SCIENCE CHINA
Information Sciences

• LETTER •

March 2020, Vol. 63 139112:1–139112:3
https://doi.org/10.1007/s11432-018-9522-6

Ciphertext-only fault analysis on the Midori lightweight cryptosystem

Wei LI[1,2,3,4], Linfeng LIAO[1], Dawu GU[2], Shan CAO[1], Yixin WU[1], Jiayao LI[1], Zhihong ZHOU[4], Zheng GUO[5], Ya LIU[6,2*] & Zhiqiang LIU[2]

[1]*School of Computer Science and Technology, Donghua University, Shanghai* 201620, *China;*
[2]*Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai* 200240, *China;*
[3]*Shanghai Key Laboratory of Scalable Computing and Systems, Shanghai* 200240, *China;*
[4]*Shanghai Key Laboratory of Integrate Administration Technologies for Information Security, Shanghai* 200240, *China;*
[5]*School of Microelectronics, Shanghai Jiao Tong University, Shanghai* 200240, *China;*
[6]*Department of Computer Science and Engineering, University of Shanghai for Science and Technology, Shanghai* 200093, *China*

Received 11 May 2018/Accepted 16 July 2018/Published online 11 February 2020

Citation Li W, Liao L F, Gu D W, et al. Ciphertext-only fault analysis on the Midori lightweight cryptosystem. Sci China Inf Sci, 2020, 63(3): 139112, https://doi.org/10.1007/s11432-018-9522-6

Dear editor,

The Midori lightweight cipher can be applied in the RFID tags and other low-resource devices to protect information on the Internet. This cipher was proposed by Banik et al. and presented at the ASIACRYPT conference [1]. It has a Substitution Permutation Network structure with a 128-bit keysize, and two block-size variants, 64 bits and 128 bits. The optimization of its design and implementation mainly consist of replacing an 8-bit S-box with two 4-bit S-boxes. In 2013, Fuhr et al. [2] proposed the ciphertext-only fault analysis (CFA) on AES. Then in 2016, Dobraunig et al. [3] validated the aforementioned CFA using a physical experiment and broke a series of nonce-based authenticated encryption schemes on AES.

In this study, we propose a CFA with six different distinguishers to successfully break Midori in a software experiment. When random faults are injected into the penultimate round, the attackers can not only use an SEI distinguisher, but also apply goodness of fit (GF), goodness of fit-square Euclidean imbalance (GF-SEI), maximum likelihood (ML), hamming weight (HW), and maximum a posteriori (MAP) distinguishers. Our CFA method requires approximately 280 ciphertexts and 132 ciphertexts with the deeper round of Midori-64 and Midori-128 in the best case, respectively. Table 1 shows the summary of ciphertext-only fault analysis on Midori. The assumption in this study is the weakest in terms of capabilities of the attackers, and thus it is the most practical for real-world applications. The fault model is random nibble for Midori-64 and random byte for Midori-128. The attackers can induce a fault for a nibble or a byte to one layer. However, the value of the fault is unknown. The target can be performed with a bitwise-AND operation by a fault.

Notations. Let $\hat{y}$ represent the faulty ciphertext. Let r represent the number of rounds with $r \in \{16, 20\}$. Let WK represent the whitening key. Let K_1 and K_2 denote the subkeys from the secret key K. Let RK_i represent the i-th round key with $0 \leqslant i \leqslant r-1$. Let A_i, B_i, C_i, and D_i represent the output of the SB, SC, MC, and AK layers, respectively, in the i-th round with $0 \leqslant i \leqslant r-1$. Let $\hat{A}_i$, $\hat{B}_i$, $\hat{C}_i$ and $\hat{D}_i$ be the faulty output of the above layers in the i-th round with $0 \leqslant i \leqslant r-1$. Let SB^{-1}, SC^{-1}, and MC^{-1}represent the inverse operation of the above layers. Let $\sum$ and $\prod$ denote the sum and multiplication of all elements. Let $\sharp$ be the number of elements.

Main procedure.

Step 1. The fault injection targets at the penul-

* Corresponding author (email: liuya@usst.edu.cn)

Table 1 Summary of our ciphertext-only fault analysis on Midori

Distinguisher version	Midori-64			Midori-128		
	Fault model	♯Faults	Time (s)	Fault model	♯Faults	Time (h)
SEI	Nibble	480	0.84	Byte	240	57.71
GF	Nibble	448	0.80	Byte	200	52.43
GF-SEI	Nibble	384	0.68	Byte	180	46.64
MAP	Nibble	400	0.78	Byte	144	38.55
ML	Nibble	296	0.57	Byte	140	35.20
HW	Nibble	280	0.52	Byte	132	35.66

timate round both for Midori-64 and Midori-128. The faulty ciphertext is derived when any plaintext is encrypted with the same secret key.

Step 2. This step is aimed at recovering the whitening key WK in the last round. A fault injection can target at either A_{r-1}, B_{r-1} or C_{r-1}. Any modification of one fault in the penultimate round provokes the faulty ciphertext. The attackers have the following:

$$\begin{aligned}&\hat{B}_{r-1}\\&\quad=\mathrm{MC}^{-1}(\mathrm{SB}^{-1}(\hat{y}\oplus WK)\oplus\mathrm{RK}_{r-1})\\&\quad=\mathrm{MC}^{-1}(\mathrm{SB}^{-1}(\hat{y}\oplus\mathrm{WK}))\oplus\mathrm{MC}^{-1}(\mathrm{RK}_{r-1})\\&\quad=\mathrm{MC}^{-1}(\mathrm{SB}^{-1}(\hat{y}\oplus\mathrm{WK})),\end{aligned}$$

where the XOR operation with $\mathrm{MC}^{-1}(\mathrm{RK}_{r-1})$ does not alter the distance of the biased distribution from the uniform distribution. The attackers can leverage various statistical analyses of the target of $\hat{B}_{r-1}$ to recover three nibbles or bytes of WK. A list of possible $\hat{B}_{r-1}$ can be deduced by the candidates of WK. Then the attackers derive the right WK by the maximum or minimum value of a distinguisher. The distinguishers, in conjunction with multiple faulty ciphertexts $\hat{y}$, enable the collection of a list of possible candidates for WK. Subsequently, the attackers continue to perform a brute-force search for three nibbles or bytes of WK, until the set of WK candidates has only one element.

Step 3. This step is aimed at recovering the secret key K of Midori. For Midori-128, the attackers can directly derive $K = \mathrm{WK}$. For Midori-64, they can apply WK to decrypt the last two rounds, and obtain the input of the $(r-2)$-th round, denoted as B_{r-2}. The attackers continues injecting random faults before C_{r-2} in the $(r-2)$-th round, and then have the following:

$$\begin{aligned}&\hat{B}_{r-2}\\&\quad=\mathrm{MC}^{-1}(\mathrm{SB}^{-1}(\mathrm{SC}^{-1}(\hat{y}'\oplus\mathrm{RK}_{r-1}))\oplus\mathrm{RK}_{r-2})\\&\quad=\mathrm{MC}^{-1}(\mathrm{SB}^{-1}(\mathrm{SC}^{-1}(\hat{y}'\oplus\mathrm{RK}_{r-1}))),\end{aligned}$$

where $\hat{y}' = \mathrm{MC}^{-1}(\mathrm{SB}^{-1}(\hat{y}\oplus\mathrm{WK}))$. Hence, they can use any of the distinguishers to derive all nibbles of the last round key RK_{r-1} with $r = 16$. The secret key K is deduced as

$$\begin{aligned}K&=K_0||K_1=K_0||(K_0\oplus\mathrm{WK})\\&=(\mathrm{RK}_{15}\oplus\alpha_{14})||(\mathrm{RK}_{15}\oplus\alpha_{14}\oplus\mathrm{WK}).\end{aligned}$$

In Step 2, six distinguishers are listed as follows:

Square Euclidean imbalance (SEI) measures the distance from an unknown distribution to a uniform distribution. The attackers do not need to know the specific distribution of a nibble or a byte, which only satisfies a non-uniform distribution. They have the following:

$$\mathrm{SEI}=\sum_{m=0}^{M-1}\left(\frac{\sharp\{\hat{B}_{r-1}|\hat{B}_{r-1}=m,\hat{B}_{r-1}\in\hat{\Upsilon}\}}{N}-\frac{1}{M}\right)^2,$$

where M denotes the total number of one nibble or byte, $m\in[0,M-1]$, N represents the number of all injecting faults, $\hat{B}_{r-1}$ denotes the faulty value of B_{r-1}, and $\hat{\Upsilon}$ represents the set of all $\hat{B}_{r-1}$. Here, $M\in\{2^4,2^8\}$. The correct WK maximizes the SEI values. Hence, the attackers can compute the maximum value of SEI to distinguish WK.

GF describes how well it fits a set of observations. It summarizes the discrepancy between the observed values and the expected values for the model in question. In our analysis, the distribution of an injected faulty nibble or byte is known. The attackers can do brute-force search on the bitwise-AND operation of two nibbles or bytes. There is

$$\mathrm{GF}=\sum_{m=0}^{M-1}\frac{(O_m-E_m)^2}{E_m},$$

where

$$\begin{aligned}O_m&=\sharp\{\hat{B}_{r-1}|\hat{B}_{r-1}=m,\hat{B}_{r-1}\in\hat{\Upsilon},\\&\qquad m\in[0,M-1]\},\\E_m&=\sharp\{\bar{B}_{r-1}|\bar{B}_{r-1}=m,\dot{B}_{r-1}\in\bar{\Upsilon},\\&\qquad m\in[0,M-1]\},\end{aligned}$$

where M denotes the total number of one nibble or byte, O_m is an expected number for m, E_m represents a theoretical number for m, $\hat{B}_{r-1}$ is the observed faulty value of B_{r-1}, $\bar{B}_{r-1}$ represents the expected value of B_{r-1}, $\hat{\Upsilon}$ denotes the set of all

observed $\hat{B}_{r-1}$, and $\bar{\Upsilon}$ represents the set of all observed $\bar{B}_{r-1}$. Here, $M \in \{2^4, 2^8\}$. The correct WK minimizes the GF values.

GF-SEI is a double distinguisher that is used to combine the advantages of the aforementioned single GF distinguisher and single SEI distinguisher. Specifically, if GF $> \chi_a^2$, then $\hat{B}_{r-1}$ can reject the known distribution. χ_a^2 can be deduced by the known degree- of- freedom df and the defined a significance level of the χ^2-distribution. Here, $M \in \{2^4, 2^8\}$, and df $= M - 1 \in \{15, 255\}$. In our analysis, the GF distinguisher is significantly effective when $N \geqslant 50$ and $E_m = \sharp\{\bar{B}_{r-1} | \bar{B}_{r-1} = m, \bar{B}_{r-1} \in \bar{\Upsilon}, m \in [0, M-1]\} \geqslant 5$. The attackers can exclude the wrong candidate for WK using a GF distinguisher, and then deduce the correct WK by using an SEI distinguisher. The correct WK first satisfies GF $\leqslant \chi_a^2$ and then maximizes the SEI value.

The MAP probability estimate is a distinguisher for an unknown quantity that is equal to the mode of the posterior distribution. It utilizes an augmented optimization objective which integrates a prior distribution over the quantity that is to be estimated. There is

$$\mathrm{MAP} = \frac{p(\hat{\Upsilon}|\mathrm{WK}^t) \cdot \pi(\mathrm{WK}^t)}{\sum_{t=0}^{T-1} p(\hat{\Upsilon}|\mathrm{WK}^t) \cdot \pi(\mathrm{WK}^t)},$$

where T represents the total number of three nibbles or bytes in a subkey, $t \in [0, T-1]$, $\pi(\mathrm{WK}^t)$ represents the prior distribution of WK^t, and $p(\hat{\Upsilon}|\mathrm{WK}^t)$ denotes the conditional probability of $\hat{\Upsilon}$ when the parameter is WK^t, respectively. Here, $T \in \{2^{12}, 2^{24}\}$. It is the correct WK that maximizes the MAP values.

ML estimation is a distinguisher for estimating the parameters of a given distribution model by determining the parameter value that maximizes the likelihood. The attackers can extend the ML distinguisher to the random fault model in the deeper round as follows:

$$\mathrm{ML} = \prod_{n=0}^{N-1} p(\hat{B}_{r-1}),$$

where N represents the number of faults, $n \in [0, N-1]$, p is the probability of the element, and $\hat{B}_{r-1}$ represents the observed faulty value of B_{r-1}. The correct WK maximizes the ML value.

HW represents the number of non-zero bits of a nibble or a byte in the fault analysis. The attackers can apply the HW distinguisher to the random fault model in the deeper round. There is

$$\mathrm{HW} = \frac{1}{N} \sum_{n=0}^{N-1} \mathrm{hw}(\hat{B}_{r-1}),$$

where N represents the number of faults, $n \in [0, N-1]$, hw denotes the hamming weight of the element, and $\hat{B}_{r-1}$ describes the observed faulty value of B_{r-1}. On the basis of the bitwise-AND operation in our fault model, the attackers can compute the minimum value of HW to distinguish WK. The experimental results are shown in Appendix A.

Conclusion. This study details the CFA with six distinguishers on Midori in the random nibble-oriented or byte-oriented fault model. The analysis can break the secret keys of Midori with at least 280 and 132 faults. This shows that the CFA is a strong threat to Midori.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 61772129, 61472250, 61672347, 61402288, 61402286, 61572192), Shanghai Natural Science Foundation (Grant Nos. 15ZR1400300, 16ZR1401100), and Opening Project of Shanghai Key Laboratory of Integrate Administration Technologies for Information Security (Grant No. AGK201703), Opening Project of Shanghai Key Laboratory of Scalable Computing and Systems, National Cryptography Development Fund (Grant No. MMJJ20180101), Fundamental Research Funds for the Central Universities, and Foundation of Science and Technology on Information Assurance Laboratory (Grant No. KJ-17-008).

Supporting information Appendix A. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Banik S, Bogdanov A, Isobe T, et al. Midori: a block cipher for low energy. In: Proceedings of International Conference on the Theory and Application of Cryptology and Information Security, Hanoi, 2015. 411–436
2 Fuhr T, Jaulmes E, Lomné V, et al. Fault attacks on AES with faulty ciphertexts only, In: Proceedings of Workshop on Fault Diagnosis and Tolerance in Cryptography, Washington, 2013. 108–118
3 Dobraunig C, Eichlseder M, Korak T, et al. Statistical fault attacks on nonce-based authenticated encryption schemes. In: Proceedings of International Conference on the Theory and Application of Cryptology and Information Security, Hanoi, 2016. 369–395