

Collisions for Hash Functions

MD4, MD5, HAVAL-128 and RIPEMD

Xiaoyun Wang¹, Dengguo Feng², Xuejia Lai³, Hongbo Yu¹

The School of Mathematics and System Science, Shandong University, Jinan250100, China¹

Institute of Software, Chinese Academy of Sciences, Beijing100080, China²

Dept. of Computer Science and Engineering, Shanghai Jiaotong University, Shanghai, China³

xywang@sdu.edu.cn¹

revised on August 17, 2004

1 Collisions for MD5

MD5 is the hash function designed by Ron Rivest [9] as a strengthened version of MD4 [8]. In 1993 Bert den Boer and Antoon Bosselaers [1] found pseudo-collision for MD5 which is made of the same message with two different sets of initial value. H. Dobbertin[3] found a free-start collision which consists of two different 512-bit messages with a chosen initial value IV'_0 .

$$IV'_0 : A'_0 = 0x12AC2375, B'_0 = 0x3B341042, C'_0 = 0x5F62B97C, D'_0 = 0x4BA763ED$$

Our attack can find many real collisions which are composed of two 1024-bit messages with the original initial value IV_0 of MD5:

$$IV_0 : A_0 = 0x67452301, B_0 = 0xefcdab89, C_0 = 0x98badcfe, D_0 = 0x10325476$$

$$M' = M + \Delta C_1, \Delta C_1 = (0,0,0,0,2^{31}, \dots, 2^{15}, \dots, 2^{31}, 0)$$

$$N'_i = N_i + \Delta C_2, \Delta C_2 = (0,0,0,0,2^{31}, \dots, -2^{15}, \dots, 2^{31}, 0)$$

(non-zeros at position 4,11 and 14)

such that

$$MD5(M, N_i) = MD5(M', N'_i).$$

On IBM P690, it takes about one hour to find such M and M' , after that, it takes only 15 seconds to 5 minutes to find N_i and N'_i , so that (M, N_i) and (M', N'_i) will produce the same hash same value. Moreover, our attack works for any given initial value.

The following are two pairs of 1024-bit messages producing collisions, the two examples have the same 1-st half 512 bits.

X_1	M	2dd31d1 c4eee6c5 69a3d69 5cf9af98 <u>87b5ca2f</u> ab7e4612 3e580440 897ffbb8 634ad55 2b3f409 8388e483 <u>5a417125</u> e8255108 9fc9cdf7 <u>f2bd1dd9</u> 5b3c3780
	N_1	d11d0b96 9c7b41dc f497d8e4 d555655a c79a7335 cfdeb0 66f12930 8fb109d1 797f2775 eb5cd530 baade822 5c15cc79 ddcb74ed 6dd3c55f d80a9bb1 e3a7cc35
X_1	$M\emptyset$	2dd31d1 c4eee6c5 69a3d69 5cf9af98 <u>7b5ca2f</u> ab7e4612 3e580440 897ffbb8 634ad55 2b3f409 8388e483 <u>5a41f125</u> e8255108 9fc9cdf7 <u>72bd1dd9</u> 5b3c3780
	N_1	d11d0b96 9c7b41dc f497d8e4 d555655a 479a7335 cfdeb0 66f12930 8fb109d1 797f2775 eb5cd530 baade822 5c154c79 ddcb74ed 6dd3c55f 580a9bb1 e3a7cc35
H	9603161f f41fc7ef 9f65ffbc a30f9dbf	
X_2	M	2dd31d1 c4eee6c5 69a3d69 5cf9af98 87b5ca2f ab7e4612 3e580440 897ffbb8 634ad55 2b3f409 8388e483 <u>5a417125</u> e8255108 9fc9cdf7 <u>f2bd1dd9</u> 5b3c3780
	N_2	313e82d8 5b8f3456 d4ac6dae c619c936 b4e253dd fd03da87 6633902 a0cd48d2 42339fe9 e87e570f 70b654ce 1e0da880 bc2198c6 9383a8b6 2b65f996 702af76f
X_2	$M\emptyset$	2dd31d1 c4eee6c5 69a3d69 5cf9af98 <u>7b5ca2f</u> ab7e4612 3e580440 897ffbb8 634ad55 2b3f409 8388e483 <u>5a41f125</u> e8255108 9fc9cdf7 <u>72bd1dd9</u> 5b3c3780
	N_2	313e82d8 5b8f3456 d4ac6dae c619c936 34e253dd fd03da87 6633902 a0cd48d2 42339fe9 e87e570f 70b654ce 1e0d2880 bc2198c6 9383a8b6 ab65f996 702af76f
H	8d5e7019 6324c015 715d6b58 61804e08	

Table 1 Two pairs of collisions for MD5

2 Collisions for HAVAL-128

HAVAL is proposed in [10]. HAVAL is a hashing algorithm that can compress messages of any length in 3,4 or 5 passes and produce a fingerprint of length 128, 160, 192 or 224 bits.

Attack on a reduced version for HAVAL was given by P. R. Kasselman and W T Penzhorn [7], which consists of last rounds for HAVAL-128. We break the full HAVAL-128 with only about the 2^6 HAVAL computations. Here we give two examples of collisions of HAVAL-128, where

$$M' = M + \Delta C, \Delta C = (2^{i-1}, 0, 0, 0, 2^{i-12}, \dots, 2^{i-8}, 0, \dots, 0)$$

with non-zeros at position 0,11,18, and $i = 0, 1, 2, \dots, 31$, such that $HAVAL(M) = HAVAL(M')$.

M_1	6377448b d9e59f18 f2aa3ccb d6cb92ba ee544a44 879fa576 1ca34633 76ca5d4f a67a8a42 8d3adc8b b6e3d814 5630998d 86ea5dcd a739ae7b 54fd8e32 acbb2b36 38183c9a b67a9289 c47299b2 27039ee5 dd555e14 839018d8 aabbd9c9 d78fc632 fff4b3a7 40000096 7f466aac fffffbc0 5f4016d2 5f4016d0 12e2b0 f4307f87
-------	--

M ₁	6377488b a67a8a42 38183c9a fff4b3a7	d9e59f18 8d3adc8b b67a9289 40000096	f2aa3cbb b6e3d814 c47299ba 7f466aac	d6cb92ba d630998d 27039ee5 fffffb0	ee544a44 86ea5dcd dd555e14 5f4016d2	879fa576 a739ae7b 839018d8 5f4016d0	1ca34633 54fd8e32 aabbd9c9 12e2b0	76ca5d4f acbb2b36 d78fc632 f4307f87
H	95b5621c	ca62817a	a48dacd8	6d2b54bf				
M ₂	6377448b a67a8a42 38183c9a fff4b3a7	d9e59f18 8d3adc8b b67a9289 40000096	f2aa3cbb b6e3d814 c47299b2 7f466aac	d6cb92ba 5630998d 27039ee5 fffffb0	ee544a44 86ea5dcd dd555e14 5f4016d2	879fa576 a739ae7b 839018d8 5f4016d0	1ca34633 54fd8e32 aabbd9c9 12e2b0	76ca5d4f acbb2b36 d78fc632 f5b16963
M ₂	6377488b a67a8a42 38183c9a fff4b3a7	d9e59f18 8d3adc8b b67a9289 40000096	f2aa3cbb b6e3d814 c47299ba 7f466aac	d6cb92ba d630998d 27039ee5 fffffb0	ee544a44 86ea5dcd dd555e14 5f4016d2	879fa576 a739ae7b 839018d8 5f4016d0	1ca34633 54fd8e32 aabbd9c9 12e2b0	76ca5d4f acbb2b36 d78fc632 f5b16963
H	b0e99492	d64eb647	5149ef30	4293733c				

Table 2 Two pairs of collision, where i=11 and these two examples differ only at the last word

3 Collisions for MD4

MD4 is designed by R. L. Rivest[8] . Attack of H. Dobbertin in Eurocrypt'96[2] can find collision with probability $1/2^{22}$. Our attack can find collision with hand calculation, such that

$$M' = M + \Delta C, \Delta C = (0, 2^{31}, -2^{28} + 2^{31}, 0, 0, 0, 0, 0, 0, -2^{16}, 0, 0, 0)$$

$$\text{and } MD4(M) = MD4(M') .$$

M ₁	4d7a9c83 c69d71b3	56cb927a f9e99198	b9d5a578 d79f805e	57a7a5ee a63bb2e8	de748a3c 45dd8e31	dcc366b3 97e31fe5	b683a020 2794bf08	3b2a5d9f b9e8c3e9
M ₁	4d7a9c83 c69d71b3	d6cb927a f9e99198	29d5a578 d79f805e	57a7a5ee a63bb2e8	de748a3c 45dc8e31	dcc366b3 97e31fe5	b683a020 2794bf08	3b2a5d9f b9e8c3e9
H	5f5c1a0d	71b36046	1b5435da	9b0d807a				
M ₂	4d7a9c83 c69d71b3	56cb927a f9e99198	b9d5a578 d79f805e	57a7a5ee a63bb2e8	de748a3c 45dd8e31	dcc366b3 97e31fe5	b683a020 f713c240	3b2a5d9f a7b8cf69
M ₂	4d7a9c83 c69d71b3	d6cb927a f9e99198	29d5a578 d79f805e	57a7a5ee a63bb2e8	de748a3c 45dc8e31	dcc366b3 97e31fe5	b683a020 f713c240	3b2a5d9f a7b8cf69
H	e0f76122	c429c56c	ebb5e256	b809793				

Table 3 Two pairs of collisions for MD4

4 Collisions for RIPEMD

RIPEMD was developed for the RIPE project (RACE Integrity Primitives Evaluation, 1988-1992). In 1995, H. Dobbertin proved that the reduced version RIPEMD with two rounds is not collision-free[4]. We show

that the full RIPEMD also isn't collision-free. The following are two pairs of collisions for RIPEMD:

$$M'_i = M_i + \Delta C, \Delta C = (0,0,0,2^{20},0,0,0,0,0,2^{18} + 2^{31},0,0,0,2^{31})$$

M ₁	579faf8e bdeaae7	9ecf579 78bc91f2	574a6aba 47bc6d7d	78413511 9abdd1b1	a2b410a4 a45d2015	ad2f6c9f 817104ff	b56202c 264758a8	4d757911 61064ea5
M ₁	579faf8e bdeaae7	9ecf579 78bc91f2	574a6aba c7c06d7d	78513511 9abdd1b1	a2b410a4 a45d2015	ad2f6c9f 817104ff	b56202c 264758a8	4d757911 e1064ea5
H	1fab152	1654a31b	7a33776a	9e968ba7				
M ₂	579faf8e bdeaae7	9ecf579 78bc91f2	574a6aba 47bc6d7d	78413511 9abdd1b1	a2b410a4 a45d2015	ad2f6c9f a0a504ff	b56202c b18d58a8	4d757911 e70c66b6
M ₂	579faf8e bdeaae7	9ecf579 78bc91f2	574a6aba c7c06d7d	78513511 9abdd1b1	a2b410a4 a45d2015	ad2f6c9f a0a504ff	b56202c b18d58a8	4d757911 670c66b6
H	1f2c159f	569b31a6	dfcaa51a	25665d24				

Table 4 The collisions for RIPEMD

5 Remark

Besides the above hash functions we break, there are some other hash functions not having ideal security. For example, collision of SHA-0 [6] can be found with about 2^{40} computations of SHA-0 algorithms, and a collision for HAVAL-160 can be found with probability $1/2^{32}$.

Note that the messages and all other values in this paper are composed of 32-bit words, in each 32-bit word the most left byte is the most significant byte.

- 1 B. den Boer, Antoon Bosselaers, Collisions for the Compression Function of MD5, Eurocrypt'93.
- 2 H. Dobbertin, Cryptanalysis of MD4, Fast Software Encryption, LNCS 1039, D. , Springer-Verlag, 1996.
- 3 H. Dobbertin, Cryptanalysis of MD5 compress, presented at the rump session of Eurocrypt'96.
- 4 Hans Dobbertin, RIPEMD with Two-round Compress Function is Not Collision-Free, J. Cryptology 10(1), 1997.
- 5 H. Dobbertin, A. Bosselaers, B. Preneel, "RIPMEMD-160: A Strengthened Version of RIPMMD," Fast Software Encryption, LNCS 1039, D.Gollmann, Ed., Springer-Verlag, 1996, pp. 71-82.
- 6 FIPS 180-1, Secure hash standard, NIST, US Department of Commerce, Washington D. C., April 1995.
- 7 P. R. Kasselman, W T Penzhorn , Cryptanalysis of reduced version of HAVAL, Vol. 36, No. 1, Electronic Letters, 2000.
- 8 R. L. Rivest, The MD4 Message Digest Algorithm, Request for Comments (RFC)1320, Internet Activities Board, Internet Privacy Task Force, April 1992.
- 9 R. L Rivest, The MD5 Message Digest Algorithm, Request for Comments (RFC)1321, Internet Activities Board, Internet Privacy Task Force, April 1992.3RIPEMD-1281
- 10 Y. Zheng, J. Pieprzyk, J. Seberry, HAVAL--A One-way Hashing Algorithm with Variable Length of Output, Auscrypt'92.