Linear Cryptanalysis of the Block Cipher PRESENT

Joo Yeon Cho

Helsinki University of Technology,
Department of Information and Computer Science,
P.O. Box 5400, FI-02015 TKK, Finland
joo.cho@tkk.fi

Abstract. PRESENT is a hardware-oriented block cipher suitable for resource constrained environment. We analyze PRESENT by a multidimensional linear cryptanalysis method. We claim that the PRESENT using 80-bit key can be attacked up to 23 round faster than key exhaustive search with around $2^{59.3}$ data complexity. Our results are superior to all the previous attacks under the known plaintext attack scenario. We demonstrate our claim by performing the linear attacks on reduced variants of PRESENT. Our results exemplify that multidimensional linear attack can improve the performance of classical linear attack significantly.

Keywords : Block Ciphers, Linear Cryptanalysis, PRESENT, Multidimensional Linear Cryptanalysis.

1 Introduction

PRESENT [3] is a lightweight SPN block cipher that is designed for resource restricted applications such as RFID and sensor networks. PRESENT was proposed by Bogdanov et al. at CHES 2007. As far as we know, three cryptanalytic studies on PRESENT have been presented so far. The first attack is a differential cryptanalysis that can recover the secret key up to 16 rounds using 2^{64} chosen texts and 2^{65} memory accesses [13]. The second attack is a differential attack using algebraic techniques that can recover a 80-bit key up to 16 rounds with similar complexity to [13] and a 128-bit key up to 19 rounds by 2^{113} computations [1]. The third attack is a statistical saturation attack that is applicable up to 24 round using 2^{57} chosen texts and 2^{57} time complexity under the condition that the parts of plaintexts are fixed to a constant value [4].

In this paper, we analyze PRESENT by a linear cryptanalysis method. We observe that PRESENT has a large number of linear approximations that hold with the same order of magnitude of correlations due to the simple structure of the round function. As shown in [6], a multidimensional linear attack can be efficiently applied to such cipher. According to our analysis, the PRESENT using 80-bit key can be attacked up to 23 round faster than key exhaustive search with around $2^{59.3}$ data complexity. Our results are superior to all the previous attacks under the known plaintext attack scenario. We demonstrate our claim by performing the linear attacks on reduced variants of PRESENT.

This paper is organized as follows. In Section 2, the structure of PRESENT is briefly described and the framework of multidimensional linear attack is presented. In Section 3, linear characteristics are derived and their capacities are computed. In Section 4, the attack algorithm using linear characteristics is described. In Section 5, our attack are applied to reduced variants of PRESENT and the experimental results are presented. Section 6 concludes this paper.

2 Preliminaries

2.1 Brief Description of PRESENT

PRESENT is a SPN block cipher that consists of 31 rounds. The block length is 64 bits and two key lengths of 80 and 128 bits are supported. Each of the 31 rounds consists of three layers: addRoundKey, SboxLayer and pLayer. The AddRoundKey is a 64-bit eXclusiveOR operation with a round key. The SboxLayer is a 64-bit nonlinear transform using a single S-box 16 times in parallel. The S-box is a nonlinear bijective mapping $S : \mathbb{F}_2^4 \mapsto \mathbb{F}_2^4$ given in Table 4. The pLayer is a bit-by-bit permutation $P : \mathbb{F}_2^{64} \mapsto \mathbb{F}_2^{64}$ given in Table 5. The design idea of SboxLayer and pLayer is adapted from Serpent [2] and DES block cipher [9], respectively.

Depending on the key size, two versions of key scheduling algorithms are provided. The structure and pseudo-code of PRESENT are illustrated in Figure 1 taken from [3]. For complete description of PRESENT we refer to the paper [3].

```
generateRoundKeys()
for $i = 1$ to 31 do
    addRoundKey(STATE,$K_i$)
    sBoxLayer(STATE)
    pLayer(STATE)
end for
addRoundKey(STATE,$K_{32}$)
```

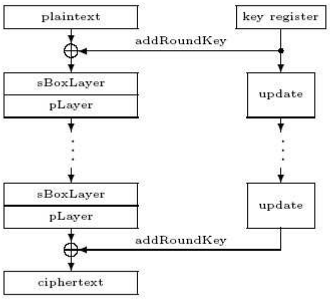

Fig. 1. Overview of PRESENT

2.2 Multidimensional Linear Cryptanalysis using Matsui's Algorithm 2

Multidimensional linear cryptanalysis is an extension of Matsui's classical linear cryptanalysis [8] in which multiple linear approximations are optimally exploited. The general framework of the multidimensional linear cryptanalysis adapting Matsui's algorithm 2 was presented by Hermelin et al. in [7]. In their paper, Hermelin et al. studied two statistic methods: the log-likelihood ratio (LLR) and the χ^2. We apply the χ^2 statistic method to PRESENT since the LLR method is not proper to PRESENT-like structure. The detailed explanation will be given in Section 4.4.

The brief framework of the χ^2 method is given below. Let m denote the dimension of linear approximations and p be the probability distribution of m-dimensional approximations. The capacity of p is defined by $C = \sum_{i=0}^{2^m-1} \frac{(p_i-u_i)^2}{u_i}$ where u is the uniform distribution. Suppose l is the length of the target key. For all values of $k \in [0, 2^l - 1]$, one obtains the empirical probability distributions q^k by measuring the frequency of m-dimensional vectors of which coordinates are parity bits of m linear independent approximations. Then the key candidates are ranked by computing the χ^2-statistic that is defined as

$$\mathcal{D}(k) = 2^m \sum_{i=0}^{2^m-1} (q_i^k - 2^{-m})^2 \tag{1}$$

which represents the l_2-distance of the q^k from the uniform distribution.

If the right key is ranked in the position of d from the top out of 2^l key candidates, we say that the attack has the advantage of $(l - \log_2 d)$ [12]. The advantage of the χ^2-method using statistic (1) is derived in Theorem 1 in [7] by

$$advantage = \frac{(NC - 4\Phi^{-2}(2P_s - 1))^2}{8(2^m - 1)}, \quad \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \tag{2}$$

where P_s is the success probability, N is the amount of data and C is the capacity.

2.3 Notations

Given a 64-bit string X, we use the notation X_i for indicating the i-th 4-bit string of X as counting X_0 at the right. Let S_i denote the i-th S-box in the SboxLayer, P denote the permutation mapping defined in Table 5 and $K^{(r)}$ denote the r-th round key. Let $\Omega^{(r)}(a, b)$ denote a linear characteristic over r rounds of PRESENT with the input mask a and the output mask b where $a, b \in \mathbb{F}_2^{64}$. We write $\Omega^{(r)}(a, b) = \Omega^{(r)}(a_i, b_i)$ if only $a_i, b_i \neq 0$ and other bits of a and b are zero. In our notation of the bit masks, we identify $\mathbb{F}_2^4$ with $\mathbb{Z}_{16}$. We use the little endian for bit notation through the paper, that is, the least significant bit is counted at the rightmost.

3 Linear Characteristics of PRESENT

Let $\pi(\alpha, \beta)$ denote a linear approximation of S-box S where $\alpha, \beta \in \mathbb{F}_2^4$ are an input and output mask of S, respectively. The correlation of $\pi(\alpha, \beta)$ is denoted by $\rho(\alpha, \beta)$. We observe that the S-box has the following properties:

S1. For $\alpha, \beta \in \{2, 4, 8\}$, $\rho(\alpha, \beta) = \pm 2^{-2}$ except that $\rho(8, 4) = 0$; and
S2. For $\alpha \in \{1, 2, 4, 8\}$, $\rho(\alpha, 1) = \rho(1, \alpha) = 0$.

We focus on the linear trails exploiting the linear approximations of which the input and output masks have a single active bit, respectively. The linear masks having more than one active bits affect at least two S-boxes in the consecutive round due to the permutation property, which yield much less correlations in the multiple rounds of PRESENT.

Definition 1. *A single-bit linear trail is a linear trail where the input and output masks of linear approximations of all intermediate S-boxes are of Hamming weight one.*

Hereafter, a linear trail means a single-bit linear trail unless specified otherwise. Let A be a subset of the SboxLayer defined as $A = \{S_5, S_6, S_7, S_9, S_{10}, S_{11}, S_{13}, S_{14}, S_{15}\}$ and B be a set of bit locations defined as $B = \{4i + 1, 4i + 2, 4i + 3 | 0 \leq i \leq 15, S_i \in A\}$. We observe that the permutation P given by Table 5 has the following properties:

P1. For $0 \leq x \leq 63$, $P \circ P \circ P(x) = x$; and

P2. If $x \in B$, then $P(x) \in B$. If $x \notin B$, then one of $P(x), P \circ P(x)$ or $P \circ P \circ P(x)$ becomes the least significant bit of an S-box.

According to Property S2, the linear trails passing through the least significant bit of S-box do not have correlations. Hence, we focus on the linear characteristics exploiting the S-boxes in the set A only.

We build the linear characteristics ending with a single S-box in order to minimize the active S-boxes in the last round. Due to the permutation, a single S-box always activates four S-boxes in the next round. Since the output of each S-box is 4 bits, our attack targets to recover the 16 bits of the last round key.

3.1 4 Round Linear Characteristics

Let $\Omega^{(4)}(a,b)$ denote a linear characteristic starting with all S-boxes in the set A and ending with S_5 where a and b are 64-bit strings. The $\Omega^{(4)}(a,b)$ is actually composed of nine linear characteristics $\Omega^{(4)}(a_i,b_5)$ starting with a single $S_i \in A$ and ending with S_5. Due to properties of the S-box and the permutation, each $\Omega^{(4)}(a_i,b_5)$ has four dominant linear trails as shown in Figure 2. Note that the $\Omega^{(4)}(a_i,b_5)$ allows the linear trail using the least significant bit since the a_i and b_5 can be arbitrary values between 0 and 15. Each linear trail of $\Omega^{(4)}(a_i,b_5)$ takes the following path:

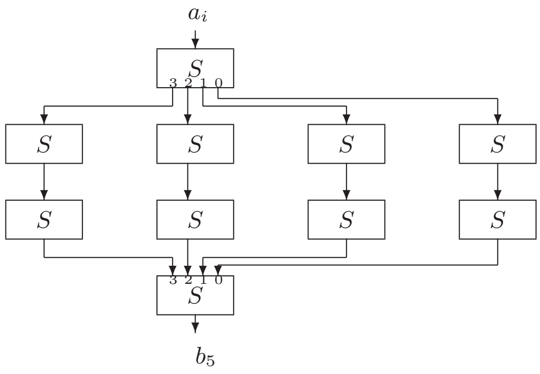

Fig. 2. Linear trails in $\Omega_4(a_i,b_5)$

$$\pi(a_i,2^j) \rightarrow \pi(2^u,2) \rightarrow \pi(2^v,2) \rightarrow \pi(2^j,b_5)$$

where $j \in \{0,1,2,3\}$ and

$$u = \begin{cases} 1, & \text{if } i = 5,9,13 \\ 2, & \text{if } i = 6,10,14 \\ 3, & \text{if } i = 7,11,15 \end{cases} \quad \text{and} \quad v = \begin{cases} 1, & \text{if } i = 5,6,7 \\ 2, & \text{if } i = 9,10,11 \\ 3, & \text{if } i = 13,14,15. \end{cases}$$

According to the correlation theorem [11], the correlation of $\Omega^{(4)}(a_i,b_5)$ given the key K is the summation of all linear trails in the characteristic. See also Section 7.9 in [5]. We estimate the correlation by the summation of correlations of dominant four linear trails as follows:

$$c^{(4)}(a_i, b_5; K) = \sum_{j=0}^{3} (-1)^{k_j} \rho(a_i, 2^j) \cdot 2^{-2} \cdot 2^{-2} \cdot \rho(2^j, b_5) \tag{3}$$

where k_j denotes the combination of round key bits located in the j-th linear trail. Since $\Omega^{(4)}(a, b)$ has nine linear characteristics and each one has four linear trails with the correlations of (3), the capacity of $\Omega^{(4)}(a, b)$ is estimated by

$$C^{(4)}(K) = 9 \times \sum_{\alpha=0}^{15} \sum_{\beta=0}^{15} \left(c^{(4)}(\alpha, \beta; K) \right)^2 . \tag{4}$$

Due to Parseval's theorem, $\sum_{\alpha=0}^{15} \rho(\alpha, \beta)^2 = 1$, for a fixed $\beta \in \mathbb{F}_2^4$. Since the S-box is bijective, also $\sum_{\beta=0}^{15} \rho(\alpha, \beta)^2 = 1$, for a fixed $\alpha \in \mathbb{F}_2^4$. Hence, we obtain the following lemma:

Lemma 1. *Let us assume that the round keys of PRESENT are statistically independent. For $a, b \in \mathbb{F}_2^{64}$, the expected value of the capacity of $\Omega^{(4)}(a, b)$ over the secret key K is estimated to be* $2^{-2.83}$.

Proof. From (3), we derive

$$\begin{aligned} E_K \left[\left(c^{(4)}(\alpha, \beta; K) \right)^2 \right] &= E_K \left[\left(\sum_{j=0}^{3} (-1)^{k_j} \cdot 2^{-4} \cdot \rho(\alpha, 2^j) \cdot \rho(2^j, \beta) \right)^2 \right] \\ &= E_K \left[\sum_{j=0}^{3} \sum_{v=0}^{3} (-1)^{k_j \oplus k_v} \cdot 2^{-8} \cdot \rho(\alpha, 2^j) \cdot \rho(2^j, \beta) \cdot \rho(\alpha, 2^v) \cdot \rho(2^v, \beta) \right] \\ &= 2^{-8} \cdot \sum_{j=0}^{3} \rho(\alpha, 2^j)^2 \cdot \rho(2^j, \beta)^2 \end{aligned}$$

since

$$E_K \left[(-1)^{k_j} (-1)^{k_v} \right] = \begin{cases} 1 & \text{if } j = v, \\ 0 & \text{if } j \neq v \Leftrightarrow k_j \neq k_v \end{cases}$$

under the assumption that the round key bits are statistically independent. Then, due to Parseval's theorem, we get

$$E_K[C^{(4)}(K)] = 9 \cdot \sum_{\alpha=0}^{15} \sum_{\beta=0}^{15} \left(2^{-8} \cdot \sum_{j=0}^{3} \rho(\alpha, 2^j)^2 \cdot \rho(2^j, \beta)^2 \right) = 9 \cdot 4 \cdot 2^{-8} = 2^{-2.83}.$$

□

Since the round keys are expanded from the user-supplied key K by the key scheduling algorithm, the assumption that the round keys are statistically independent is not always fulfilled in general. Hence, Lemma 1 provides an estimated average capacity of linear characteristic. However, as studied in Section 7.9 in [5], the assumption on statistical independence of the round keys is reasonable in practice and we verified the results of Lemma 1 by experiments. We use this assumption for the analysis of further rounds of PRESENT in the next section.

3.2 $n+4$ Round Linear Characteristic

Let us recall $B = \{4i+1, 4i+2, 4i+3 | 0 \le i \le 15, S_i \in A\}$ defined in the previous section. Let Ω_1 denote a 1-round characteristic that exploits all linear trails indexed by B, as shown in Figure 3. For a positive integer n, the $n+4$ round linear characteristic $\Omega^{(n+4)}(a,b)$ is constructed by repeating n times the Ω_1 and adding the 4 round characteristic at the bottom as shown in Figure 4. We note that the S_5 in the last round of the characteristic can be replaced by any other S-box included in A.

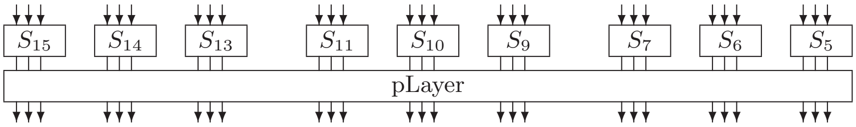

Fig. 3. 1-round linear characteristic Ω_1

In contrast to the four round characteristic, the $\Omega^{(n+4)}(a,b)$ do not allow the linear trails using the least significant bit of S-box. If the $\pi(1,\alpha)$ or $\pi(\alpha,1)$ for $\alpha \in \{2^0, 2^1, 2^2, 2^3\}$ is used at least once in the $n+4$ round characteristics, then such linear trails do not have correlation since $\rho(\alpha,1) = \rho(1,\alpha) = 0$ due to Property S2.

Corollary 1. *The expected value of the capacity of $\Omega^{(4)}(a,b)$ without the least significant bit trail is* $2^{-3.25}$.

Proof. The claim is followed by the proof of Lemma 1 with three linear trails. □

The expected capacity of $n+4$ round linear characteristics is calculated in two steps: first, the expected correlations of all linear trails over the second to the n-th round are calculated; then Corollary 1 is applied to compute the desired capacity.

Let $\theta_i^{(r)}$ denote the correlation of the i-th linear trail over the second to the r-th round where $i \in B$. As shown in Figure 4, any linear trail of the $(r+1)$-th round can be extended to three linear trails of the r-th round. Hence, given K, the $\theta_i^{(r+1)}$ is recursively expressed as

$$\forall i \in B,\ \ \theta_i^{(r+1)}(K) = \sum_{j=1}^{3} (-1)^{K_\nu^{(r+1)}} \rho(2^j, 2^{i \bmod 4})\ \theta_{IP(\nu)}^{(r)}(K),\ \nu = 4\lfloor i/4 \rfloor + j \tag{5}$$

where $K_\nu^{(r+1)}$ denotes the ν-th bit of $r+1$ round key and IP is an inverse mapping of P. Since $\{IP(\nu) | i \in B, 1 \le j \le 3\} = B$, the expected values of $\theta_i^{(n)}$ are calculated by the following algorithm:

1. Initialize $\theta_i^{(0)} = 1$ for all $i \in B$.
2. Repeat for $1 \le r \le n$,
   (a) Compute $\theta_i^{(r)}(K)$ using (5) for all possible values of $K \in \mathbb{F}_2^{27}$.
   (b) Assign $\theta_i^{(r)} = E_K(|\theta_i^{(r)}(K)|)$.

Having $\theta_i^{(n)}$ for all $i \in B$, the capacity of the $n+4$ round characteristic is computed by the following lemma:

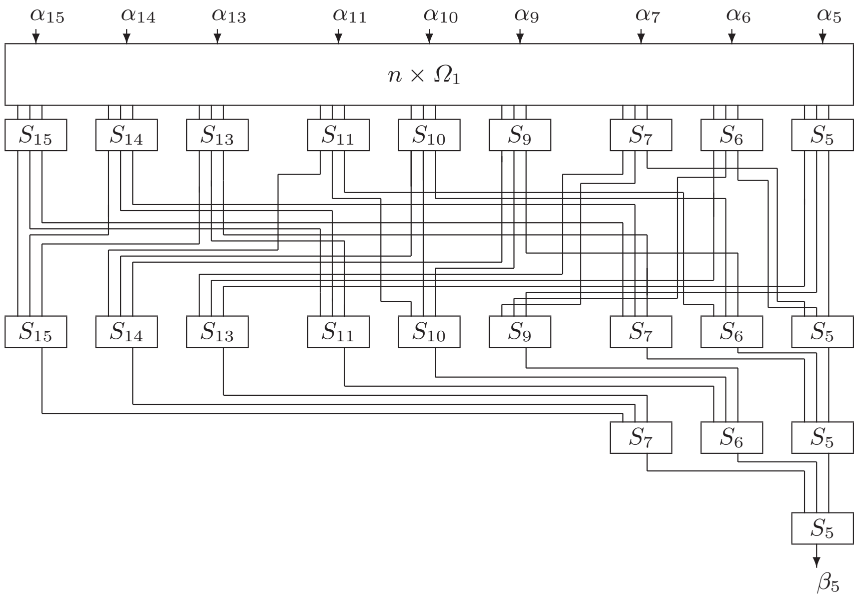

Fig. 4. $(n+4)$ rounds linear characteristic $\Omega^{(n+4)}(\alpha, \beta)$

Lemma 2. *Let us assume that the round keys of PRESENT are statistically independent. For* $a, b \in \mathbb{F}_2^{64}$*, the expected capacity of* $\Omega^{(n+4)}(a, b)$ *over the key* K *is*

$$2^{-8} \sum_{t \in B} \left(\theta_t^{(n)}\right)^2.$$

Proof. For $a_i, b_5 \in \mathbb{F}_2^4$, we have

$$c^{(n+4)}(a_i, b_5; K) = \sum_{j=1}^{3} (-1)^{k_j} \cdot \rho(a_i, 2^j) \cdot \theta_t^{(n)}(K) \cdot 2^{-2} \cdot 2^{-2} \cdot \rho(2^j, b_5), \quad t = P(4i + j).$$

Similarly to Corollary 1, we get

$$E_K \left[\left(c^{(n+4)}(\alpha, \beta; K) \right)^2 \right] = E_K \left[2^{-8} \sum_{j=1}^{3} \rho(a_i, 2^j)^2 \cdot (\theta_t^{(n)}(K))^2 \cdot \rho(2^j, b_5)^2 \right].$$

Due to Parseval's theorem, the capacity is

$$E_K[C^{(n+4)}(K)] = 9 \sum_{\alpha=0}^{15} \sum_{\beta=0}^{15} E_K \left[\left(c^{(n+4)}(\alpha, \beta; K) \right)^2 \right] = 9 \cdot 3 \cdot 2^{-8} \cdot E_K \left[\left(\theta_t^{(n)}(K) \right)^2 \right].$$

Since the mapping $x \mapsto P(x)$ for $x \in B$ is bijective, the claim follows. □

Table 1 shows the average capacities of $n+4$ round linear characteristics that are calculated by Lemma 2. By interpolating the results shown in Table 1, we can see that the capacity of $n+4$ round characteristic is estimated into the following formula:

$$C^{(n+4)} \approx 2^{-5.83 - 2.61(n-5)}, \quad n > 0. \tag{6}$$

The results of Lemma 2 means that the expected capacity of $\Omega^{(n+4)}(a,b)$ is the sum of the square of correlations of all linear trails starting from the second round and ending to the second last round. Since the multidimensional linear attack takes all the correlations of the first round and the last round, the capacity of the characteristic is only contributed by the linear trails of $n+2$ rounds.

The idea behind Lemma 2 is well consistent to the results of [10]. In Theorem 1 of [10], it is proved that the average value of the square of correlation [1] of the linear approximation is the sum of the square of correlations of all linear trails. See also [5]. This theorem can be extended to the multidimensional linear cryptanalysis as follows:

Proposition 1. *The expected capacity of an m-dimensional linear approximation is the sum of the square of the expected correlations of all the linear trails that all the $2^m - 1$ one-dimensional linear approximations have.*

In the next section, we describe how the linear characteristics are used for the linear attacks based on Matsui's algorithm 2.

round	capacity	round	capacity
5	$2^{-5.83}$	18	$2^{-39.74}$
6	$2^{-8.41}$	19	$2^{-42.36}$
7	$2^{-11.02}$	20	$2^{-44.97}$
8	$2^{-13.63}$	21	$2^{-47.58}$
9	$2^{-16.24}$	22	$2^{-50.19}$
10	$2^{-18.85}$	23	$2^{-52.80}$
11	$2^{-21.46}$	24	$2^{-55.41}$
12	$2^{-24.08}$	25	$2^{-58.02}$
13	$2^{-26.69}$	26	$2^{-60.64}$
14	$2^{-29.30}$	27	$2^{-63.25}$
15	$2^{-31.91}$	28	$2^{-65.86}$
16	$2^{-34.52}$	29	$2^{-68.47}$
17	$2^{-37.13}$	30	$2^{-71.08}$

Table 1. Evaluation of capacities of $n+4$ round characteristics

4 Multidimensional Linear Attacks on PRESENT

4.1 Selection of linear independent approximations

For a fixed n, the dimension of input masks of the $\Omega^{(n)}(a,b)$ is 4×9 and the dimension of output mask is 5. We observe that the linear trails passing more than one S-boxes in each round have much less correlations than a single-bit linear trails. Hence, for efficiency, we choose 8-dimensional linear approximations for each $\Omega^{(n)}(a_i, b_5)$ where $S_i \in A$. Since there are nine such characteristics in $\Omega^{(r)}(a,b)$, the number of linear approximations spanned for our attack is $9 \times (2^8 - 1)$ in total.

We use 8 unit vectors as the linear independent approximations of $\Omega^{(n)}(a_i, b_5)$. Even though each unit vector does not have any correlation, all the non-negligible linear approximations

[1] This is called *the potential* in [10].

can be spanned by these unit vectors. The merit of this approach is that the evaluation of the input and output parities of linear approximations is not needed; the concatenation of input and output of linear characteristic indicates the index of the multidimensional counter. Hence, the time complexity of the attack can be reduced by at least a factor of m where m is the dimension of the linear approximations.

4.2 Attack Complexity

Let n_c be the number of linear characteristics and m is the dimension of each characteristic. The number of linear approximations available for the attack is $n_c \times (2^m - 1)$ in total. Then, the amount of data required for χ^2 statistic method is obtained by modifying (2) as follows:

$$N = \left(\sqrt{advantage \cdot 8 \cdot n_c \cdot (2^m - 1)} + 4\Phi^{-2}(2P_s - 1)\right) / C^{(r)}$$

where P_s is the success probability and $C^{(r)}$ is the capacity. If the target key is l-bit, the time complexity of the attack is estimated as

$$T = N \cdot n_c \cdot 2^l + n_c \cdot 2^{l+m} \approx N \cdot n_c \cdot 2^l$$

since N is usually much larger than 2^m. The memory complexity is around $n_c \cdot 2^{l+m}$.

For the $\Omega^{(n+4)}(a, b)$, we set $m = 8$ and $n_c = 9$. Then the full advantage (16 bits) of the attack with the success probability 0.95 is achieved by the data complexity of

$$N = \left(\sqrt{16 \cdot 8 \cdot 9 \cdot (2^8 - 1)} + 4\Phi^{-2}(2P_s - 1)\right) / C^{(r)} \approx 2^{9.1} / C^{(r)}$$

with time complexity of $9 \cdot 2^{16} \cdot N = 2^{19.2}N$ and the memory of $9 \cdot 2^{24} \approx 2^{27.2}$. Without increasing the amount of data, we can obtain another 16 bits of the last round key by using a linear characteristic ending with $S_i \in A - \{S_5\}$. We also note that our linear characteristic can be converted to a reciprocal form; starting with a single S_i at round 2 and ending with nine S-boxes of A in the last round. Then, we target to recover the first round key. According to the key scheduling of PRESENT, the 64 bits of the user-supplied secret key is used in the first round key without modification. Hence, we can recover $16n$ bits of the secret key by applying linear attacks repeatedly n times. The remaining key $128 - 16n$ or $80 - 16n$ can be obtained by exhaustive key search. We compares our attacks with previous attacks against various rounds of PRESENT in Table 2.

round	data	time	source
16	2^{64}	2^{65}	Differential [13]
	$2^{41.0}$	$2^{60.2}$	Linear (this paper)
19	-	2^{113}	Differential + Algebraic [1]
	$2^{48.8}$	$2^{68.0}$	Linear (this paper)
23	$2^{59.3}$	$2^{78.5}$	Linear (this paper)

Table 2. Comparison of data and time complexity of the attacks against PRESENT under the known plaintext attack scenario

4.3 Algorithm of Linear Attack for Recovering 16 bits of the Key

1. Prepare $9 \cdot 2^{16+8}$ counters and initialize them by zero.
2. Collect N plaintext-ciphertext pairs.
3. For $K = 0, \ldots, 2^{16-1}$,
   (a) Decrypt the ciphertext partially over the last round by using K and get the output of S_5 of $r-1$ round.
   (b) Obtain nine indices by concatenating the input of $S_i \in A$ of the plaintext and the output of S_5 of the decrypted ciphertext.
   (c) increase by 1 the counters indicated by above indices.
4. Repeat updating counters with N text pairs.
5. Compute l_2 distance between the probability distribution for each K and uniform distribution.
6. Sort out the candidate keys according to their l_2 distances.
7. Perform the right key search from the top rank.

4.4 Discussion

Weakness of bit permutation Our attack is mainly based on the observation that PRESENT has a large number of linear approximations with the same magnitude of correlations. It seems that this weakness is caused by the lack of diffusion property of the bit permutation. Even though the bit permutation is desirable for efficient hardware implementation, it has a potential weakness that input bits and output bits have one-to-one correspondence. Hence, a single-bit linear approximations of an S-box of any round can be connected to another single-bit linear approximation of next round through the permutation layer. Since the S-box of PRESENT has multiple linear approximations of which linear masks have a single active bit, one can construct multiple single-bit linear trails over arbitrary rounds.

Note that this weakness does not appear in the linear transformation functions of Serpent [2] or AES [5] since any single output bit of the linear transformation is expressed as a boolean function of at least two input bits.

Correlation and Piling Up Lemma The designers of PRESENT proved in Theorem 2 of [3] that the maximum correlation of a linear approximation of four rounds of PRESENT is 2^{-6}. As a result, the maximal correlation of a 28-round linear approximation was estimated to be $(2^{-6})^7 = 2^{-42}$ by Piling Up lemma [3]. On the other hand, according to our analysis, the capacity of the 28 round PRESENT is estimated to be around $2^{-65.9}$.

The difference between the designers' estimation and our result is originated from the two facts: firstly, a linear approximation in PRESENT has multiple linear trails with same amplitude of correlations. As mentioned before, the expected value of the squared correlation of linear approximation is the sum of the square of correlations of all linear trails [10, 5]. Hence, the squared correlation of a single linear trail is largely deviated from the true value. Secondly, PRESENT has a large amount of equally-correlated linear approximations. Since the data complexity of multidimensional linear cryptanalysis is inversely proportional to the capacity of the multiple linear approximations [7], the data complexity is reduced significantly compared to the estimate by a correlation of a single approximation.

The χ^2 and LLR method Finally, we justify the reason why the LLR method is not used for the attack on PRESENT even if the LLR method showed a better performance than the χ^2 method in the attack on SERPENT [6]. As described in [6], the LLR method is more advantageous compared to the χ^2 method if the pre-computed profile of probability distribution is accurate. However, the distribution of linear approximations in PRESENT heavily depends on the key values so that the space of profile of the probability distribution becomes too large. On the other hand, the χ^2 method does not need to know the distribution accurately. We only need to detect a large deviation from the uniform distribution. It is an open problem whether there is a way to apply the LLR method efficiently for the attacks against PRESENT. According to [7], the successful multidimensional attack using LLR method could further reduce the data complexity of our attacks by a factor of around 2^3.

5 Experiments

We performed our attacks up to 9 rounds of PRESENT with randomly chosen 32 secret keys. The data and time complexity required for recovering the 16 bits of round key is displayed in Table 3. The plaintexts were randomly generated and encrypted by PRESENT. Figure 5 illustrates the average of advantage of the attacks and the required amount of data on reduced variants of PRESENT. The dashed lines represent theoretically estimations drawn by (2) and the solid lines are empirical results. We can see that the estimation of the full advantage of the attack is well matched with empirical results up to 9 rounds PRESENT. Due to the restriction of computational resources, we did not proceed our experiments for further rounds. However, based on the experimental results, we can conclude that our estimates of attack complexity against further rounds PRESENT are reasonable.

target round	characteristic	capacity	data	time
5	$\Omega^{(4)}$	$2^{-2.8}$	$2^{11.9}$	$2^{31.1}$
6	$\Omega^{(5)}$	$2^{-5.8}$	$2^{14.9}$	$2^{34.1}$
7	$\Omega^{(6)}$	$2^{-8.4}$	$2^{17.5}$	$2^{36.7}$
8	$\Omega^{(7)}$	$2^{-11.0}$	$2^{20.1}$	$2^{39.3}$
9	$\Omega^{(8)}$	$2^{-13.6}$	$2^{22.7}$	$2^{41.9}$

Table 3. Data and time complexity required for experimental attacks

6 Conclusion

One of the recent trends to prove the resistance of linear cryptanalysis is to provide a lower bound to the number of the active S-boxes involved in a linear characteristic. Even though PRESENT provides a provable security against linear cryptanalysis according to this rule, our attack shows that the resistance of the classical linear cryptanalysis does not always thwart the multidimensional linear attacks. Even though a simple and regular structure of the cipher is desirable to the hardware-oriented block ciphers, such ciphers may have a large number of linear approximations by which a multidimensional linear attack can be applied efficiently. It is interesting to see that our attack can be applied to some other ciphers that have simple structures, like AES.

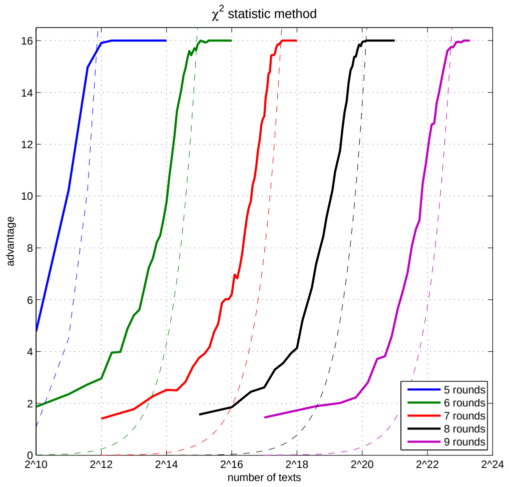

Fig. 5. Empirical evaluation of linear attacks on reduced variants of PRESENT

Acknowledgment

I wish to thank Kaisa Nyberg for very useful comments that helped to improve the paper. I am also grateful to anonymous reviewers of Asiacrypt'09 for their valuable comments.

References

1. M. Albrecht and C. Cid, *Algebraic techniques in differential cryptanalysis*, Cryptology ePrint Archive, Report 2008/177, 2008, http://eprint.iacr.org/2008/177.
2. R. Anderson, E. Biham, and L. Knudsen, *Serpent: A proposal for the Advanced Encryption Standard*, First Advanced Encryption Standard (AES) conference, 1998.
3. A. Bogdanov, L. R. Knudsen, G. Leander, C. Paar, A. Poschmann, M. J. B. Robshaw, Y. Seurin, and C. Vikkelsoe, *PRESENT: An ultra-lightweight block cipher*, Cryptographic Hardware and Embedded Systems - CHES 2007, 9th International Workshop, Vienna, Austria, September 10-13, 2007, Proceedings, Lecture Notes in Computer Science, vol. 4727, Springer, 2007, pp. 450–466.
4. B. Collard and F. Standaert, *A statistical saturation attack against the block cipher PRESENT*, Available at http://www.dice.ucl.ac.be/ fstandae/PUBLIS/62.pdf, 2009.
5. J. Daemen and V. Rijmen, *The Design of Rijndael- AES, the Advanced Encryption Standard*, Springer-Verlag, 2002.
6. M. Hermelin, J. Cho, and K. Nyberg, *Multidimensional linear cryptanalysis of reduced round Serpent*, Information Security and Privacy, 13th Australasian Conference, ACISP 2008, Wollongong, Australia, July 7-9, 2008, Proceedings, Lecture Notes in Computer Science, vol. 5107, Springer, 2008, pp. 203–215.
7. M. Hermelin, J. Y. Cho, and K. Nyberg, *Multidimensional extension of Matsui's algorithm 2*, Fast Software Encryption 2009, Lecture Notes in Computer Science, Springer, to appear.
8. M. Matsui, *Linear cryptoanalysis method for DES cipher*, Advances in Cryptology - EUROCRYPT '93, Lecture Notes in Computer Science, vol. 765, Springer, 1993, pp. 386–397.
9. National Bureau of Standards, *FIPS PUB 46-3: Data Encryption Standard (DES)*, National Institute for Standards and Technology, January 1977.

10. K. Nyberg, *Linear approximation of block ciphers*, Advances in Cryptology - EUROCRYPT '94, Workshop on the Theory and Application of Cryptographic Techniques, Perugia, Italy, May 9-12, 1994, Proceedings, Lecture Notes in Computer Science, vol. 950, Springer, 1994, pp. 439–444.
11. ______, *Correlation theorems in cryptanalysis*, Discrete Applied Mathematics **111** (2001), 177–188.
12. A. Selçuk, *On probability of success in linear and differential cryptanalysis*, Journal of Cryptology **21** (2008), no. 1, 131–147.
13. M. Wang, *Differential cryptanalysis of reduced-round PRESENT*, Progress in Cryptology - AFRICACRYPT 2008, First International Conference on Cryptology in Africa, Casablanca, Morocco, June 11-14, 2008. Proceedings, Lecture Notes in Computer Science, vol. 5023, 2008, pp. 40–49.

A The S-box and Permutation tables of PRESENT

The S-box and the permutation tables of PRESENT are given in Table 4 and Table 5, respectively.

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

Table 4. S-box table of PRESENT in hexadecimal notation

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

Table 5. Permutation table of PRESENT

B Correlation Table of S-box of PRESENT

Given an input mask α and an output mask β where $\alpha, \beta \in \mathbb{F}_2^4$, the correlation of the linear approximation $\alpha \cdot x \oplus \beta \cdot S(x) = 0$ of the S-box is measured as follows:

$$c(\alpha, \beta) = 2^{-4}(\#(\alpha \cdot x \oplus \beta \cdot S(x) = 0) - \#(\alpha \cdot x \oplus \beta \cdot S(x) = 1))$$

where the $\cdot$ notation stands for the standard inner product. The correlation table of the S-box is given in Table 6.

$\alpha \backslash \beta$	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
1	0	0	0	0	-2^{-1}	0	-2^{-1}	0	0	0	0	0	-2^{-1}	0	2^{-1}
2	0	2^{-2}	2^{-2}	-2^{-2}	-2^{-2}	0	0	2^{-2}	-2^{-2}	0	2^{-1}	0	2^{-1}	-2^{-2}	2^{-2}
3	0	2^{-2}	2^{-2}	2^{-2}	-2^{-2}	-2^{-1}	0	-2^{-2}	2^{-2}	-2^{-1}	0	0	0	-2^{-2}	-2^{-2}
4	0	-2^{-2}	2^{-2}	-2^{-2}	-2^{-2}	0	2^{-1}	-2^{-2}	-2^{-2}	0	-2^{-1}	0	0	-2^{-2}	2^{-2}
5	0	-2^{-2}	2^{-2}	-2^{-2}	2^{-2}	0	0	2^{-2}	2^{-2}	-2^{-1}	0	2^{-1}	0	2^{-2}	2^{-2}
6	0	0	-2^{-1}	0	0	-2^{-1}	0	0	-2^{-1}	0	0	2^{-1}	0	0	0
7	0	0	2^{-1}	2^{-1}	0	0	0	0	-2^{-1}	0	0	0	0	2^{-1}	0
8	0	2^{-2}	-2^{-2}	0	0	-2^{-2}	2^{-2}	-2^{-2}	2^{-2}	0	0	-2^{-2}	2^{-2}	2^{-1}	2^{-1}
9	2^{-1}	-2^{-2}	-2^{-2}	0	0	2^{-2}	-2^{-2}	-2^{-2}	-2^{-2}	-2^{-1}	0	-2^{-2}	2^{-2}	0	0
a	0	2^{-1}	0	2^{-2}	2^{-2}	2^{-2}	-2^{-2}	0	0	0	-2^{-1}	2^{-2}	2^{-2}	-2^{-2}	2^{-2}
b	-2^{-1}	0	0	-2^{-2}	-2^{-2}	2^{-2}	-2^{-2}	-2^{-1}	0	0	0	2^{-2}	2^{-2}	2^{-2}	-2^{-2}
c	0	0	0	-2^{-2}	-2^{-2}	-2^{-2}	-2^{-2}	2^{-1}	0	0	-2^{-1}	-2^{-2}	2^{-2}	2^{-2}	-2^{-2}
d	2^{-1}	2^{-1}	0	-2^{-2}	-2^{-2}	2^{-2}	2^{-2}	0	0	0	0	2^{-2}	-2^{-2}	2^{-2}	-2^{-2}
e	0	2^{-2}	2^{-2}	-2^{-1}	2^{-1}	-2^{-2}	-2^{-2}	-2^{-2}	-2^{-2}	0	0	-2^{-2}	-2^{-2}	0	0
f	2^{-1}	-2^{-2}	2^{-2}	0	0	-2^{-2}	-2^{-2}	-2^{-2}	2^{-2}	2^{-1}	0	2^{-2}	2^{-2}	0	0

Table 6. Correlation table of S-box of PRESENT: $c(\alpha, \beta)$

C Theoretical Estimation of Advantage of the Attack

The evaluation of the theoretical estimates on the advantage of the attack and data complexity against from 16 rounds to 24 rounds of PRESENT is given in Figure 6.

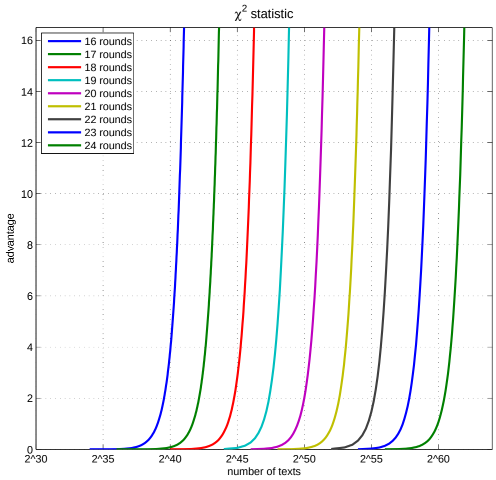

Fig. 6. Theoretical estimation of advantage of the attack against various variants of PRESENT