NTNU
Norwegian University of
Science and Technology

Distinguishers for the Compression Function and Output Transformation of Hamsi-256

Jean-Philippe Aumasson   Emilia Käsper
Lars Ramkilde Knudsen   Krystian Matusiewicz
Rune Ødegård   Thomas Peyrin   Martin Schläffer

ACISP, Sydney, Australia - 2010-07-05

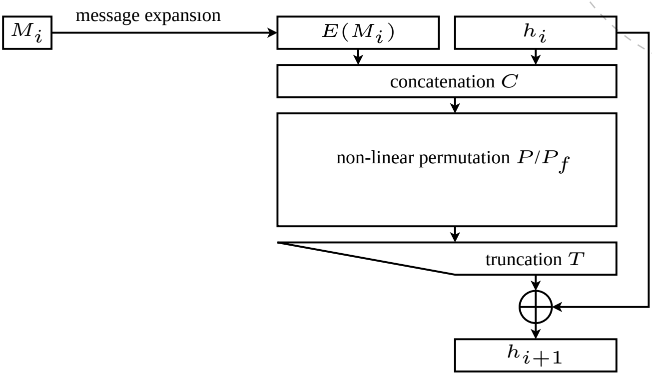
Overview

NIST Hash Competition

- Collision attacks on the deployed standards MD5 and SHA-1 [WLF+05, WY05, WYY05b, WYY05a] have weakened the confidence in the MD family of hash functions.
- The US Institute of Standards and Technology (NIST) launched a public competition to develop a future SHA-3 standard [NIS07].
- The hash function Hamsi [Kü09a] is one of 64 designs submitted to NIST in fall 2008.
- Hamsi is also one of the 14 submissions selected for the second round of the competition.

Overview

Hamsi domain extension algorithm



Message expansion

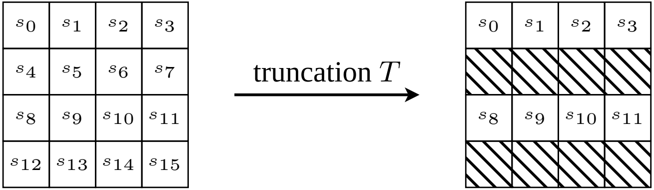
- The message expansion of Hamsi uses a linear code to expand a 32-bit word into eight words (that is, 256 bits).
- The minimum distance of the code is 83.

Concatination

$$(m_0, m_1, \ldots, m_7,\ c_0, c_1, \ldots, c_7) \xrightarrow{\text{concatenation } C}$$

m_0	m_1	c_0	c_1
c_2	c_3	m_2	m_3
m_4	m_5	c_4	c_5
c_6	c_7	m_6	m_7

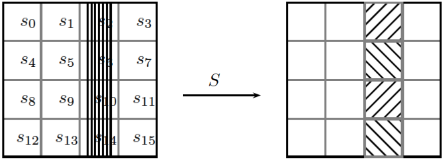
Truncation


s_0 s_1 s_2 s_3
s_4 s_5 s_6 s_7
s_8 s_9 s_{10} s_{11}
s_{12} s_{13} s_{14} s_{15}
truncation T
s_0 s_1 s_2 s_3
s_8 s_9 s_{10} s_{11}

Permutations

- The permutations P and P_f only differ in the number of rounds (3 for P and 6 for P_f), and the constants used.
- The round function is composed of three layers.
- First, constants and a counter are XORed to the whole internal state.
- Then there is a substitution layer.
- Followed by a linear layer.

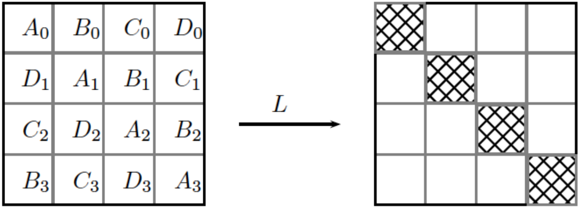
Permutation - Substitution layer



The substitution layer uses one 4-bit Sbox of the block cipher Serpent [BAK98], in a bitsliced way.

x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$S[x]$	8	6	7	9	3	12	10	15	13	1	14	4	0	11	5	2

Permutation - Linear layer.



$$
\begin{aligned}
a &:= a \lll 13 \\
c &:= c \lll 3 \\
b &:= a \oplus b \oplus c \\
d &:= (a \ll 3) \oplus c \oplus d \\
b &:= b \lll 1
\end{aligned}
$$

$$
\begin{aligned}
d &:= d \lll 7 \\
a &:= a \oplus b \oplus d \\
c &:= (b \ll 7) \oplus c \oplus d \\
a &:= a \lll 5 \\
c &:= c \lll 22
\end{aligned}
$$

Overview

Definitions

k-**sum** problem for Hamsi's compression function f

Find $x_1, \ldots, x_k$ strings of n bits such that

$$\bigoplus_{i=1}^{k} f(x_k) = 0$$

Zero-sum problem: additional requirement that $\bigoplus_{i=1}^{k} x_i = 0$
Generic method: generalized birthday in $O(k2^{n/(1+\log k)})$ [Wag02]
XHASH attack (linear algebra) for $k \approx n$ [BM97]

Observation

- The only nonlinear component of Hamsi's compression function is the layer of Sboxes.
- One round thus has degree 3.
- N rounds have degree at most 3^N, with respect to any choice of variables.
- With carefully chosen variables we can make the first round linear and the degree at most 3^{N-1} after N rounds.
- This means that the degree is at most 81 after five rounds, and that at least six rounds are necessary to reach maximal degree.

Finding *k*-sums and zero sums

- For randomly chosen 256-bit values, finding 4-sums for the compression function of Hamsi requires an effort of complexity approximately 2^{87}, using the generalized birthday method.
- Because of the low algebraic degree of Hamsi we are able to find 16-sums, 8-sums and 4-sums efficiently
- Examples found for the IV specified in [Kö9b]
- Based on the work of [Wag99, §9], we show how to find large zero sums efficiently by exploting the fact that two *halves* of Hamsi's permutation have low algebraic degree

Overview

Differential properties of the Sbox

- About half the differential transitions are impossible
- The probabilities of the non-zero differentials are either 2^{-2} or 2^{-3}.
- We construct high-probability differential paths by
  1. keeping the overall number of active Sboxes low and
  2. avoiding probability 2^{-3} differentials where possible

Differential properties of the Linear transform

- Each bit of L contributes to one of the 128 Sboxes in each round.
- To minimize the number of active Sboxes, we thus need to minimize number of differences in L.
- If we introduce a single input difference at bit position in one input word, the HW of the output differences depends on the position and word of the input difference.
- We observe that for some specific words and bit positions, the resulting HW can be quite small. This happens if one or more differences are removed by the shift operation.

Near collision

- Using our observations on the differential properties of Hamsi's Sbox and linear transform we searched manually for high-probability paths leading to near-collisions for the compression function.
- Nikolic reported near collisions [Nik09] on $(256 - 25)$ bits with 14 differences in the chaining value.
- Wang et al. reported [WWJW09] near collisions on $(256 - 23)$ bits with 16 differences in the chaining value.
- We found a $(256 - 25)$-bit collision from 6 bit differences.

Automated differential path search

- We searched for differential paths with some difference in the input and output chaining value, and no difference in the input message.
- For this purpose we constructed an automated differential path randomized search algorithm.
- The primary heuristic is to minimise the HW of the differences in each round.
- Full details of the path search are in the paper.

It.	Sbox input				Sbox output				Prob.
start					00000000	00000000	84004880	4081C400	
					2C020018	000045C0	00000000	00000000	
					00000000	00000000	84024880	4081C400	
					28020018	000045C0	00000000	00000000	
1	00000000	00000000	84004880	4081C400	04000000	00000000	04000000	40818000	(58)
	2C020018	000045C0	00000000	00000000	28020018	000040C0	04020000	00000000	
	00000000	00000000	84024880	4081C400	00000018	00004100	00000800	00804000	
	28020018	000045C0	00000000	00000000	04020000	000004C0	80024880	00004400	
2	00000000	00000000	00000000	00010000	00000000	00000000	00000000	00010000	17
	30000010	00000080	00000000	00000080	30000000	00000000	00000000	00000080	
	30000010	00000080	00000000	00010080	00000010	00000000	00000000	00000000	
	00000000	00000000	00000000	00000000	00000000	00000080	00000000	00000000	
3	00000000	00000000	00000000	00000000	00000000	00000000	00000000	00000000	3
	20000000	00000000	00000000	00000000	20000000	00000000	00000000	00000000	
	20000000	00000000	00000000	00000000	00000000	00000000	00000000	00000000	
	00000000	00000000	00000000	00000000	00000000	00000000	00000000	00000000	
4	00000000	00000000	00000000	00000008	40000000	00000000	00000000	00000000	5
	40000000	00000000	00000000	00000000	40000000	00000000	00000000	00000008	
	00000000	00000000	00000000	00000000	00000000	00000000	00000000	00000000	
	00000000	00000000	00000000	00000000	00000000	00000000	00000000	00000008	
5	04038000	00000000	00000200	00000010	80000000	00001000	00000000	00200410	33
	80000000	00001000	00000000	00000010	04038002	00001000	00000801	00000000	
	00000002	00000000	00000a01	00000000	00000000	00000000	00000000	00000000	
	00000000	00000000	00000000	00200400	84038002	00000000	00000a01	00200400	
6	08420002	F8022900	00000000	30821140	08830144	A0022100	0C051080	10C01000	90
	0903000C	00000000	04001002	00000000	0181014C	58A04845	0C051082	22406340	
	00000000	A0A26145	00041080	12807200	01800148	58A04845	08011002	22406340	
	01C0014A	00000000	08051082	10420000	00400002	58000800	00040080	20020140	
End	CD9F7546	362513EA	56FE147F	85F6B1E1					
	8D0682FD	F100928A	B44C3D06	18A0D101					
	B8871BEA	70315A82	4819C14B	26257026					
	A1DD0199	40072022	8329356A	A744E830					

6-round differential path

- We found a differential path for 6 round Hamsi with probability 2^{-148}. Ideally, each differential should have probability $\approx 2^{-256}$.
- Thus, we show that the output transformation does not behave ideally.
- The current results don't seem to affect the security claims of the full hash function.
- However, we recommend increasing the number of rounds to 8.

Overview

Summary

Higher-order and standard differential cryptanalysis applied to the compression function of Hamsi-256

- Suboptimal algebraic degree
- *k*-sums and zero-sums found efficiently
- Near collisions; we found a $(256 - 25)$-bit collision from 6 bit differences.
- We found a differential path for 6 round Hamsi with probability 2^{-148}.
- We found a truncated differential path for 6 rounds in 180 of 256 output bits with probability $2^{-120.8}$

Overview

Overview

References I

[BAK98] Eli Biham, Ross J. Anderson, and Lars R. Knudsen.
Serpent: A new block cipher proposal.
In Serge Vaudenay, editor, *FSE*, volume 1372 of *LNCS*, pages 222–238. Springer, 1998.

[BM97] Mihir Bellare and Daniele Micciancio.
A new paradigm for collision-free hashing: Incrementality at reduced cost.
In Walter Fumy, editor, *EUROCRYPT*, volume 1233 of *LNCS*, pages 163–192. Springer, 1997.

[Kö9a] Özgül Kücük.
The hash function Hamsi.
Submission to NIST, January 2009.

References II

[Kö9b] Özgül Kücük.
Reference implementation of Hamsi.
Submission to NIST, January 2009.

[Nik09] Ivica Nikolić.
Near collisions for the compression function of Hamsi-256.
CRYPTO rump session, 2009.

[NIS07] NIST.
Announcing request for candidate algorithm nominations for a new cryptographic hash algorithm (SHA-3) family.
Federal Register Notice, 72(112), November 2007.

References III

http://csrc.nist.govgroups/ST/hash/documents/FR_Notice_Nov07.pdf.

[Wag99] David Wagner.
The boomerang attack.
In Lars R. Knudsen, editor, *FSE*, volume 1636 of *LNCS*, pages 156–170. Springer, 1999.

[Wag02] David Wagner.
A generalized birthday problem.
In Moti Yung, editor, *CRYPTO*, volume 2442 of *LNCS*, pages 288–303. Springer, 2002.

References IV

[WLF+05] Xiaoyun Wang, Xuejia Lai, Dengguo Feng, Hui Chen, and Xiuyuan Yu.
Cryptanalysis of the hash functions MD4 and RIPEMD.
In Ronald Cramer, editor, *EUROCRYPT*, volume 3494 of *LNCS*, pages 1–18. Springer, 2005.

[WWJW09] Meiqin Wang, Xiaoyun Wang, Keting Jia, and Wei Wang.
New pseudo-near-collision attack on reduced-round of Hamsi-256.
Cryptology ePrint Archive, Report 2009/484, 2009.

References V

[WY05] Xiaoyun Wang and Hongbo Yu.
How to break MD5 and other hash functions.
In Ronald Cramer, editor, *EUROCRYPT*, volume 3494 of *LNCS*, pages 19–35. Springer, 2005.

[WYY05a] Xiaoyun Wang, Yiqun Lisa Yin, and Hongbo Yu.
Finding collisions in the full SHA-1.
In Victor Shoup, editor, *CRYPTO*, volume 3621 of *LNCS*, pages 17–36. Springer, 2005.

[WYY05b] Xiaoyun Wang, Hongbo Yu, and Yiqun Lisa Yin.
Efficient collision search attacks on SHA-0.
In Victor Shoup, editor, *CRYPTO*, volume 3621 of *LNCS*, pages 1–16. Springer, 2005.