complexity is primarily owing to storing the vectors N_1, thus, the memory requirement are 2^{110} bytes for counters.

For the attack on 5-round SAFER+/256, we mount the 2.75-round zero-correlation linear approximations from round 1.5 to round 4.25, and extend 1.5 rounds forward and 0.75 rounds backward. We proceed similar steps to attack 5-round SAFER+/256. The data complexity of the attack is 2^{128} known plaintexts. The total time complexity is 2^{191} encryptions and the memory complexity is about 2^{230} bytes, which is showed in Appendix C(A).

3) Key Recovery Attacks on SAFER++ : In this subsection, we will show our attacks on 5-round SAFER++/128 and 6-round SAFER++/256. For the attack on 5-round SAFER++/128, We mount the 3.75 rounds zero-correlation linear approximations from round 0.5 to round 4.25, and extend 0.5 round forward and 0.75 round backward, see Fig.5.

Attack Process. The key recovery attacks on 5 round of SAFER++/128 are proceeded with the partial-sum technique as follows.

1. Collect all the 2^{128} plaintext-ciphertext pairs (P, C). Allocate 8-bit counters $N_1[y_1]$ for 2^{112} possible values of

$$y_1 = P[1] \| C[1,4,5,8,9,12,13,16] \| M^1 \| M^2 \| M^3 \| M^4 \| M^5,$$

and initialize them to zero, where M^1, M^2, M^3, M^4, M^5 are 8-bit values with

$$M^1 = C[6] \boxplus C[7]; \quad M^2 = C[10] \boxplus C[11] \boxplus C[14];$$
$$M^3 = C[3] \boxplus C[14] \boxplus C[15]; \quad M^4 = C[6] \boxplus C[7] \boxplus C[15];$$
$$M^5 = -C[7] \boxminus 4C[10] \boxminus 4C[11] \boxminus C[14].$$

For every (P, C) pair, extract the value of y_1 and increment the corresponding counter $N_1[y_1]$.

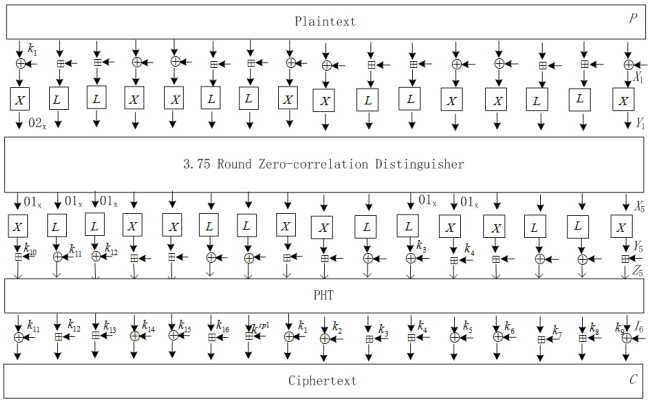


Fig. 5: Attacks on 5-round SAFER++/128.

2. Allocate 8-bit counters $N_2[y_2]$ for 2^{97} possible values of

$$y_2 = C[1,4,5,9,12,13,16] \| M^1 \| M^2 \| M^3 \| M^4 \| M^5 \| Y_1[1]_{[7]},$$

and initialize them to zero. Guess $k_1^1[1]$, that is, k_1 and compute

$$M^1 = M^1 \boxminus 4(C[8] \oplus k_1), M^4 = M^4 \boxminus 4(C[8] \oplus k_1),$$
$$M^4 = M^5 \boxplus (C[8] \oplus k_1),$$

and partially decrypt y_1 to get the value of y_2, then update the corresponding counter by $N_2[y_2] += N_1[y_1]$.

The following steps in the partial decryption phase are similar to Step 2. Thus, to be consistent, we use Table IV to show the details of each step of the partial decryption, and in the table, we denote

$$K^{(1)} = k_{14} \boxplus k_{15} \boxplus k_{16} \boxminus 4k^{sp1} \boxplus k_1 \boxminus k_4 \boxplus k_{19};$$
$$K^{(2)} = k_1 \boxplus k_2 \boxplus k_3 \boxminus 4k_4 \boxplus k_6 \boxminus k_8 \oplus k_{11};$$
$$K^{(3)} = k_{12} \boxminus k_{13} \boxplus k_5 \boxplus k_6 \boxplus k_7 \boxminus 4k_8 \oplus k_{12};$$
$$K^{(4)} = k_{14} \boxplus k_{15} \boxplus k_{16} \boxminus 4k^{sp1} \boxplus k_7 \boxminus k_8 \oplus k_3;$$
$$K^{(5)} = -k_{10} \boxplus k_{13} \boxminus k_{16} \boxplus k^{sp1} \boxminus 4k_1 \boxminus 4k_2 \boxminus 4k_3 \boxplus 16k_4 \boxminus k_6 \boxplus k_4.$$

15. After Step 14, 104 key bits have been guessed. If $N_{13}[0] \neq 2^{127}$, then, discard the guessed keys and guess another subkey until we get the correct subkey. We do exhaustive search for all keys conforming to this possible key candidate.

Complexity of the Attack. In this attack,there are 104-bit key value guessed during the encryption phase, about $2^{104} \times 2^{-31.3} = 2^{78.7}$ subkey candidates will be left. Step 1 requires 2^{128} memory accesses; Step 2 requires $2^{112} \times 2^8 = 2^{120}$ memory accesses; Step 3-14 requires about $11 \times 2^{113} \approx 2^{116.3}$ memory accesses; Step 15 requires $2^{24} \times 2^{78.7}$ 5-round SAFER++ encryption.

The total time complexity is about $2^{128} \times 1/2 \times 1/5 = 2^{124.7}$ 5-round encryptions. The data complexity of this attack is 2^{128} known plaintexts, the time complexity is about $2^{124.7}$ 5-round encryptions. The memory complexity is primarily owing to storing the vectors N_1, thus, the memory requirement are 2^{110} bytes for counters.

For the attack on 6-round SAFER++/256, we mount the 3.75-round zero-correlation linear approximations from round

1.5 to round 5.25, and extend 1.5 rounds forward and 0.75 rounds backward. We proceed similar steps to attack 6-round SAFER+/256. The data complexity of the attack is 2^{128} known plaintexts. The total time complexity is 2^{191} encryptions and the memory complexity is about 2^{228} bytes, which is showed in Appendix C(B).

V. Conclusion

In this paper, we mainly investigate the degradation properties of Boolean function from the aspects of the distributions of the differences and linear masks. Two characterizations have been shown and some degradation properties of the exponentiation type S-box of the SAFER block ciphers are discovered. Moreover, those observations are applied to evaluate the security of the SAFER block cipher family by means of zero-correlation linear cryptanalysis. Key recovery attacks on 5-round SAFER SK/128, 4(5) round of SAFER+/128(256), 5(6)-round SAFER++/128(256) have been presented. Some of our attacks are the best attacks in term of the number of rounds.

Appendix A

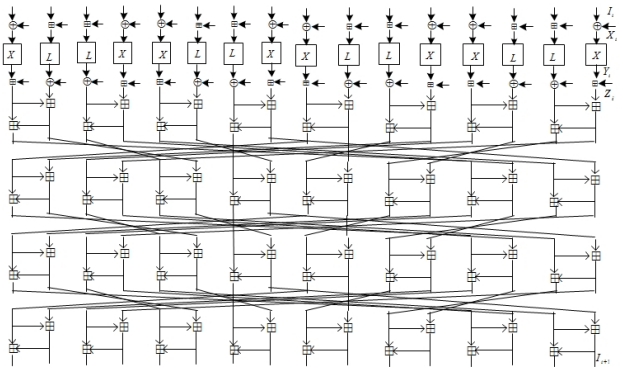
A. The matrix $M+$ of SAFER+

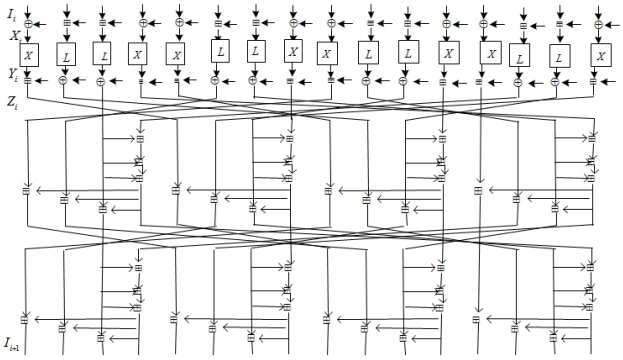
$$\begin{pmatrix}
2 & 2 & 1 & 1 & 16 & 8 & 2 & 1 & 4 & 2 & 4 & 2 & 1 & 1 & 4 & 4 \\
1 & 1 & 1 & 1 & 8 & 4 & 2 & 1 & 2 & 1 & 4 & 2 & 1 & 1 & 2 & 2 \\
1 & 1 & 4 & 4 & 2 & 1 & 4 & 2 & 4 & 2 & 16 & 8 & 2 & 2 & 1 & 1 \\
1 & 1 & 2 & 2 & 2 & 1 & 2 & 1 & 4 & 2 & 8 & 4 & 1 & 1 & 1 & 1 \\
4 & 4 & 2 & 1 & 4 & 2 & 4 & 2 & 16 & 8 & 1 & 1 & 1 & 1 & 2 & 2 \\
2 & 2 & 2 & 1 & 2 & 1 & 4 & 2 & 8 & 4 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 4 & 2 & 4 & 2 & 16 & 8 & 2 & 1 & 2 & 2 & 4 & 4 & 1 & 1 \\
1 & 1 & 2 & 1 & 4 & 2 & 8 & 4 & 2 & 1 & 1 & 1 & 2 & 2 & 1 & 1 \\
2 & 1 & 16 & 8 & 1 & 1 & 2 & 2 & 1 & 1 & 4 & 4 & 4 & 2 & 4 & 2 \\
2 & 1 & 8 & 4 & 1 & 1 & 1 & 1 & 1 & 1 & 2 & 2 & 4 & 2 & 2 & 1 \\
4 & 2 & 4 & 2 & 4 & 4 & 1 & 1 & 2 & 2 & 1 & 1 & 16 & 8 & 2 & 1 \\
2 & 1 & 4 & 2 & 2 & 2 & 1 & 1 & 1 & 1 & 1 & 1 & 8 & 4 & 2 & 1 \\
4 & 2 & 2 & 2 & 1 & 1 & 4 & 4 & 1 & 1 & 4 & 2 & 2 & 1 & 16 & 8 \\
4 & 2 & 1 & 1 & 1 & 1 & 2 & 2 & 1 & 1 & 2 & 1 & 2 & 1 & 8 & 4 \\
16 & 8 & 1 & 1 & 2 & 2 & 1 & 1 & 4 & 4 & 2 & 1 & 4 & 2 & 4 & 2 \\
8 & 4 & 1 & 1 & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 1 & 2 & 1 & 4 & 1
\end{pmatrix}$$

B. The matrix $M++$ of SAFER++

$$\begin{pmatrix}
1 & 2 & 1 & 1 & 1 & 1 & 1 & 1 & 4 & 2 & 2 & 2 & 1 & 1 & 2 & 1 \\
2 & 1 & 1 & 1 & 1 & 1 & 2 & 1 & 1 & 1 & 1 & 1 & 2 & 4 & 2 & 2 \\
2 & 2 & 4 & 2 & 2 & 1 & 1 & 1 & 1 & 2 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 2 & 1 & 1 & 1 & 1 & 2 & 1 & 2 & 1 & 1 & 1 \\
4 & 2 & 2 & 1 & 1 & 1 & 2 & 1 & 1 & 1 & 1 & 1 & 1 & 2 & 1 & 1 \\
1 & 1 & 2 & 1 & 2 & 1 & 1 & 1 & 2 & 4 & 2 & 2 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 2 & 1 & 1 & 2 & 2 & 4 & 2 & 2 & 1 & 1 & 1 \\
1 & 2 & 1 & 1 & 1 & 1 & 1 & 1 & 2 & 1 & 1 & 1 & 1 & 1 & 2 & 1 \\
1 & 1 & 2 & 1 & 4 & 2 & 2 & 2 & 1 & 2 & 2 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 2 & 4 & 2 & 2 & 1 & 1 & 1 & 1 & 2 & 1 & 1 & 1 \\
1 & 2 & 1 & 1 & 1 & 1 & 1 & 1 & 2 & 1 & 1 & 1 & 2 & 2 & 4 & 2 \\
2 & 1 & 1 & 1 & 1 & 1 & 2 & 1 & 1 & 1 & 1 & 1 & 1 & 2 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 2 & 1 & 1 & 1 & 1 & 2 & 1 & 4 & 2 & 2 & 2 \\
2 & 4 & 2 & 2 & 1 & 1 & 1 & 1 & 2 & 1 & 1 & 1 & 1 & 1 & 2 & 1 \\
1 & 1 & 1 & 1 & 2 & 2 & 4 & 2 & 1 & 1 & 1 & 1 & 1 & 2 & 1 & 1 \\
1 & 1 & 2 & 1 & 2 & 1 & 1 & 1 & 1 & 2 & 1 & 1 & 1 & 1 & 1 & 1
\end{pmatrix}$$

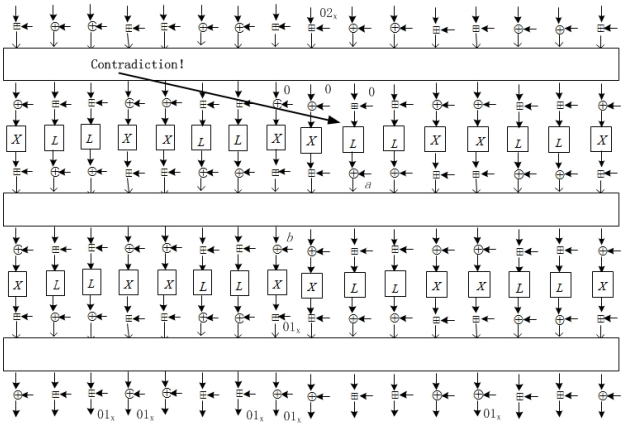
C. Round functions of SAFER+(Up) and SAFER++(Down)



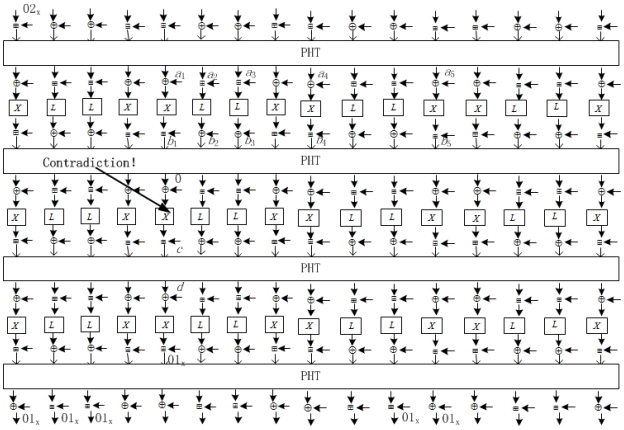


Appendix B

A. A 2.75-round distinguisher of SAFER+

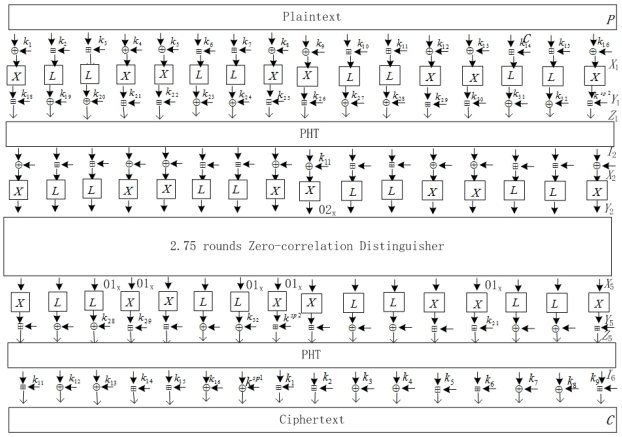


B. A 3.75-round distinguisher of SAFER++

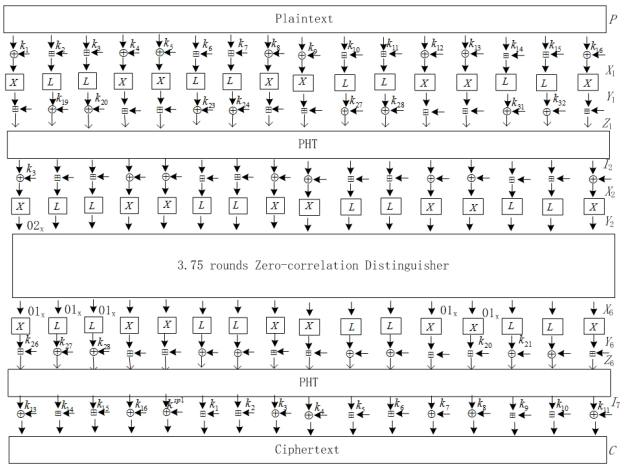


APPENDIX C

A. Key recovery attack on 5-round SAFER+



B. Key recovery attack on 6-round SAFER++



REFERENCES

[1] F. Webster and E. Tavares. On the design of S-boxes. In: Advances in Cryptology-CRYPTO'85, LNCS 219, Berlin: Springer-Verlag, 1986: 523-534.

[2] E. Biham and A. Shamir. Differential cryptanalysis of DES-like cryptosystems. Journal of Cryptology, 1991, 4(1): 3-72.

[3] C. Gupta and P. Sarkar. Construction of Perfect Nonlinear and maximally nonlinear multiple-output Boolean functions satisfying higher order strict avalanche criteria. IEEE Transactions on Information Theory, 2004, 50(11): 2886-2893.

[4] Q. Yang, W. Wong, F Liao, et al. One-way hash function construction based on chaotic map network. Chaos, Solitons and Fractals, 2009, 41(5): 2566-2574.

[5] J.L. Massey. SAFER K-64: A byte-oriented block-ciphering algorithm. Fast Software Encryption, 1993, 809: 1-17.

[6] J.L. Massey. Strengthened key schedule for the cipher SAFER. ftp://ftp.cert.dfn.de/ pub/tools/crypt/SAFER/,1995.

[7] J.L. Massey, G.H. Khachatrian, and M.K. Kuregian. 1st AES conference on nomination of SAFER+ as candidate algorithm for the advanced encryption standard. http://csrc.nist.gov/encryption/aes/, 1998.

[8] J.L. Massey, G.H. Khachatrian, and M.K. Kuregian. 1st NESSIE workshop on the SAFER++ Block encryption algorithm. http://cryptonessie.org, 2000

[9] J. Nakahara, B. Preneel, and J. Vandewalle. Linear cryptanalysis of reduced-round versions of the SAFER block cipher family. Fast Software Encryption, 2000: 244-261

[10] J. Nakahara. Cryptanalysis and Design of Block Ciphers. PhD thesis, Katholieke Universiteit, Leuven, 2003

[11] G. Piret and J. Quisauaterm. Integral cryptanalysis on reduced-round SAFER++: A way to extend the attack, NESSIE Public Report. NES/DOC/UCL/WP5/002/1, 2003

[12] Y. Yemo and I. Park. Optimization of integral cryptanalysis on reduced-round SAFER++. Joho Shori Gakkai Shinpojiumu Ronbunshu, 2003, 2003(15), 223-227.

[13] A. Biryukov, C.D. Canniere, and G. Dellkrantz. Cryptanalysis of SAFER++. Crypto'03, 2003, 195-211.

[14] J. Nakahara and B. Preneel. Impossible differential attacks on reduced-round SAFER ciphers, NESSIE Public Report. NES/DOC/KUL/WP5/30/1, 2003.

[15] B. Behnam, E. Taraneh, and R.A. Mohammad. Impossible differential cryptanalysis of SAFER++. Security and Management, 2008: 10-14.

[16] S. Zheng, C. Wang, and Y. Yang. A new impossible differential attack on SAFER ciphers. Computers and Electrical Engineering, 2010, 36(1): 180-189.

[17] J. Zhao, M. Wang, J. Chen, and Y. Zheng. New Impossible Differential Attack on SAFER Block Cipher Family. IEICE TRANS. FUNDAMENTALS, 2015, 98(3): 843-852.

[18] F. Chabaud, S. Vaudenay. Links Between Differential and Linear Cryptoanalysis. EUROCRYPT'94, LNCS 950, Springer, 1994: 356-365.

[19] C.S. Williams. Proposal for A 'Tweak' to Cylink's AES Candidate Algorithm SAFER+. AES, http://csrc.nist.gov/aes/, 1999.

[20] A. Bogdanov and V.Rijmen. Linear Hulls with Correlation Zero and Linear Cryptanalysis of Block Ciphers. Des. Codes Cryptography, 2014, 70(3): 369-383.

[21] A. Bogdanov and M. Wang. Zero Correlation Linear Cryptanalysis with Reduced Data Complexity. Fast Software Encryption, 2012, 7549: 29-48.