Differential Cryptanalysis of Reduced-Round ICEBERG

Abstract. ICEBERG is proposed by Standaert *et al.* in FSE 2004 for reconfigurable hardware implementations. It uses 64-bit block size and 128-bit key and the round number is 16. Specially, It is a SPN block cipher and all components are involutional and allow very efficient combinations of encryption/decryption. In this paper, we propose an elaborate method to identify the 6-round differentials and present the differential attack on 7-round ICEBERG with 2^{54} chosen plaintexts and $2^{90.28}$ 7-round encryptions. Then we use multiple differentials to attack 8-round ICEBERG with 2^{63} chosen plaintexts and 2^{96} 8-round encryptions. The previous linear cryptanalysis can only attack 7-round ICEBERG with the whole codebook. It means that ICEBERG is more resistant to linear cryptanalysis than differential cryptanalysis. Although our attack cannot threat ICEBERG, we give the best attack for ICEBERG and our elaborate method to identify multiple differential can be used for other similar block ciphers.

Keywords: Differential Cryptanalysis, Light-Weight, Block Cipher, ICEBERG, Involutional

1 Introduction

RFID systems and sensor networks have been aggressively deployed in a variety of applications, but their further pervasive usage is mainly limited by lots of security and privacy concerns. As RFID tags and sensor networks are low cost with limited resources, the present cryptographic primitives cannot be feasible. So the security primitives suitable for these light-weight environments must be designed. Recently, several light-weight block ciphers have been proposed such as PRESENT [1], mCRYPTON [2], HIGHT [3], SEA [4] and KTANTAN [5] etc. In general, a block cipher based on SP-network structure has the different encryption and decryption process like AES [6], which will increase the hardware costs. Although the block cipher based on the Feistel structure does not have such disadvantage, its slow avalanche effect requires the large round number to guarantee the security. In this way, how to design an involutional block cipher based on SP-network structure has become an important object in the field of light-weight block cipher.

At FSE 2004, Standaert *et al.* proposed a fast involutional block cipher with SP-network structure optimized for reconfigurable hardware implementations, named ICEBERG [7]. ICEBERG uses 64-bit text blocks and 128-bit keys and the round number is 16. Specially, all components are involutional and allow very efficient combinations of encryption/decryption. In practice, very low-cost

hardware crypto-processors and high throughput data encryption are potential applications of ICEBERG. In [8], Sun *et al.* gave the linear cryptanalysis for 7-round ICEBERG with the whole codebook and $2^{91.19}$ 7-round encryptions, which is the first attack for reduced-round of ICEBERG.

Differential cryptanalysis, proposed by Biham and Shamir [9], has been one of the most classic cryptanalytic techniques. Although the original ICEBERG proposal provided theoretical upper bounds of the probability for the differential characteristics of 16-round ICEBERG, the proposal did not give the concrete differential cryptanalysis. In addition, there is no public cryptanalytic results about ICEBERG. In this paper, we will give the concrete differential cryptanalysis for reduced-round of ICEBERG, specially, we will improve the differential cryptanalytic results for ICEBERG with multiple differential cryptanalysis, which has been put forward in [11].

In this paper, we make use of the property of the linear layer of ICEBER and design an efficient searching algorithm to identify the differential characteristic for ICEBERG. As a result, we found that the highest probability of 6-round differential $2^{-60.53}$ is much greater than the highest probability of the 6-round differential characteristic $2^{-63.32}$, so we give two attacks with multiple differentials, the first one is the structure attack on 7-round ICEBERG with one output difference and multiple input differences and it requires that $2^{90.28}$ times of 7-round encryptions and 2^{54} chosen plaintexts; and the second one is the multiple differential attack on 8-round ICEBERG with multiple output differences. Although we cannot threat ICEBERG, the attack on 8-round ICEBERG we give is the best attack. Furthermore, our method to identify the probability of differentials can be used for other block ciphers.

The paper is organized as follows. Section 2 presents the description for ICEBERG block cipher. In Section 3, we discuss the property of the linear layer of ICEBERG and identify the best 6-round differential characteristics and differentials. Section 4 presents the 7-round structure attack on ICEBERG and the 8-round multiple differential attack on ICEBERG, respectively. Section 5 concludes this paper.

2 Description of ICEBERG

ICEBERG is proposed by Standaert *et al.* on FSE 2004, and it is a fast involutional block cipher with SP-network structure optimized for reconfigurable hardware implementations [7]. Specially, all components are involutional and allow very efficient combinations of encryption/decryption. In practice, very low-cost hardware crypto-processors and high throughput data encryption are potential applications of ICEBERG. It operates on 64-bit block and uses a 128-bit key. The round number is 16. The round function ρ_K can be expressed as:

$$\rho_K : \mathbb{Z}_2^{64} \to \mathbb{Z}_2^{64} : \rho_K \equiv \epsilon_K \circ \gamma,$$

where γ is the non-linear layer and ϵ_K is the linear layer.

It is an involutional cipher since its encryption is only different from its decryption in the key schedule. Because the key schedule has little relationship with our analysis, we will not describe it here.

2.1 Non-Linear Layer γ

Non-linear layer γ is composed of non-linear substitution layers $S0$ and $S1$ and bit permutation layer $P8$. Fig. 1 depicts the non-linear layer γ. Each substitution layer consists of 16 identical S-boxes in parallel. The bit permutation layer consists of eight identical bit permutations $P8$. The γ layer can be expressed as:

$$\gamma : \mathbb{Z}_2^{64} \rightarrow \mathbb{Z}_2^{64} : \gamma \equiv S0 \circ P8 \circ S1 \circ P8 \circ S0.$$

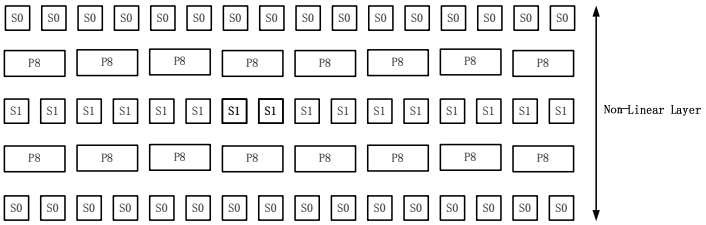

Fig. 1. The Non-Linear Layer γ

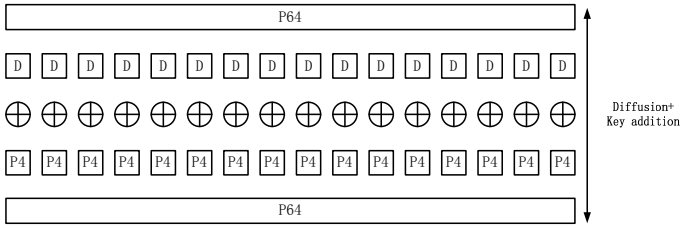

Fig. 2. The Linear Layer ϵ_K

The γ layer can be viewed as one layer consisting of the application of eight identical 8×8 S-boxes listed in Table 1.

2.2 Linear Layer ϵ_K

Fig. 2 depicts the linear layer ϵ_K. The ϵ_K can be described as:

$$\epsilon_K : \mathbb{Z}_2^{64} \rightarrow \mathbb{Z}_2^{64} : \epsilon_K \equiv P64 \circ P4 \circ \sigma_K \circ M \circ P64.$$

Table 1. The 8×8 S-box

	00	01	02	03	04	05	06	07	08	09	0a	0b	0c	0d	0e	0f
00	24	c1	38	30	e7	57	df	20	3e	99	1a	34	ca	d6	52	fd
10	40	6c	d3	3d	4a	59	f8	77	fb	61	0a	56	b9	d2	fc	f1
20	07	f5	93	cd	00	b6	62	a7	63	fe	44	bd	5f	92	6b	68
30	03	4e	a2	97	0b	60	83	a3	02	e5	45	67	f4	13	08	8b
40	10	ce	be	b4	2a	3a	96	84	c8	9f	14	c0	c4	6f	31	d9
50	ab	ae	0e	64	7c	da	1b	05	a8	15	a5	90	94	85	71	2c
60	35	19	26	28	53	e2	7f	3b	2f	a9	cc	2e	11	76	ed	4d
70	87	5e	c2	c7	80	b0	6d	17	b2	ff	e4	b7	54	9d	b8	66
80	74	9c	db	36	47	5d	de	70	d5	91	aa	3f	c9	d8	f3	f2
90	5b	89	2d	22	5c	e1	46	33	e6	09	bc	e8	81	7d	e9	49
a0	e0	b1	32	37	ea	5a	f6	27	58	69	8a	50	ba	dd	51	f9
b0	75	a1	78	d0	43	f7	25	7b	7e	1c	ac	d4	9a	2b	42	e3
c0	4b	01	72	d7	4c	fa	eb	73	48	8c	0c	f0	6a	23	41	ec
d0	b3	ef	1d	12	bb	88	0d	c3	8d	4f	55	82	ee	ad	86	06
e0	a0	95	65	bf	7a	39	98	04	9b	9e	a4	c6	cf	6e	dc	d1
f0	cb	1f	8f	8e	3c	21	a6	b5	16	af	c5	18	1e	0f	29	79

Table 2. The $P64$ Permutation

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

It consists of the 64-bit permutation layer $P64$, the parallel binary matrix multiplications M, the key addition layer σ_K, the parallel 4-bit permutation layer and the identical 64-bit permutation as in Fig. 2. $P64$ and $P4$ are listed in Table 2 and Table 3, respectively. The matrix multiplication M is based on the parallel application of a simple involutional matrix multiplication. Let $V \in \mathbb{Z}_2^{4\times 4}$ be a binary involutional matrix (i.e. such that $V^2 = I_n$):

$$V = \begin{bmatrix} 0\,1\,1\,1 \\ 1\,0\,1\,1 \\ 1\,1\,0\,1 \\ 1\,1\,1\,0 \end{bmatrix}.$$

M is then defined as:

$$M : \mathbb{Z}_{2^4}^{16} \to \mathbb{Z}_{2^4}^{16} : x \to y = M(x) \Leftrightarrow y_i = V \cdot x_i \qquad 0 \le i \le 15.$$

Then the diffusion box D is defined as performing multiplication by V.

Table 3. The $P4$ Permutation

0	1	2	3
1	0	3	2

The encryption process for R rounds is defined as follows:

$$\sigma_{RK_0^R} \circ \gamma \circ (\bigcirc_{r=1}^{R-1} \rho_{RK_1^r}) \circ \sigma_{RK_1^0}.$$

where σ_K is the key addition layer.

3 Differential Distinguishers of 6-round ICEBERG

Differential cryptanalysis, introduced by Biham and Shamir in [9], is one of the most popular and important attack towards block ciphers especially iterated ones. It uses the differential characteristic with high probability for inner rounds as a distinguisher to recover the subkey bits in fore or last few rounds. Then several works proposed that the effect of differential cryptanalysis can be strengthened with differentials [10] or multiple differential [11]. In this section, we will identify the best differential characteristics of 6-round ICEBERG and the best differentials of 6-round ICEBERG.

3.1 Differential Characteristic of 6-round ICEBERG

The way to search the best differential characteristic of an iterated SPN block cipher depends on two components. The differential distribution table of the active S-box in the non-linear layer determines the probability of the differential characteristic while the linear layer determines the least number of active S-boxes. So the higher probability of the active S-box and the fewer active S-boxes are there in one round, the better is the differential characteristic. According to the differential distribution table of 8×8 S-box for ICEBERG, we found that the probability ranges from 2^{-7} to 2^{-5}. The gap between the best and the worst differential characteristic is so small that the number of active S-boxes in the differential characteristic is the determinant factor for its probability. Therefore, we will aim at finding the best one among differential characteristics with as few active S-boxes as possible which is determined by the linear layer of ICEBERG.

Property of Linear Layer $P64$-$DP4$-$P64$

To achieve hardware efficiency, the linear layer of light-weight block cipher is always designed to be a permutation which can be implemented by wire-crossing, sometimes to be involutional. Unlike other light-weight block ciphers, the permutation $P64$ mapping on 64 bit is emplaced twice, respectively at the beginning and end of the linear layer of ICEBERG, and there are 16 same

permutations $P4$ on each 4 bit, key bitwise addition and 16 same diffusion boxes D inside the linear layer. The diffusion box D, which makes each output bit equal to the exclusive-or among the three input bits, results in the full diffusion can be obtained after two rounds. The diffusion pattern of differential characteristic for 2-round is shown in Fig. 3. In view of differential cryptanalysis, $P4$ and D can be regarded as a whole, named $DP4$ depicted in Table 4. Now we give the following three properties of the linear layer of ICEBERG,

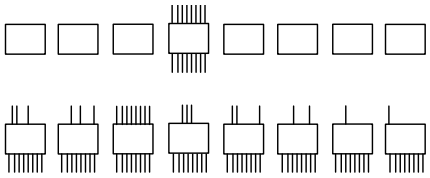

Fig. 3. Diffusion of Two Rounds ICEBERG

Table 4. $DP4$ Linear-Layer

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	13	14	3	7	10	9	4	11	6	5	8	12	1	2	15

Property 1. According to Table 2, the four input bits of each $DP4$ must be output from four different S-boxes; moreover, as $P64$ is involutional, the four output bits must be input to four different S-boxes.

Property 2. According to Table 2, each two output bits from each 8×8 S-box must input to different bytes after $P64$, and input to different nibbles of $DP4$.

Property 3. One nonzero input/output bit for $DP4$ will produce three nonzero output/input bits, or two nonzero input/output bits for $DP4$ will produce two nonzero output/input bits.

Because of the above three properties, the number of active S-boxes for two rounds is at least four. So there are three primary patterns ($1 \rightarrow 3$, $2 \rightarrow 2$, $3 \rightarrow 1$) and two auxiliary patterns ($2 \rightarrow 3$, $3 \rightarrow 2$), where $a \rightarrow b$ means that there are a nonzero input difference bits and b nonzero output difference bits. For each pattern, we take the diffusion box $DP4$ as the start point for our searching algorithm. Next, we will analyze these patterns.

- *Pattern* ($1 \rightarrow 3$): It implies there should be m ($1 \leqslant m \leqslant 8$) active $DP4(s)$ with $1 \rightarrow 3$ (1 nonzero input difference bit to 3 nonzero output difference bits). After the bottom $P64$, the $3m$ nonzero output difference bits of $DP4$s

will input to three S-boxes, each of which has m nonzero input difference bits. Whilst the m non-zoro input bits will input to the same S-box before the top $P64$ in reverse order. To sustain this condition, we need to search m-$DP4$s whose three output bits will be located in the same three bytes after $P64$. Meanwhile, deduced by the reversed $P64$, the m input bits should be just right located in the same one byte. The pattern is depicted in Fig. 4(a).

To efficiently search all the possible combinations of m-$DP4$, we divide them into several sets by m and look into the relation of these sets during their generation.

Proposition 1. *Assuming Γ_m is the set of all m-$DP4$ possible combinations. If $(\alpha_0, \alpha_1 \dots \alpha_m) \in \Gamma_{m+1} (m \geqslant 1)$, then there should be $(\alpha_0, \alpha_1 \dots \alpha_{m-1}) \in \Gamma_m$ and $(\alpha_0, \alpha_1 \dots \alpha_{m-2}, \alpha_m) \in \Gamma_m$. So Γ_{m+1} can be generated from Γ_m. Vice versa, if $(\alpha_0, \alpha_1 \dots \alpha_{m-1}) \in \Gamma_m$ and $(\alpha_0, \alpha_1 \dots \alpha_{m-2}, \beta_{m-1}) \in \Gamma_m$, then $(\alpha_0, \alpha_1 \dots \alpha_{m-1}, \beta_{m-1}) \in \Gamma_{m+1}$.*

We will produce Γ_i from *Proposition* 1. Firstly, in order to generate Γ_1, we searched the combinations for 1-$DP4$s with only one active $DP4$ locating in 16 kinds of possible positions. Then we used the elements from Γ_1 to produce Γ_2, and then we produce Γ_3 from Γ_2. In the similar way, we will stop the process until there is no elements in some Γ_m (As the data in Table 5 showing, the process stops at Γ_4). The sets of possible combinations for $Pattern(1 \to 3)$ are shown in the first row in Table 5. As we can see, $1 \leqslant m \leqslant 3$, and in each sub row we give an example in each Γ_m. For example, the first sub row means that there is a possible combination for Γ_1, in which the output difference of the active S-box S_0 in top round (we will call this as top output difference for short in following sections) is 1_x, the input difference and output difference on 0-$DP4$ are 1_x and d_x, and the input difference of the active S-boxes in the next round (for short we call bottom input difference) S_0 is 1_x, S_2 is 80_x, S_3 is 2_x.

☐ *Pattern* $(2 \to 2)$: It implies there should be $m (1 \leqslant m \leqslant 8)$ active $DP4(s)$ with $2 \to 2$ (2 nonzero input difference bits to 2 nonzero output difference bits). After bottom $P64$, the $2m$ active output bits of $DP4$s should be input to two active S-boxes, each of which has m nonzero input difference bits. At the same time the m nonzero input difference bits should input to two active S-boxes after passing the top $P64$ in reverse order. To satisfy this condition, we need to search m-$DP4$s whose two output bits are located in the same two bytes after $P64$. Meanwhile, deduced by the reversed $P64$, the $2m$ input bits should be just right located in the same two bytes. The pattern is depicted in Fig. 4(b).
In the similar way as *Pattern* $(1 \to 3)$, we can obtain $\Gamma_1, \Gamma_2 \dots \Gamma_{i-1}$ iteratively until there is no element in some Γ_m (As the data in Table 5 showing, the process terminates at Γ_6).

☐ $Pattern(3 \to 1)$: Since $DP4(x) = DP4^{-1}(x)$, $P64(x) = P64^{-1}(x)$, $Pattern(3 \to 1)$ is $Pattern(1 \to 3)$ in reverse order. It is depicted in Fig. 4(c).

☐ *Pattern* $(2 \to 3)$: Similar as the above analysis, we can finally get $2 \leqslant m \leqslant 6$. It is depicted in Fig. 4(d).

☐ *Pattern* $(3 \to 2)$: It is the reversal of $Pattern(2 \to 3)$, depicted in Fig. 4(e).

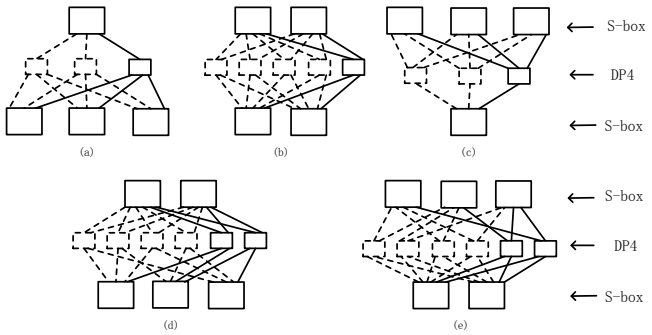

Fig. 4. Patterns for Linear Layer $P64$-$DP4$-$P64$

Table 5 gives $|\Gamma_m|$ and one example in Γ_m for each pattern. It should be noticed that the pattern is not an intact two-round differential characteristic, it begins with the output differences of the active S-boxes in the first round and ends at the input differences of the active S-boxes in the second round, as shown in Fig. 4. We name this semi-joint 2-round differential characteristic as a **node** for short in following sections which involves two members: top output difference and bottom input difference.

Search the 6-round Differential Characteristics

According to the above five patterns, eight patterns for 6-round differential characteristics can be produced in Fig. 5.

Because $S(x) = S^{-1}(x)$, $P64(x) = P64^{-1}(x)$ and $DP4(x) = DP4^{-1}(x)$, each differential characteristic in Fig. 5(a) will generate a reversed differential characteristic in Fig. 5(b) by turning its input and output differences upside down in each round. The same scenario is for (Fig. 5(c), Fig. 5(d)) and (Fig. 5(e), Fig. 5(f)). In this way, in order to search the differential characteristic with highest probability among the 5 patterns, firstly, we can concatenate 2-round nodes in the five patterns depicted in Table 5 to the 6-round nodes. Two 2-round nodes can be linked with each other only if they are on same series of active S-boxes and the joint entries whose bottom input difference is from the above node and the top output difference comes from the node below have nonzero probability. For example, the bottom input difference of the 2-round node $\{(0, 5_x), (1, 11_x)\} \to \{(0, 5_x), (1, 11_x)\}$ and the top output difference of the 2-round node $\{(0, 4_x), (1, 1_x)\} \to \{(0, 4_x), (1, 1_x)\}$ are both on the

Table 5. Patterns for $DP4$

Pattern	m	$\lvert\Gamma_m\rvert$	Δ_{out} (pos, out)	Δ_{in} (pos, in)	$DP4$ (pos, in, out)
$1 \to 3$	1	64	$(0, 1_x)$	$(0, 1_x), (2, 80_x), (3, 2_x)$	$(0, 1_x, d_x)$
	2	13	$(5, 84_x)$	$(5, 84_x), (6, 28_x), (7, 28_x)$	$(1, 2_x, e_x), (3, 2_x, e_x)$
	3	1	$(5, c4_x)$	$(5, c4_x), (6, 29_x), (7, 68_x)$	$(1, 2_x, e_x), (3, 2_x, e_x), (7, 2_x, e_x)$
$2 \to 2$	1	96	$(0, 1_x), (1, 10_x)$	$(0, 1_x), (1, 10_x)$	$(0, 3_x, 3_x)$
	2	27	$(0, 5_x), (1, 11_x)$	$(0, 5_x), (1, 11_x)$	$(0, 3_x, 3_x), (5, c_x, c_x)$
	3	12	$(6, a8_x), (7, 2c_x)$	$(6, a8_x), (7, 2c_x)$	$(1, c_x, c_x), (3, c_x, c_x), (5, 3_x, 3_x)$
	4	5	$(6, a9_x), (7, 6c_x)$	$(6, a9_x), (7, 6c_x)$	$(1, c_x, c_x), (3, c_x, c_x),$ $(5, 3_x, 3_x), (7, c_x, c_x)$
	5	1	$(6, ab_x), (7, 7c_x)$	$(6, ab_x), (7, 7c_x)$	$(1, c_x, c_x), (3, c_x, c_x),$ $(5, 3_x, 3_x), (7, c_x, c_x), (8, c_x, c_x)$
$3 \to 1$	1	64	$(0, 1_x), (2, 80_x), (3, 2_x)$	$(0, 1_x)$	$(0, d_x, 1_x)$
	2	13	$(5, 84_x), (6, 28_x), (7, 28_x)$	$(5, 84_x)$	$(1, e_x, 2_x), (3, e_x, 2_x)$
	3	1	$(5, c4_x), (6, 29_x), (7, 68_x)$	$(5, c4_x)$	$(1, e_x, 2_x), (3, e_x, 2_x), (7, e_x, 2_x)$
$2 \to 3$	2	218	$(1, 12_x), (3, 6_x)$	$(0, 1_x), (2, c0_x), (4, 1_x)$	$(0, a_x, 5_x), (2, 6_x, 9_x)$
	3	119	$(2, c4_x), (3, 2_x)$	$(2, c4_x), (3, 7_x), (4, 21_x)$	$(0, c_x, c_x), (2, 1_x, d_x), (4, 4_x, 7_x)$
	4	51	$(1, 1c_x), (3, 1c_x)$	$(1, 1e_x), (2, c0_x), (3, 1e_x)$	$(0, 2_x, e_x), (2, 4_x, 7_x)$ $(6, c_x, c_x), (8, 3_x, 3_x)$
	5	17	$(1, 3c_x), (3, 3c_x)$	$(1, 3e_x), (2, c0_x), (3, 3e_x)$	$(0, 2_x, e_x), (2, 4_x, 7_x)$ $(6, c_x, c_x), (8, 3_x, 3_x), (11, c_x, c_x)$
	6	3	$(6, eb_x), (7, 7c_x)$	$(0, 40_x), (6, eb_x), (7, 7e_x)$	$(1, c_x, c_x), (3, c_x, c_x), (5, 3_x, 3_x)$ $(7, c_x, c_x), (8, c_x, c_x), (13, 4_x, 7_x)$
$3 \to 2$	2	218	$(0, 1_x), (2, c0_x), (4, 1_x)$	$(1, 12_x), (3, 6_x)$	$(0, 5_x, a_x), (2, 9_x, 6_x)$
	3	119	$(2, c4_x), (3, 7_x), (4, 21_x)$	$(2, c4_x), (3, 2_x)$	$(0, c_x, c_x), (2, d_x, 1_x), (4, 7_x, 4_x)$
	4	51	$(1, 1e_x), (2, c0_x), (3, 1e_x)$	$(1, 1c_x), (3, 1c_x)$	$(0, e_x, 2_x), (2, 7_x, 4_x)$ $(6, c_x, c_x), (8, 3_x, 3_x)$
	5	17	$(1, 3e_x), (2, c0_x), (3, 3e_x)$	$(1, 3c_x), (3, 3c_x)$	$(0, e_x, 2_x), (2, 7_x, 4_x)$ $(6, c_x, c_x), (8, 3_x, 3_x), (11, c_x, c_x)$
	6	3	$(0, 40_x), (6, eb_x), (7, 7e_x)$	$(6, eb_x), (7, 7c_x)$	$(1, c_x, c_x), (3, c_x, c_x), (5, 3_x, 3_x)$ $(7, c_x, c_x), (8, c_x, c_x), (13, 7_x, 4_x)$

Δ_{out} means the output difference of the active S-box in this round;
Δ_{in} means the input difference of the active S-box in the next round;
The tuple (a, b) means a is the index of S-box, b is the output difference in Δ_{out} column or the input difference in Δ_{in} column;
The triple (a, b, c) means on the a-th $DP4$, the input difference is b and the output difference is c.

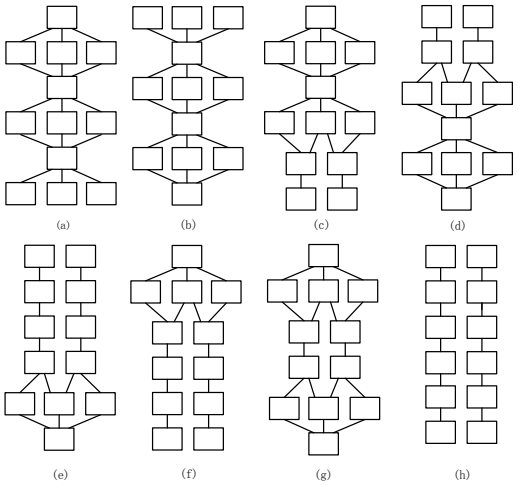

Fig. 5. 6-round Best Differential Characteristic Patterns

active S-boxes S_0 and S_1, and from the difference distribution of the S-box, $P_r\{5_x \rightarrow 4_x\} = 2^{-7}, P_r\{11_x \rightarrow 1_x\} = 2^{-5.42}$, so they can be concatenated to 3-round node $\{(0, 5_x), (1, 11_x)\} \rightarrow \{(0, 4_x), (1, 1_x)\}$ with probability $2^{-12.42}$. By concatenating 2-round nodes one by one or iterative concatenation (two 2-round nodes can concatenate to a 3-round node, then two 3-round nodes can concatenate to a 5-round node, go along with concatenating another 2-round node at end), we will get a 6-round node.

Table 6 gives examples for the best 6-round node for the eight patterns in Fig. 5. From the last column of Table 6, we can see the highest probability is $2^{-46.84}$. From column Δ_{out} and Δ_{in}, we see that there are 4 active S-boxes in the first and the last round. According to the differential distribution table, the probability of one entry is at most 2^{-5}, so the probability of the best differential characteristic for the first node pattern depicted in Table 6 is at least $2^{-66.84}$ less than 2^{-64} which has little contribution for our distinguisher. So we only identify the 6-round differential characteristics with the probability greater than 2^{-64}. By computing all the eight node patterns, the 6-round differential characteristics with the probability greater than 2^{-64} have been identified from the last node pattern $2 \rightarrow 2 \rightarrow 2 \rightarrow 2 \rightarrow 2 \rightarrow 2$.

By further analysis on the 6-round nodes $\{(6, ab_x), (7, 7c_x) \rightarrow (6, ab_x), (7, 7c_x)\}$ and $\{(6, \mathtt{b_x}), (7, 70_x) \rightarrow (6, \mathtt{b_x}), (7, 70_x)\}$, we obtain four best 6-round differential characteristics with probability $2^{-63.32}$ as follows,

$$(93ab0000\ 00000000_x) \xrightarrow{6r} (93ab0000\ 00000000_x) \xrightarrow{LL} (93ebc446\ 2010a106_x),$$

$$(93ab0000\ 00000000_x) \xrightarrow{6r} (c7ab0000\ 00000000_x) \xrightarrow{LL} (c7eb8444\ 3010a802_x),$$

$$(c7ab0000\ 00000000_x) \xrightarrow{6r} (93ab0000\ 00000000_x) \xrightarrow{LL} (93ebc446\ 2010a106_x),$$

Table 6. Results of Eight 6-Round Nodes

Node pattern	Amount	Δ_{out}	Δ_{in}	Pr
$1 \to 3 \to 1 \to 3 \to 1 \to 3$	306	$(5, c0_x)$	$(5, c0_x), (6, 9_x), (7, 60_x)$	$2^{-46.84}$
$3 \to 1 \to 3 \to 1 \to 3 \to 1$	306	$(5, c0_x), (6, 9_x), (7, 60_x)$	$(5, c0_x)$	$2^{-46.84}$
$1 \to 3 \to 1 \to 3 \to 2 \to 2$	347	$(6, 10_x)$	$(6, 40_x), (7, 2_x)$	$2^{-59.42}$
$2 \to 2 \to 3 \to 1 \to 3 \to 1$	347	$(6, 40_x), (7, 2_x)$	$(6, 10_x)$	$2^{-59.42}$
$2 \to 2 \to 2 \to 2 \to 3 \to 1$	4606	$(6, b_x), (7, 70_x)$	$(7, 2_x)$	$2^{-56.66}$
$1 \to 3 \to 2 \to 2 \to 2 \to 2$	4606	$(7, 2_x)$	$(6, b_x), (7, 70_x)$	$2^{-56.66}$
$1 \to 3 \to 2 \to 2 \to 3 \to 1$	757	$(1, 4_x)$	$(1, 40_x)$	$2^{-63.66}$
$2 \to 2 \to 2 \to 2 \to 2 \to 2$	53846	$(6, ab_x), (7, 7c_x)$	$(6, ab_x), (7, 7c_x)$	$2^{-43.32}$
		$(6, b_x), (7, 70_x)$	$(6, b_x), (7, 70_x)$	$2^{-43.32}$

Amount: the total number of nodes.
(Δ_{out}, Δ_{in}, Pr): the node with highest probability Pr, its top output difference is Δ_{out} and bottom input difference is Δ_{in}.

$$(c7ab0000\ 00000000_x) \xrightarrow{6r} (c7ab0000\ 00000000_x) \xrightarrow{LL} (c7eb8444\ 3010a802_x).$$

where "$6r$" stands for the 6-round node with input difference of the active S-boxes in the first round and output difference of the active S-boxes in the last round; the "LL" stands for the linear layer in the last round. Due to the limited space, we only list the details for the last 6-round differential characteristic in Table 7.

3.2 Differentials of 6-Round ICEBERG

From Table 6, there are 53846 6-round nodes in $2 \to 2 \to 2 \to 2 \to 2 \to 2$ pattern. Not only the active S-Boxes for the input difference and the output difference of the above four best 6-round nodes with highest probability are located in (S_6, S_7), but also other 52384 nodes are located in them. So we traverse 2^{32} different input differences and output differences to compute the probability of each differential on the given input and output difference value on the four fixed active S-boxes. As a result, we identify some differentials with higher probability in Table 8. The first column is the differential; the second column shows how many 6-round nodes contributing to the differential among all the 52384 6-round nodes; the third column is the corresponding probability of the differential.

3.3 Structures of Differentials of 6-Round ICEBERG

Now we consider the effect of a set of differentials whose active S-boxes locate in S_6 and S_7. Firstly, we collected differentials with different input differences and

Table 7. 6-Round Differential Characteristic

Round		Output Difference	Probability P_r
		S_6=ab_x, S_7=$c7_x$	
R_1	S-box	S_6=ab_x, S_7=$7c_x$	2^{-10}
R_1	LT	S_6=ab_x, S_7=$7c_x$	1
R_2	S-box	S_6=b_x, S_7=70_x	$2^{-10.83}$
R_2	LT	S_6=b_x, S_7=70_x	1
R_3	S-box	S_6=ab_x, S_7=$7c_x$	$2^{-10.83}$
R_3	LT	S_6=ab_x, S_7=$7c_x$	1
R_4	S-box	S_6=b_x, S_7=70_x	$2^{-10.83}$
R_4	LT	S_6=b_x, S_7=70_x	1
R_5	S-box	S_6=ab_x, S_7=$7c_x$	$2^{-10.83}$
R_5	LT	S_6=ab_x, S_7=$7c_x$	1
R_6	S-box	S_6=ab_x, S_7=$c7_x$	2^{-10}
R_6	LT	S_0=2_x, S_1=$a8_x$, S_2=10_x, S_3=30_x, S_4=44_x, S_5=84_x, S_6=eb_x, S_7=$c7_x$	1

Table 8. Differentials of 6-Round ICEBERG

Differential	Amount/52384	Pr
$(7c0b0000\ 00000000_x) \rightarrow (7c0b0440\ 00000104_x)$	1338	$2^{-60.53}$
$(7cab0000\ 00000000_x) \rightarrow (7c0b0440\ 00000104_x)$	2046	$2^{-60.59}$
$(7c0b0000\ 00000000_x) \rightarrow (7cab0000\ 00000000_x)$	2046	$2^{-60.59}$
$(c7ab0000\ 00000000_x) \rightarrow (7c0b0440\ 00000104_x)$	2242	$2^{-60.59}$
$(7c0b0000\ 00000000_x) \rightarrow (c7eb8444\ 3010a802_x)$	2242	$2^{-60.59}$

fixed given output difference on active S-boxes S_6 and S_7 whose probabilities are greater than the average probability 2^{-64} as a set. Then we compared all the 2^{16} sets with different output differences to obtain the set of differentials with highest probability. Some of them with higher probability are shown in Table 9. Because the number of input differences is large, we have not listed the input differences in Table 9. The first column is the sets of the differentials with fixed output differences; the second column shows the amount of various of differentials' input differences; the third column is the sum of probabilities of all the differentials in the set.

4 Attacks against Reduced-Round ICEBERG

4.1 Structure Attack to 7-Round ICEBERG

In this section, we will use the set with 10981 6-round differentials in the first row in Table 9 to proceed the structure attack on 7-round ICEBERG. Firstly,

Table 9. Sets of Differentials of 6-Round ICEBERG

Set of Differentials	Amount	Pr
(xxxx0000 00000000_x) $\rightarrow$ ($7c0b0440$ 00000104_x)	10981	$2^{-49.77}$
(xxxx0000 00000000_x) $\rightarrow$ ($c7eb8444$ $3010a802_x$)	10784	$2^{-49.82}$
(xxxx0000 00000000_x) $\rightarrow$ ($7cab0000$ 00000000_x)	10501	$2^{-49.86}$
(xxxx0000 00000000_x) $\rightarrow$ ($93ebc446$ $2010a106_x$)	10473	$2^{-49.87}$

Amount: the number of different input differences.

we construct 2^{41} structures, in each structure, the 16-bit plaintext input to S-boxes S_6 and S_7 traverses all possible values and other plaintext bits will take the fixed value. So our attack will use $2^{41} \cdot 2^{16} = 2^{57}$ chosen plaintexts. The sum of the probability for all differentials is $2^{-49.77}$. In each structure, there are $2^{15} \cdot 10981 \approx 2^{28.42}$ plaintext pairs satisfying any of the 10981 input differences. So the expected number of right pairs is $2^{41} \cdot 2^{15} \cdot 2^{-49.77} = 2^{6.23} \approx 75.06$. According to the input difference in round 7 ($7c0b0440$ 00000104_x), we know there are six active S-boxes, which results in 48 subkey bits to be guessed. Considering all the possible output differences of the six active S-boxes and the two non-active S-boxes, the filtering probability β and the average number of subkey values counted per pair α should be computed. We list the number of the output differences and the number of subkey values counted according to the different cases in Table 10, which can be obtained from the differential distribution tables of S-boxes. As we see in each row of Table 10, the number of possible subkey values for the given input difference of the active S-box varies on 2, 4, 6 and 8. In column 3 of Table 10, we classify the number of the output differences by the number of candidate subkey values. Column 4 is the average number of subkeys suggested by the given input difference. As a result, the filtering ratio $\beta = (\frac{94}{256})^2 \cdot (\frac{99}{256})^3 \cdot \frac{102}{256} \cdot 2^{-16} \approx 2^{-24.33}$, the average increment on per counter for wrong pairs will be $2^{41} \cdot 2^{15} \cdot 10981 \cdot 2^{-24.33} \cdot 2^{8.34} \cdot 2^{-48} \approx 2^{5.43} \approx 43.11$.

Table 10. Number of Candidate Subkey Values

Input Difference	Number of Output Differences	Number of Candidate Subkeys 2	4	6	8	Average Number
1_x	94	65	25	3	1	$2^{1.45}$
4_x	99	75	20	3	1	$2^{1.37}$
40_x	102	80	18	4	0	$2^{1.33}$
b_x	94	70	15	8	1	$2^{1.45}$
$7c_x$	99	76	19	2	2	$2^{1.37}$
ab_x	99	76	18	4	1	$2^{1.37}$
70_x	105	84	19	2	0	$2^{1.29}$

Next, we will give the attacking procedure as follows:

- For each structure:
  - a: Insert all the ciphertexts into the hash table according to the 16-bit ciphertext bits of the non-active S-boxes in the last round.
  - b: For each entry with collision (a pair of ciphertext with equal 16-bit values) check whether the plaintexts difference (in round 1) is one of the 10981 differentials's input differences.
  - c: Filter wrong pairs which satisfy none of 48-bit of the ciphertext differences on the six active S-boxes of the 10981 differentials.
  - d: For each possible subkey in round 7, decrypt the last round to obtain the output difference of round 6, and check whether the difference equals to the output difference of the differentials. If a pair passes the above test, add one to the counter related to the subkey value. The average increment on per counter for wrong pairs will be 43.11.
- 2: Collect all the subkeys whose counter has at least 75 hits. With the high probability the correct subkey is in this list.
- 3: Exhaustively search the remaining 80-bit subkey key and we can obtain the whole 128-bit master key.

In step (a), the time complexity is 2^{16} memory accesses. In step (b), about 2^{15} pairs remain through the filter of step (a), so the time complexity is 2^{16} memory accesses. So for all structures, the two steps require 2^{58} memory accesses. For all structures, the time complexity of step (c) is negligible, and the time complexity of step (d) is about $2^{41} \cdot 2^{15} \cdot 10981 \cdot 2^{-24.33} \cdot 2^{48} \approx 2^{93.09}$ one-round decryptions, which equals to $2^{90.28}$ 7-round encryptions.

The signal to noise ratio is:

$$S/N = \frac{2^{-49.77} \cdot 2^{48}}{10981 \cdot 2^{-24.33} \cdot 2^{8.34}} \approx 1.74.$$

The success rate is computed with the method in [12] as follows,

$$\begin{aligned} Ps &= \Phi\left(\frac{\sqrt{\mu S_N} - \Phi^{-1}(1-2^{-a})}{\sqrt{S_N+1}}\right) \\ &= \Phi\left(\frac{\sqrt{75.06 \cdot 1.74} - \Phi^{-1}(1-2^{-48})}{\sqrt{1.74+1}}\right) = 98.65\%, \end{aligned}$$

To recover the 48 subkey bits, the time complexity is about $2^{90.28}$ 7-round encryptions. The remaining 80-bit key can be exhaustively searched within 2^{80} 7-round encryptions.

In all, the data complexity is 2^{57} chosen plaintexts, and the time complexity is $2^{90.28}$ 7-round encryptions. The memory requirements are 2^{48} counters. The success rate is 98.65%.

4.2 Multiple Differential Attack against 8-Round ICEBERG

As we see in Table 7 and Table 8, the best differential characteristic or differential for 6-round has at least six active S-boxes for the output difference because

of the fast diffusion of the linear layer of ICEBERG in the last round, which make it infeasible to proceed the 8-round attack. For example, the best 6-round differential $(7c0b0000\ 00000000_x) \rightarrow (7c0b0440\ 00000104_x)$ will result in six active S-boxes in the following round. But it should be noticed that the difference before the linear layer of the last round is $(7c0b0000\ 00000000_x)$, with only two active bytes.

If the number of active bytes passing through the linear layer remains two, it will be helpful to extend more rounds for the differential. So the differentials of 6-round ICEBERG which can be used to proceed the filtering efficiently in round 8 should have two properties: higher probability and the two active bytes before linear layer in the last round should confirm to $Pattern(2 \rightarrow 2)$. Recall that all of the 6-round differentials of ICEBERG whose probabilities are higher than 2^{-64} confirm to $2 \rightarrow 2 \rightarrow 2 \rightarrow 2 \rightarrow 2 \rightarrow 2$. At last, we searched in this pattern and found 28 differentials satisfying the above properties, whose input difference or output difference doesn't change after passing the linear layer. We list the best four differentials in Table 11. The first column is the differential; the second column shows how many 6-round nodes contributing to the differential among all the 52384 6-round nodes; the third column is the corresponding probability of the differential.

Table 11. Differentials of 6-Round ICEBERG Whose Output Differences Confirm to $Pattern(2 \rightarrow 2)$

Differential	Amount/52384	Pr
$(7cab0000\ 00000000_x) \rightarrow (7cab0000\ 00000000_x)$	3382	$2^{-60.64}$
$(700b0000\ 00000000_x) \rightarrow (7cab0000\ 00000000_x)$	2761	$2^{-60.96}$
$(7cab0000\ 00000000_x) \rightarrow (700b0000\ 00000000_x)$	2761	$2^{-60.96}$
$(700b0000\ 00000000_x) \rightarrow (700b0000\ 00000000_x)$	2817	$2^{-61.26}$

We use the four differentials from round 2 to round 7 to recover total 32 subkey bits in round 0 and round 8, so the input difference of the 8-th round will be $(7cab0000\ 00000000_x)$ or $(700b0000\ 00000000_x)$. According to Table 10, we can see the size of the set of possible output differences of round 8 for $(7cab0000\ 00000000_x)$ will be $99 \cdot 99 \approx 2^{13.26}$, for $(700b0000\ 00000000_x)$ it will be $94 \cdot 105 \approx 2^{13.27}$. Because there are totally 42 shared output differences for (ab_x, b_x) and 40 ones for $(7c_x, 70_x)$, the size of the set of possible output differences in round 8 for the input differences set $\{(7cab0000\ 00000000_x)$, $(700b0000\ 00000000_x)\}$ is $99 \cdot 99 + 94 \cdot 105 - 42 \cdot 40 = 17991 \approx 2^{14.13}$. So we construct 2^{47} structures different on bits from 0 to 46. Since the linear layer is involutional, the set of input differences in round 1 is the same as the one in round 8, the size of the set of chosen plaintexts differences should be $2^{14.13}$. In each structure, there are $2^{15+14.13} = 2^{29.13}$ pairs.

The right pairs for each differential can be computed respectively. We take $(7cab0000\ 00000000_x) \rightarrow (7cab0000\ 00000000_x)$ as an example. In each structure,

the expected number of pairs with input differences $\Delta S_6 = \Delta x_i$ and $\Delta S_7 = \Delta y_i$ in the first round should be $2^{15} \cdot p_i \cdot q_i$, if we assume $p_i = P_r\{\Delta x_i \rightarrow ab_x\}$ and $q_i = P_r\{\Delta y_i \rightarrow 7c_x\}$. Since all the input differences are considered, $\sum_i p_i \cdot q_i = 1$. So the expected number of right pairs for $(7cab0000\ 00000000_x) \rightarrow (7cab0000\ 00000000_x)$ should be $2^{47} \cdot 2^{15} \cdot 2^{-60.64} = 2^{1.36}$. So the total expected number of right pairs for the four differentials should be $2^{47} \cdot 2^{15} \cdot (2^{-60.64} + 2 \cdot 2^{-60.96} + 2^{-61.26}) \approx 2^{3.06} \approx 8.36$.

Since there are six non-active S-boxes in round 8, the filter ratio is $2^{-48} \cdot 2^{14.13} \cdot 2^{-16} = 2^{-49.87}$, there will be $2^{47} \cdot 2^{15+14.13} \cdot 2^{-49.87} = 2^{26.26}$ pairs remained after the ciphertext differences filter. Considering the 42 shared output differences for (ab_x, b_x) and 40 ones for $(7c_x, 70_x)$, the average number of counted subkey values for each pair is $\frac{23472+16800+17472+15264+14000+21600+21840}{17991} \approx 7.25$, which is computed out by a tiny programme we wrote. So the expected increment on each counter for wrong subkey values will be $2^{26.26} \cdot (7.25)^2 \cdot 2^{-32} \approx 2^{-0.02} \approx 0.98$, while the value of the counter corresponding to the right key is at least 8.36.

The signal noise $S/N = 8.36/0.98 = 8.51$. The success rate is computed as follows,

$$\begin{aligned} Ps &= \Phi\left(\frac{\sqrt{\mu S_N} - \Phi^{-1}(1-2^{-a})}{\sqrt{S_N+1}}\right) \\ &= \Phi\left(\frac{\sqrt{8.36 \cdot 8.51} - \Phi^{-1}(1-2^{-32})}{\sqrt{8.51+1}}\right) = 76.10\%. \end{aligned}$$

After recover the 32-bit subkey, we can exhaustively search the remaining 96-bit subkey to get 128-bit master key. In this attack, the data complexity is 2^{63} chosen plaintexts and the time complexity is about 2^{96} times of 8-round encryptions. The memory requirements are 2^{32} counters.

5 Summary

As a block cipher for reconfigurable hardware implementations, ICEBERG is a SP-network structure involutional block cipher, so the property with very low-cost hardware crypto-processors and high throughput data encryption will result in the potential applications of ICEBERG. In this paper, we elaborately analyze the property of the linear layer of ICEBERG and design an efficient searching algorithm to identify the 6-round differential characteristics. Then we present the first differential analysis of 7-round ICEBERG using structure attack. Our attack requires $2^{90.28}$ 7-round encryptions and 2^{54} chosen plaintexts. Then we give multiple differential attack against 8-round ICEBERG, which requires 2^{63} chosen plaintexts and 2^{96} 8-round encryptions. We have improved the previous linear cryptanalysis for 7-round ICEBERG and it shows that ICEBERG can resist linear cryptanalysis more than differential cryptanalysis.

References

1. 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. In CHES 2007, volume 4727 of LNCS, pp. 450-466. Springer, 2007.
2. C. Lim and T. Korkishko. mCrypton - A Lightweight Block Cipher for Security of Lowcost RFID Tags and Sensors. In T. Kwon, J. Song and M. Yung, editors, Workshop on Information Security Applications - WISA'05, volume 3786 of LNCS, pp. 243-258. Springer, 2005.
3. D. Hong, J. Sung, S. Hong, J. Lim, S. Lee, B. S Koo, C. Lee, D. Chang, J. Lee, K. Jeong, H. Kim, J. Kim and S. Chee. HIGHT: A New Block Cipher Suitable for Low-Resource Device. In L. Goubin and M. Matsui, editors, In CHES 2006, volume 4249 of LNCS, pp. 46-59. Springer, 2006.
4. F. Standaert, G. Piret, N. Gershenfeld and J. Quisquater. SEA: a Scalable Encryption Algorithm for Small Embedded Applications. In J. Domingo-Ferrer, J. Posegga and D. Schreckling, editors, Smart Card Research and Applications, In CARDIS 2006, volume 3928 of LNCS, pp. 222-236. Springer, 2006.
5. C. De Canniere, O. Dunkelman and M. Knezevic. Katan and Ktantan-a Family of Small and Efficient Hardware Oriented Block Ciphers. In CHES 2009, volume 5747 of LNCS, pp. 272-288. Springer, 2009.
6. J. Daemen, V. Rijmen. The Design of Rijndael: AES - The Advanced Encryption Standard. Springer, Heidelberg, 2002.
7. F. Standaert, G. Piret, G. Rouvroy, J. Quisquater, and J. Legat. ICEBERG : An Involutional Cipher Efficient for Block Encryption in Reconfigurable Hardware. In FSE 2004, volume 3017 of LNCS, pp. 279-299, Springer, 2004.
8. Y. Sun and M. Q. Wang. Linear Cryptanalysis of Reduced-Round of ICEBERG, Submitted to ISPEC 2012.
9. E. Biham and A. Shamir. Differential Cryptanalysis of DES-like Cryptosystems. Journal of Cryptology, 4(1), pp. 3–72, 1991.
10. X. Lai, J.L. Massey and S. Murphy. Markov Ciphers and Differential Cryptanalysis. Advances in Cryptology-Eurocrypt'91, LNCS 547, pp. 17–38, Springer-Verlag, 1991.
11. C. Blondeau and B. Gérard. Multiple Differential Cryptanalysis: Theory and Practice. FSE 2011, LNCS 6733, pp. 35-54, Springer-Verlag, 2011.
12. A. A. Selcuk and A. Bicak. On Probability of Success in Linear and Differential Cryptanalysis, Security in Communication Networks (SCN 2002), LNCS 2576, pp.174-185, Springer-Verlag, 2003.