强物理不可克隆函数的侧信道混合攻击

刘 威,蒋烈辉,常 瑞
(战略支援部队信息工程大学网络空间安全学院,河南郑州 450002)

摘 要: 物理不可克隆函数(Physical Unclonable Function,PUF)凭借其固有的防篡改、轻量级等特性,在资源受限的物联网安全领域拥有广阔的应用前景,其自身的安全问题也日益受到关注. 多数强 PUF 可通过机器学习方法建模,抗机器学习的非线性结构 PUF 难以抵御侧信道攻击. 本文在研究强 PUF 建模的基础上,基于统一符号规则分类介绍了现有的强 PUF 侧信道攻击方法如可靠性分析、功耗分析和故障注入等,重点论述了各类侧信道/机器学习混合攻击方法的原理、适用范围和攻击效果,文章最后讨论了 PUF 侧信道攻击面临的困境和宜采取的对策.

关键词: 强物理不可克隆函数; 侧信道混合攻击; 机器学习; 可靠性; 功耗分析; 故障注入
中图分类号: TP309 **文献标识码:** A **文章编号:** 0372-2112 (2019)12-2639-08
电子学报 URL: http://www.ejournal.org.cn **DOI:** 10.3969/j.issn.0372-2112.2019.12.025

Side-Channel Hybrid Attacks on Strong Physical Unclonable Function

LIU Wei, JIANG Lie-hui, CHANG Rui
(*Institute of Cyberspace Security, Information Engineering University of Strategic Support Force, Zhengzhou, Henan* 450002, *China*)

Abstract: Due to the tamperproof and lightweight nature, physical unclonable function are proposed to provide security with low cost for the internet of things. Security of PUF itself has attracted much more attention. Almost all strong PUF can be modeled using machine learning techniques, while the complicated PUF with non-linear structure, which are resistant to machine learning modeling, are vulnerable to side channel attacks. According to the unified symbol rules, the paper presents existing side channel attack methods on strong PUFs, such as reliability analysis, power analysis and fault injection. The principles, performance and applications of side channel/machine learning hybrid attack methods are elaborated and analyzed. In the end, the temporary predicaments and countermeasures of side channel attack on PUF are discussed.

Key words: strong PUF; side-channel hybrid attack; machine learning; reliability; power analysis; fault injection

1 引言

物理不可克隆函数作为轻量级的安全策略,可用于知识产权保护,轻量级安全认证和密钥生成. 凭借其无需存储密钥、防篡改等特点[1~3],在资源受限的物联网领域获得了广泛关注. 自 2002 年 Gassend 等首次提出硅基 PUF[1]以来,研究者设计了多种类型的 PUF 来利用制造差异产生的熵源. PUF 通常按激励响应对(Challenge Response Pair, CRP)的数量分为弱 PUF (Weak PUF)和强 PUF (Strong PUF)[2]. 强 PUF 拥有大量 CRP,通常与基本组件数量呈指数关系. 常见的强 PUF 包括光学 PUF、忆阻器交叉 PUF[4]和基于仲裁器的 PUF 等,仲裁器 PUF、轻量级安全 PUF[5](Lightweight Secure PUF, LSPUF)、组合 PUF[6]和异或仲裁器 PUF[7](XOR Arbiter PUF,简称异或 PUF)等同属基于仲裁器的 PUF.

尽管在提出之初 PUF 就被定位为无法克隆,但随后的很多实验研究和理论分析均表明 PUF 能够被建模. 攻击者在获取足够的 CRP 后,利用密码分析对组合 PUF 建模的准确率达到 90% 以上[8]. 伴随着人工智能的研究热潮,机器学习方法也被广泛应用于攻击强 PUF. 然而,机器学习方法也存在局限性,当非线性部件数量较多或 PUF 规模较大时,攻击难度呈指数级增长.

侧信道攻击是从密码系统硬件电路中提取信息的有效方法. 与传统方法寻找算法理论的脆弱性不同,侧信道攻击主要利用系统物理实现时带入的脆弱性. 对结构复杂的强 PUF,仅依靠侧信道攻击所获取的信息也很难直接建模,但与机器学习方法结合后,侧信道攻击

收稿日期:2018-10-20;修回日期:2019-06-12;责任编辑:覃怀银
基金项目:国家自然科学基金(No.61871405, No.61802431)

能提供除 CRP 以外的信息,可有效降低机器学习对 PUF 建模时的计算量,缩短攻击时间,因此,机器学习和侧信道结合的混合攻击成为目前强 PUF 攻击的主流方法[9,10].

2 PUF 机器学习建模

2.1 描述符号规则

由于文献中参数定义各不相同,为统一描述模型和攻击方法,对文中常用的符号参数说明如下:

$\boldsymbol{C}$:PUF 激励向量

$\boldsymbol{R}$:PUF 响应,如响应只有 1 位也可用 r 表示

Rel:可靠性指标

Rep:可重复度,由可靠性定义衍生的参数

h:可靠度,由可靠性定义衍生的参数

$\boldsymbol{W}$:延迟向量,由仲裁器 PUF 内部结构参数决定,可唯一描述仲裁器 PUF

$\boldsymbol{\Phi}$:特征向量,由激励向量运算得到,分析中等价于激励向量

ΔD_i:第 i 级输出的两路信号延迟差

ΔD_{PUF}:激励产生的 PUF 延迟差,符合均值为 0 的高斯分布,标准差为 σ_V

$\Delta D_{\mathrm{noise}}$:累积噪声产生的延迟,符合均值为 0 的高斯分布,标准差为 σ_N

Δt_{Arb}:由仲裁器偏差引入的延迟

ε:无穷小量

2.2 典型强 PUF 的数学模型

仲裁器 PUF 采用多路选择器级联结构,如图 1 所示,激励向量 $\boldsymbol{C}$ 的每一位对应控制一级多路器,用于选择平行路径或者交叉路径,输入信号沿两条不同路径传输产生延迟差,最终由仲裁器比较传输时延,形成响应 $\boldsymbol{R}$.

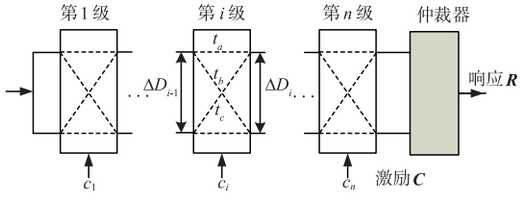

图1 仲裁器PUF电路结构

设第 i 级多路器中信号经平行路径的传输时延记为 t_a 和 t_c,沿交叉路径传输的时延记为 t_b 和 t_d,$\delta_{c_i,i}$是输入为 c_i 时信号通过第 i 级时产生的延迟差.

$$\delta_{c_i,i}=\begin{cases}t_a-t_c, & c_i=0\\ t_b-t_d, & c_i=1\end{cases} \tag{1}$$

$$\Delta D_i=\Delta D_{i-1}(1-2c_i)+\delta_{c_i,i} \tag{2}$$

响应 r 由最终延迟差 ΔD_n 决定,如式(3)所示.

$$r=\begin{cases}1, & \Delta D_n>0\\ 0, & \Delta D_n<0\end{cases} \tag{3}$$

以上分析表明仲裁器 PUF 可由 $2n$ 个延迟差值 δ 来描述,经式(4)运算简化后只需 $n+1$ 个参数,记为延迟向量 $\boldsymbol{W}=(w_1,w_2,\cdots w_{n+1})$. $\boldsymbol{W}$ 由各级延迟参数决定,可唯一描述仲裁器 PUF.

$$w_i=\begin{cases}\delta_{0,1}-\delta_{1,1}, & i=1\\ \delta_{0,i-1}+\delta_{1,i-1}+\delta_{0,i}-\delta_{1,i}, & 1\leq i\leq n-1\\ \delta_{0,n}+\delta_{1,n}, & i=n\end{cases} \tag{4}$$

延迟差 ΔD_n 可以由延迟向量 $\boldsymbol{W}$ 与特征向量 $\boldsymbol{\Phi}=(\varphi_1,\varphi_2,\cdots\varphi_{n+1})$ 相乘得到,计算公式为 $\Delta D_n=\boldsymbol{W}^{\mathrm{T}}\boldsymbol{\Phi}$,其中特征向量由激励向量 $\boldsymbol{C}$ 按式(5)计算得到.

$$\varphi_i=\begin{cases}\prod_{l=i}^{n}(1-2c_l), & 1\leq i\leq n\\ 1, & i=n+1\end{cases} \tag{5}$$

虽然延迟参数无法通过测量直接得到,但是获取足够数量的 CRP 后,可用最小二乘法求解出 $\boldsymbol{W}$,得到数学模型[11].

忆阻器的制造过程差异可以用作硬件安全原语的源熵,Rose 和 Meade 设计了忆阻器交叉 PUF,Liu Yuntao 等分析给出了其数学模型[10],如式(6)所示.

$$U(D,\boldsymbol{\Phi})=\boldsymbol{\Phi}^{\mathrm{T}}D \tag{6}$$

其中 $\boldsymbol{\Phi}=(c_1,c_2,\cdots,c_n,1)^{\mathrm{T}}$ 是特征向量. 忆阻器交叉 PUF 的数学模型与仲裁器 PUF 的模型类似,相应的,建模方法也在两类 PUF 间通用.

2.3 机器学习攻击方法

近年来,机器学习被广泛应用于硬件木马检测、芯片防伪和轻量级认证协议攻击,也成为攻击强 PUF 的重要方法[12,13]. 逻辑回归(Logistic Regression,LR)、支持向量机(Support Vector Machine,SVM)、进化策略(Evolution Strategy,ES)、深度学习(Deep Learning,DL)和集成学习(Ensemble Learning,EL)等均能攻击一种或数种 PUF[11,14~16].

以使用逻辑回归攻击仲裁器 PUF[11]为例,定义概率函数 $\mathrm{p}(\boldsymbol{C},r/\boldsymbol{W})=\sigma(rf)=(1+\mathrm{e}^{-rf})^{-1}$,表示延迟向量为 $\boldsymbol{W}$ 条件下,输入激励 $\boldsymbol{C}$ 时得到响应 r 的概率. 如式(7)所示,对训练集 M 中的 CRP 使用梯度下降法,得出延迟向量即完成建模.

$$\begin{aligned}\widehat{\boldsymbol{W}}&=\mathrm{argmin}_{\boldsymbol{W}}l(M,\boldsymbol{W})\\ &=\mathrm{argmin}_{\boldsymbol{W}}\sum_{(C,r)\in M}-\ln(\sigma(rf(\boldsymbol{W},\boldsymbol{C})))\end{aligned} \tag{7}$$

Ruhrmair 分别基于仿真和实际数据对 LR、SVM 和 ES 在攻击标准、异或、轻量级和前馈等一系列基于仲裁器的 PUF 上的应用进行了研究,重点分析了逻辑回归攻击强 PUF 的性能表现,结果表明基于仲裁器的 PUF

基本上无法抵御机器学习攻击,但是攻击复杂度会随着激励位数、基础仲裁器 PUF 个数的增加而迅速增长[11].

Vijayakumar A 等针对抽象 PUF 模型进行了基于支持向量机、逻辑回归、Bagging and Boosting 和进化策略等算法的攻击实验,得出集成学习如随机森林、Adaboost 等比传统算法更有效的结论[17].

2.4　抗机器学习攻击的 PUF 结构

为了抵御机器学习攻击,PUF 设计者开展了大量的研究改进工作.

防御思路之一是混淆,即隐藏 PUF 的真实激励或响应. 例如 RFEP(Reverse Fuzzy Extractor Protocol)协议[18]、Slender PUF[19](SLPUF)等在认证应用中仅使用辅助数据或响应的部分序列. 但辅助数据也会泄露相应的信息,Becker 使用进化策略从泄露的不可靠性信息中获得了准确的 PUF 模型,从而破解了 RFEP 和 SLPUF. Gassend 等提出的受控 PUF(Controlled PUF)采取了防止攻击者随意输入激励,将响应经过移位处理后输出等策略[20],但机器学习使用其纠错码信息也可以成功建模.

另一种安全强化方法是加入非线性元素,如异或 PUF. 如图 2 所示,异或 PUF 中的仲裁器 PUF 使用相同的激励,得到的响应经异或运算后输出. 组合 PUF 与异或 PUF 结构相似,唯一区别是组合 PUF 中各基础仲裁器 PUF 的激励不同. Sahoo 等提出了两种不同的密码分析方法,成功攻击了组合 PUF 和 LSPUF[8].

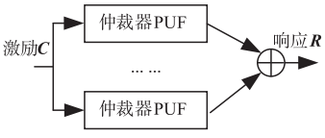

图2　异或PUF电路结构

上述两种思路可以融合在一起,形成抗机器学习能力更强的 PUF 设计,如张吉良等提出的 DMOS-PUF[21],以预先存储的 PUF 响应作为密钥,使用随机数选择密钥,通过异或运算对 PUF 的激励响应进行混淆,进一步提高了安全性.

异或 PUF 规模越大,其抗机器学习建模的能力越强. 当单独使用的机器学习算法遇到结构复杂、规模庞大的 PUF,其运算时间往往超出了实用范围[22].

3　PUF 侧信道分析

密码算法的物理安全(即实现安全)并不等价于算法设计时的理论安全,实现过程中产生的时间、功耗、电磁辐射等信息泄露称为侧信道泄露. 利用侧信道泄露信息进行密码分析的方法称为旁路攻击或侧信道分析. 虽然 PUF 被认为能抵御克隆、物理侵入等物理攻击,但是一系列研究表明 PUF 并不具备抗侧信道分析的特性.

3.1　可靠性分析

可靠性(reliability)是 PUF 的安全属性之一,常用于 PUF 性能评估. 如果不同环境下,PUF 对同一激励的响应一致,则认为该 PUF 具备可靠性[23]. PUF 的可靠性 Rel 通常借助位错误率(Bit Error Rate,BER)衡量,计算方法如式(8).

$$\mathrm{Rel}=1-\mathrm{BER}=1-\frac{1}{m}\sum_{t=1}^{m}\frac{\mathrm{HD}(R_i,R_{i,t})}{l} \tag{8}$$

其中 l 为响应的位数,m 为测试次数,$\mathrm{HD}(R_i,R_{i,t})$ 为第 t 次测试的响应与原始响应间的汉明距离. 可靠性的理想值为 1,但实际环境中,PUF 不可避免地会受到温度、电压、老化等诸多因素影响,使得实际测量值小于 1. 各类 PUF 的可靠性差别不大,平均为 94.5%. 其中,仲裁器 PUF 的可靠性通常在 95% 左右[24]. 对攻击者而言,PUF 响应的可靠性也能提供有价值的信息[25],单纯使用可靠性信息对 PUF 进行建模,称为可靠性分析.

Delvaux 基于可靠性提出了可重复度(repeatability)的概念,特指 PUF 在噪声影响下的短期可靠性[25](不包括器件老化的长期效应). 可重复度利用重复测量中随机噪声的作用对仲裁器 PUF 建模[26],实现了 Ruhrmair 提出的思路[14]. 尽管性能上不及机器学习,但该方法仅使用可靠性信息,开辟了强 PUF 侧信道攻击的先河.

可重复度 Rep 是重复使用同一激励 n 次,得到响应值为 1 的次数与总次数 n 的比值. 考虑仲裁器引入的延迟后,PUF 输出由 $\Delta D_{\mathrm{PUF}}-\Delta t_{\mathrm{Arb}}$ 决定. 结合噪声的概率密度函数,可重复度 Rep 与 PUF 参数的关系如式(9):

$$\mathrm{Rep}(\Delta D_{\mathrm{PUF}})=\frac{1}{2}\mathrm{erfc}\left(\frac{\Delta t_{\mathrm{Arb}}-\Delta D_{\mathrm{PUF}}}{\sqrt{2}\sigma_N}\right) \tag{9}$$

其中 $\mathrm{erfc}(t)=(2/\sqrt{\pi})\int_t^{\infty}\mathrm{e}^{-z^2}\mathrm{d}z$,$\Delta D_{\mathrm{PUF}}(\mathrm{Rep})=\Delta t_{\mathrm{Arb}}-\sqrt{2}\sigma_N\,\mathrm{erfc}^{-1}(2\mathrm{Rep})$.

当 $\Delta D_{\mathrm{PUF}}=\Delta t_{\mathrm{Arb}}$ 时,Rep 值为 1/2,Rep 偏离 1/2 越远,说明该响应的可重复度越高. Rep 在[0.1,0.9]取值范围内同激励 C_i 成线性关系[26],采集得到线性区域内的 n 个(C_i,Rep_i)组合后,用最小二乘法建立延迟模型,建模过程与利用激励响应对联立方程组类似. 但通常仲裁器 PUF 的响应都比较稳定,只有 10% 左右的取值会在[0.1,0.9]范围内. 因此该方法效率比机器学习低,且不能用于异或 PUF.

3.2　功耗分析和电磁分析

功耗分析是一种利用设备运行时泄露的功耗进行密码分析的方法,通过功耗分析能推导出设备进行的操作和操作涉及的参数.

Ruhrmair 等基于 PSpice 仿真提出了利用简单功耗分析获取 PUF 响应信息的方法[27]. 仲裁器 PUF 通常用触发器实现仲裁,响应为"1"时触发器发生翻转,其功耗远高于响应为"0"不发生翻转时. 测量供电引脚上的电压/电流值,得到功耗轨迹,就能够确定响应值. 简单功耗分析适用于内部包含多路仲裁器 PUF 且响应无法从外界直接读取的复杂结构 PUF. 这类 PUF 的功耗值会随着内部响应"1"的数量线性增长,通过简单计算可得到响应统计值. 对组合 PUF 可采取分治策略[27],通过不断设计调整激励,分析功耗得到响应为"1"的 PUF 总数,可以分辨出每路仲裁器 PUF 的激励响应. 相当于把组合 PUF 拆分成许多独立的仲裁器 PUF,通过联立方程即可获得模型参数.

实际使用时,PUF 仅占整个芯片的很小一部分,加上噪声干扰,使得简单功耗分析几乎无法从芯片的实测功耗轨迹中提取 PUF 的功耗信息. 此外,简单功耗分析对异或 PUF 无效,因为异或 PUF 中的仲裁器 PUF 共用激励,简单功耗分析得到的响应总数无法区分出单个的仲裁器 PUF.

电磁分析是利用设备运行时对外的电磁辐射进行密码分析的方法,主要通过在设备附近放置电磁探头,测量其工作期间辐射的电磁信号,研究电磁场与内部结构、运行算法之间的相关性,获取内部参数. 电磁分析与功耗分析类似,2011 年 Merli 等首次对 RO PUF 实施了电磁侧信道攻击[28].

4 侧信道混合攻击

常见的侧信道攻击方法无法实现对异或 PUF 的建模,而单纯使用机器学习,其攻击时间和准确率往往达不到实用要求. 为了实现对非线性结构 PUF 的建模,将侧信道攻击与机器学习方法结合,形成混合攻击[29~32],利用侧信道分析得到的额外结构信息,同 CRP 一起提供给机器学习算法,混合攻击方法能够提高预测准确率、缩短攻击时间.

4.1 基于可靠性的 CMA-ES 混合攻击

单纯可靠性分析不但使用范围受限,而且攻击效果也不如机器学习. 为了攻击异或 PUF 等非线性结构的复杂 PUF,Becker 提出了进化策略协同可靠性分析的侧信道混合攻击方法[30]. 攻击使用的可靠度参数 h 由 Rel 衍生而来,计算方法如式(10). 其中 r_j 是选定特征向量 $\boldsymbol{\Phi}$ 重复测试对象 m 次,得到的 m 个响应中的第 j 个.

$$h = \left| \frac{m}{2} - \sum_{j=1}^{m} r_j \right| \tag{10}$$

基于可靠性的 CMA-ES 算法运算过程同传统的 CMA-ES 算法类似,除了增加了参数 ε 以外,其他步骤都一样. 运行时按照式(11)对模型计算猜测可靠度 $\hat{h}$.

$$\hat{h} = \begin{cases} 1, & |\boldsymbol{W}^{\mathrm{T}}\boldsymbol{\Phi}| \geq \varepsilon \\ 0, & |\boldsymbol{W}^{\mathrm{T}}\boldsymbol{\Phi}| < \varepsilon \end{cases} \tag{11}$$

计算可靠度向量 $\boldsymbol{H} = (h_1, \cdots, h_n)$ 和猜测可靠度向量 $\hat{\boldsymbol{H}} = (\hat{h}_1, \cdots, \hat{h}_n)$ 的皮尔逊相关系数(Pearson correlation coefficient),选出相关系数高的模型作为下一轮的亲本,直至收敛到近似最优解.

异或 PUF 的可靠性由 n 路仲裁器 PUF 共同决定,每路的权重都相同,任意一路响应翻转都将导致结果变化,因此基于可靠性攻击异或 PUF 可以采取分治策略. 对其中 1 路仲裁器 PUF 建模时,将其他 $n-1$ 路视为噪声,用以相关系数为适应度函数的 CMA-ES 算法建模,重复 n 次后就得到异或 PUF 的模型. CMA-ES 属于非确定性算法,当不能收敛到最优解时,需重启学习过程. CMA-ES 算法的概率特性使得每次运行都能收敛到不同的 PUF 模型[31],正好规避了其他机器学习方法攻击异或 PUF 时每次都收敛到同一模型的缺陷. 该混合攻击方法的复杂度仅随着异或逻辑的数量线性增长,使攻击复杂度由指数级降至多项式级.

4.2 功耗分析混合攻击

单纯的功耗分析不适用于异或 PUF,Ruhrmair 以异或 PUF 为目标,提出了结合逻辑回归的功耗侧信道混合攻击方法[19]. 设 $r_i(C) \in \{0,1\}$ 表示激励为 C 时异或 PUF 中第 i 个仲裁器 PUF 的输出,则输出为"1"的仲裁器 PUF 数量 $n = \sum_i r_i(\boldsymbol{C})$. 使用梯度优化算法,让 $f(\boldsymbol{W},\boldsymbol{C}) = \sum_i \boldsymbol{\Theta}(\boldsymbol{W}_i^{\mathrm{T}}\boldsymbol{\Phi})$ 和实际输出 n 间的均方误差最小.

$$l(M,\boldsymbol{W}) = \sum_{(\boldsymbol{c},t)\in M} (f(\boldsymbol{W},\boldsymbol{C}) - n)^2 \tag{12}$$

相应的梯度为:

$$\nabla l(M,\boldsymbol{W}) = \sum_{(\boldsymbol{c},t)\in M} 2(f(\boldsymbol{W},\boldsymbol{C}) - n)\nabla f(\boldsymbol{W}) \tag{13}$$

$$\nabla f(\boldsymbol{W}_j) = \sigma(\boldsymbol{W}_j^{\mathrm{T}}\boldsymbol{\Phi}_j)(1 - \sigma(\boldsymbol{W}_j^{\mathrm{T}}\boldsymbol{\Phi}_j))\boldsymbol{\Phi}_j \tag{14}$$

使用 RProp 算法来寻找使 l 最小的解 $\hat{\boldsymbol{W}}$. 运用侧信道信息的模型复杂度仅取决于方向,与权重向量的长度无关,降低了算法求解的复杂度. Ruhrmair 使用两阶段优化方法,首先,使用侧信道信息基于梯度优化建模,直到能准确预测 95% 以上的异或输出,然后使用标准的逻辑回归算法继续优化模型[32].

Becker 通过对比 PUF 的实际功耗曲线与不同准确率预测模型的功耗曲线间的皮尔逊相关系数,得出结论:模型与真实 PUF 的功耗相关系数同模型的准确率呈线性关系,即模型的准确率越高,功耗相关系数值越大. 功耗相关分析的结果可以用于判断模型的优劣. Becker 随后使用相关功耗分析[31]结合 CMA-ES 成功建模了 LSPUF. 使用进化策略随机产生最初的预测模型,

以功耗相关系数作为适应度函数判断模型的准确性,将准确率高的模型留作亲本,最终收敛到近似最优解.进化策略的优势在于受噪声影响较小,这种方法可以推广应用于异或 PUF 建模,此外,Becker 还指出增加仲裁器 PUF 的个数对该混合攻击的影响有限,无法起到提高安全性的作用.

4.3　故障注入混合攻击

故障攻击(fault attack)属于主动攻击,利用错误信息和故障行为对 PUF 进行建模的方法称为 PUF 故障攻击[33]. 故障注入是实施故障分析的必要条件,由于 PUF 自身固有的防篡改特性,侵入式和半侵入式故障注入难以发挥作用,通常使用非侵入式故障注入,通过外界干扰引发 PUF 产生响应位翻转等故障. 理论上,改变电压、磁场、温度等均可能导致 PUF 响应故障,但考虑到实施的难易程度,最常用的仍然是调整电压和温度[29].

Delvaux 在基于可重复性建模 PUF 时,就提出可以通过改变供电电压、工作温度等环境条件诱发响应错误,增加不稳定 CRP 的数量,有针对性的产生可用的 CRP,提高攻击效率[34]. Becker 在攻击 Controlled PUF 时也使用了故障注入[30]. 其基本思路为:在激励不变的情况下,逐步调整供电电压后观测响应情况,如果诱发了位翻转故障则记故障参数 f 为 1,否则记 f 为 0. 发生位翻转是因为在特定激励下 PUF 对应的延迟差接近于 0,即 $|\Delta D_n| < \varepsilon$. 定义激励为 $\boldsymbol{C}$ 时 PUF 模型对应产生的延迟差为 $\Delta\hat{D}_i$,模型的故障参数为 $\hat{f}_i$. 如果 $|\Delta\hat{D}_i| < \varepsilon$ 记 $\hat{f}_i = 1$,否则 $\hat{f}_i = 0$. 以模型故障向量 $\hat{\boldsymbol{F}} = (\hat{f}_1, \cdots, \hat{f}_n)$ 与实测故障向量 $\boldsymbol{F} = (f_1, \cdots, f_n)$ 的相关系数作为适应度函数,判定 CMA-ES 算法每轮得到模型是否合适. 目前的故障注入方法均作为可靠性侧信道攻击的辅助手段,用于加速收集可靠性信息.

4.4　侧信道混合攻击策略及性能

混合攻击的组合方式和攻击对象见表 1.

表 1　混合攻击方式与对象

侧信道	机器学习	攻击对象
可靠性[31]	CMA-ES	XOR PUF
功耗[30]	CMA-ES	LW PUF、Controlled PUF
功耗[32]	LR	XOR PUF、LW PUF
时序[32]	LR	XOR PUF、LW PUF
可靠性 + 电压故障注入[30]	CMA-ES	Controlled PUF 、SL PUF
功耗[10]	optimization-theoretic ML	Arbiter PUF、Memristor Crossbar PUF

攻击方法的性能主要从模型准确率、所需 CRP 数量和攻击时间等评价指标上. 模型准确率是通过测试数据集中的 CRP 检验模型时,响应正确的数量与 CRP 总数之间的比值.

表 2　针对异或 PUF 等混合攻击性能对比

攻击方法	对象及复杂度	CRP 数量(10^3)	准确率	时间
可靠性 + CMA-ES[31]	128bit 8XOR	300	89.1%	3.4h
	128bit 16XOR	500	80.1%	34h
功耗分析 + LR[32]	128bit 8XOR	51.6	98%	13min
	128bit 16XOR	103.2	97.5%	148min
功耗分析 + OT[10]	128bit 5XOR	5.7	95%	3.3h
	256bit 5XOR	11.2	95%	29.7h

由于运行算法的平台各不相同,因此时间参数只用于评估复杂度带来的影响. 各类方法攻击的 PUF 类型各不相同,表 2 选取攻击难度最大的异或 PUF 和前馈 PUF 为例,比较对应方法的性能. 从对比可以看出,功耗加逻辑回归混合攻击的所需 CRP 数量和准确率指标上优于可靠度加进化策略的混合攻击. 但 CMA-ES 也具有其独特优势,在 PUF 的电路结构未知时,只要拥有足够规模的 CRP,仍然能够通过 CMA-ES 算法推测出结构模型和内部参数.

5　PUF 侧信道攻击的困境与对策

5.1　侧信道攻击面临的困境

尽管侧信道分析为攻击复杂结构的 PUF 开辟了新途径,但是实际应用中侧信道攻击也面临着诸多困难.

随着防侧信道泄露方法的应用和抗侧信道攻击结构的提出[35~37],传统方法面临失效的危险. 例如增加电路的对称性,使用特殊的电路结构如电流受限型延迟单元[35]等防止侧信道信息泄露的方法被应用于 PUF 设计. 新的 PUF 结构一方面从运行机制上阻断了实施侧信道攻击的可能性,如 Gao Yansong 等提出的 FSM-PUF 通过禁止重复相同激励限制可靠性信息的提取,以此避免可靠性分析[36],另一方面传统密码算法被用来同 PUF 结合以抵御侧信道混合攻击,如 W Yu 等提出的掩膜 AES PUF[37]. 这些方法都大大提高了侧信道攻击的难度.

一些侧信道信息测量手段和半入侵攻击方式需要特殊设备或条件,不具备普适性. 如 Ruhrmair 提出的时序侧信道攻击,其延迟测量电路利用 FPGA 内部逻辑实

现,无法适用于大多数 PUF. Shahin Tajik 等提出的激光故障注入[33]属于半入侵式攻击,施加激光脉冲诱发故障时要去除目标芯片底部封装,将硅衬底尽量变薄,实施难度较大. 光子发射分析也由于需要专业设备,限制了其应用.

尽管目前 PUF 安全分析已经将抗侧信道分析能力纳入评价体系,如 Fukushima 等采用侧信道分析与模型分析结合评估 PUF 的安全性[38],但是还局限于以实验方式测试,缺乏模型支持和理论研究,限制了侧信道安全评价的适用范围. 对 PUF 抗侧信道能力的建模分析和理论研究也是当前急需解决的问题.

5.2 宜采取的对策建议

面对上述挑战,运用新的侧信道信息获取手段、研究新的侧信道分析方法成为研究人员关注的焦点.

当前侧信道混合攻击中使用的信息还局限于功耗、可靠性等,分析方法也只用了简单分析和相关性分析. 新的侧信道信息获取途径如光子发射分析[39]、剩磁衰变[40]等可以与机器学习结合发挥作用,复杂的侧信道分析方法如模板分析、立方体分析等也有待应用于强 PUF 攻击. 以模板分析为例,现有的攻击场景通常设定为已知攻击对象的结构类型,但实际上情况往往更复杂,攻击者不能确定对象类型,也就无法选用合适的攻击方法. 模板分析可以通过对功耗等特征的相关性分析,建立模板,根据侧信道信息分析确定对象类型,进而获取模型参数.

目前 PUF 设计中使用的侧信道防御对策通常只能应对某一类或几类侧信道攻击,无法防御所有的攻击类型. 研究多侧信道组合的混合攻击方法成为新的突破口. 例如,增加电路结构的对称性可以抵御功耗分析,但对可靠性分析没有效果;2018 年 Sahoo 等提出了一种新 PUF 结构 rMPUF[41]以抵御基于可靠性的 CMA-ES 混合攻击,但是,该结构没有对功耗侧信道攻击做防范. 将不同的侧信道分析方法组合,能获得更多的目标信息,最终拓宽攻击范围和提高预测性能.

6 结论

PUF 作为一种硬件安全原语,以其轻量级、防篡改等特性受到研究者的广泛关注. 但是无论是已经商业化应用的强 PUF 还是新提出的强 PUF 结构,均缺乏足够的侧信道安全性,侧信道分析与机器学习结合的混合攻击对当前几乎所有基于仲裁器的 PUF 均构成严重威胁. 研究者在进行 PUF 设计或安全性评估时,需充分考虑抗侧信道混合攻击的能力.

参考文献

[1] Gassend B, Clarke D, Dijk M V, et al. Silicon physical random functions[A]. Proceedings of 9th ACM Conf Computer and Communication Security [C]. New York: ACM Press, 2002. 148 – 160.

[2] Herder C, Yu M D, Koushanfar F, et al. Physical unclonable functions and applications: a tutorial[J]. Proceedings of the IEEE, 2014, 102(8): 1126 – 1141.

[3] Jiliang Zhang, Gang Qu, Yongqiang Lyu, Qiang Zhou. A survey on silicon PUFs and recent advances in ring oscillator PUFs[J]. Journal of Computer Science and Technology (JCST), 2014, 29(4): 664 – 678.

[4] Rose G S, Meade C A. Performance analysis of a memristive crossbar PUF design[A]. Proceedings of the 52nd Annual Design Automation Conference[C]. New York: ACM Press, 2015. 1 – 6.

[5] Majzoobi M, Koushanfar F, Potkonjak M. Lightweight secure PUFs[A]. IEEE/ACM International Conference on Computer-Aided Design[C]. New York: IEEE, 2008. 670 – 673.

[6] Sahoo D P, Saha S, Mukhopadhyay D, et al. Composite PUF: A new design paradigm forphysically unclonable functions on FPGA[A]. IEEE International Symposium on Hardware-Oriented Security and Trust [C]. New York: IEEE, 2014. 50 – 55.

[7] Nguyen P H, Sahoo D P. Lightweight andsecure PUFs: a survey[A]. International Conference on Security, Privacy, and Applied Cryptography Engineering[C]. Berlin: Springer International Publishing, 2014. 1 – 13.

[8] Sahoo D P, Nguyen P H, Mukhopadhyay D, et al. Acase of lightweight PUF constructions: cryptanalysis and machine learning attacks[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2015, 34(8): 1334 – 1343.

[9] Xu X, Burleson W. Hybrid side-channel/machine-learning attacks on PUFs: a new threat? [A]. Proceedings of Design, Automation and Test in Europe Conference and Exhibition[C]. New York: IEEE, 2014. 349 – 354.

[10] Liu Y, Xie Y, Bao C, et al. Acombined optimization-theoretic and side-channel approach for attacking strong physical unclonable functions[J]. IEEE Transactions on Very Large Scale Integration Systems, 2017, 26(1): 73 – 81.

[11] Rührmair U, Sölter J, Sehnke F, et al. PUFmodeling attacks on simulated and silicon data[J]. IEEE Transactions on Information Forensics & Security, 2013, 8(11): 1876 – 1891.

[12] Tobisch J, Becker G T. On the scaling of machine learning attacks on PUFs with application to noise bifurcation[A]. International Workshop on Radio Frequency Identification: Security and Privacy Issues [C]. Berlin: Springer, 2015. 17 – 31.

[13] J Ye, Q Guo, Y Hu, H Li, X Li. Modeling attacks on strong physical unclonable functions strengthened by random number and weak PUF[A]. 2018 IEEE 36th VLSI Test Symposium[C]. New York:IEEE,2018. 1 – 6.

[14] Rührmair U, Holcomb D E. PUFs at a glance[A]. Proceedings of Design, Automation & Test in Europe Conference & Exhibition[C]. New York:IEEE,2014. 1 – 6.

[15] Ganji F, Tajik S, Fäßler F, et al. Strong machine learning attack against PUFs with no mathematical model[A]. Proceedings of Cryptographic Hardware and Embedded Systems[C]. Berlin:Springer,2016. 389 – 411.

[16] Guo Q, Ye J, Gong Y, et al. Efficientattack on non-linear current mirror PUF with genetic algorithm[A]. Asian Test Symposium[C]. New York:IEEE,2016. 49 – 54.

[17] Vijayakumar A, Patil V C, Prado C B, et al. Machine learning resistant strong PUF: possible or a pipe dream? [A]. IEEE International Symposium on Hardware Oriented Security and Trust[C]. New York: IEEE, 2016. 19 – 24.

[18] Herrewege A V, Katzenbeisser S, Maes R, et al. Reverse-fuzzy extractors: enabling lightweight mutual authentication for PUF-enabled RFIDs[A]. Financial Cryptography and Data Security[C]. Berlin:Springer,2012. 374 – 389.

[19] Majzoobi M, Rostami M, Koushanfar F, et al. Slender PUF protocol: a lightweight, robust, and secure authentication by substring matching[A]. Proceedings of IEEE Symposium on Security and Privacy Workshops[C]. New York: IEEE Computer Society,2012. 33 – 44.

[20] Gassend B, Clarke D, Dijk M V, et al. Controlled physical random functions[A]. Proceedings of Computer Security Applications Conference[C]. New York:IEEE,2007. 149 – 160.

[21] Jiliang Zhang, Lu Wan, Qiang Wu, Gang Qu. DMOS-PUF: dynamic multi-key-selection obfuscation for strong PUFs against machine learning attacks[EB/OL]. https://arxiv. org/abs/1806. 02011, 2018, arXiv: 1806. 02011.

[22] Ruhrmair U, Schlichtmann U, Burleson W. Special session: How secure are PUFs really? on the reach and limits of recent PUF attacks[A]. Design, Automation and Test in Europe Conference and Exhibition[C]. New York: IEEE,2014. 346 – 349.

[23] Chang C H, Zheng Y, Zhang L. Aretrospective and a look forward: fifteen years of physical unclonable function advancement[J]. IEEE Circuits & Systems Magazine,2017, 17(3):32 – 62.

[24] 叶靖,胡瑜,李晓维. 非确定性仲裁型物理不可克隆函数设计[J]. 计算机辅助设计与图形学学报,2017(1):166 – 171.

YE Jing, HU Yu, LI Xiao-wei. Nondeterministic logic based arbiter physical unclonable function[J]. Journal of Computer-Aided Design and Computer Graphics, 2017 (1):166 – 171. (in Chinese)

[25] Becker G T. Onthe pitfalls of using arbiter-PUFs as building blocks[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2015, 34 (8): 1295 – 1307.

[26] Delvaux J, Verbauwhede I. Side channel modeling attacks on 65nm arbiter PUFs exploiting CMOS device noise [A]. Proceedings of IEEE International Symposium on Hardware-Oriented Security and Trust[C]. New York: IEEE,2013. 137 – 142.

[27] Mahmoud A, RÄuhrmair U, Majzoobi M, Koushanfar F. Combined modeling and side channel attacks on strong PUFs[J]. IACR Cryptology ePrint Archive, 2013: 632 – 640.

[28] Merli D, Heyszl J, Heinz B, et al. Localized electro-magnetic analysis of RO PUFs[A]. IEEE International Symposium on Hardware-Oriented Security and Trust[C]. New York:IEEE,2013. 19 – 24.

[29] Kumar R, Burleson W. Hybrid modeling attacks on current-based PUFs[A]. IEEE International Conference on Computer Design[C]. New York:IEEE,2014. 493 – 496.

[30] Becker G T. R Kumar. Active and passive side-channel attacks on delay based PUF designs[J/OL]. IACR Cryptology ePrint Archive, 2014, https://eprint. iacr. org/2014/287,2014.

[31] Becker G T. Thegap between promise and reality: on the insecurity of XOR arbiter PUFs[A]. Cryptographic Hardware and Embedded Systems[C]. Berlin:Springer,2015. 535 – 555.

[32] Xu X, Mahmoud A, Majzoobi M, et al. Efficient power and timing side channels for physical unclonable functions [A]. International Workshop on Cryptographic Hardware and Embedded Systems[C]. New York:Springer-Verlag, 2014. 476 – 492.

[33] Tajik S, Lohrke H, Ganji F, et al. Laser fault attack on physically unclonable functions[A]. The Workshop on Fault Diagnosis & Tolerance in Cryptography[C]. New York:IEEE Computer Society,2015. 85 – 96.

[34] Delvaux J, Verbauwhede I. Fault injection modeling attacks on 65 nm arbiter and RO sum PUFs via environmental changes[J]. IEEE Transactions on Circuits & Systems I Regular Papers,2014,61(6):1701 – 1713.

[35] Cao Y, Zhao X, Ye W. A compact and low power RO PUF with high resilience to the EM side-channel attack and the SVM modelling attack of wireless sensor networks[J]. Sensors,2018,18(2):322 – 332.

[36] Gao Y, Ma H, Al-Sarawi S F, et al. PUF-FSM: a controlled strong PUF[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2018, 37(5): 1104 – 1108.

[37] Yu Weize, Chen Jia. Masked AES PUF: a new PUF against hybrid side-channel/machine-learning attacks[J]. Electronics Letters, 2018, 54(10): 618 – 620.

[38] Fukushima, Souissiy, et al. Delay PUF assessment method based on side-channel and modeling analyzes: the final piece of all-in-one assessment methodology[A]. Trustcom/Big Data SE/ISPA[C]. New York: IEEE, 2016. 201 – 207.

[39] Tajik S, Dietz E, Frohmann S, et al. Photonic side-channel analysis of arbiter PUFs[J]. Journal of Cryptology, 2017, 30(2): 550 – 571.

[40] Zeitouni S, Oren Y, Wachsmann C, et al. Remanence decay side-channel: the PUF case[J]. IEEE Transactions on Information Forensics & Security, 2016, 11(6): 1106 – 1116.

[41] Sahoo D P, Mukhopadhyay D, Chakraborty R S, et al. Amultiplexer-based arbiter PUF composition with enhanced reliability and security[J]. IEEE Transactions on Computers, 2018, 67(3): 403 – 417.

作者简介

刘 威 男,1982 年 1 月出生于湖北省随州市. 现为战略支援部队信息工程大学讲师. 主要研究方向为嵌入式系统安全.
E-mail: newwayclwy@sina.com

蒋烈辉 男,1967 年 5 月出生于浙江省东阳市. 现为战略支援部队信息工程大学教授、博士生导师. 主要研究方向为计算机体系结构、逆向工程.
E-mail: jiangliehui@163.com

常 瑞 女,1981 年 10 月出生于河南省郑州市. 现为战略支援部队信息工程大学副教授、硕士生导师. 主要研究方向为嵌入式系统安全.
E-mail: crix1021@163.com