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 \( IV'_0 \).

\[ IV'_0 : A'_0 = \text{0x12AC2375, } B'_0 = \text{0x3B341042, } C'_0 = \text{0x5F62B97C, } D'_0 = \text{0x4BA763ED} \]

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

\[ IV_0 : A_0 = \text{0x67452301, } B_0 = \text{0xefcdab89, } C_0 = \text{0x98badcfe, } D_0 = \text{0x10325476} \]

\[ M' = M + \Delta C_1, \Delta C_1 = (0,0,0,0,2^{31},...,2^{15},...,2^{31},0) \]

\[ N'_i = N_i + \Delta C_2, \Delta C_2 = (0,0,0,0,2^{31},...,2^{15},...,2^{31},0) \]

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.
### Table 1  Two pairs of collisions for MD5

![Table Image]

### 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

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

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

$$M' = M + \Delta C, \Delta C = (0.2^{31} - 2^{28} + 2^{31}, 0, 0, 0, 0, 0, 0, 0, 0,-2^{16} 0, 0, 0)$$

and $MD4(M) = MD4(M')$.

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 is not collection-free. The following are two pairs of collisions for RIPEMD:

\[ M_i = M_i + \Delta C, \Delta C = (0,0,0,2^{20},0,0,0,0,0,0,2^{18} + 2^{31},0,0,0,0,2^{31}) \]

<table>
<thead>
<tr>
<th>( M_1 )</th>
<th>579fa8e</th>
<th>9ecf579</th>
<th>574a6aba</th>
<th>78413511</th>
<th>a2b410a4</th>
<th>adf6c9f</th>
<th>b56202c</th>
<th>4d757911</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>bdeaa7</td>
<td>78bc91f2</td>
<td>47bc6d7d</td>
<td>9abdd1b1</td>
<td>a45d2015</td>
<td>817104ff</td>
<td>264758a8</td>
<td>61064ea5</td>
</tr>
<tr>
<td>( M_1 )</td>
<td>579fa8e</td>
<td>9ecf579</td>
<td>574a6aba</td>
<td>78513511</td>
<td>a2b410a4</td>
<td>adf6c9f</td>
<td>b56202c</td>
<td>4d757911</td>
</tr>
<tr>
<td></td>
<td>bdeaa7</td>
<td>78bc91f2</td>
<td>c7c06d7d</td>
<td>9abdd1b1</td>
<td>a45d2015</td>
<td>817104ff</td>
<td>264758a8</td>
<td>e1064ea5</td>
</tr>
<tr>
<td>( H )</td>
<td>1fab152</td>
<td>165a31b</td>
<td>7a33376a</td>
<td>9e968ba7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( M_2 )</td>
<td>579fa8e</td>
<td>9ecf579</td>
<td>574a6aba</td>
<td>78413511</td>
<td>a2b410a4</td>
<td>adf6c9f</td>
<td>b56202c</td>
<td>4d757911</td>
</tr>
<tr>
<td></td>
<td>bdeaa7</td>
<td>78bc91f2</td>
<td>47bc6d7d</td>
<td>9abdd1b1</td>
<td>a45d2015</td>
<td>a0a504ff</td>
<td>b18d58a8</td>
<td>e70c66b6</td>
</tr>
<tr>
<td>( M_2 )</td>
<td>579fa8e</td>
<td>9ecf579</td>
<td>574a6aba</td>
<td>78513511</td>
<td>a2b410a4</td>
<td>adf6c9f</td>
<td>b56202c</td>
<td>4d757911</td>
</tr>
<tr>
<td></td>
<td>bdeaa7</td>
<td>78bc91f2</td>
<td>c7c06d7d</td>
<td>9abdd1b1</td>
<td>a45d2015</td>
<td>a0a504ff</td>
<td>b18d58a8</td>
<td>670c66b6</td>
</tr>
<tr>
<td>( H )</td>
<td>1f2c159f</td>
<td>569b31a6</td>
<td>dfca51a</td>
<td>25665d24</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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.
3. H. Dobbertin, Cryptanalysis of MD5 compress, presented at the rump session of Eurocrypt'96.