SM4 密码算法的唯密文故障分析

李 玮[1),2),3)] 汪梦林[1)] 谷大武[2)] 温云华[1)] 李嘉耀[1)] 张雨希[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 200240)

[3)](上海交通大学上海市可扩展计算与系统重点实验室 上海 200240)

摘 要 SM4 算法是我国公布的首个商用密码算法,用于保护 WAPI 无线局域网标准中的数据传输,2012 年成为国家密码行业标准,2021 年作为 ISO/IEC 国际标准正式发布.在唯密文攻击中,攻击者除了所截获的密文,没有其它可利用的信息,因而获取的信息最少、攻击难度较大.目前,国内外未有公开发表的 SM4 算法抵抗唯密文攻击的结果.本文采用巴氏系数-汉明重量、詹森香农散度-汉明重量和詹森香农散度-汉明重量-极大似然估计等新型区分器对 SM4 算法实施唯密文故障分析.仿真实验表明,该方法最少仅需要 136 个随机故障,可以破译 SM4 算法的 128 比特主密钥,且在准确度、成功率、耗时和复杂度等方面攻击效果佳.该结果为无线局域网中密码算法的安全性分析和实现提供了重要参考.

关键词 故障分析;SM4;唯密文攻击;分组密码;密码分析

中图法分类号 TP309 **DOI 号** 10.11897/SP.J.1016.2022.01814

Ciphertext-Only Fault Analysis of the SM4 Cryptosystem

LI Wei[1),2),3)] WANG Meng-Lin[1)] GU Da-Wu[2)] WEN Yun-Hua[1)] LI Jia-Yao[1)] ZHANG Yu-Xi[1)]

[1)](*School of Computer Science and Technology, Donghua University, Shanghai* 201620)

[2)](*Department of Computer and Science and Engineering, Shanghai Jiao Tong University, Shanghai* 200240)

[3)](*Shanghai Key Laboratory of Scalable Computing and System, Shanghai Jiao Tong University, Shanghai* 200240)

Abstract The SM4 cryptosystem is the first commercial cryptographic algorithm announced by the government of China. It can be used to protect the data transmission in the wireless local area network (LAN) authentication and privacy infrastructure (WAPI). It has been the standard of the national cryptographic industry since 2012 and the international standard of ISO/IEC since 2021. It is the 32-round block cipher with the generalized Feistel network (GFN). It has the 128-bit block size and 128-bit master key. Since the publication of the SM4, there is plenty of cryptanalysis to evaluate its security, including zero-correlation linear analysis, linear analysis, algebraic side-channel analysis, algebraic analysis, integral analysis, impossible differential analysis, rectangle analysis, differential analysis, differential fault analysis and algebraic fault analysis, which have the basic assumptions of chosen plaintext attacks (CPA) or known plaintext attacks (KPA). The CPA needs to acquire the ciphertexts with the corresponding plaintexts, and the KPA demands a great number of known plaintexts and ciphertexts. Compared with the CPA and KPA, the ciphertext-only attack (COA) only requires the ciphertexts. It can do the security analysis of the cryptosystem without other information. Thus, the COA has the flexible applications and can

收稿日期:2021-05-20;在线发布日期:2022-01-21.本课题得到国家自然科学基金项目(61772129,61932014,62102077)、国家密码发展基金项目(MMJJ20180101)、上海市自然科学基金、上海市扬帆计划(21YF1401200)、上海市可扩展计算与系统重点实验室开放课题和中央高校基本科研业务费专项资金资助.**李 玮**,博士,教授,博士生导师,中国计算机学会(CCF)会员,主要研究领域为对称密码算法的设计与分析.E-mail:liwei.cs.cn@gmail.com.**汪梦林**,硕士研究生,主要研究方向为分组密码的故障分析.**谷大武**,博士,教授,博士生导师,主要研究领域为密码学与计算机安全.**温云华**,博士,讲师,主要研究方向为密码分析.**李嘉耀**(通信作者),博士研究生,主要研究方向为对称密码的安全性分析.E-mail:gaajiulei@gmail.com.**张雨希**,硕士研究生,主要研究方向为分组密码的安全性分析.

effectively verify the security of the cryptosystem. Up to now, there is no literature about the security of SM4 against the COA. On the designing structure of SM4, this paper proposes the novel ciphertext-only fault analysis, which adopts a random byte-oriented fault model and utilizes a series of distinguishers of Square Euclidean Imbalance (SEI), Hamming Weight (HW), Maximum Likelihood estimate (MLE), Goodness of Fit (GF), Goodness of Fit-Square Euclidean Imbalance (GF-SEI), Maximum Likelihood estimate-Square Euclidean Imbalance (MLE-SEI), Bhattacharyya Coefficient-Hamming Weight (BC-HW), Jensen Shannon distance-Hamming Weight (JSD-HW) and Jensen Shannon distance-Hamming Weight-Maximum Likelihood estimate (JSD-HW-MLE) for statistical analysis. This paper simulates the ciphertext-only fault analysis of SM4 cryptosystem on the computer software and uses indicators, such as accuracy, the number of faults, success probability, latency and complexities to measure the efficiency of different distinguishers. The accuracy is measured by the mean absolute error (MAE), which indicates the ability of different distinguishers to filter the subkey or the master key. When the success probability is highest, the number of faults can reflect the attacking ability. Fewer the number of faults, the stronger the attacking ability. The success probability refers to the probability of the ciphertext-only fault analysis in breaking the SM4 cryptosystem with different distinguishers. The distinguisher is perfect when the success probability is at least 99%. The latency means the time required to restore the master key of SM4 cryptosystem. The complex consists of time complexity, data complexity and memory complexity. The simulation experiments show that SM4 cannot resist the attack of the ciphertext-only fault analysis, the novel distinguishers of BC-HW, JSD-HW and JSD-HW-MLE only require 140, 140 and 136 faults to break SM4 cryptosystem, and the attack performs better in terms of accuracy, success probability, latency and complexities. The results provide an important reference for the security analysis and implementation of cryptographic algorithms.

Keywords　fault analysis; SM4; the ciphertext-only attack; block cipher; cryptanalysis

1　引　言

无线局域网具有移动性强、灵活性大、易扩展等优点，近年来在军事通信、金融证券、交通运输、医疗卫生等领域得到广泛应用[1]. 由于无线电波不要求建立物理的连接通道，且无线信号是发散传输，因此，无线局域网易遭到窃听、伪造和篡改等攻击，从而造成信息泄漏[2]. SM4 算法是我国提出的首个商用分组密码算法，用于保护 WAPI 无线局域网协议中的数据安全[3]. 该算法于 2012 年被国家密码管理局确立为国家密码行业标准，2021 年作为国际标准 ISO/IEC 18033-3:2010/AMD1:2021，由国际标准化组织 ISO/IEC 正式发布，标志着我国商用密码科技水平和国际标准化能力的不断提升. 大量研究表明，SM4 算法可以抵抗差分分析、线性分析、零相关线性分析、矩形分析、不可能差分分析、积分分析和代数分析等传统密码分析的攻击[4-10].

近年来，故障分析对密码算法的安全性提出了严峻挑战. 攻击者在密码算法加解密时诱导故障，利用故障泄露的信息恢复密钥. 故障注入的方式包括涡流磁场、激光照射、异常时钟或计算机软件模拟实现. 1997 年，Boneh 等学者通过诱导随机硬件故障攻击 RSA 算法，因此产生了故障分析[11]. 后来，根据攻击条件不同，故障分析被划分为差分故障分析、不可能差分故障分析、中间相遇故障分析、代数故障分析和唯密文故障分析等多个种类[12-16]. 目前，故障分析已成为检测分组密码实现安全性的一种重要密码分析技术.

根据攻击者获取信息的能力不同，密码分析包括唯密文攻击、已知明文攻击、选择明文攻击等. 其中，唯密文攻击表示攻击者除了所截获的密文，没有其他可利用的信息. 因此攻击者获取的信息最少、攻击难度最大.

唯密文故障分析是指攻击者在密码算法运行时诱导随机故障，从而得到错误密文，并利用区分器统计错误中间状态的分布规律，进而恢复子密钥及主密钥，完成密码破译. 现有区分器包括平方欧氏距

离、汉明重量、极大似然估计、拟合优度、拟合优度-平方欧氏距离和极大似然估计-平方欧氏距离等[16-19]. 目前,针对 SM4 算法的密码分析方法均集中于已知明文攻击和选择明文攻击,而未有公开发表的 SM4 算法抵抗唯密文攻击的结果.

本文采用唯密文故障分析破译 SM4 算法,构造了巴氏系数-汉明重量、詹森香农散度-汉明重量和詹森香农散度-汉明重量-极大似然估计等新型区分器. 与已有区分器相比,三种新型区分器在准确度、故障数、成功率、耗时和复杂度等方面的攻击效果更佳. 实验结果为无线局域网中密码算法的安全性分析和实现提供了重要参考.

2 相关工作

SM4 算法自发布以来,引起来了国内外研究学者的广泛关注. 2007 年,Liu 等学者使用积分分析评估其安全性,给出 13 轮的积分分析结果[4]. 同年,Lu 利用 12 轮不可能差分路径,将不可能差分分析扩展到 16 轮[5]. 2009 年,Erickson 等学者针对该算法进行 5 轮的代数分析[6]. 次年,Liu 等学者采用 18 轮线性特征,实现了 22 轮的多维线性分析[7]. 2013 年,Kong 等学者采用 16 轮矩形区分器,并利用哈希表降低了 18 轮矩形分析的时间复杂度[8]. 2015 年,马猛等学者证明 SM4 算法不能抵御 14 轮多维零相关线性分析[9]. 2018 年,赵艳敏等学者构造了 19 轮有效差分特征,使用两个辅助矩阵降低了 23 轮差分分析的存储复杂度[10].

与上述传统密码分析相比,以故障分析为代表的旁路分析破译速度快、威胁大. 2006 年,张蕾等学者首次提出了针对 SM4 算法的差分故障攻击,通过在最后一轮运算中诱导 8 个随机字节故障,获得该轮子密钥并推导出主密钥[20];2008 年,李玮等学者在密钥编排方案中诱导故障,利用 8 个错误密文即可破译该算法[21];2011 年,Li 等学者在第 28 轮运算中仅需注入 1 个随机字节故障,可以快速恢复出主密钥[22]. 2010 年,Hu 等学者使用代数故障分析检验该算法的安全性,仅使用 1 个故障即可破译主密钥[23]. 2013 年,刘会英等学者采集 SM4 算法加密时泄露的功耗信息,使用代数旁路分析恢复了主密钥[24]. 表 1 列出了 SM4 算法的已有密码分析,攻击假设均为已知明文攻击或选择明文攻击.

表 1 SM4 算法的密码分析比较

分析方法	攻击假设	攻击轮数	参考文献
零相关线性分析	已知明文攻击	14	[9]
线性分析	已知明文攻击	22	[7]
代数旁路分析	已知明文攻击	32	[24]
代数分析	选择明文攻击	5	[6]
积分分析	选择明文攻击	13	[4]
不可能差分分析	选择明文攻击	16	[5]
矩形分析	选择明文攻击	18	[8]
差分分析	选择明文攻击	23	[10]
差分故障分析	选择明文攻击	32	[20-22]
代数故障分析	选择明文攻击	32	[23]
唯密文故障分析	唯密文攻击	32	本文

不同于差分故障分析、代数故障分析等攻击方法,唯密文故障分析建立在唯密文攻击的基本假设上,此时攻击者仅需要获取随机错误密文,无需借助其它信息,即可完成密码破译. 2013 年,Fuhr 等学者使用平方欧氏距离、汉明重量和极大似然估计区分器对 AES 算法进行唯密文故障分析,各区分器至少需要 320、288 和 224 个随机字节故障才能够破译该算法[16]. 2016 年,Dobraunig 等学者结合硬件实现,利用唯密文故障分析破译了基于 AES 密码原语的认证加密算法[25]. 近年来,李玮等学者针对 LED、LBlock 和 MIBS 等算法提出了拟合优度、拟合优度-平方欧氏距离、极大似然估计-平方欧氏距离等区分器,进一步提高了唯密文故障分析效率[17-19].

本文探究了 SM4 算法的新型唯密文故障分析,构造出巴氏系数-汉明重量、詹森香农散度-汉明重量和詹森香农散度-汉明重量-极大似然估计等三个新型区分器. 表 2 比较了各区分器破译 AES 算法、

表 2 唯密文故障分析破译 AES、LED、LBlock、MIBS 和 SM4 算法主密钥的结果比较

区分器	英文简称	AES-128[16]		LED-128[17]		LBlock-80[18]		MIBS-80[19]		SM4-128	
		故障数	成功率/%	故障数	成功率/%	故障数	成功率/%	故障数	成功率/%	故障数	成功率/%
平方欧氏距离	SEI	320	≥99	560	≥99	408	≥99	324	≥99	280	≤55
汉明重量	HW	288	≥99	312	≥99	—	—	330	≥99	152	≥99
极大似然估计	MLE	224	≥99	320	≥99	232	≥99	210	≥99	164	≥99
拟合优度	GF	—	—	480	≥99	376	≥99	330	≥99	256	≤94
拟合优度-平方欧氏距离	GF-SEI	—	—	424	≥99	288	≥99	258	≥99	276	≤88
极大似然估计-平方欧氏距离	MLE-SEI	—	—	—	—	176	≥99	276	≥99	168	≥99
巴氏系数-汉明重量	BC-HW	—	—	—	—	—	—	—	—	140	≥99
詹森香农散度-汉明重量	JSD-HW	—	—	—	—	—	—	—	—	140	≥99
詹森香农散度-汉明重量-极大似然估计	JSD-HW-MLE	—	—	—	—	—	—	—	—	136	≥99

LED 算法、LBlock 算法、MIBS 算法和 SM4 算法的结果，本文提出的新型区分器不仅降低了故障数，而且成功率均达到 99%以上，提升了攻击效果.

3　SM4 算法介绍

3.1　术语说明

记 Z_2^e 为 e 比特的二进制向量集；

记 $X\in(Z_2^{32})^4$ 为明文输入，$Y\in(Z_2^{32})^4$ 为密文输出，τ 为非线性变换，L 和 L' 为线性变换，T 为 τ 和 L 的组合，且 $T=L(\tau(\cdot))$，S 表示 S 盒，$(X_i,X_{i+1},X_{i+2},X_{i+3})\in(Z_2^{32})^4$ 为加密算法第 $i+1$ 轮的输入且 $i\in[0,31]$；

记 $MK\in(Z_2^{32})^4$ 为 128 比特主密钥，$rk_i\in Z_2^{32}$ 为第 $i+1$ 轮子密钥，$(K_i,K_{i+1},K_{i+2},K_{i+3})\in(Z_2^{32})^4$ 为密钥编排方案第 $i+1$ 轮的输入，$CK_i\in(Z_2^8)^4$ 为固定参数；

记 f 为故障数，hw_n 为二进制字符串的汉明重量，Q_n 和 P_n 分别为错误中间状态的实际概率和理论概率，S_n 为 P_n 和 Q_n 的平均值，V_n 和 U_n 为错误中间状态的实际个数和理论个数，n 为错误中间状态值，$n\in[0,N]$ 且 $N=255$；

$\oplus$ 表示按位异或，$\&$ 表示按位与，Σ 表示连加，Π 表示连乘，$\sim$ 表示故障注入后的错误状态，0x 表示十六进制.

3.2　算法介绍

SM4 算法是具有广义 Fesitel 结构的分组密码，其结构如图 1 所示. 分组长度和密钥长度均为 128 比特，迭代 32 轮. 加解密结构一致，子密钥顺序相反. 加密算法和密钥编排方案如算法 1、2 所示.

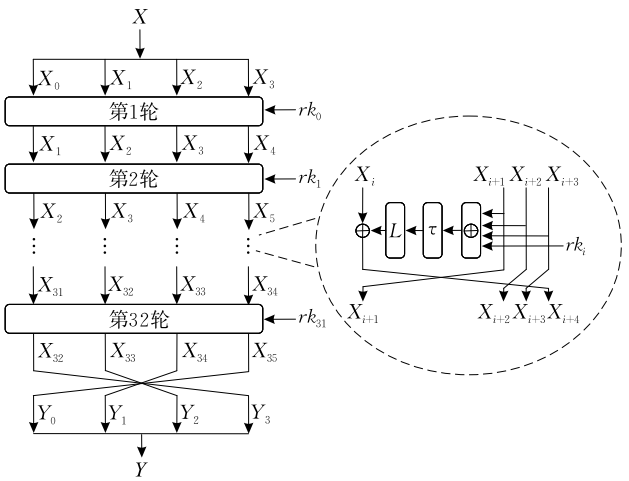

图 1　SM4 算法的结构

算法 1.　加密算法.

输入：$X,rk_0,rk_1,\cdots,rk_{31}$

输出：Y

1. $X=(X_0,X_1,X_2,X_3)$；
2. FOR $i=0$ TO 31 DO
3. 　$X_{i+4}=X_i\oplus L(\tau(X_{i+1}\oplus X_{i+2}\oplus X_{i+3}\oplus rk_i)$；
4. END FOR
5. $(Y_0,Y_1,Y_2,Y_3)=(X_{35},X_{34},X_{33},X_{32})$；
6. $Y=(Y_0,Y_1,Y_2,Y_3)$；
7. RETURN Y.

算法 2.　密钥编排方案.

输入：$MK,CK_0,CK_1,\cdots,CK_{31}$

输出：$rk_0,rk_1,\cdots,rk_{31}$

1. $MK=(MK_0,MK_1,MK_2,MK_3)$；
2. $K_0=MK_0\oplus$0xa3b1bac6；
3. $K_1=MK_1\oplus$0x56aa3350；
4. $K_2=MK_2\oplus$0x677d9197；
5. $K_3=MK_3\oplus$0xb27022dc；
6. FOR $i=0$ TO 31 DO
7. 　$K_{i+4}=K_i\oplus L'(\tau(K_{i+1}\oplus K_{i+2}\oplus K_{i+3}\oplus CK_i))$；
8. 　$rk_i=K_{i+4}$；
9. END FOR
10. RETURN $rk_0,rk_1,\cdots,rk_{31}$.

4　唯密文故障分析

4.1　攻击假设和故障模型

唯密文故障分析基于唯密文攻击假设，即攻击者监测加密通信的能力最弱，仅能获取随机密文. 结合 SM4 算法的构造，本文通过按位“与”注入单字节随机故障，使得中间状态各比特位的“0”和“1”出现的概率由原来的 0.5 和 0.5，分别变为 0.75 和 0.25. 通过统计可知，故障注入位置的中间状态所呈现的不均匀分布如图 2 所示. 以 $n=132$ 为例，因其二进制中包含 6 个“0”和 2 个“1”，所以 132 出现的概率为

$$0.75^6\times0.25^2\approx0.011.$$

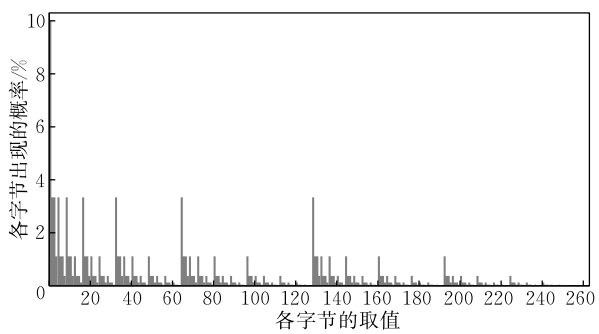

图 2　按位“与”后的字节值分布

4.2 攻击步骤

结合唯密文攻击假设和随机字节故障模型，针对 SM4 算法进行唯密文故障分析方法如下：

(1) 故障诱导. 攻击者利用相同主密钥加密随机明文时，以按位“与”的方式在某一轮注入随机字节故障，从而获得错误密文. 以图 3 最后四轮的故障扩散路径为例，此时攻击者将单字节故障注入到第 29 轮的输入 X_{31} 中.

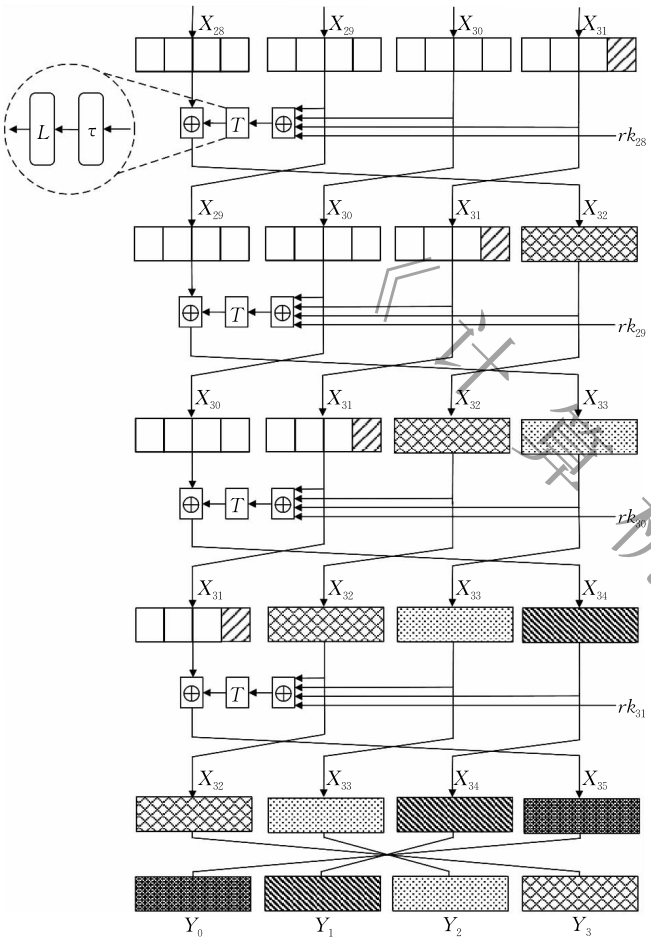

图 3 最后四轮运算的故障扩散路径

(2) 子密钥恢复. 攻击者利用错误密文和密钥候选值，得到错误中间状态值. 如图 3 所示，攻击者假设子密钥 rk_{31}，解密错误密文并获取故障注入后的中间状态值：

$$\tilde{X}_{31}=\tilde{X}_{35}\oplus L(\tau(\tilde{X}_{32}\oplus\tilde{X}_{33}\oplus\tilde{X}_{34}\oplus rk_{31}))$$
$$=\tilde{Y}_0\oplus L(\tau(\tilde{Y}_3\oplus\tilde{Y}_2\oplus\tilde{Y}_1\oplus rk_{31})).$$

攻击者再利用 4.3 节的区分器计算该中间状态的实际分布律，若与图 2 的理论分布律相符，则 rk_{31} 可能是正确子密钥. 在恢复 rk_{31} 的唯一值后，攻击者将故障导入第 28 轮获取随机错误密文，使用正确的 rk_{31} 解密错误密文，从而获得第 31 轮的错误输出(X_{31}，X_{32}，X_{33}，X_{34})，重复上述推导可以恢复子密钥 rk_{30}：

$$\tilde{X}_{30}=\tilde{X}_{34}\oplus L(\tau(\tilde{X}_{31}\oplus\tilde{X}_{32}\oplus\tilde{X}_{33}\oplus rk_{30})).$$

同理，攻击者可以恢复子密钥 rk_{29} 和 rk_{28}.

(3) 主密钥恢复. 攻击者根据最后四轮子密钥，可以依次推导所有的 rk_i 值，从而恢复主密钥 MK，求解过程如算法 3 所示.

算法 3. 主密钥的恢复.

输入：rk_{28}，rk_{29}，rk_{30}，rk_{31}，CK_0，CK_1，…，CK_{31}

输出：MK

1. $K_{32}=rk_{28}$；
2. $K_{33}=rk_{29}$；
3. $K_{34}=rk_{30}$；
4. $K_{35}=rk_{31}$；
5. FOR i=31 TO 0 DO
6. $\quad K_i=K_{i+4}\oplus L'(\tau(K_{i+1}\oplus K_{i+2}\oplus K_{i+3}\oplus CK_i))$；
7. END FOR
8. $MK_0=K_0\oplus$0xa3b1bac6；
9. $MK_1=K_1\oplus$0x56aa3350；
10. $MK_2=K_2\oplus$0x677d9197；
11. $MK_3=K_3\oplus$0xb27022dc；
12. $MK=(MK_0,MK_1,MK_2,MK_3)$；
13. RETURN MK.

4.3 区分器

4.3.1 现有区分器

2013 年，Fuhr 等学者首次将平方欧氏距离、汉明重量和极大似然估计区分器应用于 AES 算法的唯密文故障分析[16]；后来，李玮等学者针对 LED 等算法提出了拟合优度、拟合优度-平方欧氏距离和极大似然估计-平方欧氏距离区分器[17-19]. 具体如下：

(1) 平方欧氏距离区分器

数学家 Euclid 提出了欧氏距离，它的平方值称为平方欧氏距离，通常用来比较两点之间的距离. 在 AES 算法的密码分析中，平方欧氏距离区分器用于计算错误中间状态的实际分布与均匀分布之间的距离，它的最大值对应于正确子密钥，表达式为

$$\mathrm{SEI}=\sum_{n=0}^{N}\left(Q_n-\frac{1}{256}\right)^2.$$

(2) 汉明重量区分器

汉明重量由数学家 Reed 于 1954 年提出，用于计算二进制字符串与等长零字符串的汉明距离，等价于该字符串中“1”的个数[26]. 由于两个比特进行“与”操作后，结果为“0”的概率是 75%，结果为“1”的概率是 25%，所以故障被注入后，错误中间状态的各比特出现“0”的概率最大，此时汉明重量区分器的最小值对应于正确子密钥，表达式为

$$\mathrm{HW}=\frac{1}{f}\sum_{n=0}^{N}(hw_n\cdot V_n).$$

(3) 极大似然估计区分器

极大似然估计由数学家 Wilks 于 1938 年提出，用来估计概率模型的参数[27]. 在唯密文故障分析中，极大似然估计区分器要求攻击者已知故障模型对应的理论分布，即统计受故障影响的中间状态的理论概率乘积，其最大值对应于正确子密钥，表达式为

$$\mathrm{MLE}=\prod_{n=0}^{N}(P_n)^{V_n}.$$

(4) 拟合优度区分器

拟合优度由数学家 Pearson 于 1900 年提出，用于检验实验结果是否遵循理论分布[28]. 拟合优度区分器用于计算实际分布与理论分布的差距，其值越小表示两种分布的差异越小，可能对应正确子密钥，表达式为

$$\mathrm{GF}=\sum_{n=0}^{N}\frac{(U_n-V_n)^2}{U_n}.$$

(5) 拟合优度-平方欧氏距离区分器

该区分器是结合了两种单重区分器优点的双重区分器[17]. 攻击者通过拟合优度区分器挑选出区分器值较小的密钥候选值，表达式为

$$\mathrm{GF}=\sum_{n=0}^{N}\frac{(U_n-V_n)^2}{U_n}.$$

然后，攻击者使用平方欧氏距离区分器进一步筛选，保留区分器最大值，即对应正确子密钥，表达式为

$$\mathrm{SEI}=\sum_{n=0}^{N}\left(Q_n-\frac{1}{256}\right)^2.$$

(6) 极大似然估计-平方欧氏距离区分器

该区分器在单重区分器的基础上进行了改进[18-19]. 首先，攻击者采用极大似然估计区分器减小候选子密钥的猜测范围，即挑选出使区分器值较大的密钥候选值，表达式为

$$\mathrm{MLE}=\prod_{n=0}^{N}(P_n)^{V_n}.$$

再利用平方欧氏距离区分器找出使区分器值最大的密钥候选值，即对应正确子密钥，表达式为

$$\mathrm{SEI}=\sum_{n=0}^{N}\left(Q_n-\frac{1}{256}\right)^2.$$

4.3.2　新型区分器

本节提出了巴氏系数-汉明重量、詹森香农散度-汉明重量和詹森香农散度-汉明重量-极大似然估计区分器并应用于 SM4 算法的唯密文故障分析中.

(1) 巴氏系数-汉明重量区分器

1943 年，统计学家 Bhattacharyya 提出了巴氏系数，用于度量两个统计样本之间的重叠量，确定两个样本的接近度[29]. 本文使用巴氏系数区分器统计分析错误中间状态，当两个统计样本不相交时，值为 0；当样本完全重叠时，值为 1；样本部分重叠时，值介于 0 到 1 之间. 攻击者先找出使巴氏系数区分器值较大的密钥候选值，表达式为

$$\mathrm{BC}=\sum_{n=0}^{N}\sqrt{P_n\cdot Q_n}.$$

然后，攻击者挑选出使汉明重量区分器值最小的正确子密钥，表达式为

$$\mathrm{HW}=\frac{1}{f}\sum_{n=0}^{N}(hw_n\cdot V_n).$$

(2) 詹森香农散度-汉明重量区分器

1991 年，数学家 Lin 提出詹森香农散度，用于测量两个概率分布之间的相似度[30]. 本文利用詹森香农散度区分器统计分析错误中间状态，其值越小，表示理论分布与实际分布差异越小，即对应正确子密钥. 攻击者先使用詹森香农散度区分器筛选出使区分器值较小的密钥候选值，表达式为

$$\mathrm{JSD}=\sqrt{\frac{1}{2}\cdot\sum_{n=0}^{N}\left(P_n\cdot\log\frac{P_n}{S_n}+Q_n\cdot\log\frac{Q_n}{S_n}\right)}.$$

然后，攻击者筛选出使汉明重量区分器值最小的正确子密钥，表达式为

$$\mathrm{HW}=\frac{1}{f}\sum_{n=0}^{N}(hw_n\cdot V_n).$$

(3) 詹森香农散度-汉明重量-极大似然估计区分器

新型三重区分器先使用詹森香农散度区分器缩小密钥的搜索空间，表达式为

$$\mathrm{JSD}=\sqrt{\frac{1}{2}\cdot\sum_{n=0}^{N}\left(P_n\cdot\log\frac{P_n}{M_n}+Q_n\cdot\log\frac{Q_n}{M_n}\right)}.$$

接着，利用汉明重量区分器筛除使区分器值较大的候选子密钥，表达式为

$$\mathrm{HW}=\frac{1}{f}\sum_{n=0}^{N}(hw_n\cdot V_n).$$

最后，使用极大似然估计区分器挑选出唯一正确的子密钥，表达式为

$$\mathrm{MLE}=\prod_{n=0}^{N}(P_n)^{V_n}.$$

5　实验分析

本实验使用 Python 语言进行编程，通过计算机软件模拟 SM4 算法的故障导入，采用 GPU 并行计算完成破译过程，并采用不同的指标对实验结果

进行分析对比.

5.1 准确度

平均绝对值误差(Mean Absolute Error,*MAE*)可以度量区分器的准确度.故障数相同时,不同区分器挑选出的密钥候选值的 *MAE* 值越小,表示该区分器的准确度越高.*MAE* 表达式为

$$MAE=\frac{1}{\theta}\left(\sum_{\sigma=1}^{\theta}|h(\sigma)-1|\right),$$

其中,θ 表示重复实验的次数,第 σ 次实验获取的密钥候选值的理论个数和实际个数分别为 1 和 $h(\sigma)$.图 4 列出了各区分器破译 SM4 算法的一轮子密钥时,故障数与 *MAE* 值的关系.如图 4 所示,SEI 区分器的 *MAE* 值始终大于 0,即准确度较低.与现有区分器相比,新型区分器 BC-HW、JSD-HW 和 JSD-HW-MLE 的 *MAE* 曲线的下降趋势明显且快速逼近于 0,即准确度高.当故障数相同时,JSD-HW 区分器的 *MAE* 值更逼近于 0,如附录表 A1 所示.

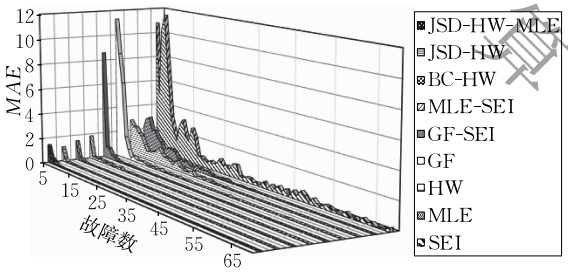

图 4 各区分器破译 SM4 算法的准确度

5.2 故障数

不同区分器破译密码算法所需要的故障数越少,代表攻击成本越低.图 5 表示当成功率最高时,各区分器破译 SM4 算法的 128 比特主密钥所需要的故障数.如图 5 所示,对于 SEI、GF 和 GF-SEI 等区分器,当故障数分别为 280、256 和 276 时,破译 SM4 算法成功率最高;对于 HW、MLE、MLE-SEI、BC-HW、JSD-HW 和 JSD-HW-MLE 等区分器,分别需要 152、164、168、140、140 和 136 个故障,就可以恢复 SM4 算法的 128 比特主密钥且成功率不低于 99%.其中,JSD-HW-MLE 区分器所需要的故障数最少,即攻击成本最低,如附录表 A2 所示.

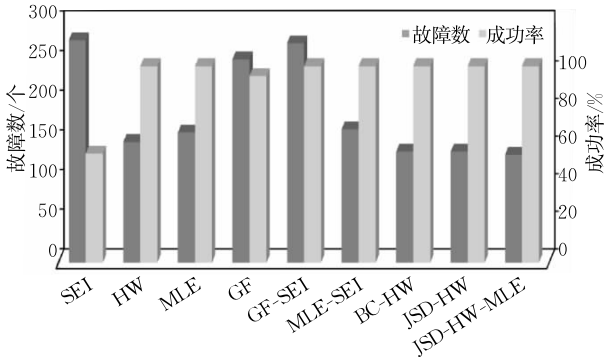

图 5 各区分器破译 SM4 算法的故障数

5.3 成功率

成功率是指不同区分器破译密码的概率.图 6 表示不同区分器破译 SM4 算法的一轮子密钥时,成功率随故障数变化的趋势.如图 6 所示,SEI、GF 和 GF-SEI 等区分器的成功率最高为 55%、94% 和 88%,HW、MLE、MLE-SEI、BC-HW、JSD-HW 和 JSD-HW-MLE 等区分器的成功率至少为 99%.与现有区分器相比,三种新型区分器 BC-HW、JSD-HW 和 JSD-HW-MLE 仅需要 140、140 和 136 个故障即能够以至少 99% 的成功率破译 SM4 算法,攻击效率更高.

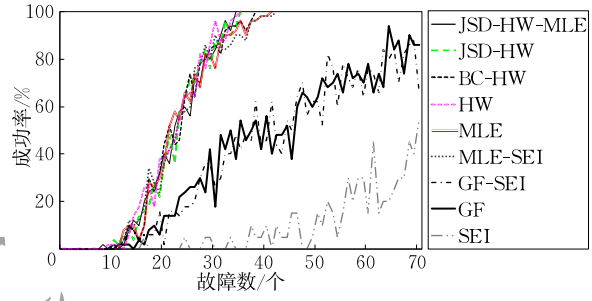

图 6 各区分器破译 SM4 算法的成功率

5.4 耗时

耗时是指破译 SM4 算法消耗的时间,包括故障诱导、密钥搜索和统计分析等过程.图 7 为各区分器恢复 SM4 算法一轮子密钥的时间堆积图.当成功率最高时,SEI、HW、MLE、GF、GF-SEI、MLE-SEI、BC-HW、JSD-HW 和 JSD-HW-MLE 等区分器,分别消耗 108.32、26.12、47.16、44.96、44.92、45.04、20.04、20.16 和 19.68 min,即可破译 SM4 算法.新型区分器 BC-HW、JSD-HW 和 JSD-HW-MLE 破译 SM4 算法的耗时较少且成功率不低于 99%,如附录表 A3 所示.

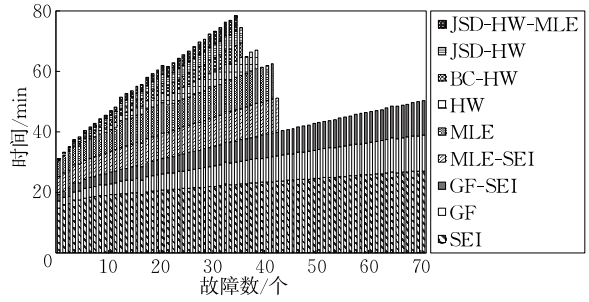

图 7 各区分器破译 SM4 算法的时间堆积

5.5　复杂度分析

时间、数据和存储复杂度能够估量破译 SM4 算法所需的时间量、数据量和存储量. 时间复杂度的计算公式为

$$f\cdot\psi\cdot\omega,$$

数据复杂度的计算公式为

$$f\cdot\psi,$$

存储复杂度的计算公式为

$$2^{7}\cdot f+2^{2}\cdot\eta+2^{\phi-\eta+2}.$$

其中，f 为故障数，ψ 表示候选密钥个数且 $\psi=2^{32}$，ω 为不同区分器的时间复杂度，η 表示候选密钥比特数，ϕ 表示子密钥比特数.

以三重区分器 JSD-HW-MLE 为例，由实验结果可知，当其成功率到达 99%时，恢复主密钥需要的故障数为 136，因此，此区分器的时间复杂度为

$$f\cdot\psi\cdot\omega=136\cdot2^{32}\cdot35\approx2^{44.22}.$$

根据其它实验结果，表 3 比较了不同区分器破译 SM4 算法的复杂度. 从表中结果可知，新型区分器 BC-HW、JSD-HW 和 JSD-HW-MLE 的时间、数据和存储复杂度均优于已有区分器，其中 JSD-HW-MLE 的复杂度最优.

表 3　各区分器破译 SM4 算法的复杂度比较

区分器	时间复杂度	数据复杂度	存储复杂度
SEI	$2^{48.48}$	$2^{40.13}$	$2^{17.13}$
HW	$2^{44.54}$	$2^{39.25}$	$2^{16.25}$
MLE	$2^{47.58}$	$2^{39.36}$	$2^{16.36}$
GF	$2^{49.17}$	$2^{40.00}$	$2^{17.00}$
GF-SEI	$2^{49.29}$	$2^{40.11}$	$2^{17.11}$
MLE-SEI	$2^{47.61}$	$2^{39.39}$	$2^{16.39}$
BC-HW	$2^{44.30}$	$2^{39.13}$	$2^{16.13}$
JSD-HW	$2^{44.30}$	$2^{39.13}$	$2^{16.13}$
JSD-HW-MLE	$2^{44.22}$	$2^{39.09}$	$2^{16.09}$

6　结束语

本文提出了针对 SM4 算法的唯密文故障分析，设计并实现了 BC-HW、JSD-HW 和 JSD-HW-MLE 等新型区分器，均能降低攻击代价，提高攻击效率. 研究表明，SM4 算法易受唯密文故障分析的影响. 因此，在无线局域网环境中，建议采取有效硬件措施保护相关加密设备，防止攻击者对其进行故障诱导.

参　考　文　献

[1] Murakami T, Ishihara K, Abeysekera H, et al. An actual stadium verification of WLAN using a distributed smart antenna system (D-SAS). IEICE Transactions on Communications, 2021, E104.B(1): 109-117

[2] Yan Z C, Yang C, You W, et al. Achieving secure and convenient WLAN sharing in personal. IET Information Security, 2020, 14(6): 733-744

[3] Office of State Commercial Cipher Administration. Block cipher for WLAN products — SMS4. http://www.oscca.gov.cn/UpFile/200621016423197990.pdf(in Chinese)
(国家商用密码管理办公室. 无线局域网产品使用的 SMS4 密码算法. http://www.oscca.gov.cn/UpFile/2006210164-23197990.pdf)

[4] Liu F, Ji W, Hu L, et al. Analysis of the SMS4 block cipher//Proceedings of the Australasian Conference on Information Security and Privacy. Townsville, Australia, 2007: 158-170

[5] Lu J Q. Attacking reduced-round versions of the SMS4 block cipher in the Chinese WAPI standard//Proceedings of the International Conference on Information and Communications Security. Zhengzhou, China, 2007: 306-318

[6] Erickson J, Ding J T, Christensen C. Algebraic cryptanalysis of SMS4: Gröbner basis attack and SAT attack compared//Proceedings of the International Conference on Information Security and Cryptology. Seoul, Korea, 2009: 73-86

[7] Liu Z Q, Gu D W, Zhang J. Multiple linear cryptanalysis of reduced-round SMS4 block cipher. Chinese Journal of Electronics, 2010, 19(3): 389-393

[8] Kong X L, Wang W, Xu Q L. Improved rectangle attack on SMS4 reduced to 18 rounds//Proceedings of the International Conference on Computational Intelligence and Security. Leshan, China, 2013: 575-578

[9] Ma Meng, Zhao Ya-Qun, Liu Qing-Cong, et al. Multidimensional zero-correlation linear cryptanalysis on SMS4 algorithm. Journal of Cryptologic Research, 2015, 2(5): 458-466(in Chinese)
(马猛，赵亚群，刘庆聪等. SMS4 算法的多维零相关线性分析. 密码学报，2015，2(5)：458-466)

[10] Zhao Yan-Min, Liu Yu, Wang Mei-Qin. Improved differential attack on 23-round SMS4. Journal of Software, 2018, 29(9): 2821-2828(in Chinese)
(赵艳敏，刘瑜，王美琴. 对 SMS4 密码算法改进的差分攻击. 软件学报，2018，29(9)：2821-2828)

[11] Boneh D, DeMillo R A, Lipton R J. On the importance of checking cryptographic protocols for faults//Proceedings of the International Conference on the Theory and Applications of Cryptographic Techniques. Konstanz, Germany, 1997: 37-51

[12] Dusart P, Letourneux G, Vivolo O. Differential fault analysis on AES//Proceedings of the International Conference on Applied Cryptography and Network Security. Kunming, China, 2003: 293-306

[13] Blömer J, Seifert J P. Fault based cryptanalysis of the advanced encryption standard (AES)//Proceedings of the International Conference on Financial Cryptography. Guadeloupe, France, 2003: 162-181

[14] Derbez P, Fouque P A, Leresteux D. Meet-in-the-middle and impossible differential fault analysis on AES//Proceedings of the International Workshop on Cryptographic Hardware and Embedded Systems. Nara, Japan, 2011: 274-291

[15] Courtois N T, Jackson K, Ware D. Fault-algebraic attacks on inner rounds of DES//Proceedings of the Strategies Telecom and Multimedia. Sophia Antipolls: French Riviera, 2010: 22-24

[16] Fuhr T, Jaulmes E, LomnéV, et al. Fault attacks on AES with faulty ciphertexts only//Proceedings of the Workshop on Fault Diagnosis and Tolerance in Cryptography. Los Alamitos, USA, 2013: 108-118

[17] Li W, Liao L F, Gu D W, et al. Ciphertext-only fault analysis on the LED lightweight cryptosystem in the Internet of Things. IEEE Transactions on Dependable and Secure Computing, 2019, 16(3): 454-461

[18] Li Wei, Wu Yi-Xin, Gu Da-Wu, et al. Ciphertext-only fault analysis of the LBlock lightweight cipher. Journal of Computer Research and Development, 2018, 55(10): 2174-2184(in Chinese)
(李玮, 吴益鑫, 谷大武等. LBlock 轻量级密码算法的唯密文故障分析. 计算机研究与发展, 2018, 55(10): 2174-2184)

[19] Li Wei, Cao Shan, Gu Da-Wu, et al. Ciphertext-only fault analysis of the MIBS lightweight cryptosystem in the Internet of Things. Journal of Computer Research and Development, 2019, 56(10): 2216-2228(in Chinese)
(李玮, 曹珊, 谷大武等. 物联网中 MIBS 轻量级密码的唯密文故障分析. 计算机研究与发展, 2019, 56(10): 2216-2228)

[20] Zhang Lei, Wu Wen-Ling. Differential fault analysis on SMS4. Chinese Journal of Computers, 2006, 29(9): 1596-1602(in Chinese)
(张蕾, 吴文玲. SMS4密码算法的差分故障攻击. 计算机学报, 2006, 29(9): 1596-1602)

[21] Li Wei, Gu Da-Wu. Differential fault analysis on the SMS4 cipher by inducing faults to the key schedule. Journal on Communications, 2008, 29(10): 135-142(in Chinese)
(李玮, 谷大武. 基于密钥编排故障的 SMS4 算法的差分故障分析. 通信学报, 2008, 29(10): 135-142)

[22] Li R L, Sun B, Li C, et al. Differential fault analysis on SMS4 using a single fault. Information Processing Letters, 2011, 111(4): 156-163

[23] Hu Z H, Liao X Y. SMS4 algorithm algebra fault attack//Proceedings of the International Symposium on Electronic Commerce and Security. Nanchang, China, 2010: 118-120

[24] Liu Hui-Ying, Zhao Xin-Jie, Wang Tao, et al. Research on hamming weight-based algebraic side-channel attacks on SMS4. Chinese Journal of Computers, 2013, 36(6): 1183-1193(in Chinese)
(刘会英, 赵新杰, 王韬等. 基于汉明重的 SMS4 密码代数旁路攻击研究. 计算机学报, 2013, 36(6): 1183-1193)

[25] Dobraunig C, Eichlseder M, Korak T, et al. Statistical fault attacks on nonce-based authenticated encryption schemes//Proceedings of the International Conference on the Theory and Application of Cryptology and Information Security. Hanoi, Vietnam, 2016: 369-395

[26] Reed I. A class of multiple-error-correcting codes and the decoding scheme. Transactions of the IRE Professional Group on Information Theory, 1954, 4(4): 38-49

[27] Wilks S S. The large-sample distribution of the likelihood ratio for testing composite hypotheses. The Annals of Mathematical Statistics, 1938, 9(1): 60-62

[28] Pearson K. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 1900, 50(302): 157-175

[29] Bhattacharyya A. On a measure of divergence between two statistical populations defined by their probability distributions. Bulletin Calcutta Mathematical Society, 1943, 35(4): 99-109

[30] Lin J H. Divergence measures based on the Shannon entropy. IEEE Transactions on Information Theory, 1991, 37(1): 145-151

附 录. 实验数据.

明文:随机生成

SM4 算法主密钥:01 23 45 67 89 ab cd ef fe dc ba 98 76 54 32 10

表 A1 各区分器恢复 SM4 算法一轮子密钥的 *MAE* 值

故障数	SEI	HW	MLE	GF	GF-SEI	MLE-SEI	BC-HW	JSD-HW	JSD-HW-MLE
0	4 294 967 295.00	4 294 967 295.00	4 294 967 295.00	4 294 967 295.00	4 294 967 295.00	4 294 967 295.00	4 294 967 295.00	4 294 967 295.00	4 294 967 295.00
1	1 677 215.00	1 677 215.00	1 677 215.00	1 677 215.00	1 677 215.00	1 677 215.00	1 677 215.00	1 677 215.00	1 677 215.00
2	1 677 216.40	65 183.23	65 578.24	1 112 705.36	65 620.36	65 355.14	65 520.38	65 696.34	65 463.16
3	65 898.26	263.24	262.14	49 081.08	6279.68	246.16	211.28	276.24	259.24
4	250.04	11.38	18.28	358.34	334.06	15.64	12.16	10.14	10.26
5	12.80	2.84	2.18	14.30	8.60	1.82	1.54	1.16	1.44
6	4.90	2.43	2.92	8.50	0.90	0.20	0.26	0.24	0.52
7	10.40	2.38	3.10	6.16	0.80	0.22	0.16	0.08	0.34
8	14.55	1.98	2.04	1.76	0.36	0.16	0.06	0.04	0.10
9	7.00	1.80	2.18	0.80	0.26	0.22	0.02	0.04	0.04

（续　表）

故障数	SEI	HW	MLE	GF	GF-SEI	MLE-SEI	BC-HW	JSD-HW	JSD-HW-MLE
10	3.25	1.16	1.56	0.14	0.08	0.16	0.00	0.04	0.02
11	2.20	0.68	1.36	0.40	0.06	0.10	0.00	0.02	0.00
12	1.45	0.78	0.76	0.20	0.00	0.18	0.00	0.02	0.00
13	2.80	0.82	1.02	0.18	0.08	0.12	0.00	0.00	0.00
14	3.30	0.88	1.08	0.24	0.04	0.00	0.00	0.00	0.00
15	2.05	1.32	1.02	0.16	0.06	0.14	0.02	0.00	0.00
16	1.70	0.92	0.92	0.10	0.02	0.14	0.00	0.00	0.00
17	2.85	0.70	0.40	0.08	0.02	0.02	0.00	0.00	0.00
18	2.10	0.68	0.52	0.14	0.00	0.00	0.00	0.00	0.00
19	0.75	0.52	0.38	0.06	0.02	0.06	0.00	0.00	0.00
20	0.80	0.32	0.52	0.16	0.04	0.04	0.00	0.00	0.00
21	0.60	0.76	0.24	0.00	0.00	0.02	0.00	0.00	0.00
22	0.45	0.40	0.20	0.08	0.00	0.04	0.00	0.00	0.00
23	0.70	0.28	0.44	0.06	0.04	0.02	0.00	0.00	0.00
24	0.55	0.18	0.28	0.04	0.00	0.00	0.00	0.00	0.00
25	0.70	0.24	0.22	0.04	0.02	0.00	0.00	0.00	0.00
26	1.10	0.38	0.30	0.08	0.00	0.00	0.00	0.00	0.00
27	0.65	0.02	0.22	0.00	0.00	0.00	0.00	0.00	0.00
28	0.75	0.16	0.10	0.06	0.00	0.00	0.00	0.00	0.00
29	0.55	0.18	0.06	0.04	0.00	0.02	0.00	0.00	0.00
30	1.05	0.04	0.10	0.00	0.00	0.00	0.00	0.00	0.00
31	1.05	0.18	0.10	0.00	0.02	0.00	0.00	0.00	0.00
32	0.30	0.08	0.14	0.02	0.02	0.00	0.00	0.00	0.00
33	0.25	0.04	0.04	0.02	0.00	0.00	0.00	0.00	0.00
34	0.40	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00
35	0.35	0.00	0.00	0.00	0.00	0.02	0.00	0.00	—
36	0.15	0.06	0.06	0.02	0.00	0.02	—	—	—
37	0.25	0.00	0.02	0.00	0.00	0.00	—	—	—
38	0.15	0.00	0.00	0.02	0.00	0.00	—	—	—
39	0.50	—	0.06	0.02	0.00	0.00	—	—	—
40	0.05	—	0.04	0.00	0.00	0.00	—	—	—
41	0.45	—	0.00	0.06	0.00	0.00	—	—	—
42	0.20	—	—	0.00	0.00	0.00	—	—	—
43	0.55	—	—	0.00	0.00	—	—	—	—
44	0.25	—	—	0.04	0.00	—	—	—	—
45	0.50	—	—	0.00	0.00	—	—	—	—
46	0.10	—	—	0.00	0.00	—	—	—	—
47	0.60	—	—	0.00	0.00	—	—	—	—
48	0.30	—	—	0.00	0.00	—	—	—	—
49	0.75	—	—	0.00	0.00	—	—	—	—
50	0.50	—	—	0.00	0.00	—	—	—	—
51	0.10	—	—	0.00	0.00	—	—	—	—
52	0.50	—	—	0.00	0.00	—	—	—	—
53	0.20	—	—	0.00	0.00	—	—	—	—
54	0.15	—	—	0.00	0.00	—	—	—	—
55	0.35	—	—	0.00	0.00	—	—	—	—
56	0.30	—	—	0.00	0.00	—	—	—	—
57	0.15	—	—	0.04	0.00	—	—	—	—
58	0.20	—	—	0.00	0.00	—	—	—	—
59	0.10	—	—	0.00	0.00	—	—	—	—
60	0.35	—	—	0.00	0.00	—	—	—	—
61	0.05	—	—	0.00	0.00	—	—	—	—
62	0.20	—	—	0.00	0.00	—	—	—	—
63	0.00	—	—	0.00	0.00	—	—	—	—
64	0.35	—	—	0.00	0.00	—	—	—	—
65	0.10	—	—	0.02	0.00	—	—	—	—
66	0.25	—	—	0.00	0.00	—	—	—	—
67	0.05	—	—	0.00	0.00	—	—	—	—
68	0.10	—	—	0.00	0.00	—	—	—	—
69	0.10	—	—	0.00	0.00	—	—	—	—
70	0.30	—	—	0.00	0.00	—	—	—	—

表 A2 各区分器恢复 SM4 算法一轮子密钥的成功率 （单位：%）

故障数	SEI	HW	MLE	GF	GF-SEI	MLE-SEI	BC-HW	JSD-HW	JSD-HW-MLE
0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	2
9	0	0	0	0	0	0	2	0	0
10	0	2	0	0	2	0	2	4	0
11	0	0	0	2	0	2	2	2	0
12	0	4	8	2	0	4	0	6	2
13	0	10	10	2	10	12	6	8	8
14	0	8	10	0	8	10	14	4	12
15	0	4	4	4	4	12	18	14	10
16	5	10	10	2	6	18	26	16	20
17	0	30	28	8	6	34	26	18	24
18	0	24	26	10	6	28	18	28	26
19	0	32	30	6	16	24	38	22	32
20	0	44	36	14	2	36	38	36	40
21	0	52	52	14	16	42	42	48	36
22	0	44	58	14	16	46	46	36	52
23	0	46	54	22	14	56	56	58	58
24	5	64	66	24	18	58	64	62	60
25	0	74	62	26	18	60	66	72	56
26	0	64	74	26	20	76	64	70	72
27	0	78	68	30	30	78	82	78	70
28	5	80	82	24	36	86	76	84	82
29	5	84	82	42	36	80	86	76	80
30	0	86	76	18	26	90	96	88	84
31	0	82	88	48	30	84	88	86	90
32	0	92	90	42	40	84	88	96	96
33	5	94	90	50	40	86	96	90	92
34	0	94	92	38	48	90	98	94	100
35	0	96	96	54	42	90	100	100	—
36	0	92	90	46	48	88	—	—	—
37	10	96	96	50	46	96	—	—	—
38	5	100	96	54	62	96	—	—	—
39	5	—	98	42	44	98	—	—	—
40	10	—	98	56	46	98	—	—	—
41	0	—	100	40	62	98	—	—	—
42	10	—	—	48	44	100	—	—	—
43	5	—	—	48	40	—	—	—	—
44	5	—	—	52	54	—	—	—	—
45	15	—	—	38	50	—	—	—	—
46	15	—	—	60	64	—	—	—	—
47	0	—	—	66	70	—	—	—	—
48	0	—	—	64	56	—	—	—	—
49	5	—	—	60	62	—	—	—	—
50	15	—	—	64	62	—	—	—	—
51	10	—	—	72	52	—	—	—	—
52	20	—	—	68	82	—	—	—	—
53	15	—	—	70	76	—	—	—	—
54	5	—	—	74	60	—	—	—	—
55	15	—	—	66	78	—	—	—	—
56	30	—	—	78	72	—	—	—	—
57	20	—	—	72	68	—	—	—	—
58	30	—	—	74	76	—	—	—	—
59	30	—	—	70	64	—	—	—	—
60	15	—	—	78	76	—	—	—	—
61	45	—	—	66	74	—	—	—	—
62	15	—	—	74	66	—	—	—	—
63	20	—	—	68	84	—	—	—	—
64	20	—	—	94	80	—	—	—	—
65	25	—	—	84	84	—	—	—	—
66	30	—	—	88	70	—	—	—	—
67	30	—	—	74	82	—	—	—	—
68	45	—	—	90	86	—	—	—	—
69	40	—	—	86	88	—	—	—	—
70	55	—	—	86	66	—	—	—	—

表 A3　各区分器恢复 SM4 算法一轮子密钥的时间　　（单位：min）

故障数	SEI	HW	MLE	GF	GF-SEI	MLE-SEI	BC-HW	JSD-HW	JSD-HW-MLE
0	14.79	0.18	5.48	2.23	2.22	5.78	0.16	0.16	0.15
1	15.90	0.33	5.65	2.29	2.31	5.87	0.32	0.29	0.29
2	16.08	0.48	5.83	2.63	2.55	6.03	0.50	0.53	0.52
3	17.24	0.63	6.00	2.81	2.93	5.96	0.58	0.58	0.57
4	17.37	0.76	6.14	2.93	2.89	6.10	0.70	0.70	0.70
5	18.13	0.93	6.32	3.12	3.05	6.23	0.86	0.85	0.84
6	18.43	1.07	6.44	3.23	3.14	6.34	0.98	0.98	0.97
7	18.57	1.20	6.58	3.37	3.26	6.47	1.10	1.09	1.09
8	18.94	1.35	6.73	3.52	3.39	6.59	1.23	1.22	1.22
9	18.90	1.54	6.91	3.70	3.56	6.76	1.40	1.39	1.39
10	19.05	1.74	7.11	3.85	3.69	6.88	1.62	1.57	1.55
11	19.19	1.89	7.26	3.99	3.81	7.01	1.78	1.67	1.67
12	19.37	2.07	8.90	4.17	3.99	7.19	1.95	1.84	1.85
13	19.51	2.21	9.04	4.32	4.11	7.32	2.09	1.97	1.97
14	19.65	2.37	8.91	4.45	4.24	7.44	2.22	2.08	2.10
15	19.84	2.55	9.00	4.63	4.41	7.62	2.40	2.27	2.27
16	19.21	2.69	9.53	4.78	4.54	7.75	2.39	2.39	2.39
17	20.15	2.88	9.73	4.95	4.70	7.92	2.58	2.56	2.55
18	20.28	3.02	9.86	5.09	4.82	8.04	2.68	2.68	2.68
19	20.42	3.19	10.07	5.22	4.94	8.16	2.83	2.83	2.82
20	20.60	3.37	10.22	5.40	5.11	8.33	3.00	3.00	2.99
21	20.72	3.55	8.70	5.53	5.23	8.45	3.15	3.14	3.13
22	20.85	3.69	8.84	5.65	5.34	8.58	3.27	3.27	3.28
23	21.02	3.90	9.04	5.82	5.50	8.78	3.47	3.46	3.47
24	21.14	4.03	9.19	5.96	5.62	8.88	3.59	3.58	3.59
25	21.27	4.20	9.31	6.08	5.74	8.98	3.82	3.71	3.72
26	21.44	4.37	9.54	6.25	5.90	9.15	3.89	3.89	3.88
27	21.57	4.53	9.67	6.37	6.02	9.27	4.17	4.01	4.01
28	21.69	4.67	9.76	6.50	6.14	9.40	4.16	4.14	4.13
29	21.87	4.87	10.00	6.67	6.30	9.56	4.33	4.33	4.33
30	21.98	4.97	10.28	6.80	6.42	9.68	4.40	4.42	4.43
31	22.12	5.07	10.38	6.92	6.54	9.80	4.58	4.59	4.61
32	22.62	5.20	10.51	7.21	6.78	10.05	4.65	4.67	4.66
33	22.41	5.32	10.69	7.22	6.82	10.09	4.75	4.76	4.77
34	22.59	5.49	10.79	7.38	6.98	10.26	4.91	4.93	4.92
35	22.71	5.89	10.96	7.50	7.09	10.37	5.01	5.04	—
36	22.81	5.77	11.01	7.64	7.20	10.49	—	—	—
37	22.97	6.53	11.23	7.79	7.36	10.66	—	—	—
38	23.09	6.53	11.38	7.90	7.47	10.77	—	—	—
39	23.22	—	11.49	8.02	7.58	10.88	—	—	—
40	23.38	—	11.53	8.19	7.73	11.03	—	—	—
41	23.48	—	11.79	8.29	7.84	11.11	—	—	—
42	23.60	—	—	8.40	7.95	11.26	—	—	—
43	23.76	—	—	8.56	8.10	—	—	—	—
44	23.87	—	—	8.68	8.21	—	—	—	—
45	23.99	—	—	8.79	8.32	—	—	—	—
46	24.14	—	—	8.95	8.48	—	—	—	—
47	24.26	—	—	9.09	8.59	—	—	—	—
48	24.38	—	—	9.18	8.71	—	—	—	—
49	24.52	—	—	9.33	8.85	—	—	—	—
50	24.63	—	—	9.44	8.94	—	—	—	—
51	24.80	—	—	9.59	9.09	—	—	—	—
52	24.89	—	—	9.70	9.21	—	—	—	—
53	25.01	—	—	9.81	9.32	—	—	—	—
54	25.16	—	—	9.96	9.47	—	—	—	—
55	25.26	—	—	10.07	9.56	—	—	—	—
56	25.36	—	—	10.18	9.66	—	—	—	—
57	25.51	—	—	10.33	9.82	—	—	—	—
58	25.62	—	—	10.43	9.92	—	—	—	—
59	25.73	—	—	10.54	10.03	—	—	—	—
60	25.89	—	—	10.69	10.18	—	—	—	—
61	25.99	—	—	10.80	10.27	—	—	—	—
62	26.08	—	—	10.89	10.36	—	—	—	—
63	26.26	—	—	11.06	10.54	—	—	—	—
64	26.64	—	—	11.24	10.70	—	—	—	—
65	26.47	—	—	11.27	10.74	—	—	—	—
66	26.59	—	—	11.39	10.87	—	—	—	—
67	26.70	—	—	11.51	10.98	—	—	—	—
68	26.87	—	—	11.66	11.13	—	—	—	—
69	26.97	—	—	11.77	11.23	—	—	—	—
70	27.08	—	—	11.88	11.34	—	—	—	—

LI Wei, Ph. D., professor, Ph. D. supervisor. Her main research interests include the design and analysis of symmetric ciphers.

WANG Meng-Lin, M. S. candidate. Her main research interest is fault analysis of block ciphers.

GU Da-Wu, Ph. D., professor, Ph. D. supervisor. His main research interests are cryptology and computer security.

WEN Yun-Hua, Ph. D., lecturer. Her main research interest is cryptanalysis.

LI Jia-Yao, Ph. D. candidate. His main research interest is security analysis of symmetric ciphers.

ZHANG Yu-Xi, M. S. candidate. Her main research interest is security analysis of block ciphers.

Background

Our work is supported by the National Natural Science Foundation of China under Grant Nos. 61772129, 61932014, 62102077, the National Cryptography Development Fund under Grant No. MMJJ20180101, the Natural Science Foundation of Shanghai, Shanghai Sailing Program under Grant No. 21YF1401200, the Opening Project of Shanghai Key Laboratory of Scalable Computing and Systems, and the Fundamental Research Funds for the Central Universities.

The SM4 cryptosystem is the first commercial block cipher announced by the government of China, which is used for security protocol of WAPI in the wireless local area network to protect the data security. The traditional cryptanalysis of the SM4 cryptosystem includes zero-correlation linear analysis, linear analysis, algebraic side-channel analysis, algebraic analysis, integral analysis, impossible differential analysis, rectangle analysis, differential analysis, differential fault analysis and algebraic fault analysis, etc.

Compared to traditional cryptanalysis, side-channel attacks have higher attack efficiency and make a severe challenge to the security of cryptographic algorithms. Fault analysis belongs to side-channel attacks, and was first proposed by Boneh et al. in 1997 to break RSA by injecting random hardware faults. Later, a variety of attacking methods are derived from fault analysis including differential fault analysis, impossible differential fault analysis, meet-in-the-middle fault analysis, algebraic fault analysis, and ciphertext-only fault analysis, etc. The existing side-channel attacks against SM4 cryptosystem include differential fault analysis, algebraic fault analysis and algebraic side-channel analysis.

At present, the basic assumptions of the existing security analysis of SM4 cryptosystem focus on the chosen-plaintext attack and the known-plaintext attack. The ciphertext-only fault analysis has the basic assumption of the ciphertext-only attack, which requires the weakest ability of attackers. In 2013, Ruhr et al. proposed the ciphertext-only fault analysis of AES in the software implementation. In 2016, Dobraunig et al. proved that attackers can inject faults into actual hardware and utilize ciphertext-only fault analysis to attack the AES-based authentication encryption algorithms. Later, Li et al. expanded the ciphertext-only fault analysis with new distinguishers on lightweight block ciphers including LED, LBlock and MIBS.