Saturnin-Short 轻量级认证加密算法的统计无效故障分析

李玮[1,2,3,4]，刘春[1]，谷大武[2]，孙文倩[1]，高建宁[1]，秦梦洋[1]

（1. 东华大学计算机科学与技术学院，上海 201620；2. 上海交通大学计算机科学与工程系，上海 200204；
3. 上海市可扩展计算与系统重点实验室，上海 200204；4. 上海市信息安全综合管理技术研究重点实验室，上海 200093）

摘 要：面向随机单字节故障模型和唯密文攻击假设，提出了一种针对 Saturnin-Short 算法的统计无效故障分析方法。该方法基于统计分布和无效状态分析，通过结合故障注入前后中间状态的变化，设计并采用了概率对称卡方–极大似然估计和调和中项–汉明重量新型区分器，最少仅需 1 097 个无效故障并以不低于 99%的成功率恢复 Saturnin-Short 算法的 256 bit 原始密钥。实验分析表明，所提区分器不仅降低了故障注入数，而且减少了攻击时间和复杂度。因此，Saturnin-Short 算法不能抵抗统计无效故障分析的攻击。研究结果为其他轻量级认证加密算法的安全性分析提供了重要参考。

关键词：Saturnin-Short；认证加密；统计无效故障分析；密码分析

中图分类号：TP309.7

文献标志码：A

DOI: 10.11959/j.issn.1000−436x.2023084

Statistical ineffective fault analysis of the lightweight authenticated cipher algorithm Saturnin-Short

LI Wei[1,2,3,4], LIU Chun[1], GU Dawu[2], SUN Wenqian[1], GAO Jianning[1], QIN Mengyang[1]

1. School of Computer Science and Technology, Donghua University, Shanghai 201620, China
2. Department of Computer and Science and Engineering, Shanghai Jiao Tong University, Shanghai 200204, China
3. Shanghai Key Laboratory of Scalable Computing and System, Shanghai 200204, China
4. Shanghai Key Laboratory of Integrate Administration Technologies for Information Security, Shanghai 200093, China

Abstract: On the random single byte-oriented fault model and the assumption of ciphertext-only attack, a statistical ineffective fault analysis of the Saturnin-Short cipher was proposed. The analysis combined the statistical distribution with the ineffective analysis, and discussed the difference between intermediate states before and after fault injections. A variety of dual distinguishers was designed, such as the probabilistic symmetric Chi-square-maximum likelihood estimate, and harmonic mean-Hamming weight. It only required at least 1 097 ineffective faults to recover the 256 bit secret key with a success rate of at least 99%. The experimental results show that the proposed distinguishers can not only decrease fault injections, but also reduce the attacking time and complexities. Therefore, the Saturnin-Short cipher cannot resist against the statistical ineffective fault analysis. It provides an important reference for the security analysis of other lightweight authenticated ciphers.

Keywords: Saturnin-Short, authenticated cipher, statistical ineffective fault analysis, cryptanalysis

---

收稿日期：2022–12–26；**修回日期：**2023–03–29

基金项目：国家自然科学基金资助项目（No.61772129, No.61932014, No.62102077）；国家密码发展基金资助项目（No.MMJJ20180101）；信息安全国家重点实验室开放课题资助项目（No.2021-MS-05）；上海市扬帆计划基金资助项目（No.21YF1401200, No.23YF1401000）；中央高校基本科研业务费专项资金资助项目（No.2232022D-25）

Foundation Items: The National Natural Science Foundation of China (No.61772129, No.61932014, No.62102077), The National Cryptography Development Foundation of China (No.MMJJ20180101), The Open Fund Program for State Key Laboratory of Information Security of China (No.2021-MS-05), Shanghai Sailing Program (No.21YF1401200, No.23YF1401000), The Fundamental Research Funds for the Central Universities (No.2232022D-25)

0 引言

随着信息社会的快速发展，越来越多的智能卡、微控制器等设备接入物联网环境中，用于收集、传输和处理大量重要数据。这些数据涉及个人隐私和商业秘密等，不可避免地面临着泄露、篡改和伪造等多种安全风险，对物联网的安全造成极大威胁[1-2]。由于物联网中的微型设备受到计算、存储和功耗等诸多方面的制约，传统的密码算法难以有效平衡安全、效率和灵活性。因此，近年来国内外研究学者提出并设计了一系列轻量级密码算法来保障物联网的信息保密性、完整性和认证性[3-6]。

2020 年，Canteaut 等[7]提出了一种具有后量子安全的轻量级认证加密算法 Saturnin-Short，采用了分组密码 Saturnin 算法实现加密和解密，可以同时保证数据的保密性和认证性。算法可以根据实际需要选择合适的加（解）密轮数，范围为 10～31 轮。Canteaut 等通过 Walsh 变换计算出 S 盒的线性度，并推导出 8 轮线性轨迹的最高平方相关值为 $2^{-441.5}$。同时，他们结合分析 S 盒的差分均匀度、非平凡差分、不可能差分变换等，获得缩减轮 8 轮差分分析、7 轮不可能差分分析以及 7.5 轮中间相遇分析的攻击结果[7]。目前，该算法可以抵抗常见的传统密码分析方法，如线性分析、差分分析、不可能差分分析和中间相遇分析，这些分析方法的基本假设集中于已知明文攻击或选择明文攻击。目前，国内外尚未有 Saturnin-Short 轻量级认证加密算法基于唯密文攻击的相关研究成果，表 1 对比了该算法的相关安全性分析。

表 1 Saturnin-Short 密码算法的安全性分析对比

分析类型	基本假设	攻击最高轮数/轮	方法
线性分析	已知明文攻击	8	文献[7]
差分分析	选择明文攻击	8	文献[7]
不可能差分分析	选择明文攻击	7	文献[7]
中间相遇分析	选择明文攻击	7.5	文献[7]
统计无效故障分析	唯密文攻击	31	本文方法

在物联网领域中，密码算法通常使用硬件或以硬件为表现形式的软件来实现，并运行在密码载体设备中。此时，攻击者可以利用异常电流、异常时钟、激光照射等物理方式干扰密码变换的正常过程，从而收集到有用的数据来分析和破译密码，这种攻击方式称为故障分析[8]。近年来，故障分析已逐渐成为检测密码算法实现安全的重要指标之一，其常见种类有统计故障分析、无效故障分析、差分故障分析、代数故障分析、中间相遇故障分析和持久故障分析等[9-17]。其中，统计故障分析是指通过统计分布，结合受故障影响后的中间状态分布律，完成破译密码[9,18-19]。无效故障分析是指利用受无效故障影响的信息和中间状态之间的依赖关系来恢复密码的原始密钥[20-21]。2018 年，Dobraunig 等[22]在国际密码硬件和嵌入式系统（CHES, cryptographic hardware and embedded systems）会议上提出了统计无效故障分析方法，充分结合了统计故障分析和无效故障分析的特点，仅需使用无效故障产生的正确输出，通过推导中间状态的理论分布律及差异性，即可完成破译密码算法。近年来，该方法充分结合瞬时故障、持久故障、无效故障和有效故障等多种故障类型，有效地绕过具有检测机制的故障防御策略，对密码的安全实现具有巨大的威胁力[23-24]。

在密码算法的安全性分析中，基本假设作为重要前提，表明了攻击者能力的强弱和实现的难易程度，如唯密文攻击、已知明文攻击、选择明文攻击和选择密文攻击等。其中，唯密文攻击对攻击者能力要求最弱，仅需截获密文即可，在实际应用中更易实现。统计无效故障分析的基本假设是唯密文攻击。2018 年，Dobraunig 等[25]针对 AES 算法及工作模式的统计无效故障分析，采用随机单字节故障模型，使用平方欧氏距离（SEI, square Euclidean imbalance）区分器，仅需 602 个无效故障和 1 808 个总故障，就可以恢复出 128 bit 密钥。2020 年，Gruber 等[26]在轻量级认证加密算法 GIMLI 的统计无效故障分析中，采用随机单字节故障模型，使用平方欧氏距离区分器，仅需要 5 803 个无效故障和 58 027 个总故障，即可破译 256 bit 密钥。

目前，国内外尚未有轻量级认证加密算法 Saturnin-Short 抵御统计无效故障分析的公开成果。本文基于唯密文攻击的基本假设，提出了面向随机单字节故障模型的统计无效故障分析方法，设计了概率对称卡方–极大似然估计（PSC-MLE, probabilistic symmetric Chi-square-maximum likelihood estimate）和调和中项–汉明重量（HM-HW, harmonic mean-Hamming weight）新型区分器。如表 2 所示，该方法能以不低于 99% 的成功率恢复 Saturnin-Short 密码的 256 bit 原始密钥，并且新型

区分器在无效故障数、总故障数、耗时、成功率和复杂度方面均具有较大优势。

1　Saturnin-Short 算法简介

1.1　符号说明

本文符号说明如表 3 所示。

1.2　Saturnin-Short 认证加密算法

Saturnin-Short 是于 2020 年提出的一种具有后量子安全的轻量级认证加密算法[7]。该算法的输入包括不超过 128 bit 的明文、256 bit 原始密钥、128 bit Nonce 以及加（解）密轮数 R，输出为明文和密文或空集，其中，$10 \leqslant R \leqslant 31$。256 bit 中间状态数据均采用 4×4×4 的立方体结构，数据单元为 4 bit 。

Saturnin-Short 认证加密算法包含以下 4 个部分。

1) 填充。当明文输入不足 128 bit 时，继续加比特 1 和连续多个比特 0，直到达到 128 bit。

2) 加密。128 bit 填充明文与 128 bit Nonce 连接，与 256 bit 原始密钥相互作用，经过 R 轮加密后，输出为 256 bit 密文。

3) 解密。256 bit 密文与原始密钥相互作用，经过 R 轮解密后，输出待认证 Nonce 和待认证明文。

4) 认证。将加密输入中的 Nonce 和解密输出的待认证 Nonce 进行比较，若两者相同，认证通过且输出明文和密文；否则，认证不通过且输出空集。

Saturnin-Short 认证加密算法如算法 1 所示。

算法 1　Saturnin-Short 认证加密算法

输入　X, N, K, R

输出　$\overline{X}, Y$ 或 $\varnothing$

1)　$\hat{X} = \mathrm{Padding}(X)$；　　//填充

2)　$Y = \mathrm{Saturnin}_R(K, R, N \| \hat{X})$；　　//加密

3)　$\overline{N} \| \overline{X} = \mathrm{Saturnin}_R^{-1}(K, R, Y)$；　　//解密

4)　if $\overline{N} = N$ then　　//认证

5)　　return $\overline{X}, Y$；

6)　else

7)　　return $\varnothing$；

8)　end if

1.3　加密和解密

轻量级认证加密算法 Saturnin-Short 采用分组密码 Saturnin 完成加密和解密。Saturnin 密码的分组、密钥和输出长度均为 256 bit，采用典型的代换置换网络结构。

Saturnin-Short 算法的加密过程如图 1 所示，其中，$F_1, F_2, \cdots, F_R$ 表示轮变换，具体包括 S 盒变换（SC, S-box convert）、列混淆（MC, mixed column）、纵

表 2　**Saturnin-Short 密码不同区分器下恢复 256 bit 原始密钥所需故障数、耗时、成功率和复杂度对比**

区分器	无效故障数/个	总故障数/个	耗时/s	成功率	时间复杂度	数据复杂度	存储复杂度
平方欧氏距离（SEI）	∞	∞	∞	—	∞	∞	∞
极大似然估计（MLE）	1 485	14 861	8.152	⩾99%	$2^{14.64}$	$2^{18.54}$	$2^{18.55}$
汉明重量（HW）	1 454	14 542	7.787	⩾99%	$2^{14.61}$	$2^{18.51}$	$2^{18.52}$
概率对称卡方–极大似然估计（PSC-MLE）	1 228	12 296	5.821	⩾99%	$2^{14.37}$	$2^{18.26}$	$2^{18.28}$
调和中项–汉明重量（HM-HW）	1 097	10 981	4.532	⩾99%	$2^{14.21}$	$2^{18.10}$	$2^{18.12}$

表 3　**符号说明**

符号	说明
X、$\hat{X}$、$\overline{X}$、Y、N、$\overline{N}$	明文、填充明文、待认证明文、密文、Nonce 和待认证 Nonce
K、K_{rot}、R、C_0、C_r	原始密钥、旋转密钥、加（解）密轮数、初始常数和第 r 轮的轮常数，其中，$1 \leqslant r \leqslant R$，$10 \leqslant R \leqslant 31$
$\mathrm{Saturnin}_R$、$\mathrm{Saturnin}_R^{-1}$、Padding	加密过程、解密过程和填充过程
SC、MC、$\mathrm{SR}_{\mathrm{slice}}$、$\mathrm{SR}_{\mathrm{sheet}}$	S 盒变换、列混淆、纵切片行移位和横切片行移位
SC^{-1}、MC^{-1}、$\mathrm{SR}_{\mathrm{slice}}^{-1}$、$\mathrm{SR}_{\mathrm{sheet}}^{-1}$	SC、MC、$\mathrm{SR}_{\mathrm{slice}}$ 和 $\mathrm{SR}_{\mathrm{sheet}}$ 的逆操作
AK、$\mathrm{AK}_{\mathrm{rot}}$、IC、LFSR	密钥加、旋转密钥加、初始化轮常数和线性反馈移位
T、T^{ℓ}、K^{ℓ}、K_{rot}^{ℓ}	第 R–1 轮 $\mathrm{AK}_{\mathrm{rot}}$ 运算后输出的中间状态值及 T、K 和 K_{rot} 的第 ℓ 字节，$0 \leqslant \ell \leqslant 63$
⊕、mod、‖、∅	异或、模、连接和空集

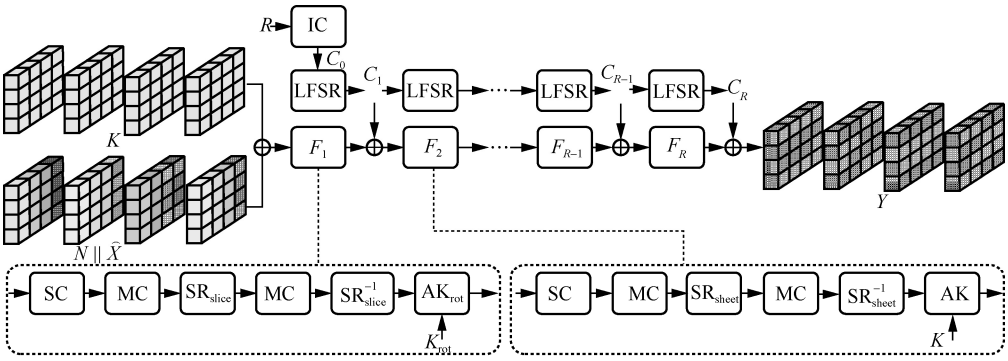


图 1 Saturnin-Short 算法的加密过程

切片行移位（SR_{slice}, shift row slice）、横切片行移位（SR_{sheet}, shift row sheet）、密钥加（AK, add key）、旋转密钥加（AK_{rot}, add rotated key）、初始化常数（IC, initialization constant）、线性反馈移位寄存器（LFSR, linear feedback shift register），变换过程如图 2、图 3、表 4 所示[7]。

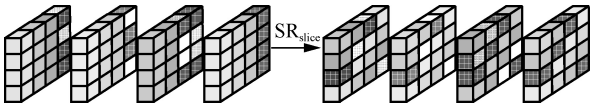


图 2 纵切片行移位

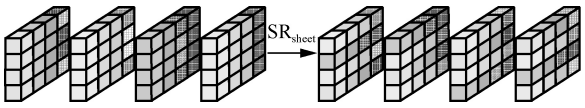


图 3 横切片行移位

表 4 **S 盒变换**

输入	偶数索引	奇数索引
0	0	0
1	6	9
2	14	13
3	1	2
4	15	15
5	4	1
6	7	11
7	13	7
8	9	6
9	8	4
10	12	5
11	5	3
12	2	8
13	10	12
14	3	10
15	11	14

算法 2 给出了 Saturnin-Short 算法的加密过程，解密过程为加密的逆过程。

算法 2 Saturnin-Short 算法加密过程

输入 $\hat{X}, N, K, R$

输出 Y

1) $S = (N \| \hat{X}) \oplus K$；
2) $C_0 = \mathrm{IC}(R)$；
3) for $r = 1$ to R do
4) $C_r = \mathrm{LFSR}(C_{r-1})$；
5) $S = \mathrm{MC}(\mathrm{SC}(S))$；
6) if $r \bmod 2 = 1$ then
7) $S = \mathrm{AK}_{\mathrm{rot}}(\mathrm{SR}_{\mathrm{slice}}^{-1}(\mathrm{MC}(\mathrm{SR}_{\mathrm{slice}}(S))))$；
8) else
9) $S = \mathrm{AK}(\mathrm{SR}_{\mathrm{sheet}}^{-1}(\mathrm{MC}(\mathrm{SR}_{\mathrm{sheet}}(S))))$；
10) end if
11) $S = S \oplus C_r$；
12) end for
13) $Y = S$；
14) return Y

2 统计无效故障分析

2.1 基本假设和故障模型

本文的基本假设为唯密文攻击，故障模型采用随机单字节。也就是说，攻击者使用相同原始密钥对明文进行加密，可以获得随机密文。在攻击过程中，单字节故障按位“与”的方式被注入加密过程中，由于算法采用的 S 盒输入为半字节，因此单字节故障可看作两个半字节连接，图 4 给出了受无效故障影响后的半字节分布律。

攻击者在对 Saturnin-Short 算法进行统计无效故障分析时，根据算法输出的结果可以将故障分成有效故障和无效故障。

1) 若输出为空，即算法检测到篡改攻击，则注入的故障为有效故障。

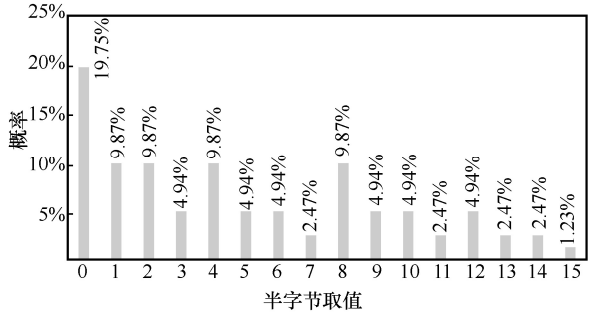


图 4　受无效故障影响后的半字节分布律

2) 若输出为非空，即算法没有检测到攻击，则注入的故障为无效故障。

2.2　区分器

2018 年，Dobraunig 等[25]使用 SEI 区分器对 AES 密码及工作模式进行统计无效故障分析。2020 年，Gruber 等[26]利用该攻击并使用相同区分器对 GIMLI 密码进行破译。本文使用平方欧氏距离、极大似然估计（MLE, maximum likelihood estimate）、汉明重量（HW, Hamming weight）区分器，并提出了 PSC-MLE 和 HM-HW 双重区分器。

1) 平方欧氏距离区分器

欧氏距离是由数学家 Euclid 提出的一种计算距离的度量方法，其平方值为平方欧氏距离。平方欧氏距离区分器可计算实际分布的样本值与均匀分布的理论值之间距离，并选取最大距离值对应的候选密钥，可以应用于统计无效故障分析 AES 和 GIMLI 的分析中[25-26]。该区分器表达式为

$$\mathrm{SEI}=\sum_{\varepsilon=0}^{w-1}\left(\frac{O_\varepsilon}{l}-\frac{1}{w}\right)^2$$

其中，l 表示注入无效故障数量，w 表示半字节所有可能取值的个数，O_ε 表示中间状态值为 ε 的个数。理论上，中间状态值的取值是均匀分布的。由于注入随机故障后正确密钥对应的中间状态值的分布是偏离均匀分布的，因此，当 SEI 取最大值时，对应的候选密钥为正确密钥。

2) 极大似然估计区分器

极大似然估计是 Gauss 等于 1821 年提出的一种参数估计方法，1922 年 Fisher 等[27]推广并深化该方法。极大似然估计区分器通过似然函数来计算每一组样本值理论应该出现的概率，并选出最大概率值对应的候选密钥，最早被 Fuhr 等[9]应用于统计故障分析 AES 算法中，未用于统计无效故障分析中。该区分器表达式为

$$\mathrm{MLE}=\prod_{\lambda=1}^{l}p(\varepsilon_\lambda)$$

其中，l 表示注入无效故障数量，ε_λ 表示第 λ 个故障推导出来的中间状态值，$p(\varepsilon_\lambda)$ 表示中间状态值为 ε_λ 的理论概率。当 MLE 取最大值时，对应的候选密钥为正确密钥。

3) 汉明重量区分器

1954 年，Reed[28]提出了一个计算二进制字符串的汉明重量的概念。汉明重量区分器可计算中间状态值的二进制中非零的个数，并选出最小个数对应的候选密钥，首次被 Fuhr 等[9]应用于 AES 算法的统计故障分析中，未用于统计无效故障分析中。该区分器表达式为

$$\mathrm{HW}=\sum_{\lambda=1}^{l}\mathrm{hw}(\varepsilon_\lambda)$$

其中，l 表示注入无效故障数量，ε_λ 表示第 λ 个故障推导出来的中间状态值，$\mathrm{hw}(\varepsilon_\lambda)$ 表示 ε_λ 的汉明重量。由于注入随机故障会破坏中间状态值 0 和 1 的平衡，受无效故障影响的半字节呈现 0 比 1 多的情况，因此，当汉明重量取最小值时，对应的候选密钥为正确密钥。

4) 概率对称卡方–极大似然估计区分器

概率对称卡方是一种计算分布之间距离的度量方法[29]。概率对称卡方区分器可计算实验样本值分布与受无效故障影响的理论值分布之间的距离程度，其表达式为

$$\mathrm{PSC}(k_w)=\sum_{\varepsilon=0}^{w-1}\frac{2(p(\varepsilon)-q(\varepsilon))^2}{p(\varepsilon)+q(\varepsilon)}$$

其中，w 表示半字节所有可能取值的个数，$p(\varepsilon)$ 表示中间状态值为 ε 的理论概率，$q(\varepsilon)$ 表示中间状态值为 ε 的实际概率。当 $\mathrm{PSC}(k_w)$ 取最小值时，对应的候选密钥为正确密钥。

由于仅由正确密钥推导得到的实验样本值的分布才会接近受无效故障影响的理论值分布，因此使用概率对称卡方区分器时，所得计算值越小的实验样本值所对应的候选密钥是正确密钥的概率越高。概率对称卡方–极大似然估计区分器是本文提出的结合 PSC 区分器和 MLE 区分器的双重区分器。该区分器依次计算待测样本的 PSC 值和 MLE 值，从若干组 MLE 较优值中取 PSC 最优值，对应的候选密钥为正确密钥。

5) 调和中项–汉明重量区分器

调和平均数是应用在数学中计算平均数方法

之一。调和中项区分器是通过调和平均数公式，计算样本值与受无效故障影响的理论值之间的调和平均数，并选取最大的调和平均数对应的候选密钥，其表达式为

$$\mathrm{HM}=\sum_{\varepsilon=0}^{w-1}\frac{p(\varepsilon)q(\varepsilon)}{p(\varepsilon)+q(\varepsilon)}$$

其中，w 表示半字节所有可能取值的个数，$p(\varepsilon)$ 表示中间状态值为 ε 的理论概率，$q(\varepsilon)$ 表示中间状态值为 ε 的实际概率。当 HM 取最大值时，对应的候选密钥为正确密钥。

由于调和中项区分器在计算调和平均数不需要考虑总体单位数，因此在筛选密钥过程中具有很好的表现。调和中项–汉明重量区分器是本文提出的结合 HM 区分器和 HW 区分器的双重区分器。该区分器依次计算待测样本的 HW 值和 HM 值，从若干组 HW 较优值中取 HM 最优值，对应的候选密钥为正确密钥。

表 5 给出了本文使用的不同区分器的取值和说明。

表 5　　不同区分器的取值和说明

区分器	取值范围	筛选过程
SEI	最大值	评估实际分布和均匀分布之间的距离，选出距离相差最大的统计样本
MLE	最大值	评估中间状态的出现概率，选出概率最大的统计样本
HW	最小值	评估中间状态的汉明重量，选出汉明重量最小的统计样本
PSC-MLE	PSC 最小值和 MLE 最大值	先后使用 PSC 和 MLE，选出距离程度最小且出现概率最大的统计样本
HM-HW	HM 最大值和 HW 最小值	先后使用 HM 和 HW，选出调和平均数最大且汉明重量最小的统计样本

2.3 攻击过程

攻击过程具体包括以下 4 个步骤。

步骤 1　攻击者采用相同原始密钥 K 对明文进行加密，并注入随机单字节故障，生成多个随机密文。故障注入的位置为加密过程的倒数第二轮 $\mathrm{AK}_{\mathrm{rot}}$ 运算后，故障扩散路径如图 5 所示，其中，阴影部分表示受故障影响的单元。

步骤 2　攻击者利用认证后的密文 Y，推导出原始密钥 K 和倒数第二轮 $\mathrm{AK}_{\mathrm{rot}}$ 运算后输出中间状态值 T 的关系为

$$T=\mathrm{SC}^{-1}(\mathrm{MC}^{-1}(\mathrm{SR}_{\mathrm{sheet}}^{-1}(\mathrm{MC}^{-1}(\mathrm{SR}_{\mathrm{sheet}}((Y\oplus C_R)\oplus K)))))\oplus C_{R-1}=\mathrm{SC}^{-1}(\mathrm{MC}^{-1}(\mathrm{SR}_{\mathrm{sheet}}^{-1}(\mathrm{MC}^{-1}(\mathrm{SR}_{\mathrm{sheet}}(Y\oplus C_R))))\oplus \mathrm{MC}^{-1}(\mathrm{SR}_{\mathrm{sheet}}^{-1}(\mathrm{MC}^{-1}(\mathrm{SR}_{\mathrm{sheet}}(K)))))\oplus C_{R-1}=\mathrm{SC}^{-1}(T_Y\oplus T_K)\oplus C_{R-1}$$

其中，

$$T_Y=\mathrm{MC}^{-1}(\mathrm{SR}_{\mathrm{sheet}}^{-1}(\mathrm{MC}^{-1}(\mathrm{SR}_{\mathrm{sheet}}(Y\oplus C_R))))$$

$$T_K=\mathrm{MC}^{-1}(\mathrm{SR}_{\mathrm{sheet}}^{-1}(\mathrm{MC}^{-1}(\mathrm{SR}_{\mathrm{sheet}}(K))))$$

攻击者利用步骤 1 中的随机密文 Y 能求解出 T_Y。因此，针对原始密钥 K 的变换值 T_K，攻击者均可以收集到对应的中间状态值 T。

步骤 3　攻击者通过分析图 5 的故障扩散路径，利用 T_Y 和 T_K 的 2 个 4 bit，分别推导出受故障影响后中间状态值 T 的 2 个 4 bit 为

$$T^t=\mathrm{SC}^{-1}(T_Y^t\oplus T_K^t)\oplus C_{R-1}^t$$

$$T^{t+4}=\mathrm{SC}^{-1}(T_Y^{t+4}\oplus T_K^{t+4})\oplus C_{R-1}^{t+4}$$

其中，$(t,t+4)\in\{(0,4),(1,5),(2,6),(3,7),(8,12),(9,13),(10,14),(11,15),(16,20),(17,21),(18,22),(19,23),(24,28),(25,29),(26,30),(27,31),(32,36),(33,37),(34,38),(35,39),(40,44),(41,45),(42,46),(43,47),(48,52),(49,53),(50,54),(51,55),(56,60),(57,61),(58,62),(59,63)\}$。

攻击者选定 2.2 节中的区分器后，根据输入每一个

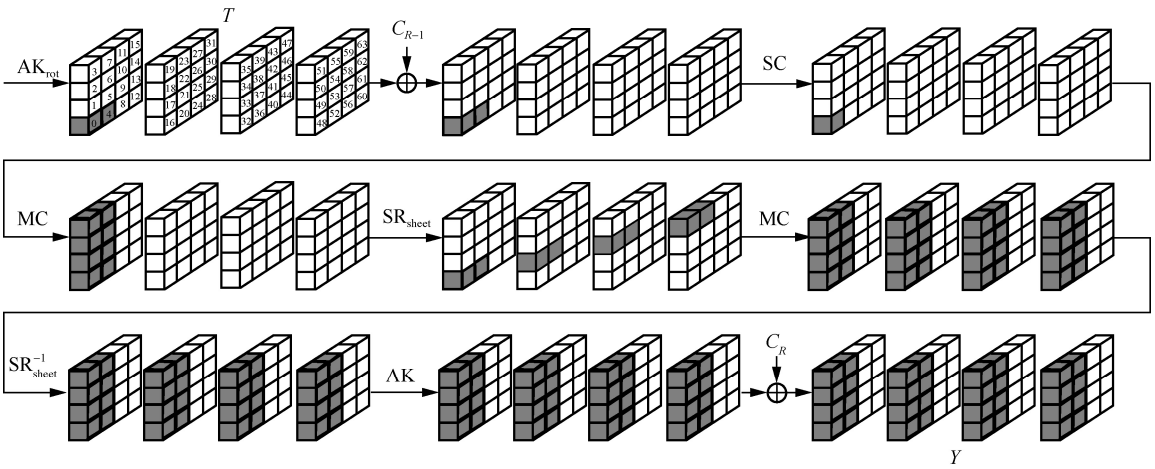


图 5　Saturnin-Short 算法加密最后两轮运算的单字节故障扩散路径

候选密钥推导出来的一组中间状态值，可以获得对应该候选密钥的区分器值，并选择区分器最大值或者最小值对应的候选密钥字节，即对应 T_K 的 8 bit。

步骤 4　攻击者通过多次随机故障注入并重复步骤 1～步骤 3，逐个字节恢复出 256 bit 的 T_K。256 bit 的原始密钥 K 为

$$K = \mathrm{SR}_{\mathrm{sheet}}^{-1}(\mathrm{MC}(\mathrm{SR}_{\mathrm{sheet}}(\mathrm{MC}(T_K))))$$

3　实验分析

本文实验使用的计算机设备配置为 Intel(R) Core(TM) i5-7200U CPU，使用 C++语言编程实现 Saturnin-Short 认证加密算法的统计无效故障分析过程，注入故障的动作由计算机软件模拟实现。本文共进行了 10 000 次实验，每次实验包括故障注入以及恢复原始密钥的过程。

3.1　故障数

在统计无效故障分析中，总故障数指攻击者在攻击过程中实际注入的故障总数，包括无效故障数和有效故障数。无效故障数和有效故障数分别是注入故障后生成的正确密文和错误密文的数量。无效故障数和总故障数是衡量统计无效故障分析优劣的重要指标，若攻击者注入的无效故障数和总故障数越少，则在实际应用中越易实现。表 6 和表 2 分别列出了各区分器以不低于 99%的成功率恢复 8 bit 和 256 bit 原始密钥所需的故障数。可以看出，由于受无效故障影响后的中间状态值的可能取值的个数呈现相似规律，SEI 区分器对每一组中间状态值统计分析过程中采用相同的均匀概率分布，此时 SEI 区分器不能起到区分作用，因此在实际应用中不推荐使用 SEI 区分器。实验结果表明，使用 MLE、HW、PSC-MLE 和 HM-HW 区分器，分别仅需要 1 485、1 454、1 228 和 1 097 个无效故障，即可以不低于 99%的成功率恢复出 256 bit 原始密钥。在所有区分器中，本文提出的 HM-HW 双重区分器所需无效故障数和总故障数最少。

表 6 各区分器恢复 8 bit 原始密钥所需故障数

区分器	无效故障数/个	总故障数/个	成功率
SEI	∞	∞	—
MLE	46	464	≥99%
HW	45	455	≥99%
PSC-MLE	38	384	≥99%
HM-HW	34	343	≥99%

3.2　成功率

成功率指破译密钥的成功概率。各区分器在不同无效故障数下恢复 256 bit 原始密钥的成功率如图 6 所示。实验结果表明，MLE、HW、PSC-MLE 和 HM-HW 区分器均能达到 99%及以上的成功率，而 SEI 区分器的成功率最高不超过 5%，最低为 0。由图 6 可知，在恢复原始密钥全部 256 bit 情况下，HM-HW 双重区分器仅使用 1 097 个无效故障，就能最先达到 99%的成功率，表现更佳。

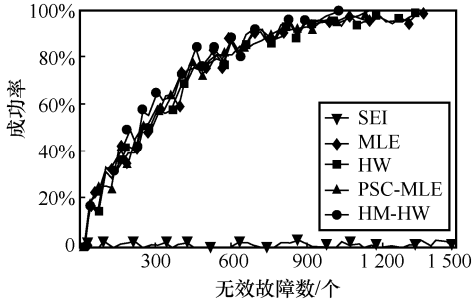


图 6　各区分器不同无效故障数下恢复 256 bit 原始密钥的成功率

3.3　耗时

耗时指利用区分器对中间状态值进行统计分析并筛选正确原始密钥时所消耗的时间。各区分器在不同无效故障数下恢复 256 bit 原始密钥的耗时如图 7 所示。表 7 给出了各区分器在不同无效故障数下恢复原始密钥的耗时。表 2、图 7 和表 7 表明，以 99%及以上成功率恢复 256 bit 原始密钥，MLE、HW、PSC-MLE、HM-HW 区分器耗时分别为 8.152 s、7.787 s、5.821 s 和 4.532 s。在所有区分器中，HM-HW 双重区分器破译密码的耗时最少。

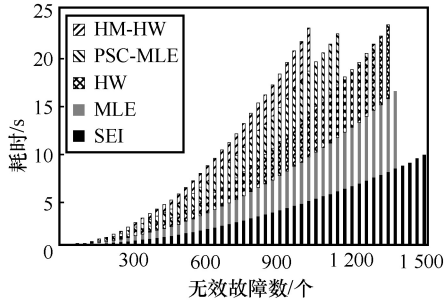


图 7　各区分器在不同无效故障数下恢复 256 bit 原始密钥的耗时

表 7　各区分器在不同无效故障数下恢复原始密钥的耗时

区分器	耗时/s				
	300 个	600 个	900 个	1 200 个	1 500 个
SEI	0.48	1.696	3.68	6.208	9.504
MLE	0.479	1.692	3.64	6.194	—
HW	0.472	1.669	3.59	6.191	—
PSC-MLE	0.47	1.662	3.575	6.084	—
HM-HW	0.465	1.643	3.534	—	—

3.4 复杂度

时间复杂度、数据复杂度和存储复杂度可用于衡量破译密码时所需的时间量、数据量和存储量，计算式分别为

$$\eta m + 2^u (l + w)$$

$$v(m + 2^u l)$$

$$v(l + 2^u)$$

其中，η 为加密变换次数，m 为总故障数，u 为枚举密钥的比特数，l 为注入无效故障数，w 为半字节所有可能取值的个数，v 为分组长度，$2^u l$ 为枚举密钥处理所有密文的次数，$2^u w$ 为区分器工作次数，vm 为遍历密文的数据量，$v2^u l$ 为遍历枚举密钥的数据量，vl 为存储的密文比特数，$v2^u$ 为存储候选密钥的比特数。表 8 分别给出了各区分器以 99%成功率恢复 256 bit 原始密钥的时间复杂度、数据复杂度和存储复杂度。在所有区分器中，HM-HW 双重区分器的时间复杂度、数据复杂度和存储复杂度均最低。

表 8 各区分器以 99%成功率恢复 256 bit 原始密钥的复杂度分析

区分器	时间复杂度	数据复杂度	存储复杂度
SEI	∞	∞	∞
MLE	$2^{19.98}$	$2^{18.23}$	$2^{13.95}$
HW	$2^{19.96}$	$2^{18.19}$	$2^{13.93}$
PSC-MLE	$2^{19.71}$	$2^{17.95}$	$2^{13.75}$
HM-HW	$2^{19.55}$	$2^{17.79}$	$2^{13.64}$

结合图 6、图 7、表 6~表 8 的实验结果来看，在以不低于 99%的成功率恢复原始密钥的情况下，与原有 SEI、MLE 和 HW 区分器相比，所提 PSC-MLE 和 HM-HW 区分器能够在更短时间内以更少故障数、更低复杂度破译 Saturnin-Short 密码，其中，HM-HW 区分器破译效果更好。

4 结束语

本文针对 Saturnin-Short 认证加密算法抵抗统计无效故障分析的安全性进行了研究，在基于统计分布和无效状态分析的基础上，通过结合故障注入前后中间状态值的变化，提出并讨论了概率对称卡方–极大似然估计和调和中项–汉明重量新型区分器。新型区分器不仅破译密码的成功率不低于 99%，而且能够在更短时间内以更少故障数和更低复杂度破译 Saturnin-Short 算法。研究表明，Saturnin-Short 密码算法易受统计无效故障分析的威胁。因此，在实际应用中使用该算法时，建议采取必要的有效措施抵御统计无效故障分析的攻击。下一步工作将结合 Saturnin-Short 算法内部的更深轮进行安全分析。

参考文献：

[1] CHEHAB M, MOURAD A. LP-SBA-XACML: lightweight semantics based scheme enabling intelligent behavior-aware privacy for IoT[J]. IEEE Transactions on Dependable and Secure Computing, 2022, 19(1): 161-175.

[2] ALRAWI O, LEVER C, ANTONAKAKIS M, et al. SoK: security evaluation of home-based IoT deployments[C]//Proceedings of 2019 IEEE Symposium on Security and Privacy. Piscataway: IEEE Press, 2019: 1362-1380.

[3] BANIK S, ISOBE T, LIU F K, et al. Orthros: a low-latency PRF[J]. IACR Transactions on Symmetric Cryptology, 2021(1): 37-77.

[4] BEIERLE C, LEANDER G, MORADI A, et al. CRAFT: lightweight tweakable block cipher with efficient protection against DFA attacks[J]. IACR Transactions on Symmetric Cryptology, 2019(1): 5-45.

[5] NAITO Y, MATSUI M, SUGAWARA T, et al. SAEB: a lightweight blockcipher-based AEAD mode of operation[J]. IACR Transactions on Cryptographic Hardware and Embedded Systems, 2018(2): 192-217.

[6] NAITO Y, SASAKI Y, SUGAWARA T. Lightweight authenticated encryption mode suitable for threshold implementation[C]//Annual International Conference on the Theory and Applications of Cryptographic Techniques. Berlin: Springer, 2020: 705-735.

[7] CANTEAUT A, DUVAL S, LEURENT G, et al. Saturnin: a suite of lightweight symmetric algorithms for post-quantum security[J]. IACR Transactions on Symmetric Cryptology, 2020(1): 160-207.

[8] BONEH D, DEMILLO R A, LIPTON R J. On the importance of checking cryptographic protocols for faults[C]//Advances in Cryptology - EUROCRYPT '97. Berlin: Springer, 1997: 37-51.

[9] FUHR T, JAULMES E, LOMNÉ V, et al. Fault attacks on AES with faulty ciphertexts only[C]//Proceedings of 2013 Workshop on Fault Diagnosis and Tolerance in Cryptography. Piscataway: IEEE Press, 2013: 108-118.

[10] CLAVIER C. Secret external encodings do not prevent transient fault analysis[C]//International Workshop on Cryptographic Hardware and Embedded Systems. Berlin: Springer, 2007: 181-194.

[11] BIHAM E, SHAMIR A. Differential fault analysis of secret key cryptosystems[C]//Advances in Cryptology – CRYPTO'97. Berlin: Springer, 1997: 513-525.

[12] COURTOIS N, WARE D A, JACKSON K. Fault-algebraic attacks on inner rounds of DES[C]//Strategies Telecom and Multimedia. Montreuil: Computer Science, 2010: 22-24.

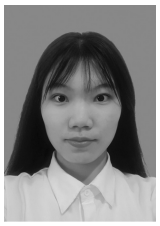
[13] DERBEZ P, FOUQUE P A, LERESTEUX D. Meet-in-the-middle and impossible differential fault analysis on AES[C]//International Workshop on Cryptographic Hardware and Embedded Systems. Berlin: Springer, 2011: 274-291.

[14] ZHANG F, LOU X X, ZHAO X J, et al. Persistent fault analysis on block ciphers[J]. IACR Transactions on Cryptographic Hardware and

Embedded Systems, 2018(3): 150-172.
[15] JANA A, PAUL G. Differential fault attack on PHOTON-beetle[C]// Proceedings of the 2022 Workshop on Attacks and Solutions in Hardware Security. New York: ACM Press, 2022: 25-34.
[16] ZHANG F, FENG T X, LI Z Q, et al. Free fault leakages for deep exploitation: algebraic persistent fault analysis on lightweight block ciphers[J]. IACR Transactions on Cryptographic Hardware and Embedded Systems, 2022(2): 289-311.
[17] 王永娟, 樊昊鹏, 代政一, 等. 侧信道攻击与防御技术研究进展[J]. 计算机学报, 2023, 46(1): 202-228.
WANG Y J, FAN H P, DAI Z Y, et al. Research progress of side channel attack and defense technology[J]. Chinese Journal of Computers, 2023, 46(1): 202-228.
[18] DOBRAUNIG C, EICHLSEDER M, KORAK T, et al. Statistical fault attacks on nonce-based authenticated encryption schemes[C]// International Conference on the Theory and Application of Cryptology and Information Security. Berlin: Springer, 2016: 369-395.
[19] 李玮, 汪梦林, 谷大武, 等. 轻量级密码算法TWINE的唯密文故障分析[J]. 通信学报, 2021, 42(3): 135-149.
LI W, WANG M L, GU D W, et al. Ciphertext-only fault analysis of the TWINE lightweight cryptogram algorithm[J]. Journal on Communications, 2021, 42(3): 135-149.
[20] SUGAWARA T, SHOJI N, SAKIYAMA K, et al. Exploiting bitflip detector for non-invasive probing and its application to ineffective fault analysis[C]//Proceedings of 2017 Workshop on Fault Diagnosis and Tolerance in Cryptograph. Piscataway: IEEE Press, 2017: 49-56.
[21] CLAVIER C, WURCKER A. Reverse engineering of a secret AES-like cipher by ineffective fault analysis[C]//Proceedings of 2013 Workshop on Fault Diagnosis and Tolerance in Cryptography. Piscataway: IEEE Press, 2013: 119-128.
[22] DOBRAUNIG C, EICHLSEDER M, KORAK T, et al. SIFA: exploiting ineffective fault inductions on symmetric cryptography[J]. IACR Transactions on Cryptographic Hardware and Embedded Systems, 2018(3): 547-572.
[23] BAGHERI N, SADEGHI S, RAVI P, et al. SIPFA: statistical ineffective persistent faults analysis on Feistel ciphers[J]. IACR Transactions on Cryptographic Hardware and Embedded Systems, 2022(3): 367-390.
[24] VAFAEI N, ZAREI S, BAGHERI N, et al. Statistical effective fault attacks: the other side of the coin[J]. IEEE Transactions on Information Forensics and Security, 2022, 17: 1855-1867.
[25] DOBRAUNIG C, EICHLSEDER M, GROSS H, et al. Statistical ineffective fault attacks on masked AES with fault countermeasures[C]//International Conference on the Theory and Application of Cryptology and Information Security. Berlin: Springer, 2018: 315-342.
[26] GRUBER M, PROBST M, TEMPELMEIER M. Statistical ineffective fault analysis of GIMLI[C]//Proceedings of 2020 IEEE International Symposium on Hardware Oriented Security and Trust. Piscataway: IEEE Press, 2020: 252-261.
[27] FISHER R A. On the mathematical foundations of theoretical statistics[J]. Philosophical Transactions of the Royal Society of London, 1922, 222(594): 309-368.
[28] REED I. A class of multiple-error-correcting codes and the decoding scheme[J]. Transactions of the IRE Professional Group on Information Theory, 1954, 4(4): 38-49.
[29] CHA S H. Comprehensive survey on distance/similarity measures between probability density functions[J]. International Journal of Mathematical Models & Methods in Applied Sciences, 2007, 1(4): 300-307.

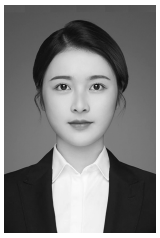
[作者简介]

李玮（1980– ），女，安徽寿县人，博士，东华大学教授、博士生导师，主要研究方向为对称密码的设计与分析。



刘春（2000– ），女，江西萍乡人，东华大学硕士生，主要研究方向为轻量级密码的安全性分析。

谷大武（1970– ），男，河南漯河人，博士，上海交通大学教授、博士生导师，主要研究方向为密码学和计算机安全。



孙文倩（2000– ），女，安徽铜陵人，东华大学硕士生，主要研究方向为对称密码的故障分析。

高建宁（1999– ），男，宁夏西吉人，东华大学硕士生，主要研究方向为对称密码的安全性分析。

秦梦洋（2000– ），男，河南许昌人，东华大学硕士生，主要研究方向为轻量级密码的故障分析。