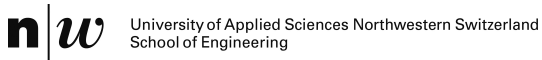


# Preimages of HAVAL and MD5

Jean-Philippe Aumasson



(Joint work with Willi Meier and Florian Mendel)

# Definition

“Given  $y$ , find  $x$  such that  $h(x) = y$ ”

Given what?

1. **random range** element
2. **random image**
3. **any fixed** value (“weak” images)

Attacks here work for 1, 2, 3

# Previous preimage attacks

Google (*collision attack*)  $\approx 500\,000$

Google (*preimage attack*)  $\approx 15\,000$

- ▶ MD2 (2004)
- ▶ Parallel FFT-Hashing (2007)
- ▶ Syndrome-based hash (2007)
- ▶ HAS-V (2007)
- ▶ Tiger (2007)
- ▶ MD4 (2008)
- ▶ GOST (2008)
- ▶ Snefru (2008)
- ▶ SHA-0/1 (2008)

# Content of this talk

**MD5**: how to invert. . .

- ▶ 22 steps?
- ▶ 31 steps?
- ▶ 47 steps?

What about **HAVAL**?

From the **compression to hash** function

# MD5

- ▶ 1991: publication (Rivest)
- ▶ 1993: **collision** attack (compression function)
- ▶ 2005: **collision** attack (hash function)
- ▶ 2005+: faster, chosen-prefix, meaningful **collisions**

No preimage attack.



# MD5 compression function

## Input

- ▶ chain value  $H_0H_1H_2H_3$
- ▶ message  $M_0M_1 \dots M_{15}$

## Algorithm

- ▶ copy  $H_0H_1H_2H_3$  into  $A_0B_0C_0D_0$
- ▶ for  $i = 1 \dots 64$

$$A_i = D_{i-1}$$

$$B_i = f_i(A_{i-1}, B_{i-1}, C_{i-1}, D_{i-1}, M_{\sigma(i)})$$

$$C_i = B_{i-1}$$

$$D_i = C_{i-1}$$

- ▶ return  $(A_0 + A_{64})\|(B_0 + B_{64})\|(C_0 + C_{64})\|(D_0 + D_{64})$

# Unrolling (**index**, function, message word)

|           |  |           |  |
|-----------|--|-----------|--|
| <b>1</b>  | $f_1(A_0, B_0, C_0, D_0, 0)$                 | <b>17</b> | $f_{17}(A_{16}, B_{16}, C_{16}, D_{16}, 1)$  |
| <b>2</b>  | $f_2(A_1, B_1, C_1, D_1, 1)$                 | <b>18</b> | $f_{18}(A_{17}, B_{17}, C_{17}, D_{17}, 6)$  |
| <b>3</b>  | $f_3(A_2, B_2, C_2, D_2, 2)$                 | <b>19</b> | $f_{19}(A_{18}, B_{18}, C_{18}, D_{18}, 11)$ |
| <b>4</b>  | $f_4(A_3, B_3, C_3, D_3, 3)$                 | <b>20</b> | $f_{20}(A_{19}, B_{19}, C_{19}, D_{19}, 0)$  |
| <b>5</b>  | $f_5(A_4, B_4, C_4, D_4, 4)$                 | <b>21</b> | $f_{21}(A_{20}, B_{20}, C_{20}, D_{20}, 5)$  |
| <b>6</b>  | $f_6(A_5, B_5, C_5, D_5, 5)$                 | <b>22</b> | $f_{22}(A_{21}, B_{21}, C_{21}, D_{21}, 10)$ |
| <b>7</b>  | $f_7(A_6, B_6, C_6, D_6, 6)$                 | <b>23</b> | $f_{23}(A_{22}, B_{22}, C_{22}, D_{22}, 15)$ |
| <b>8</b>  | $f_8(A_7, B_7, C_7, D_7, 7)$                 | <b>24</b> | $f_{24}(A_{23}, B_{23}, C_{23}, D_{23}, 4)$  |
| <b>9</b>  | $f_9(A_8, B_8, C_8, D_8, 8)$                 | <b>25</b> | $f_{25}(A_{24}, B_{24}, C_{24}, D_{24}, 9)$  |
| <b>10</b> | $f_{10}(A_9, B_9, C_9, D_9, 9)$              | <b>26</b> | $f_{26}(A_{25}, B_{25}, C_{25}, D_{25}, 14)$ |
| <b>11</b> | $f_{11}(A_{10}, B_{10}, C_{10}, D_{10}, 10)$ | <b>27</b> | $f_{27}(A_{26}, B_{26}, C_{26}, D_{26}, 3)$  |
| <b>12</b> | $f_{12}(A_{11}, B_{11}, C_{11}, D_{11}, 11)$ | <b>28</b> | $f_{28}(A_{27}, B_{27}, C_{27}, D_{27}, 8)$  |
| <b>13</b> | $f_{13}(A_{12}, B_{12}, C_{12}, D_{12}, 12)$ | <b>29</b> | $f_{29}(A_{28}, B_{28}, C_{28}, D_{28}, 13)$ |
| <b>14</b> | $f_{14}(A_{13}, B_{13}, C_{13}, D_{13}, 13)$ | <b>30</b> | $f_{30}(A_{29}, B_{29}, C_{29}, D_{29}, 2)$  |
| <b>15</b> | $f_{15}(A_{14}, B_{14}, C_{14}, D_{14}, 14)$ | <b>31</b> | $f_{31}(A_{30}, B_{30}, C_{30}, D_{30}, 7)$  |
| <b>16</b> | $f_{16}(A_{15}, B_{15}, C_{15}, D_{15}, 15)$ | <b>32</b> | $f_{32}(A_{31}, B_{31}, C_{31}, D_{31}, 12)$ |



# First 2 rounds

- |    |   |    |   |
|----|---|----|---|
| 1  | $f(A_0, B_0, C_0, D_0, 0)$              | 17 | $g(A_{16}, B_{16}, C_{16}, D_{16}, 1)$  |
| 2  | $f(A_1, B_1, C_1, D_1, 1)$              | 18 | $g(A_{17}, B_{17}, C_{17}, D_{17}, 6)$  |
| 3  | $f(A_2, B_2, C_2, D_2, 2)$              | 19 | $g(A_{18}, B_{18}, C_{18}, D_{18}, 11)$ |
| 4  | $f(A_3, B_3, C_3, D_3, 3)$              | 20 | $g(A_{19}, B_{19}, C_{19}, D_{19}, 0)$  |
| 5  | $f(A_4, B_4, C_4, D_4, 4)$              | 21 | $g(A_{20}, B_{20}, C_{20}, D_{20}, 5)$  |
| 6  | $f(A_5, B_5, C_5, D_5, 5)$              | 22 | $g(A_{21}, B_{21}, C_{21}, D_{21}, 10)$ |
| 7  | $f(A_6, B_6, C_6, D_6, 6)$              | 23 | $g(A_{22}, B_{22}, C_{22}, D_{22}, 15)$ |
| 8  | $f(A_7, B_7, C_7, D_7, 7)$              | 24 | $g(A_{23}, B_{23}, C_{23}, D_{23}, 4)$  |
| 9  | $f(A_8, B_8, C_8, D_8, 8)$              | 25 | $g(A_{24}, B_{24}, C_{24}, D_{24}, 9)$  |
| 10 | $f(A_9, B_9, C_9, D_9, 9)$              | 26 | $g(A_{25}, B_{25}, C_{25}, D_{25}, 14)$ |
| 11 | $f(A_{10}, B_{10}, C_{10}, D_{10}, 10)$ | 27 | $g(A_{26}, B_{26}, C_{26}, D_{26}, 3)$  |
| 12 | $f(A_{11}, B_{11}, C_{11}, D_{11}, 11)$ | 28 | $g(A_{27}, B_{27}, C_{27}, D_{27}, 8)$  |
| 13 | $f(A_{12}, B_{12}, C_{12}, D_{12}, 12)$ | 29 | $g(A_{28}, B_{28}, C_{28}, D_{28}, 13)$ |
| 14 | $f(A_{13}, B_{13}, C_{13}, D_{13}, 13)$ | 30 | $g(A_{29}, B_{29}, C_{29}, D_{29}, 2)$  |
| 15 | $f(A_{14}, B_{14}, C_{14}, D_{14}, 14)$ | 31 | $g(A_{30}, B_{30}, C_{30}, D_{30}, 7)$  |
| 16 | $f(A_{15}, B_{15}, C_{15}, D_{15}, 15)$ | 32 | $g(A_{31}, B_{31}, C_{31}, D_{31}, 12)$ |

# Inverting 22 steps

- |           |   |           |   |
|-----------|---|-----------|---|
| <b>1</b>  | $f(A_0, B_0, C_0, D_0, 0)$              | <b>17</b> | $g(A_{16}, B_{16}, C_{16}, D_{16}, 1)$  |
| <b>2</b>  | $f(A_1, B_1, C_1, D_1, 1)$              | <b>18</b> | $g(A_{17}, B_{17}, C_{17}, D_{17}, 6)$  |
| <b>3</b>  | $f(A_2, B_2, C_2, D_2, 2)$              | <b>19</b> | $g(A_{18}, B_{18}, C_{18}, D_{18}, 11)$ |
| <b>4</b>  | $f(A_3, B_3, C_3, D_3, 3)$              | <b>20</b> | $g(A_{19}, B_{19}, C_{19}, D_{19}, 0)$  |
| <b>5</b>  | $f(A_4, B_4, C_4, D_4, 4)$              | <b>21</b> | $g(A_{20}, B_{20}, C_{20}, D_{20}, 5)$  |
| <b>6</b>  | $f(A_5, B_5, C_5, D_5, 5)$              | <b>22</b> | $g(A_{21}, B_{21}, C_{21}, D_{21}, 10)$ |
| <b>7</b>  | $f(A_6, B_6, C_6, D_6, 6)$              |           |   |
| <b>8</b>  | $f(A_7, B_7, C_7, D_7, 7)$              |           |   |
| <b>9</b>  | $f(A_8, B_8, C_8, D_8, 8)$              |           |   |
| <b>10</b> | $f(A_9, B_9, C_9, D_9, 9)$              |           |   |
| <b>11</b> | $f(A_{10}, B_{10}, C_{10}, D_{10}, 10)$ |           |   |
| <b>12</b> | $f(A_{11}, B_{11}, C_{11}, D_{11}, 11)$ |           |   |
| <b>13</b> | $f(A_{12}, B_{12}, C_{12}, D_{12}, 12)$ |           |   |
| <b>14</b> | $f(A_{13}, B_{13}, C_{13}, D_{13}, 13)$ |           |   |
| <b>15</b> | $f(A_{14}, B_{14}, C_{14}, D_{14}, 14)$ |           |   |
| <b>16</b> | $f(A_{15}, B_{15}, C_{15}, D_{15}, 15)$ |           |   |

# Inverting 22 steps

Pick  $M_0 \dots M_{11}$

From **1** to **12**, compute  $A_{12} B_{12} C_{12} D_{12}$

From **22** to **17**, compute  $A_{16} B_{16} C_{16} D_{16}$

Choose  $M_{12}$  such that  $B_{13} = A_{16}$

Choose  $M_{13}$  such that  $B_{14} = D_{16}$

Choose  $M_{14}$  such that  $B_{15} = C_{16}$

Choose  $M_{15}$  such that  $B_{16} = B_{16}$

**Cost:** 22 steps

# 31 steps: same idea...

- |    |   |    |   |
|----|---|----|---|
| 1  | $f(A_0, B_0, C_0, D_0, 0)$              | 17 | $g(A_{16}, B_{16}, C_{16}, D_{16}, 1)$  |
| 2  | $f(A_1, B_1, C_1, D_1, 1)$              | 18 | $g(A_{17}, B_{17}, C_{17}, D_{17}, 6)$  |
| 3  | $f(A_2, B_2, C_2, D_2, 2)$              | 19 | $g(A_{18}, B_{18}, C_{18}, D_{18}, 11)$ |
| 4  | $f(A_3, B_3, C_3, D_3, 3)$              | 20 | $g(A_{19}, B_{19}, C_{19}, D_{19}, 0)$  |
| 5  | $f(A_4, B_4, C_4, D_4, 4)$              | 21 | $g(A_{20}, B_{20}, C_{20}, D_{20}, 5)$  |
| 6  | $f(A_5, B_5, C_5, D_5, 5)$              | 22 | $g(A_{21}, B_{21}, C_{21}, D_{21}, 10)$ |
| 7  | $f(A_6, B_6, C_6, D_6, 6)$              | 23 | $g(A_{22}, B_{22}, C_{22}, D_{22}, 15)$ |
| 8  | $f(A_7, B_7, C_7, D_7, 7)$              | 24 | $g(A_{23}, B_{23}, C_{23}, D_{23}, 4)$  |
| 9  | $f(A_8, B_8, C_8, D_8, 8)$              | 25 | $g(A_{24}, B_{24}, C_{24}, D_{24}, 9)$  |
| 10 | $f(A_9, B_9, C_9, D_9, 9)$              | 26 | $g(A_{25}, B_{25}, C_{25}, D_{25}, 14)$ |
| 11 | $f(A_{10}, B_{10}, C_{10}, D_{10}, 10)$ | 27 | $g(A_{26}, B_{26}, C_{26}, D_{26}, 3)$  |
| 12 | $f(A_{11}, B_{11}, C_{11}, D_{11}, 11)$ | 28 | $g(A_{27}, B_{27}, C_{27}, D_{27}, 8)$  |
| 13 | $f(A_{12}, B_{12}, C_{12}, D_{12}, 12)$ | 29 | $g(A_{28}, B_{28}, C_{28}, D_{28}, 13)$ |
| 14 | $f(A_{13}, B_{13}, C_{13}, D_{13}, 13)$ | 30 | $g(A_{29}, B_{29}, C_{29}, D_{29}, 2)$  |
| 15 | $f(A_{14}, B_{14}, C_{14}, D_{14}, 14)$ | 31 | $g(A_{30}, B_{30}, C_{30}, D_{30}, 7)$  |
| 16 | $f(A_{15}, B_{15}, C_{15}, D_{15}, 15)$ |    |   |

# $M_{12}$ input only once. . .

- |    |   |    |   |
|----|---|----|---|
| 1  | $f(A_0, B_0, C_0, D_0, 0)$              | 17 | $g(A_{16}, B_{16}, C_{16}, D_{16}, 1)$  |
| 2  | $f(A_1, B_1, C_1, D_1, 1)$              | 18 | $g(A_{17}, B_{17}, C_{17}, D_{17}, 6)$  |
| 3  | $f(A_2, B_2, C_2, D_2, 2)$              | 19 | $g(A_{18}, B_{18}, C_{18}, D_{18}, 11)$ |
| 4  | $f(A_3, B_3, C_3, D_3, 3)$              | 20 | $g(A_{19}, B_{19}, C_{19}, D_{19}, 0)$  |
| 5  | $f(A_4, B_4, C_4, D_4, 4)$              | 21 | $g(A_{20}, B_{20}, C_{20}, D_{20}, 5)$  |
| 6  | $f(A_5, B_5, C_5, D_5, 5)$              | 22 | $g(A_{21}, B_{21}, C_{21}, D_{21}, 10)$ |
| 7  | $f(A_6, B_6, C_6, D_6, 6)$              | 23 | $g(A_{22}, B_{22}, C_{22}, D_{22}, 15)$ |
| 8  | $f(A_7, B_7, C_7, D_7, 7)$              | 24 | $g(A_{23}, B_{23}, C_{23}, D_{23}, 4)$  |
| 9  | $f(A_8, B_8, C_8, D_8, 8)$              | 25 | $g(A_{24}, B_{24}, C_{24}, D_{24}, 9)$  |
| 10 | $f(A_9, B_9, C_9, D_9, 9)$              | 26 | $g(A_{25}, B_{25}, C_{25}, D_{25}, 14)$ |
| 11 | $f(A_{10}, B_{10}, C_{10}, D_{10}, 10)$ | 27 | $g(A_{26}, B_{26}, C_{26}, D_{26}, 3)$  |
| 12 | $f(A_{11}, B_{11}, C_{11}, D_{11}, 11)$ | 28 | $g(A_{27}, B_{27}, C_{27}, D_{27}, 8)$  |
| 13 | $f(A_{12}, B_{12}, C_{12}, D_{12}, 12)$ | 29 | $g(A_{28}, B_{28}, C_{28}, D_{28}, 13)$ |
| 14 | $f(A_{13}, B_{13}, C_{13}, D_{13}, 13)$ | 30 | $g(A_{29}, B_{29}, C_{29}, D_{29}, 2)$  |
| 15 | $f(A_{14}, B_{14}, C_{14}, D_{14}, 14)$ | 31 | $g(A_{30}, B_{30}, C_{30}, D_{30}, 7)$  |
| 16 | $f(A_{15}, B_{15}, C_{15}, D_{15}, 15)$ |    |   |

# 31 steps

Pick  $M_0 \dots M_{11}, M_{13}, M_{14}, M_{15}$

From **1** to **12**, compute  $A_{12}B_{12}C_{12}D_{12}$

From **22** to **14**, compute  $A_{13}B_{13}C_{13}D_{13}$

If  $A_{13} = D_{12}$ ,  $C_{13} = B_{12}$ , and  $D_{13} = C_{12}$ :

then choose  $M_{12}$  such that  $B_{13} = B_{12}$

**Cost:**  $\approx 2^{96} \times 31$  steps

# 47 steps: $M_2$ input only twice...

|    |                |    |                |    |                |
|----|----------------|----|----------------|----|----------------|
| 1  | $f(\dots, 0)$  | 17 | $g(\dots, 1)$  | 33 | $h(\dots, 5)$  |
| 2  | $f(\dots, 1)$  | 18 | $g(\dots, 6)$  | 34 | $h(\dots, 8)$  |
| 3  | $f(\dots, 2)$  | 19 | $g(\dots, 11)$ | 35 | $h(\dots, 11)$ |
| 4  | $f(\dots, 3)$  | 20 | $g(\dots, 0)$  | 36 | $h(\dots, 14)$ |
| 5  | $f(\dots, 4)$  | 21 | $g(\dots, 5)$  | 37 | $h(\dots, 1)$  |
| 6  | $f(\dots, 5)$  | 22 | $g(\dots, 10)$ | 38 | $h(\dots, 4)$  |
| 7  | $f(\dots, 6)$  | 23 | $g(\dots, 15)$ | 39 | $h(\dots, 7)$  |
| 8  | $f(\dots, 7)$  | 24 | $g(\dots, 4)$  | 40 | $h(\dots, 10)$ |
| 9  | $f(\dots, 8)$  | 25 | $g(\dots, 9)$  | 41 | $h(\dots, 13)$ |
| 10 | $f(\dots, 9)$  | 26 | $g(\dots, 14)$ | 42 | $h(\dots, 0)$  |
| 11 | $f(\dots, 10)$ | 27 | $g(\dots, 3)$  | 43 | $h(\dots, 3)$  |
| 12 | $f(\dots, 11)$ | 28 | $g(\dots, 8)$  | 44 | $h(\dots, 6)$  |
| 13 | $f(\dots, 12)$ | 29 | $g(\dots, 13)$ | 45 | $h(\dots, 9)$  |
| 14 | $f(\dots, 13)$ | 30 | $g(\dots, 2)$  | 46 | $h(\dots, 12)$ |
| 15 | $f(\dots, 14)$ | 31 | $g(\dots, 7)$  | 47 | $h(\dots, 15)$ |
| 16 | $f(\dots, 15)$ | 32 | $g(\dots, 12)$ |    |                |

# Differences propagation, general case

Pick random  $A_0, B_0, C_0, D_0$  and  $M$

$$1 \quad f(A_0, B_0, C_0, D_0, 0)$$

$$2 \quad f(A_1, B_1, C_1, D_1, 1)$$

$$3 \quad f(A_2, B_2, C_2, D_2, 2)$$

Modify  $C_0$  to  $C_0^*$

$$X \quad 1 \quad f(A_0, B_0, C_0^*, D_0, 0)$$

$$X \quad 2 \quad f(A_1, B_1, C_1, C_0^*, 1)$$

$$X \quad 3 \quad f(C_0^*, B_2, C_2, D_2, 2)$$

$\Rightarrow$  all first steps affected ( $X$ =state modified)



## Difference in $C_0$ + chosen IV

Pick random  $A_0, C_0, D_0$  and  $M$  and set  $B_0 = 0$

$$1 \quad f(A_0, B_0, C_0, D_0, 0)$$

$$2 \quad f(A_1, B_1, C_1, D_1, 1)$$

$$3 \quad f(A_2, B_2, C_2, D_2, 2)$$

Modify  $C_0$  to  $C_0^*$

$$\checkmark \quad 1 \quad f(A_0, 0, C_0^*, D_0, 0)$$

$$\checkmark \quad 2 \quad f(A_1, B_1, 0, C_0^*, 1)$$

$$X \quad 3 \quad f(C_0^*, B_2, C_2, 0, 2)$$

$\Rightarrow$  only step 3 affected

## Difference in $M_2$

Pick random  $A_0, B_0, C_0, D_0$  and  $M$

$$1 \quad f(A_0, B_0, C_0, D_0, 0)$$

$$2 \quad f(A_1, B_1, C_1, D_1, 1)$$

$$3 \quad f(A_2, B_2, C_2, D_2, 2)$$

Modify  $M_2$

$$\checkmark \quad 1 \quad f(A_0, B_0, C_0, D_0, 0)$$

$$\checkmark \quad 2 \quad f(A_1, B_1, C_1, C_0, 1)$$

$$X \quad 3 \quad f(A_2, B_2, C_2, D_2, 2)$$

$\Rightarrow$  only step 3 affected

# Absorbing differences

Pick random  $A_0, C_0, D_0$  and  $M$  and set  $B_0 = 0$

$$1 \quad f(A_0, B_0, C_0, D_0, 0)$$

$$2 \quad f(A_1, B_1, C_1, D_1, 1)$$

$$3 \quad f(A_2, B_2, C_2, D_2, 2)$$

Modify  $C_0$  to  $C_0^*$  and  $M_2$

$$\checkmark \quad 1 \quad f(A_0, 0, C_0^*, D_0, 0)$$

$$\checkmark \quad 2 \quad f(A_1, B_1, 0, C_0^*, 1)$$

$$\checkmark \quad 3 \quad f(C_0^*, B_2, C_2, 0, 2)$$

$\Rightarrow$  nothing changes!

# Application to 47-step MD5: key steps

|           |                |           |                |           |                |
|-----------|----------------|-----------|----------------|-----------|----------------|
| <b>1</b>  | $f(\dots, 0)$  | <b>17</b> | $g(\dots, 1)$  | <b>33</b> | $h(\dots, 5)$  |
| <b>2</b>  | $f(\dots, 1)$  | <b>18</b> | $g(\dots, 6)$  | <b>34</b> | $h(\dots, 8)$  |
| <b>3</b>  | $f(\dots, 2)$  | <b>19</b> | $g(\dots, 11)$ | <b>35</b> | $h(\dots, 11)$ |
| <b>4</b>  | $f(\dots, 3)$  | <b>20</b> | $g(\dots, 0)$  | <b>36</b> | $h(\dots, 14)$ |
| <b>5</b>  | $f(\dots, 4)$  | <b>21</b> | $g(\dots, 5)$  | <b>37</b> | $h(\dots, 1)$  |
| <b>6</b>  | $f(\dots, 5)$  | <b>22</b> | $g(\dots, 10)$ | <b>38</b> | $h(\dots, 4)$  |
| <b>7</b>  | $f(\dots, 6)$  | <b>23</b> | $g(\dots, 15)$ | <b>39</b> | $h(\dots, 7)$  |
| <b>8</b>  | $f(\dots, 7)$  | <b>24</b> | $g(\dots, 4)$  | <b>40</b> | $h(\dots, 10)$ |
| <b>9</b>  | $f(\dots, 8)$  | <b>25</b> | $g(\dots, 9)$  | <b>41</b> | $h(\dots, 13)$ |
| <b>10</b> | $f(\dots, 9)$  | <b>26</b> | $g(\dots, 14)$ | <b>42</b> | $h(\dots, 0)$  |
| <b>11</b> | $f(\dots, 10)$ | <b>27</b> | $g(\dots, 3)$  | <b>43</b> | $h(\dots, 3)$  |
| <b>12</b> | $f(\dots, 11)$ | <b>28</b> | $g(\dots, 8)$  | <b>44</b> | $h(\dots, 6)$  |
| <b>13</b> | $f(\dots, 12)$ | <b>29</b> | $g(\dots, 13)$ | <b>45</b> | $h(\dots, 9)$  |
| <b>14</b> | $f(\dots, 13)$ | <b>30</b> | $g(\dots, 2)$  | <b>46</b> | $h(\dots, 12)$ |
| <b>15</b> | $f(\dots, 14)$ | <b>31</b> | $g(\dots, 7)$  | <b>47</b> | $h(\dots, 15)$ |
| <b>16</b> | $f(\dots, 15)$ | <b>32</b> | $g(\dots, 12)$ |           |                |

# The attack

## Stage 1: MITM

Pick  $M$  and IV with  $B_0 = 0$ ,

1. store  $(A_{29}, B_{29}, C_{29}, D_{29})$  for all  $2^{32}$   $C_0$ 's (forward)
2. store  $(A_{30}, B_{30}, C_{30}, D_{30})$  for all  $2^{32}$   $C_{47}$ 's (backward)

Find entries such that

$$A_{30} = D_{29}$$

$$D_{30} = C_{29}$$

$$C_{30} = B_{29}$$

$\equiv$  96-bit equality;  $2^{64}$  choices  $\Rightarrow$  repeat  $2^{32}$  times

# The attack

**Stage 2:** correction

Modify  $M_2$  such that

$$B_{30} = g(A_{29}, B_{29}, C_{29}, D_{29}, 2)$$

Modify  $C_0$  accordingly

$\Rightarrow$  96-bit preimage ( $C_0 + C_{47}$  is random) with prob.  $2^{-32}$

**Total cost:**  $2^{96}$  trials for a 128-bit preimage

# Summary

Preimages for MD5's **compression function** with

- ▶ chosen message except  $M_1$  and  $M_2$
- ▶ IV with  $B_0 = 0$  and random  $C_0$
- ▶ storage for  $2^{36}$  bytes (64 Gb)
- ▶  $2^{96}$  compressions

By bruteforce:

- ▶ random message
- ▶ chosen IV
- ▶ negligible memory
- ▶  $2^{128}$  compressions

# HAVAL

- ▶ 1992: publication (Zheng, Pieprzyk, Seberry)
- ▶ 2003: **collision** attack (3-pass)
- ▶ 2006: **collision** attack (4- and 5-pass)
- ▶ 2008: (partial) **second-preimage** attack (3-pass)

Function similar to MD5 with

- ▶ 256-bit chain values
- ▶ 1024-bit blocks
- ▶ 3, 4, or 5 rounds



# Preimages for 3-pass HAVAL

Same strategy as for MD5

- ▶ identify **absorption** properties in the initial steps
- ▶ MITM
- ▶ modify a  $M_i$  to complete the MITM
- ▶ correct initial steps

⇒ 2 attacks in  $2^{224}$  and storage of  $2^{69}$  bytes  
(vs.  $2^{256}$  and negligible memory)

# Extension to the iterated hash

Attacks presented for the compression function  
(with IV partially random, no padding)

Restrictions for the hash function:

- ▶ **padding**: not a problem, because message chosen
- ▶ **fixed IV**: makes direct application impossible

# Iterated hash: basic MITM

Given image **H**:

1. compute a **list of images** from the fixed IV

$$(M_i, \text{compress}(\text{IV}, M_i))_i$$

2. compute a **list of preimages**

$$(H_i, M'_i)_i, \text{compress}(H_i, M'_i) = \mathbf{H}$$

Find entries such that

$$\text{compress}(\text{IV}, M_i) = H_j$$

**Cost:**  $2^{113}$  trials +  $2^{36}$  bytes for MD5,  $2^{241}$  +  $2^{69}$  for HAVAL

# Iterated hash: tree technique

Mendel & Rijmen (ICISC'07), Leurent (FSE'08)

Build a tree using **multi-target preimages**

**Cost:**  $2^{102}$  trials +  $2^{39}$  bytes for MD5,  $2^{230} + 2^{71}$  for HAVAL

# Conclusion

|              | compressions | bytes    |
|--------------|--------------|----------|
| 47-step MD5  | $2^{102}$    | $2^{39}$ |
| 3-pass HAVAL | $2^{230}$    | $2^{71}$ |

Independent work by Sasaki & Aoki @ACISP 2008:  
44-step MD5 (starting from step 6) in  $2^{96}$

# Questions?

Is it effectively faster than bruteforce?

→ Arguably yes (but not  $2^{26}$  times faster)

Same strategy applies to MD4?

→ No (because no  $M_i$  at very start and very end)

Same strategy applies to SHA-0/1/2?

→ No (nontrivial message expansion), cf. next talk