基于最大均值差异的能量侧信道泄露量化评估

洪 亮 翟元洁 王嘉熙 郑 健 胡 伟
(西北工业大学网络空间安全学院 西安 710072)

摘 要 能量侧信道分析是通过对密码设备运行时的能量消耗进行分析,推导出运行时的操作及操作涉及的敏感中间值.对密码设备进行能量泄露量化评估是分析密码设备信息泄露程度的重要手段,目前主流的评估方案主要关注于能量迹上单个样本点的泄露,并未充分考虑高阶攻击模型下的泄露评估问题,对于采用掩码防御措施的密码芯片来说,一旦发生泄露,通常表现为多变量联合泄露,因此采用传统的单样本点方法进行泄露评估会存在假阴性的问题.本文研究多点联合泄露评估问题,引入最大均值差异方法,提取能量迹的多变量联合特征,构建基于最大均值差异的能量泄露量化评估模型,提供了一种有效的能量侧信道泄露量化评估方法.通过实现无防御对策和有防御对策的 AES 算法,使用 DPA contest v2、ASCAD v1 和自采能量迹数据集进行实验,结果表明,基于最大均值差异的泄露量化评估方法能够有效降低单样本点检测方法出现假阴性的风险,HAC、MTD 和 Bartlett-F 检验的对照结果也进一步验证了该方法的有效性.

关键词 能量侧信道;信息泄露;量化评估;最大均值差异;掩码;AES
中图法分类号 TP309 **DOI 号** 10.11897/SP.J.1016.2024.01355

Quantitative Assessment of Power Side-Channel Leakage Based on MMD

HONG Liang ZHAI Yuan-Jie WANG Jia-Xi ZHENG Jian HU Wei
(*School of Cyberspace Security, Northwestern Polytechnical University, Xi'an* 710072)

Abstract Power side-channel analysis is aimed at extracting the internal operations and associated sensitive intermediate values of cryptographic devices from their power consumption patterns. Quantitatively assessing power leakage is essential for comprehending the extent of information leakage. However, current power leakage assessment approaches often focus primarily on a single leakage point, which may be inadequate for addressing the challenges posed by higher-order attack models. Additionally, cryptographic implementations utilizing masking countermeasures frequently exhibit leakage involving multiple variables, complicating detection using traditional single-point methods and leading to false negatives. To tackle this challenge, this study investigates multi-point joint leakage assessment by employing the Maximum Mean Discrepancy (MMD) method to extract the multivariate joint characteristics of power traces. The primary contribution of this paper is to assess the power-side channel leakage of AES by determining whether the distribution of power trajectory samples corresponding to two sets of keys is identical and quantitatively evaluating the degree of leakage in the encryption process of cryptographic devices. Firstly, the Maximum Mean Discrepancy, representing the largest difference in expectations over functions in the unit ball of a reproducing kernel Hilbert space (RKHS), is introduced as a side-channel

收稿日期:2023-09-11;在线发布日期:2024-04-08.本课题得到国家重点研发计划"处理器集成电路设计脆弱性检测与形式化验证"(2021YFB3100901)、国家自然科学基金"基于联合信息流分析的细粒度标准化硬件安全模型与度量研究"(62074131)、航天 772 所"同芯计划"项目资助.**洪 亮**,博士,副教授,中国计算机学会(CCF)专业会员,主要研究方向为无人系统安全、硬件安全.E-mail:hongliang@nwpu.edu.cn.**翟元洁**,硕士研究生,主要研究方向为密码学、侧信道分析与防护.**王嘉熙**,硕士,主要研究方向为密码学、侧信道分析与防护.**郑 健**,硕士,主要研究方向为硬件安全、侧信道分析.**胡 伟**(通信作者),博士,教授,中国计算机学会(CCF)高级会员,主要研究方向为硬件安全、形式化安全验证方法、密码学和可重构计算.E-mail:weihu@nwpu.edu.cn.

evaluation metric derived from transfer learning. By calculating the difference between the distributions of power trace samples, it assesses the disparity in distribution between two sets of power trace samples to evaluate the security of cryptographic devices. Secondly, building upon MMD, the Side-Channel Leakage Assessment (MMD-SCLA) scheme is proposed, which integrates multiple-point joint leakage characteristics of power traces to comprehensively quantify device security. This approach addresses the shortcomings of TVLA's single-variable quantification assessment and reduces the risk of false negatives in TVLA. To demonstrate the effectiveness of MMD-SCLA, publicly available datasets (DPA contest v2, ASCAD v1) and self-collected datasets are utilized for experimentation. To quantify the level of power leakage of the AES algorithm under various defense strategies, random delay and Gaussian noise defense mechanisms are implemented on the self-collected dataset. The TVLA, HAC, and Bartlett-F test methods are employed as comparison schemes. By integrating the MTD metrics, HAC metrics, TVLA t-values, and MMD-SCLA values across three AES power traces datasets without defense strategies, it is observed that among the implementations of AES in three cryptographic devices, SASEBO GII exhibits the smallest MTD value and the largest MMD-SCLA value. Consequently, compared to the power leakage assessment results of STM32F407 and SAKURA-X, it poses the highest risk of information leakage. The security ranking of cryptographic devices, from highest to lowest, is SAKURA-X, STM32F407, and SASEBO GII. In experiments comparing the suppression of power leakage under different defense strategies, the results of four evaluation metrics (HAC metrics, MMD-SCLA values, t-values, and Bartlett-F values) indicate that adopting the first-order masking defense strategy yields the highest security, followed by random delay methods, with Gaussian noise being the lowest. Additionally, experimental results also demonstrate that the MMD-based quantitative leakage assessment method eliminates the false negatives in the traditional TVLA methods. In summary, this work evaluates multivariate leakage analysis under different defense countermeasures, providing an effective tool for assessing side-channel power leakage. The results are also valuable for other symmetric cryptographic algorithms involving power leakage, such as SM4.

Keywords power side channels; information leakage; quantization assessment; maximum mean error; masks; AES

1 引 言

密码侧信道攻击是一种代表性的低开销密码分析手段.对密码设备进行能量泄露量化分析是为了评估密码设备信息泄露程度.当前研究主要集中在侧信道的攻击手段和相应的防御对策层面.能量分析攻击方法主要包括简单能量分析(Simple Power Analysis,SPA)[1]、差分能量分析(Differential Power Analysis,DPA)[2]、相关能量分析(Correlation Power Analysis,CPA)[3]、模板攻击(Template Attack,TA)[4]、碰撞攻击(Collision Attack)[5]和互信息分析(Mutual Information Analysis,MIA)[6]等.而防御对策主要包括掩码、随机延时、时钟抖动和高斯噪声等.侧信道量化评估是评估密码设备泄露程度和防御措施有效性的主要方式,是领域内一个极具价值和挑战性的前沿研究方向.

评估密码硬件设备应对侧信道攻击能力的一个量化指标是密钥解析能量迹数量(Measurements To Disclosure,MTD)[7],即恢复正确密钥所需的最小能量迹数量.该指标在应用中会依赖于所选择的评估方法,比如 DPA 和 CPA 在相同的加密设备上得出的 MTD 值会有所不同.

测试向量泄露评估(Test Vector Leakage Assessment,TVLA)方法解决了 MTD 依赖多个假设条件的问题[8],使用 t 值作为密码设备的量化指标,通过统计分析的方法,检验不同能量迹是否可以区分.但受限于一次只检测能量迹中的一个样本点,忽视了

各个样本点之间的关系，在高阶信息泄露场景中会存在假阴性问题. 在单变量 TVLA 基础上衍生出多变量或高阶 TVLA 评估技术，突破了基本的统计假设条件[9]. 为从原理上解决 TVLA 的局限性，皮尔森χ^2-test 通过检验两组变量独立性来判定待测能量迹之间是否有差异性[10]，在信噪比较高以及泄露信息分布在多个统计矩下的情况相比 TVLA 表现更优，并作为 TVLA 评估技术补充使用.

随着机器学习的引入，通过监督学习来训练模型作为能量迹数据集的区分器，当区分器能够以不可忽略的概率区分两组数据的差异性时，判断其存在能量泄露. 基于机器学习的评估方法存在数据集不平衡、模型过拟合、概率适应性等一系列问题，目前也是作为 TVLA 和χ^2-test 评估技术的补充使用[11].

随着能量侧信道攻击方法的发展，低阶攻击向高阶攻击演变，因而密码硬件设备的能量侧信道安全评估方法必须适应这一变化. 本文针对这一问题，贡献如下：

(1) 引入迁移学习中的最大均值差异方法(Maximum Mean Discrepancy，MMD)作为侧信道评估指标，通过计算能量迹样本分布的距离，获取两类数据的分布差异性来评估密码设备的安全性；

(2) 基于 MMD 提出了一种基于最大均值差异的能量侧信道量化分析方法(Side-Channel Leakage Assessment based on MMD，MMD-SCLA)，该方法在多变量泄露场景下，结合多点联合泄露特征，对密码设备的安全性进行综合性泄露评估；

(3) 以 TVLA、整体评估标准(Holistic Assessment Criterion，HAC)和基于 Bartlett 和多分类 F 检验(简称 Bartlett-F 检验)的侧信道泄露评估方法作为对照方案，将 CPA 攻击所得的 MTD 指标作为参考，对所提出的 MMD-SCLA 方法进行了仿真和物理实验验证，结果表明了基于最大均值差异的多变量泄露评估方法的有效性.

本文第 2 节介绍能量泄露量化评估技术和 AES 防御对策的研究工作；第 3 节详细阐述基于 MMD 的能量泄露量化评估方法；第 4 节开展 TVLA、HAC、MMD-SCLA 和 Bartlett-F 检验四种方法的对照实验，并进行综合分析；第 5 节总结本文所做的工作并展望未来研究方向.

2　相关工作

2.1　能量泄露量化评估

Goodwill 等人[12] 2011 年提出了基于 TVLA 的能量泄露评估方法，TVLA 方法主要利用威尔逊 t 检验来评估密码芯片抵抗侧信道攻击的能力，主要思想为检验两个能量迹数据集是否存在显著的统计学差异，评估方法分为非特定性检验和特定性检验. Schneider 等人[13]在 2016 年形式化了 Goodwill 的方法，采用 t 检验作为 TVLA 的假设检验核. 该方法成为标准的 TVLA 方法，并且该方法的评估结果准确预估了 DPA 成功恢复密钥的概率. Standaert[14] 在 2018 年指出，特定性 TVLA 检验方法在应用到整个能量迹过程中，仅关注单个样本点，会造成假阴性的评估结果，尤其是在被测设备具有掩码保护的情况下.

为了解决 TVLA 方法的局限性，Moradi 等人[10] 在 2018 年提出了基于χ^2-test 的评估技术. 该技术适用于噪声低或单个统计矩区分不明显的情况. 一旦高斯噪声特别显著时，会得出和 TVLA 方法一样的错误结论. Wegener 等人[15] 后续提出了基于深度学习的泄露评估技术(Deep Learning Leakage Assessment)，使用监督学习方法来训练分类器，当分类器能够以大于二项分布的概率将不同密钥对应的能量迹分类成功时，则代表存在能量泄露. Althoff 等人[16] 于 2019 年提出了整体侧信道泄露评估框架. 该框架提出了 HAC，通过判断两组能量迹的可交换性评估能量泄露程度，无需预设假设检验的前提条件，不同密码设备和高维数据场景下具有很强的鲁棒性，但该方法并未对有防护措施的密码设备进行量化分析. Bache 等人[17] 于 2023 年提出了一个多变量泄露评估框架，使用能量迹预处理和兴趣点识别方法，将置信区间的概念应用到多变量泄露检测中，从而降低了测量设置引入的不确定性，提高了泄露检测的效率.

在国内，中国信息安全测评中心陈佳哲等人[18] 在 2016 年运用 t 检验对 DES 和 3DES 算法进行了侧信道泄露评估. 中国科学院曹雨晨等人[19] 在 2020 年提出了多源融合信息泄漏检测方法，综合利用多种类型侧信道信息，有效提高了泄露检测效率. 武汉大学王娅茹等人[20] 在 2021 年提出了基于 Bartlett 和多分类 F 检验的侧信道泄露评估方法. Bartlett 检验主要用于均值差异小于方差差异的能量样本；多分类 F 检验用于均值差异大于方差差异的能量样本，将样本按明文汉明权重分类，有效降低了漏检和假阴性的概率. 信息工程大学郑震等人[21] 在 2023 年将双样本 KS 检验实施在不同样本量下，使用非特定 TVLA 得到的 t 值序列上以评估泄露，可减小能

量迹样本量较小情况下 TVLA 出现假阴性的概率.

综上,单变量 TVLA 泄露检测使用能量迹中的单个样本点进行分析,忽略了多个样本点之间的关系.多变量 TVLA 泄露检测在统计检验过程中作了更多假设,增加了假阳性的风险,可靠性不如单变量 TVLA.χ^2-test 泄露检测虽然解决了泄露出现在高阶矩如方差、偏态、峰度的问题,但对噪声敏感,每一次的评估数值极其不稳定,可能会得出被测设备安全性强的错误结论.基于深度学习的泄露评估技术,需要采集大量的能量迹数据进行针对性训练,并且对能量迹数据的质量有特定要求,量化评估花费时间长、计算代价高.

2.2 能量泄露的防御对策

目前国内外主流的防御技术主要包括掩码技术和隐藏技术两大类[22].掩码技术的主要思想是通过引入随机数的方式来掩盖敏感中间值和真实能量消耗之间的关系.隐藏技术是将算法执行过程随机化来消除密码设备能量消耗的数据依赖性.

在掩码技术研究当中,包括布尔掩码、算术掩码、总线掩码和动态局部重构等多种防御对策.布尔掩码和算术掩码对策方面,DPA contest 竞赛公布的 RSM2.0 方案中[23],采用改进型旋转 S 盒掩码隐藏 S 盒能量泄露产生位置.Jayakumar[24] 通过 Verilog 在FPGA上实现了掩码 S 盒,并测试了 S 盒代换使用布尔掩码时抵抗 DPA 攻击的能力.Benadjila 等人[25] 公开的 ASCAD 数据集对 S 盒采用一阶布尔掩码,并应用深度学习技术进行了侧信道分析.Asfand Hafeez 等人[26] 提出了一种改进的乘法掩码对策,解决了乘法掩码无法处理零值的问题.Vosoughi 等人[27] 提出一种新型总线掩码技术,在特定操作上(如 S 盒)增加 1 和 0 之间的翻转,降低能量特征和中间值之间的相关性.Kian 等人[28] 利用 FPGA 设备的动态部分重构特性混淆泄露信息,不仅克服了一阶 DPA 攻击,还提高了对二阶 DPA 攻击的抵抗能力.对于掩码对策,本文选用 ASCAD 中 ATmega8515 的一阶布尔掩码对策的可变密钥数据集进行泄露分析.

在隐藏技术研究当中,包括时间维度和振幅维度两类.时间维度最广泛使用的是随机插入伪操作和乱序操作,振幅维度则通过增加噪声、降低信号或者降低操作执行时信噪比,其中最为常见的是高斯噪声.时间维度方面,Avirneni 等人[29] 通过改变电压和频率消除敏感中间值和能量消耗之间的联系,RSM2.0 软件方案中通过引入乱序索引数组,对 AES 算法中的 16 个非线性 S 盒变换的执行顺序进行随机化保护.Couroussé 等人[30] 使用随机的寄存器分配、指令选择、噪声插入等方法实现加密过程中的代码多态功能,掩盖了能量特征和密钥之间的相关性.Chen 等人[31] 提出新的随机执行方法,通过随机和动态执行数据切断能量消耗特征和操作之间的联系.Chong 等人[32] 提出了一个具有相对时序的异步逻辑 AES 实现,并在 Sakura-X 和 Arty-A7 验证了其抗侧信道攻击的能力.振幅维度方面,Alipour 等人[33] 通过产生相关噪声来实现隐藏对策,该对策能够保护 AES 算法免遭 CPA 攻击和差异化深度学习分析.量化分析能量泄露时,如能量迹加入随机延时或高斯噪声,会混淆能量消耗特征与密码算法中间值之间的联系,泄露分析结果会出现假阴性风险.本文模拟隐藏对策时是在 STM32F407 开发板上采集能量迹,并仿真随机延时和高斯噪声两种对策进行泄露分析.

3 AES 能量泄露量化评估方法

3.1 问题描述

Prouff 等人[34] 提出的二阶差分能量分析攻击方法,已经探讨了如何使用多变量联合特征进行侧信道攻击.以图 1 为例,展示了一种非线性双变量联合泄露信息的情况.图 1 是在 FPGA 上采集的能量迹中兴趣点 i 和 j 的核密度能量值分布图,$L(t_i)$ 和 $L(t_j)$ 是 i 和 j 关于时间的能量分布函数,两处阴影对应着不同的密钥,双变量分布中的每一个分布经 TVLA 评估是安全的,但利用支持向量机(SVM)等机器学习算法能够对该分布进行分类.因而 TVLA 在面向多变量泄露场景下进行侧信道评估时,其结果存在着假阴性的可能,所以密码硬件设备的能量侧信道安全评估方法必须适应这一变化.

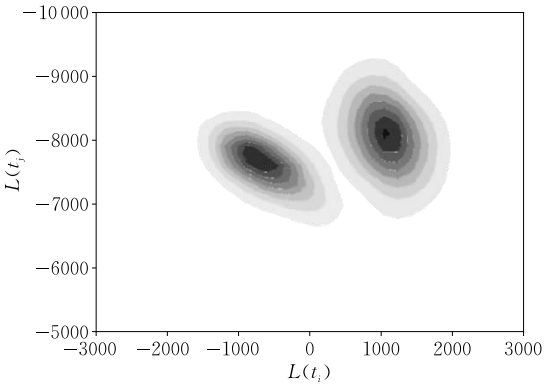

图 1 能量迹兴趣点 i 和 j 的核密度能量分布图

以 AES-128 算法为例，加密轮数为 10 轮，每轮都包含轮密钥加和字节代换操作，能量侧信道量化评估通常针对这两个操作引起的功耗变化展开分析. 通过收集待测设备使用不同密钥对明文数据加密产生的能量迹，分析两类能量迹是否具有不可区分性，以此判断是否存在数据依赖性，即信息泄露.

上述问题可归纳为“判断两个随机变量是否服从同一分布”. MMD 常用于迁移学习中，其核心思想是当两个随机变量的任意阶矩都相同，那么其服从的分布是一致的；当所服从的分布不相同时，使得这两个分布之间差距最大的矩，则定义为这两个分布的距离. MMD 通过计算再生核希尔波特空间(Reproducing Kernel Hilbert Space，RKHS)中两个样本分布的距离，即将分布映射到 RKHS 下，获取两类数据均值作差，求解其均值差异性[35]. MMD 值越小，则代表两类样本分布相似度越高，越大则代表两个样本分布相似度越低.

依据文献[36]关于 MMD 方法的定义，设$\mathcal{H}$是一个具有内积$\langle\cdot,\cdot\rangle_{\mathcal{H}}$和$\|\cdot\|_{\mathcal{H}}$相关的范数运算(定义$\mathcal{H}$上的范数运算为$\|h\|_{\mathcal{H}}=\langle h,h\rangle_{\mathcal{H}}^{1/2}$) RKHS.

假设 $X=[x_1,\cdots,x_n]$，$Y=[y_1,\cdots,y_m]$，X 和 Y 是独立同分布的随机变量样本，式(1)如下：

$$\mathrm{MMD}[\mathcal{F},p,q]=\sup_{f\in\mathcal{F}}(\mathbb{E}_p[\boldsymbol{f}(x)]-\mathbb{E}_q[\boldsymbol{f}(y)]) \tag{1}$$

其中，p 和 q 是两个样本的分布. 设随机变量 X 和 Y 是通过 p 和 q 采集的样本数据，样本个数分别为 n 和 m，MMD 经验估计可表示为式(2)：

$$\mathrm{MMD}[\mathcal{F},X,Y]=\sup_{f\in\mathcal{F}}\left(\frac{1}{n}\sum_{i=1}^{n}f(x_i)-\frac{1}{m}\sum_{i=1}^{m}f(y_i)\right) \tag{2}$$

其中$\mathbb{E}$为期望，$f(\cdot)$表示映射函数，$X\sim p$，$Y\sim q$，且$|X|=n$，$|Y|=m$，$\mathcal{F}$是函数域，函数域$\mathcal{F}$的选择需满足两个性质：

(1) 满足 $\mathrm{MMD}[\mathcal{F},p,q]:=0\Leftrightarrow p\underline{\underline{d}}q$，其中$\underline{\underline{d}}$代表分布相同；

(2) $\mathcal{F}$应为$\mathcal{H}$子集，使得 $\mathrm{MMD}[\mathcal{F},p,q]$有限.

要寻找合适$\mathcal{F}$满足上述两个性质，有定义 1.

定义 1.　设 k 为正定核函数，其定义域和值域为任意的实数集合；$\boldsymbol{x}_i$ 是样本数据集的一个样本，用向量表示，a_i则是实数集下的常数，对于 $\forall\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n\in\mathcal{H}$，$\forall a_1,\cdots,a_n\in\mathbb{R}$存在式(3)：

$$\sum_{i,j=1}^{n}a_ia_jk(\boldsymbol{x}_i,\boldsymbol{x}_j)\geqslant 0 \tag{3}$$

设$\mathcal{X}$为非空数据集，$\boldsymbol{\phi}(x)=k(x,\cdot)$，则$\mathcal{H}$存在唯一实值函数 Hilbert 空间 $H(k)$，满足：

(1) $\forall x\in\mathcal{H}$，$\boldsymbol{\phi}(x)\in H(k)$；

(2) $\forall f\in\mathcal{H}$，$\forall x\in\mathcal{X}$，$\langle\boldsymbol{f},\boldsymbol{\phi}(x)\rangle_{H(k)}=\boldsymbol{f}(x)$.

其中$H(k)$为核函数k的RKHS，$\boldsymbol{f}(x)$表示RKHS的向量 $\boldsymbol{f}$ 与空间向量$\boldsymbol{\phi}(x)$的点积，常见的特征核包括高斯核、指数核、拉普拉斯核等，当$\mathcal{F}$为再生核希尔伯特空间的单位球时，由于再生核希尔伯特空间具有良好的几何性质，空间中所有的点都在单位球上，在这个空间中，两个样本的距离可以用核函数值表示，MMD 可实现快速收敛[35-37]. 本文所选用的特征核为高斯核 $k(x,y)=\exp\left(-\dfrac{\|x-y\|^2}{2\sigma^2}\right)$，其中参数 σ 控制高斯核函数的作用范围，其值越大，高斯核函数的局部影响范围就越大. 由定义 1 和 RKHS 性质，式(1)推导可得到式(4)：

$$\begin{aligned}&\mathrm{MMD}[\mathcal{F},p,q]\\&=\sup_{\|f\|_{\mathcal{H}}\leqslant 1}(\mathbb{E}_p[\boldsymbol{f}(x)]-\mathbb{E}_q[\boldsymbol{f}(y)])\\&=\sup_{\|f\|_{\mathcal{H}}\leqslant 1}(\mathbb{E}_p[\langle\boldsymbol{f},\boldsymbol{\phi}(x)\rangle_{\mathcal{H}}]-\mathbb{E}_q[\langle\boldsymbol{f},\boldsymbol{\phi}(y)\rangle_{\mathcal{H}}])\\&=\sup_{\|f\|_{\mathcal{H}}\leqslant 1}[\langle\boldsymbol{f},\mu_p-\mu_q\rangle_{\mathcal{H}}]\\&=\|\mu_p-\mu_q\|_{\mathcal{H}}\end{aligned} \tag{4}$$

由于μ无法直接计算，采用均值代替计算，对式(4)进行平方简化运算得到式(5)：

$$\begin{aligned}\mathrm{MMD}^2[\mathcal{F},p,q]=&\mathbb{E}_{x,x'}[k(x,x')]-2\mathbb{E}_{x,y}[k(x,y)]+\\&\mathbb{E}_{y,y'}[k(y,y')]\end{aligned} \tag{5}$$

其中 x 和 x'表示服从 p 的随机变量，y 和 y'表示服从 q 的随机变量，$k(\cdot,\cdot)$代表高斯核函数对每一个样本进行映射. 综合式(1)、(2)和(5)可得式(6)：

$$\begin{aligned}&\mathrm{MMD}^2[\mathcal{F},X,Y]=\left\|\frac{1}{n}\sum_{i=1}^{n}f(x_i)-\frac{1}{m}\sum_{j=1}^{m}f(y_j)\right\|_{\mathcal{H}}^2\\&=\left\|\frac{1}{n^2}\sum_{i=1}^{n}\sum_{i'=1}^{n}k(x_i,x_i')-\frac{2}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}k(x_i,y_i)+\right.\\&\left.\frac{1}{m^2}\sum_{j=1}^{m}\sum_{j'=1}^{m}k(y_j,y_j')\right\|_{\mathcal{H}}^2\\&=tr(\boldsymbol{KL})\end{aligned} \tag{6}$$

$$\boldsymbol{K}=\begin{bmatrix}K_{s,s}&K_{s,t}\\K_{t,s}&K_{t,t}\end{bmatrix},\ \boldsymbol{L}=\begin{cases}\dfrac{1}{n^2}, & x_i\in S\\[2mm]\dfrac{1}{m^2}, & y_j\in T\\[2mm]\dfrac{-1}{nm}, & \text{其他}\end{cases}.$$

其中 $\boldsymbol{K}$ 代表高斯核函数计算得到的核矩阵，MMD 求解分布差异性问题转化成求解矩阵的迹.

3.2　能量泄露量化评估模型

通过上一节分析，将不同密钥对应的能量迹的分布差异性作为安全性量化指标，来反映密码设备是否存在数据依赖性. 因此，评估模型的零假设是不

同密钥对应的能量迹服从的分布不一致,通过计算两组样本分布的最大均值差异来判断接受或者拒绝零假设,MMD 值则代表了能量侧信道泄露的程度.

3.2.1 形式化定义

$\boldsymbol{L}_F$和$\boldsymbol{L}_R$是由两组不同密钥产生的能量迹构成的集合,其中$\boldsymbol{L}_F$是一组基于伪随机明文及固定密钥的能量迹集合,$\boldsymbol{L}_R$是一组基于伪随机明文及随机密钥的能量迹集合.二者皆为 $N\times M$ 的矩阵,N 为能量迹曲线个数,M 为每条能量迹曲线包含的样本点数量.将集合$\boldsymbol{L}_F$和$\boldsymbol{L}_R$的第 i 条能量迹和第 j 个样本点分别记为 F_{ij} 和 R_{ij},其中 $i\in[1,N]$,$j\in[1,M]$.

根据 Liu 等人在文献[38]中提出的基于 MMD 的假设检验方法,给出定义 2.

定义 2. 设 A 和 B 是随机变量 X 和 Y 服从的分布.$X=(x_1,x_2,\cdots)$,$Y=(y_1,y_2,\cdots)$,其中$X\sim A$,$Y\sim B$,X 和 Y 是能量迹所构成的样本集合.对 A 和 B 进行假设检验,零假设和备择假设分别表示为

$$H_0: A\neq B,\quad H_1: A=B.$$

当 $\mathrm{MMD}(X,Y)<\alpha$ 时,则拒绝零假设,$(x_1,x_2,x_3,\cdots)\underline{d}(y_1,y_2,y_3,\cdots)$,$A$ 和 B 分布相同,被测设备是安全的.反之,A 和 B 分布不同,被测设备不安全.其中 MMD 代表两个分布的差异值,α 代表拒绝零假设的阈值.根据文献[35]和[39]的建议,α 取为 0.

3.2.2 能量泄露量化评估过程

MMD-SCLA 进行侧信道能量泄露检测的过程如下,其流程如图 2 所示.

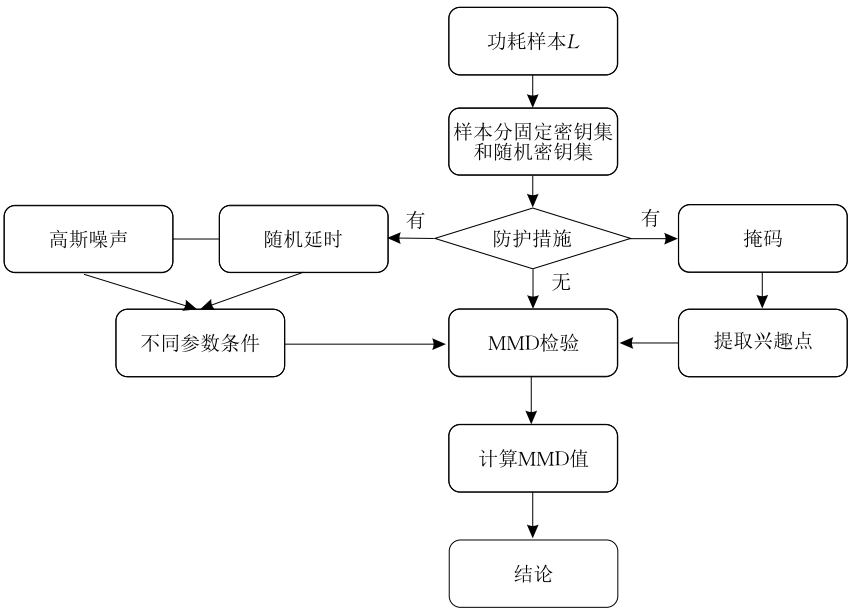

图 2 MMD-SCLA 量化评估过程

(1)首先将能量迹集合 $\boldsymbol{L}$ 分成固定密钥$\boldsymbol{L}_F$和随机密钥 $\boldsymbol{L}_R$两组,每组集合数量为 N,每条能量迹对应 M 个样本点.

(2)判断加密过程中实施的防护措施所属类型,处于掩码对策条件下,提取能量迹中的样本点进行分析;处于高斯噪声和随机延时条件下,设置参数对不同程度的噪声和延时条件进行能量泄露分析.

(3)数据预处理之后进行 MMD 检验,求解两个能量迹集合对应分布是否相同.

(4)当 MMD-SCLA 泄露评估值逐渐趋于 0 时,则被测设备不存在能量泄露.而随着泄露评估值不断增大时,被测设备能量泄露越严重.

为了衡量分布差异性,Mul 表示不同高斯核函数 σ 值间隔倍数,Num 表示高斯核数量,核矩阵的迹即评估结果记为V.MMD 泄露评估的过程如算法 1 所示.

算法 1. MMD-SCLA.

输入:$\boldsymbol{L}_F,\boldsymbol{L}_R,F_{ij},R_{ij},(i\in[1,N],j\in[1,M]),Mul,Num$

输出:V

//将不同能量迹数据组成的矩阵按列合并

1. $\boldsymbol{Total}\leftarrow\boldsymbol{L}_F\|\boldsymbol{L}_R$;

//矩阵 $\boldsymbol{Total}$ 整体复制 $2N$ 次

2. $\boldsymbol{Total}_0\leftarrow OverallExpand(\boldsymbol{Total})$;

//矩阵 $\boldsymbol{Total}$ 每一行复制 $2N$ 次

3. $\boldsymbol{Total}_1\leftarrow RowExpand(\boldsymbol{Total})$;

//求解 $\boldsymbol{Total}$ 任意行向量距离矩阵 $\boldsymbol{d}$

4. $\boldsymbol{d}\leftarrow EuclideanDistance(\boldsymbol{Total}_0,\boldsymbol{Total}_1)$;

//求不同核的 σ 值列表

5. FOR $i=0$ TO Num

6. $\quad SigmaList_i\leftarrow\sigma*(Mul**i)$

7. END FOR

//根据高斯核函数求核矩阵 $\boldsymbol{K}$

8. FOR $j=0$ TO $SigmaList$

9. $K_j \leftarrow GuassianKernel(\boldsymbol{d}, SigmaList_j)$;

10. END FOR

//根据核矩阵 $\boldsymbol{K}$ 的迹求解最终的 MMD 泄露值

11. $V \leftarrow (K_{SS}+K_{tt}-K_{st}-K_{ts})/N^2$;

其中，第 1 步将两类密钥对应的能量迹矩阵按列合并；2～3 步执行矩阵整体扩充和行扩充；第 4 步根据两个扩充矩阵，求解任意两个不同行向量之间的距离(i,j)，i 和 j 分别代表 ***Total*** 中第 i 行和第 j 行向量之间的欧式距离，得出高斯核指数的分子；5～7 步通过高斯核数量和间隔倍数求解不同高斯核的 σ 值；8～10 步根据高斯核函数求解核矩阵；11 步将核矩阵分成四部分，并根据核矩阵的迹得出最终的 MMD 泄露值 V.

3.3　MMD-SCLA 方法有效性分析.

TVLA 评估技术要求能量迹符合高斯分布且各个样本之间独立，Schneider 等人[40]证明了实际的能量迹不一定服从高斯分布，因此在进行泄露测试时应该尽可能的不依赖数据分布假设. 对于均值近似相同，方差不同的样本进行泄露评估，TVLA 评估技术可能存在假阴性问题. 比如仿真生成两组能量迹数据集，每组包含 1000 条能量迹，每条能量迹有 2000 个样本点. 设定两组样本均值分别为 0 和 0.005，方差分别为 10 和 80，样本均值近似相等，样本间的均值差异明显小于方差差异，对样本进行 TVLA 评估结果如图 3 所示.

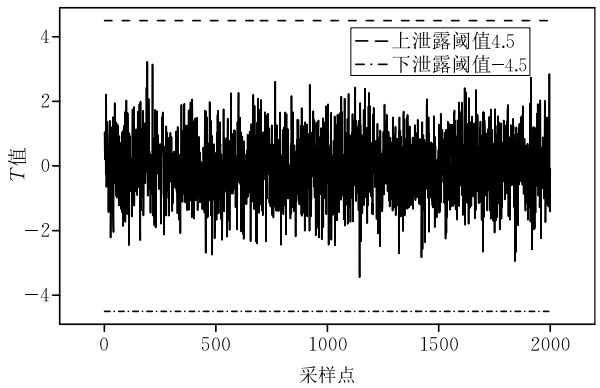

图 3　仿真能量迹 TVLA 评估结果

由结果可知 TVLA 泄露峰值为 3.23，小于阈值，并未发现能量泄露点. 而使用 MMD-SCLA 对两组能量迹进行泄露评估所得结果为 1.14，两组样本的分布存在差异，即样本存在能量泄露. 由上述实验可知，当两类样本的均值差异明显小于方差差异时，TVLA 存在假阴性风险，所得结论与文献[20]一致.

基于 MMD 的泄露评估方法，不依赖能量迹数据分布的假设条件，弥补了 TVLA 单变量量化评估的缺陷性，降低了 TVLA 泄露评估方法出现假阴性问题的风险.

相比 HAC，MMD-SCLA 将矩阵通过高斯核函数映射到再生核希尔伯特空间中进行距离计算，每一对样本的内积结果需存储在 $n\times n$ 大小的矩阵中，MMD-SCLA 的时间复杂度和空间复杂度均为 $O(n^2)$，在使用核函数优化的情况下，时间复杂度会变成 $O(n\times m)$，m 为核函数计算的次数. 而 HAC 使用欧式距离计算两类密钥所对应的能量迹矩阵，需遍历 n^2 个元素，将形成的距离矩阵进行排序并重新赋值，结果同样需要存储在 $n\times n$ 大小的矩阵中，其时间复杂度和空间复杂度均为 $O(n^2)$. 因此 MMD-SCLA 和 HAC 在计算开销和存储开销上是相同的.

该方法相比采用 KL 散度(Kullback-Leibler Divergence，KLD)量化不同密钥对应分布的差异性[41]时，克服了 KL 散度的不对称问题. 此外，采用 KL 散度的量化方案在当分布差异性较大时，KL 散度值会出现无穷大的结果.

4　实　验

本文以 AES-128 算法为评估对象，使用公开数据集和自采数据集进行实验. 为了量化不同防御对策下 AES 算法的能量泄露程度，在自采数据集上实现了随机延时和高斯噪声两种防御对策. 以 TVLA、HAC 和 Bartlett-F 检验方法对不同防御对策能量迹进行泄露评估作为对照方案，量化被测设备处在不同防御对策场景下能量泄露的脆弱性程度. 本文使用 AES 数据集汇总如表 1 所示.

表 1　数据集描述

数据集	芯片设备	防御措施
DPA contest v2	SASEBO GII	无
STM32F407	STM32F407	无
SAKURA-X	SAKURA-X	无
ASCAD v1	ATmega8515	一阶掩码
高斯噪声自采数据集	STM32F407	高斯噪声
随机延时自采数据集	STM32F407	随机延时

4.1　无防御对策的 AES 能量迹数据集设计.

无防御对策数据集由公开数据集 DPA contest v2、自采数据集 STM32F407 和自采数据集 SAKURA-X 构成. 其中 DPA contest v2 公开数据集是在 SASEBO GII 板上进行采集，密码算法未采用防御对策，通常

用于衡量侧信道攻击方法的有效性.

自采数据集分为 AES 软件实现数据集 STM32-F407 和硬件实现数据集 SAKURA-X. 构建过程为:首先在单片机和 FPGA 上实现无防护的 AES 加密算法,加密过程采用的密钥和明文参考 RAMBUS 报告发布的能量迹采集方案中的"固定密钥和随机密钥数据集"[42].

$\boldsymbol{L}_F$代表固定密钥和伪随机明文加密产生的数据集,固定密钥 K_{fix}=0x811E3731B0120A7842781E22B25CDDF9,初始明文 P_0=0xAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA,伪随机明文生成过程如式(7)所示.

$$P_{j+1}=\mathrm{AES}(K_{fix},P_j),\ 0\leq j<n \tag{7}$$

$\boldsymbol{L}_R$代表随机密钥和伪随机明文加密产生的数据集,初始随机密钥 K_0=0x53535353535353535353535353535353,初始伪随机明文 P_0=0xCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC,伪随机密钥生成过程如式(8)所示,伪随机明文生成过程遵循式(7).

$$K_{i+1}=\mathrm{AES}(K_{fix},K_i),\ 0\leq i<n \tag{8}$$

其中单片机采集能量迹的样本点个数为 1600,FPGA 上采集能量迹的样本点个数为 12000.

软件实现 AES 的能量迹采集设备使用 32 位的 STM32F407IGT6 MCU 单片机,开发板上 J5、J4、J3、J2 共 4 个 SMA 端子可用于电压能量迹采集. 在微控制单元(Microcontroller Unit,MCU)供电的干路上串联一个 1Ω 的采样电阻,采集电阻中经过的电流在电阻两端的压降,即可采集能量迹. J4 和 J3 分别是采样电阻靠近线性稳压源一端和 MCU 芯片一端,J2 是 J4 与 J3 差分放大信号输出端,J5 是 J3 端信号放大输出端. 示波器采用 PICO 3206D 便携数字示波器记录能量迹,在采集 AES 无防御对策的能量迹过程中,整个实验平台连接如图 4 所示. STM32F407 开发板和示波器使用 USB3.0 接口连接 PC;无源高阻探头一端连接开发板上的 PD7 引脚,另一端连接示波器通道 A 采集触发信号. SMA 连接线一端连接 J2,另一端连接示波器通道 B 采集功耗信号. 在使用示波器采集能量迹的过程当中,上升沿触发,触发阈值为 1000 mv,触发延迟设置为 0,同时设置一次采集的波形数量为 100,触发前记录点个数 0,触发后记录点个数为 1600,采样率为 500 MHz,通道 A 的垂直范围为±5 V. 通道 B 的垂直范围为±2 V. 实验平台连接示意图如图 4 所示.

在实际能量迹采集过程当中,即使没有人为添加噪声,也会存在小范围窗口抖动和高斯噪声,因此无噪声并且无防御对策的能量迹条件,一般代指没有人工添加噪声和防御对策,环境本身噪声处于低水平,不足以影响评估者对于脆弱性的分析.

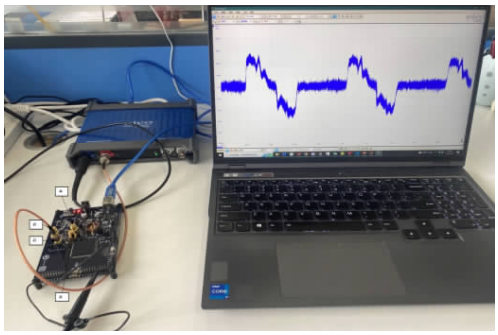

图 4 实验平台连接示意图

4.2 有防御对策的 AES 能量迹数据集设计

实验模拟以下三种场景:高斯噪声、随机延时和掩码. 掩码能量迹由公开数据集 ASCAD v1 构成,高斯噪声和随机延时能量迹在自采数据集 STM32F407 上仿真实现.

4.2.1 随机延时和高斯噪声实现

模拟随机延时的步骤如下. 对于能量迹集合中每条能量迹 L_i的每个样本点 P_i.

(1) 从伯努利分布 $Bernoulli(p)$随机抽样.

(2) 如果采样值为 1,从均匀分布 $Uniform(0,a-b)$抽取一个数 m.

(3) 根据 m 从均匀分布 $Uniform(m,m+b)$中再抽取一个随机数 n.

(4) 在能量迹中插入以下三个样本点$[P_i,P_i+R,P_i+1]$共 n 次,R 是一个固定的抖动幅度,该抖动仿真了加入随机延时对策之后的功耗值.

本文通过对正态分布 $N(0,\sigma^2)$进行随机抽样模拟高斯噪声并将其添加到能量迹上,以此来实现高斯噪声防御对策. 随机延时和高斯噪声能量迹与无防御措施的原始能量迹对比如图 5 和图 6 所示.

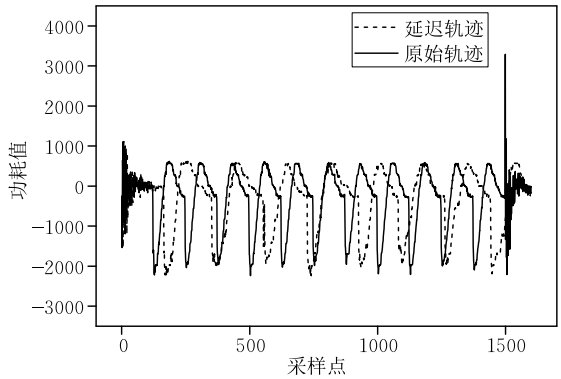

图 5 随机延时能量迹对比

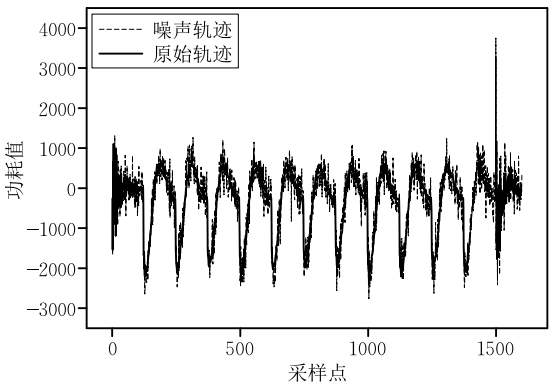

图 6　高斯噪声能量迹对比

4.2.2　ASCAD 数据集的掩码技术

本文所使用的公开数据集 ASCAD v1 采用了一阶布尔掩码防御对策，来消除能量迹和敏感中间值之间的相关性.

在公开数据集 ASCAD 中，将数据集分为“固定密钥和随机密钥数据集”. 固定密钥 K_{fix} = 0x00112233445566778899aabbccddeeff，随机密钥 K_{random} 取值范围在 0x00000000000000000000000000000000 ～ 0xfffffa6ca116534e67e9c9cc8ca53676 之间，能量迹数量均为 100 000.

该数据集从 Atmega8515 芯片设备上采集，在进行字节代换时选取其中的第三状态字节与掩码进行异或操作，实现带掩码的字节代换. 通过提取不同密钥对应的能量迹形成两组集合，每个集合的一条能量迹有 250 000 个样本点，图 7 是其中一条能量迹的展示.

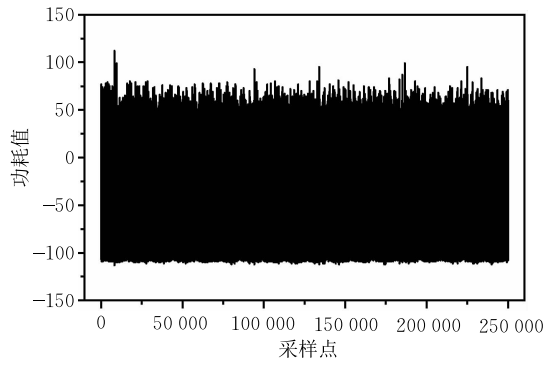

图 7　ASCAD v1 能量迹

4.3　MMD-SCLA 量化评估实验

4.3.1　无防御对策数据集 MMD-SCLA 评估

对于公开数据集 DPA contest v2 的 32 个密钥，能量消耗模型采用汉明重量模型，32 个密钥的汉明重量分布如图 8 所示. 汉明重量最小的密钥为 0x00000000000000003243f6a8885a308d3(即序号为 0 的密钥)，汉明重量最大的密钥为 0xd1310ba698dfb5ac2ffd72dbd01adfb7(即序号为 5 的密钥). 将密钥相同的能量迹分为一组，共分为 32 组，每组有 20 000 条能量迹，每一组随机抽取 1000 条能量迹，使用 MMD-SCLA 分别对密钥 0 和密钥 5 与其他 31 组密钥分别组成的能量迹集合进行泄露评估实验，结果如图 9 和图 10 所示. 由图可知，密钥 0、密钥 5 与密钥 30 所对应的能量迹集合，泄露程度最高，MMD 泄露评估值分别为 6.96 和 7.13.

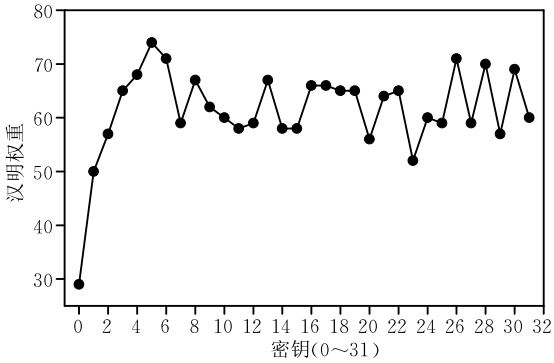

图 8　32 个密钥的汉明重量分布

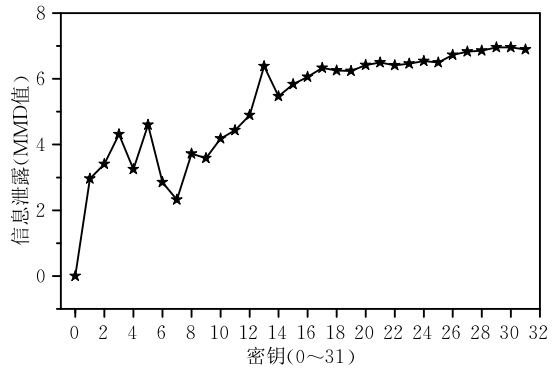

图 9　密钥 0 能量迹泄露评估

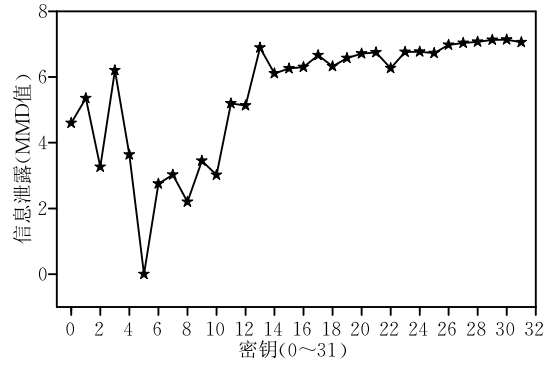

图 10　密钥 5 能量迹泄露评估

使用 MMD-SCLA 对自采数据集 STM32F407 和 SAKURA-X 进行泄露评估. 无防御对策下软件实现的 AES 算法，其 MMD 泄露评估值为 5.21，硬件实现的 AES 算法的泄露评估值为 2.26.

4.3.2 有防御对策数据集 MMD-SCLA 评估

使用随机延时和高斯噪声两种自采数据集和公开掩码数据集 ASCAD,在上述基准下进行有防御对策条件下的泄露评估.

(1)随机延时

随机延时通过时间随机性使密码设备加密过程中采集的能量迹不对齐,增大能量分析攻击的利用难度.延时参数越大,能量迹随机性越强.因此,通过实验对延时参数设置固定间隔,分析不同随机延时条件对于能量泄露程度的影响.实验设置延时参数区间为[0.01,0.20],步长为 0.01.当延时参数分别为 0.01、0.05、0.10、0.20 时的延时效果如图 11 所示.针对不同延时水平下的能量迹,使用 MMD-SCLA 进行泄露评估,实验结果如图 12 所示.根据实验结果分析,随机延时参数设置为 0.02 时,泄露峰值为 0.63.由于随机延时导致两组能量迹数据集的分布差异性小,因此采用随机延时防御对策条件下的泄露值较未采用的更小.

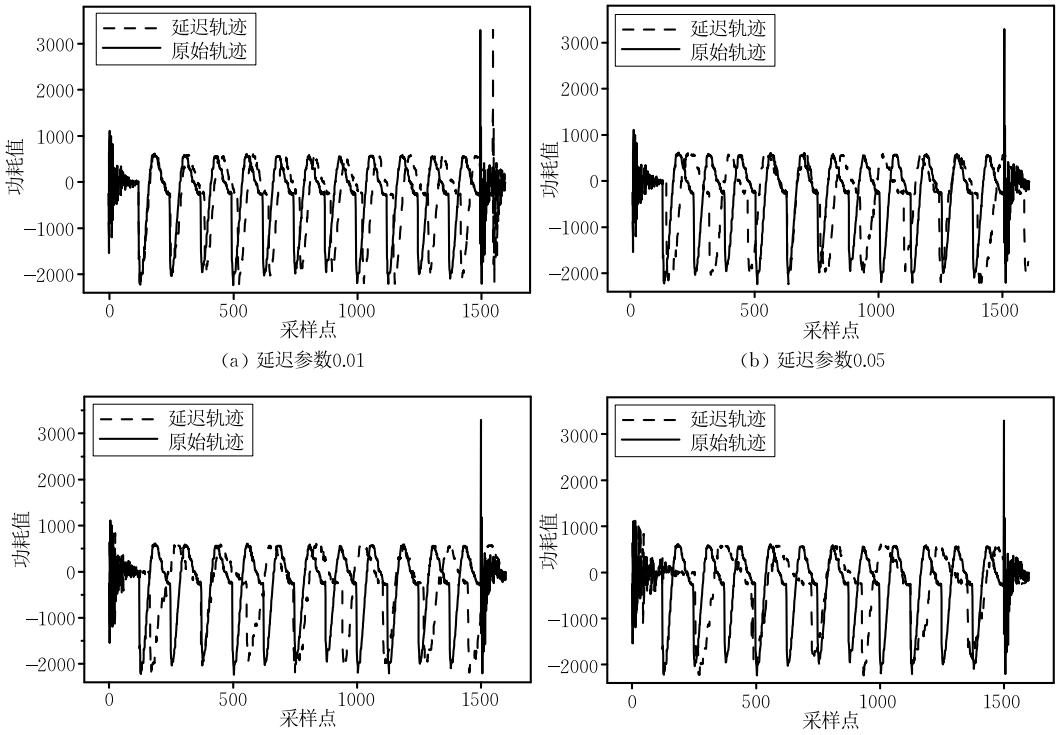

图 11 随机延时对策下的能量迹

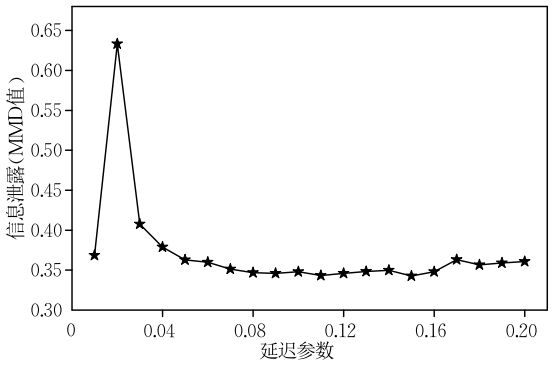

图 12 随机延时对策下 MMD-SCLA 泄露评估

(2)高斯噪声

针对评估高斯噪声数据集,首先设置高斯噪声参数 σ 区间为[20,400],步长为 20,当参数 σ 分别为 100、200、300、400 时,能量迹如图 13 所示.针对不同高斯噪声参数下的能量迹,使用 MMD-SCLA 进行泄露评估实验,实验结果如图 14 所示.参数为 20 和 400 时,对应泄露值为 4.33 和 0.15,泄露峰值坐标为(40,5.35).在模拟高斯噪声过程中,根据文献[43]和[44]对能量迹不做任何滤波处理,目的是验证噪声对 MMD-SCLA 评估的影响.由于是对正态分布 $N(0,\sigma^2)$ 进行随机抽样模拟添加高斯噪声,所以随着参数 σ 的增大,MMD-SCLA 泄露评估值呈非线性下降趋势.当高斯噪声大到一定程度时,原始能量迹形状被明显扭曲,能量迹分布最终呈现相同的正态分布,两组能量迹的 MMD-SCLA 泄露评估值趋近于 0.

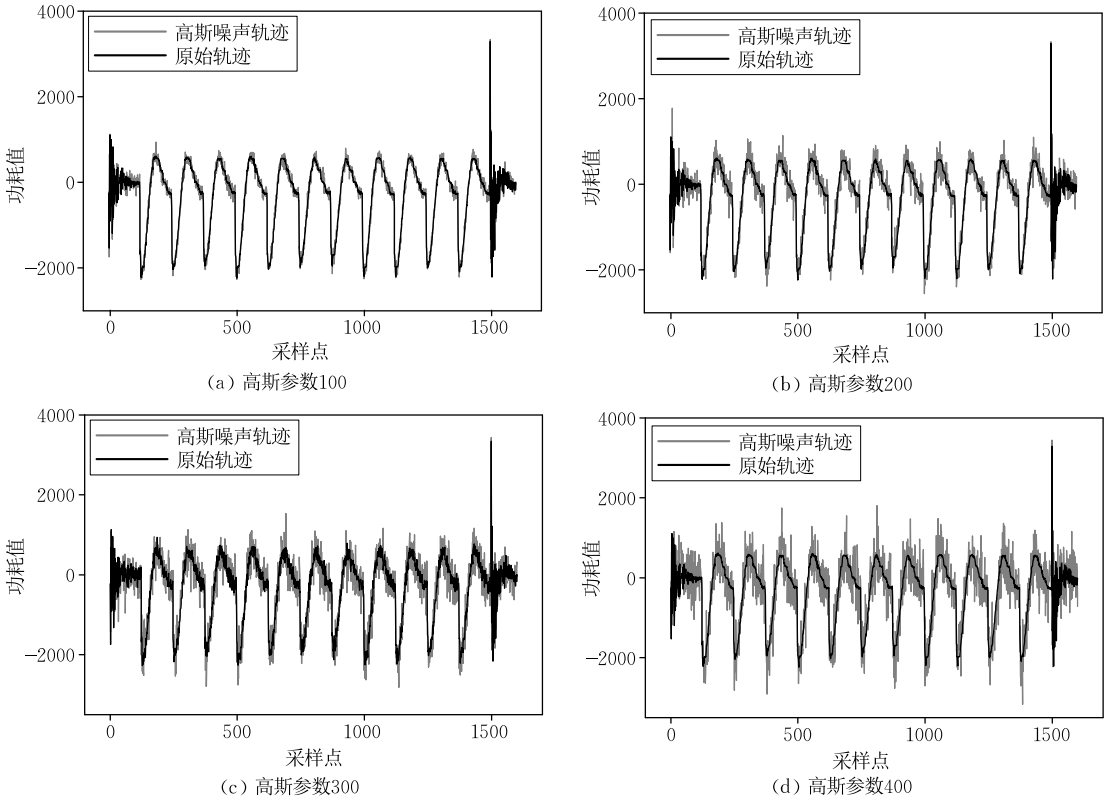

图 13　高斯噪声对策下的能量迹

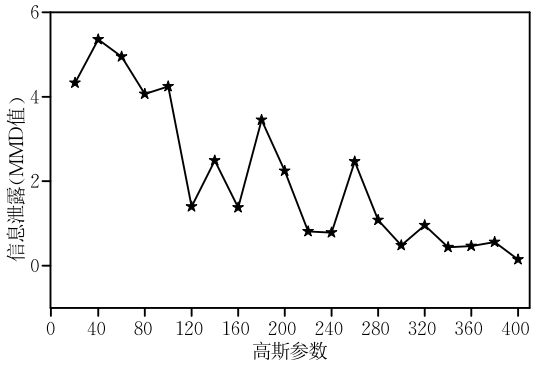

图 14　高斯噪声对策下 MMD-SCLA 泄露评估

(3)一阶掩码

ASCAD v1 数据集具有 10 万条固定密钥能量迹，和 20 万条随机密钥能量迹. 将固定密钥和随机密钥对应的能量迹分成两组，各自抽取 1000 条分成两个集合，针对每条能量迹的 250 000 个样本点，间隔设置为 125，每个间隔范围内对应的功耗值求平均，形成 2000 个新的样本点组成，每组能量迹集合维度为(1000，2000). 提取结果如图 15 所示. 使用 MMD 方法对处理后的数据集进行泄露评估实验，20 次重复实验后的结果如图 16 所示.

根据实验结果可以发现采取掩码防御对策的 AES加密能量迹，进行MMD检验的泄露峰值为 0.47.

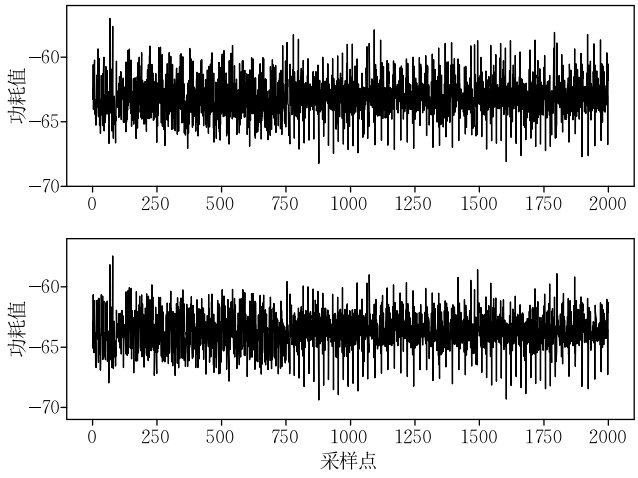

图 15　ASCAD v1 样本点能量迹

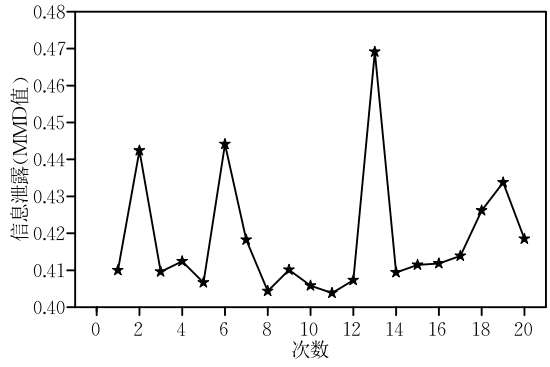

图 16　掩码对策下 MMD-SCLA 泄露评估

4.4 对照方案泄露评估实验及分析

4.4.1 无防御对策数据集对比实验

能量迹数据集与 4.3.1 节保持一致，以 TVLA 单变量泄露检验以及 HAC 多变量泄露检验作为对照实验，使用 CPA 对无防御对策 AES 进行侧信道攻击，得出的 MTD 作为侧信道定量指标，进一步验证 MMD-SCLA 泄露评估方法的有效性.

(1) MTD

下面针对三个无防御对策的数据集进行 CPA 分析，采用逐个密钥字节单独攻击的手段，使用汉明重量模型对 AES 算法的最后一轮的最后一个字节进行攻击. 对 DPA contest v2 数据集进行分析，MTD 值为 58；对于自采数据集 STM32F407，MTD 值为 97；对于自采数据集 SAKURA-X，MTD 值为 269. 单个密钥字节皮尔森相关性系数结果如图 17～图 19 所示.

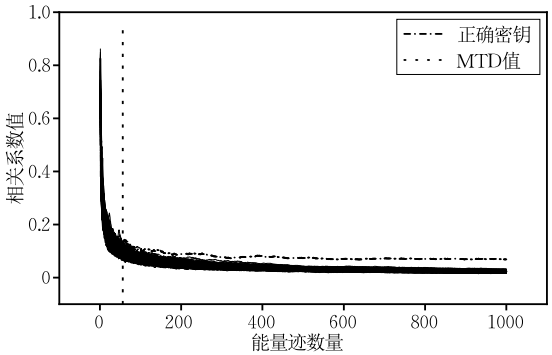

图 17 DPA contest v2 CPA 攻击结果

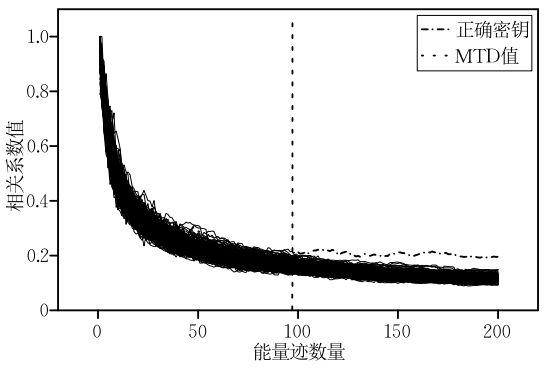

图 18 STM32F407 CPA 攻击结果

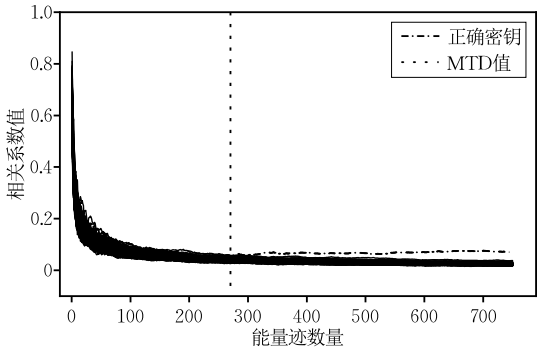

图 19 SAKURA-X CPA 攻击结果

(2) TVLA

使用 TVLA 进行泄露分析，结果如图 20～图 22 所示. TVLA 评估结果峰值分别为 81.39、253.19、268.97. 泄露阈值采用文献[12]的建议，取为±4.5. 当采样点的 t 值的绝对值大于 4.5 时，判断该处存在能量泄露. t 值越大，泄露程度越严重.

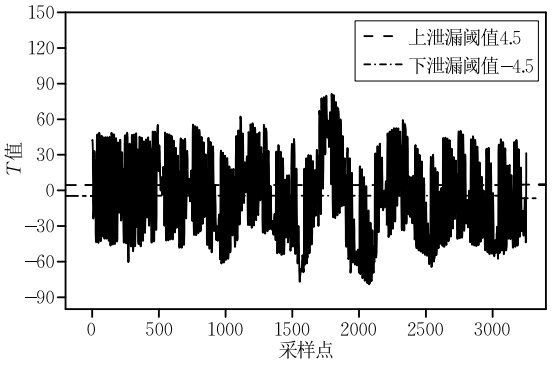

图 20 DPA contest v2 TVLA 评估结果

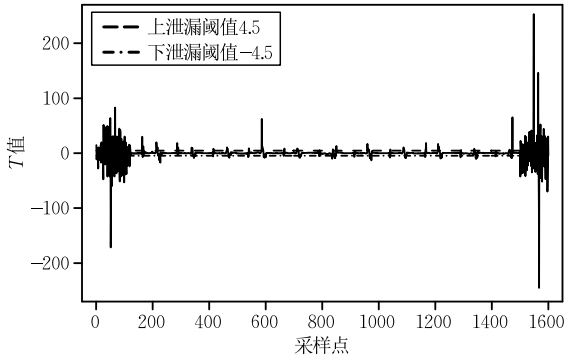

图 21 STM32F407 TVLA 评估结果

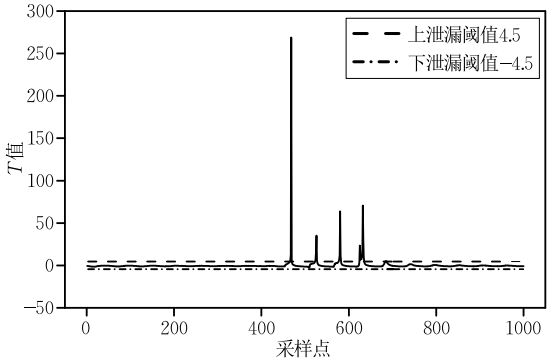

图 22 SAKURA-X TVLA 评估结果

(3) HAC

针对无防御对策的数据集使用 HAC 进行泄露分析结果如图 23～图 25 所示，评估结果分别为 0.92、0.99、0.72. 图中横轴 0 代表最近邻居，1 代表最远邻居. 纵轴是第 k 邻居节点在固定密钥和随机密钥对应的联合能量迹中的比例. 在中间的深灰色中心线 z 是平均斜率，平均斜率越大，可交换性就越低，AES 加密过程中能量泄露严重，被测设备安全

性较低，浅灰色线和黑色线分别是置信区间下限和上限.

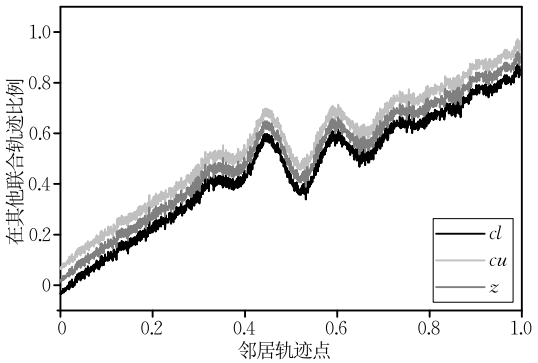

图 23　DPA contest v2 HAC 评估结果

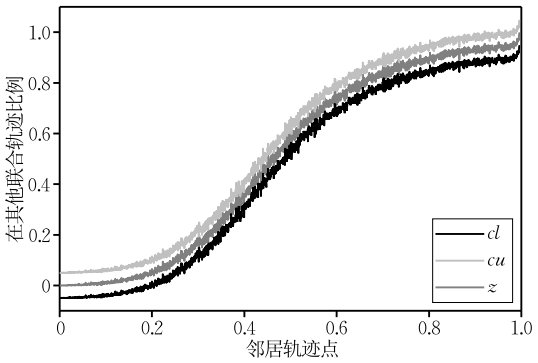

图 24　STM32F407 HAC 评估结果

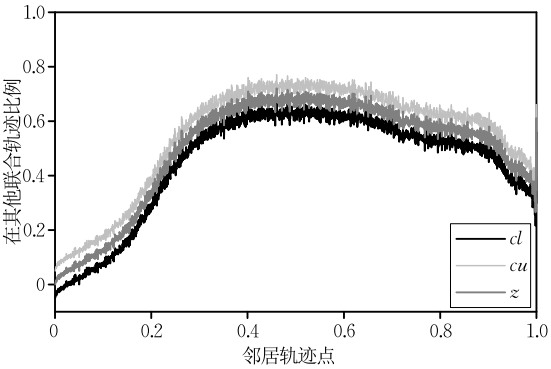

图 25　SAKURA-X HAC 评估结果

4.4.2　有防御对策数据集对比实验

能量迹数据集与 4.3.2 节保持一致，以 TVLA 单变量泄露检验、HAC 多变量泄露检验以及 Bartlett-F 检验作为对照实验.

(1) TVLA

根据 4.3 节 MMD-SCLA 能量泄露分析峰值结果，对 STM32F407 能量迹数据集设置随机延时参数为 0.02 作为延时条件下的实验数据集，设置高斯噪声参数为 40 作为噪声条件下的实验数据集，以及 ASCAD v1 掩码数据集作为掩码条件下的实验数据集，使用 TVLA 对三种能量泄露场景进行评估结果如图 26 和图 27 所示.

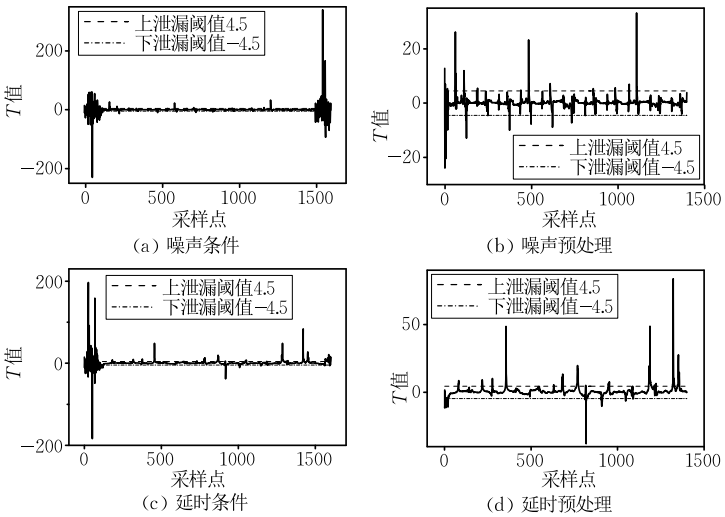

图 26　噪声条件和延时条件 TVLA 评估

分析 TVLA 评估的三种结果，在高斯噪声和随机延时条件下，TVLA 泄露评估取得 t 值的峰值对应坐标分别为(1547，339.70)，(25，196.93). 由能量迹曲线可知，第 1547 和第 25 个样本点不属于 AES 算法 10 轮加密过程的任一轮次. 由于高斯噪声和随机延时峰值点位置没有进行正常加密操作，所以对能量迹进行预处理，截取能量迹中实际加密的 1400 个样本点进行泄露评估，对应泄露峰值坐标为(1109，33.24)，(1320，83.73)(见图 26). 掩码条件下评估结果如图 27 所示，TVLA 泄露评估 t 值得峰值对应坐标(487，2.87)，未检测出泄露.

(2) HAC

使用 HAC 进行泄露评估，HAC 斜率分别为 0.83 和 0.56(见图 28、图 29). 对 ASCAD 掩码数据集进行泄露评估，HAC 斜率为 0.54，结果如图 30 所示.

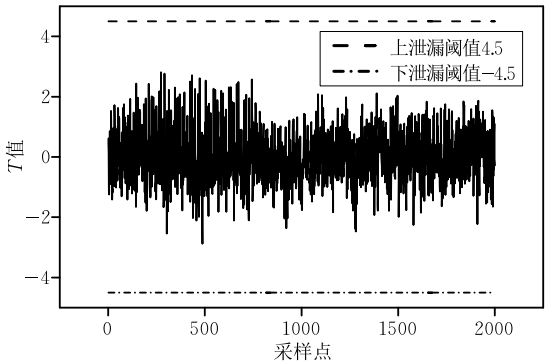

图 27 掩码条件 TVLA 评估

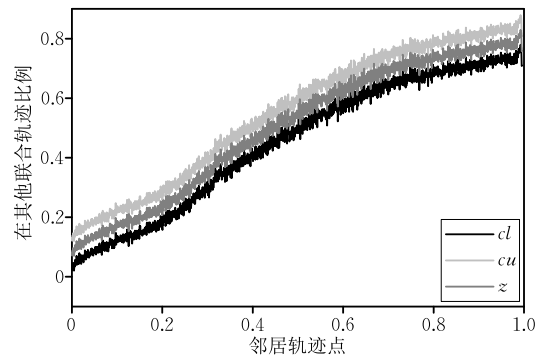

图 28 噪声条件 HAC 评估

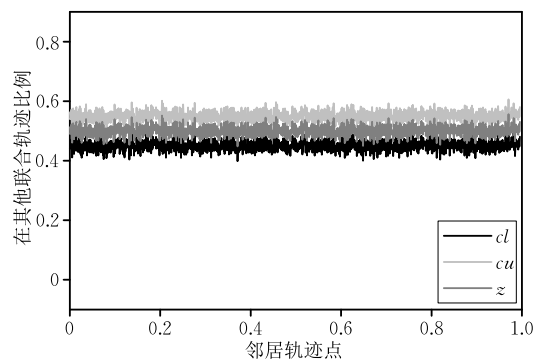

图 29 延时条件 HAC 评估

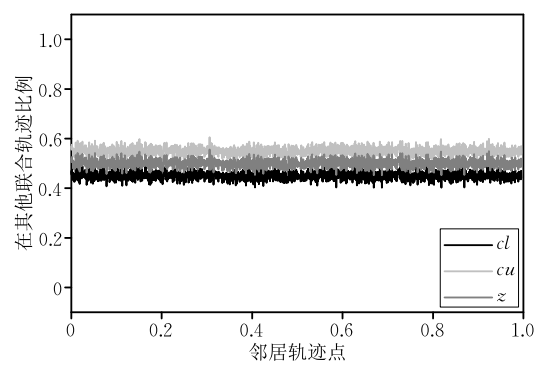

图 30 掩码条件 HAC 评估

(3) Bartlett-F 检验

Bartlett-F 检验使用 p 值作为评估指标，p 值越小，泄露程度越大. 文献[20]指出，当 p 值小于阈值 1×10^{-5} 时，即 $-\lg(p)$ 值大于 5，判断设备存在泄露. 在高斯噪声和随机延时条件下，Bartlett-F 泄露评估结果取得的峰值对应坐标分别为(1250，14.62)，(248，37.31)，如图 31 和图 32 所示.

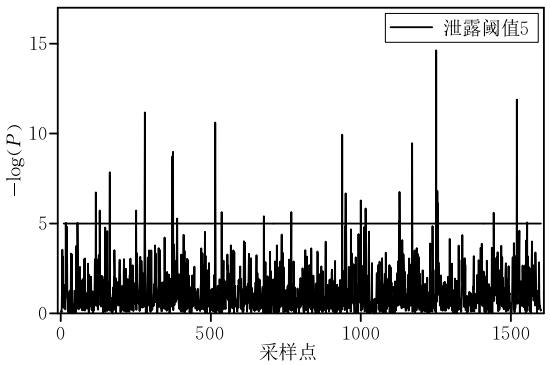

图 31 噪声条件 Bartlett-F 评估

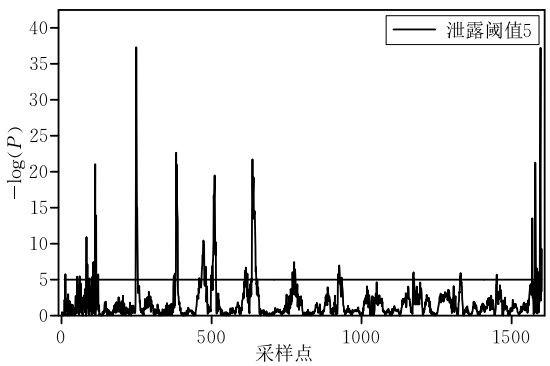

图 32 延时条件 Bartlett-F 评估

对 ASCAD 掩码数据集进行多变量泄露评估，处理方法根据文献[20]使用组合函数对原始 ASCAD 数据集进行预处理获取组合能量迹，Bartlett-F 泄露评估结果如图 33 所示，峰值对应的坐标为(3651，8.44).

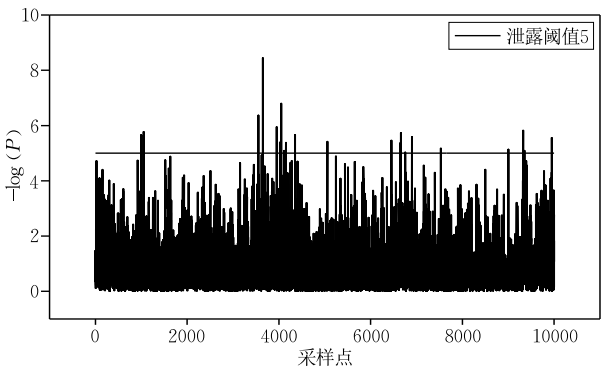

图 33 掩码条件 Bartlett-F 评估

4.4.3 实验结果对比分析

综合三个无防御对策 AES 数据集的 MTD 指标、HAC 指标、TVLA 的 t 值以及 MMD-SCLA 泄露评估峰值，能量泄露量化评估实验结果如表 2 所示. MMD-SCLA 的泄露评估值与 CPA 攻击所需能量迹数量的结果相对应，即 MTD 指标越小，能量泄

露程度越高，MMD-SCLA 泄露评估值越大. 反之 MTD 指标越大，能量泄露程度越低，MMD-SCLA 泄露评估值越小.

表 2　无防御对策评估方法对比分析

芯片设备	CPA MTD	TVLA	HAC	MMD-SCLA
SASEBO GII	58	81.39	0.92	6.96
STM32F407	97	253.19	0.99	5.21
SAKURA-X	269	268.97	0.72	2.26

当 MMD-SCLA 泄露评估值增大时，就意味着对应能量迹分布差异性变大，被测设备的能量泄露程度就越高. 而当 MMD-SCLA 泄露评估值逐渐趋于 0 时，对应能量迹的分布差异性越小，被测设备能量泄露程度越低. 在三个芯片设备的 AES 实现中，SASEBO GII 的 MTD 值最小，MMD-SCLA 泄露评估值最大，因此相比 STM32F407 和 SAKURA-X 的能量泄露评估结果，其信息泄露的风险最大. 密码设备的安全性排名由高到低为 SAKURA-X、STM32F407 和 SASEBO GII.

综合有防御对策的 HAC 指标，TVLA 的 t 值，Bartlett-F 的 p 值以及 MMD-SCLA 泄露评估峰值，能量泄露量化评估实验结果如表 3 所示. 对比分析 MMD-SCLA 和 TVLA 两种方法的结果，在随机延时和高斯噪声对策下的能量泄露，使用 MMD-SCLA 泄露评估方法降低了单变量评估出现的假阴性的风险，成功提取高斯噪声对策下多个样本点泄露特征，检测出其存在泄露. 在掩码对策下，TVLA 和 MMD-SCLA 泄露评估结果一致，该对策的安全性最强.

表 3　不同防御对策泄露评估方法对比分析

硬件	防御对策	TVLA	Bartlett-F	HAC	MMD-SCLA
STM32F407	高斯噪声	33.24	14.62	0.83	5.35
STM32F407	随机延时	83.73	37.31	0.56	0.63
ATmega 8515	一阶掩码	2.87	8.44	0.54	0.47

对于高斯噪声对策评估结果分析，HAC 斜率、MMD-SCLA 泄露评估值、TVLA 评估值和 Bartlett-F 评估值相对较高，泄露程度严重，攻击者利用能量迹恢复密钥的难度较小. 对于随机延时对策，HAC 斜率和 MMD-SCLA 泄露评估值相对较小. 对于掩码防御对策，四种方法的评估结果均相对较小，代表设备应用掩码后泄露程度较低，攻击者利用的难度增大，密码设备相对安全. 使用 MMD-SCLA 通过量化分布差异性与 HAC 通过邻居节点量化分布是否存在可交换性作为对照，其结果所代表密码设备的能量泄露程度具有一致性. 针对密码设备采取的不同防御对策，安全性排名由高到低依次为一阶掩码，随机延时和高斯噪声.

综上所述，使用 MMD-SCLA 进行评估避免了单变量 TVLA 评估丢失兴趣点之间的联系，对照实验进一步验证了 MMD-SCLA 方法的有效性.

5　总　结

本文研究了密码设备能量侧信道安全性问题，提出了基于 MMD 的能量泄露量化评估方法，降低了单变量 TVLA 方法面对多变量联合泄露场景下存在假阴性的问题. 以 TVLA、HAC 和 Bartlett-F 评估方法作为对照方案，MTD 指标作为参考，对 AES 无防御对策和有防御对策下的不同场景进行量化分析. 在高斯噪声防御对策下，TVLA 分析结果能量泄露较少，而 HAC 和 MMD-SCLA 分析结果能量泄露程度严重，MMD-SCLA 相比 TVLA 单变量泄露评估更准确. 通过 HAC 指标、MTD 指标和 Bartlett-F 指标的实验结果，验证了 MMD-SCLA 方法的有效性. MMD-SCLA 可广泛应用于评估不同防御对策条件下的多变量泄露分析，与特定的能量消耗模型或攻击方法无关，对不同密码设备的安全性进行综合性评估，为侧信道能量泄露评估提供了有效手段. 文中以 AES 密码设备为例，对于其他涉及能量泄露的密码算法如 SM4 也具有通用意义.

参　考　文　献

[1] Kocher P C. Timing attacks on implementations of Diffie-Hellman, RSA, DSS, and other systems//Proceedings of the Annual International Cryptology Conference. Berlin, Germany: Springer, 1996: 104-113

[2] Mangard S, Oswald E, Standaert F X. One for all-all for one: Unifying standard differential power analysis attacks. IET Information Security, 2011, 5(2): 100-110

[3] Brier E, Clavier C, Olivier F. Correlation power analysis with a leakage model//International Workshop on Cryptographic Hardware and Embedded Systems. Berlin, Germany: Springer, 2004: 16-29

[4] Chari S, Rao J R, Rohatgi P. Template attacks//Proceedings of the International Workshop on Cryptographic Hardware and Embedded Systems. Berlin, Germany: Springer, 2002: 13-28

[5] Schramm K, Leander G, Felke P, et al. A collision-attack on AES: Combining side channel-and differential-attack//Proceedings of the 6th International Workshop Cambridge on Cryptographic Hardware and Embedded Systems. MA, USA, 2004: 163-175

[6] Steuer R, Kurths J, Daub C O, et al. The mutual information: Detecting and evaluating dependencies between variables. Bioinformatics, 2002, 18(2): 231-240

[7] Tiri K, Hwang D, Hodjat A, et al. Prototype IC with WDDL and differential routing—DPA resistance assessment//Proceedings of the International Workshop on Cryptographic Hardware and Embedded Systems. Berlin, Germany: Springer, 2005: 354-365

[8] Cooper J, DeMulder E, Goodwill G, et al. Test vector leakage assessment (TVLA) methodology in practice//Proceedings of the International Cryptographic Module Conference. Gaithersburg, USA, 2013: 13

[9] Bronchain O, Schneider T, Standaert F X. Multi-tuple leakage detection and the dependent signal issue. IACR Transactions on Cryptographic Hardware and Embedded Systems, 2019, (2): 318-345

[10] Moradi A, Richter B, Schneider T, et al. Leakage detection with the x2-test. IACR Transactions on Cryptographic Hardware and Embedded Systems, 2018, 2018(1): 209-237

[11] Chen Hua, Xi Wei, Fan Li-Min, et al. Side channel analysis and evaluation on cryptographic products. Journal of Electronics & Information Technology, 2020, 42(8): 1836-1845(in Chinese)
(陈华, 习伟, 范丽敏等. 密码产品的侧信道分析与评估. 电子与信息学报, 2020, 42(8): 1836-1845)

[12] Goodwill G, Jun B, Jaffe J, Rohatgi P. A testing methodology for side-channel resistance validation//Proceedings of the NIST Non-Invasive Attack Testing Workshop. Boston, USA, 2011: 115-136

[13] Schneider T, Moradi A. Leakage assessment methodology. Journal of Cryptographic Engineering, 2016, 6(2): 85-99

[14] Standaert F X. How (not) to use welch's t-test in side-channel security evaluations//Proceedings of the International Conference on Smart Card Research and Advanced Applications. Cham: Springer, 2018: 65-79

[15] Moos T, Wegener F, Moradi A. DL-LA: Deep Learning Leakage Assessment: A modern roadmap for SCA evaluations. IACR Transactions on Cryptographic Hardware and Embedded Systems, 2021(3): 552-598

[16] Althoff A, Blackstone J, Kastner R. Holistic power side-channel leakage assessment: Towards a robust multidimensional metric//Proceedings of the 2019 IEEE/ACM International Conference on Computer-Aided Design (ICCAD). Westminster, USA, 2019: 1-8

[17] Bache F, Wloka J, Sasdrich P, et al. Multivariate TVLA-efficient side-channel evaluation using confidence intervals. IEEE Transactions on Computers, 2023, 72(8): 1-12

[18] Chen Jia-Zhe, Li He-Xin, Wang Ya-Nan, Wang Yu-Hang. Evaluating side-channel information leakage in 3DES using the t-test. Journal of Tsinghua University (Science and Technology), 2016, 56(5): 499-503(in Chinese)
(陈佳哲, 李贺鑫, 王亚楠, 王宇航. 运用 t 检验评估 3DES 算法的侧信道信息泄露. 清华大学学报(自然科学版), 2016, 56(5): 499-503)

[19] Cao Yu-Chen, Zhou Yong-Bin. Multi-channel fusion leakage detection. Journal of Cyber Security, 2020, 5(6): 40-52(in Chinese)
(曹雨晨, 周永彬. 多源融合信息泄漏检测方法. 信息安全学报, 2020, 5(6): 40-52)

[20] Wang Ya-Ru, Tang Ming. Side channel leakage assessment with the Bartlett and multi-classes F-test. Journal on Communications, 2021, 42(12): 35-43(in Chinese)
(王娅茹, 唐明. 基于 Bartlett 和多分类 F 检验侧信道泄露评估. 通信学报, 2021, 42(12): 35-43)

[21] Zheng Zhen, Yan Ying-Jian, Cai Jue-Song, Liu Yan-Jiang. Non-specific TVLA method based on two-sample KS test. Journal on Communications, 2023, 44(5): 137-147(in Chinese)
(郑震, 严迎建, 蔡爵嵩, 刘燕江. 基于双样本 KS 检验的非特定 TVLA 方法. 通信学报, 2023, 44(5): 137-147)

[22] Mangard S, Oswald E, Popp T. Power Analysis Attacks. Translated by Feng Deng-Guo, Zhou Yong-Bin, Liu Jing-Ye. Beijing: Science Press, 2010(in Chinese)
(Mangard S, Oswald E, Popp T. 能量分析攻击. 冯登国, 周永彬, 刘继业译. 北京: 科学出版社, 2010)

[23] Hall of fame in DPA contest v4.2 competition. http://www.dpa-contest.org/v4/42_hall_of_fame.php

[24] Jayakumar M. FPGA-masked S-box implementation for AES engine//Proceedings of the VLSI Design: Circuits, Systems and Applications. Singapore: Springer, 2018: 223-233

[25] Benadjila R, Prouff E, Strullu R, et al. Deep learning for side-channel analysis and introduction to ASCAD database. Journal of Cryptographic Engineering, 2020, 10(2): 163-188

[26] Asfand Hafeez M, Mazyad Hazzazi M, Tariq H, et al. A low-overhead countermeasure against differential power analysis for aes block cipher. Applied Sciences, 2021, 11(21): 10314

[27] Vosoughi M A, Wang L, Köse S. Bus-invert coding as a low-power countermeasure against correlation power analysis attack//Proceedings of the 2019 ACM/IEEE International Workshop on System Level Interconnect Prediction (SLIP). Las Vegas, USA, 2019: 1-5

[28] Kian A, Hosseintalaee H, Jahanian A. Protecting the FPGA IPs against higher-order side channel attacks using dynamic partial reconfiguration//Proceedings of the 20th International Symposium on Computer Architecture and Digital Systems (CADS). Rasht, Iran, 2020: 1-4

[29] Avirneni N D P, Somani A K. Countering power analysis attacks usingreliable and aggressive designs. IEEE Transactions on Computers, 2013, 63(6): 1408-1420

[30] Couroussé D, Barry T, Robisson B, et al. Runtime code polymorphism as a protection against side channel attacks//Proceedings of the IFIP International Conference on Information Security Theory and Practice. Cham: Springer, 2016: 136-152

[31] Chen S, Ge W, Yang J, et al. A power analysis attack countermeasure based on random execution//Proceedings of the 17th IEEE International Conference on Trust, Security and Privacy in Computing and Communications/12th IEEE International Conference on Big Data Science and Engineering (TrustCom/BigDataSE). New York, USA, 2018: 1474-1479

[32] Chong K S, Ng J S, Chen J, et al. Dual-hiding side-channel-attack resistant FPGA-based asynchronous-logic AES: Design, countermeasures and evaluation. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 2021, 11(2): 343-356

[33] Alipour A, Papadimitriou A, Beroulle V, et al. On the performance of non-profiled differential deep learning attacks against an AES encryption algorithm protected using a correlated noise generation based hiding countermeasure//Proceedings of the 2020 Design, Automation & Test in Europe Conference & Exhibition (DATE). Grenoble, France, 2020: 614-617

[34] Prouff E, Rivain M, Bevan R. Statistical analysis of second order differential power analysis. IEEE Transactions on Computers, 2009, 58(6): 799-811

[35] Gretton A, Borgwardt K M, Rasch M J, et al. A kernel two-sample test. The Journal of Machine Learning Research, 2012, 13(1): 723-773

[36] Fukumizu K, Gretton A, Sun X, et al. Kernel measures of conditional dependence//Proceedings of the 20th International Conference on Neural Information Processing Systems (NIPS 2007). Vancouver, Canada, 2007: 489-496

[37] Sriperumbudur B K, Gretton A, Fukumizu K, et al. Hilbert space embeddings and metrics on probability measures. The Journal of Machine Learning Research, 2010, 11(3): 1517-1561

[38] Liu F, Xu W, Lu J, et al. Learning deep kernels for non-parametric two-sample tests//Proceedings of the 37th International Conference on Machine Learning. Vienna, Austria, 2020: 6316-6326

[39] Sutherland D J, Tung H Y, Strathmann H, et al. Generative models and model criticism via optimized maximum mean discrepancy. arXiv preprint arXiv:1611.04488, 2016

[40] Schneider T, Moradi A, Standaert F X, et al. Bridging the gap: Advanced tools for side-channel leakage estimation beyond Gaussian templates and histograms//Proceedings of the 23rd International Conference on Selected Areas in Cryptography. St. John's, Canada, 2017: 58-78

[41] Deng J, Wang Y, Guo J, et al. A similarity measure based on Kullback-Leibler divergence for collaborative filtering in sparse data. Journal of Information Science, 2019, 45(5): 656-675

[42] Becker G, Cooper J, De Mulder E, et al. Test Vector Leakage Assessment (TVLA) Derived Test Requirements (DTR) with AES//Proceedings of the International Cryptographic Module Conference. Gaithersburg, USA, 2013: 1-8

[43] Wu L, Picek S. Remove some noise: On pre-processing of side-channel measurements with autoencoders. IACR Transactions on Cryptographic Hardware and Embedded Systems, 2020(4): 389-415

[44] Swaminathan S, Chmielewski L, Perin G, et al. Deep learning-based side-channel analysis against AES inner rounds// Applied Cryptography and Network Security Workshops: ACNS 2022 Satellite Workshops. Rome, Italy, 2022: 165-182

HONG Liang, Ph. D., associate professor. His current research interests include unmanned system security and hardware security.

ZHAI Yuan-Jie, M. S. candidate. Her current research interests include cryptography, side-channel analysis and protection.

WANG Jia-Xi, M. S. His current research interests include cryptography, side-channel analysis and protection.

ZHENG Jian, M. S. His current research interests include hardware security and side-channel analysis.

HU Wei, Ph. D., professor. His current research interests include hardware security, formal security verification methods, cryptography and reconfigurable computing.

Background

Quantitative power leakage assessment of cryptographic devices is essential for analyzing the extent of information leakage from cryptographic devices. Researchers have focused their research on quantitative power leakage assessment techniques in four different directions: TVLA leakage assessment techniques quantify the variability of the mean values of two sets of power leakage traces collected for a specific device under test; χ^2 leakage assessment techniques quantify the variability between leaked power traces according to a probability distribution; deep learning leakage assessment techniques use neural networks as differentiators to distinguish between power traces with a non-negligible probability of there is leakage; the HAC comprehensive leakage assessment framework quantifies the degree of leakage by the exchangeability of power trajectory distributions. The focus of this paper is to evaluate the power-side channel leakage of AES by determining whether the distribution of power trajectory samples corresponding to two sets of keys is the same and assess the degree of leakage in the encryption process of cryptographic devices quantitatively. The MMD method is used to establish a quantitative assessment model of power leakage and to verify the effectiveness of our method by using power leakage under different scenarios with no defensive countermeasures and with defensive countermeasures (mask, delay, Gaussian noise). The experimental results of the TVLA and HAC methods for leakage assessment were used as comparison schemes. The MTD metric obtained from the CPA attack was used as a reference to experimentally validate the effectiveness of the multivariate leakage assessment method based on the maximum mean difference, avoids probabilistic adaptation and overfitting in the assessment process with deep learning, and enables leakage detection in both vertical and horizontal directions. Furthermore, our work can evaluate multivariate leakage analysis under different defense countermeasures, providing an effective tool for assessing side-channel power leakage. The result is also helpful for other symmetric cryptographic algorithms involving power leakage, such as SM4.

This research is supported by the National Key Research and Development Program of China (2021YFB3100901) and the National Natural Science Foundation of China under the project "Fine-Grained Standardized Hardware Security Model and Metric Based on Joint Information Flow Analysis" (No. 62074131), and the project funded by "Tongxin Project" of Institute of Aerospace 772.