On a Conditional Collision Attack on NaSHA-512

S. Markovski[1], A. Mileva[2], V. Dimitrova[1] and D. Gligoroski[3]

[1]University "Ss Cyril and Methodius", Faculty of Sciences, Institute of Informatics, P. O. Box 162, Skopje, Republic of Macedonia ({smile,vesnap}@ii.edu.mk)

[2] University "Goce Delčev" , Faculty of Informatics, Štip, Republic of Macedonia (aleksandra.mileva@ugd.edu.mk)

[3] NTNU, Department of Telematics Trondheim, Norway (danilog@item.ntnu.no)

Abstract

A collision attack on NaSHA-512 was proposed by L. Ji et al. The claimed complexity of the attack is 2^{192}. The proposed attack is realized by using a suitable differential pattern. In this note we show that the correct result that can be inferred from their differential pattern is in fact a conditional one. It can be stated correctly as follows: A collision attack on NaSHA-512 of complexity $k = 1, 2, \dots, 2^{320}$ can be performed with an unknown probability of success p_k, where $0 \le p_1 \le p_2 \le p_{2^{320}} \le 1$. Consequently, the attack proposed by L. Ji et al. can be considered only as a direction how a possible collision attack on NaSHA-512 could be realized. The birthday attack remains the best possible attack on NaSHA-512.

1 Introduction

Recently, a collision attack on NaSHA-512 hash function was proposed by L. Ji, X. Liangyu and G. Xu [1]. NaSHA(m,k,r) is a new family of hash functions [2] proposed for SHA-3, and the attack is on its 512-bit hash version. The attackers claim that their attack is of complexity 2^{192}, but they do not give a profound analysis of their estimation. Here we show that if a collision attack on NaSHA-512 of complexity 2^{192} can be performed, then a system E of three quasigroup equations with five unknowns will have a solution. There are no theoretical results for solvability of quasigroup equations, so no one can check if that system E of quasigroup equations has a solution, especially having in mind that the quasigroups are of order 2^{64}. On the other side, in the set of quasigroups of order 4, we have examples of systems of equations of kind similar as E with empty set of solutions, that can be effectively checked. Hence, the attack proposed in [1] can be taken only as conditional one.

In order to make this note readable, we use the same notation, as well as the citations, from [1]. So, we recommend to the reader to follow both [1] and this note.

2 Short description of NaSHA-(512,2,6)

We give a short description of NaSHA-(512,2,6) at first.

Let denote the 1024-bit initial chaining value of NaSHA-(512,2,6) by $H = H_1||H_2||\dots||H_{16}$ and let denote a 1024-bit message block by $M = M_1||M_2||\dots||M_{16}$, where H_i and M_i are 64-bits words. Then, the state of the compression function is defined to be the 2048-bit word

$$S = M_1||H_1||M_2||H_2||\dots||M_{16}||H_{16},$$

represented as 32 64-bit words $S = S_1||S_2||\dots||S_{32}$. Then NaSHA transform the word S into the word $S' = \mathcal{MT}(LinTr_{512}^{32}(S))$, where $LinTr_{512}$ and $\mathcal{MT}$ are defined as

$$LinTr_{512}(S_1||S_2||\dots||S_{31}||S_{32}) = (S_7 \oplus S_{15} \oplus S_{25} \oplus S_{32})||S_1||S_2||\dots||S_{31},$$

$$\mathcal{MT} = \rho(\mathcal{RA}_{l_1}) \circ \mathcal{A}_{l_2}.$$

The definition of $\rho(\mathcal{RA}_{l_1})$ is irrelevant for the attack, and the transformation $\mathcal{A}_{l_2}$ is defined iteratively by

$$\mathcal{A}_{l_2}(x_1,\dots,x_{32}) = (z_1,\dots,z_{32}) \Leftrightarrow z_j = \begin{cases} (l_2 + x_1) * x_1, & j = 1 \\ (z_{j-1} + x_j) * x_j, & 2 \le j \le 32 \end{cases} \tag{1}$$

Here, l_2 is a constant, $\oplus$ denotes the bitwise xoring, $+$ denotes the addition modulo 2^{64} and $*$ denotes a quasigroup operation defined by an extended Feistel network $F_{A,B,C}$ as $x * y = F_{A,B,C}(x \oplus y) \oplus y$. If there is another message block for processing, every second 64-bit word from S' goes as chaining value in the next iteration. If the processed block is the last one, every forth 64-bit word from S' goes as hash result.

The extended Feistel network $F_{A,B,C}$ is a permutation of the set $\{0,1\}^{64}$ and is defined in NASHA by

$$F_{A,B,C}(L||R) = (R \oplus A)||(L \oplus B \oplus f_{a_1,b_1,c_1,a_2,b_2,c_2,a_3,b_3,c_3,\alpha,\beta,\gamma}(R \oplus C))$$

where $a_1, b_1, c_1, a_2, b_2, c_2, a_3, b_3, c_3$ are 8-bit words, α, β, γ are 16-bit words, A, B, C are 32-bit words, L, R are 32-bit variables and f is a suitably defined function. So, the quasigroup operation $*$ in NaSHA used in transformation $\mathcal{A}_{l_2}$ depends on 15 parameters $a_1, b_1, c_1, a_2, b_2, c_2, a_3, b_3, c_3, \alpha, \beta, \gamma, A, B, C$. These parameters and the constant l_2 are different in every iteration of the compression function and depend on the processed message block. They are obtained from the equalities:

$$l_2 = S_3 + S_4,$$

$$a_1||b_1||c_1||a_2||b_2||c_2||a_3||b_3 = S_5 + S_6, \quad c_3 = a_1,$$

$$\alpha||\beta||\gamma||\alpha_2 = S_7 + S_8,$$

$$A||B = S_{11} + S_{12}, \quad C||A_2 = S_{13} + S_{14},$$

the values α_2 and A_2 are irrelevant for the attack.

3 Setting the attack parameters

The attack is based on a differential pattern obtained by using the difference `0x00000000FFFFFFFF`, where `0` $= 0000,$ `F` $= 1111$. Several very clever observations are obtained.

1) Let x be any 64-bit word. Denote by $(x)_i$ the i-th bit of x and construct a new 64-bit word a by $(a)_{64\ldots33} = \neg(x)_{64\ldots33}$, $(a)_{32} = 1$ and $(a)_{31\ldots1} = 0$. Note thata $a = a(x)$ is a function of x. Define a difference $\Delta x =$ `0x00000000FFFFFFFF`. Then for the word $x' = x \oplus \Delta x$ the following equality is true:

$$(a + x) * x = (a + x') * x',$$

where $\oplus$ denotes the 64-bit XOR, $+$ denotes the addition modulo 2^{64} and $*$ denotes the quasigroup operation defined by an extended Feistel network $F_{A,B,C}$. Here $A,\ B,\ C$ are parameters that are computed from the input message and the chaining values.

2) If the parameters $a_1, b_1, c_1, a_2, b_2, c_2, a_3, b_3, c_3, \alpha, \beta, \gamma$ are known, i.e., the function f is defined, then the parameters $A,\ B,\ C$ can be chosen such that the following equalities hold true:

$$(a+x)*x = a = (a+x')*x'.$$

3) The initial chaining value of NaSHA is $H = H_1||H_2||\dots||H_{16}$ and let take an input message $M = M_1||M_2||\dots||M_{16}$, where H_i and M_i are 64-bits words. Only the words M_i can be chosen in a suitable way a collision attack to be realized. The idea of the attack is to find two different 1024-bits input messages M and M' such that

$$\mathcal{A}_{l_2}(LinTr^{32}_{512}(M_1||H_1||M_2||H_2||\dots||M_{16}||H_{16})) =$$
$$= \mathcal{A}_{l'_2}(LinTr^{32}_{512}((M'_1||H_1||M'_2||H_2||\dots||M'_{16}||H_{16})).$$

The values of l_2 and l'_2 are defined after $LinTr^{32}_{512}$ is applied.

4) Let denote

$$LinTr^{32}_{512}(M_1||H_1||M_2||H_2||\dots||M_{16}||H_{16}) = S_1||S_2||\dots||S_{32},$$

$$LinTr^{32}_{512}(M'_1||H_1||M'_2||H_2||\dots||M'_{16}||H_{16}) = S'_1||S'_2||\dots||S'_{32}.$$

Then, M (as well as M') can be recovered from $S_1||S_2||\dots||S_{32}$ by using $LinTr^{-1}_{512}$. Recall that now in NaSHA l_2 and l'_2 are defined by $l_2 = S_3 + S_4,\ l'_2 = S'_3 + S'_4$.

4 Collision attacks on NaSHA

5) Take an arbitrary 64-bits word x and the differential

$$\Delta x = \texttt{0x00000000FFFFFFFF}.$$

Note that x can be chosen at 2^{64} manners.

6) Suppose that the input messages M and M' satisfy the conditions $M_1 = M'_1, M_2 = M'_2, M_3 = M'_3 \oplus \Delta x, M_4 = M'_4, M_5 = M'_5 \oplus \Delta x, M_6 = M'_6 \oplus \Delta x, M_7 = M'_7 \oplus \Delta x, M_8 = M'_8, M_9 = M'_9 \oplus \Delta x, M_{10} = M'_{10} \oplus \Delta x, M_{11} = M'_{11} \oplus \Delta x, M_{12} = M'_{12}, M_{13} = M'_{13}, M_{14} = M'_{14}, M_{15} = M'_{15} \oplus \Delta x, M_{16} =$

$M'_{16} \oplus \Delta x$. Then we have that $S_9 = S'_9 \oplus \Delta x, S_{10} = S'_{10} \oplus \Delta x, S_{17} = S'_{17} \oplus \Delta x, S_{18} = S'_{18} \oplus \Delta x, S_{19} = S'_{19} \oplus \Delta x, S_{20} = S'_{20} \oplus \Delta x, S_{21} = S'_{21} \oplus \Delta x, S_{29} = S'_{29} \oplus \Delta x, S_{31} = S'_{31} \oplus \Delta x$.

7) Now choose the values for the words S_i and S'_i in a suitable manner. By using $LinTr^{-1}_{512}$ corresponding messages M and M' will be obtained.

7.1) Take $S_9 = x' = x \oplus \Delta x,\ S_{10} = S_{17} = S_{18} = S_{19} = S_{20} = S_{21} = S_{29} = S_{30} = S_{31} = x$ and $S'_9 = x, S'_{10} = S'_{17} = S'_{18} = S'_{19} = S'_{20} = S'_{21} = S'_{29} = S'_{31} = x' = x \oplus \Delta x$.

7.2) Take $S_5 = S'_5 = y_5, S_6 = S'_6 = y_6, S_7 = S'_7 = y_7, S_8 = S'_8 = y_8, S_{11} = S'_{11} = y_{11}, S_{14} = S'_{14} = y_{14}$, where y_i are unknown (free) words.

7.3) By using the equality (1) of [1], the words $S_1, S_2, S_3, S_4, S_{12}, S_{13}, S_{15}, S_{16}, S_{22}, S_{23}, S_{24}, S_{25}, S_{26}, S_{27}, S_{28}, S_{32}$ can be expressed by the initial chaining value H, the word x and the unknown words $y_5, y_6, y_7, y_8, y_{11}, y_{14}$. Hence, they are functions of $x, y_5, y_6, y_7, y_8, y_{11}, y_{14}$.

7.4) The parameters $a_1, b_1, c_1, a_2, b_2, c_2, a_3, b_3, c_3, \alpha, \beta, \gamma, A, B, C$ and the constants $l_2,\ l'_2$ now can be expressed as functions of $x, y_5, y_6, y_7, y_8, y_{11}, y_{14}$ as well:

$$l_2 = l'_2 = S_3(x, y_5, y_6, y_7, y_8, y_{11}, y_{14}) + S_4(x, y_5, y_6, y_7, y_8, y_{11}, y_{14}),$$

$$a_1||b_1||c_1||a_2||b_2||c_2||a_3||b_3 = y_5 + y_6,$$

$$\alpha||\beta||\gamma||\alpha_2 = y_7 + y_8,$$

$$A||B = y_{11} + S_{12}(x, y_7, y_8),$$

$$C||A_2 = S_{13}(x, y_6) + y_{14}.$$

7.5) The parameters A, B, C of $F_{A,B,C}$ have to be determined in such a way the equality $(a + x) * x = a$ to be satisfied. For that aim at first fixed values to y_5, y_6, y_7, y_8 have to be given, and after that the values for y_{11} and y_{14} can be computed. Note that now $S_{11} = y_{11}$ and $S_{14} = y_{14}$ are functions of x, y_5, y_6, y_7, y_8.

8) Note that after the values of y_5, y_6, y_7 and y_8 are chosen, all the words S_i and S'_i are determined. We have to check if the equalities

$\mathcal{A}_{l_2}(S_1||S_2||\ldots||S_{32}) = \mathcal{A}_{l'_2}(S'_1||S'_2||\ldots||S'_{32}) = z_1||z_2||\ldots||z_{32}$
hold for some z_i.

The differential pattern of the attack is defined in such a way that

$z_8||z_9||z_{10} = a||a||a,$

$z_{16}||\ldots||z_{21} = a||a||a||a||a||a,$

$z_{28}||\ldots||z_{31} = a||a||a||a.$

Then only the values of $z_1, \ldots, z_7,\ z_{11}, \ldots, z_{15},\ z_{22}, \ldots, z_{27}$ and z_{32} have to be found.

8.1) We can compute $z_1 = (l_2 + S_1) * S_1, z_2 = (z_1 + S_2) * S_2, z_3 = (z_2 + S_3) * S_3, \ldots, z_7 = (z_6 + S_7) * S_7$. Note that $z_1, \ldots, z_7$ are functions of x, y_5, y_6, y_7, y_8.

Now, **the equality** $z_8 = a$, i.e., $(z_7 + S_8) * S_8 = a$, **has to be satisfied**, in order the transformations $\mathcal{A}_{l_2}$ and $\mathcal{A}_{l'_2}$ to be fulfilled.

8.2) If $z_8 = a$ holds true, we can compute $z_{11} = (a + S_{11}) * S_{11}, z_{12} = (z_{11}+S_{12})*S_{12}, \ldots, z_{15} = (z_{14}+S_{15})*S_{15}$. Note that $z_{11}, \ldots, z_{15}$ are functions of x, y_5, y_6, y_7, y_8.

Now, **the equality** $z_{16} = a$, i.e., $(z_{15} + S_{16}) * S_{16} = a$, **has to be satisfied**, in order the transformations $\mathcal{A}_{l_2}$ and $\mathcal{A}_{l'_2}$ to be fulfilled.

8.3) If $z_8 = a$ and $z_{16} = a$ hold true, we can compute $z_{22} = (a + S_{22}) * S_{22}, z_{23} = (z_{22} + S_{232}) * S_{23}, \ldots, z_{27} = (z_{26} + S_{27}) * S_{27}$. Note that $z_{22}, \ldots, z_{27}$ are functions of x, y_5, y_6, y_7, y_8.

Now, **the equality** $z_{28} = a$, i.e., $(z_{27} + S_{28}) * S_{28} = a$, **has to be satisfied**, in order the transformations $\mathcal{A}_{l_2}$ and $\mathcal{A}_{l'_2}$ to be fulfilled.

8.4) If $z_8 = a$, $z_{16} = a$ and $z_{28} = a$ hold true, we can compute $z_{32} = (a + S_{32}) * S_32$.

5 Solvability of quasigroup equations

In order the above attack to be successful, for some values of the variables x, y_5, y_6, y_7, y_8 the following equalities have to be satisfied: $z_8 = a$, $z_{16} = a$ and $z_{28} = a$. Then we have that the next proposition is true:

Proposition 1 *If there is a collision on NaSHA-512 obtained by the attack as explained in 1) – 8), then the system E of three quasigroup equations with fife variables*

$$\left\{ \begin{array}{l} (z_7(x, y_5, y_6, y_7, y_8) + S_8(x, y_5, y_6, y_7, y_8)) * S_8(x, y_5, y_6, y_7, y_8) = a(x) \\ (z_{15}(x, y_5, y_6, y_7, y_8) + S_{16}(x, y_5, y_6, y_7, y_8)) * S_{16}(x, y_5, y_6, y_7, y_8) = a(x) \\ (z_{27}(x, y_5, y_6, y_7, y_8) + S_{28}(x, y_5, y_6, y_7, y_8)) * S_{28}(x, y_5, y_6, y_7, y_8) = a(x) \end{array} \right.$$

has a solution, where z_i are obtained iteratively as in 8).

There are not known any results for solving systems of quasigroup equations, except checking all possible solutions. So, for the system E we have to make 2^{320} checks to find a solution, if any. Of course, it can not be realized, at least with today computing power. Next we give two examples of systems of quasigroup equations in the set of quasigroups of order 4 that have empty set of solutions.

Example 1 The system of two quasigroup equations with 3 unknowns x, y, a:

$$((1+x+y)*(1+y)+2+x+y)*y=a,$$
$$((3+x+y)*y+x+y)*(x+y+1)=a$$

has no solution in the quasigroup

$*$	0	1	2	3
0	0	2	1	3
1	3	0	2	1
2	1	3	0	2
3	2	1	3	0

Example 2 The system of three quasigroup equations with 5 unknowns x, y, z, u, a:

$$r(x,y,z,u) := \Big\{\{[(1+x+y+z)*(2+x+z+u)+3+x+u]*(1+y)+ 2+z+u\}*(1+z)\Big\}*u=a,$$

$$s(x,y,z,u) := \Big\{\{[(3+x+y)*(z+u)+1+x+y+z]*(x+z)+ 1+x+z+u\}*(1+x+y)+1+x+u\Big\}*(y+z)=a,$$

$$t(x,y,z,u) := \Big\{\{[(1+y+u)*(y+z)+z+u]*(x+z+u)+ 2+y+z\}*z+z+1\Big\}*(3+y+u)=a$$

has no solution in the quasigroup

$*$	0	1	2	3
0	0	1	2	3
1	3	2	1	0
2	2	3	0	1
3	1	0	3	2

6 Conclusion

The attack given in [1] is a very sophisticated one and a lot of effort is given to be realized. Nevertheless, it is not a valuable attack on NaSHA-512. We do not know if the system of quasigroup equations $E: \; z_8 = a, \; z_{16} = a, \; z_{28} = a$ with fife unknowns has a solution in a quasigroup of order 2^{64}. The attacker are stating that there is a collision of NaSHA-512 of complexity 2^{192}, but one can state that there is a collision of complexity 2^{64} as well. The proper statement that can be inferred from the attack designed as in [1] is the following: **For each $k = 1, 2, \ldots, 2^{320}$ there is a collision attack on NaSHA-512 of complexity k that can be realized with probability p_k. The Probabilities p_k are not known and $0 \leq p_1 \leq p_2 \leq \cdots \leq p_{2^{320}} \leq 1$.**

Still, the best attack on NaSHA-512 is the birthday attack.

References

[1] L. Ji, X. Liangyu and G. Xu, *Collison attack on NaSHA-512* http://eprint.iacr.org/2008/519

[2] Smile Markovski and Aleksandra Mileva, *Algorithm Specications of NaSHA*, 2008 http://inf.ugd.edu.mk/images/stories/file/Mileva/Nasha.htm