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

Bozhan Su, Wenling Wu, Shuang Wu, and Le Dong

State Key Laboratory of Information Security,
Institute of Software, Chinese Academy of Sciences, Beijing 100190, P.R. China
Graduate University of Chinese Academy of Sciences, Beijing 100049, 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 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 20-round compression functions of Skein-256, Skein-512 and a 24-round compression function of Skein-1024 have a complexity of 2^{97}, 2^{52} and 2^{452} respectively.

Keywords: 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.

S.-H. Heng, R.N. Wright, and B.-M. Goi (Eds.): CANS 2010, LNCS 6467, pp. 124–139, 2010.

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, we select some intermediate state as a starting point and place a low Hamming weight difference in it. 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	20	2^{97}	-	130-bit near-collision	✓
Skein-512	20	2^{52}	-	266-bit near-collision	✓
Skein-1024	24	2^{452}	-	512-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].

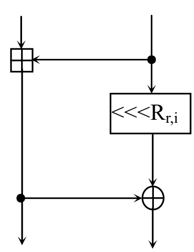

Fig. 1. The MIX function

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.

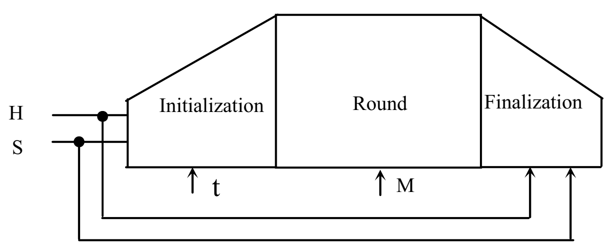

Fig. 2. Overall Structure of Compression Function of BLAKE

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

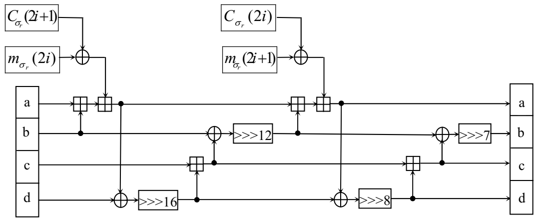

Fig. 3. The G function of BLAKE-32 for index i

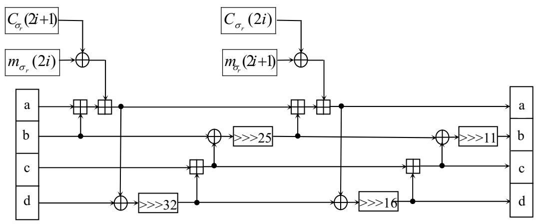

Fig. 4. The G function of BLAKE-64 for index i

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:

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

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.

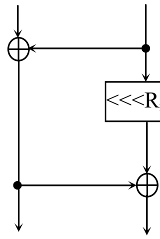

Fig. 5. linearized MIX function in Threefish

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.

3.1 Near Collisions for the 20-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 and t_0 gives a difference $(\Delta, \Delta, 0, 0)$ at the third subkey, and $(0, 0, 0, 0)$ after the fourth. The difference in the state of round 8 is canceled out at the third subkey which is then turned into an eight-round local collision from round 9 to round 16. After 20 rounds, the Hamming weight of the difference becomes too large to obtain near collisions. In the 20-th round, after adding the final subkey and feedforward value, one obtains a collision on $256 - 126 = 130$ bits. Table 2 shows the corresponding differential trail of the key and the tweak from the 0-th round to the 19-th round. Table 3 presents the corresponding trail from the 0-th round to the 19-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 and t_0

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

Table 3. Differential trail used for near collision of a 20-round compression function of Skein-256, with probability of 2^{-97}

Rd	Difference				Pr
0	b0dff57c25c19314	a5b2b6692bd196c8	861349393b7673c0	3c708bb2d1caf2d2	-
1	e82d8c56764c8096	956d43150e1005dc	601166d49d04b503	3a63c28beabc8112	M
2	0a44a5491af1e45a	7d40cf43785c854a	5090945bd4b01c4b	5a72a45f77b83411	M
3	2708680a86a06010	77046a0a62ad6110	86e030002608280a	0ae23004a308285a	M
4	5004000044050100	500c0200e40d0100	8400000405000050	8c02000485000050	M
5	0008000020080000	80080200a0080000	0802000000000000	0802000080000000	2^{-58}
6	0000020000000000	8000020080000000	0000000000000000	0000000080000000	2^{-8}
7	0000000080000000	8000000080000000	0000000080000000	0000000080000000	2^{-3}
8	8000000000000000	8000000000000000	0000000000000000	0000000000000000	2^{-2}
	no differences in round 9 - 16				1
17	0000000000002000	8000000000000000	8000000000008000	0000000000000000	1
18	8008000000008008	8000000000002000	8000000000002040	8000000000008000	2^{-2}
19	000000102040a040	000800000000a008	008808800800a008	000000000000a040	2^{-7}
20	a156edfd2dd5925c	25bab6790b919680	8e0f41291b36718c	3cf88332d9caf29a	2^{-17}

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 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
		0	0	0	0	0	Δ	0	Δ
6	24	k_6	k_7	k_8	k_0	k_1	k_2+t_0	k_3+t_1	k_4
		0	0	0	0	0	0	0	Δ
7	28	k_7	k_8	k_0	k_1	k_2	k_3+t_1	k_4+t_2	k_5
		0	0	0	0	0	0	0	0
8	32	k_8	k_0	k_1	k_2	k_3	k_4+t_2	k_5+t_0	k_6
		0	0	0	0	Δ	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
		0	0	0	Δ	Δ	0	Δ	0
10	40	k_1	k_2	k_3	k_4	k_5	k_6+t_1	k_7+t_2	k_8
		0	0	Δ	Δ	0	Δ	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}
		0	Δ	0	0	0	0	0	0	0	0	0	0	0	Δ	Δ	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
		Δ	0	0	0	0	0	0	0	0	0	0	0	0	Δ	0	Δ
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
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
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	Δ	Δ
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	Δ	Δ	Δ	0
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	Δ	0	Δ	Δ	0
6	24	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	k_3+t_0	k_4+t_1	k_5
		0	0	0	0	0	0	0	0	0	0	Δ	0	Δ	Δ	Δ	0

3.2 Near Collisions for the 20-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 $>>> 7$, 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.

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

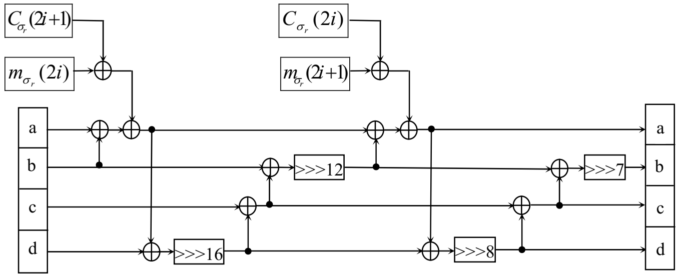

Fig. 6. linearized G function in BLAKE-32

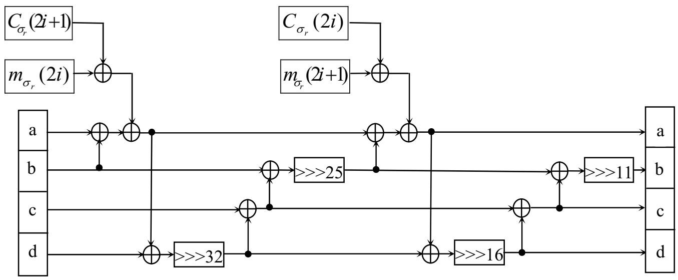

Fig. 7. linearized G function in BLAKE-64

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 $\binom{1024}{1} + \binom{1024}{2} \approx 2^{19}$ 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 $m[11] = 0x8000000080000000$ 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 20-round compression functions of Skein-256, Skein-512 and the 24-round compression function of 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.

Acknowledgment

The authors would like to thank the anonymous referees for their valuable comments. Furthermore, this work is supported by the National Natural Science Foundation of China (No. 60873259, and No. 60903212) and the Knowledge Innovation Project of The Chinese Academy of Sciences.

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 (2007), http://csrc.nist.gov/groups/ST/hash/documents/FR_Notice_Nov07.pdf (17/10/2008)
2. Wang, X., Lai, X., Feng, D., Chen, H., Yu, X.: Cryptanalysis of the Hash Functions MD4 and RIPEMD. In: Cramer, R. (ed.) EUROCRYPT 2005. LNCS, vol. 3494, pp. 1–18. Springer, Heidelberg (2005)
3. Wang, X., Yu, H.: How to Break MD5 and Other Hash Functions. In: Cramer, R. (ed.) EUROCRYPT 2005. LNCS, vol. 3494, pp. 19–35. Springer, Heidelberg (2005)
4. Wang, X., Yu, H., Yin, Y.L.: Efficient Collision Search Attacks on SHA-0. In: Shoup, V. (ed.) CRYPTO 2005. LNCS, vol. 3621, pp. 1–16. Springer, Heidelberg (2005)
5. Wang, X., Yin, Y.L., Yu, H.: Finding Collisions in the Full SHA-1. In: Shoup, V. (ed.) CRYPTO 2005. LNCS, vol. 3621, pp. 17–36. Springer, Heidelberg (2005)
6. Biham, E., Dunkelman, O.: A Framework for Iterative Hash Functions - HAIFA. In: Second NIST Cryptographic Hash Workshop, Santa Barbara, California, USA, August 24-25 (2006)
7. Chabaud, F., Joux, A.: Differential Collisions in SHA-0. In: Krawczyk, H. (ed.) CRYPTO 1998. LNCS, vol. 1462, pp. 56–71. Springer, Heidelberg (1998)
8. Indesteege, S., Preneel, B.: Practical Collisions for EnRUPT. In: Dunkelman, O. (ed.) FSE 2009. LNCS, vol. 5665, pp. 122–138. Springer, Heidelberg (2009)
9. Brier, E., Khazaei, S., Meier, W., Peyrin, T.: Linearization Framework for Collision Attacks: Application to Cubehash and MD6. In: Matsui, M. (ed.) ASIACRYPT 2009. LNCS, vol. 5912, pp. 560–577. Springer, Heidelberg (2009)
10. Rijmen, V., Oswald, E.: Update on SHA-1. In: Menezes, A. (ed.) CT-RSA 2005. LNCS, vol. 3376, pp. 58–71. Springer, Heidelberg (2005)
11. Aumasson, J.-P., Henzen, L., Meier, W., Phan, R.C.-W.: SHA-3 proposal BLAKE, version 1.3 (2008), http://131002.net/blake/blake.pdf
12. Aumasson, J.-P., Çalik, Ç., Meier, W., Özen, O., Phan, R.C.-W., Varici, K.: Improved Cryptanalysis of Skein. In: Matsui, M. (ed.) ASIACRYPT 2009. LNCS, vol. 5912, pp. 542–559. Springer, Heidelberg (2009)
13. Aumasson, J.-P., Guo, J., Knellwolf, S., Matusiewicz, K., Meier, W.: Differential and Invertibility Properties of BLAKE. In: Beyer, I. (ed.) FSE 2010. LNCS, vol. 6147, pp. 318–332. Springer, Heidelberg (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. Bernstein, D.J.: ChaCha, a variant of Salsa20 (January 2008), http://cr.yp.to/chacha/chacha-20080128.pdf

A Differential Trails of Reduced-Round Skein and BLAKE

Table 6. Differential trail used for near collision of 20-round Skein-512, with probability of 2^{-52}

Rd	Difference	Pr
20	0000000010004800 0020001000004000 0002201000080000 0000200000080000 8000000020000200 8000000020000200 0000088000080000 8000008000080000	-
21	0002001000000000 0000001000000000 8000000000000000 8000000000000000 0000080000000000 0000080000000000 0020001010000800 0000001000000800	2^{-35}
22	0000000000000000 0000000000000000 0000000000000000 0000000000000000 0020000010000000 0020000000000000 0002000000000000 0002000000000000	2^{-7}
23	0000000000000000 0000000000000000 0000000010000000 0000000010000000 0000000000000000 0000000000000000 0000000000000000 0000000000000000	2^{-3}
24	0000000000000000 0000000000000000 0000000000000000 0000000000000000 0000000000000000 0000000000000000 0000000000000000 8000000000000000	2^{-1}
	no differences in round 25 - 32	1
33	0000000000000000 0000000000000000 8000000000000000 0000000000000000 0000000000000000 8000000000000000 0000000000000000 0000000000000000	1
34	8000000000000000 0000000000000000 8000000000000000 0000000000000000 0000000000000000 8000000000002000 0000000000000000 8000000000000000	1
35	8000000000000000 8000000000000000 8000000000002000 8000004000000000 8000000000000000 8002000800002000 8000000000000000 8000000000000000	2^{-1}
36	0000004000002000 0000080000000000 0002000800002000 0080000000000000 0000000000000000 0022008802002008 0000000000000000 0000804000002100	2^{-5}
37	8082000800002000 0000084000042000 8022008802002008 c000806100002180 8000804000002100 882280a802882228 0000084000002000 8082000820202000	M
38	402280e902000188 818a084884040000 082200e802880328 8092480860210104 8082084820200000 8220a0e22200a108 8082084800040000 c62180eb03840188	M
39	88b048e062a9022c 50a080a187071598 02a2a8aa0220a108 66afce920f875994 46a388a303800188 02f22ceb1270d019 c1a888a186040188 84b468c0f2bb4b2d	M
40	640d66381da7b09c 78b069d6e2bbcfe4 c453845811f8d191 f5206eb3bfd667bf c51ce06154bf48a5 5d535664dae2a341 5810c0c1e5a617b4 9837aa1b38d18c0c	M

Table 7. Differential trail used for near collision of Skein-1024, of probability 2^{-452}

Rd	Difference				Pr
0	8140008142000042	8040008100000042	0000000000080040	0000000000000040	-
	0000000000000080	0000000000000080	4100000100488224	4000000100480200	
	0001000000024040	0001000000020040	0010208010000000	0010008010000000	
	2000000000000000	a000000000000000	8000044000008002	0000040000000002	
1	8100000042000000	0100000002000000	0000000000080000	0000000000080000	2^{-87}
	0100000000008024	0100000000000020	0000000000000000	0000000000000000	
	0000200000000000	0000200000000000	0000000000000000	0000000000000000	
	0000004000008000	0000000000008000	0000000000004000	0000000000004000	
2	8000000040000000	0000000040000000	0000000000000000	0000000000000000	2^{-12}
	0000000000000000	0000000000000000	0000000000008004	0000000000000004	
	0000000000000000	0000000000000000	0000004000000000	0000004000000000	
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
3	8000000000000000	0000000000000000	0000000000000000	0000000000000000	2^{-4}
	0000000000008000	0000000000008000	0000000000000000	0000000000000000	
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
4	8000000000000000	0000000000000000	0000000000000000	0000000000000000	2^{-1}
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
	0000000000000000	8000000000000000	0000000000000000	8000000000000000	
	no differences in round 5 - 12				1
13	0000000000000000	0000000000000000	0000000000000000	0000000000000000	1
	0000000000000000	0000000000000000	0000000000000000	0000000020000000	
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
14	0000000000000000	0000000000000000	0000000000000000	0000000000000000	2^{-1}
	0000000020000000	0000000000000000	0000000000000000	0000000000000000	
	0000000000000000	0000000020010000	0000000000000000	0000000000000000	
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
15	0000000000000000	0001000820010000	0000000000000000	0000000000000000	2^{-3}
	0000000000000000	0000000000000000	0000000020000000	0000000000000000	
	0000000000000000	0000000000000000	0000000000000000	0000000000000000	
	0000000000000000	0000000020000000	0000000020010000	0000000000000000	
16	0001000820010000	0000000000000000	0000000000000000	0000000020000004	2^{-8}
	0000000020000000	0000000000000000	0000000000000000	0000000020010000	
	0000000000000000	0000000020000000	0000000020000000	0000000000000000	
	0000000020010000	0000000000000000	0000000000000000	0201104822010000	
17	0001000820010000	0000002020000000	0000000020000004	0000000020210000	2^{-42}
	0000000020010000	0000000020000000	0000000020000000	c221104862230904	
	0000000020000000	8000000020011000	0000000020010000	0000040020008004	
	8201104822010000	0000000020000000	0000000020000000	0001000820010000	
18	0001002800010000	a001000000015002	0000000000210004	8211104802010000	2^{-47}
	c221104842230904	1000840200118004	0000000000010000	0001001800830010	
	0000040000018004	404000c066121880	8201104802010000	0001010800210004	
	0001000800010000	0000008000010000	8000000000011000	0001002800010808	
19	a000002800005002	d80866c167139b85	8211104802200004	0001008c00000800	2^{-84}
	0001001800820010	c000194000221004	d221944a42328900	8051002a1010180a	
	8200114002200004	2002000860800010	0001008800000000	a30014c82230000c	
	8001002800001808	d201144e6222898c	404004c066139884	a002a02d40025002	
20	780866e96713cb87	a20a014962a43054	821010c402200804	464e7644ebae4385	2^{-163}
	527094605222910a	21221440e8000140	c001195800a01014	bac2a04d2351cdc6	
	a30114402230000c	72408160f022c52a	5200146662229184	8ad010c482200814	
	e042a4ed2611c886	d005d95819e01036	a202114862a00014	7904bec58560bb3c	
21	da0267a005b7fbd3	d9579a22fe406202	c45e6680e98e4b81	b14418b56264592e	M
	7ac3b91523f1ddd2	5cda01cae860d880	73528020ba22904a	3a5e8142f9819499	
	58d004a2e0029990	eb6b5ff67e908df4	b0477db53ff1d8b0	58b2ef57b509410d	
	5b06af8de7c0bb28	7352a8249e601857	d1419520d212c526	9262cee411b56916	
22	0355fd82fbf799d1	050ea433779acb2a	751a7e358bea12af	a8355a6433003106	M
	490c016243a304d3	c95e84bf600e3895	2619b8dfcb910552	29f176e0063a6413	
	e8f592e28af899bd	edc5d39649b4c8df	285407a979a0a37f	fc8b1a8f4efa707a	
	43235bc4c3a7ac30	a64562de0179658a	b3bb5b549e921464	997703c299f54086	
23	065b59b18c6d52fb	99820cd285b33f4c	dd2f2451b8ea23a9	733e937e94f329ad	M
	0fe8ce3fcdab6141	3d1ef6d41b30ee3e	805285dd23ad3c46	afffb2170a55bae5	
	d4df1d26375ad305	96ec0443901360cf	e566391ac2dec9ba	de3f4ed2ed4c6099	
	2acc5896076754e2	9c18617261f28c41	05304174c34c5162	bd963c248eea00e2	
24	9fee540b09e9742f	63fb10d5c082c5c8	ae11bf272c2e139c	88b2be9fe5aeef4f	M
	2fad3fc229cf87db	4dc84784c08d0ee2	32f638ebd6897253	067c7ad0439f7753	
	c4a688375301a8c3	81b79521741b2223	36d439ed66a2d8a3	85f11291bf6796f7	
	38b482904da65194	6b71411a3e2c0f92	bea1c00ba749b3ce	9b8060686fe0cc74	

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	