基于时频特征的多源融合信息泄露检测方法

冯 祺[1,2], 周永彬[1,2], 明经典[1,2], 张 倩[1,2]

[1] 中国科学院信息工程研究所信息安全国家重点实验室 北京 中国 100093
[2] 中国科学院大学网络空间安全学院 北京 中国 100049

摘要 密码芯片在运行过程中会同时产生能量消耗、电磁辐射等多种信息泄露, 而信息泄露利用对密码设备的实际安全性造成严重威胁。泄露检测是评估密码设备信息泄露风险威胁的一项重要技术, 主要通过假设检验的方式检测密码设备是否存在与敏感数据相关的信息泄露。仅对其中一种特定类型的信息泄露进行检测容易忽视多种信息泄露之间存在的内在关联性, 故难以充分刻画密码设备的实际安全性。多源融合信息泄露检测是试图克服这一重要技术缺陷的新方向。本文提出基于时频特征的多源融合信息泄露检测方法, 在确定性和非确定性检测两种场景, 基于时频特征的多源融合信息泄露检测方法充分利用假设检验 t-test、Hotelling's T2-test、F-test、Wilk's Lambda-test 的特性, 并将这四种假设检验方法与信息泄露的时域和频域特征进行融合, 深入挖掘与敏感信息相关的信息泄露。本文通过频率信息泄露点密度、信噪比、维数等多种因素与检测出泄露所需侧信息数量的关系, 分析了基于时频特征的多源融合信息泄露检测方法的可行性与适用性。实验结果表明, 在采样点数量相同的情况下, 与已有检测方法相比, 本文新方法的误报率降低 99.33%～99.97%; 在确定性检测情况下, 与已有检测方法相比, 本文新方法检测出泄露所需侧信息数量降低 15%～52%; 在非确定性检测情况下, 与已有检测方法相比, 本文新方法检测出泄露所需侧信息数量降低 29%～64%。

关键词 密码设备; 侧信道分析; 泄露检测; 时频特征; 多源融合
中图法分类号 TP309.2 DOI 号 10.19363/J.cnki.cn10-1380/tn.2026.03.14

Time-Frequency Characteristics Based Multi-Channel Fusion Leakage Detection

FENG Qi[1,2], ZHOU Yongbin[1,2], MING Jingdian[1,2], ZHANG Qian[1,2]

[1] State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China.
[2] School of Cyber Security, University of Chinese Academy of Sciences, Beijing 100049, China.

Abstract Various information leakages such as power and electromagnetic are generated during the running of cryptographic devices, and The utilization of information leakage poses a serious threat to the actual security of cryptographic device. Leakage detection is an important technology to assess the risk of leakage of cryptographic device, it is to find the evidence of dependency between leakages and sensitive data through hypothesis testing. Detecting only one specific type of information leakage ignores the inherent correlation between multiple information leakages, so it is difficult to fully characterize the actual security of cryptographic devices. Multi-channel fusion leakage detection is a new direction to overcome this technical defect. This paper proposes time-frequency characteristics based multi-channel fusion leakage detection. In both of specific and non-specific scenarios, time-frequency characteristics based multi-channel fusion leakage detection fully utilize the characteristics of hypothesis testing t-test, Hotelling's T2 test, F-test, and Wilk's Lambda test, and combine these four hypothesis testing methods with the time-domain and frequency-domain characteristics of information leakage to deeply explore information leakages related to sensitive data. This paper analyzes the feasibility and applicable scenarios of time-frequency characteristics based multi-channel fusion leakage detection by examining the relationship between multiple factors such as frequency information leakage density, signal-to-noise ratio, dimension, etc and the number of measures required to detect. The experimental results show that the false positive rate of the new method proposed in this paper is reduced by 99.33%-99.97% compared with the existing detection methods when the number of sampling points is the same. In the case of specific test, compared with the existing detection methods, the number of measures required to detect by the new method in this paper is reduced by 15%-52%. In the case of non-specific test, compared with the existing detection methods, the number of measures required to detect by the new method in this paper is reduced by 29%-64%.

Key words cryptographic device; side channel analysis; leakage detection; time-frequency characteristics; multi-channel fusion

通讯作者: 周永彬, 博士, 研究员, Email: zhouyongbin@iie.ac.cn。
本课题得到国家自然科学基金(No. 61632020, No. U1936209, No. 62002353)和北京市自然科学基金(No. 4192067)资助。
收稿日期: 2020-12-18; 修改日期: 2021-02-01; 定稿日期: 2023-08-09

1 引言

为保护敏感或秘密信息不被非法窃取, 加密算法通常被实现到不同类型的电子设备中, 如智能卡、移动终端、FPGA 等。本文将实现密码算法的电子设备统称为密码设备[1]。密码设备运行时通常会同时产生执行时间[2]、能量消耗[3]、电磁辐射[4]等多种信息泄露, 利用这些信息泄露进行密码分析的攻击通常称为侧信道分析。密码设备安全性受到侧信道分析的严重威胁, 例如, 2020 年 Standaert 团队的工作[5], 对基于 nRF52832 处理器的无线系统进行侧信道攻击, 在办公室环境采集距离 15 m 的电磁信息泄露, 并对其分析, 成功恢复出 AES-128 加密密钥。

事实上, 密码设计人员和认证机构早已关注如何评估密码实现的侧信道安全性。目前主要有两种评估方法: 攻击依赖性检测和一致性检测。其中, 攻击依赖性检测[6-7]旨在评估密码设备抵御侧信道攻击的能力。例如, 通用准则(Common Criteria, CC)[8]框架下的评估认证属于此种评估方法, 此时评估者需要实施多种已知的侧信道攻击, 包括简单能量分析、高阶差分能量分析等。这种评估方法存在两方面显著问题: 首先, 攻击方法和泄露模型严重依赖于专家经验, 故难以确定适用于待检设备的最优攻击方法; 其次, 新的攻击方法和泄露模型总会或早或晚被提出, 这就要求评估机构尽快更新攻击方法, 逐渐增多的攻击方法和泄露模型必然导致评估时间开销的增大。

一致性检测旨在判断密码设备产生的侧信息在指定的安全水平下能否通过某种统计检验, 例如, ISO/IEC17825[9]采用的测试向量泄露评估(Test Vector Leakage Assessment, TVLA)即属于此类评估方法。因此, 泄露检测尝试回答密码设备产生的侧信息中是否存在(可被检测到的)信息泄露。这种检测并不考虑任何具体的攻击方法, 因此无须通过精确刻画泄露特征并利用泄露模型对秘密信息进行恢复; 另一方面, 一致性泄露检测不需要依赖对具体侧信道攻击方法的相关专家知识, 降低了检测实施的技术门槛并缩短了检测评估时间。

泄露检测分为确定性检测(Specific Test)和非确定性检测(Non-Specific Test)。非确定性检测是根据明文划分侧信息, 例如, 基于(fixed-versus-random)t-test 的 TVLA[10-11]是一种非确定性检测。该方法检测不同敏感数据信息泄露分布的均值是否存在显著差异, 即对信息泄露分布的一种统计矩进行比较, 而基于 χ^2-test 的检测方法[12]可以检测到泄露的整体分布, 利用分布的多种统计矩。因此针对高阶掩码方案的检测, 相比于 t-test, χ^2-test 需要更少的侧信息检测出泄露。基于互信息[13-15]的检测方法也具有检测任何统计矩的能力, 因此当存在泄露的统计矩未知时, 可以利用基于 χ^2-test 或基于互信息的检测方法。Fixed vs. random t-test 主要检测固定明文对应的信息泄露分布均值和随机明文对应的信息泄露分布均值是否显著不同。这种二划分策略与攻击方法、泄露模型无关, 所以它检测出的泄露点不一定可以恢复正确的密钥, 因此检测出用于恢复正确密钥的特征点是泄露检测的一项补充任务。

确定性检测根据依赖于密钥的中间值划分侧信息, 具有检测出特征点的潜力[16]。与非确定性检测相比, 由于确定性检测是检测中间值的信息泄露, 所以泄露点密度[17]通常较低。Bhasin 等人[18]提出的基于 F-test 的检测方法适用于确定性检测。相比于基于 t-test 的检测方法, 基于 F-test 的检测方法需要更多的侧信息检测出泄露, 具有较高的数据复杂度, 因为基于 F-test 的检测方法是根据中间值对比信息泄露多种划分的差异, 而基于 t-test 的检测方法是对比信息泄露二划分的差异, 具有更简单的估计任务。由于确定性检测的增益效果与非确定性检测互补, 故针对确定性检测仍有其他研究, 如 Durvaux 等人[19]提出基于相关性的检测方法。

通常情况下, 基于假设检验的泄露检测总会不可避免地存在误报率和漏报率, 其中误报率是指错误地检测到信息泄露的概率, 而漏报率则指未检测到信息泄露的概率。TVLA 建议单次检测的显著水平 α (即误报率)为 0.00001, 相应地, 整体检测的误报率 α_{overall} (即对于采样点数量为 n_l 的侧信息, 至少一个采样点存在误报的概率为 $\alpha_{\text{overall}}=1-(1-\alpha)^{n_l}$)则随着采样点数量的增多而增大。高误报率的检测意味着一个非泄露的设备难以通过检测。针对该问题, Ding 等人[16]引入 Šidák 校正[20]方法, 该校正方法是根据采样点数量 n_l 对误报率 α 进行调整, 以保持整体误报率 α_{overall} 不变。可是经过 Šidák 校正后, TVLA 的漏报率 β 随着采样点数量的增多而升高。一个高漏报率的检测意味着可能高估密码设备的安全性。上述问题本质上是因采样点的数量较多而进行多次检测引起的。

从泄露检测方法角度来看, 已有 TVLA 研究工

作只是从一方面评估密码设备安全性, 而综合安全性评估不足。例如, 异步时序逻辑电路可抑制能量信息泄露, 却增强了电磁信息泄露[21]。因此, 研究多源融合泄露检测, 以期实现对密码设备更全面的侧信道安全性评估, 是一种现实技术途径。为系统地研究多源融合泄露检测, 根据不同的融合层次将多源融合泄露检测分为三种策略: 数据级融合泄露检测、特征级融合泄露检测以及决策级融合泄露检测。数据级融合泄露检测是通过 TVLA 对多源侧信息融合检测。特征级融合泄露检测是利用 TVLA 对多源侧信息产生的特征集合融合检测。其中, 融合检测分为两种: 1) 直接利用假设检验方法对多源信息泄露融合检测; 2) 多源信息泄露自身融合后进行检测。决策级融合泄露检测是通过 TVLA 对多源信息泄露分别检测, 而后将检测结果融合取得最终的泄露检测结果。

文献[22]提出三种数据级多源融合泄露检测方法, 分别为多源时频融合信息泄露检测、多源简单融合信息泄露检测及基于多元 T 检验的多源信息泄露检测。前两种方法利用 t-test 对能量和电磁侧信息融合后的时域信息进行检测。基于多元 T 检验的多源信息泄露检测利用 Hotelling's T^2-test 对能量和电磁时域信息相同时刻采样点进行二维检测。

文献[22]的工作主要有以下局限性: 1) 该工作只针对非确定性检测场景进行研究, 故难以评估多源融合泄露检测方法在不同检测场景下的检测能力。2) 针对某些密码实现的信息泄露检测, 文献[22]的方法具有较大的误报率, 如本实验掩码型防护 AES 硬件实现的信息泄露检测, 其误报率为 0.99329。为此, 本文在确定性和非确定性两种检测场景下开展基于时频特征的多源融合信息泄露检测方法研究, 该方法以能量和电磁侧信息符合相同的泄露模型为前提, 融合利用能量和电磁信息泄露的时频特征。本文与文献[22]工作的区别如表 1 所示。相比于文献[22], 本文不仅研究 t-test、Hotelling's T^2-test 在多源融合泄露检测方法中的应用, 而且研究 F-test、Wilk's Lambda-test 的应用, 以检测用于恢复正确密钥的特征点。此外, 本文从频率信息的泄露点密度、维数、信噪比等多个参数出发, 分析基于时频特征的多源融合信息泄露检测方法的可行性和适用性。

本文提出多源时频信息数据级融合泄露检测将能量和电磁频率信息在复数域相乘, 而后转换为时域信息进行检测, 由于复数域上的频率信息包括振幅和相位信息[23], 故相比于仅利用振幅信息的多源时频融合信息泄露检测方法, 该方法进一步提高了融合后信息泄露的信噪比, 降低了数据复杂度。

本文其余部分组织如下: 第 2 节介绍背景知识; 第 3 节提出基于时频特征的多源融合信息泄露检测方法; 第 4 节为实验研究, 主要与 TVLA、基于多元 T 检验的多源信息泄露检测方法等具有显著检测能力的方法[22]进行对比, 验证本文所提方法的检测能力; 第 5 节对本文进行总结和展望。

表 1 本文与文献[22]工作的区别

Table 1 The differences between our work and paper[22]

	文献[22]方法	本文方法
适用场景	非确定性检测	确定性和非确定性检测
泄露特征	能量和电磁泄露时域特征	能量和电磁泄露时频特征
融合层次	数据级融合	数据级、特征级、决策级融合
数学工具	t-test、Hotelling's T^2-test	F-test、Wilk's Lambda-test、t-test、Hotelling's T^2-test
量化参数	信噪比	频率信息泄露点密度、维数、相关系数、信噪比

2 背景知识

2.1 信息泄露模型

密码设备的能量和电磁信息泄露具有相似的泄露模型 L, 其公式如下:

$$L = \omega \times f(\upsilon) + \mathcal{N}$$

其中, ω 为常数系数, υ 为目标中间值变量, $\mathcal{N}$ 为随机噪声, f 为目标中间值变量的信息泄露转换函数, 常用的转换函数有汉明重量和汉明距离等。根据具体的密码设备选择 ω, $\mathcal{N}$ 以及函数 f。

2.2 TVLA

在 TVLA 框架的检测步骤中, 首先对待检设备(Device Under Test, DUT)采集一组侧信息, 而后对该组侧信息中的每一个采样点进行单变量检测, 旨在判断该组侧信息是否存在泄露。记采集到的侧信息为 $L = [L_1, \ldots, L_i, \ldots, L_{n_l}]$, 其中, L_i 为第 i 个采样点侧信息, n_l 为该组侧信息的采样点数量。对于第 i 个采样点, 泄露检测的零假设为“H_0: 不存在信息泄露”, 而备择假设则为“$\mathrm{H}_{\mathrm{alt}}$: 存在信息泄露”:

$$\mathrm{H}_0 : L_i = \mathcal{N}_i \text{ vs. } \mathrm{H}_{\mathrm{alt}} : L_i = V + \mathcal{N}_i$$

其中, V 表示泄露信号, $\mathcal{N}_i$ 表示噪声。

在 H_0 假设下, 当第 i 个采样点的统计量 s_i 被观测到的概率 p_i-value 较低时, 表明拒绝 H_0 假设的置信度较高, 因此认为第 i 个采样点存在泄露。在

TVLA 中, 如果存在至少一个采样点拒绝 H_0 假设, 即最小的 p_i-value 小于α, 则认为 DUT 存在泄露, 这种决策方式称为 minimum p-value (mini-p)策略。

不同的假设检验方法可以应用于 TVLA 框架中, 以适用于根据中间值划分侧信息的确定性检测和根据明文划分侧信息的非确定性检测。

2.3 非确定性检测

常用于非确定性检测的假设检验方法有 Welch's t-test 和其多维扩展 Hotelling's T^2-Test。

Welch's t-test. 在非确定性 t-test 检测中, 将第 i 个采样点侧信息 L_i 划分为两组侧信息 $L_{i,A}$ (由固定明文产生)、$L_{i,B}$ (由随机明文产生), 通过统计量 t_i 对比这两组侧信息, 其计算方法如下:

$$t_i = \frac{\bar{L}_{i,A} - \bar{L}_{i,B}}{\sqrt{s_{i,A}^2/N_A - s_{i,B}^2/N_B}},$$

$$\mathrm{df}_i = \frac{(s_{i,A}^2/N_A + s_{i,B}^2/N_B)^2}{(s_{i,A}^2/N_A)^2/(N_A-1) + (s_{i,B}^2/N_B)^2/(N_B-1)},$$

其中, $\bar{L}_{i,A}$ 和 $\bar{L}_{i,B}$ 分别为两组侧信息的样本均值, $s_{i,A}^2$ 和 $s_{i,B}^2$ 分别为两组侧信息的样本方差, N_A 和 N_B 分别为两组侧信息的样本数量。在 H_0 假设下, 统计量 t_i 服从自由度为 df_i 的 t 分布。此时, 可通过统计量 t_i 计算出 p_i-value, 其计算公式如下:

$$p_i\text{-value} = 2\times(1-\mathrm{CDF}(t_i, \mathrm{df}_i)),$$

其中, $\mathrm{CDF}(\cdot, \mathrm{df}_i)$ 表示 t 分布的累积分布函数。统计量阈值 th_i 与α具有对应关系, 因此可通过 t 分布累积分布函数计算获得。H_{alt} 假设为真时, t_i 服从非中心参数为 δ_i 的非中心 t 分布, 故漏报率 β_i 计算如下:

$$\begin{aligned}
\beta_i &= \mathrm{NCTCDF}(\mathrm{th}_i, \mathrm{df}_i, \delta_i),\\
\mathrm{th}_i &= \mathrm{CDF}^{-1}(1-\alpha/2, \mathrm{df}_i),\\
\delta_i &= \frac{(\bar{\mu}_{i,A} - \bar{\mu}_{i,B})}{\sqrt{\sigma_{i,A}^2/N_A + \sigma_{i,B}^2/N_B}},
\end{aligned} \tag{1}$$

其中, $\mathrm{NCTCDF}(\cdot, \mathrm{df}_i, \delta_i)$ 是非中心 t 分布的累积分布函数, $\bar{\mu}_{i,A}$、$\bar{\mu}_{i,B}$ 分别为总体 A、B 的均值, $\sigma_{i,A}$、$\sigma_{i,B}$ 分别为总体 A、B 的方差。

Hotelling's T^2-Test. Hotelling's T^2-Test 是 t-test 的多维扩展, 不仅考虑了多个维度的信息泄露, 而且考虑了信息泄露之间的相关性。T^2-Test 的统计量 T^2 服从自由度为 (n_l, n_A+n_B-2) 的 F 分布, 其计算公式如下:

$$T^2 \sim \frac{1}{\zeta}\mathrm{F}(n_l, N_A+N_B-2),$$

$$T^2 = \frac{N_A N_B}{N_A+N_B}(\bar{\boldsymbol{L}}_A - \bar{\boldsymbol{L}}_B)^{\mathrm{T}}\boldsymbol{S}^{-1}(\bar{\boldsymbol{L}}_A - \bar{\boldsymbol{L}}_B),$$

$$\zeta = \frac{(N_A+N_B-2)n_l}{(N_A+N_B-1-n_l)},$$

$$\boldsymbol{S} = \frac{\sum_{t=1}^{N_A}(\boldsymbol{l}_{t,A}-\bar{\boldsymbol{L}}_A)(\boldsymbol{l}_{t,A}-\bar{\boldsymbol{L}}_A)^{\mathrm{T}} + \sum_{t=1}^{N_B}(\boldsymbol{l}_{t,B}-\bar{\boldsymbol{L}}_B)(\boldsymbol{l}_{t,B}-\bar{\boldsymbol{L}}_B)^{\mathrm{T}}}{N_A+N_B-2},$$

其中, $\bar{\boldsymbol{L}}_A$、$\bar{\boldsymbol{L}}_B$ 分别为两组侧信息 A、B 的样本均值向量, $\boldsymbol{l}_{t,A}$、$\boldsymbol{l}_{t,B}$ 分别为 A、B 的第 t 个样本向量, $\boldsymbol{S}$ 为估计的协方差矩阵。因此 p-value 计算如下:

$$p\text{-value} = 1-\mathrm{CDF}(\zeta T^2, n_l, N_A+N_B-2), \tag{2}$$

其中, $\mathrm{CDF}(\cdot, n_l, N_A+N_B-2)$ 是 F 分布的累积分布函数, 由式 (2) 可知, p-value 与信息泄露的数量 (N_A, N_B)、维数 n_l 以及信息泄露之间的相关系数有关。当 H_{alt} 假设为真时, 统计量服从非中心参数为 δ 的非中心 F 分布, 故漏报率 计算如下。

$$\begin{aligned}
\beta &= \mathrm{NCFCDF}(\mathrm{th}, n_l, N_A+N_B-2, \delta),\\
\mathrm{th} &= \mathrm{CDF}^{-1}(1-\alpha, n_l, N_A+N_B-2)/\zeta,\\
\delta &= \frac{N_A N_B}{N_A+N_B}(\bar{\boldsymbol{\mu}}_{\boldsymbol{A}} - \bar{\boldsymbol{\mu}}_{\boldsymbol{B}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\bar{\boldsymbol{\mu}}_{\boldsymbol{A}} - \bar{\boldsymbol{\mu}}_{\boldsymbol{B}}),
\end{aligned} \tag{3}$$

其中, $\mathrm{NCFCDF}(\cdot, n_l, N_A+N_B-2, \delta)$ 是非中心 F 分布的累积分布函数, $\bar{\boldsymbol{\mu}}_{\boldsymbol{A}}$、$\bar{\boldsymbol{\mu}}_{\boldsymbol{B}}$ 是总体 A、B 的均值, $\boldsymbol{\Sigma}$ 为两个总体的协方差矩阵。

2.4 确定性检测

常用于确定性检测的假设检验方法有 F-test, 并引入其多维扩展 Wilk's Lambda-test。

F-test. 根据中间值变量 υ 和采用的泄露模型将侧信息划分为 g 组。在 H_0 假设下, 统计量 F_i 服从自由度为 $(g-1, N-g)$ 的 F 分布:

$$F_i \sim \mathrm{F}(g-1, N-g),$$

$$F_i = \frac{\frac{1}{g-1}\sum_{s=1}^{g} N_s(\bar{L}_{i,s} - \bar{L}_i)^2}{\frac{1}{N-g}\sum_{s=1}^{g}\sum_{t=1}^{N_s}(\bar{L}_{i,s,t} - \bar{L}_{i,s})^2},$$

其中, N 表示样本数量, 第 s 组侧信息共有 N_s 个样本, 其样本均值为 $\bar{L}_{i,s}$。$\bar{L}_{i,s,t}$ 表示 s 组侧信息的第 t 个样本, $\bar{L}_i$ 表示第 i 个采样点所有样本的均值。p_i-value 为 H_0 假设下大于统计量 F_i 的概率, 其计算公式如下:

$$p_i\text{-value}=1-\mathrm{CDF}(F_i,g-1,N-g),$$

其中, $\mathrm{CDF}(F_i,g-1,N-g)$ 表示 F 分布的累积分布函数, 当 $\mathrm{H_{alt}}$ 假设为真时, 统计量 F_i 服从非中心参数为 δ_i 的非中心 F 分布[24], 漏报率 β_i 计算如下:

$$\begin{aligned}&\beta_i=\mathrm{NCFCDF}(\mathrm{th}_i,g-1,N-g,\delta_i),\\&\mathrm{th}_i=\mathrm{CDF}^{-1}(1-\alpha,g-1,N-g),\\&\delta_i=\frac{\sum_{s=1}^{g}N_s(\bar{L}_{i,s}-\bar{L}_i)^2\big/N}{\sigma_i^{\,2}},\end{aligned}\tag{4}$$

Wilk's Lambda-test. Wilk's Lambda-test 检验是 F-test 的多维扩展。在 $\mathrm{H_0}$ 假设下, 统计量 Λ 服从自由度为 $(\mathrm{df}_1,\mathrm{df}_2)$ 的 F 分布:

$$\begin{aligned}&F_\tau=\frac{1-\Lambda^{1/q}}{\Lambda^{1/q}}\cdot\frac{\mathrm{df}_1}{\mathrm{df}_2}\sim\mathrm{F}(\mathrm{df}_1,\mathrm{df}_2),\\&\Lambda=\frac{\left|\sum_{s=1}^{g}\sum_{t=1}^{N_s}(\boldsymbol{L}_{s,t}-\bar{\boldsymbol{L}}_s)(\boldsymbol{L}_{s,t}-\bar{\boldsymbol{L}}_s)^{\mathrm{T}}\right|}{\left|\sum_{s=1}^{g}\sum_{t=1}^{N_s}(\boldsymbol{L}_{s,t}-\bar{\boldsymbol{L}})(\boldsymbol{L}_{s,t}-\bar{\boldsymbol{L}})^{\mathrm{T}}\right|},\\&\mathrm{df}_1=n_l(g-1),\\&\mathrm{df}_2=q\left[N-g-\frac{n_l-g+2}{2}\right]-\frac{n_l(g-1)-2}{2},\\&q=\sqrt{\frac{n_l^2(g-1)^2-4}{n_l^2+(g-1)^2-5}},\end{aligned}$$

此时, p-value 计算如下:

$$p\text{-value}=1-\mathrm{CDF}(F_\tau,\mathrm{df}_1,\mathrm{df}_2),\tag{5}$$

其中, $\mathrm{CDF}(\cdot,\mathrm{df}_1,\mathrm{df}_2)$ 为 F 分布的累积分布函数。由式(5)可知, p-value 受到信息泄露的数量 N、维数 n_l 以及信息泄露之间相关系数的影响。当 $\mathrm{H_{alt}}$ 假设为真时, 统计量 Λ 服从非中心参数为 δ 的非中心 F 分布[25], 故漏报率 β 计算如下:

$$\begin{aligned}&\beta=\mathrm{NCFCDF}(\mathrm{th},\mathrm{df}_1,\mathrm{df}_2,\delta),\\&\mathrm{th}=\mathrm{CDF}^{-1}(1-\alpha,\mathrm{df}_1,\mathrm{df}_2),\\&\delta=\frac{1-\lambda^{1/q}}{\lambda^{1/q}}\cdot N,\ \lambda=\frac{\left|\boldsymbol{\Sigma}\right|}{\left|\boldsymbol{\Sigma}+\boldsymbol{\Sigma}_{\boldsymbol{\mu}}\right|},\\&\boldsymbol{\Sigma}_{\boldsymbol{\mu}}=\frac{1}{n}\sum_{s=1}^{g}N_s(\bar{\boldsymbol{\mu}}_{\boldsymbol{s}}-\bar{\boldsymbol{\mu}})(\bar{\boldsymbol{\mu}}_{\boldsymbol{s}}-\bar{\boldsymbol{\mu}})^{\mathrm{T}},\end{aligned}\tag{6}$$

其中, $\boldsymbol{\Sigma}$ 表示组内协方差矩阵, $\boldsymbol{\Sigma}_{\boldsymbol{\mu}}$ 表示组间协方差矩阵。$\bar{\boldsymbol{\mu}}_{\boldsymbol{s}}$ 表示 s 组总体的均值向量, $\bar{\boldsymbol{\mu}}$ 表示总体的均值向量。

2.5 短时傅里叶变换

本文使用短时傅里叶变换(Short Time Fourier Transform, STFT)获得能量和电磁泄露的时频信息。与傅里叶变换主要区别在于, 短时傅里叶变换通过一个窗口函数将时间序列信号划分为数个较短的等长信号, 再分别计算每个较短信号的傅里叶变换, 以获得时频分布。离散短时傅里叶变换计算公式如下:

$$\mathrm{STFT}\{x[n]\}(m,w)=\sum_{n=-\infty}^{\infty}x[n]w[n-m]\mathrm{e}^{-jwn}$$

其中, $x[n]$表示时间序列信号, $w[n]$表示窗口函数, 常用窗口函数如 Rectangula, Hann, Hamming, Blackman-Harris 等[26]。

3 基于时频特征的多源融合泄露检测方法

本文从融合角度设计了三个级别的基于时频特征的多源融合信息泄露检测方法(MCTFF-TVLA), 分别为多源时频信息决策级融合泄露检测方法(MCTFF-TVLA-Dc)、多源时频信息特征级融合泄露检测方法(MCTFF-TVLA-Ft)以及多源时频信息数据级融合泄露检测方法(MCTFF-TVLA-Dt),其中的多源主要包括能量和电磁信息泄露。

能量或电磁侧信息主要由泄露部分和噪声组成, 其中泄露部分是由电路中的数值转换和执行的操作引起的, 噪声是由电源、时钟发生器等元件造成的[1]。由于泄露部分和噪声的频率分布不同, 因此侧信息经过短时傅里叶变换后, 泄露部分和噪声将分布在不同的频率上。为提高泄露检测的能力, 可选择较大信噪比的频率分量进行检测。本文采用文献[27]建议的频率选择方法, 即选取频率信息中方差相对最大或最小的频率分量, 这是由于不同类型信号出现或消失引起了方差的突然变化。根据文献[27]频率选择方法, 能量和电磁信息泄露分别选择 p、(n_f-p)个较大信噪比的频率分量, 以提升信息泄露检测能力。

3.1 多源时频信息决策级融合泄露检测

本文提出多源时频信息决策级融合泄露检测方法, 其框架如图 1 所示, 首先对能量迹和电磁迹分别进行 STFT 变换, 获得能量和电磁频谱, 并将其转换为振幅和相位信息, 再从中选择出较大信噪比的频率信息分量。而后, 对选出的能量和电磁频率信息分量逐一按照单位时间进行单维泄露检测。为保持每一单位时间的误报率 α 为 0.00001, 对同一时间内每一个频率信息分量的显著水平 $\alpha_{\text{per-test}}$ 进行 Šidák 校正。具体实现细节见算法 1。

算法 1.多源时频信息决策级融合泄露检测;

输入: 能量迹 $\boldsymbol{L}_{\mathbf{Power}}$ 和电磁迹 $\boldsymbol{L}_{\mathbf{EM}}$;

输出: 是否存在泄露;

1: 对能量迹和电磁迹分别做短时傅里叶变换:

$$\boldsymbol{L}'_{\mathbf{Power}} = \mathrm{STFT}(\boldsymbol{L}_{\mathbf{Power}});$$

$$\boldsymbol{L}'_{\mathbf{EM}} = \mathrm{STFT}(\boldsymbol{L}_{\mathbf{EM}});$$

2: 从 $\boldsymbol{L}'_{\mathbf{Power}}$ 和 $\boldsymbol{L}'_{\mathbf{EM}}$ 中共选择出 n_f 个频率分量:

$$\boldsymbol{L}'_{\mathbf{Power1}} = \mathrm{select}(\boldsymbol{L}'_{\mathbf{Power}});$$

$$\boldsymbol{L}'_{\mathbf{EM1}} = \mathrm{select}(\boldsymbol{L}'_{\mathbf{EM}});$$

3:计算频率 $\boldsymbol{L}'_{\mathbf{Power1}}$ 和 $\boldsymbol{L}'_{\mathbf{EM1}}$ 的振幅和相位信息:

$$\boldsymbol{P}_{\mathbf{amplitude}} = \mathrm{abs}(\boldsymbol{L}'_{\mathbf{Power1}}); \boldsymbol{P}_{\mathbf{phase}} = \mathrm{angle}(\boldsymbol{L}'_{\mathbf{Power1}});$$

$$\boldsymbol{E}_{\mathbf{amplitude}} = \mathrm{abs}(\boldsymbol{L}'_{\mathbf{EM1}}); \boldsymbol{E}_{\mathbf{phase}} = \mathrm{angle}(\boldsymbol{L}'_{\mathbf{EM1}});$$

4: 对同一单位时间内的振幅信息和相位信息进行组合, 构成 m 组多维侧信息 $\boldsymbol{L}^j_{\mathbf{fusion}}$:

$$\boldsymbol{L}^j_{\mathbf{fusion}} = [\boldsymbol{P}^j_{\mathbf{amplitude}}, \boldsymbol{P}^j_{\mathbf{phase}}, \boldsymbol{E}^j_{\mathbf{amplitude}}, \boldsymbol{E}^j_{\mathbf{phase}}], j = 1,2,\ldots,m;$$

5: 对频率信息分量的 $\alpha_{\text{per-test}}$ 进行 Šidák 校正:

$$\alpha_{\text{per-test}} = 1-(1-\alpha)^{1/2n_f};$$

6: 依据具体检测场景选择 F-test 或 t-test 检验方法, 并对 $\boldsymbol{L}^j_{\mathbf{fusion}}$ 的每一维频率信息进行检测;

7: 如果 mini $p \le \alpha_{\text{per-test}}$, 返回存在泄露, 否则返回未存在泄露。

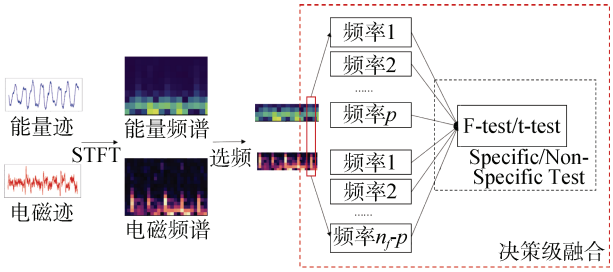

图 1 多源时频信息决策级融合泄露检测框架

Figure 1 The framework of multi-channel time-frequency fusion leakage detection in decision level

3.2 多源时频信息特征级融合泄露检测

多源时频信息决策级融合泄露检测方法是对频率信息分别进行检测, 而未利用同一时间频率信息之间的相关关系。针对这一问题, 本文提出多源时频信息特征级融合泄露检测方法, 其框架如图 2 所示, 该方法利用多元统计分析假设检验方法对能量和电磁所选频率信息同时进行多维检测。相比于单维频率信息的检测, 它可能利用更少的侧信息检测出泄露。以该框架为基础, 非确定性检测利用 Hotelling's T^2-Test 检验方法, 确定性检测利用 Wilk's Lambda-test 检验方法。具体实现细节见算法 2。

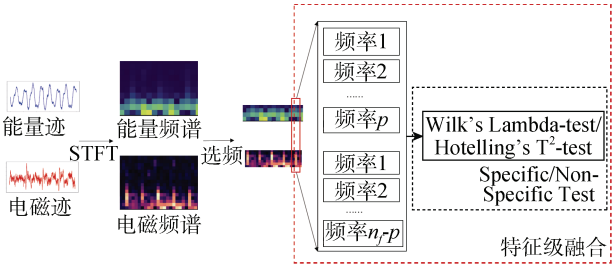

图 2 多源时频信息特征级融合泄露检测框架

Figure 2 The framework of multi-channel time-frequency fusion leakage detection in feature level

算法 2. 多源时频信息特征级融合泄露检测;

输入: 能量迹 $\boldsymbol{L}_{\mathbf{Power}}$ 和电磁迹 $\boldsymbol{L}_{\mathbf{EM}}$;

输出: 是否存在泄露;

1: 对能量迹和电磁迹分别做短时傅里叶变换:

$$\boldsymbol{L}'_{\mathbf{Power}} = \mathrm{STFT}(\boldsymbol{L}_{\mathbf{Power}});$$

$$\boldsymbol{L}'_{\mathbf{EM}} = \mathrm{STFT}(\boldsymbol{L}_{\mathbf{EM}});$$

2: 从 $\boldsymbol{L}'_{\mathbf{Power}}$ 和 $\boldsymbol{L}'_{\mathbf{EM}}$ 中共选择出 n_f 个频率分量:

$$\boldsymbol{L}'_{\mathbf{Power1}} = \mathrm{select}(\boldsymbol{L}'_{\mathbf{Power}});$$

$$\boldsymbol{L}'_{\mathbf{EM1}} = \mathrm{select}(\boldsymbol{L}'_{\mathbf{EM}});$$

3: 分别计算 $\boldsymbol{L}'_{\mathbf{Power1}}$ 和 $\boldsymbol{L}'_{\mathbf{EM1}}$ 的振幅和相位信息:

$$\boldsymbol{P}_{\mathbf{amplitude}} = \mathrm{abs}(\boldsymbol{L}'_{\mathbf{Power1}}); \boldsymbol{P}_{\mathbf{phase}} = \mathrm{angle}(\boldsymbol{L}'_{\mathbf{Power1}});$$

$$\boldsymbol{E}_{\mathbf{amplitude}} = \mathrm{abs}(\boldsymbol{L}'_{\mathbf{EM1}}); \boldsymbol{E}_{\mathbf{phase}} = \mathrm{angle}(\boldsymbol{L}'_{\mathbf{EM1}});$$

4: 对同一时间的振幅信息和相位信息进行组合, 构成 m 组多维频率信息 $\boldsymbol{L}^j_{\mathbf{fusion}}$:

$$\boldsymbol{L}^j_{\mathbf{fusion}} = [\boldsymbol{P}^j_{\mathbf{amplitude}}, \boldsymbol{P}^j_{\mathbf{phase}}, \boldsymbol{E}^j_{\mathbf{amplitude}}, \boldsymbol{E}^j_{\mathbf{phase}}],$$

$$j = 1,2,\ldots,m;$$

5: 依据确定性或非确定性检测场景选择 Wilk's Lambda-test 或 Hotelling's T^2-Test 检验方法, 并对 $\boldsymbol{L}^j_{\mathbf{fusion}}$ 进行多维检测;

6: 如果 mini $p \le \alpha$, 则返回存在泄露, 否则返回未存在泄露。

3.3 多源时频信息数据级融合泄露检测

当密码设计者需要修复泄露时, 定位中间值的信息泄露在时域中的位置是必要的, 基于此需求, 本文提出多源时频信息数据级融合泄露检测方法, 其框架如图 3 所示, 通过 STFT 将能量和电磁信息转换到频域信息进行融合, 再将融合后的频域信息通过短时傅里叶逆变换(Inverse Short Time Fourier Transform, ISTFT)转换成时域信息进行检测。频率信息融合时, 由于能量和电磁信息在频域上的泄露分布不一定相同, 因此该方法对选择后的频率信息进

行哈达玛积操作, 即频率分量元素对应相乘, 从而增大泄露信号和噪声之间的差异。此外, 该方法不仅利用了能量和电磁频域上振幅信息, 也利用了相位信息, 进一步降低检测的数据复杂度。具体实现细节见算法 3。

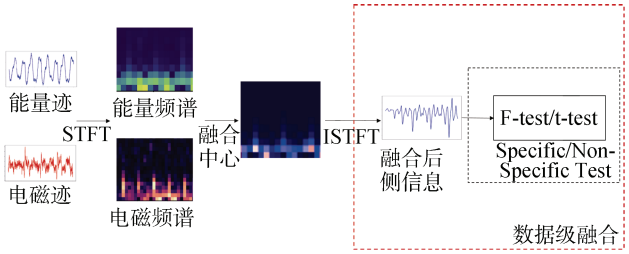

图 3　多源时频信息数据级融合泄露检测框架

Figure 3　The framework of multi-channel time-frequency fusion leakage detection in data level

算法 3.多源时频信息数据级融合泄露检测;

输入: 能量迹 $\boldsymbol{L}_{\mathbf{Power}}$ 和电磁迹 $\boldsymbol{L}_{\mathbf{EM}}$;

输出: 是否存在泄露;

1: 对能量迹和电磁迹分别做短时傅里叶变换:

$$\boldsymbol{L}'_{\mathbf{Power}} = \mathrm{STFT}(\boldsymbol{L}_{\mathbf{Power}});$$

$$\boldsymbol{L}'_{\mathbf{EM}} = \mathrm{STFT}(\boldsymbol{L}_{\mathbf{EM}});$$

2:对 $\boldsymbol{L}'_{\mathbf{Power}}$ 和 $\boldsymbol{L}'_{\mathbf{EM}}$ 进行频率选择:

$$\boldsymbol{L}'_{\mathbf{Power1}}, \boldsymbol{L}'_{\mathbf{Power2}} = \mathrm{select}(\boldsymbol{L}'_{\mathbf{Power}});$$

$$\boldsymbol{L}'_{\mathbf{EM1}}, \boldsymbol{L}'_{\mathbf{EM2}} = \mathrm{select}(\boldsymbol{L}'_{\mathbf{EM}});$$

3: 分别计算 $\boldsymbol{L}'_{\mathbf{Power1}}$ 和 $\boldsymbol{L}'_{\mathbf{EM1}}$ 的哈达玛积及 $\boldsymbol{L}'_{\mathbf{Power2}}$ 和 $\boldsymbol{L}'_{\mathbf{EM2}}$ 的哈达玛积:

$$\boldsymbol{L}'_1 = \boldsymbol{L}'_{\mathbf{Power1}} \circ \boldsymbol{L}'_{\mathbf{EM1}};$$

$$\boldsymbol{L}'_2 = \boldsymbol{L}'_{\mathbf{Power2}} \circ \boldsymbol{L}'_{\mathbf{EM2}};$$

4: 对 $\boldsymbol{L}'_1$ 和 $\boldsymbol{L}'_2$ 进行逆短时傅里叶变换, 获得融合后的侧信息:

$$\boldsymbol{L}_{\mathbf{fusion}} = \mathrm{ISTFT}(\boldsymbol{L}'_1, \boldsymbol{L}'_2);$$

5: 依据确定性或非确定性检测场景选择 F-test 或 t-test 检验方法, 并对 $\boldsymbol{L}_{\mathbf{fusion}}$ 进行检测;

6: 如果 $\mathrm{mini}\, p \le \alpha$, 则返回存在泄露, 否则返回未存在泄露。

3.4　高阶泄露检测

对于无保护密码实现产生的信息泄露, 大部分情况可以通过对比不同敏感数据对应侧信息均值的差异检测到信息泄露, 而掩码型方案实现会消除这种差异。针对掩码型方案实现的泄露检测需要利用与掩码相关的中间值的联合信息泄露, 因此在泄露检测前对侧信息进行预处理。中心乘积[28-29]是一种常用的预处理方法, 即对侧信息上的任意 d 个采样点进行乘积组合, 如公式(7)所示。如果 d 个采样点包含目标敏感数据的 d 个共享因子信息, 则可以检测出泄露。

$$L' = \prod_{i=0}^{d-1} (L_i - \bar{L}_i) \tag{7}$$

对于 n_l 个采样点, 可重复地选取 d 个采样点, 其组合点总数为 n'_l, 其计算方法如公式(8)所示。因为组合点数量迅速增长, 对误报率产生严重影响, 所以 n'_l 为评估掩码型方案误报率的重要参数。

$$n'_l = \begin{pmatrix} n_l + d - 1 \\ \\ d \end{pmatrix} = \frac{(n_l + d - 1)!}{(n_l - 1)! d!} \tag{8}$$

受到上述方法的启发, 对于掩码型防护密码实现信息泄露检测, MCTFF-TVLA-Dt 方法可以对融合后的时域信息进行多点组合, 以检测高阶泄露。而 MCTFF-TVLA-Ft 和 MCTFF-TVLA-Dc 方法则是将任意 d 个单位时间上的所选频率信息进行乘积组合, 以达到检测高阶泄露的效果。

4　实验与结果分析

4.1　模拟实验

MCTFF-TVLA 方法因利用多维频率信息具有降低数据复杂度的优势, 为验证其方法的有效性, 我们通过模拟实验评估该方法的增益效果。这种增益依赖于具体的实验场景, 例如, 并非所选频率信息均存在泄露, 而非泄露点可能会降低检测能力。因此模拟实验通过分析频率信息的泄露点密度 ϕ、维数 n_f、信噪比 SNR 以及频率信息间的相关系数 ρ 四个主要参数与检测出泄露所需侧信息数量 N 的关系, 以评估 MCTFF-TVLA 检测方法的适用场景。同时通过与 TVLA、基于多元 T 检验的多源信息泄露检测方法进行对比, 分析 MCTFF-TVLA 方法所具有的潜在优势。

频率信息的泄露点密度 ϕ 定义如下:

$$\phi = \frac{n_o}{n_f} \tag{9}$$

其中, n_o 表示频率信息中泄露点数量, n_f 表示频率信息的维数, 当泄露点密度 ϕ 为 1 时, 表示所选频率信息均存在泄露, 当泄露点密度为 0 时, 表示所选频率信息不存在泄露。

4.1.1　模拟实验模型及参数设置

实验选择 AES-128 第一轮 S 盒输出作为目标中

间值变量υ，记p表示明文字节，k_c表示正确密钥字节，则中间值变量υ可表示为$S(p\oplus k_c)$。实验选择汉明重量 HW 作为中间值的信息泄露转换函数f，即利用汉明重量模型刻画维数为n_f的能量和电磁频率信息泄露$\boldsymbol{L}$:

$$\boldsymbol{L}=\omega\times\mathrm{HW}(\upsilon)+\mathcal{N}_{0,\boldsymbol{\Sigma}} \tag{10}$$

其中，$\boldsymbol{\omega}=[\omega_i\ \forall i\in[1,\ldots,n_f]]$是一个维数为$n_f$的常量向量，当$i^{\text{th}}$频率信息不存在泄露时，设置$\omega_i=0$，当$i^{\text{th}}$频率信息存在泄露时，$\omega_i$从区间[1,10]随机选取一个整数，因此通过$\boldsymbol{\omega}$可以控制能量和电磁频率信息的泄露点密度$\phi$。

$\mathcal{N}_{0,\boldsymbol{\Sigma}}$表示均值为 0、协方差矩阵为$\boldsymbol{\Sigma}$的多元高斯分布噪声。协方差矩阵$\boldsymbol{\Sigma}$依赖于$i^{\text{th}}$频率信息的噪声方差$\mathrm{var}_{\text{noise}}^{i}$以及$p^{\text{th}}$和$q^{\text{th}}$($p,q\in[1,\ldots,n_f]$)频率信息之间的相关系数$\rho_{p,q}$，因此$\mathrm{var}_{\text{noise}}^{i}$和$\rho_{p,q}$需要被设置。在模拟环境中，第$i$个点的信噪比定义为信号方差与噪声方差之比(见公式(11))。首先可计算出$\omega_i^2\,\mathrm{var}(\mathrm{HW}(\upsilon))$，$\omega_i$是从区间[1,10]随机选择的一个整数，对于 AES-128 算法，目标中间值变量υ是从区间[0,…,255]随机选择的。模拟实验中 SNR 是已知的，因此$\mathrm{var}_{\text{noise}}^{i}$可以被计算得出。在模拟实验中，假设频率信息之间正相关，故$\rho_{p,q}$从区间[0,1]中进行选择，通过$\mathrm{var}_{\text{noise}}^{i}$和$\rho_{p,q}$可以计算出$\boldsymbol{\Sigma}$。从而根据公式(10)模拟出维数为$n_f$、侧信息数量为$N$的能量和电磁频率信息泄露。

$$\mathrm{SNR}=\frac{\omega_i^2\,\mathrm{var}(\mathrm{HW}(\upsilon))}{\mathrm{var}_{\text{noise}}^{i}} \tag{11}$$

在模拟实验中，第一步给定显著水平α值，一般设置$\alpha=10^{-5}$[12,17]，第二步设置漏报率β为 0.1，漏报率β指频率信息存在泄露时，未被检测出泄露的概率，换言之，$1-\beta$表示频率信息存在泄露时，被正确检测出泄露的概率，又称为检测率[17]。根据公式(1)、(3)、(4)、(6)可知，泄露点需要更多的侧信息N达到更高的检测率$1-\beta$。模拟实验重点研究频率信息泄露点密度ϕ、相关系数矩阵ρ、信噪比 SNR 和维数n_f四项参数在漏报率β达到 0.1 时与检测出泄露所需侧信息数量N的关系。针对同一组参数，统计 50 次达到给定漏报率所需侧信息的数量，并取均值作为该组参数的最终结果。

4.1.2 泄露点维数对数据复杂度的影响

在给定频率信息泄露点密度、信噪比和相关系数的情况下，研究频率信息维数对泄露检测所需侧信息数量的影响。首先假设频率信息之间相互独立，在给定泄露点密度ϕ为 1，信噪比 SNR 为 0.01 的情况下，观察确定性检测和非确定性检测的频率信息维数n_f与检测出泄露所需侧信息数量N的关系，如图 4 所示，横坐标表示频率信息维数n_f，纵坐标表示泄露检测所需侧信息的数量N。

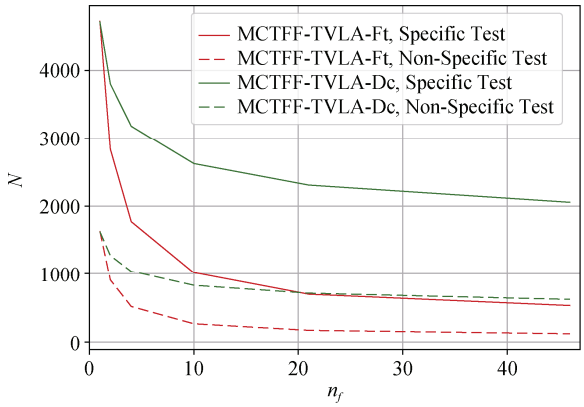

图 4 n_f 和 N 的关系图

Figure 4 The relationship between n_f and N

值得注意的是，在确定性和非确定性检测下，MCTFF-TVLA-Ft 和 MCTFF-TVLA-Dc 方法检测出泄露所需侧信息数量N均随着频率信息维数n_f的增大而逐渐降低，原因在于随着所选频率信息维数n_f的增加，泄露信息不断增加，使得检测出泄露所需侧信息的数量不断减少。相比于 MCTFF-TVLA-Dc 方法，MCTFF-TVLA-Ft 方法的检测能力提升效果更加明显，原因在于 MCTFF-TVLA-Ft 检测方法利用 Wilk's Lambda-test 和 Hotelling's T^2-Test 的多维检测特点，不仅结合了所有频率信息上的差异，而且考虑了频率信息间的相关性，而 MCTFF-TVLA-Dc 检测方法只是利用 F-test 和 t-test 检测频率信息中均值信号差异最大的泄露点，因此在频率信息泄露点密度较高的场景，如无保护 AES 实现产生的信息泄露，MCTFF-TVLA-Ft 方法检测出泄露所需侧信息数量低于 MCTFF-TVLA-Dc 方法所需的侧信息数量。例如，对于本实验中的确定性检测，当频率信息维数n_f为 4 时，MCTFF-TVLA-Ft 方法需要 1773 条侧信息检测出泄露，而 MCTFF-TVLA-Dc 方法需要 3186 条侧信息检测出泄露。当频率信息维数n_f为 10 时，MCTFF-TVLA-Ft 方法需要 1016 条侧信息检测出泄

露, 而 MCTFF-TVLA-Dc 方法需要 2628 条侧信息。

通过该实验可知, 虽然基于多元 T 检验的多源信息泄露检测方法也是进行多维检测, 但它只是对能量和电磁侧信息两个维数的时间采样点进行检测, 无法充分发挥多维检测的优势。

4.1.3 泄露点密度 ϕ 对数据复杂度的影响

假设频率信息之间相互独立, 在给定频率信息维数 n_f 为 10, 信噪比 SNR 均为 0.01 的情况下, 确定性检测和非确定性检测的频率信息泄露点密度 ϕ 与检测出泄露所需侧信息数量 N 的关系如图 5 所示, 横坐标表示频率信息泄露点密度 ϕ, 纵坐标仍表示检测出泄露所需侧信息的数量 N。

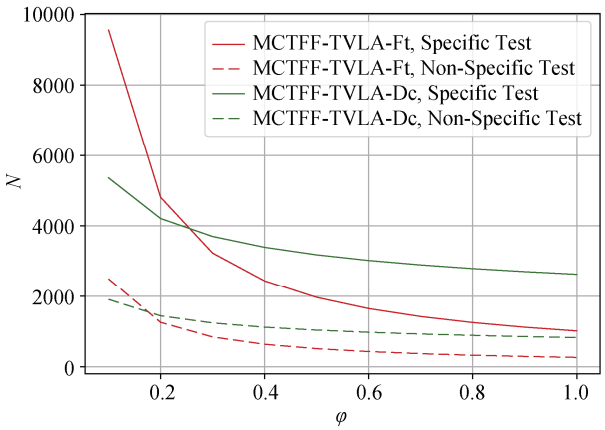

图 5 ϕ 与 N 的关系图

Figure 5 The relationship between ϕ and N

当泄露点密度为 1 时, MCTFF-TVLA-Ft 方法在确定性和非确定性检测下泄露检测所需侧信息数量正如预期, 均小于 MCTFF-TVLA-Dc 方法所需侧信息数量。

随着泄露点密度的降低, 这两种检测方法检测出泄露所需的侧信息的数量逐渐增加。对于基于 Wilk's Lambda-test 的 MCTFF-TVLA-Ft 方法, 降低的泄露点密度增加了组间协方差矩阵 $\boldsymbol{\Sigma}_\mu$ 中 0 元素的个数, 使得 δ 变小, 因此降低了零假设和备择假设下统计量分布的距离, 这需要通过增大侧信息数量 N 达到给定的漏报率 β。同理, 基于 Hotelling's T^2-Test 的 MCTFF-TVLA-Ft 方法通过增大侧信息数量 N 弥补减小的 δ 值, 以达到给定的漏报率 β。对于基于 F-test 和 t-test 的 MCTFF-TVLA-Dc 方法, 降低的泄露点密度 ϕ 使得整体漏报的概率增大, 因此需要通过增加侧信息数量以保持漏报率不变。具体地, 漏报率 β 为所有泄露点均漏报的概率, 如公式(12)所示, 随着泄露点密度 ϕ 降低, 泄露点维数 n_o 减小, 导致漏报率 β 增大。

$$\beta = \prod_{i=1}^{n_o} \beta_i \tag{12}$$

从图 5 中可知, 相比于 MCTFF-TVLA-Dc 方法, 泄露点密度对 MCTFF-TVLA-Ft 方法的影响更大, 以至于在低泄露点密度场景, 如掩码型防护 AES 实现产生的信息泄露, MCTFF-TVLA-Dc 方法检测出泄漏所需的信息数量低于 MCTFF-TVLA-Ft 方法。

4.1.4 信噪比对数据复杂度的影响

在频率信息之间相互独立的条件下, 设置频率信息泄露点密度 ϕ 为 1, 维数 n_f 为 10, 确定性检测和非确定性检测的频率信息信噪比 SNR 与检测出泄露所需侧信息数量 N 的关系如图 6 所示, 横坐标表示频率信息信噪比, 纵坐标表示检测出泄露所需侧信息的数量 N。

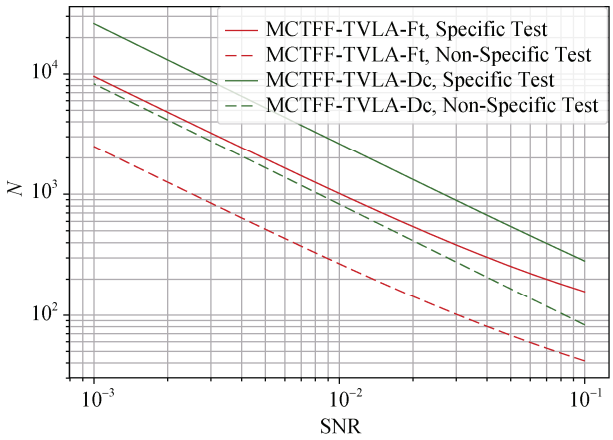

图 6 SNR 和 N 的关系图

Figure 6 The relationship between SNR and N

正如 4.1.3 节中讨论, 在相同信噪比 SNR、高泄露点密度场景下, MCTFF-TVLA-Ft 方法检测出泄露所需侧信息的数量低于 MCTFF-TVLA-Dc 方法。

值得注意的是, 随着频率信息信噪比的增加, MCTFF-TVLA-Ft 和 MCTFF-TVLA-Dc 方法在确定性和非确定性检测下泄露检测所需的侧信息的数量均在减少。例如, 在确定性检测下, 当信噪比为 0.01 时, MCTFF-TVLA-Ft 方法检测出泄露需要 908 条侧信息, 而当信噪比为 0.1 时, MCTFF-TVLA-Ft 方法检测出泄露仅需要 157 条侧信息。高信噪比的侧信息更有助于降低检测出泄露需要的数据复杂度。因此, 相比于仅在时域信息上进行泄露检测的 TVLA 和基于多元 T 检验的多源信息泄露检测方法, MCTFF-TVLA 方法利用了较高信噪比的低频信息, 具有更强的信息泄露检测能力。

4.1.5 泄露点间相关性对数据复杂度的影响

为分析频率信息间相关系数对泄露检测所需侧

信息数量N的影响, 实验设置4种不同相关系数矩阵$\boldsymbol{\Sigma}_i, i=1,2,3,4$, 如图 7 所示, 相关系数矩阵$\boldsymbol{\Sigma}_1$表示相邻频率信息变量之间的相关系数为 0, 即所选频率信息之间是相互独立的, 而相关系数矩阵$\boldsymbol{\Sigma}_2$、$\boldsymbol{\Sigma}_3$和$\boldsymbol{\Sigma}_4$表示相邻频率信息变量之间具有相关性, 且相关性的大小和范围按比例依次增大。其他参数包括频率信息泄露点密度ϕ、维数n_f、信噪比 SNR 分别设置为 1、10 和 0.01。针对不同的相关系数矩阵, MCTFF-TVLA-Ft 和 MCTFF-TVLA-Dc 方法在确定性和非确定性检测下检测出泄露所需的侧信息数量N如表 2 所示。

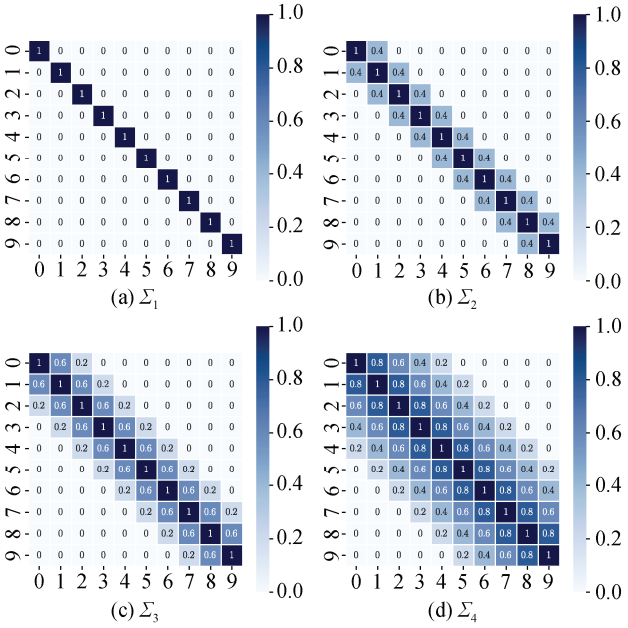

图 7　维数 n_f 为 10 的相关系数矩阵

Figure 7　Correlation matrices for n_f=10

表 2　$\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, \boldsymbol{\Sigma}_3, \boldsymbol{\Sigma}_4$ 与 N 的关系

Table 2　The relationship between $\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, \boldsymbol{\Sigma}_3, \boldsymbol{\Sigma}_4$ and N

检测场景	泄露检测方法	$\boldsymbol{\Sigma}_1$	$\boldsymbol{\Sigma}_2$	$\boldsymbol{\Sigma}_3$	$\boldsymbol{\Sigma}_4$
Specific Test	MCTFF-TVLA-Ft	1016	1669	2238	3547
	MCTFF-TVLA-Dc	2628	2628	2628	2628
Non-Specific Test	MCTFF-TVLA-Ft	268	438	587	928
	MCTFF-TVLA-Dc	829	829	829	829

值得注意的是, 在不同相关系数矩阵下, MCTFF-TVLA-Ft 方法检测出泄露所需的侧信息数量是不同的, 而 MCTFF-TVLA-Dc 方法检测出泄露所需侧信息数量均相同。原因在于 MCTFF-TVLA-Ft 方法利用 Wilk's Lambda-test 和 Hotelling's T^2-Test 额外考虑频率信息间的相关系数, 而 MCTFF-TVLA-Dc 方法利用 F-test 和 t-test 对各个频率信息分别进行单维检测。

正如预期, 当频率信息间相互独立时, 即相关系数矩阵为Σ_1时, MCTFF-TVLA-Ft 方法检测出泄露所需的侧信息数量低于 MCTFF-TVLA-Dc 方法检测出泄露所需的侧信息数量, 而对于相关性矩阵$\boldsymbol{\Sigma}_i, i=1,2,3,4$, MCTFF-TVLA-Ft 方法检测出泄露所需侧信息数量 N 依次增多, 例如, 在确定性检测下, 当相关系数矩阵为Σ_1时, MCTFF-TVLA-Ft 方法需要 1016 条侧信息检测出泄露, 而当相关系数矩阵为Σ_4时, MCTFF-TVLA-Ft 方法需要 3547 条侧信息检测出泄露, 以至于高于 MCTFF-TVLA-Dc 方法检测出泄露所需侧信息数量。原因在于当频率信息相互独立时, 各频率均提供不同的信息, 随着频率信息间的相关性增大, 各频率提供的信息相似度提高, 导致侧信息中的信息含量降低。因此, 在泄露检测时, 尽量选择相关系数较低的频率信息。

4.1.6　窗口函数对频率信息间相关系数的影响

通过分析窗口函数对频率信息间相关系数的影响, 选择适用于侧信息 STFT 变换的窗口函数, 以提高信息泄露的检测能力。

实验选择 Rectangular、Hanning、Hamming、Blackman-Harris 4 种常用的窗口函数分别对侧信息进行 STFT 变换, 并计算低频信息间相关系数, 这 4 种窗口函数生成频率信息之间的相关系数区间分别为[0,0.43]、[0.03,0.96]、[0,0.96]、[0.12,0.97]。相比于 Hanning、Hamming、Blackman-Harris 窗口函数, 由 Rectangular 窗口函数生成的频率信息间相关系数的取值范围更小, 有助于提高MCTFF-TVLA-Ft方法的检测能力(如 4.1.5 节讨论)。因此在实际实验中采用 Rectangular 窗口函数对侧信息进行 STFT 变换。

4.1.7　多源融合泄露检测对误报率的影响

在 TVLA 步骤中, 当最小的 p-value 小于显著水平α时, 则认为 DUT 存在泄露。一般设置侧信息上每个采样点检测的显著水平α为 0.00001, 换言之, 侧信息上每个采样点的误报率为 0.00001。可是整体检测的误报率α_{overall}随着采样点数量的增多而增大, 如图 8 (a)所示, 当采样点数量n_l为 100000 时, TVLA 的误报率α_{overall}达到 0.63212。

为解决上述问题, Šidák 校正是一种可选方法, 它是根据采样点数量n_l对各采样点的误报率α进行调整, 以保持整体误报率α_{overall}不变, 如公式(13)所示。

$$\alpha = 1-(1-\alpha_{\text{overall}})^{1/n_l} \tag{13}$$

由上式可知误报率α随着信息采样点数量n_l的增多而降低。在其他条件不变时, F-test 和 t-test 的漏

报率 β 均与误报率 α 成反比(依据公式(1)、(4), 因此, 经过 Šidák 校正, TVLA 的漏报率 β 随着采样点数量的增多而升高, 如图 8 (b)、(c)所示。例如, 对于确定性检测, 设置 SNR、HW( )、N 分为 0.01、3 和 300, 当采样点数量 n_l 是 1 时, TVLA 的漏报率 β 为 0.18, 采样点数量 n_l 为 10000 时, TVLA 的漏报率 β 达到 0.79。

上述问题是由于对大量的采样点进行多次检测引起的, 因此, 对多信道侧信息进行 TVLA 泄露检测, 将进一步增加采样点的数量, 使问题更加严重。而多源融合泄露检测消除了由多信道侧信息带来采样点数量增加的缺点。假设能量和电磁侧信息采样点数量相同, 那么 MCTFF-TVLA-Da 可将检测次数降为总采样点数量的一半。此时, 误报率最多降低 0.25, 如图 8(a)中虚线所示。然而当总采样点数量较多时, MCTFF-TVLA-Da 依然具有较大的误报率, 如总采样点数量为 5×10^5 时, 其误报率 α_{overall} 为 0.91792。文献[22]方法对能量和电磁信息泄露融合后的时域信息进行检测, 且采用 mini-p 策略, 因此其误报率与 MCTFF-TVLA-Da 相同。

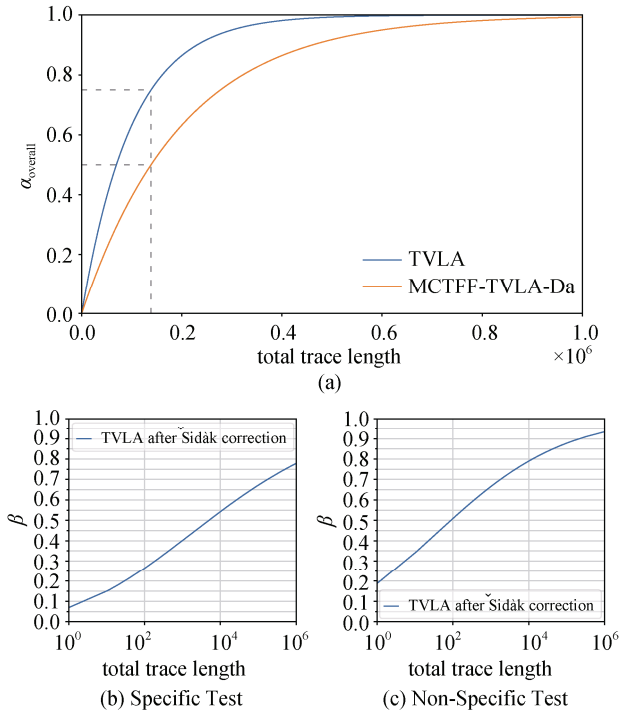

图 8 n_l 对 α_{overall} 或 β 的影响

Figure 8 The influence of n_l on α_{overall} or β

由于 MCTFF-TVLA-Ft 和 MCTFF-TVLA-Dc 方法可将 STFT 窗口长度内的误报率 α 保持不变, 因此可进一步降低整体检测误报率 α_{overall}。模拟实验无法确定窗口长度, 故在真实实验评估二者的误报率。

4.2 真实实验

实验主要研究无保护 AES 实现和一阶掩码型防护 AES 实现的信息泄露检测, 并执行了无保护 AES 硬件实现、掩码型防护 AES 硬件实现、掩码型防护 AES 软件实现三种实际实验, 通过检测出泄露所需数据复杂度和误报率评估 MCTFF-TVLA 方法的检测能力和有效性。同时与其他方法进行对比, 具体有 TVLA-Power(对能量信息泄露进行 TVLA 检测)、TVLA-EM(对电磁信息泄露进行 TVLA 检测)、MCTVLA(基于多元 T 检验的多源信息泄露检测方法)、SFTVLA-SUM(简单融合信息泄露检测方法)以及 TFFTVLA(多源时频融合信息泄露检测方法)。

4.2.1 无保护 AES 硬件实现

检测的目标设备是嵌有 Kintex-7 FPGA 芯片的 SAKURA-X 开发板, 其运行频率为 20MHz。利用 Picoscope 5000 示波器以 5GSa/s 的采样率对 AES-128 加密算法运行时的能量迹和电磁迹进行同时采集, 其中, 能量迹是通过 SAKURA-X 开发板进行采集的, 电磁迹是通过型号为 RS-H 50-1 的近场电磁探针垂直放在 FPGA 芯片上采集的。能量迹和电磁迹的采样点数量均为 3500。为减少噪声, 重复采集同一组数据的侧信息 20 次并取均值。

首先通过分析窗口长度对频率信息信噪比 SNR 的影响, 选择适用于 STFT 的窗口长度, 以提高 MCTFF-TVLA 方法的检测能力。实验利用最后一轮 S 盒输入和相应密文的汉明距离计算频率信息的信噪比 SNR, 选择窗口步长为窗口长度的 1/2。图 9 显示能量和电磁侧信息 STFT 所用不同窗口长度对应频率信息信噪比 SNR。当窗口长度为 300 时, 关于能量和电磁频率信息的 SNR 最大。如 4.1.4 节讨论, SNR 越高检测出泄露所需侧信息数量越少, 高 SNR 有助于提高检测能力。因此, 在硬件实现泄露检测场景下, STFT 窗口长度设置为 300。

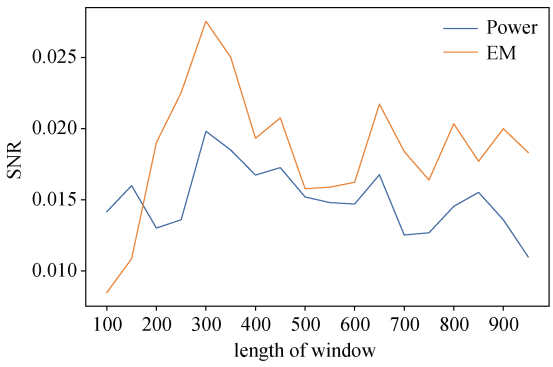

图 9 STFT 窗口长度与信噪比 SNR 的关系

Figure 9 The relationship between the window length of STFT and SNR

确定性检测结果如图 10(a)所示, 所有的检测方法均成功检测出泄露, 即 p-value 超过了阈值(红色虚线)。值得注意的是, MCTFF-TVLA-Ft 方法需要 3050 条侧信息检测出泄露, 而 MCTFF-TVLA-Dc 方法需要 4000 条侧信息, 相比于 MCTFF-TVLA-Dc 方法, MCTFF-TVLA-Ft 方法具有更高的检测效率, 该结果与模拟实验 4.1.3 节中频率信息具有较高泄露点密度的检测结果是一致的, 具体地, 在选频环节, 实验选择了 8 维频率信息, 通过 F-test 对其进行检测, 其中有 6 维频率信息达到阈值, 意味着频率信息泄露点密度为 0.75, 较高的泄露点密度使得 MCTFF-TVLA-Ft 方法具有更强的检测能力。在图10(a)中, MCTFF-TVLA-Ft 和 MCTFF-TVLA-Dc 的检测能力均高于其他四种检测方法, 原因在于 MCTFF-TVLA-Ft 和 MCTFF-TVLA-Dc 方法检测具有较高信噪比的低频信息, 故这两种方法的 p-value 先于其他方法到达阈值。

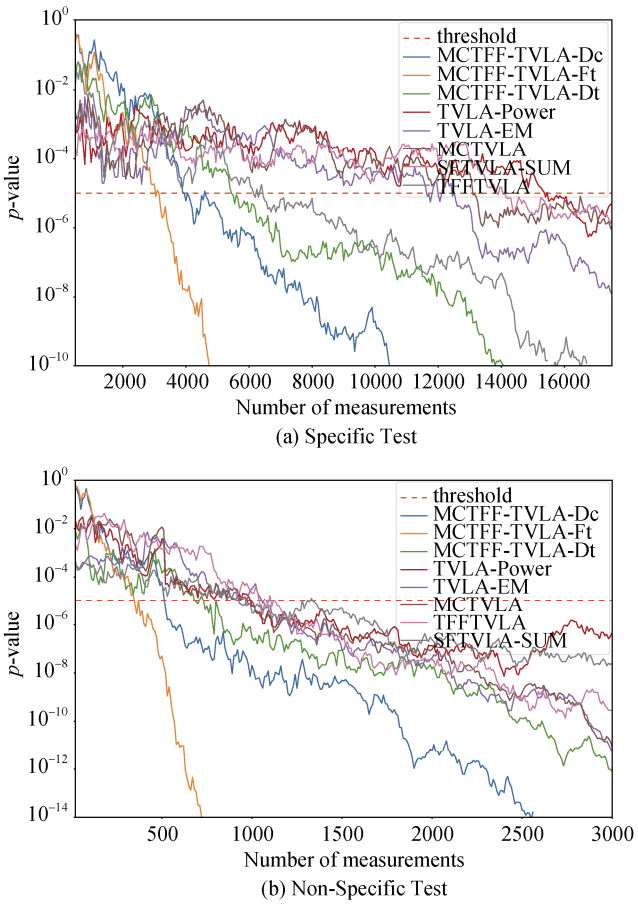

图 10 无保护 AES 硬件实现的泄露检测结果对比

Figure 10 The comparison of leakage detection result of unprotected AES hardware implementation

在图 10(a)中, MCTFF-TVLA-Dt 需要 5500 条侧信息检测出泄露, 而 TFFTVLA 需要 6300 条, 相比于 TFFTVLA, MCTFF-TVLA-Dt 进一步提升了检测效率, 验证了 MCTFF-TVLA-Dt 方法有助于提高时域信息的信噪比。TVLA-Power、TVLA-EM 方法分别需要 15300、11800 条侧信息检测出泄露, 而 MCTVLA 检测方法需要 13100 条侧信息, 其 p-value 并未先于 TVLA-EM 方法到达阈值, 其原因在于 MCTVLA 方法只检测能量和电磁泄露相同时刻的采样点, 即同时检测两个维数的泄露, 如图 4 所示, 检测能力提升不明显, 而且同一时刻的能量和电磁信息不一定同时存在泄露, 故 MCTVLA 方法无法发挥多维检测的优势。SFTVLA-SUM 方法检测泄露所需侧信息数量处于 TVLA-Power、TVLA-EM 方法之间, 这是由于能量和电磁信息泄露的信噪比存在差异造成的[22]。

非确定性检测结果如图 10 (b)所示, 各方法的检测结果对比与确定性检测相似, MCTFF-TVLA-Ft、MCTFF-TVLA-Dc、MCTFF-TVLA-Dt 分别需要 360、520、700 条侧信息检测出泄露。而 TVLA-Power、TVLA-EM、MCTVLA、SFTVLA-SUM、TFFTVLA 方法均需要约 1000 条侧信息检测出泄露, 验证了本文所提方法在非确定性检测下的有效性。

对于同一种检测方法, 相比于多划分的确定性检测, 二划分的非确定性检测的数据复杂度更低, 例如, MCTFF-TVLA-Ft 方法在确定性检测下需要 3050 条侧信息检测出泄露, 而在非确定性检测下仅需 360 条侧信息, 验证了相比于多划分的确定性检测, 二划分的非确定性检测更易检测到泄露。

对于无保护 AES 硬件实现, 相比于现有的检测方法, MCTFF-TVLA 方法在确定性和非确定性检测下泄露检测所需侧信息数量分别降低 52%和 64%(依据公式(14)), 有效提高了泄露检测能力。TVLA 方法对能量和电磁信息进行检测的误报率 α_{overall} 达到 0.06761, TFFTVLA、MCTFF-TVLA-Dt、MCTVLA、SFTVLA-SUM 方法误报率 α_{overall} 为 0.03439, 而 MCTFF-TVLA-Ft、MCTFF-TVLA-Dc 方法可将误报率 α_{overall} 降低至 0.00023。相比于现有的检测方法, MCTFF-TVLA 方法的误报率降低 99.33%。

$$P_{\text{reduced}} = 1-\frac{\min\left(N(\text{MCTFF-TVLA-Dc/Ft/Dt})\right)}{\min\left(N(\text{TFFTVLA,MCTVLA,TVLA-EM/Power})\right)} \tag{14}$$

4.2.2 掩码型防护 AES 硬件实现

检测的目标设备仍是嵌有 Kintex-7 FPGA 芯片的 SAKURA-X 开发板, 其运行频率为 20MHz。采集

掩码型防护 AES 算法最后三轮运行中的能量和电磁信息泄露, 其采样点数量均为 1000 点。采集配置与无保护 AES 硬件实现相同。此时, STFT 窗口长度仍设置为 300。

硬件掩码型防护策略按照文献[30]提出的方案 DPA Contest v4.2 进行实现。因为 AES 算法受到两个共享因子的一阶掩码型防护, 所以在泄露检测前, 需要对侧信息进行两两组合, 以检测设备的二阶泄露, 经过组合后的侧信息共有 500500 个组合点(依据公式 8)。

在确定性(50000 条侧信息)和非确定性(7000 条侧信息)检测下, MCTFF-TVLA-Dt 和 TFFTVLA 方法均无法检测到泄露, 可能由于中心乘积的组合方式不适合这两种检测方法。其他检测方法均成功地检测出泄露。尽管两两组合后的侧信息可以降低掩码的影响产生二阶泄露, 但其信噪比低于无保护 AES 硬件实现一阶泄露的信噪比, 故同一检测方法在掩码型防护 AES 硬件实现场景下所需的侧信息数量高于在无保护 AES 硬件实现场景下泄露检测所需的侧信息数量。

确定性检测结果如图 11(a)所示, 值得注意的是, MCTFF-TVLA-Ft 方法需要 26500 条侧信息检测出泄露, 而 MCTFF-TVLA-Dc 方法需要 25000 条侧信息, 相比于 MCTFF-TVLA-Ft, MCTFF-TVLA-Dc 方法具有更高的检测效率, 该结果与模拟实验中频率信息具有较低泄露点密度的检测结果是一致的, 具体地, 在选频环节, 实验选择了 4 维频率信息, 经过两两组合后, 同一单位时间的频率信息共有 16 维, 通过 F-test 对其进行检测, 其中有 2 维频率信息达到阈值, 意味着频率信息泄露点密度为 0.125, 较低的泄露点密度使得 MCTFF-TVLA-Dc 方法具有更强的检测能力。MCTFF-TVLA-Ft 和 MCTFF-TVLA-Dc 方法的检测效率仍高于其他方法, 表明所选频率信息二阶泄露信噪比高于时域信息二阶泄露信噪比。

掩码型防护 AES 硬件实现非确定性检测结果如图 11 (b)所示, MCTFF-TVLA-Ft、MCTFF-TVLA-Dc 方法检测出泄露所需侧信息的数量是近似的, 均需 900 条侧信息检测出泄露, 表明所选频率信息依然受到低泄露点密度的影响。

对于掩码型防护 AES 硬件实现, 相比于现有泄露检测方法, MCTFF-TVLA 方法在确定性和非确定性检测下泄露检测所需侧信息数量分别降低了 15%和 31%(依据公式(14)), 有效提高了泄露检测效率。TVLA 方法对能量和电磁信息泄露进行检测的误报率 α_{overall} 达到 0.99995, MCTVLA 和 SFTVLA-SUM 方法误报率 α_{overall} 为 0.99329, 而 MCTFF-TVLA-Ft、MCTFF-TVLA-Dc 的误报率 α_{overall} 可降至 0.00027, 相比于现有检测方法, MCTFF-TVLA 方法误报率降低了 99.97%, 有效控制了整体检测误报率。

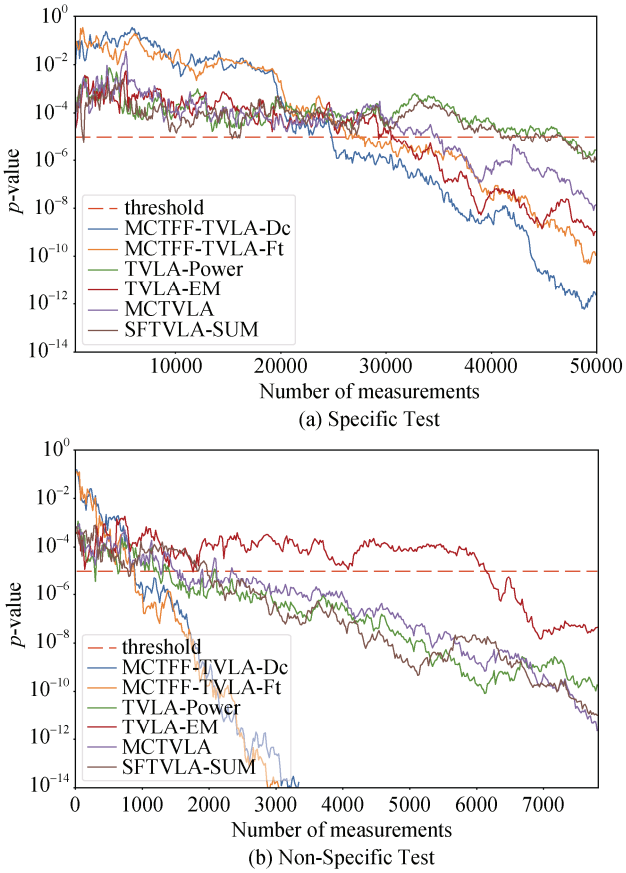

图 11 掩码型保护 AES 硬件实现的泄露检测结果对比

Figure 11 The comparison of leakage detection result of masked AES hardware implementation

4.2.3 掩码型防护 AES 软件实现

检测的目标设备为 ATmega163 智能卡。实验选择 AES-128 RSM [30](DPA Contest v4.2)掩码型方案, 其实现源码由 DPA Contest v4[31]提供。Picoscope 5000 示波器以 1GSa/s 的采样率采集 AES-128 加密算法第一轮运行时的能量和电磁信息泄露, 其采样点数量均为 20 万个。为提高检测效率, 对能量和电磁侧信息执行步长为 10 的压缩, 得到采样点数量均为 2 万的能量和电磁侧信息, 其中, 能量迹是通过 SASEBO-W 开发板采集的, 电磁迹是通过型号为 RS-H 400-1 电磁探针垂直放在智能卡芯片上采集的。

首先确定有效的 STFT 窗口长度, 利用掩码型 S 盒输出的汉明重量计算频率信息信噪比 SNR, 窗口步长仍为窗口长度的 1/2。STFT 所用不同窗口长度对应的能量和电磁频率信息 SNR 如图 12 所示, 窗口

长度对电磁频率信息的 SNR 影响较小, 而当窗口长度为200时, 能量频率信息泄露的SNR最大, 因此窗口长度设置为 200。

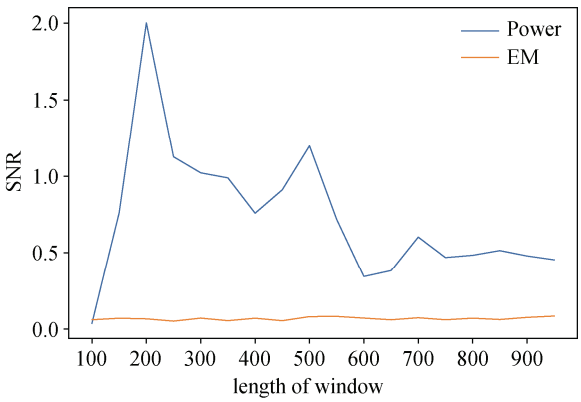

图 12 STFT 窗口长度与信噪比 SNR 的关系
Figure 12 The relationship between the window length of STFT and SNR

掩码型防护 AES 软件实现的确定性检测和非确定性检测如图 13 (a)、(b)所示, 各检测方法对比结果与掩码型保护 AES 硬件实现相似, 值得注意的是, 在确定性检测和非确定性检测下, MCTFF-TVLA-Dc 方法的检测效率均高于 MCTFF-TVLA-Ft, 该检测结果与模拟实验中具有较低泄露点密度的检测结果是一致的。实验结果表明 MCTFF-TVLA-Dc 方法更适合检测掩码型防护 AES 硬件和软件实现的信息泄露, 而 MCTFF-TVLA-Ft 方法更适合无保护 AES 硬件实现的泄露检测。

对于掩码型防护 AES 软件实现, 相比于现有的检测方法, MCTFF-TVLA 方法在确定性和非确定性检测下泄露检测所需侧信息数量分别降低了 23%和 29%(依据公式(14)), 验证了该方法的有效性。TVLA 方法对能量和电磁信息进行检测误报率 α_{overall} 达到 0.99326, MCTVLA 和 SFTVLA-SUM 的误报率 α_{overall} 为 0.91791, 而 MCTFF-TVLA-Ft 和 MCTFF- TVLA-Dc 方法的检测误报率 α_{overall} 可降至 0.00025, 相比于现有的检测方法, MCTFF-TVLA 方法的误报率降低了 99.97%, 有效控制了泄露检测误报率。

5 结论

本文提出了基于时频特征的多源融合信息泄露检测方法, 有效融合利用能量和电磁信息泄露。本文的贡献在于基于时频特征的多源融合信息泄露检测方法有效降低了误报率和检测出泄露所需要的侧信息数量, 提高了检测能力。评估者可以依据实际情况选择更有效的检测方法, 对于无保护 AES 实现的信息泄露, 傅里叶变换可将能量和电磁泄露集中在所选的频率信息上, 故MCTFF-TVLA-Ft方法在该场景下具有更好的检测能力。而掩码型防护 AES 实现的信息泄露检测, MCTFF-TVLA-Dc 方法具有更好的检测能力, 原因在于多点组合稀释了能量和电磁的泄露点。由于 MCTFF-TVLA-Dt 方法对侧信息的时域信息进行泄露检测, 更适用于定位信息泄露在时域中的位置, 但是该方法在掩码型防护实现的场景下未能检测到泄露, 需要进一步确定原因。

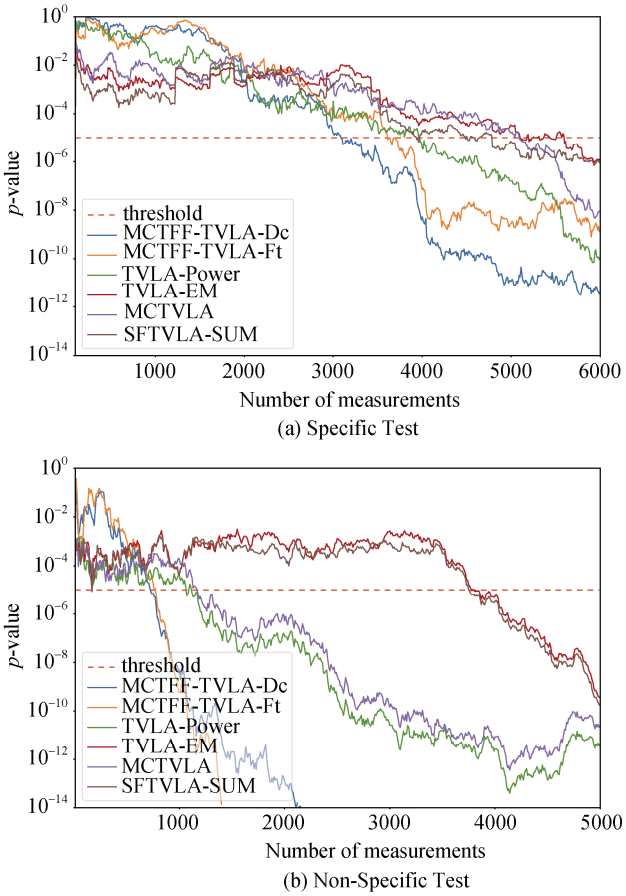

图 13 掩码型保护 AES 软件实现的泄露检测结果对比
Figure 13 The comparison of leakage detection result of masked AES software implementation

下一步工作将进一步探索基于时频特征的多源融合信息泄露检测方法在其他有保护型密码实现中的应用(例如内嵌随机延迟[32-33]等防护措施)并设计出更加通用的多源融合泄露检测方法。

为检测 AES-128 RSM 掩码型方案软件实现的二阶泄露, 对能量和电磁侧信息进行两两组合, 致使侧信息组合点数超过 10^8 个(依据公式(8)), 超出我们的磁盘容量, 因此通过 F-test 选取 500 个掩码和受掩码保护值的特征点进行组合, 得到 25 万个组合点的侧信息。

参考文献

[1] Mangard S, Oswald E, Popp T. *Power analysis attacks: revealing the secrets of smart cards*[M]. New York: Springer, 2007.

[2] Ocher P C. Timing Attacks on Implementations of Diffie-Hellman, RSA, DSS, and other Systems[M]. *Advances in Cryptology — CRYPTO '96*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1996: 104-113.

[3] Kocher P, Jaffe J, Jun B. Differential Power Analysis[M]. *Advances in Cryptology — CRYPTO' 99*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999: 388-397.

[4] Quisquater J J, Samyde D. ElectroMagnetic Analysis (EMA): Measures and Counter-Measures for Smart Cards[M]. *Smart Card Programming and Security*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2001: 200-210.

[5] Camurati G, Francillon A, Standaert F X. Understanding Screaming Channels: From a Detailed Analysis to Improved Attacks[J]. *IACR Transactions on Cryptographic Hardware and Embedded Systems*, 2020: 358-401.

[6] Standaert F X, Malkin T G, Yung M. A Unified Framework for the Analysis of Side-Channel Key Recovery Attacks[C]. *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, 2009: 443-461.

[7] Whitnall C, Oswald E. A Comprehensive Evaluation of Mutual Information Analysis Using a Fair Evaluation Framework[C]. *Annual Cryptology Conference*, 2011: 316-334.

[8] The Common Criteria. https://www.commoncriteriaportal.org/. Sep. 2016.

[9] ISO/IEC 17825: Information technology – Security techniques – Testing methods for the mitigation of non-invasive attack classes against cryptographic modules. *International Organization for Standardization, Geneva, CH*, 2016.

[10] Gilbert Goodwill B J, Jaffe J, Rohatgi P. A testing methodology for side-channel resistance validation[C]. *NIST Non-Invasive Attack Testing Workshop*. 2011, 7: 115-136.

[11] Cooper J, DeMulder E, Goodwill G, et al. Test Vector Leakage Assessment (TVLA) methodology in practice[C]. *International Cryptographic Module Conference*. 2013, 20.

[12] Moradi A, Richter B, Schneider T, et al. Leakage Detection with the X2-Test[J]. *IACR Transactions on Cryptographic Hardware and Embedded Systems*, 2018: 209-237.

[13] Chatzikokolakis K, Chothia T, Guha A. Statistical Measurement of Information Leakage[C]. *International Conference on Tools and Algorithms for the Construction and Analysis of Systems*, 2010: 390-404.

[14] Chothia T, Guha A. A Statistical Test for Information Leaks Using Continuous Mutual Information[C]. *2011 IEEE 24th Computer Security Foundations Symposium*, 2011: 177-190.

[15] Mather L, Oswald E, Bandenburg J, et al. Does my Device Leak Information? an a Priori Statistical Power Analysis of Leakage Detection Tests[C]. *International Conference on the Theory and Application of Cryptology and Information Security*, 2013: 486-505.

[16] Ding A A, Zhang L W, Durvaux F, et al. Towards Sound and Optimal Leakage Detection Procedure[C]. *International Conference on Smart Card Research and Advanced Applications*, 2018: 105-122.

[17] Bronchain O, Schneider T, Standaert F X. Multi-Tuple Leakage Detection and the Dependent Signal Issue[J]. *IACR Transactions on Cryptographic Hardware and Embedded Systems*, 2019: 318-345.

[18] Bhasin S, Danger J L, Guilley S, et al. Side-Channel Leakage and Trace Compression Using Normalized Inter-Class Variance[C]. *The Third Workshop on Hardware and Architectural Support for Security and Privacy*, 2014: 1-9.

[19] Durvaux F, Standaert F X. From Improved Leakage Detection to the Detection of Points of Interests in Leakage Traces[C]. *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, 2016: 240-262.

[20] Šidák Z. Rectangular Confidence Regions for the Means of Multivariate Normal Distributions[J]. *Journal of the American Statistical Association*, 1967, 62(318): 626-633.

[21] Fournier J J A, Moore S, Li H Y, et al. Security Evaluation of Asynchronous Circuits[C]. *International Workshop on Cryptographic Hardware and Embedded Systems*, 2003: 137-151.

[22] Cao Y C, Zhou Y B. Multi-Channel Fusion Leakage Detection[J]. *Journal of Cyber Security*, 2020, 5(6): 40-52.
(曹雨晨, 周永彬. 多源融合信息泄露检测方法[J]. *信息安全学报*, 2020, 5(6): 40-52.)

[23] Van de Vegte J. *Fundamentals of digital signal processing*[M]. Upper Saddle River, NJ: Prentice Hall, 2001.

[24] Faul F, Erdfelder E, Lang A G, et al. G*Power 3: A Flexible Statistical Power Analysis Program for the Social, Behavioral, and Biomedical Sciences[J]. *Behavior Research Methods*, 2007, 39(2): 175-191.

[25] Steyn H S Jr, Ellis S M. Estimating an Effect Size in One-Way Multivariate Analysis of Variance (MANOVA)[J]. *Multivariate Behavioral Research*, 2009, 44(1): 106-129.

[26] Barenghi A, Pelosi G, Teglia Y. Improving First Order Differential Power Attacks through Digital Signal Processing[C]. *The 3rd international conference on Security of information and networks*, 2010: 124-133.

[27] Cao Y C, Zhou Y B, Zhang H L. Multi-Channel Time-Frequency Fusion Attacks[J]. *International Journal of Information and Computer Security*, 2021, 16(1/2): 84.

[28] Prouff E, Rivain M, Bevan R. Statistical Analysis of Second Order Differential Power Analysis[J]. *IEEE Transactions on Computers*, 2009, 58(6): 799-811.

[29] Ding A A, Zhang L W, Fei Y S, et al. A Statistical Model for Higher Order DPA on Masked Devices[C]. *International Workshop on Cryptographic Hardware and Embedded Systems*, 2014: 147-169.

[30] Bhasin S, Bruneau N, Danger J L, et al. Analysis and Improvements of the DPA Contest V4 Implementation[C]. *International Conference on Security, Privacy, and Applied Cryptography Engineering*, 2014: 201-218.

[31] DPA Contest v4. http://www.dpacontest.org/v4/index.php. Jul. 2015.

[32] Coron J S, Kizhvatov I. An Efficient Method for Random Delay Generation in Embedded Software[C]. *International Workshop on*

Cryptographic Hardware and Embedded Systems, 2009: 156-170.

[33] Coron J S, Kizhvatov I. Analysis and Improvement of the Random Delay Countermeasure of CHES 2009[C]. *International Workshop on Cryptographic Hardware and Embedded Systems*, 2010: 95-109.

冯祺 于 2013 年在吉林大学计算机科学与技术专业获得学士学位。现在中国科学院信息工程研究所攻读硕士学位。研究领域为硬件安全与应用密码学。研究兴趣包括: 侧信道攻击方法、安全性检测等。Email: fengqi@iie.ac.cn

明经典 于 2016 年在北京交通大学通信工程专业获得学士学位。现在中国科学院信息工程研究所攻读博士学位。研究领域为硬件安全与应用密码学。研究兴趣包括: 侧信道攻击与防御、安全性评估等。Email: mingjingdian@iie.ac.cn

周永彬 于 2004 年 3 月获得计算机应用技术专业博士学位。现任中国科学院信息工程研究所研究员。研究领域为网络与信息安全理论及技术。主要研究兴趣包括: 密码学、密码工程、网络系统安全、数据安全与隐私保护等。Email: zhouyongbin@iie.ac.cn

张倩 于 2014 年在清华大学微电子与纳电子学系获得工程硕士学位。现在中国科学院信息工程研究所攻读博士学位, 现任中国科学院信息工程研究所助理研究员。研究领域为硬件安全与应用密码学。研究兴趣包括: 密码硬件设计分析、侧信道分析方法与防御机制等。Email: zhangqian@iie.ac.cn