Near-Collisions on the Reduced-Round Compression Functions of Skein and BLAKE

Bozhan Su, Wenling Wu, Shuang Wu, Le Dong

State Key Laboratory of Information Security,
Institute of Software, Chinese Academy of Sciences, Beijing 100190, P. R. China
{*subozhan, wwl, wushuang, dongle*}*@is.iscas.ac.cn*

Abstract. The SHA-3 competition organized by NIST [1] aims to find a new hash standard as a replacement of SHA-2. Till now, 14 submissions have been selected as the second round candidates, including Skein and BLAKE, both of which have components based on modular addition, rotation and bitwise XOR (ARX). In this paper, we propose improved near-collision attacks on the reduced-round compression functions of Skein and a variant of BLAKE. The attacks are based on linear differentials of the modular additions. The computational complexity of near-collision attacks on a 4-round compression function of BLAKE-32, 4-round and 5-round compression functions of BLAKE-64 are 2^{21}, 2^{16} and 2^{216} respectively, and the attacks on a 24-round compression functions of Skein-256, Skein-512 and Skein-1024 have a complexity of 2^{60}, 2^{230} and 2^{395} respectively.

Key words: Hash function, Near-collision, SHA-3 candidates, Skein, BLAKE

1 Introduction

Hash function, a very important component in cryptology, is a function of creating a short digest for a message of arbitrary length. The classical security requirements for such a function are preimage resistance, second-preimage resistance and collision resistance. In other words, it should be impossible to find a collision in less hash computations than birthday attack, or a (second)-preimage in less hash computations than brute force attack.

In recent years, the popular hash functions (MD4, MD5, RIPEMD, SHA-0 and SHA-1) have been seriously attacked [2–5]. As a response to advances in the cryptanalysis of hash functions, NIST launched a public competition to develop a new hash function called SHA-3. Till now, 14 submissions have been selected as the second round candidates.

Skein and BLAKE are two of the second round candidates of SHA-3. Skein uses the UBI chaining mode, while BLAKE uses HAIFA approach. Both of them are of the ARX (Addition-Rotate-XOR) type. More specifically, their design primitives use only addition, rotation and XOR.

Previous works studied the linear differential trails of non-linear operations such as boolean functions and modular additions. Linear differential trails can be constructed to find near-collisions of these hash functions [7, 9, 10, 13]. Recently, linear differential attacks have been applied to many SHA-3 candidates, such as EnRUPT, CubeHash, MD6, and BLAKE [8–10].

In this paper, we further study the linear differential techniques and propose near-collision attacks on the reduced-round compression functions of Skein and BLAKE. Our strategy to find optimal linear differential trails can be described in three steps. First, linear approximations of reduced-round compression functions of Skein and BLAKE is constructed. In this step, all the addition modulo 2^{64} components of Skein and BLAKE are approximated by bitwise XOR of the inputs. Second, a difference with low hamming weight in some intermediate state as a starting point is placed. Third, the difference above propagates in both forward and backward directions until the probability becomes too small to obtain near collisions. Table 1 summarizes our attack along with the previously known ones on the reduced-round compression functions of Skein and BLAKE.

Table 1. Comparison of results on the reduced-round compression functions of Skein and BLAKE

Target	Rounds	Time	Memory	Type	Authors
Skein-512	17	2^{24}	-	434-bit near-collision	[12]
Skein-256	24	2^{60}	-	236-bit near-collision	✓
Skein-512	24	2^{230}	-	374-bit near-collision	✓
Skein-1024	24	2^{395}	-	740-bit near-collision	✓
BLAKE-32	4	2^{56}	-	232-bit near-collision	[13]
BLAKE-32	4	2^{21}	-	152-bit near-collision	✓
BLAKE-64	4	2^{16}	-	396-bit near-collision	✓
BLAKE-64	5	2^{216}	-	306-bit near-collision	✓

The paper is organized as follows. In Section 2, we describe Skein and BLAKE hash functions. In Section 3, the linear differential technique is applied to Skein and present near-collisions for Skein's compression function with reduced-round Threefish-256, Threefish-512 and Threefish-1024. In Section 4, we apply the linear differential technique to BLAKE and obtain near-collisions for reduced-round compression functions of BLAKE. Finally, Section 5 summarizes this paper.

2 Description of Skein and BLAKE

2.1 Skein

Skein is a family of hash functions based on the tweakable block cipher Threefish, which has equal block and key size of either 256, 512, or 1,024 bits. The MMO

(Matyas-Meyer-Oseas) mode is used to construct the Skein compression function from Threefish. The format specification of the tweak and a padding scheme defines the so-called Unique Block Iteration (UBI) chaining mode. UBI is used for IV generation, message compression, and as output transformation.

Threefish consists of a number of similar rounds, which is based on three simple operations: Addition modulo 2^{64}, Rotation and XOR. The intermediate state of Threefish is organized as a number of 64-bit words. The letter Δ stands for a difference in the most significant bit (MSB), i.e., $\Delta = 0x8000000000000000$. Subkeys are derived from the cipher key K and tweak $T = (t_0, t_1)$ through a simple key schedule.

Let N_w denote the number of words in the key and the plaintext block, N_r be the number of rounds. For Threefish-256, $N_w = 4$ and $N_r = 72$. Let $v_{d,i}$ be the value of the ith word of the encryption state after d rounds. The procedure of Threefish-256 encryption is:

1. $(v_{0,0}, v_{0,1}, \cdots, v_{0,N_w-1}) := (p_0, p_1, \cdots, p_{N_w-1})$, where (p_0, p_1, p_2, p_3) is the 256-bit plaintext.

2. For each round, we have

$$e_{d,i} := \begin{cases} (v_{d,i} + k_{d/4,i}) \bmod 2^{64} & \text{if } d \bmod 4{=}0, \\ v_{d,i} & \text{otherwise.} \end{cases}$$

Where $k_{d/4,i}$ is the i-th word of the subkey added to the d-th round. For $i = 0, 1, \cdots, N_w - 1$, $d = 0, 1, \cdots, N_r - 1$.

3. Mixing and word permutations followed:

$$\begin{aligned} (f_{d,2j}, f_{d,2j+1}) :=& \text{MIX}_{d,j}(e_{d,2j}, e_{d,2j+1}), & j = 0, \cdots, N_w/2 - 1, \\ v_{d+1,i} :=& f_{d,\pi(i)}, & i = 0, \cdots, N_w - 1, \end{aligned}$$

where the MIX operation depicted in Figure 1 transforms two of these 64-bit words and is common to all Threefish variants, with $R_{d,i}$ rotation constant depending on the Threefish block size, the round index d and the position of the two 64-bit words i in the Threefish state. The permutation $\pi(.)$ and the rotation constant $R_{d,i}$ can be referred to [14].

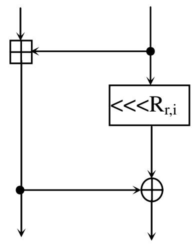

Fig. 1. The MIX function

After N_r rounds, the ciphertext $C = (c_0, c_1, \cdots, c_{N_w-1})$ is given as follows:

$$c_i := (v_{N_r,i} + k_{N_r/4,i}) \bmod 2^{64} \qquad \text{for } i = 0, 1, \cdots, N_w - 1.$$

The s-th keying ($d = 4s$) uses subkeys $k_{s,0}, \cdots, k_{s,N_w-1}$. These are derived from the key $k_0, \cdots, k_{N_w-1}$ and from the tweak t_0, t_1 as follows:

$$\begin{aligned}
k_{s,i} &:= k_{(s+i) \bmod (N_w+1)} && \text{for } i = 0, \cdots, N_w - 4 \\
k_{s,i} &:= k_{(s+i) \bmod (N_w+1)} + t_{s \bmod 3} && \text{for } i = N_w - 3 \\
k_{s,i} &:= k_{(s+i) \bmod (N_w+1)} + t_{(s+1) \bmod 3} && \text{for } i = N_w - 2 \\
k_{s,i} &:= k_{(s+i) \bmod (N_w+1)} + s && \text{for } i = N_w - 1
\end{aligned}$$

where $k_{N_w} := \lfloor 2^{64}/3 \rfloor \oplus \bigoplus_{i=0}^{N_w-1} k_i$ and $t_2 := t_0 \oplus t_1$.

2.2 BLAKE

The BLAKE family of hash functions is designed by Aumasson et al. [11] and follows HAIFA structure [6] with internal wide-pipe design strategy. Two versions of BLAKE are available: a 32-bit version (BLAKE-32) for message digests of 224 bits and 256 bits operates on 32-bit words, and a 64-bit version (BLAKE-64) for message digests of 384 bits and 512 bits operates on 64-bit words.

BLAKE operates on a large inner state v which is represented as a 4×4 matrix of words. The compression function consists of three steps: Initialization, 14 iterations of Rounds and Finalization as illustrated in Figure 2.

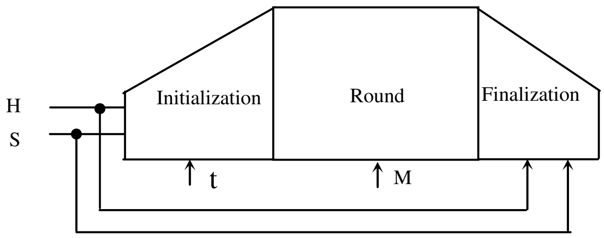

Fig. 2. Overall Structure of Compression Function of BLAKE

During the First step, the inner state v is initialized from 8 words of the chaining value $h = h_0, \cdots, h_7$, 4 words of the salt S and 2 words of block index (t_0, t_1) as follows:

$$\begin{pmatrix} v_0 & v_1 & v_2 & v_3 \\ v_4 & v_5 & v_6 & v_7 \\ v_8 & v_9 & v_{10} & v_{11} \\ v_{12} & v_{13} & v_{14} & v_{15} \end{pmatrix} \longleftarrow \begin{pmatrix} h_0 & h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 & h_7 \\ s_0 \oplus c_0 & s_1 \oplus c_1 & s_2 \oplus c_2 & s_3 \oplus c_3 \\ t_0 \oplus c_4 & t_0 \oplus c_5 & t_1 \oplus c_6 & t_1 \oplus c_7 \end{pmatrix}$$

Then, a series of 14 rounds is performed. Each round is based on the stream cipher ChaCha [15] and consists of the eight round-dependent transformations $G_0, \cdots, G_7$. Figure 3 and Figure 4 show the G function of BLAKE-32 and BLAKE-64 for index i respectively, where σ_r is a fixed permutation used in round r, M_{σ_r} are message blocks and C_{σ_r} are round-dependent constants. The $G_i(0 \leq i \leq 7)$ function takes 4 registers and 2 message words as input and outputs the updated 4 registers. A column step and diagonal step update the four columns and the four diagonals of matrix v respectively as follows:

$$G_0(v_0, v_4, v_8, v_{12}) \quad G_1(v_1, v_5, v_9, v_{13}) \quad G_2(v_2, v_6, v_{10}, v_{14}) \quad G_3(v_3, v_7, v_{11}, v_{15})$$
$$G_4(v_0, v_5, v_{10}, v_{15}) \quad G_5(v_1, v_6, v_{11}, v_{12}) \quad G_6(v_2, v_7, v_8, v_{13}) \quad G_7(v_3, v_4, v_9, v_{14})$$

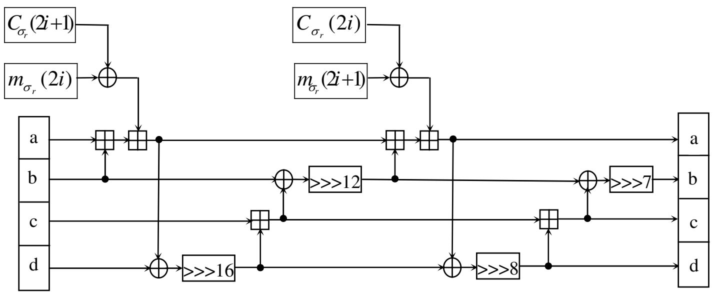

Fig. 3. The G function of BLAKE-32 for index i

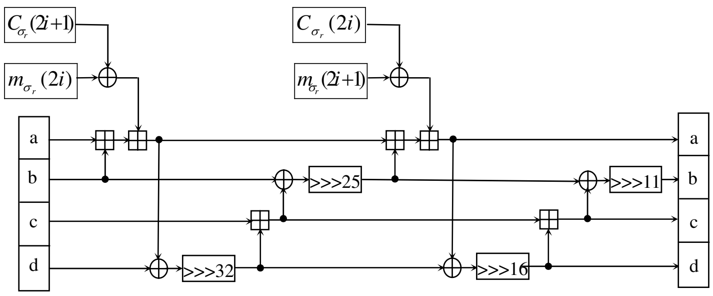

Fig. 4. The G function of BLAKE-64 for index i

In the last step, the new chaining value $h' = h'_0, \cdots, h'_7$ is computed from the internal state v and the previous chain value h (Finalization step):

$$
\begin{array}{l|l}
h'_0 \leftarrow h_0 \oplus s_0 \oplus v_0 \oplus v_8 & h'_4 \leftarrow h_4 \oplus s_4 \oplus v_4 \oplus v_{12} \\
h'_1 \leftarrow h_1 \oplus s_1 \oplus v_1 \oplus v_9 & h'_5 \leftarrow h_5 \oplus s_5 \oplus v_5 \oplus v_{13} \\
h'_2 \leftarrow h_2 \oplus s_2 \oplus v_2 \oplus v_{10} & h'_6 \leftarrow h_6 \oplus s_6 \oplus v_6 \oplus v_{14} \\
h'_3 \leftarrow h_3 \oplus s_3 \oplus v_3 \oplus v_{11} & h'_7 \leftarrow h_7 \oplus s_7 \oplus v_7 \oplus v_{15}
\end{array}
$$

3 Near-Collisions for the Reduced-Round Compression Function of Skein

Skein is based on the UBI (Unique Block Iteration) chaining mode that uses Threefish block cipher to build a compression function. The compression function outputs $E_k(t, m) \oplus m$, where E is Threefish.

Since the MIX function is the only non-linear component in the Threefish block cipher, the first step is to linearize the MIX function to obtain linear approximations of the Compression Function of Skein. To Linearize the MIX function, We replace the modular addition with XOR. The linearized MIX function is illustrated in Figure 5.

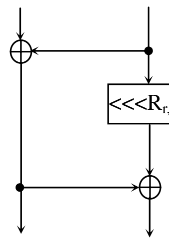

Fig. 5. linearized MIX function in Threefish

3.1 Near Collisions for the 24-Round Compression Function of Skein-256

After linearizing the Compression Function of Skein-256, we need to choose the starting point. Since Skein-256 has 72 rounds, there are $72 \approx 2^6$ possible choices. Then we place one or two bits of differences in the message blocks and certain round of the intermediate state at the starting point. Since compression function of Skein-256 uses 256-bit message and 256-bit state, there are $\binom{512}{1} + \binom{512}{2} \approx 2^{17}$ choices of positions for the one or two bits above. Therefore, the search space is less than 2^{23}, which can be searched exhaustively.

Our aim is to find one path with the highest probability in the search space. As introduced in [9], we can calculate probability of one differential trail by counting hamming weight of the differences. We search for 24-round differential trail and the results are introduced as follows.

The difference Δ in k_2, k_3, t_0 and t_1 gives a difference $(0, 0, 0, \Delta)$ at the third subkey, and $(0, 0, 0, 0)$ after the fourth. The difference in the state of round 20 is canceled out at the third subkey which is then turned into an eight-round local collision from round 21 to round 28. After 24 rounds, the hamming weight of the difference becomes too large to obtain near collisions. In the 35-th round, after adding the final subkey and feedforward value, one obtains a collision on $256 - 20 = 236$ bits. Table 2 shows the corresponding differential trail of the key and the tweak from the 12-th round to the 35-th round. Table 3 presents the corresponding trail from the 12-th round to the 35-th round. In the table, the probability for all rounds are given, except for the first round, which are indicated with M as we will use message modification techniques to make sure the first round of the trail fulfills.

Table 2. Details of the subkeys and of their differences of Skein-256, given a difference in k_2, k_3, t_0 and t_1.

Rd	d	$k_{s,0}$	$k_{s,1}$	$k_{s,2}$	$k_{s,3}$
3	12	k_3	$k_4 + t_2$	$k_0 + t_0$	k_1
		Δ	Δ	Δ	0
4	16	k_4	$k_0 + t_0$	$k_1 + t_1$	k_2
		0	Δ	0	Δ
5	20	k_0	$k_1 + t_1$	$k_2 + t_2$	k_3
		0	0	0	Δ
6	24	k_1	$k_2 + t_2$	$k_3 + t_0$	k_4
		0	0	0	0
7	28	k_2	$k_3 + t_0$	$k_4 + t_1$	k_0
		Δ	0	0	0
8	32	k_3	$k_4 + t_1$	$k_0 + t_2$	k_1
		Δ	0	Δ	0
9	36	k_4	$k_0 + t_2$	$k_1 + t_0$	k_2
		0	Δ	Δ	Δ

The message modification are applied to the most expensive part in our trail, namely the first round. Freedom degrees in chaining value and the message can be used to fulfill the first round of the trail. We use techniques introduced in [9] to derive sufficient conditions for each modular addition of the first round of the trail. Then the message block and the chaining value are chosen according to the conditions.

Table 3. Differential trail used for near collision of a 24-round compression function of Skein-256, with probability of 2^{-60}.

Rd	Difference	Pr
12	2a0344037023028a 60c217767a8a8080 ee8002206ae20266 7e23020a22014e01	-
13	c0a3442714a300aa 4ac153750aa9820a 4ea102204ac10264 10a3002a48e34c67	M
14	8a2246035a02028a 8a6217521e0a82a0 1e02020a0a620642 5e02020a02224e03	M
15	8040414144008002 004051514408802a 4000000000404041 4000000008404841	M
16	0000000000080028 8000101000080028 0010104008002800 0000000008000800	M
17	0000101000000020 0000101000000000 0010104000000000 8010104000002000	2^{-27}
18	0000000000000020 0000000000000020 0000004000002000 8000000000002000	2^{-7}
19	0000000000000000 0000000000000000 0000004000000000 8000004000000000	2^{-3}
20	0000000000000000 0000000000000000 0000000000000000 8000000000000000	2^{-1}
	no differences in round 21 - 28	1
29	0000000000000000 8000000000000000 8000000080000000 0000000000000000	1
30	0000000000000000 8000000000000000 8000000000000000 8000000080000000	2^{-1}
31	0000000000000000 8000000000000000 8000000000000000 0000000080000000	2^{-1}
32	8000000000000000 8000000000000000 8000000000000000 8000000080000000	2^{-1}
33	8000000080000000 8000000000000000 8000000000000000 8000000080000000	2^{-2}
34	0000000080002000 0000000080000000 0000800080008000 0000000080000000	2^{-2}
35	2000a00020008000 0000000000002000 0000800020002000 0000800000008000	2^{-5}
36	200008002800a000 2000a0002000a000 80008000a0008000 000000002000a000	2^{-10}

3.2 Near Collisions for the 24-Round Compression Functions of Skein-512 and Skein-1024

Ideas for near collision attacks on the reduced-round compression functions of Skein-512 and Skein-1024 are similar to the one of Skein-256. So we skip explanations here. In Table 4 and Table 5, we propose difference in the key schedule of Skein-512 and Skein-1024. The differential trails for them are illustrated in Table 6 and Table 7 in the appendix.

4 Near Collisions for the Reduced-Round Compression Function of BLAKE

4.1 Linearizing G function of BLAKE-32 and BLAKE-64

In order to linearize the G function, modular additions are replaced with XORs. Near collision attack for a 4-round compression function of BLAKE-32 in [13] also uses the linearization technique. The cyclic rotation constants in BLAKE-32 are 16,12,8,7. Notice that three of the constants 16,12 and 8 have a greatest common divisor 4, so difference $0xAAAAAAAA$ is cyclic invariant with these rotation constants, where A is a 4-bit value. In the linearized BLAKE-32, if all differences in registers are restricted to this pattern, cyclic rotations difference $>>> 16, >>> 12$ and $>>> 8$ can be removed. If zero differences pass through

Table 4. Details of the subkeys and of their differences of Skein-512, given a difference in k_4, k_5 and t_0 (leading to a differences in t_2).

Rd	d	$k_{s,0}$	$k_{s,1}$	$k_{s,2}$	$k_{s,3}$	$k_{s,4}$	$k_{s,5}$	$k_{s,6}$	$k_{s,7}$
5	20	k_5	k_6	k_7	k_8	k_0	$k_1 + t_2$	$k_2 + t_0$	$k_3 + 5$
		Δ	0	0	0	0	Δ	Δ	0
6	24	k_6	k_7	k_8	k_0	k_1	$k_2 + t_0$	$k_3 + t_1$	$k_4 + 6$
		0	0	0	0	0	Δ	0	Δ
7	28	k_7	k_8	k_0	k_1	k_2	$k_3 + t_1$	$k_4 + t_2$	$k_5 + 7$
		0	0	0	0	0	0	0	Δ
8	32	k_8	k_0	k_1	k_2	k_3	$k_4 + t_2$	$k_5 + t_0$	$k_6 + 8$
		0	0	0	0	0	0	0	0
9	36	k_0	k_1	k_2	k_3	k_4	$k_5 + t_0$	$k_6 + t_1$	$k_7 + 9$
		0	0	0	0	Δ	0	0	0
10	40	k_1	k_2	k_3	k_4	k_5	$k_6 + t_1$	$k_7 + t_2$	$k_8 + 10$
		0	0	0	Δ	Δ	0	Δ	0

Table 5. Details of the subkeys and of their differences of Skein-1024, given a difference in k_0, k_2 and t_1 (leading to a differences in t_2).

Rd	d	$k_{s,0}$	$k_{s,1}$	$k_{s,2}$	$k_{s,3}$	$k_{s,4}$	$k_{s,5}$	$k_{s,6}$	$k_{s,7}$	$k_{s,8}$	$k_{s,9}$	$k_{s,10}$	$k_{s,11}$	$k_{s,12}$	$k_{s,13}$	$k_{s,14}$	$k_{s,15}$
0	0	k_0	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	k_9	k_{10}	k_{11}	k_{12}	$k_{13} + t_0$	$k_{14} + t_1$	k_{15}
		Δ	0	Δ	0	0	0	0	0	0	0	0	0	0	0	0	0
1	4	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	k_9	k_{10}	k_{11}	k_{12}	k_{13}	$k_{14} + t_1$	$k_{15} + t_2$	k_0
		0	Δ	0	0	0	0	0	0	0	0	0	0	0	Δ	Δ	0
2	8	k_2	k_3	k_4	k_5	k_6	k_7	k_8	k_9	k_{10}	k_{11}	k_{12}	k_{13}	k_{14}	$k_{15} + t_2$	$k_0 + t_0$	k_1
		Δ	0	0	0	0	0	0	0	0	0	0	0	0	Δ	0	Δ
3	12	k_3	k_4	k_5	k_6	k_7	k_8	k_9	k_{10}	k_{11}	k_{12}	k_{13}	k_{14}	k_{15}	$k_0 + t_0$	$k_1 + t_1$	k_2
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	16	k_4	k_5	k_6	k_7	k_8	k_9	k_{10}	k_{11}	k_{12}	k_{13}	k_{14}	k_{15}	k_0	$k_1 + t_1$	$k_2 + t_2$	k_3
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	Δ	Δ
5	20	k_5	k_6	k_7	k_8	k_9	k_{10}	k_{11}	k_{12}	k_{13}	k_{14}	k_{15}	k_0	k_1	$k_2 + t_2$	$k_3 + t_0$	k_4
		0	0	0	0	0	0	0	0	0	0	0	0	Δ	Δ	Δ	0

$>>> 7$, the only possible difference pattern in registers is either $0xAAAAAAAA$ or zero which can be indicated as 1-bit value. So the linear differential trails with this difference pattern form a small space of size 2^{32}, which can be searched by brute force. The linear differential trail in [13] is the best one in this space. But this attack doesn't work on BLAKE-64, because the cyclic rotation constants are different. BLAKE-64 uses the number of rotations 32, 25, 16 and 11. Two of them are not multiples of 4, which implies more restrictions of the differential trail.

To obtain near collisions for a reduced-round compression function of BLAKE-64 and improve the previous near-collision attack on a reduced-round compression function of BLAKE-32 in [13], we have to release the restrictions. This can be done in two ways: using non-linear differential trail instead of linear one, or still using linear differential trail but releasing restrictions on the differential pattern. In this paper, we use linear differential trail and try to release restrictions on the differential pattern. Instead of using cyclic invariant differences, we use a random difference of hamming weight less than or equal to two in the intermediate states.

Since we intend to release restrictions on the differential pattern, the cyclic invariant differential pattern in previous works is not used. So the cyclic rotations can not be removed.

Figure 6 and Figure 7 show the linearized G function of BLAKE-32 and BLAKE-64 respectively.

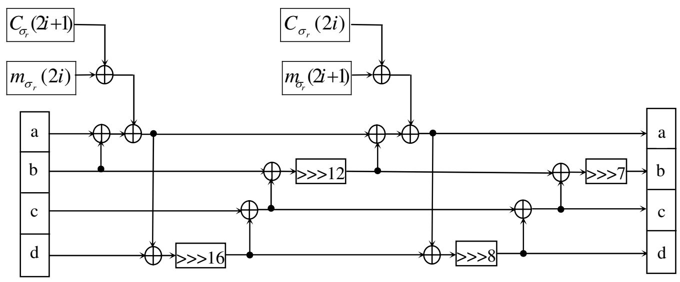

Fig. 6. linearized G function in BLAKE-32

4.2 Searching for Differential Trails with High Probability

We need to choose the starting point after linearizing G function. Since BLAKE-32 has 10 rounds and BLAKE-64 has 14 rounds, there are less than 2^4 possible choices. Then we place one or two bits of differences in the message blocks and certain round of the intermediate state at the starting point. Because compression function of BLAKE-32 uses 512-bit message and 512-bit state and compression function of BLAKE-64 uses 1024-bit message and 1024-bit state, there are

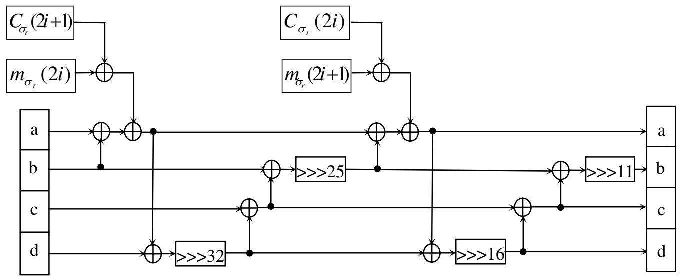

Fig. 7. linearized G function in BLAKE-64

$\binom{1024}{1} + \binom{1024}{2} \approx 2^{19}$ and $\binom{2048}{1} + \binom{2048}{2} \approx 2^{21}$ choices of positions for the pair of bits on BLAKE-32 and BLAKE-64 respectively. Therefore, the search spaces for BLAKE-32 and BLAKE-64 are less than 2^{23} and 2^{25} respectively, which can be explored exhaustively.

Our aim is to find one path with the highest probability in the search space. Furthermore, following Section 3.1, we calculate probability of one differential trail by counting hamming weight in the differences. We search for differential trails of 4-round compression function of BLAKE-32, 4-round and 5-round compression functions of BLAKE-64. And the results are introduced in the following sections.

4.3 Near Collision for 4-Round Compression Function of BLAKE-32

We search with the configuration where differences are in $m[0] = 0x80008000$ and $v[0, 2, 4, 8, 10]$ and find that a starting point at round 4 leads to a linear differential trail whose total hamming weight is 21. We don't need to count for the last round, since it can be fulfilled by message modifications with similar techniques used in attacks on Skein.

So, This trail can be fulfilled with probability of 2^{-21}. Complexity of this attack is 2^{21} with no memory requirements. With assumption that no differences in the salt value, this configuration has a final collision on $256 - 104 = 152$ bits after the finalization. Table 8 in the appendix demonstrates how differences propagate in intermediate chaining values from round 4 to 7.

4.4 Near Collision for the 4-Round Compression Function of BLAKE-64

We search with the configuration where differences are in $m[11] = 0x80000000$ 80000000 and $v[0, 2, 4, 8, 10]$ and find that a starting point at round 7 leads to a linear differential trail whose total hamming weight is equal to 16. We don't need to count for the last round, since it can be fulfilled by message modifications with similar techniques used in attacks on Skein.

So, This trail can be fulfilled with probability of 2^{-16}. Complexity of this attack is 2^{16} with no memory requirements. With assumption that no differences in the salt value, this configuration has a final collision on $512 - 116 = 396$ bits after the finalization. Table 9 in the appendix demonstrates how differences propagate in intermediate chaining values from round 7 to 10.

4.5 Near Collision for the 5-Round Compression Function of BLAKE-64

Then we search for 5-round differential trails, with the configuration where differences are placed in $m[11] = 0x8000000080000000$ and $v[0, 2, 4, 8, 10]$. We find that a starting point at round 7 leads to a linear differential trail whose total hamming weight is 216. This trail with probability of 2^{-216} is illustrated in Table 10 of the appendix, which leads to a $512 - 206 = 306$-bit collision after feedforward. The message modifications are also applied to the last round.

5 Conclusion

In this paper, we revisited the linear differential techniques and applied it to two ARX-based hash functions: Skein and BLAKE. Our attacks include near-collision attacks on the 24-round compression functions of Skein-256, Skein-512 and Skein-1024, the 4-round compression function of BLAKE-32, and the 4-round and 5-round compression functions of BLAKE-64. Future works might apply some non-linear differentials for integer addition besides XOR differences to improve our results.

References

1. National Institute of Standards and Technology: Announcing Request for Candidate Algorithm Nominations for a New Cryptographic Hash Algorithm (SHA-3) Family. Federal Register, 27(212):62212-62220(Nov. 2007) Available: `http://csrc.nist.gov/groups/ST/hash/documents/FR_Notice_Nov07.pdf`(2008/10/17)
2. Xiaoyun Wang, Xuejia Lai, Dengguo Feng, Hui Chen, Xiuyuan Yu: Cryptanalysis of the Hash Functions MD4 and RIPEMD. EUROCRYPT 2005, LNCS 3494, pp. 1-18, Springer Verlag, 2005
3. Xiaoyun Wang, Hongbo Yu: How to Break MD5 and Other Hash Functions. EUROCRYPT 2005, LNCS 3494, pp. 19-35, Springer Verlag, 2005
4. Xiaoyun Wang, Hongbo Yu, Yiqun Lisa Yin: Efficient Collision Search Attacks on SHA-0. CRYPTO 2005, LNCS 3621, pp. 1-16, Springer Verlag, 2005
5. Xiaoyun Wang, Yiqun Lisa Yin, Hongbo Yu: Finding Collisions in the Full SHA-1. CRYPTO 2005, LNCS 3621, pp. 17-36, Springer Verlag, 2005
6. E. Biham and O. Dunkelman. A Framework for Iterative Hash Functions - HAIFA. In Second NIST Cryptographic Hash Workshop, Santa Barbara, California, USA, August 24-25, 2006, August 2006
7. Florent Chabaud, Antoine Joux, Differential Collisions in SHA-0, Advanced in Cryptology, proceedings of CRYPTO 1998, LNCS 1462, pp. 56-71, Springer Verlag, 1998
8. Sebastiaan Indesteege, Bart Preneel, Practical Collisions for EnRUPT, FSE 2009, LNCS 5665, pp. 122-138, Springer Verlag, 2009

9. Eric Brier, Shahram Khazaei, Willi Meier, Thomas Peyrin, Linearization Framework for Collision Attacks: Application to Cubehash and MD6, Advanced in Cryptology, proceedings of ASIACRYPT 2009, LNCS 5912, pp. 560-577, Springer Verlag, 2009
10. Vincent Rijmen, Elisabeth Oswald, Update on SHA-1, Topics in Cryptology, proceedings of CT-RSA 2005, LNCS 3376, pp. 58-71, Springer Verlag, 2005
11. Jean-Philippe Aumasson, L. Henzen, W. Meier, and R. C.-W. Phan. SHA-3 proposal BLAKE, version 1.3. Available online at http://131002.net/blake/blake.pdf, 2008
12. Jean-Philippe Aumasson, Çagdas Çalik, Willi Meier, Onur Özen, Raphael C.-W. Phan, Kerem Varici: Improved Cryptanalysis of Skein. ASIACRYPT 2009, LNCS 5912, pp. 542-559, Springer Verlag, 2009
13. Jean-Philippe Aumasson, Jian Guo, Simon Knellwolf, Krystian Matusiewicz, and Willi Meier, Differential and Invertibility Properties of BLAKE, FSE 2010, to appear in LNCS, Springer 2010
14. Ferguson, N., Lucks, S., Schneier, B., Whiting, D., Bellare, M., Kohno, T., Callas, J., Walker, J.: The Skein Hash Function Family. Submission to NIST (2008)
15. D. J. Bernstein. ChaCha, a variant of Salsa20. Available online at http://cr.yp.to/chacha/chacha-20080128.pdf, January 2008.

A Differential Trails of Reduced-Round Skein and BLAKE

Table 6. Differential trail used for near collision of 24-round Skein-512, with probability of 2^{-230}.

Rd	Difference				Pr
20	177363f900ab3668	36ed5b708e227114	55bc1c3e7881275c	4e65052fe03ee6b3	-
	8ca8e770541856b3	36a6043068ef74e1	821adaa76647acf8	d0857e4c77f10cb0	
21	1bd9191198bfc1ef	0af0294dc0abc1a1	3a0ee3403cf72252	2e074b0908d70142	M
	d29fa4eb11b6a048	a21e22e38124a488	a19e38898e89477c	811420858c004114	
22	1409a84934202310	1400884920202110	70818608909204c0	6181000c80a00440	M
	208a180c02890668	080a080c02002648	1129305c5814004e	110110541804004c	
23	1100860410320080	1100040410220080	2880100000892020	2080100040882200	M
	0028200840100002	00a0200800100000	0009200014000200	0001000010000200	
24	0800000040010220	0080000440000220	0088000040000002	0008000000000002	M
	0008200004000000	8008200000000000	0000820000100000	8000820000000000	
25	0080000040000000	0000000040000000	0000000004000000	0000000004000000	2^{-43}
	0000000000100000	0000000000100000	0880000400010000	0080000000010000	
26	0000000000000000	0000000000000000	0000000000000000	0000000000000000	2^{-8}
	0800000400000000	0800000000000000	0080000000000000	0080000000000000	
27	0000000000000000	0000000000000000	0000000400000000	0000000400000000	2^{-3}
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
28	0000000000000000	0000000000000000	0000000000000000	0000000000000000	2^{-1}
	0000000000000000	0000000000000000	0000000000000000	8000000000000000	
	no differences in round 29 - 36				1
37	0000000000000000	0000000000000000	8000000000000000	0000000000000000	1
	0000000000000000	8000000000000000	0000000000000000	0000000000000000	
38	8000000000000000	0000000000000000	8000000000000000	0000000000000000	1
	0000000000000000	8000000000002000	0000000000000000	8000000000000000	
39	8000000000000000	8000000000000000	8000000000002000	8000020000000000	2^{-1}
	8000000000000000	8000000008000000	8000000000000000	8000000000000000	
40	0000020000002000	0000000100000000	0000000008000000	0000020000000000	2^{-3}
	0000000000000000	0000020008002000	0000000000000000	0000020004002010	
41	8000020008000000	0002020100002000	8000020008002000	80100a0004012012	2^{-24}
	8000020004002010	0002020008000000	0000020100002000	800102000c000000	
42	001008000c010012	8400000148002004	800200000c002010	800100410c812006	2^{-26}
	800100010c002000	000200000c200010	8002000108002000	4035082018010810	
43	0003004100810016	8412980104082816	8003000100202010	c0b300010a85381c	2^{-47}
	c037082110012810	4203010100002010	8410080144012016	400e004300031914	
44	40b000000aa5180c	b413905054a42025	8234092010010800	c592826243021882	2^{-74}
	c41e084244023902	8236082430311810	8411984004892800	308981660aa70d06	

Table 7. Differential trail used for near collision of Skein-1024, of probability 2^{-395}.

Rd	Difference				Pr
0	19784dd0abac34ae	195468f0130f00ce	1866a2c424af0b54	fc2f300ca644975c	-
	724160f9fbe7774d	354b6cea52cf6b59	b7e8d028e7ee826b	c80d060ce08aa6aa	
	9e01dc1568d478f3	6c62c73d18ea1df5	9c52d04d61b020b8	90f0436baf866419	
	c56a33799988135a	4620157d0e931057	fc472494ac63eae4	7839420c8263b374	
1	802c2520b8a33460	90a426309a23906a	644992c882eb9c08	dc0982c082ca8b08	M
	7fe5d624076424c1	8a75cc2a06056541	470a0c13a9281c14	4808081729281800	
	0ca29326ce3644a1	2cb0b22284625484	834a2604971b030d	824806001000038d	
	047e66982e005990	0c66e64166434521	f2631b28703e6506	703f2a2076ba6008	
2	108803102280a40a	90a0038028009409	b840100800211700	9841100800000508	M
	0f02040480000414	0d02000480000410	f5901a0e01614180	5510300601614100	
	01022004871b0080	03022000840b9080	081880d948431cb1	8818005151400c81	
	825c31080684050e	805c30080404000a	201221044a541025	201220804810002c	
3	802800900a803003	0020008008801003	2001000000211208	0001000000201208	M
	a0802a0800000080	80a2220800000000	0200040000000004	0200040000000000	
	8000808819031030	0000008018030030	0200010002800504	0200000000800100	
	0000018402441009	0000008000041009	0200000403109000	0000000403109000	
4	8008001002002000	8000001002002000	2000000000010000	2000000000000000	M
	0000000000000004	0000000000000004	2022080000000080	6012000000000080	
	0000010002000404	0000010002000004	0000010402400000	0000000402400000	
	0200000000000000	8200000000000000	8000800801001000	0000800001000000	
5	8008000000000000	0008000002000000	0000000000010000	0000000000010000	2^{-71}
	4030080000000000	0010080000000000	0000000000000000	0000000000000000	
	0000010000000000	0000010000000000	0000000000000000	0000000000000000	
	0000000800001000	0000000000001000	0000000000000400	0000000000000400	
6	8000000002000000	0000000002000000	0000000000000000	0000000000000000	2^{-11}
	0000000000000000	0000000000000000	4020000000000000	4000000000000000	
	0000000000000000	0000000000000000	0000000800000000	0000000800000000	
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
7	8000000000000000	0000000000000000	0000000000000000	0000000000000000	2^{-4}
	0020000000000000	0020000000000000	0000000000000000	0000000000000000	
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
8	8000000000000000	0000000000000000	0000000000000000	0000000000000000	2^{-1}
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
	0000000000000000	8000000000000000	0000000000000000	8000000000000000	
	no differences in round 9 - 16				1
17	0000000000000000	0000000000000000	0000000000000000	0000000000000000	1
	0000000000000000	0000000000000000	0000000000000000	0000000020000000	
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
18	0000000000000000	0000000000000000	0000000000000000	0000000000000000	2^{-1}
	0000000020000000	0000000000000000	0000000000000000	0000000000000000	
	0000000000000000	1000000020000000	0000000000000000	0000000000000000	
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
19	0000000000000000	1000020020000100	0000000000000000	0000000000000000	2^{-3}
	0000000000000000	0000000000000000	0000000020000000	0000000000000000	
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
	0000000000000000	0000000020000000	1000000020000000	0000000000000000	
20	1000020020000100	0000000000000000	0000000000000000	0000000020000200	2^{-8}
	0000000020000000	0000000000000000	0000000000000000	1000000020000000	
	0000000000000000	0000000020000000	0000000020000000	0000000000000000	
	1000000020000000	0000000000000000	0000000000000000	1002022020040140	
21	1000020020000100	0000000220000000	0000000020000200	1000020020000000	2^{-42}
	1000000020000000	0000000020000000	0000000020000000	9800820024004100	
	0000000020000000	1000800020010000	1000000020000000	8000080020000200	
	9002022020040140	0000000020000000	0000000020000000	1000020020000100	
22	1000020200000100	1008800200110001	1000020000000200	90020220000401c0	2^{-39}
	9800820004004100	d010080100040200	1000000000000000	1002022000040140	
	9000080000000200	9004420414016102	9002022000040140	9000020040000a00	
	1000020000000100	1000000000200000	1000800000010000	1002220200000100	
23	0008820000110101	00406a455417732b	80020020000403c0	8000024000200100	2^{-74}
	0002022000040140	40100020000c0b41	48108a0104044300	0000a22204450500	
	0002002040040b40	0a02823011040160	0000020000200100	80120722400c0b40	
	0002a20200010100	48c09a0905044702	00044a0414016302	800a821100101109	
24	0048e8455406722a	aa28a11141401c20	00020260002402c0	59de3128076d6216	2^{-141}
	4810282300414600	e8020762a4640bc1	4012020000080a01	a02cdb150115500b	
	80120522402c0a40	6010282611516428	48c2380b05054602	0082066200240bc0	
	800ec8151411720b	015212002008060a	0a00821051000a20	6a0dbc52272d726a	

Table 8. Differential trail used for near collision of 4-round BLAKE-32, with probability of 2^{-21}.

Rd	Difference	Pr
4	88008800 00000000 80008000 00000000	-
	88008800 00000000 00000000 00000000	
	80008000 00000000 80008000 00000000	
	00000000 00000000 00000000 00000000	
5	00000000 00000000 80008000 00000000	2^{-12}
	00000000 00000000 00000000 00000000	
	00000000 00000000 00000000 00000000	
	00000000 00000000 00000000 00000000	
6	00000000 00000000 00000000 00000000	2^{-1}
	00000000 00000000 00000000 00000000	
	00000000 00000000 00000000 00000000	
	00000000 00000000 00000000 00000000	
7	80088008 00000000 00000000 00000000	2^{-8}
	00000000 11101110 00000000 00000000	
	00000000 00000000 88008800 00000000	
	00000000 00000000 00000000 08000800	
8	28222822 18981898 11111111 19181918	M
	33123312 44414441 02230223 32233223	
	91919191 10101010 28222822 08080808	
	89918991 08800880 89918991 08880888	

Table 9. Differential trail used for near collision of 4-round BLAKE-64, with probability of 2^{-16}.

Rd	Difference	Pr
7	8100000081000000 0000000000000000 8000000080000000 0000000000000000	-
	8100000081000000 0000000000000000 0000000000000000 0000000000000000	
	8000000080000000 0000000000000000 8000000080000000 0000000000000000	
	0000000000000000 0000000000000000 0000000000000000 0000000000000000	
8	0000000000000000 0000000000000000 8000000080000000 0000000000000000	2^{-12}
	0000000000000000 0000000000000000 0000000000000000 0000000000000000	
	0000000000000000 0000000000000000 0000000000000000 0000000000000000	
	0000000000000000 0000000000000000 0000000000000000 0000000000000000	
9	0000000000000000 0000000000000000 0000000000000000 0000000000000000	2^{-1}
	0000000000000000 0000000000000000 0000000000000000 0000000000000000	
	0000000000000000 0000000000000000 0000000000000000 0000000000000000	
	0000000000000000 0000000000000000 0000000000000000 0000000000000000	
10	8000000080000000 0000000000000000 0000000000000000 0000000000000000	2^{-3}
	0000000000000000 0000001000000010 0000000000000000 0000000000000000	
	0000000000000000 0000000000000000 0000800000008000 0000000000000000	
	0000000000000000 0000000000000000 0000000000000000 0000800000008000	
11	8240204082402040 a8402040a8402040 0850085008500850 2850200028502000	M
	0a0002000a000200 0004400400044004 0010080000100800 0a110a010a110a01	
	8850081088500810 2010285020102850 2240000022400000 a0002840a0002840	
	2840a0002840a000 0040000000400000 2840200028402000 2040804020408040	

Table 10. Differential trail used for near collision of 5-round BLAKE-64, with probability of 2^{-216}.

Rd	Difference				Pr
7	8100000081000000	0000000000000000	8000000080000000	0000000000000000	-
	8100000081000000	0000000000000000	0000000000000000	0000000000000000	
	8000000080000000	0000000000000000	8000000080000000	0000000000000000	
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
8	0000000000000000	0000000000000000	8000000080000000	0000000000000000	2^{-12}
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
9	0000000000000000	0000000000000000	0000000000000000	0000000000000000	2^{-1}
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
10	8000000080000000	0000000000000000	0000000000000000	0000000000000000	2^{-3}
	0000000000000000	0000001000000010	0000000000000000	0000000000000000	
	0000000000000000	0000000000000000	0000800000008000	0000000000000000	
	0000000000000000	0000000000000000	0000000000000000	0000800000008000	
11	8240204082402040	a8402040a8402040	0850085008500850	2850200028502000	2^{-200}
	0a0002000a000200	0004400400044004	0010080000100800	0a110a010a110a01	
	8850081088500810	2010285020102850	2240000022400000	a0002840a0002840	
	2840a0002840a000	0040000000400000	2840200028402000	2040804020408040	
12	8a14284d8a14284d	8285222482852224	c2a442e0c2a442e0	4881023048810230	M
	001d0aac001d0aac	1b001a111b001a11	4aa500044aa50004	0c284c3c0c284c3c	
	6ab4c0e56ab4c0e5	c26048d1c26048d1	2851a04d2851a04d	0a6122d00a6122d0	
	0081aa700081aa70	28c0209128c02091	2885223428852234	0091a8950091a895	