On the Second Conditional Collision Attack on NaSHA-384/512

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

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

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

Abstract

Recently, a new collision attack on NaSHA-384/512 have been proposed by Z. Ji and D. Li [2]. The claimed complexity of the attack is 2^{128} with probability of $(1-\frac{2}{2^{64}-1})^2$ ($\gg \frac{1}{2}$). We show that the claimed probability of their attack is not correct. The attack is based on an assumption that a system E of two quasigroup equations has a solution. The attacker do not give any evidence why the system E has a solution, and their attack is based only on their believes that they can find a solution after making 2^{128} checks. Unless the attacker provide a proof that the system E do have a solution and that the solution can be found after 2^{128} checks, their attack is irrelevant.

1 Introduction

Recently, a new collision attack on NaSHA-384/512 have been proposed by Z. Ji and D. Li [2]. NaSHA(m,k,r) is a new family of hash functions [3] proposed for SHA-3, and the attack is on its 384-bit and 512-bit hash versions. The claimed complexity of the attack is 2^{128} with probability of $(1-\frac{2}{2^{64}-1})^2$ ($\gg \frac{1}{2}$).

What is actually presented in their paper is a system E of two quasigroup equations with fife variables that has a small probability to have a solution, i.e. the probability is $(1 - \frac{2}{2^{64}-1})^2$. Moreover, they only calculate the probability each equation separately to have a solution, which is $(1 - \frac{2}{2^{64}-1})$ each. But the system E is a system of two mutually dependable equations of fife variables, and it is not true that the probability p to solve this system is $(1 - \frac{2}{2^{64}-1})^2$. The probability p is quite unknown, there is no theory for solving quasigroup equations, and a system of quasigroup equations may have no solutions at all. Unless the attackers provide a proof that the system E do have a solution and that the solution can be found after 2^{128} checks, their attack is irrelevant.

We note that the attack in [2] is based on the same idea as the attack of [1], and all of the arguments given in [4] can be applied to the attack of [2]. Nevertheless, here we present the attack of [2] and give comments of it.

2 Short description of NaSHA-(384,2,6) and NaSHA-(512,2,6)

We give a short description of NaSHA-(384,2,6) and 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_1,B_1,C_1} as $x * y = F_{A_1,B_1,C_1}(x \oplus y) \oplus y$. 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. For NaSHA-(384,2,6) is the same, but final hash is modulo 2^{384}.

The extended Feistel network F_{A_1,B_1,C_1} is a permutation of the set $\{0,1\}^{64}$ and is defined in NaSHA by

$$F_{A_1,B_1,C_1}(L||R) = (R \oplus A_1)||(L \oplus B_1 \oplus f_{a_1,b_1,c_1,a_2,b_2,c_2,a_3,b_3,c_3,\alpha,\beta,\gamma}(R \oplus C_1))$$

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_1, B_1, C_1 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 $a_1, b_1, c_1, a_2, b_2, c_2, a_3, b_3, c_3, \alpha, \beta, \gamma, A_1, B_1, C_1$. 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:

$$l_2 = S_3 + S_4,$$

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

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

$$A_1||B_1 = S_{11} + S_{12}, \quad C_1||A_2 = S_{13} + S_{14},$$

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 `0xFFFF00000000FFFF`. Several observations are obtained.

1) Let $x =$ `0xFFFFFFFF00008000` and $a =$ `0x7FFF80017FFF8000` be 64-bit words. Then for the word $y = x \oplus \Delta x$ the following equality is true:

$$(a + x) * x = (a + y) * y$$

and

$$a_L = ((a + x) * x)_L$$

where $\oplus$ denotes the 64-bit XOR, $+$ denotes the addition modulo 2^{64}, a_L means the left half bits of a 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:

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

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 us 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 $LinTr^{-1}_{512}$.

4 Collision attacks on NaSHA

5) Take x = `0xFFFFFFFF00008000`, a = `0x7FFF80017FFF8000`, Δx = `0xFFFF00000000FFFF` and $y = x + \Delta x$.

6) Suppose that the input messages M and M' satisfy the conditions $M_1 = M'_1, M_2 = M'_2 \oplus \Delta x, M_3 = M'_3 \oplus \Delta x, M_4 = M'_4 \oplus \Delta x, M_5 = M'_5, M_6 = M'_6 \oplus \Delta x, M_7 = M'_7 \oplus \Delta x, M_8 = M'_8, M_9 = M'_9 \oplus \Delta x, M_{10} = M'_{10}, M_{11} = M'_{11} \oplus \Delta x, M_{12} = M'_{12}, M_{13} = M'_{13}, M_{14} = M'_{14}, M_{15} = M'_{15} \oplus \Delta x, M_{16} = M'_{16}$. Then we have that $S_{11} = S'_{11} \oplus \Delta x, S_{12} = S'_{12} \oplus \Delta x, S_{25} = S'_{25} \oplus \Delta x, S_{28} = S'_{28} \oplus \Delta x, S_{29} = S'_{29} \oplus \Delta x, S_{32} = S'_{32} \oplus \Delta x$.

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 $S_{12} = y,\ S_{11} = S_{25} = S_{26} = S_{27} = S_{28} = S_{29} = S_{30} = S_{31} = S_{32} = x$ and $S'_9 = x, S'_{11} = S'_{25} = S'_{26} = S'_{27} = S'_{28} = S'_{29} = S'_{30} = S'_{31} = S'_{32} = y$.

7.2) Take $S_9 = S'_9 = y_9, S_{10} = S'_{10} = y_{10}, S_{13} = S'_{13} = y_{13}, S_{14} = S'_{14} = y_{14}, S_{22} = S'_{22} = y_{22}, S_{24} = S'_{24} = y_{24}$, where y_i are unknown (free) words.

7.3) By using the equality of [2], the words $S_1, S_2, S_3, S_4, S_5, S_6, S_7, S_8, S_{15}, S_{16}, S_{17}, S_{18}, S_{19}, S_{20}, S_{21}, S_{23}$ can be expressed by the initial chaining value H, the word x and the unknown words $y_9, y_{10}, y_{13}, y_{14}, y_{22}, y_{24}$. Hence, they are functions of $y_9, y_{10}, y_{13}, y_{14}, y_{22}, y_{24}$.

7.4) The parameters $a_1, b_1, c_1, a_2, b_2, c_2, a_3, b_3, c_3, \alpha, \beta, \gamma, A_1, B_1, C_1$ and the constants $l_2,\ l'_2$ now can be expressed as functions of $y_9, y_{10}, y_{13}, y_{14}, y_{22}, y_{24}$ as well:

$$l_2 = l'_2 = S_3(y_9, y_{10}, y_{13}, y_{14}, y_{22}, y_{24}) + S_4(y_9, y_{10}, y_{13}, y_{14}, y_{22}, y_{24}),$$

$$a_1||b_1||c_1||a_2||b_2||c_2||a_3||b_3 = S_5(y_9, y_{10}, y_{13}, y_{14}, y_{22}, y_{24}) + S_6(y_9, y_{10}, y_{13}, y_{14}, y_{22}, y_{24}),$$

$$\alpha||\beta||\gamma||\alpha_2 = S_7(y_9, y_{10}, y_{13}, y_{14}, y_{22}, y_{24}) + S_8(y_9, y_{10}, y_{13}, y_{14}, y_{22}, y_{24}),$$

$$A_1||B_1 = S_{11} + S_{12},$$

$$C_1||A_2 = y_{13} + y_{14}.$$

7.5) The parameters A_1 and B_1 are fixed and $a_L = ((a + x) * x)_L = ((a + y) * y)_L$, so only the parameter C_1 of F_{A_1,B_1,C_1} have to be determined in such a way the equality $a_R = ((a+x)*x)_R = ((a+y)*y)_R$ to be satisfied. For that aim at first fixed values to $y_9, y_{10}, y_{14}, y_{22}, y_{24}$ have to be given, and after that the values for y_{13} can be computed. Note that now $S_{13} = y_{13}$ is function of $y_9, y_{10}, y_{14}, y_{22}, y_{24}$.

8) Note that after the values of $y_9, y_{10}, y_{14}, y_{22}$ and y_{24} are chosen, all the words S_i and S'_i are determined. We have to check if the equalities

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

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

$z_{10}||z_{11}||z_{12} = a||a||a,$

$z_{24}||\dots||z_{32} = a||a||a||a||a||a||a||a||a.$

Then only the values of $z_1, \dots, z_9$ and $z_{13}, \dots, z_{23}$ have to be found.

8.1) We can compute $z_1 = (l_2 + S_1) * S_1, z_2 = (z_1 + S_2) * S_2, z_3 = (z_2 + S_3) * S_3, \dots, z_9 = (z_8 + S_9) * S_9$. Note that $z_1, \dots, z_9$ are functions of $y_9, y_{10}, y_{14}, y_{22}, y_{24}$.

Now, **the equality** $z_{10} = a$, i.e., $(z_9+S_{10})*S_{10} = a$, **has to be satisfied**, in order the transformations $\mathcal{A}_{l_2}$ and $\mathcal{A}_{l'_2}$ to be fulfilled.

8.2) If $z_{10} = a$ holds true, we can compute $z_{13} = (a + S_{13}) * S_{13}, z_{14} = (z_{13}+S_{14})*S_{14}, \dots, z_{23} = (z_{22}+S_{23})*S_{23}$. Note that $z_{13}, \dots, z_{23}$ are functions of $y_9, y_{10}, y_{14}, y_{22}, y_{24}$.

Now, **the equality** $z_{24} = a$, i.e., $(z_{23} + S_{24}) * S_{24} = a$, **has to be satisfied**, in order the transformations $\mathcal{A}_{l_2}$ and $\mathcal{A}_{l'_2}$ to be fulfilled.

Proposition 1 *If there is a collision on NaSHA-384/512 obtained by the attack as explained in 1) – 8), then the system E of two quasigroup equations with fife variables*

$$\begin{cases} (z_9(y_9, y_{10}, y_{14}, y_{22}, y_{24}) + S_{10}(y_9, y_{10}, y_{14}, y_{22}, y_{24})) * S_{10}(y_9, y_{10}, y_{14}, y_{22}, y_{24}) = a \\ (z_{23}(y_9, y_{10}, y_{14}, y_{22}, y_{24}) + S_{24}(y_9, y_{10}, y_{14}, y_{22}, y_{24})) * S_{16}(y_9, y_{10}, y_{14}, y_{22}, y_{24}) = a \end{cases}$$

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

As far as we know, there is no efficient method for solving systems of quasigroup equations, except checking all possible solutions. So, for the system E we have to make up to 2^{320} checks to find a solution, if any. Of course, that is infeasible with the current technology.

The attackers are not solving this system. They only choose $y_9, y_{10}, y_{14}, y_{22}, y_{24}$ randomly and after calculating the parameters of quasigroup operations, only check if the system E has a solution. They only calculate the probability each equation separately to have a solution, which is $(1 - \frac{2}{2^{64}-1})$ each. But the system E is a system of two mutual dependable equations of fife variables, and probability to solve this system in total is not $(1 - \frac{2}{2^{64}-1})^2$. This probability is unknown, because there are cases when this kind of systems do not have solutions (see the Example 1 and 2 below).

Example 1 The system of two equations with 4 unknowns

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

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

has no solutions in the quasigroup

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

Example 2 The system of two equations with 4 unknowns

$$x * y = 0,$$
$$(1 + x + 2z) * y = 0$$

has no solutions in the quasigroup

$*$	0	1	2	3	4	5	6	7
0	7	4	6	2	1	0	3	5
1	1	2	5	7	6	3	4	0
2	3	5	4	0	2	6	7	1
3	4	3	0	1	5	2	6	7
4	0	1	3	4	7	5	2	6
5	6	7	2	3	0	1	5	4
6	2	6	1	5	4	7	0	3
7	5	0	7	6	3	4	1	2

5 Conclusion

The attack given in [2] is very similar to the previous attack given in [1]. Nevertheless, it is not a valuable attack on NaSHA-384/512, because we do not know if the system of quasigroup equations $E: \ z_{10} = a, \ z_{24} = a$ with fife unknown variables has a solution. So, the claimed probability $(1-\frac{2}{2^{64}-1})^2$ $(\gg \frac{1}{2})$ that after 2^{128} checks, a solution of the system of equations E can be found is not true. In fact, it is highly probable that the system E does not have solutions at all.

References

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

[2] Z. Ji and D. Li, *Collison attack on NaSHA-384/512* http://eprint.iacr.org/2009/026

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

[4] S. Markovski, A. Mileva, V. Dimitrova and D. Gligoroski, *On a Conditional Collision Attack on NaSHA-512* http://eprint.iacr.org/2009/034