SCIENCE CHINA
Information Sciences

• HIGHLIGHT •

March 2018, Vol. 61 038105:1–038105:2
doi: 10.1007/s11432-017-9119-6

Improved cryptanalysis of step-reduced SM3

Yanzhao SHEN[1,2], Dongxia BAI[3] & Hongbo YU[3,4*]

[1]*Key Laboratory of Cryptologic Technology and Information Security, Ministry of Education, Shandong University, Jinan* 250100*, China;*
[2]*School of Mathematics, Shandong University, Jinan* 250100*, China;*
[3]*Department of Computer Science and Technology, Tsinghua University, Beijing* 100084*, China;*
[4]*Science and Technology on Communication Security Laboratory, Chengdu* 610041*, China*

Received 15 May 2017/Accepted 27 May 2017/Published online 23 August 2017

Citation Shen Y Z, Bai D X, Yu H B. Improved cryptanalysis of step-reduced SM3. Sci China Inf Sci, 2018, 61(3): 038105, doi: 10.1007/s11432-017-9119-6

SM3 is the Chinese hash standard and is standardized in GB/T 32905-2016 [1]. As a hash function, it must fulfill three security requirements, collision resistance, preimage resistance, and second preimage resistance. During the ongoing evaluation, it is believed that whenever the hash function behaves differently from a random function, it is considered as the hash function's weakness. In recent years, the analysis has not only been limited to the classical security requirements, but also in the near-collision, boomerang distinguisher, and (semi-)free-start collision. Most of the previous preimage attacks on SM3 [2, 3] are either without padding or padding is not present from the first step. The best boomerang attack on SM3 covers 37 steps [4, 5]. In this article, we focus on the preimage attack from the first step, with message padding. A preimage attack on 30-step SM3 is proposed. Furthermore, we improve the 37-step boomerang attack and extend it to the 38-step boomerang attack. A summary of the previous results and along with our owns is given in Table 1.

Brief description of SM3. SM3 was proposed by Wang et al. [1]. The structure of SM3 resembles the structure of SHA-256. It is based on the Merkle-Damgård design which uses a 512-bit long message block and outputs a 256-bit long hash value. It first calls the message padding procedure. Then, using the compression function, it updates the initial value and produces the hash value. Let CF_1 and CF_2 be the sub-part of the compression function. Here, CV_i denotes the input chaining variables of the i-th step.

The message padding procedure ensures a padded message length is a multiple of 512 bit. For an l-bit message, the bit "1" is appended to the end of the message, followed by k "0" bit, where k is the smallest non-negative solution to the equation $l + k + 1 \equiv 448 \bmod 512$. Then, the 64-bit block is appended, which is equal to the number l expressed using a binary representation. The words m_{14} and m_{15} in the last message block represent the message length. For example, the message length of a 2-block padded message is less than 960, i.e., m_{14} and the most significant 22 bit of m_{15} must be "0". Whenever an attack uses these bits, it contains at least 3 message blocks. In work [3], all bits of m_{14} on a 29-step attack and the most significant 4 bit of m_{15} on a 30-step attack are not "0". The message blocks are more than 2^{19} and their complexities are much more than 2^{256}, which is the bound of preimage attack. Therefore, both preimage attacks are invalid.

30-step preimage attack on SM3. Assume LD_i are the linear space used in CF_i, where $i \in \{1, 2\}$. CF_1 is from step 0 to step 14 and LD_1 is fixed in the most significant 5 bit of m_2. CF_2 is from step

* Corresponding author (email: yuhongbo@mail.tsinghua.edu.cn)
The authors declare that they have no conflict of interest.

Table 1 Summary of the attacks on SM3

Attacks	Steps	Padding	Complexity	Ref.
Preimage	28(1-28)	Yes	$2^{241.5}$	[2]
Preimage	30(7-36)	Yes	2^{249}	[2]
Preimage	29(1-29)	No	2^{245}	[3]
Preimage	30(1-30)	No	$2^{251.1}$	[3]
Preimage	30(1-30)	Yes	$2^{255.3}$	Ours
Boomerang	34(1-34)	No	$2^{31.4}$	[4, 5]
Boomerang	35(1-35)	No	$2^{33.6}$	[4, 5]
Boomerang	36(1-36)	No	$2^{73.4}$	[4, 5]
Boomerang	37(1-37)	No	2^{192}	[4, 5]
Boomerang	37(1-37)	No	2^{125}	Ours
Boomerang	38(1-38)	No	2^{208}	Ours

15 to step 29 and LD_2 is fixed in the most significant 5 bit of m_{13}. By adding 7 equations on m_0, m_1, m_3, m_4, m_6, m_7 and m_{10}, the expanded message words $w_i (14 \leqslant i \leqslant 28)$ are independent of m_{13}. To satisfy the message padding, $m_{14} = 0x0$, $m_{15} = 0x3bf$ (the message length is $959 = 512 + 447$ bit) and the least significant bit of m_{13} is set to 1. The mask vector is fixed as $T = (0x0, 0x0, 0x0, 0x4000, 0x0, 0x0, 0x0, 0x30003)$. Randomly fixing a value to CV_{14}, the error probability is approximately 0.58. The 30-step preimage attack process introduced by Knellwolf and Khovratovich [6] can be implemented. A 1-block pseudo-preimage is obtained with complexity $2^{252.55}$. It can be converted to a preimage (2 message blocks, with padding) with complexity $2^{255.3}$.

Improved boomerang attacks on SM3. For the 37-step boomerang attack we use the Type III boomerang attack (see [7]). The complexity for the best algorithm is 2^{128}. For the 38-step boomerang attack we use the Type I boomerang attack, and the generic complexity is 2^{256}.

By comparing the probabilities of several alternative differential characteristics, we find that the differences diffuse slower in CF_1^{-1} than in CF_2, i.e., the probability of the differential characteristic is higher in CF_1^{-1} than in CF_2. We carefully choose the steps where the bit differences are fixed in the message words. In light of this idea, the 37-step boomerang attack on SM3 is mounted with two high probability differential characteristics. In this case CF_1^{-1} is from step 17 to 0 and CF_2 is from step 18 to 36. Compared with the two differential characteristics used in the previous 37-step boomerang attack [4, 5], where CF_1^{-1} is from step 16 to 0 and CF_2 is from step 17 to 36, our differential characteristics hold with higher probabilities of course leading to a lower attack complexity. We start from the intermediate values of the boomerang distinguisher quartet and use the message modification technique to modify the chaining values and message words. This fulfills the conditions of intermediate steps to improve the complexity of the attack. An example for the intermediate steps of a 37-step boomerang distinguisher is provided to demonstrate that the attack is compatible.

The 38-step boomerang attack is obtained by adding one step before the differential characteristic of CF_1^{-1}. The differential characteristic of CF_1^{-1} is from step 18 to 0. The intermediate connection part between CF_1^{-1} and CF_2 is almost the same with the 37-step boomerang attack. Therefore, it is also compatible and correct.

Conclusion. This article presents a preimage attack and boomerang attacks on the Chinese hash function standard SM3. We first study the effect of the message padding and find that some of the previous preimage attacks are invalid. A preimage attack on a 30-step SM3 with complexity $2^{255.3}$ is proposed. We also propose high probability differential characteristics for the 37-step SM3 compression function and improve the 37-step boomerang attack from complexity 2^{192} to 2^{125}. Then we extend it to a 38-step. The examples of boomerang distinguishers for intermediate steps of these attacks are also given. Our attacks on SM3 penetrate the highest number of steps.

Acknowledgements This work was supported by National Basic Research Program of China (973 Program) (Grant No. 2013CB834205) and National Natural Science Foundation of China (Grant No. 61373142).

References

1 Standardization Administration of the People's Republic of China. Information security techniques — SM3 cryptographic hash algorithm. GB/T 32905-2016. http://www.sac.gov.cn/gzfw/ggcx/gjbzgg/201614/
2 Zou J, Wu W, Wu S, et al. Preimage attacks on step-reduced SM3 hash function. In: Proceedings of the 14th International Conference on Information Security and Cryptology. Berlin: Springer-Verlag, 2011. 375–390
3 Wang G, Shen Y. Preimage and pseudo-collision attacks on step-reduced SM3 hash function. Inf Proc Lett, 2013, 113: 301–306
4 Bai D, Yu H, Wang G, et al. Improved boomerang attacks on SM3. In: Proceedings of the 18th Australasian Conference on Information Security and Privacy. Berlin: Springer-Verlag, 2013. 251–266
5 Bai D, Yu H, Wang G, et al. Improved boomerang attacks on round-reduced SM3 and keyed permutation of BLAKE-256. IET Inf Secur, 2014, 9: 167–178
6 Knellwolf S, Khovratovich D. New preimage attacks against reduced SHA-1. In: Proceedings of the 32nd Annual Cryptology Conference. Berlin: Springer-Verlag, 2012. 367–383
7 Yu H, Chen J, Wang X. The boomerang attacks on the round-reduced Skein-512. In: Proceedings of the 19th International Conference on Selected Areas in Cryptography. Berlin: Springer-Verlag, 2012. 287–303