SCIENCE CHINA
Information Sciences

• LETTER •

September 2021, Vol. 64 199105:1–199105:2
https://doi.org/10.1007/s11432-018-9878-1

Improved Guess and Determine attack on the MASHA stream cipher

Lin DING[1,2*], Dawu GU[1], Lei WANG[1,3], Chenhui JIN[2] & Jie GUAN[2]

[1]*Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai* 200240*, China;*
[2]*Zhengzhou Information Science and Technology Institute, Zhengzhou* 450000*, China;*
[3]*Westone Cryptologic Research Center, Beijing* 100070*, China*

Received 14 December 2018/Revised 8 March 2019/Accepted 26 April 2019/Published online 21 May 2021

Citation Ding L, Gu D W, Wang L, et al. Improved Guess and Determine attack on the MASHA stream cipher. Sci China Inf Sci, 2021, 64(9): 199105, https://doi.org/10.1007/s11432-018-9878-1

Dear editor,

K2 is a secure and high-performance stream cipher and has been standardized by ISO/IEC 18033-4. The MASHA stream cipher is a successor of K2 with integrated MAC functionality proposed in 2011. Based on optimizing the guess and determination process, we present an improved Guess and Determine attack on MASHA with time complexity of 2^{224}, reducing the time complexity of the existing attack by a factor of 2^{96}. To the best of our knowledge, this is the best attack on MASHA so far.

Introduction. Generally, a stream cipher is mainly a pseudorandom keystream generator, which produces a pseudorandom keystream sequence to encrypt the plaintext messages. Clock-controlled keystream generator is a well-known method for building stream ciphers. Various clock-controlled stream ciphers have been proposed, e.g., A5/1 [1], LILI-128 [2], MICKEY [3], and some recent cryptanalytic developments on clock-controlled stream ciphers can be found in [4, 5]. In 2007, Kiyomoto et al. [6] presented a new design of irregular clocking for word-oriented stream ciphers, called dynamic feedback control, and proposed the K2 stream cipher. It has an excellent performance in hardware and software. The designers believe that the dynamic feedback control mechanism is potentially effective against several different types of attacks, not only existing attacks but also novel attacks. The K2 stream cipher has been standardized by ISO/IEC 18033-4 in 2011.

Based on the K2 stream cipher, a new high-speed stream cipher with integrated MAC functionality, called MASHA [7] (message authenticated streaming-encryption heterogeneous algorithm) was proposed in 2011. The cipher uses a 128-bit key in conjunction with a 192-bit IV to provide a 128-bit security level for both encryption and message authentication. Based on a deep security analysis on MASHA, the designers claimed that MASHA is secure against all known attacks, e.g., Guess and Determine Attack, which is an attack strategy that has been successfully applied to many stream ciphers [8, 9]. Besides the security analysis by the designers, no attack on MASHA has been published so far. In the specification of MASHA , the designers presented a Guess and Determine attack on MASHA with the time complexity of 2^{320}.

A brief description of the MASHA stream cipher. We recall the MASHA stream cipher briefly, for more details refer to [7]. MASHA can be divided into three parts: two feedback shift registers, FSR-A and FSR-B, and a finite state machine (FSM). The total size of the internal state is 640 bits. The symbols $\oplus$ and $+$ denote bitwise exclusive-or operation and addition modulo 2^{32}, respectively. The structure of MASHA is shown in Figure 1 in [7]. Since our attack is not concerned with the initialization and MAC generation of MASHA, thus here we omit these two parts in the following description.

Denote by A_t^i the i-th register of FSR-A at time t. FSR-A consists of five 32-bit registers and operates with the following recurrence functions:

$$A_{t+1}^4 = \alpha_0 A_t^0 \oplus A_t^1 \oplus A_t^3, \tag{1}$$

$$A_{t+1}^i = A_t^{i+1} \ (0 \leqslant i \leqslant 3). \tag{2}$$

Denote by B_t^i the i-th register of FSR-B at time t. FSR-B consists of eleven 32-bit registers and operates with the following recurrence functions:

$$B_{t+1}^{10} = A_t^3 + M\mathrm{Sub}\left(\alpha_1 B_t^0 [\oplus, +]_{cl0_t} B_t^1 \right.$$
$$\left. \oplus B_t^4 [\oplus, +]_{cl1_t} B_t^6 + B_t^{10} \oplus C_t^{\mathrm{H}}\right), \tag{3}$$

$$B_{t+1}^4 = A_t^1 + M\mathrm{Sub}\left(B_t^5 \oplus C_t^{\mathrm{L}}\right), \tag{4}$$

$$B_{t+1}^i = B_t^{i+1} \ (0 \leqslant i \leqslant 3, 5 \leqslant i \leqslant 9), \tag{5}$$

where the control bits $cl0_t = A_t^2[31]$ and $cl1_t = A_t^2[30]$ are defined as the most significant bit and the second most significant bit of A_t^2, respectively. C_t^{H} and C_t^{L} denote the higher and lower 32-bits of ciphertext C_t, respectively. MSub is a nonlinear transformation and consists of a 32×32 S-box Sub prepended by the AES MixColumn operations.

* Corresponding author (email: dinglin_cipher@163.com)

The FSM consists of four internal 32-bit registers $R1$, $R2$, $L1$ and $L2$. The internal registers are updated as follows:

$$R1_{t+1} = \mathrm{Sub}(L2_t + B_t^5), \tag{6}$$

$$L1_{t+1} = \mathrm{Sub}(R2_t + B_t^0), \tag{7}$$

$$R2_{t+1} = \mathrm{Sub}(R1_t), \tag{8}$$

$$L2_{t+1} = \mathrm{Sub}(L1_t). \tag{9}$$

At time t, a 64-bit keystream $z_t = (z_t^{\mathrm{H}}, z_t^{\mathrm{L}})$ is generated as follows:

$$z_t^{\mathrm{L}} = A_t^4 \oplus (R2_t + R1_t), \tag{10}$$

$$z_t^{\mathrm{H}} = A_t^0 \oplus (L2_t + L1_t), \tag{11}$$

where z_t^{H} and z_t^{L} denote the higher and lower 32-bits of z_t, respectively.

At time t, MASHA encrypts a 64-bit plaintext message $P_t = (P_t^{\mathrm{H}}, P_t^{\mathrm{L}})$ to the ciphertext $C_t = (C_t^{\mathrm{H}}, C_t^{\mathrm{L}})$ as follows:

$$C_t^{\mathrm{L}} = z_t^{\mathrm{L}} \oplus P_t^{\mathrm{L}}, \tag{12}$$

$$C_t^{\mathrm{H}} = z_t^{\mathrm{H}} \oplus P_t^{\mathrm{H}}. \tag{13}$$

Improved Guess and Determine attack on MASHA. In the specification of MASHA [7], the designers propose a simple Guess and Determine attack on the MASHA stream cipher with time complexity of 2^{320}. The main idea of their attack is to guess all components of FSR-A and five components of the remaining cipher, and then determine the remaining unknown components. In our attack on MASHA, we also guess all components of FSR-A, which is the same with their attack. Unlike their attack, we only guess two components of the remaining cipher by optimizing the guess and determination process of their attack. Thus, a total of seven components, i.e., $A_t^0, A_t^1, A_t^2, A_t^3, A_t^4, R1_t, L1_t$, should be guessed in our attack.

For convenience, $A \xrightarrow{(*)} B$ denotes the deduction of B from A by equation (*). The determination process of our attack can be divided into two phases as follows.

Phase one. For a given guess, all bits of the following unknown components can be immediately determined by exploiting the relationships of the cipher.

- $C_t, P_t \xrightarrow{(12,13)} z_t$.
- $z_t^{\mathrm{L}}, A_t^4, R1_t \xrightarrow{(10)} R2_t$.
- $z_t^{\mathrm{H}}, A_t^0, L1_t \xrightarrow{(11)} L2_t$.
- $R1_t \xrightarrow{(8)} R2_{t+1}$.
- $L1_t \xrightarrow{(9)} L2_{t+1}$.
- $A_t^0, A_t^1, A_t^2, A_t^3, A_t^4 \xrightarrow{(1,2)} A_{t+1}^0, A_{t+1}^1, A_{t+1}^2, A_{t+1}^3, A_{t+1}^4$.
- $C_{t+1}, P_{t+1} \xrightarrow{(12,13)} z_{t+1}$.
- $z_{t+1}^{\mathrm{L}}, A_{t+1}^4, R2_{t+1} \xrightarrow{(10)} R1_{t+1}$.
- $z_{t+1}^{\mathrm{H}}, A_{t+1}^0, L2_{t+1} \xrightarrow{(11)} L1_{t+1}$.
- $R1_{t+1}, L2_t \xrightarrow{(6)} B_t^5$.
- $L1_{t+1}, R2_t \xrightarrow{(7)} B_t^0$.

Phase two. The last eight steps of Phase one can be repeated for $i = 1, 2, \ldots, 5$ to determine more unknown components.

- $R1_{t+i} \xrightarrow{(8)} R2_{t+i+1}$.
- $L1_{t+i} \xrightarrow{(9)} L2_{t+i+1}$.
- $A_{t+i}^0, A_{t+i}^1, A_{t+i}^2, A_{t+i}^3, A_{t+i}^4 \xrightarrow{(1,2)} A_{t+i+1}^0, A_{t+i+1}^1, A_{t+i+1}^2, A_{t+i+1}^3, A_{t+i+1}^4$.
- $C_{t+i+1}, P_{t+i+1} \xrightarrow{(12,13)} z_{t+i+1}$.
- $z_{t+i+1}^{\mathrm{L}}, A_{t+i+1}^4, R2_{t+i+1} \xrightarrow{(10)} R1_{t+i+1}$.
- $z_{t+i+1}^{\mathrm{H}}, A_{t+i+1}^0, L2_{t+i+1} \xrightarrow{(11)} L1_{t+i+1}$.
- $R1_{t+i+1}, L2_{t+i} \xrightarrow{(6)} B_{t+i}^5$.
- $L1_{t+i+1}, R2_{t+i} \xrightarrow{(7)} B_{t+i}^0$.

Up to the end of Phase two, we have obtained $R2_t, L2_t$, $B_{t+i}^0, B_{t+i}^5, i = 0, 1, \ldots, 5$. It is easy to know that $B_t^i = B_{t+i}^0$ $(i = 1, \ldots, 4)$ and $B_t^{5+i} = B_{t+i}^5$ $(i = 1, \ldots, 5)$ hold, which means we have obtained all eleven components $B_t^0, \ldots, B_t^{10}$ of FSR-B. Thus, up to now, we have recovered all 640-bit internal state of the MASHA stream cipher. And then we tested the correctness of the recovered internal state by producing a keystream and comparing it with the observed keystream. If the keystreams agree, the recovered state is correct. Otherwise, we repeat the above process until the correct internal state is found.

In our Guess and Determine attack on the MASHA stream cipher, a total of seven 32-bit words, i.e., $A_t^0, A_t^1, A_t^2, A_t^3, A_t^4, R1_t, L1_t$, should be guessed. Then the remaining components can be determined using the process above. Thus, the time complexity of our attack on MASHA is 2^{224}. In the attack, we only utilize 14 keystream words, i.e., $z_t, \ldots, z_{t+6}$, in the determination process, and then about another 6 keystream words are required to verify whether the found internal state is correct or not. Thus, our attack requires about 20 keystream words in total.

Conclusion. So far, no attack on MASHA has been published, besides the security analysis by its designers. In this study, based on optimizing the guess and determination process of the designers' Guess and Determine attack, we propose an improved Guess and Determine attack on MASHA with time complexity of 2^{224}, which improves their attack by a factor of 2^{96}. To the best of our knowledge, this is the best attack on MASHA so far.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 61602514, 61802437, 61272488, 61202491, 61572516, 61272041, 61772547), National Cryptography Development Fund (Grant No. MMJJ20170125), and National Postdoctoral Program for Innovative Talents (Grant No. BX201700153).

References

1 Biham E, Dunkelman O. Cryptanalysis of the A5/1 GSM stream cipher. In: Proceedings of INDOCRYPT 2000, Calcutta, 2000. 43–51
2 Simpson L, Dawson E, Golic J, et al. LILI keystream generator. In: Proceedings of SAC 2000, Ontario, 2000. 248–261
3 Babbage S, Dodd M. The MICKEY stream ciphers. In: New Stream Cipher Designs. Berlin: Springer, 2008. 191–209
4 Li L, Liu X H, Wang Z, et al. An improved attack on clock-controlled shift registers based on hardware implementation. Sci China Inf Sci, 2013, 56: 112107
5 Hu J, Li R L, Tang C J. A real-time inversion attack on the GMR-2 cipher used in the satellite phones. Sci China Inf Sci, 2018, 61: 032113
6 Kiyomoto S, Tanaka T, Sakurai K. K2: a stream cipher algorithm using dynamic feedback control. In: Proceedings of SECRYPT 2007, Barcelona, 2007. 204–213
7 Kiyomoto S, Henricksen M, Yap W, et al. MASHA—low cost authentication with a new stream cipher. In: Proceedings of ISC 2011, Xi'an, 2011. 63–78
8 Feng X, Liu J, Zhou Z, et al. A byte-based Guess and Determine attack on SOSEMANUK. In: Proceedings of ASIACRYPT 2010, Singapore, 2010. 146–157
9 Li R, Li H, Li C, et al. A low data complexity attack on the GMR-2 cipher used in the satellite phones. In: Proceedings of FSE 2013, Singapore, 2013. 485–501