Integral Attacks on Reduced-Round PRESENT

Shengbao Wu[1,2] and Mingsheng Wang[3]

[1] Trusted Computing and Information Assurance Laboratory, Institute of Software, Chinese Academy of Sciences, Beijing 100190, PO Box 8718, China
[2] Graduate School of Chinese Academy of Sciences, Beijing 100190, China
[3] State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China
wangmingsheng@iie.ac.cn

Abstract. Integral attack is a powerful technique to recover the secret key of block ciphers by usually exploiting the fact that specific parts of the output after several round encryptions has a zero-sum property in a set of chosen plaintexts. In FSE 2008, bit-based integral attack proposed by Z'aba et al. revealed that integral attacks may be not only suitable for byte-based block ciphers but also still applied to bit-based block ciphers. In this work, we show that integral attack against bit-based block ciphers can be improved not only by the theorem of higher-order differential attack but also by using specific algebraic properties of Sboxes, and the order of plaintexts in a set, which is important in bit-based integral attack, is not required here. We focus on the block cipher PRESENT. Based on some algebraic properties of its Sbox, we propose two integral distinguishers: a 5 round (4-th order) integral distinguisher and a 7 round (16-th order) integral distinguishers, which can be used to attack 10 (out of 31) round PRESENT. As far as we know, it is the first time that a 7 round integral distinguisher of PRESENT is reported. Algebraic techniques used in this paper may be also applied to other block ciphers to improve their known integral attacks.

Keywords: Integral Attack, PRESENT, Higher Order Differential Attack, Boolean Function.

1 Introduction

The integral attack is one of the most popular cryptanalytic tools for block ciphers. It was first known as "Square attack" due to its efficiency in attacking the Rijndael-predecessor Square [8]. Later, several variants of Square attack have been proposed, including saturation attack [13] and multiset attack [5]. In FSE 2002, Knudsen and Wagner introduced the definition of integral and unified these kinds of attacks as integral attack [11].

The basic idea of integral attack is to analyze the propagation of sums of (many) values. Thus, it can be seen as a dual to the differential cryptanalysis. When applying integral attack to a block cipher, an attacker first selects a d-th order integral, that is, he/she chooses a set of 2^d plaintexts, where d bit positions

S. Qing et al. (Eds.): ICICS 2013, LNCS 8233, pp. 331–345, 2013.

take on all values through the set, and the other bits are chosen to be arbitrary constants. Then, he/she traces the evolvement of the sum of this set of plaintexts through the encryption algorithm and builds an integral distinguisher as long as possible. Finally, the integral distinguisher will be used to verify key guesses. In practice, a zero-sum property in specific parts of the ciphertext is often used as the integral distinguisher.

In quite a long time, integral attack has not been thought suitable for bit-based block ciphers, such as Noekeon [9], Serpent [1] and PRESENT [2]. Until 2008, Z'aba et al. proposed the bit-based integral attack [16], which was applied to Noekeon, Serpent and PRESENT reduced up to 5, 6 and 7 rounds, respectively. Although the bit-based integral attack does not pose a serious threat to known block ciphers, it reveals that integral attacks may be not only suitable for byte-based block ciphers but also still applied to bit-based block ciphers.

Many cryptanalysis methods may be not so powerful as nowadays in their first proposals. However, with the studies getting further, they became more and more powerful. On the one hand, new techniques may be introduced to improve known cryptanalysis methods. For example, the partial-sum techniques proposed by Ferguson et al. [10] enhance the ability of integral attack. On the other hand, a cryptanalysis method may be improved using the theorem of other cryptanalysis methods if they have some links. For example, the data complexity of a zero-correlation attack [4] may be improved using the theorem of integral attack [3].

Integral attack and higher-order differential attack also have some links in constructing distinguishers. To construct a d-th order integral distinguisher with a zero-sum property is equivalent to show that the algebraic degree of specific parts of the ciphertext is at most $d - 1$, if XOR difference is considered in the higher-order differential attack. This technique has been used, for instance, in [15].

In this paper, we show that integral attack against bit-based block ciphers can be improved not only by the theorem of higher-order differential attack but also by using specific algebraic properties of Sboxes. What is more, the order of plaintexts in a set, which is important in bit-based integral attack, is not required here.

We focus on the bit-based block cipher PRESENT. Firstly, we analyze the algebraic properties of PRESENT's Sbox. We observe that the rightmost coordinate of the Sbox is quadratic while other three coordinates has algebraic degree 3. Combined it with the properties of diffusion layer, we find that, for the rightmost bit of the output, the growth rate of its algebraic degree is slower than other bits. Than, we propose two integral distinguishers: the first one uses that the rightmost bit of the output after 5 rounds has a zero-sum property in the 4 rightmost bits of the input. Similarly, the second distinguisher is based on the fact that the rightmost bit of the output after 7 rounds has a zero-sum property in the 16 rightmost bits of the input. Our distinguishers improve the 3.5 round (4-th order) integral distinguisher proposed by Z'aba et al. and the 5 round (32-th order) integral distinguisher proposed by Zhang et al [17].

Finally, we applied our distinguishers to recover the keys up to 10 (out of 31) rounds of PRESENT. All known integral attacks on reduced-round PRESENT are summarized in Table 1.

Table 1. Summary of integral attacks on reduced-round PRESENT

Rounds	Key Size	Data	Time	Memory	Attacks & Source
5	all	$N \cdot 2^{32}$ CP†	-	-	(32-th order) integral distinguisher [17]
5	80	$2^{6.4}$ CP	$2^{25.7}$	-	Bit-Pattern Based Integral [16]
6	80	$2^{22.4}$ CP	$2^{41.7}$	-	Bit-Pattern Based Integral [16]
7	128	$2^{24.3}$ CP	$2^{100.1}$	2^{77}	Bit-Pattern Based Integral [16]
7	80	$2^{8.3}$ CP	2^{60}	2^{17}	Section 5
8	80	$2^{10.1}$ CP	$2^{72.6}$	2^{66}	Section 5
9	80	$2^{20.3}$ CP	2^{60}	2^{17}	Section 5
10	128	$2^{22.4}$ CP	$2^{99.3}$	2^{81}	Section 5

† N is the number of sets required in a key-recover attack.

Even though we only restrict our attention on PRESENT in this work, the algebraic techniques used in constructing longer integral distinguishers here may be also applied to other block ciphers to improve their known integral attacks.

Outline of This Paper. In Section 2, we introduce the encryption process of PRESENT, the definition of boolean functions and the basic idea of integral attack. In Section 3, we analyze the properties of PRESENT's S-box and present some observations on the degree of boolean functions. The integral distinguishers are constructed in Section 4 and attacks based on them are given in Section 5. Finally, we conclude this paper.

2 Preliminaries

In this section, we briefly describe the block cipher PRESENT, boolean functions and the integral attack.

2.1 Description of PRESENT

PRESENT [2], proposed by A. Bogdanov et.al in CHES 2007, is a 31-round ultra-lightweight block cipher with block length 64 bits. It has two versions supporting key length 80 bits and 128 bits, which will be denoted as PRESENT-80 and PRESENT-128, respectively. The underlying structure of PRESENT is a typical SP-network which has three layers in every round: AddRoundKey, SBoxLayer and PLayer. In the AddRoundKey layer, a round key with 64 bits is XORed to the current state. Then, one 4-bit Sbox is applied 16 times in parallel in the SBoxLayer. Finally, a fully wired permutation P on the 64-bit state is employed in the PLayer. The outline of one round PRESENT is shown in Fig. 1. Notice that there is an AddRoundKey layer after round 31. The Sbox

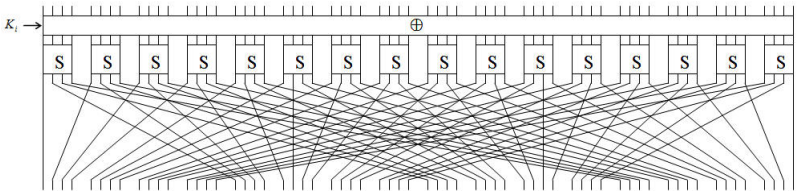

Fig. 1. One round PRESENT

Table 2. The Sbox of PRESENT

x	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
$S(x)$	C	5	6	B	9	0	A	D	3	E	F	8	4	7	1	2

and P permutation used in PRESENT are illustrated in Table 2 and Table 3, respectively.

The key schedule of PRESENT-80 is given below. Firstly, the 80-bit key is stored in a key register K and represented as $k_{79}k_{78}\ldots k_0$. In round i, the most significant 64-bit keys are extracted as the subkey $K^{(i)}$, that is, $K^{(i)} = k_{79}k_{78}\ldots k_{16}$. Then, key register $K = k_{79}k_{78}\ldots k_0$ is updated as follows:

$$\begin{aligned}[k_{79}k_{78}\ldots k_1k_0] &= [k_{18}k_{17}\ldots k_{20}k_{19}],\\ [k_{79}k_{78}k_{77}k_{76}] &= S[k_{79}k_{78}k_{77}k_{76}],\\ [k_{19}k_{18}k_{17}k_{16}k_{15}] &= [k_{19}k_{18}k_{17}k_{16}k_{15}] \oplus round\_counter.\end{aligned}$$

We omit the key schedule of PRESENT-128 here since we do not use it in this paper.

Table 3. The PLayer of PRESENT. Bit i of state is moved to bit position $P(i)$.

i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$P(i)$	0	16	32	48	1	17	33	49	2	18	34	50	3	19	35	51
i	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
$P(i)$	4	20	36	52	5	21	37	53	6	22	38	54	7	23	39	55
i	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47
$P(i)$	8	24	40	56	9	25	41	57	10	26	42	58	11	27	43	59
i	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63
$P(i)$	12	28	44	60	13	29	45	61	14	30	46	62	15	31	47	63

2.2 Boolean Functions

A *boolean function f of n variables* is a map from $\mathbb{F}_2^n \to \mathbb{F}_2$. It can be expressed as a polynomial in $\mathbb{F}_2[x_1, \ldots, x_n]/(x_1^2 - x_1, \ldots, x_n^2 - x_n)$, called *algebraic normal form*. That is,

$$f(x_1, x_2, \ldots, x_n) = \bigoplus_{\Lambda \subseteq \{1,2,\ldots,n\}} a_\Lambda \prod_{k \in \Lambda} x_k. \quad (1)$$

In the subsequent discussions, we denote by $\mathcal{B}_2[x_1, x_2, \ldots, x_n]$ the set of all boolean functions with variables $x_1, x_2, \ldots, x_n$. The *algebraic degree* (or *degree*) of f, denoted by $deg(f)$, is the number of variables in the highest order term with nonzero coefficient. For a further step, the *degree* of a vectorial boolean function from $\mathbb{F}_2^n \to \mathbb{F}_2^m$ is defined as the highest degree of its coordinates.

2.3 Integral Attack

Let $E = E_1 \circ E_0$ be the encryption function of an r round block cipher, where E_0 is the first k rounds of E and E_1 is the last $r - k$ rounds. It can be written formally as

$$Y = E(X, K) = E_1(E_0(X, K_0), K_1),$$

or equivalently,

$$E_1^{-1}(Y, K_1) = E_0(X, K_0), \quad (2)$$

where E_1^{-1} is the inverse function of E_1, K is the master key, K_0 and K_1 are subkeys in the first k rounds and the last $r - k$ rounds, respectively.

In integral attacks, an attacker first selects a set of 2^d plaintext, where d bit positions of X take on all values through the set and the other bits of X are chosen to be arbitrary constants. Then, a zero-sum property of the set of plaintexts propagating through k round encryptions is proved, that is, an attacker demonstrates that

$$\bigoplus_{X \in \Lambda} E_0(X, K_0) = 0, \quad (3)$$

where Λ is the set of 2^d plaintexts. Finally, the subkey K_1 in the last $r-k$ rounds is guessed and equation

$$\bigoplus_{X \in \Lambda, Y = E(X,K)} E_1^{-1}(Y, K_1) = 0 \quad (4)$$

is used to verify the guess. The remaining key bits in the master key K will be obtained by exhausting method.

Notice that, the integral distinguisher (3) can be built upon a specific parts of the output of E_0, that is, the zero-sum property may be only valid in some specific bits. Based on the theorem of higher-order differential attack, (3) can be proved by showing that some specific bits of the output of E_0 have degree at most $d - 1$.

3 Degree of Boolean Functions and PRESENT's Sbox

In this section, we discuss some properties for evaluating the degree of boolean functions and then analyze the algebraic properties of PRESENT's Sbox.

Some trivial bounds for operations between (vectorial) boolean functions are summarized in Proposition 1.

Proposition 1. *Suppose $f, g \in \mathcal{B}_2[x_1, x_2, \ldots, x_n]$ are two boolean functions, h is a vectorial boolean function from $\mathbb{F}_2^n \to \mathbb{F}_2^n$, then the degree of composed function $f \circ h$, product $f \cdot g$ and sum $f \oplus g$ can be evaluated as*

$$\begin{aligned} deg(f \circ h) &\leq deg(f)deg(h), \\ deg(f \cdot g) &\leq deg(f) + deg(g), \\ deg(f \oplus g) &\leq max\{deg(f), deg(g)\}. \end{aligned}$$

Moreover, $deg(f), deg(g), deg(f \circ h), deg(f \cdot g)$ and $deg(f \oplus g)$ are less than or equal to n.

These bounds are so loose that they are unfitted in some cases. Here, we analyze the product of boolean functions and show a tighter degree bound in a specific situation. First, we introduce the definition of *m-partition.*

Definition 1. *Nonempty sets $U_1, \ldots, U_m$ is called an m-partition of $U = \{x_1, x_2, \ldots, x_n\}$, if $U = U_1 \cup \cdots \cup U_m$ and $U_i \cap U_j = \emptyset$ for $1 \leq i < j \leq m$.*

Let n_i be the number of variables in U_i, then $n = n_1 + \cdots + n_m$. Our observation is given below.

Proposition 2. *Suppose $U_1, \ldots, U_m$ is an m-partition of $U = \{x_1, x_2, \ldots, x_n\}$, $f_1, f_2, \ldots, f_k$ is a list of boolean functions satisfying:*

1. *For each f_i, there exists a $j \in \{1, 2, \ldots, m\}$ such that $f_i \in \mathcal{B}_2[U_j]$,*
2. *$deg(f_i) \leq n_j - 1$,*

then, for any $k \leq 2m - 1$, we have

$$deg(f_1 \cdot f_2 \cdots f_k) \leq n - 1.$$

Proof. This can be explained as an allocation problem. Now, we have k tokens $f_1, \ldots, f_k$ and m boxes $\mathcal{B}_2[U_1], \ldots, \mathcal{B}_2[U_m]$. When throwing $k \leq 2m - 1$ tokens to m boxes, there must exist a box with the condition that it contains no more than one token. Without loss of generality, suppose this box is $\mathcal{B}_2[U_1]$.

If it's empty, then all f_is do not involve variables in U_1, which implies

$$deg(f_1 \cdot f_2 \cdots f_k) \leq n - n_1 \leq n - 1.$$

If it contains a token f_i, then, from Proposition 1, we have

$$\begin{aligned} deg(f_1 \cdot f_2 \cdots f_k) &\leq deg(f_i) + deg(f_1 \cdots f_{i-1} \cdot f_{i+1} \cdots f_k) \\ &\leq (n_i - 1) + (n - n_i) \leq n - 1. \end{aligned}$$

□

This property will be used for constructing integral distinguishers of PRESENT, combining with the subsequent observations on PRESENT's Sbox.

Proposition 3. *The Sbox of PRESENT is a permutation* $S : \mathbb{F}_2^4 \to \mathbb{F}_2^4$. *It can be expressed as a vectorial boolean function with four coordinates. Suppose its input is a vector* $x = (x_3, x_2, x_1, x_0)$ *and output is a vector* $y = (y_3, y_2, y_1, y_0)$, *where* $x_i, y_i \in \mathbb{F}_2$ *and* $0 \leq i \leq 3$. *Then, the algebraic normal form of PRESENT's Sbox is:*

$$\begin{cases} y_3 = 1 \oplus x_0 \oplus x_1 \oplus x_3 \oplus x_1x_2 \oplus x_0x_1x_2 \oplus x_0x_1x_3 \oplus x_0x_2x_3; \\ y_2 = 1 \oplus x_2 \oplus x_3 \oplus x_0x_1 \oplus x_0x_3 \oplus x_1x_3 \oplus x_0x_1x_3 \oplus x_0x_2x_3; \\ y_1 = x_1 \oplus x_3 \oplus x_1x_3 \oplus x_2x_3 \oplus x_0x_1x_2 \oplus x_0x_1x_3 \oplus x_0x_2x_3; \\ y_0 = x_0 \oplus x_2 \oplus x_3 \oplus x_1x_2; \end{cases}$$

The correctness of Proposition 3 can be easily checked. From Proposition 3, we immediately have

Corollary 1. *The degree of PRESENT's Sbox S is 3. However, its rightmost coordinate is only quadratic, that is,* $deg(y_0) = 2$.

Corollary 2. *Let* $f = c_f \oplus f_0 \oplus f_2 \oplus f_3 \oplus f_1f_2$ *and* $g = c_g \oplus g_0 \oplus g_2 \oplus g_3 \oplus g_1g_2$, *where* $f_i, g_i \in \mathcal{B}_2[x_i]$ *for* $0 \leq i \leq 3$ *and* $c_f, c_g \in \mathbb{F}_2$ *are constants, then we have* $deg(f \cdot g) \leq 3$.

According to the PLayer of PRESENT (see Fig. 1), all 16 quadratic coordinates of Sboxes are translated to the rightmost 16 bits after one round encryption. Thus, we have

Observation. *The growth rate of algebraic degree for the bits in the right side of the output is slower than those in the left side.*

After several rounds of encryption, the effect is finally accumulated to the rightmost bit of the output. Therefore, in the subsequent discussions, we only consider the degree for the rightmost bit of the output.

4 Integral Distinguishers of PRESENT

In this section, we proposed two integral distinguishers of PRESENT. We denote by $X^{(i)}$ the state entering round i, $Y^{(i)}$ the state before the SBoxLayer, $Z^{(i)}$ the state after the SBoxLayer and $K^{(i)}$ the subkey of round i. Thus, $Y^{(i)} = X^{(i)} \oplus K^{(i)}$. Each state and subkey can be represented as a vector of 64 bits, for example, $X^{(i)} = (x_{63}^{(i)}, x_{62}^{(i)}, \ldots, x_0^{(i)})$, where $x_0^{(i)}$ is the least significant (rightmost) bit of $X^{(i)}$. Additionally, let $x_{[j-k]}^{(i)}$ be the consecutive $j-k+1$ bits of $X^{(i)}$ from bit k to bit j, and $x_{[j,\ldots,k]}^{(i)}$ represents several separate bits $x_j^{(i)}, \ldots, x_k^{(i)}$ of $X^{(i)}$.

4.1 A 5 Round (4-th Order) Integral Distinguisher

In this subsection, we show a 4-th order integral distinguisher of PRESENT, which provides us a 5-round integral distinguisher.

Proposition 4. *Choose a set of 2^4 values in the input of round 2, where all values of bits $x^{(2)}_{[48,32,16,0]}$ of input $X^{(2)}$ are chosen and other bits are chosen to be arbitrary constants. Then, the rightmost bit of $X^{(6)}$, that is, the bit $x^{(6)}_0$, has a zero-sum property.*

Proof. Consider $x^{(6)}_0$ as a boolean function of $X^{(2)}$, then we only need to prove that $x^{(6)}_0 \in \mathcal{B}_2[x^{(2)}_{48}, x^{(2)}_{32}, x^{(2)}_{16}, x^{(2)}_0]$ has degree at most 3. The proof process is shown in the phase T1 of Fig. 2.

In round 2, since $x^{(2)}_{[3-1]}$ are fixed, then $z^{(2)}_{[3-0]}$ are affine functions with only one variable $x^{(2)}_0$, that is, $z^{(2)}_{[3-0]} \in \mathcal{B}_2[x^{(2)}_0]$. Similarly, we have $z^{(2)}_{[19-16]} \in \mathcal{B}_2[x^{(2)}_{16}]$, $z^{(2)}_{[35-32]} \in \mathcal{B}_2[x^{(2)}_{32}]$ and $z^{(2)}_{[51-48]} \in \mathcal{B}_2[x^{(2)}_{48}]$. Other bits of $Z^{(2)}$ are constants.

In round 3, we have $x^{(3)}_{[48,32,16,0]} \in \mathcal{B}_2[x^{(2)}_0]$, $x^{(3)}_{[52,36,20,4]} \in \mathcal{B}_2[x^{(2)}_{16}]$, $x^{(3)}_{[56,40,24,8]} \in \mathcal{B}_2[x^{(2)}_{32}]$ and $x^{(3)}_{[60,44,28,12]} \in \mathcal{B}_2[x^{(2)}_{48}]$.

In round 4, we have $x^{(4)}_{[12,8,4,0]} \in \mathcal{B}_2[x^{(2)}_0]$, $x^{(4)}_{[13,9,5,1]} \in \mathcal{B}_2[x^{(2)}_{16}]$, $x^{(4)}_{[14,10,6,2]} \in \mathcal{B}_2[x^{(2)}_{32}]$ and $x^{(4)}_{[15,11,7,3]} \in \mathcal{B}_2[x^{(2)}_{48}]$.

In summary, bits marked with red color (resp. green color, blue color and purple color) in Fig. 2 are affine functions with only one variable $x^{(2)}_0$ (resp. $x^{(2)}_{16}$, $x^{(2)}_{32}$ and $x^{(2)}_{48}$). Other bits are not considered here since they are independent of $x^{(6)}_0$.

In round 5, from the expression of Sbox, we have

$$y^{(5)}_i = k^{(5)}_i \oplus y^{(4)}_{4i} \oplus y^{(4)}_{4i+2} \oplus y^{(4)}_{4i+3} \oplus y^{(4)}_{4i+1} \cdot y^{(4)}_{4i+2},$$

where $y^{(4)}_{4i+j} \in \mathcal{B}_2[x^{(2)}_{16j}]$ for $0 \le j \le 3$ and $0 \le i \le 3$.

Finally,

$$\begin{aligned} deg(x^{(6)}_0) &= deg(y^{(5)}_0 \oplus y^{(5)}_2 \oplus y^{(5)}_3 \oplus y^{(5)}_1 \cdot y^{(5)}_2) \\ &\le max\{deg(y^{(5)}_0), deg(y^{(5)}_2), deg(y^{(5)}_3), deg(y^{(5)}_1 \cdot y^{(5)}_2)\} \\ &\le max\{2, 2, 2, 3\} = 3, \end{aligned}$$

where the final inequation comes from Corollary 2. □

A 5-round integral distinguisher is obtained by adding one round to the upper side of the distinguisher given in Proposition 4.

Theorem 1. *Choose a set of 2^4 values in the plaintext, where all values of bits $x^{(1)}_{[3,2,1,0]}$ of input $X^{(1)}$ are chosen and other bits are chosen to be arbitrary constants. Then, the rightmost bit of $X^{(6)}$, that is, the bit $x^{(6)}_0$, has a zero-sum property.*

Proof. It's based on the fact that $x^{(1)}_{[3-0]} \to x^{(2)}_{[48,32,16,0]}$ is a permutation (see the phase T2 of Fig. 2 with bold line). We have

$$\bigoplus_{x^{(2)}_{[48,32,16,0]} \in \mathbb{F}_2^4} R_{K^{(5)}} \circ \cdots \circ R_{K^{(2)}}(x^{(2)}_{[48,32,16,0]}, c') = \bigoplus_{x^{(1)}_{[3-0]} \in \mathbb{F}_2^4} R_{K^{(5)}} \circ \cdots \circ R_{K^{(1)}}(x^{(1)}_{[3-0]}, c), \tag{5}$$

where $R_{K^{(i)}}$ is the round function with key $K^{(i)}$, c is the constant chosen in the plaintext and c' is the constant deduced from c by one round encryption. □

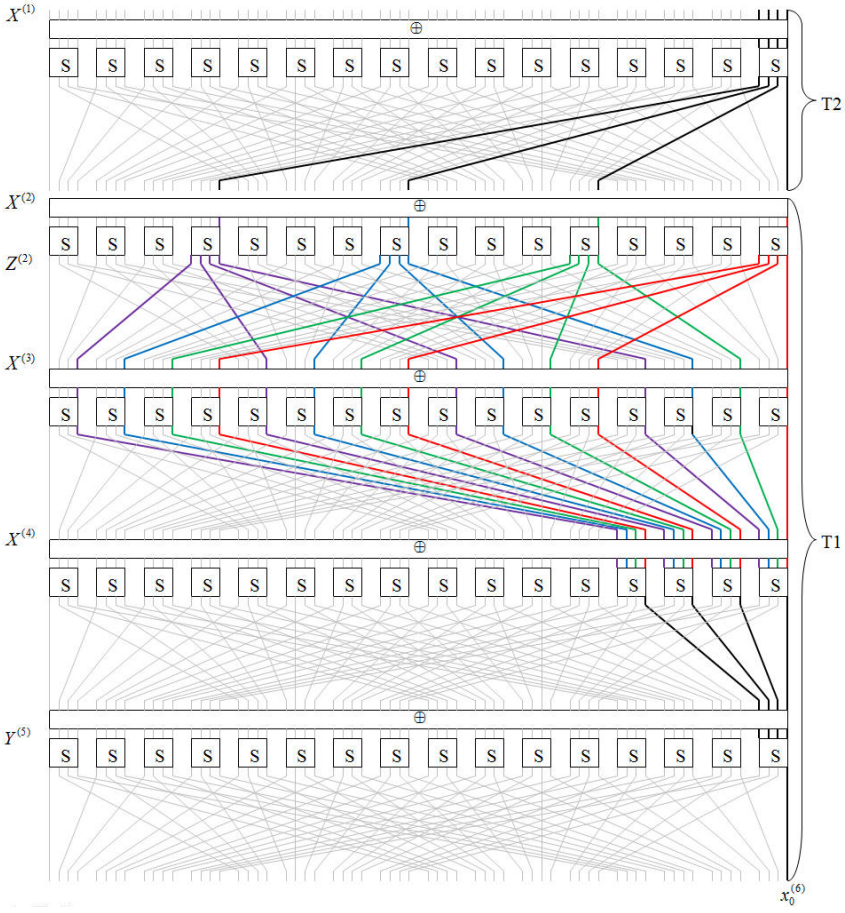

Fig. 2. 4-th order integral distinguisher of PRESENT

4.2 A 7 Round (16-th Order) Integral Distinguisher

In this subsection, we show a 16-th order integral distinguisher of PRESENT, which provides us a 7-round integral distinguisher.

Proposition 5. *Choose a set of 2^{16} values in the input of round 3, where all values of bits $x_{4j}^{(3)}$ $(0 \le j \le 15)$ of input $X^{(3)}$ are chosen and other bits are chosen to be arbitrary constants. Then, the rightmost bit of $X^{(8)}$, that is, the bit $x_0^{(8)}$, has a zero-sum property.*

Proof. Consider $x_0^{(8)}$ as a boolean function of $X^{(3)}$, then we only need to prove that $x_0^{(8)} \in \mathcal{B}_2[x_{60}^{(3)}, x_{56}^{(3)}, \ldots, x_4^{(3)}, x_0^{(3)}]$ has degree at most 15. The proof process is shown in the phase T1 of Fig. 3.

In round 3, we have $z_{[4j+3,4j+2,4j+1,4j]}^{(3)} \in \mathcal{B}_2[x_{4j}^{(3)}]$ $(0 \le j \le 15)$.

In round 4, we have $y_i^{(4)} \in \mathcal{B}_2[x_{4j}^{(3)}]$ $(0 \le i \le 63)$, where $j = i \bmod 16$.

In round 5, from the expression of PRESENT's Sbox, we have $y_i^{(5)} \in \mathcal{B}_2[x_{16j+12}^{(3)}, x_{16j+8}^{(3)}, x_{16j+4}^{(3)}, x_{16j}^{(3)}]$ and $deg(y_i^{(5)}) \le 3$, where $0 \le i \le 63$ and $j = i \bmod 4$.

In summary, bits marked with red color (resp. green color, blue color and purple color) in Fig. 3 are boolean functions in $\mathcal{B}_2[x_{12}^{(3)}, x_8^{(3)}, x_4^{(3)}, x_0^{(3)}]$ (resp. $\mathcal{B}_2[x_{28}^{(3)}, x_{24}^{(3)}, x_{20}^{(3)}, x_{16}^{(3)}]$, $\mathcal{B}_2[x_{44}^{(3)}, x_{40}^{(3)}, x_{36}^{(3)}, x_{32}^{(3)}]$ and $\mathcal{B}_2[x_{60}^{(3)}, x_{56}^{(3)}, x_{52}^{(3)}, x_{48}^{(3)}]$) and have degree at most 3.

Now, we consider $x_0^{(8)}$ as a boolean function of $Y^{(5)}$. Then, $x_0^{(8)} \in \mathcal{B}_2[y_{[63-0]}^{(5)}]$ has representation

$$x_0^{(8)} = \bigoplus_{\Lambda \subseteq \{0,1,2,\ldots,63\}} a_\Lambda \prod_{k \in \Lambda} y_k^{(5)}. \tag{6}$$

Notice that $x_0^{(8)}$ is also a boolean function in $\mathcal{B}_2[x_{60}^{(3)}, \ldots, x_4^{(3)}, x_0^{(3)}]$. Thus, in the following discussions, we have to show that each term $a_\Lambda \prod_{k \in \Lambda} y_k^{(5)} \in \mathcal{B}_2[x_{60}^{(3)}, \ldots, x_4^{(3)}, x_0^{(3)}]$ with $a_\Lambda \neq 0$ has degree at most 15.

First, we show that $deg(\prod_{k \in \Lambda} y_k^{(5)}) \le 15$ if $|\Lambda| \le 7$. Here, $|\Lambda|$ is the number of elements in set Λ. Denote by $U = \{x_{4j}^{(3)} | 0 \le j \le 15\}$ and $U_k = \{x_{16k+12}^{(3)}, x_{16k+8}^{(3)}, x_{16k+4}^{(3)}, x_{16k}^{(3)}\}$ for $0 \le k \le 3$, then $U_0, \ldots, U_3$ is a 4-partition of U. Notice that $y_i^{(5)}$ $(0 \le i \le 63)$ satisfies the condition of Proposition 2, which implies that $deg(\prod_{k \in \Lambda} y_k^{(5)}) \le 15$ if $|\Lambda| \le 7$. Therefore, we only need to check the terms $a_\Lambda \prod_{k \in \Lambda} y_k^{(5)}$ with $a_\Lambda \neq 0$ and $|\Lambda| \ge 8$.

Secondly, we show that a_Λ is always zero in (6) if $|\Lambda| > 8$. According to Proposition 3 and Proposition 1, we have $deg(y_{[15-0]}^{(6)}) \le 2$, $deg(y_{[3-0]}^{(7)}) \le 4$ and $deg(x_0^{(8)}) \le 8$ if $y_{[15-0]}^{(6)}, y_{[3-0]}^{(7)}$ and $x_0^{(8)}$ are viewed as boolean functions in $\mathcal{B}_2[y_{[63-0]}^{(5)}]$. Thus, in (6), all $a_\Lambda = 0$ if $|\Lambda| \ge 9$, which implies that we only need to consider the terms with $|\Lambda| = 8$.

Thirdly, we show that only one term in (6) may have $|\Lambda| = 8$. According to the algebraic normal form of PRESENT's Sbox, $x_0^{(8)} \in \mathcal{B}_2[y_{[63-0]}^{(5)}]$ can be expressed as follows.

$$\begin{aligned} x_0^{(8)} &= y_0^{(7)} \oplus y_2^{(7)} \oplus y_3^{(7)} \oplus y_1^{(7)} y_2^{(7)} \backsimeq y_1^{(7)} y_2^{(7)} \\ &= (k_1^{(7)} \oplus y_4^{(6)} \oplus y_6^{(6)} \oplus y_7^{(6)} \oplus y_5^{(6)} y_6^{(6)})(k_2^{(7)} \oplus y_8^{(6)} \oplus y_{10}^{(6)} \oplus y_{11}^{(6)} \oplus y_9^{(6)} y_{10}^{(6)}) \\ &\backsimeq y_5^{(6)} y_6^{(6)} y_9^{(6)} y_{10}^{(6)} \backsimeq y_{21}^{(5)} y_{22}^{(5)} y_{25}^{(5)} y_{26}^{(5)} y_{37}^{(5)} y_{38}^{(5)} y_{41}^{(5)} y_{42}^{(5)}, \end{aligned}$$

where $\backsimeq$ means that the terms with degree 8 can only appear in these products. Thus, the remaining work is to prove that $y_{21}^{(5)} y_{22}^{(5)} y_{25}^{(5)} y_{26}^{(5)} y_{37}^{(5)} y_{38}^{(5)} y_{41}^{(5)} y_{42}^{(5)}$ is a boolean function with degree at most 15 in $\mathcal{B}_2[x_{60}^{(3)}, x_{56}^{(3)}, \ldots, x_4^{(3)}, x_0^{(3)}]$.

Finally, we show that $y_{21}^{(5)} y_{22}^{(5)} y_{25}^{(5)} y_{26}^{(5)} y_{37}^{(5)} y_{38}^{(5)} y_{41}^{(5)} y_{42}^{(5)} \in \mathcal{B}_2[x_{60}^{(3)}, \ldots, x_4^{(3)}, x_0^{(3)}]$ has degree less than 15. Since $y_{[41,37,25,21]}^{(5)} \in \mathcal{B}_2[x_{28}^{(3)}, x_{24}^{(3)}, x_{20}^{(3)}, x_{16}^{(3)}]$ and $y_{[42,38,26,22]}^{(5)} \in \mathcal{B}_2[x_{44}^{(3)}, x_{40}^{(3)}, x_{36}^{(3)}, x_{32}^{(3)}]$, we have

$$deg(y_{21}^{(5)} y_{22}^{(5)} y_{25}^{(5)} y_{26}^{(5)} y_{37}^{(5)} y_{38}^{(5)} y_{41}^{(5)} y_{42}^{(5)}) \leq 8.$$

In summary, $x_0^{(8)} \in \mathcal{B}_2[x_{60}^{(3)}, x_{56}^{(3)}, \ldots, x_4^{(3)}, x_0^{(3)}]$ has degree

$$deg(x_0^{(8)}) = max\{deg(\prod_{k \in \Lambda, |\Lambda| \leq 7} y_k^{(5)}), deg(y_{21}^{(5)} y_{22}^{(5)} y_{25}^{(5)} y_{26}^{(5)} y_{37}^{(5)} y_{38}^{(5)} y_{41}^{(5)} y_{42}^{(5)})\} \leq 15.$$

□

A 7-round integral distinguisher is obtained by adding two rounds to the upper side of the distinguisher given in Proposition 5.

Theorem 2. *Choose a set of 2^{16} values in the plaintext, where all values of bits $x_{[15-0]}^{(1)}$ of input $X^{(1)}$ are chosen and other bits are chosen to be arbitrary constants. Then, the rightmost bit of $X^{(8)}$, that is, the bit $x_0^{(8)}$, has a zero-sum property.*

Proof. It's based on the fact that $x_{[15-0]}^{(1)} \rightarrow x_{[60,56,\ldots,4,0]}^{(3)}$ is a permutation. This permutation is shown in phase T2 of Fig. 3 with bold line. □

5 Integral Attack on Reduced-Round PRESENT

In this section, we attack reduced-round PRESENT using the 4-th order integral distinguisher and 16-th order integral distinguisher.

The general attack procedure is given as follows.

1. Choose a set of 2^n ($n = 4$ or $n = 16$) plaintexts to construct a structure, where the rightmost n bits take all possible values of $\mathbb{F}_2^n$ while other bits are chosen to be arbitrary constants over $\mathbb{F}_2$. Obtain the corresponding ciphertexts after r-round encryption.

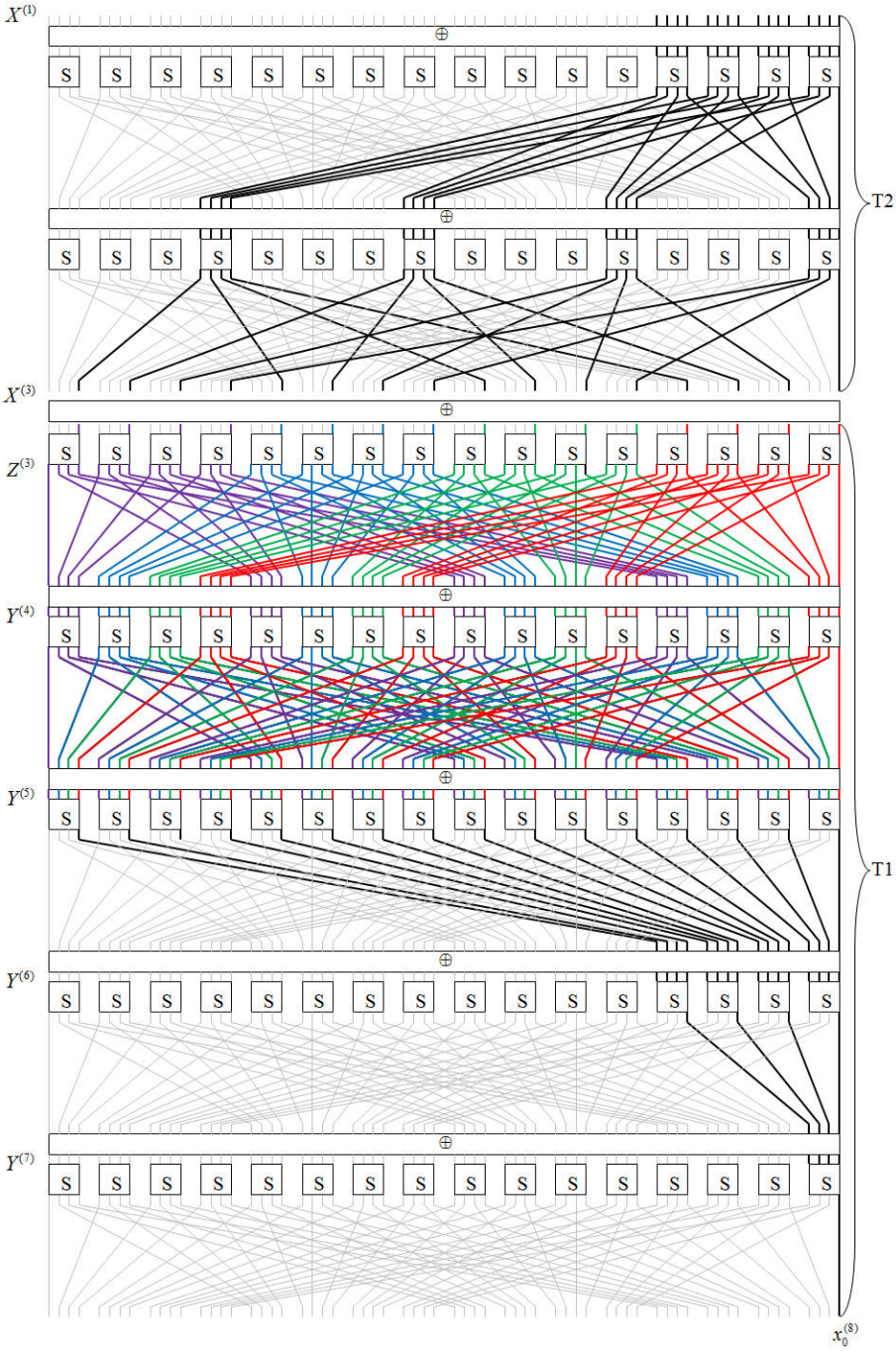

Fig. 3. 16-th order differential characteristics of PRESENT

2. For every guessing of the corresponding subkeys in the last $(r-m)$ rounds, decrypt the ciphertexts to obtain the one bit state $y_0^{(m+1)}$ after the m-th round, where m is the length of integral distinguishers.
3. Check whether $\bigoplus_{\Lambda} y_0^{(m+1)} (= \bigoplus_{\Lambda} x_0^{(m+1)})$ is zero, where Λ with $|\Lambda| = 2^n$ is the set of chosen plaintexts. If the equation is not satisfied, we know the guessed subkey is wrong. Then, we guess another subkey and repeat until the correct subkey is found.
4. Recover the remaining key bits in the master key by exhausting method.

Suppose we need to guess k bit subkey in the last $(r-m)$ rounds, the complexity of this attack can be estimated as follows. Step 1 needs about 2^n plaintexts which requires 2^n encryptions. In step 2 and step 3, a subkey needs about $\frac{r-m}{r} \times 2^n$ encryptions. For a wrong subkey guess, equation $\bigoplus_{\Lambda} y_0^{(m+1)} = 0$ holds with probability $\frac{1}{2}$. Therefore, to discard all the wrong k-bit subkey guesses, we need about k plaintext structures. Suppose the master key has $|K|$ bits, then the time complexity of step 4 is about $2^{|K|-k}$ r-round encryptions.

Thus, the data complexity is about $k \times 2^n$ chosen plaintexts. The time complexity in recovering these k key bits is about

$$\sum_{i=1}^{k} (2^n + 2^n \times \frac{r-m}{r} 2^k \times (\frac{1}{2})^{i-1}) \approx \frac{r-m}{r} \times 2^{n+k+1} \tag{7}$$

r-round encryptions. So, the final time complexity is $max\{\frac{r-m}{r} \times 2^{n+k+1}, 2^{|K|-k}\}$. A total of 2^k bits are required to keep track of possible values for the k key bits, so the memory complexity is 2^{k-3} bytes.

To attack $r = m+2$ round PRESENT, we need to guess 4 key bits $k_{[48,32,16,0]}^{(m+2)}$ in $K^{(m+2)}$, 16 key bits $k_{4j}^{(m+3)}$ for $0 \le j \le 15$ in $K^{(m+3)}$ to obtain state $y_0^{(m+1)}$. Thus, when considering the 4-th order integral attack, we have $n = 4$, $m = 5$, $r = 7$ and $k = 20$. In this case, we can recover 20 bit keys of 7-round PRESENT with data complexity $20 \times 2^4 \approx 2^{8.3}$, time complexity $\frac{2}{7} \times 2^{4+20+1} \approx 2^{23.2}$ 7-round encryptions and memory 2^{17} bytes. To recover the master key of PRESENT-80, we have to exhaust the remaining 60 key bits. Thus, the final time complexity is 2^{60}. Similarly, when considering 16-th order differential attack, we can recover 20 bit keys of 9-round PRESENT with data complexity $2^{20.3}$, time complexity $2^{34.8}$ 9-round encryptions and memory 2^{17} bytes. To recover the master key of PRESENT-80, the final time complexity is 2^{60}.

To attack $r = m+3$ round PRESENT, we need to guess all of 64 key bits in $K^{(m+4)}$ additionally, totally guessing 84 key bits. Utilizing the 16-th order integral, we can attack 10 rounds of PRESENT-128 with data complexity $2^{22.4}$, time complexity $2^{99.3}$ 10-round encryptions and memory 2^{81} bytes. The remaining 44 key bits can be obtained easily by exhausting method. To attack 8 rounds of PRESENT-80 using 4-th order integral, we need some properties of the key schedule. After examining the key schedule for 80-bit keys, we find that bits $k_{4j}^{(m+3)}$ for $j = 0$ or $5 \le j \le 15$ in $K^{(m+3)}$ and bits $k_{[48,16,0]}^{(m+2)}$ in $K^{(m+2)}$ are given from guessing all of $K^{(m+4)}$. Thus, in total we need to guess

$64 + (16 - 12) + (4 - 3) = 69$ bits of key, which leads to an attack of 8-round PRESENT-80 with data complexity $2^{10.1}$, time complexity $2^{72.6}$ 8-round encryptions and memory 2^{66} bytes. The remaining 11 key bits can be obtained easily by exhausting method.

6 Discussions and Conclusions

In this paper, we discuss the integral attack against bit-based block ciphers. We focus on the block cipher PRESENT and show that integral attack can be improved not only by the theorem of higher-order differential attack but also by using specific algebraic properties of Sboxes. What is more, the order of plaintexts in a set, which is important in bit-based integral attack, is not required here.

Combined with the algebraic properties of PRESENT's Sbox and its diffusion layer, we proposed two integral distinguishers: one 5 round (4-th order) integral distinguisher and one 7 round (16-th order) integral distinguisher, where the latter is the longest integral distinguisher of PRESENT as far as we know. Based on the integral distinguishers proposed in this paper, 10 (out of 31) rounds of PRESENT can be attacked.

Although the number of attack rounds in this paper are not so impressive as other statistical attacks [6,7,14], it is the first time that some new algebraic properties for constructing integral distinguishers of PRESENT are reported. For a further step, the algebraic techniques used in constructing longer integral distinguishers here may be also applied to other block ciphers to improve their known integral attacks.

Acknowledgements. We are grateful to the anonymous reviewers for their valuable comments on this paper. This work was supported by the National Basic Research Program of China (Grant No. 2013CB834203 and Grant No. 2013CB338002) and the National Natural Science Foundation of China (Grant No. 11171323).

References

1. Anderson, R., Biham, E., Knudsen, L.: Serpent: A Proposal for the Advanced Encryption Standard. NIST AES Proposal (1998), `http://www.cl.cam.ac.uk/rja14/serpent.html`
2. Bogdanov, A., Knudsen, L.R., Leander, G., Paar, C., Poschmann, A., Robshaw, M.J.B., Seurin, Y., Vikkelsoe, C.: PRESENT: An ultra-lightweight block cipher. In: Paillier, P., Verbauwhede, I. (eds.) CHES 2007. LNCS, vol. 4727, pp. 450–466. Springer, Heidelberg (2007)
3. Bogdanov, A., Leander, G., Nyberg, K., Wang, M.: Integral and Multidimensional Linear Distinguishers with Correlation Zero. In: Wang, X., Sako, K. (eds.) ASIACRYPT 2012. LNCS, vol. 7658, pp. 244–261. Springer, Heidelberg (2012)
4. Bogdanov, A., Rijmen, V.: Linear Hulls with Correlation Zero and Linear Cryptanalysis of Block Ciphers. Designs, Codes and Cryptography (2012)

5. Biryukov, A., Shamir, A.: Structural Cryptanalysis of SASAS. In: Pfitzmann, B. (ed.) EUROCRYPT 2001. LNCS, vol. 2045, pp. 394–405. Springer, Heidelberg (2001)
6. Cho, J.Y.: Linear Cryptanalysis of Reduced-Round PRESENT. In: Pieprzyk, J. (ed.) CT-RSA 2010. LNCS, vol. 5985, pp. 302–317. Springer, Heidelberg (2010)
7. Collard, B., Standaert, F.X.: A Statistical Saturation Attack against the Block Cipher PRESENT. In: Fischlin, M. (ed.) CT-RSA 2009. LNCS, vol. 5473, pp. 195–210. Springer, Heidelberg (2009)
8. Daemen, J., Knudsen, L., Rijmen, V.: The block cipher Square. In: Biham, E. (ed.) FSE 1997. LNCS, vol. 1267, pp. 149–165. Springer, Heidelberg (1997)
9. Daemen, J., Peeters, M., Van Assche, G., Rijmen, V.: Nessie Proposal: NOEKEON. In: First Open NESSIE Workshop (2000), `http://gro.noekeon.org/`
10. Ferguson, N., Kelsey, J., Lucks, S., Schneier, B., Stay, M., Wagner, D., Whiting, D.: Improved cryptanalysis of Rijndael. In: Schneier, B. (ed.) FSE 2000. LNCS, vol. 1978, pp. 213–230. Springer, Heidelberg (2001)
11. Knudsen, L., Wagner, D.: Integral Cryptanalysis. In: Daemen, J., Rijmen, V. (eds.) FSE 2002. LNCS, vol. 2365, pp. 112–127. Springer, Heidelberg (2002)
12. Lai, X.: Higher order derivatives and differential cryptanalysis. In: Proc. Symposium on Communication, Coding and Cryptography, in Honor of J. L. Massey on the Occasion of his 60th Birthday, Kluwer Academic Publishers, Dordrecht (1994)
13. Lucks, S.: Attacking seven rounds of Rijndael under 192-bit and 256-bit keys. In: Proc. 3rd AES Candidate Conf., pp. 215–229 (2000)
14. Wang, M.: Differential Cryptanalysis of reduced-round PRESENT. In: Vaudenay, S. (ed.) AFRICACRYPT 2008. LNCS, vol. 5023, pp. 40–49. Springer, Heidelberg (2008)
15. Yu, X., Wu, W., Li, Y., Zhang, L.: Cryptanalysis of Reduced-Round KLEIN Block Cipher. In: Wu, C.-K., Yung, M., Lin, D. (eds.) Inscrypt 2011. LNCS, vol. 7537, pp. 237–250. Springer, Heidelberg (2012)
16. Z'aba, M.R., Raddum, H., Henricksen, M., Dawson, E.: Bit-pattern based integral attack. In: Nyberg, K. (ed.) FSE 2008. LNCS, vol. 5086, pp. 363–381. Springer, Heidelberg (2008)
17. Zhang, W., Su, B., Wu, W., Feng, D., Wu, C.: Extending Higher-Order Integral: An Efficient Unified Algorithm of Constructing Integral Distinguishers for Block Ciphers. In: Bao, F., Samarati, P., Zhou, J. (eds.) ACNS 2012. LNCS, vol. 7341, pp. 117–134. Springer, Heidelberg (2012)