Full-Round Differential Attack on the Original Version of the Hash Function Proposed at PKC'98

Donghoon Chang^{1*}, Jaechul Sung², Soohak Sung³, Sangjin Lee¹, and Jongin ${\rm Lim}^1$

¹ Center for Information Security Technologies(CIST), Korea University, Anam Dong, Sungbuk Gu, Seoul, Korea {pointchang, sangjin, jilim}@cist.korea.ac.kr ² Korea Information Security Agency(KISA), Karag-dong, Songpa-gu, Seoul, Korea sjames@kisa.or.kr ³ Paichai University, Daejeon, Korea, sungsh@mail.paichai.ac.kr

Abstract. Shin et al.[4] proposed a new hash function with 160-bit output length at PKC'98. Recently, at FSE 2002, Han et al.[5] cryptanalyzed the hash function proposed at PKC'98 and suggested a method finding a collision pair with probability 2^{-30} , supposing that boolean functions satisfy the SAC(Strict Avalanche Criterion). This paper improves their attack and shows that we can find a collision pair from the original version of the hash function with probability $2^{-37.13}$ through the improved method. Furthermore we point out a weakness of the function comes from shift values dependent on message.

1 Introduction

MD4, MD5, RIPEMD-160, HAVAL, SHA-1 are well known dedicated hash functions. Dobbertin[2][3] showed that there are serious weakness in MD4 and MD5. Haval[11] was attacked partially. Shin et al.[4] proposed a new hash function with 160-bit output length at PKC'98.

Recently, at FSE 2002, Han et al.[5] pointed out that, unlike the designer's attention, some of the boolean functions of the hash function proposed at PKC'98 do not satisfy the SAC(Strict Avalanche Criterion). And they analyzed the hash function proposed at PKC'98 and found a collision pair with probability 2^{-30} , supposing that the boolean functions satisfy the SAC. However only one of boolean functions used in the hash function satisfies the SAC. So, their attack introduced at FSE 2002 can not be applied to the hash function itself.

This paper improves the method proposed at FSE 2002 and shows that we can find a collision pair "from the original version" of the hash function with

^{*} This work was supported by both Ministry of Information and Communication and Korea Information Security Agency, Korea, under project 2002-130

K. Nyberg and H. Heys (Eds.): SAC 2002, LNCS 2595, pp. 160-174, 2003.

[©] Springer-Verlag Berlin Heidelberg 2003

probability $2^{-37.13}$ through the improved method. And we point out the problem of the hash function comes from shift values dependent on message. Next, we show that reduced versions for 3-pass HAVAL are attacked by our attack method.

2 The Hash Function Proposed at PKC'98

In this section, we briefly describe the hash function proposed at PKC'98, and we introduce notations used in this paper.

word	32-bit string
block	512-bit string used as input of compression function in the hash function
+	addition modulo 2^{32} operation between two words
$X^{<< s}$	left rotation X by s bits
$X \land Y$	bitwise logical AND operation of X and Y
$X \lor Y$	bitwise OR operation of X and Y
$X{\oplus}Y$	bitwise XOR operation of X and Y

2.1 Input Block Length and Padding

An input message is processed by 512-bit block. The proposed hash function pads a message by appending a single bit 1 next to the least significant bit of the message, followed by zero or more bit 0s until the length of the message is 448 modulo 512, and then appends to the message the 64-bit original message length modulo 2^{64} .

2.2 Initial Value(IV)

The initial values of five chaining variables (A,B,C,D,E) used in processing message are as follows.

А	В	С	D	Е
0x67452301	0xefcdab89	0x98badcef	0x10325476	0xc3d2e1f0

2.3 Constants

The following numbers are used as $constants(K_i \text{ is used in round } i)$.

 $K_1=$ 0 $K_2=$ 0x5a827999 $K_3=$ 0x6ed9eba1 $K_4=$ 0x8f1bbcdc

2.4 Expansion of Message Variables

Eight message variables, $X_{16}, X_{17}, \dots, X_{23}$ are additionally generated from original sixteen input message words, X_0, X_1, \dots, X_{15} as follows. So twenty-four message words are applied to the compression function.

$$X_{16+i} = (X_{0+i} \oplus X_{2+i} \oplus X_{7+i} \oplus X_{12+i})^{<<1} \quad (i = 0, 1, \cdots, 7)$$
(1)

2.5 Ordering of Message Words

This hash function consists of four rounds. Each round has 24 steps.

Table 1. The definition of permutation ρ

																								23
$\rho(i)$	4	21	17	1	23	18	12	10	5	16	8	0	20	3	22	6	11	19	15	2	7	14	9	13

The ordering of message words is determined by ρ as follows.

Round	1	2	3	4
Permutation	id	ρ	ρ^2	ρ^3

2.6 Boolean Functions

The boolean functions used at each round are as follows. Only the function f_2 satisfies the SAC, while it is not the case for f_0 and f_1 .

 $f_0(x_1, x_2, x_3, x_4, x_5) = (x_1 \land x_2) \oplus (x_3 \land x_4) \oplus (x_2 \land x_3 \land x_4) \oplus x_5$ (2)

 $f_1(x_1, x_2, x_3, x_4, x_5) = x_2 \oplus ((x_4 \wedge x_5) \vee (x_1 \wedge x_3))$ (3)

$$f_2(x_1, x_2, x_3, x_4, x_5) = x_1 \oplus (x_2 \wedge (x_1 \oplus x_4)) \oplus (((x_1 \wedge x_4) \oplus x_3) \wedge x_5)$$
(4)

2.7 Operation in One Step

The operation in one step is defined as follows. The function f_0 is applied to the first round($0\sim23$ step), the function f_1 is applied to the second round($24\sim47$ step), the function f_2 is applied to the third round($48\sim71$ step) and the function f_1 is applied to the fourth round($72\sim95$ step).

$$A_{i} = (f(A_{i}, B_{i}, C_{i}, D_{i}, E_{i}) + X + K)^{<

$$(5)$$$$

$$A_{i+1} = E_i, B_{i+1} = A_i, C_{i+1} = B_i, D_{i+1} = C_i, E_{i+1} = D_i$$
(6)

The operation in one step can be described in the figure 1.

2.8 Shift Operation

The shift values used in each step, $S_i (i = 0, 1, 2, \dots, 95)$, are determined depending on message words as follows.

$$S_i = X_{R(i \mod 24)} \mod 32 \tag{7}$$

Next table shows R function per round.

Round	1	2	3	4
R function	ρ^3	ρ^2	ρ	id

For example, let's solve S_{20} , shift value at step 20. The step 20 is in the first round, so R function is ρ^3 . Therefore $S_{20} = X_{\rho^3(20)} \mod 32 = X_8 \mod 32$.

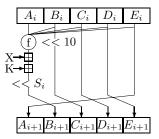


Fig. 1. The operation in one step

2.9 Each Step Operation through the Table

The following table shows each step operation from step 0 to step 6. A_i, B_i, C_i, D_i, E_i mean chaining variables after a equation (5) and before a equation (6) at *i* step.

Table 2. The message input order, shift value and chaining values per step

step	А	В	С	D	Е	Input	Shift value(mod 32)
0	A_0	B_0	C_0	D_0	E_0	X_0	X ₁₃
1	A_1	B_1	C_1	D_1	E_1	X_1	X_{22}
2	A_2	B_2	C_2	D_2	E_2	X_2	X_2
3	A_3	B_3	C_3	D_3	E_3	X_3	X_{14}
4	A_4	B_4	C_4	D_4	E_4	X_4	X_3
5	A_5	B_5	C_5	D_5	E_5	X_5	X_6
6	A_6	B_6	C_6	D_6	E_6	X_6	X_7

 \Box : The updated part per step

3 The Analysis of the Original Version of the Hash Function Proposed at PKC'98

3.1 The Analysis of Input and Output Difference for Boolean Functions

Boolean function f_2 satisfies the SAC. This means that when it has a difference at only one bit, output bit becomes zero with 1/2 probability. At FSE 2002, under supposing that three boolean functions satisfy the SAC, Han et al. proposed a method finding a collision pair with probability 2^{-30} . But in fact, boolean functions f_0 and f_1 do not satisfy the SAC. In case of f_0 , when only last input bit has a difference, output difference becomes always 1(table 3-5). Also, in case of f_1 , when only second input bit has a difference, output difference becomes always 1(table 3-13). These mean the followings: if a message word having a difference is an input of a round using the function f_0 , the difference avalanche occurs after four steps because fifth chaining variable have a difference. If message word having a difference is input of a round using the function f_1 , the difference avalanche occurs after one step because second chaining variable have a difference. And the effect of difference avalanche is larger according as the number of steps increases. Therefore the method proposed by Han et al. at FSE 2002 cannot be applied to the original version of the hash function. But, in case of f_0 , if we give simultaneously differences to fifth chaining variable and other chaining variable, the output bit becomes zero with some probability(table 3-8,9,11). This is similar to f_1 . We will use this fact to attack the original version of the hash function proposed at PKC'98.

No.	Func.	Input	Prob.	No.	Func.	Input	Prob.
1	f_0	x_1	1/2	17	f_1	x_{2}, x_{3}	3/8
2	f_0	x_2	1/2	18	f_1	x_2, x_5	3/8
3	f_0	x_3	3/4	19	f_1	x_{1}, x_{4}	3/8
4	f_0	x_4	3/4	20	f_1	x_1, x_2	3/8
5	f_0	x_5	0	21	f_1	x_1, x_3	5/8
6	f_0	x_1, x_2	1/2	22	f_1	x_{2}, x_{4}	3/8
7	f_0	x_1, x_4	1/4	23	f_1	x_3, x_5	3/8
8	f_0	x_{3}, x_{5}	1/4	24	f_1	x_{4}, x_{5}	5/8
9	f_0	x_1, x_5	1/2	25	f_1	x_{3}, x_{4}	3/8
10	f_0	x_{3}, x_{4}	3/4	26	f_1	x_{1}, x_{5}	3/8
11	f_0	x_4, x_5	1/4	27	f_1	x_2, x_3, x_4	5/8
12	f_1	x_1	5/8	28	f_2	x_1	1/2
13	f_1	x_2	0	29	f_2	x_2	1/2
14	f_1	x_3	5/8	30	f_2	x_3	1/2
15	f_1	x_4	5/8	31	f_2	x_4	1/2
16	f_1	x_5	5/8	32	f_2	x_5	1/2

Table 3. The probability that output difference bit becomes 0 with the change of location of input difference bit of boolean function f_i

3.2 The Analysis of Input and Output Difference for Step Operation

At each step, such as equation (5), message word(X) having a difference has influence on a updating chaining variable(A_i). Therefore, we need to take the following facts into consideration for difference analysis. The step operation uses addition modulo 2^{32} . If addition modulo 2^{32} is substituted by \oplus , we must consider a probability according to a carry. If k=31, with respect to most significant bit, addition modulo 2^{32} and \oplus play the same role. If $k \neq 31$, a carry happens with probability 1/2. So, the analysis of input and output difference of step operation is done by equations (8), (9), (10), (11). α and β denote arbitrary words, and p denotes the probability satisfying a equation.

$$case \ 1: k = 31$$

$$((\alpha \oplus 1^{<\kappa}) + \beta) \oplus (\alpha + \beta) = 1^{<\kappa} \qquad (p = 1) \tag{8}$$

$$((\alpha \oplus 1 \vee \alpha) + (\beta \oplus 1 \vee \alpha)) \oplus (\alpha + \beta) \equiv 0 \qquad (p \equiv 1)$$

$$(3)$$

$$case \ 2: k \neq 31$$

$$\begin{array}{l} ((\alpha \oplus 1^{<< k}) + \beta) \oplus (\alpha + \beta) = 1^{<< k} & (p = 1/2) \\ ((\alpha \oplus 1^{<< k}) + (\beta \oplus 1^{<< k})) \oplus (\alpha + \beta) = 0 & (p = 1/2) \end{array}$$
(10) (11)

3.3 The Analysis of Expansion of Message Words

By (1), $(12)\sim(19)$ are obtained as follows.

$$X_{16} = (X_0 \oplus X_2 \oplus X_7 \oplus X_{12})^{<<1}$$
(12)

$$X_{17} = (X_1 \oplus X_3 \oplus X_8 \oplus X_{13})^{<<1}$$
(13)
$$X_{17} = (X_2 \oplus Y_2 \oplus Y_2 \oplus Y_{23})^{<<1}$$
(14)

$$\begin{aligned}
\Lambda_{18} &= (\Lambda_2 \oplus \Lambda_4 \oplus \Lambda_9 \oplus \Lambda_{14})^{++} \\
X_{19} &= (X_3 \oplus X_5 \oplus X_{10} \oplus X_{15})^{<<1}
\end{aligned} \tag{14}$$

$$X_{19} = (X_4 \oplus X_6 \oplus X_{11} \oplus X_{16})^{<<1}$$
(16)

$$X_{21} = (X_5 \oplus X_7 \oplus X_{12} \oplus X_{17})^{<<1}$$
(17)

$$X_{22} = (X_6 \oplus X_8 \oplus X_{13} \oplus X_{18})^{<<1}$$
(18)

$$X_{23} = (X_7 \oplus X_9 \oplus X_{14} \oplus X_{19})^{<<1}$$
(19)

Through the expansion of message words, message words which are affected by each $X_i (0 \le i \le 15)$ are shown on Table 4 below. This fact was shown by Han et al. at FSE 2002. When we take a look at the table 4, if we give a difference to X_0 , X_{16} and X_{20} also have some differences.

Table 4. The effect of expansion of message words

X_0	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
X_{16}	X_{17}	X_{16}	X_{17}	X_{18}	X_{19}	X_{20}	X_{16}	X_{17}	X_{18}	X_{19}	X_{20}	X_{16}	X_{17}	X_{18}	X_{19}
X_{20}	X_{21}	X_{18}	X_{19}	X_{20}	X_{21}	X_{22}	X_{20}	X_{21}	X_{22}	X_{23}		X_{20}	X_{21}	X_{22}	X_{23}
		X_{20}	X_{21}	X_{22}	X_{23}		X_{21}	X_{22}	X_{23}			X_{21}	X_{22}	X_{23}	
		X_{22}	X_{23}				X_{23}								

3.4 The Analysis of Weakness of Shift Value Dependent on Message Words

Generally, in case of MDx-hash functions, shift values are fixed. If shift values can be selected arbitrarily by attacker, a hash function using these shift values can be attacked easily by our attack method (described in Appendix A), regardless of the expansion of message words and ordering of message words and boolean functions and step operation. The original version of the hash function proposed at PKC'98 uses shift values dependent on message words. This principle of the design makes an attacker select shift values. So, in section 5, based on this fact, we will attack the original version of the hash function proposed at PKC'98.

4 The Selection of Message Blocks and Shift Values for Maximizing an Attack Probability

4.1 The Selection of Message Block Pair

When we take a look at the table 4, for the expansion of message words, it is impossible to give a difference to only one message word pair. But we can give differences to two message word pairs. Four cases exist.

 $(X_8, X_{13}), (X_9, X_{14}), (X_{10}, X_{15}), (X_{11}, X_{20})$. Out of four cases, we select (X_{11}, X_{20}) for maximizing the probability of attack, because (X_{11}, X_{20}) makes the smallest sum of four cases–the sums of the differences of steps of inputting $(X_8, X_{13}), (X_9, X_{14}), (X_{10}, X_{15}), (X_{11}, X_{20})$ per round. The table of four cases is as follows.

Table 5. The differences of steps of inputting each word pair per round

Round	(X_8, X_{13})	(X_9, X_{14})	(X_{10}, X_{15})	(X_{11}, X_{20})
1 Round	5	5	5	9
2 Round	13	1	11	4
3 Round	3	13	13	3
4 Round	20	18	4	7
Sum	41	37	33	23

4.2 The Selection of Differences of Message Block Pair and Shift Values

The first standard giving differences to block pair is to give a difference to most significant bit for changing a modular addition with XOR operation. The second is to apply the same shift values to message block pair.

We select message block pair and shift values as follows.

$$\begin{array}{l} \cdot X = (X_i)_{0 \leq i \leq 15} , \ \widetilde{X} = (\widetilde{X_i})_{0 \leq i \leq 15} : \text{message block pair} \\ \cdot X_{11} \oplus \widetilde{X_{11}} = 1^{<<30} (\widetilde{X_{20}} \oplus 1^{<<31} = X_{20}) \\ \cdot \widetilde{X_i} = X_i (i \neq 11, 20) \\ \cdot \text{ Value of 5 low-order bits of message words} \\ lsb_5(X_1, X_4, X_7, X_{12}, X_{15}, X_{16}, X_{17}, X_{20}, X_{23}) = (10, \ 0, \ 21, \ 21, \ 11, \ 0, \ 10, \\ 0, \ 13) \end{array}$$

4.3 Input Order of Message Words and Shift Values per Round through the Table

Table 6 shows input order of message words and shift values per step.(M: order of message words, \cdot : non-fixed value)

step	Μ	$_{\rm shift}$									
0	0	•	24	4	13	48	23	0	72	13	
1	1		25	21		49	14		73	22	10
2	2		26	17		50	19	10	74	2	
3	3	•	27	1		51	21	10	75	14	•
4	4	•	28	23	•	52	13	13	76	3	0
5	5		29	18	11	53	15		77	6	
6	6	21	30	12	0	54	20	21	78	7	
7	7		31	10		55	8		79	5	21
8	8	11	32	5		56	18		80	15	
9	9		33	16	•	57	11	0	81	0	
10	10		34	8		58	5		82	18	
11	11	13	35	0	0	59	4		83	23	
12	12		36	20	21	60	7	0	84	10	21
13	13	•	37	3	10	61	1		85	21	
14	14	0	38	22		62	9		86	16	
15	15	0	39	6	21	63	12		87	20	11
16	16	0	40	11		64	0		88	4	0
17	17	10	41	19	•	65	2	•	89	17	10
18	18	21	42	15		66	6	11	90	12	
19	19	•	43	2	10	67	17		91	19	
20	20	•	44	7	•	68	10	21	92	8	0
21	21	•	45	14	•	69	22	•	93	9	•
22	22	•	46	9	0	70	16		94	11	•
23	23	10	47	13	•	71	3	•	95	1	13

Table 6. The message input order and shift values per step

5 Full-Round Differential Attack

From now on, we start to attack the original version of the hash function proposed at PKC'98 based on sections 3 and 4. Table 7 shows the full-round differential attack. Concretely we analyze steps generating differences. See Appendix A. $\triangle A_i$, $\triangle B_i$, $\triangle C_i$, $\triangle D_i$, $\triangle E_i$ mean the differences of chaining variables after a equation (5) and before a equation (6) at *i* step.

The total probability is about $2^{-37.13}$. Therefore we can find a collision pair of the original version of the hash function proposed at PKC'98 with probability $2^{-37.13}$.

6 Finding a Collision Pair in Practice by Simulation

For finding a collision pair, we executed a program written in visual C 6.0 and running on a set of 10 PCs under Windows. From this, we found a collision pair in one computer with about 10 hours. The collision pair is as follows.

Step	$\triangle A$	$\triangle B$	$\triangle C$	$\triangle D$	$\triangle E$	$\triangle X$	p	Step	$\triangle A$	$\triangle B$	$\triangle C$	$\triangle D$	$\triangle E$	$\bigtriangleup X$	p
0	0	0	0	0	0	0	1	48	0	0	0	0	0	0	1
1	0	0	0	0	0	0	1	49	0	0	0	0	0	0	1
2	0	0	0	0	0	0	1	50	0	0	0	0	0	0	1
3	0	0	0	0	0	0	1	51	0	0	0	0	0	0	1
4	0	0	0	0	0	0	1	52	0	0	0	0	0	0	1
5	0	0	0	0	0	0	1	53	0	0	0	0	0	0	1
6	0	0	0	0	0	0	1	54	0	$1^{<20}$	0	0	0	$1^{<31}$	1
7	0	0	0	0	0	0	1	55	0	$1^{<30}$	0	0	0	0	1/2
8	0	0	0	0	0	0	1	56	0	$1^{<30}$	0	0	0	0	1/2
9	0	0	0	0	0	0	1	57	0	$1^{<30}$	0	0	0	$1^{<30}$	1/4
10	0	0	0	0	0	0	1	58	0	$1^{<30}$	0	0	0	0	1/2
11	0	0	0	0	$_{1} < 11$	$1^{<30}$	1/2	59	0	0	0	0	0	0	1/2
12	0	0	0	0	$1^{<21}$	0	1/2	60	0	0	0	0	0	0	1
13	0	0	0	0	$1^{<21}$	0	3/4	61	0	0	0	0	0	0	1
14	0	0	0	0	$1^{<21}$	0	3/4	62	0	0	0	0	0	0	1
15	$1^{<21}$	0	0	0	1 < 21	0	1/2	63	0	0	0	0	0	0	1
16	$1^{<31}$	0	0	0	$1^{<21}$	0	1/4	64	0	0	0	0	0	0	1
17	$1^{<31}$	0	0	0	$1^{<31}$	0	3/8	65	0	0	0	0	0	0	1
18	$1^{<31}$	0	0	0	1 < 31	0	3/4	66	0	0	0	0	0	0	1
19	$1^{<31}$	0	0	0	$1^{<31}$	0	1/4	67	0	0	0	0	0	0	1
20	0	0	0	0	1 < 31	$1^{<31}$	1/2	68	0	0	0	0	0	0	1
21	0	0	0	0	0	0	1/2	69	0	0	0	0	0	0	1
22	0	0	0	0	0	0	1	70	0	0	0	0	0	0	1
23	0	0	0	0	0	0	1	71	0	0	0	0	0	0	1
24	0	0	0	0	0	0	1	72	0	0	0	0	0	0	1
25	0	0	0	0	0	0	1	73	0	0	0	0	0	0	1
26	0	0	0	0	0	0	1	74	0	0	0	0	0	0	1
27	0	0	0	0	0	0	1	75	0	0	0	0	0	0	1
28	0	0	0	0	0	0	1	76	0	0	0	0	0	0	1
29	0	0	0	0	0	0	1	77	0	0	0	0	0	0	1
30	0	0	0	0	0	0	1	78	0	0	0	0	0	0	1
31	0	0	0	0	0	0	1	79	0	0	0	0	0	0	1
32	0	0	0	0	0	0	1	80	0	0	0	0	0	0	1
33	0	0	0	0	0	0	1	81	0	0	0	0	0	0	1
34	0	0	0	0	0	0	1	82	0	0	0		0	0	1
35					0 1<20	0 1<31		83				0			
36	0	0	0	0		1~31	1	84	0	0	0	0	0	0	1
37	0	0	0	1<30	1<30	0	1/2	85	0	0	0	0	0	0	1
38	0	0	0	$1^{<8}$	1 < 30	0	3/8	86	0	0	0	0	0	0	1
39	0	0	0	$1^{<8}$	1 < 30	0	25/64	87	0	0	0	$1^{<10}$	0	$1^{<31}$	1
40	0	0	0	$1^{<8}$	1 < 30	1 < 30	15/128	88	0	0	$1^{<10}$	$1^{<20}$	0	0	1/2
41	0	0	0	$^{1 < 8}$	0	0	25/64	89	0	$1^{<20}$	$1^{<20}$	$1^{<20}$	0	0	5/16
42	0	0	0	0	0	0	5/8	90	0	$1^{<30}$	$1^{<20}$	$1^{<20}$	0	0	5/8
43	0	0	0	0	0	0	1	91	0	$1^{<30}$	$1^{<20}$	$1^{<20}$	0	0	25/64
44	0	0	0	0	0	0	1	92	0	$1^{<30}$	$1^{<20}$	0	0	0	15/64
45	0	0	0	0	0	0	1	93	0	1 < 30	0	0	0	0	25/64
46	0	0	0	0	0	0	1	94	0	0	0	0	0	$1^{<30}$	3/16
47	0	0	0	0	0	0	1	95	0	0	0	0	0	0	1
· · · · ·	ı						· · · · · ·		11 I.						

 Table 7. The Full-Round Differential Attack

 \Box :The updated part per step

$X_0 = \texttt{0xdf407f1a}$	$X_1 = \texttt{0x99c0464a}$	$X_2 = \texttt{0x3380a1fa}$	$X_3 = \texttt{0x0d40be50}$
$X_4 = \texttt{0x6580c1c0}$	$X_5 = \texttt{0xb8803020}$	$X_6 = \texttt{0xf5c09a9e}$	$X_7 = \texttt{0x388077d5}$
$X_8 = \texttt{0x1f005106}$	$X_9 = \texttt{0xb080db94}$	$X_{10} = \texttt{0xb700244c}$	$X_{11} = 0$ x3480cc5e
$X_{12} = \texttt{0xb5c00895}$			$X_{15} = \texttt{0xba008ecb}$
	$\widetilde{X_{11}} = X_{11} \oplus 1^{<<3}$	$\widetilde{X}_i = X_i (i \neq 11)$	

The output of compression function for two message blocks is as follows.

Oxfe684dca 0x33524aa4 0x15ce9f59 0xd200e689 0x7b01f656

And a collision pair on the compression function leads to a collision on the full hash function by simply appending identical blocks or the padding fields for same-length messages.

7 On the Security of Reduced Versions of 3-Pass HAVAL

HAVAL is a dedicated hash function of the MD family which was proposed by Zheng et al[11]. Kasselman et al. found collisions for the last two passes of 3-pass HAVAL[6]. Park et al., at ACISP 2002, found a 256-bit collision of the first two passes of 3-pass HAVAL and of the last two passes of 3-pass HAVAL[7]. In this paper, we attack reduced versions of 3-pass HAVAL using our method different from two previous attack methods(1-2 round: attack probability is 2^{-18} , 2-3 round: attack probability is 2^{-50}). Concretely, we select two messages as follow. For 1-2 round of HAVAL, $W_{29} \oplus \widetilde{W}_{29} = 1^{<<31}, \widetilde{W}_{31} \oplus 1^{<<20} = W_{31}$. For 2-3 round of HAVAL, $W_7 \oplus \widetilde{W}_7 = 1^{<<31}, \widetilde{W}_{31} \oplus 1^{<<30} = W_{31}$.(See Appendix B: Table 8, 9)

8 Conclusion

This paper shows that we can find a collision pair from the original hash function with probability $2^{-37.13}$ not transforming boolean functions. This means that a weakness of the hash function proposed at PKC'98 comes from the shift values of the hash function. That is, shift values depending on messages can be a factor of reducing the security of hash functions. Also, our attack method is applied to the reduced version of HAVAL because HAVAL has the weakness of message input order.

Therefore shift values have to be carefully chosen for the security of hash functions. And in case that shift values are fixed, the message input order have to be carefully chosen, also.

Acknowledgements

We thank to the anonymous referees for their useful comments.

References

- F. Chabaud and A. Joux, Differential Collisions in SHA-0, Advances in CRYPTO'98, LNCS 1462, Springer-Verlag, 1998, pp. 56–71.
- 2. H. Dobbertin, *Cryptanalysis of MD4*, Fast Software Encryption, LNCS 1039, Springer-Verlag, 1996, pp. 53–69.
- 3. H. Dobbertin, Cryptanalysis of MD5 Compress, May. 1996. www-cse.ucsd.edu/users/bsy/dobbertin.ps
- Sanguk Shin, Kyunghyune Rhee, Daehyun Ryu, Sangjin Lee, A New Hash Function Based on MDx-family and Its Application to MAC, Public Key Cryptography'98, pp. 234–246. 1998.
- Daewan Han, Sangwoo Park, Seongtaek Chee, Cryptanalysis of a Hash Function Proposed at PKC'98, Fast Software Encryption 2002, LNCS 2365, pp. 252–262.
- P.R. Kasselman and W.T. Penzhorn, Cryptanalysis of reduced version of HAVAL, Electronics Letters 6th January 2000 Vol.36 No.1, pp. 30–31.
- S.W. Park, S.H. Sung, S.T. Chee, J.I. Lim, On the Security of Reduced Versions of 3-Pass HAVAL, ACISP 2002, LNCS 2384, pp. 406–419.
- 8. R. Rivest, *The MD4 message digest algorithm*, RFC 1320, Internet Activities Board, Internet Privacy Task Force, Apr. 1992.
- 9. R. Rivest, *The MD5 message digest algorithm*, RFC 1321, Internet Activities Board, Internet Privacy Task Force, Apr. 1992.
- 10. Federal Information Processing Standards Publication 180-1, April 17, 1995.
- Y. Zheng, J. Pieprzyk and J. Sebberry, HAVAL A one-way hashing algorithm with variable length of output, Advances in Cryptology-Auscrypt'92, LNCS 718, Springer-Verlag, 1993, pp. 83–104.

Appendix A

1. The Analysis of 11~21 Steps of Table 7

The first round use the function f_0 . The values of Table 7 are calculated as follows.

$$\Delta E_{11} = [f_0(E_{10}, A_{10}, B_{10}, C_{10}, D_{10}) + X_{11} + K_1]^{<$$

Above equality holds with probability 1/2 by (10).

$$\Delta D_{12} = [f_0(D_{11}, E_{11}, A_{11}, B_{11}, C_{11}) + X_{12} + K_1]^{<$$

Above equality holds with probability 1/2 by 2 of Table 3.

 $\triangle C_{13} = [f_0(C_{12}, D_{12}, E_{12}, A_{12}, B_{12}) + X_{13} + K_1]^{<<S_{13}} \oplus [f_0(C_{12}, D_{12}, E_{12} \oplus 1^{<<21}, A_{12}, A_{12},$

$$B_{12}) + X_{13} + K_1]^{<$$

Above equality holds with probability 3/4 by 3 of Table 3.

$$\Delta B_{14} = [f_0(B_{13}, C_{13}, D_{13}, E_{13}, A_{13}) + X_{14} + K_1]^{<$$

Above equality holds with probability 3/4 by 4 of Table 3.

 $\Delta A_{15} = [f_0(A_{14}, B_{14}, C_{14}, D_{14}, E_{14}) + X_{15} + K_1]^{<<S_{15}=0} \oplus [f_0(A_{14}, B_{14}, C_{14}, D_{14}, E_{14} \oplus 1^{<<21}) + X_{15} + K_1]^{<<S_{15}=0} = 1^{<<21} \qquad (p = 1/2)$

Above equality holds with probability 1/2 by (10) and by 5 of Table 3.

$$\Delta E_{16} = [f_0(E_{15}, A_{15}, B_{15}, C_{15}, D_{15}) + X_{16} + K_1]^{<$$

Above equality holds with probability 1/4 by (10) and by 6 of Table 3.

 $\Delta D_{17} = [f_0(D_{16}, E_{16}, A_{16}, B_{16}, C_{16}) + X_{17} + K_1]^{\langle \langle S_{17} = 10 \rangle} \oplus [f_0(D_{16}, E_{16} \oplus 1^{\langle \langle 21 \rangle}, A_{16} \oplus 1^{\langle \langle 31 \rangle}, B_{16}, C_{16}) + X_{17} + K_1]^{\langle \langle S_{17} = 10 \rangle} = 0$ (p = 3/8)

 $\oplus 1^{-51}, B_{16}, C_{16} + X_{17} + K_1 = 0$ Above equality holds with probability 3/8 by 2, 3 of Table 3.

$$\begin{split} & \triangle C_{18} = [f_0(C_{17}, D_{17}, E_{17}, A_{17}, B_{17}) + X_{18} + K_1]^{<$$

 $\Delta B_{19} = [f_0(B_{18}, C_{18}, D_{18}, E_{18}, A_{18}) + X_{19} + K_1]^{<< S_{19}} \oplus [f_0(B_{18}, C_{18}, D_{18}, E_{18} \oplus 1^{<<31}, A_{18} \oplus 1^{<<31}) + X_{19} + K_1]^{<< S_{19}} = 0 \qquad (p = 1/4)$

Above equality holds with probability 1/4 by 11 of Table 3.

$$\Delta A_{20} = [f_0(A_{19}, B_{19}, C_{19}, D_{19}, E_{19}) + X_{20} + K_1]^{<
$$\Delta E_{21} = [f_0(E_{20}, A_{20}, B_{20}, C_{20}, D_{20}) + X_{21} + K_1]^{<$$$$

2. The Analysis of 36~42 Steps of Table 7

The probabilities of following equalities also are calculated like the above method.

$$\begin{split} & \triangle E_{36} = [f_1(E_{35}, A_{35}, B_{35}, C_{35}, D_{35}) + X_{20} + K_2]^{<< S_{36} = 21} \oplus [f_1(E_{35}, A_{35}, B_{35}, C_{35}, D_{35}) + (X_{20} \oplus 1^{<<31}) + K_2]^{<< S_{36} = 21} = 1^{<<20} (p=1) \\ & \triangle D_{37} = [f_1(D_{36}, E_{36}, A_{36}, B_{36}, C_{36}) + X_3 + K_2]^{<< S_{37} = 10} \oplus [f_1(D_{36}, E_{36} \oplus 1^{<<20}, A_{36}, B_{36}, C_{36}) + X_3 + K_2]^{<< S_{37} = 10} \oplus [f_1(D_{36}, E_{36} \oplus 1^{<<20}, A_{36}, B_{36}, C_{36}) + X_3 + K_2]^{<< S_{37} = 10} \oplus [f_1(C_{37}, D_{37}, E_{37}, A_{37}, B_{37}) + X_{22} + K_2]^{<< S_{38}} \oplus [f_1(C_{37}, D_{37} \oplus 1^{<<30}, E_{37} \oplus 1^{<<30}, A_{37}, B_{37}) + X_{22} + K_2]^{<< S_{38}} = 0 (p=3/8) \\ & \triangle B_{39} = [f_1(B_{38}, C_{38}, D_{38}, E_{38}, A_{38}) + X_6 + K_2]^{<< S_{39} = 21} \oplus [f_1(B_{38}, C_{38}, D_{38} \oplus 1^{<<30}, A_{38}) + X_6 + K_2]^{<< S_{39} = 21} \oplus [f_1(A_{39}, B_{39}, C_{39}, D_{39}, E_{39}) + X_{11} + K_2]^{<< S_{40}} \oplus [f_1(A_{39}, B_{39}, C_{39}, D_{39}, E_{39}) + X_{11} + K_2]^{<< S_{40}} \oplus [f_1(A_{39}, B_{39}, C_{39}, D_{39}, E_{39}) + X_{11} + K_2]^{<< S_{40}} \oplus [f_1(A_{49}, B_{40}, C_{40}, D_{40}) + X_{19} + K_2]^{<< S_{41}} \oplus [f_1(E_{40} \oplus 1^{<<30}, A_{40}, B_{40}, C_{40}, D_{40} \oplus 1^{<<8}) + X_{19} + K_2]^{<< S_{41}} \oplus [f_1(D_{41} \oplus 1^{<<8}, E_{41}, A_{41}, B_{41}, C_{41}) + X_{15} + K_2]^{<< S_{42}} \oplus [f_1(D_{41} \oplus 1^{<<8}, E_{41}, A_{41}, B_{41}, C_{41}) + X_{15} + K_2]^{<< S_{42}} \oplus [f_1(D_{41} \oplus 1^{<<8}, E_{41}, A_{41}, B_{41}, C_{41}) + X_{15} + K_2]^{<< S_{42}} \oplus [f_1(D_{41} \oplus 1^{<<8}, E_{41}, A_{41}, B_{41}, C_{41}) + X_{15} + K_2]^{<< S_{42}} \oplus [f_1(D_{41} \oplus 1^{<<8}, E_{41}, A_{41}, B_{41}, C_{41}) + X_{15} + K_2]^{<$$

3. The Analysis of 54~59 Steps of Table 7

The probabilities of following equalities also are calculated like the above method.

$$\begin{split} & \triangle B_{54} = [f_2(B_{53},C_{53},D_{53},E_{53},A_{53}) + X_{20} + K_3]^{<< S_{54}=21} \oplus [f_2(B_{53},C_{53},D_{53},E_{53},A_{53}) + (X_{20} \oplus 1^{<<31}) + K_3]^{<< S_{54}=21} = 1^{<<20} (p=1) \\ & \triangle A_{55} = [f_2(A_{54},B_{54},C_{54},D_{54},E_{54}) + X_8 + K_3]^{<< S_{55}} \oplus [f_2(A_{54},B_{54} \oplus 1^{<<20},C_{54}, D_{54},E_{54}) + X_8 + K_3]^{<< S_{55}} \oplus [f_2(E_{55},A_{55},B_{55},C_{55},D_{55}) + X_{18} + K_3]^{<< S_{56}} \oplus [f_2(E_{55},A_{55},B_{55} \oplus 1^{<<30}, C_{55},D_{55}) + X_{18} + K_3]^{<< S_{57}=0} \oplus [f_2(D_{56},E_{56},A_{56},B_{56},C_{56}) + X_{11} + K_3]^{<< S_{57}=0} \oplus [f_2(D_{56},E_{56},A_{56},B_{56},C_{56}) + X_{11} + K_3]^{<< S_{57}=0} = 0 (p=1/2) \\ & \triangle D_{57} = [f_2(C_{57},D_{57},E_{57},A_{57},B_{57}) + X_5 + K_3]^{<< S_{57}=0} = 0 (p=1/4) \\ & \triangle C_{58} = [f_2(C_{57},D_{57},E_{57},A_{57},B_{57}) + X_5 + K_3]^{<< S_{58}} \oplus [f_2(C_{57},D_{57},E_{57},A_{57},B_{57}) + X_5 + K_3]^{<< S_{58}} \oplus [f_2(B_{58},C_{58},D_{58},D_{58},E_{58},A_{58}) + X_4 + K_3]^{<< S_{59}} \oplus [f_2(B_{58},\Phi_1^{<<30},C_{58},D_{58},E_{58},A_{58}) + X_4 + K_3]^{<< S_{59}} \oplus [f_2(B_{58},\Phi_1^{<<30},C_{58},$$

4. The Analysis of 87~95 Steps of Table 7

The probabilities of following equalities also are calculated like the above method.

$$\begin{split} & \Delta D_{87} = [f_1(D_{86}, E_{86}, A_{86}, B_{86}, C_{86}) + X_{20} + K_4]^{<< S_{87}=11} \oplus [f_1(D_{86}, E_{86}, A_{86}, B_{86}, C_{86}) + (X_{20} \oplus 1^{<<31}) + K_4]^{<< S_{87}=11} = 1^{<<10} (p=1) \\ & \Delta C_{88} = [f_1(C_{87}, D_{87}, E_{87}, A_{87}, B_{87}) + X_4 + K_4]^{<< S_{88}=0} \oplus [f_1(C_{87}, D_{87} \oplus 1^{<<10}, E_{87}, A_{87}, B_{87}) + X_4 + K_4]^{<< S_{88}=0} = 1^{<<10} (p=1/2) \\ & \Delta B_{89} = [f_1(B_{88}, C_{88}, D_{88}, E_{88}, A_{88}) + X_{17} + K_4]^{<< S_{89}=10} \oplus [f_1(B_{88}, C_{88} \oplus 1^{<<10}, D_{88} \oplus 1^{<<20}, E_{88}, A_{88}) + X_{17} + K_4]^{<$$

Appendix B

Step	$\triangle A$	$\triangle B$	$\triangle C$	$\triangle D$	$\triangle E$	$\triangle F$	$\triangle G$	$\triangle H$	$\bigtriangleup X$	p	Step	$\triangle A$	$\triangle B$	$\triangle C$	$\triangle D$	$\triangle E$	$\triangle F$	$\triangle G$	$\triangle H$	$\triangle X$	p
0	0	0	0	0	0	0	0	0	0	1	32	0	0	0	0	0	0	0	0	0	1
1	0	0	0	0	0	0	0	0	0	1	33	0	0	0	0	0	0	0	0	0	1
2	0	0	0	0	0	0	0	0	0	1	34	0	0	0	0	0	0	0	0	0	1
3	0	0	0	0	0	0	0	0	0	1	35	0	0	0	0	0	0	0	0	0	1
4	0	0	0	0	0	0	0	0	0	1	36	0	0	0	0	0	0	0	0	0	1
5	0	0	0	0	0	0	0	0	0	1	37	0	0	0	0	0	0	0	0	0	1
6	0	0	0	0	0	0	0	0	0	1	38	0	$1^{<31}$	0	0	0	0	0	0	$1^{<31}$	1
7	$1^{<31}$	0	0	0	0	0	0	0	1 < 31	1	39	0	$1^{<31}$	0	0	0	0	0	0	0	1/2
8	$1^{<31}$	0	0	0	0	0	0	0	0	1/2	40	0	1 < 31	0	0	0	0	0	0	0	1/2
9	1 < 31	0	0	0	0	0	0	0	0	1/2	41	0	1 < 31	0	0	0	0	0	0	0	1/2
10	1 < 31	0	0	0	0	0	0	0	0	1/2	42	0	1 < 31	0	0	0	0	0	0	0	1/2
11	1 < 31	0	0	0	0	0	0	0	0	1/2	43	0	1 < 31	0	0	0	0	0	0	0	1/2
12	1 < 31	0	0	0	0	0	0	0	0	1/2	44	0	$1^{<31}$	0	0	0	0	0	0	0	1/2
13	1<31	0	0	0	0	0	0	0	0	1/2	45	0	$1^{<31}$	0	0	0	0	0	0	0	1/2
14	$1^{<31}$	0	0	0	0	0	0	0	0	1/2	46	0	$_1 {<} 20$	0	0	0	0	0	0	0	1/2
15	$_{1} < 20$	0	0	0	0	0	0	0	0	1/2	47	0	1 < 20	0	0	0	0	0	0	0	1/2
16	$1^{<20}$	0	0	0	0	0	0	0	0	1/2	48	0	$1^{<20}$	0	0	0	0	0	0	0	1/2
17	$1^{<20}$	0	0	0	0	0	0	0	0	1/2	49	0	1 < 20	0	0	0	0	0	0	0	1/2
18	1 < 20	0	0	0	0	0	0	0	0	1/2	50	0	$1^{<20}$	0	0	0	0	0	0	0	1/2
19	$1^{<20}$	0	0	0	0	0	0	0	0	1/2	51	0	$1^{<20}$	0	0	0	0	0	0	0	1/2
20	$1^{<20}$	0	0	0	0	0	0	0	0	1/2	52	0	$1^{<20}$	0	0	0	0	0	0	0	1/2
21	$1^{<20}$	0	0	0	0	0	0	0	0	1/2	53	0	$1^{<20}$	0	0	0	0	0	0	0	1/2
22	$1^{<20}$	0	0	0	0	0	0	0	0	1/2	54	0	$_1 < 9$	0	0	0	0	0	0	0	1/2
23	$1^{<9}$	0	0	0	0	0	0	0	0	1/2	55	0	$1^{<9}$	0	0	0	0	0	0	0	1/2
24	$1^{<9}$	0	0	0	0	0	0	0	0	1/2	56	0	$1^{<9}$	0	0	0	0	0	0	0	1/2
25	1 < 9	0	0	0	0	0	0	0	0	1/2	57	0	$1^{<9}$	0	0	0	0	0	0	0	1/2
26	1 < 9	0	0	0	0	0	0	0	0	1/2	58	0	$1^{<9}$	0	0	0	0	0	0	0	1/2
27	$1^{<9}$	0	0	0	0	0	0	0	0	1/2	59	0	$1^{<9}$	0	0	0	0	0	0	0	1/2
28	1<9	0	0	0	0	0	0	0	0	1/2	60	0	$1^{<9}$	0	0	0	0	0	0	0	1/2
29	1<9	0	0	0	0	0	0	0	0	1/2	61	0	$1^{<9}$	0	0	0	0	0	0	0	1/2
30	$1^{<9}$	0	0	0	0	0	0	0	0	1/2	62	0	0	0	0	0	0	0	0	$1^{<30}$	1/4
31	0	0	0	0	0	0	0	0	$1^{<30}$	1/4	63	0	0	0	0	0	0	0	0	0	1

 Table 8. Attack on the first two passes of 3-pass HAVAL

 \Box :The updated part per step

Step	$\triangle A$	$\triangle B$	$\triangle C$	$\triangle D$	$\triangle E$	$\triangle F$	$\triangle G$	$\triangle H$	$\bigtriangleup X$	p	Step	$\triangle A$	$\triangle B$	$\triangle C$	$\triangle D$	$\triangle E$	$\triangle F$	$\triangle G$	$\triangle H$	$\triangle X$	p
32	0	0	0	0	0	0	0	0	0	1	64	0	0	0	0	0	0	0	0	0	1
33	0	0	0	0	0	0	0	0	0	1	65	0	0	0	0	0	0	0	0	0	1
34	0	0	0	0	0	0	0	0	0	1	66	0	0	0	0	0	0	0	0	0	1
35	0	0	0	0	0	0	0	0	0	1	67	0	0	0	0	0	0	0	0	0	1
36	0	0	0	0	0	0	0	0	0	1	68	0	0	0	0	0	0	0	0	0	1
37	0	0	0	0	0	0	0	0	0	1	69	0	0	0	0	0	0	0	0	0	1
38	0	0	0	0	0	0	0	0	0	1	70	0	0	0	0	0	0	0	0	0	1
39	0	0	0	0	0	0	0	0	0	1	71	0	0	0	0	0	0	0	0	0	1
40	0	0	0	0	0	0	0	0	0	1	72	0	0	0	0	0	0	0	$_{1} < 31$	1 < 31	1
41	0	0	0	0	0	0	0	0	0	1	73	0	0	0	0	0	0	0	$1^{<31}$	0	1/2
42	0	0	0	0	0	0	0	0	0	1	74	0	0	0	0	0	0	0	$1^{<31}$	0	1/2
43	0	0	0	0	0	0	0	0	0	1	75	0	0	0	0	0	0	0	$1^{<31}$	0	1/2
44	0	0	0	0	0	0	0	0	0	1	76	0	0	0	0	0	0	0	$1^{<31}$	0	1/2
45	0	0	0	0	0	0	0	0	0	1	77	0	0	0	0	0	0	0	$1^{<31}$	0	1/2
46	0	0	0	0	0	0	0	0	0	1	78	0	0	0	0	0	0	0	1 < 31	0	1/2
47	0	0	0	0	0	0	0	0	0	1	79	0	0	0	0	0	0	0	$1^{<31}$	0	1/2
48	0	0	0	0	0	0	0	0	0	1	80	0	0	0	0	0	0	0	0	$1^{<20}$	1/4
49	0	0	0	0	0	0	0	0	0	1	81	0	0	0	0	0	0	0	0	0	1
50	0	0	0	0	0	0	0	0	0	1	82	0	0	0	0	0	0	0	0	0	1
51	0	0	0	0	0	0	0	0	0	1	83	0	0	0	0	0	0	0	0	0	1
52	0	0	0	0	0	0	0	0	0	1	84	0	0	0	0	0	0	0	0	0	1
53	0	0	0	0	0	0	0	0	0	1	85	0	0	0	0	0	0	0	0	0	1
54	0	$1^{<31}$	0	0	0	0	0	0	1 < 31	1	86	0	0	0	0	0	0	0	0	0	1
55	0	$1^{<31}$	0	0	0	0	0	0	0	1/2	87	0	0	0	0	0	0	0	0	0	1
56	0	$1^{<31}$	0	0	0	0	0	0	0	1/2	88	0	0	0	0	0	0	0	0	0	1
57	0	$1^{<31}$	0	0	0	0	0	0	0	1/2	89	0	0	0	0	0	0	0	0	0	1
58	0	$1^{<31}$	0	0	0	0	0	0	0	1/2	90	0	0	0	0	0	0	0	0	0	1
59	0	$1^{<31}$	0	0	0	0	0	0	0	1/2	91	0	0	0	0	0	0	0	0	0	1
60	0	$1^{<31}$	0	0	0	0	0	0	0	1/2	92	0	0	0	0	0	0	0	0	0	1
61	0	$1^{<31}$	0	0	0	0	0	0	0	1/2	93	0	0	0	0	0	0	0	0	0	1
62	0	0	0	0	0	0	0	0	$1^{<20}$	1/4	94	0	0	0	0	0	0	0	0	0	1
63	0	0	0	0	0	0	0	0	0	1	95	0	0	0	0	0	0	0	0	0	1

 Table 9. Attack on the last two passes of 3-pass HAVAL

 \Box :The updated part per step