侧信道攻击物理不可克隆函数研究综述

S191000899

廖慧芝

摘要

物理不可克隆函数（PUF）已经成为传统加密技术的一种轻量级替代品。与传统加密算法相比，PUF 不需要将密钥存储在非易失性存储器，这使得 PUF 更具安全优势。机器学习是对 PUF 的一种有效的攻击方式，已被广泛研究，但其应用于较为复杂的 PUF 实例时，机器学习建模攻击的效率和精度达到了极限。为了解决这个问题 Ahmed Mahmoud, Ulrich Ruhrmair 等于 2013 年提出可将侧信道攻击与机器学习建模攻击相结合，使得攻击更为有效。本文将简要介绍对 PUF 的一些经典侧信道攻击技术，以及利用侧信道分析与机器建模学习技术相结合的攻击方法。最后讨论利用侧信道技术对 PUF 进行攻击时，攻击者可能要面临的机遇和挑战。

关键词: 物理不可克隆函数 侧信道攻击 机器学习建模

1.简介

现在的应用和设计趋向于向小型、分布式、移动和无线方向发展。在芯片集成方面，密码保护是必不可少的，因为几乎所有的应用程序都需要处理敏感的数据，但由于上述的趋势，实施密码保护受到了一定的限制。由于需要考虑能耗、功率、芯片面积等因素，因此密码保护趋于轻量级加密。此外由于具有移动性， 攻击者可能可以轻易获得对芯片的物理访问，因此无论是侵入式还是非侵入式的硬件攻击都会对芯片安全形成重大威胁。

传统的加密方法，严重依赖于设备对私密信息的存储能力，通常是通过二进制将密钥存储在非易失性存储器中，容易受到硬件方面的攻击。又由于存储的永久性，攻击者在攻击时间上不会受到限制。尽管可以利用硬件入侵检测电路来提高额外的保护，但实际实施可能会受到限制，可能造成较高的成本消耗，不符合轻量级的要求。

物理不可克隆函数逐渐成为一类众所周知的安全原语，并在此基础上，已提出了各种安全协议。物理不可克隆函数即 PUF 通过提取随机工艺差异来唯一标志和认证物理实体。通过输入相同的激励信号，PUF 能够重复产生随机响应。由于制造工艺偏差的不可控制性和随机性，预测或读取 PUF 的行为通常比获取存储

在 NVM 中的数字密钥更加困难。

PUF 依赖于”激励-响应”对，简称 CRP，其原理是通过提取芯片再制造过程中不可避免的工艺偏差，实现激励与响应之间对应的函数关系，可以将 PUF 看作一个黑盒，接受一定长度的输入后产生一定长度可以测量的输出，通常将接收的输入叫激励，将产生的输出叫响应。需要注意的是，这里的“黑盒”和数学意义上的函数相似，但又有一定的区别，二者都是接收输入后进行一系列运算最终得出具体的输出，但当给定两个相同的函数相同的输入时，产生的输出完全一样， 而给定两个结构的 PUF 两个相同的输入时，输出结果存在差别，此外函数内部是通过一系列数学运算后得到最终结果，而物理不可克隆函数则是利用了制造过程中的差异。

PUF 有强 PUF 和弱 PUF 之分。弱 PUF 指的是具有有限数量的激励-响应对的 PUF，一个弱 PUF 可以被一个（或非常少量）固定激励所询问，在这个激励下，它产生对应的响应，具体响应取决于其内部的物理差异。而强 PUF，与弱 PUF 相反，强 PUF 具有大量的 CRP，也就是说其是一个无序的物理系统，具有非常复杂的输入输出行为。这个系统允许很多可能的激励，并且必须对输入的激励和系统中存在的差异做出响应。

对于强 PUF，在理想情况下，其激励响应对（CRP）的数量会随着所需的芯片面积成指数倍增长。利用强 PUF 指数倍的 CPR，可以提供一种身份验证机制。理想情况下，由于极为复杂的 CPR 映射关系，可为这种身份验证机制提供较好的安全性能，但实际的 PUF 设计容易受到建模攻击的影响。目前利用 ES、SVM、LR 等机器学习算法已对PUF 实施了较高精度的建模攻击。而在处理更为复杂的 XOR PUF 和轻量级 PUF 时，建模攻击在攻击的效率和精度上却已经达到了瓶颈。为了进一步辅助机器建模攻击，已经提出利用额外的辅助渠道来增强 PUF 模型的构建。

侧信道攻击，是从硬件系统（通常是密码系统）的物理实现中提取信息。 侧信道分析方法不是蛮力地从算法理论缺陷层面寻找硬件系统的攻击突破口，而是从系统的实际实现中探寻硬件系统的安全弱点。常见的侧信道信息主要包括定时信息，功耗，电磁（EM）泄漏和声音等。从实现策略来看，一些侧信道攻击需要对实施了加密技术的系统内部进行相关操作，比如半侵入式攻击方法，而其他的侧信道攻击，如差分功率分析法之类的方法，则相当于黑箱攻击。对于 PUF 这样的复杂加密功能的硬件系统，仅仅利用侧信道直接对 PUF 行为进行建模是

不可行的。但是利用侧信道分析方法获取额外的侧信道信息，将可以很好地辅助机器学习建模攻击，大大提高机器建模学习攻击的效率和精度。以电源侧信道为例，一些常用的侧信道攻击方法实施电源侧信道攻击的主要目的是提取内部操作产生的各种功耗，如简单功率分析（SPA）、差分功率分析（DPA）和高阶差分功率分析。即对于使用电源 VDD 的硬件实现，其功耗 P 可以表示为：p=VDD*I, 其中 I 代表实时电流。因此通过观察和收集电路电路轨迹，可以推到出与内部密钥功能密切相关的功耗 p。

本文将简要介绍对 PUF 的一些经典侧信道攻击技术，常见的侧信道攻击技术可以分为被动攻击、主动攻击和混合攻击三类，针对三类攻击，本文将简要介绍已有的对 PUF 的攻击的研究。最后，本文将讨论利用侧信道技术对 PUF 进行攻击时，攻击者可能要面临的机遇和挑战。

2.侧信道攻击分类

2.1 被动攻击

被动攻击是侧信道攻击的主要组成部分，利用这种攻击方法对硬件系统实施攻击，攻击者仅仅被动地观察并从攻击目标中收集信息，而不会对目标系统进行更改。在必要的情况下，攻击者可能会通过改变硬件系统的运行条件以使得目标系统可执行特定的行为，以帮助攻击者获取一些所需的信息，如时间、功耗或电磁信息等。如上所述，被动攻击是基于统计方法的一种攻击方式。因此，实施被动侧信道攻击后续需要进行后续的分析操作以及使用相关的工具， 例如 Matalab。

被动侧信道攻击方式主要有基于功率分析的侧信道攻击、基于定时的侧信道攻击以及基于电磁辐射分析的侧信道攻击方式。本文主要介绍基于功率的侧信道攻击和基于定时的侧信道攻击方式。

(1) 基于功率分析的侧道攻击

2013 年, Mahmoud 等人在[20]中提出利用功耗分析方法对PUF 进行被动攻击，利用功耗侧信道攻击方法提取XOR-PUF 和轻量级PUF 的子响应的功耗信息。其中，XOR-PUF 和轻量级PUF 的基本机制是将具有XOR 功能的子响应模式编码为公共响应，以防止原始的 CRP 信息泄露。上述两个 PUF 的基本组件是仲裁器 PUF，它采

用锁存器作为仲裁器来确定响应。根据[20]，基于锁存器的仲裁器在生成“1”的响应位时比生成“0”的响应位所消耗的功率更多，输出最终响应位为”1”会使得整个 XOR / Lightweight PUF 的功耗增加，同时生成更多的“1”子响应。因此，使用功率轨迹分析法，利用侧信道可以收集整个 PUF 的当前功耗轨迹，并借助统计处理工具将当前轨迹转换为功耗（每个电流轨迹以下的面积）。如果用“0”到”1”响应将充电的 XOR 门表示为活动的 XOR 门，则额外的功耗量与活动的 XOR 门的数量成线性比例。因此，利用每个公共响应的功耗信息，推到出子响应中“1”的比例是可行的，这将大大提高了 PUF 模型预测的成功率。

(2) 基于定时分析的侧信道攻击

用时信息是另一个常用的侧信道分析的参数。通过观察执行加密操作所需的时间变化，就可以比较简单地实施定时侧信道攻击。利用定时分析所获得的侧信道信息，可能可以确定整个加密系统的密钥。这种攻击需要使用到基于计时的统计分析方式，其可用性在不同领域得到了证明[16，27，28]。即使目前还没有发布太多对 PUF 的定时侧信道攻击，收集定时侧信道信息的目的还在于获取相关的子响应位的附加信息，以带有 N 个子 PUF 的轻量级 PUF 为例，输出网络的 XOR 功能实际上是由几个基本的 XOR 门组成。由于每个 XOR 门之间的工艺差异，我门可以假设不同的子响应模式将具有不同的时序特征（脉冲传播时延）。因此通过扫描 PUF 电路的频率，可以对公共响应进行分类，然后引用子响应模式。

2.2 主动攻击

与被动攻击不同，主动攻击试图更改系统的资源配置，或者影响系统的正常运行，从而获取侧信道信息。通过操纵目标系统或者目标系统实施正常功能以外的运作环境，攻击者可以观察并且收集目标系统的性能变化。通过分析进行修改后的目标系统的输入输出对，攻击者可以了解目标系统的内部工作机制。主动攻击中常用的方法主要有故障注入攻击方式。利用故障注入攻击方式，攻击者可以分析出目标系统的加密密钥、更改程序的流程以破坏完整性检查等。最近发表的对 PUF 的主动攻击是 Delvaux 和 Verbauwhede [7]。在[7]中，PUF 的可重复性模型是基于受 CMOS 噪音源影响的 PUF 的短期可靠性而建立的，这里，需要区分噪音和可变性的概念。正常的功能电子电路既不需要噪音也不需要可变性。但是PUF 是一类特殊的安全性原语，其主要利用测量过程中的可变性来实现。因此， 噪音成为理想的“故障注入”的候选对象。对于 PUF 电路，噪音主要来自于温度

或电压的变化等，所有的噪音都可能影响 PUF 电路的 CRP 可重复性，即使得 PUF 电路的 CRP 可重复性降低。在[30]中，Georg T. Becker, Raghavan Kumar 以 Controlled PUF 为例，主要以 Controlled PUF 结构中的 128 阶 Arbiter PUF 为研究对象，通过改变 Controlled PUF 的工作电压，使得 Controlled PUF 结构中的 128 阶 Arbiter PUF 中部分部分响应位发生翻转，通过观察知，当一个激励的脉冲信号经过Arbiter PUF 两条竞争传输路径后到达最终Arbiter 的延迟差小于一个阈值时，这个激励产生的响应位容易发生翻转。攻击者利用更改电压的方式， 进行故障注入攻击，获取 Arbiter PUF 电路的延迟信息，以此辅助筛选利用进化策略（ES）的机器学习算法所产生的 PUF 模型。

2.3 混合攻击

机器学习建模攻击即使是一种有效的攻击方法，但当攻击 256 阶或大于 256 阶且具有 6 个 XOR 或更多 XOR 门的 XOR Arbiter PUF 或 Lightweight PUF 时，单纯使用 ML 建模对 PUF 进行攻击，其攻击效率和精度也达到了极限[16] [17] [15]。为了解决这个问题，Mahmoud 等人提出了将功率侧信道攻击与常规 ML 建模攻击相结合的对 PUF 的混合攻击。Ruhrmair 在[21]中提出，对于侧信道攻击中的功率分析侧信道攻击，其攻击效果不明显，但如果适当地与建模技术结合，则可以大大提供攻击者的攻击效率。在[20]中，“good”的子响应如“全为 1”和“全为 0”作为筛选的侧信道信息，以此整个 XOR 仲裁器 PUF 和轻量级 PUF 就被分解为多个单独的仲裁器 PUF。每个单独的仲裁器 PUF 收集了足量的 CRPs，当前的 ML 技术可对目标 PUF 进行建模，只要获取每个子 PUF，整个 XOR 仲裁器 PUF 和轻量级 PUF 将收到攻击。

3.挑战

由于 PUF 结构本身的复杂性以及对实施侧道攻击的严格要求，PUF 攻击者需要面临一些挑战。 同时，各种辅助渠道策略的可用性也为这项研究提供了机会。

3.1 实际实施困难

上文中，列出的一些已发布的边信道攻击，这节将分析实施上文列出的这些攻击方法时攻击者需要面临的挑战。根据侧信道攻击的定义，与整个测量值相比， 提取的侧信道信息通常只占很小的一部分。这对 PUF 进行侧信道攻击是一种潜在

挑战，因为如果攻击者无法测量到这种细微的参数特征，则侧信道攻击将变得毫无价值。以基于功率分析的侧信道攻击为例，在[20]中，Ruhrmair 等人使用的 CRP 主要来自 SPICE 模拟，即使通过仿真，也可以比较精确地提取出 XOR 门的额外功耗。在实际实现中，从 FPGA 或ASIC 芯片上的 PUF 收集的功率走线会混入大量噪声。在这种情况下，所有功率轨迹之间都没有显着差异，因此以此推断它们所关联的子响应是不可行的。造成这个问题的原因有可能有两个，其一是在真正的硅 PUF 中，XOR 门通常消耗不超过整个设计的 5%的硅资源，带电的 XOR 门所消耗的功率比整个电路少得多； 此外环境噪声和测量噪声对提取的功率轨迹有很大影响。

3.2 防御措施

针对以上问题，所提出各种攻击方法不仅可以验证 PUF 协议的质量，还可以激发人们寻找对策以完善它。常规的针对侧信道攻击的防御措施主要包括：1. 减少信息泄漏，如平衡值的处理； 2.增加电路操作的噪声，以此掩盖侧信道信息的细微的变化。针对已发布的针对 PUF 的侧道攻击方法，研究者们还制定了许多相应的对策。在[20]中提出了一种针对 XOR 仲裁器 PUF 和轻型 PUF 的功率侧信道攻击的方法。仲裁器产生响应位为“ 1”和为“ 0”仲裁器所消耗的功耗不同， 因此作者提出，可以在每个仲裁器 PUF 的末尾使用两个交叉的仲裁器，利用改进的仲裁器 PUF，产生了恒定数量的“ 1”和“ 0”响应，因此消耗了等量的功率。

4.机遇

即使攻击者在 PUF 上实施侧信道攻击会遇到诸多挑战，但这项研究也存在一些机会。 从上面的介绍中，不难看出，已经提出了所有经过验证的侧信道攻击并提出了相应的对策，从而可以更好地保护 PUF。 但是，基于复杂的内部变化，不同的 PUF 将依赖于不同的硅特性，因此很难提出全面的对策。 因此，目前还没有万全的方案可以保护 PUF 原语免受所有现有的侧信道攻击。

5.总结

本文主要介绍了一些已被提出的典型的对 PUF 的侧信道攻击方式。利用侧信道攻击方法辅助机器建模学习攻击方法 PUF 进行攻击，将大大提高攻击者的攻击能力，但在实际在对 PUF 实施侧信道攻击时，攻击者也面临着诸多挑战。

对于已有的所有的 PUF 防御机制的设计，针对各种防御机制已有相应的防御措施，但对于提供一种可以防御所有类型的侧信道攻击的 PUF 防御机制，仍然是未来值得研究的一个热点。

参考文献

[1] Jorge Guajardo, Sandeep S. Kumar, Geert Jan Schrijen, Pim Tuyls: FPGA Intrinsic PUFs and Their Use for IP Protection, in CHES 2007:63-80

[2] Sandeep S. Kumar, Jorge Guajardo, Roel Maes, Geert Jan Schrijen, Pim Tuyls: The Butterfly PUF: Protecting IP on every FPGA, HOST 2008: 67-70

[3] Pim Tuyls, Geert Jan Schrijen, Boris Skoric, Jan van Geloven, Nynke Verhaegh, Rob Wolters: Read-Proof Hardware from Protective Coat- ings, CHES 2006: 369-383

[4] Christina Bruzska, Marc Fischlin, Heike Schroder, Stefan Katzenbeiss- er: Physically Unclonable Functions in the Universal Composition Framework. CRYPTO2011.

[5] Erdinc ztrk, Ghaith Hammouri, Berk Sunar: Towards robust low cost authentication for pervasive devices. In PerCom, pages 170-178. IEEE Computer Society,2008.

[6] Mehrdad Majzoobi, Farinaz Koushanfar, Miodrag Potkonjak: Lightweight Secure PUFs. IC-CAD 2008: 607-673.

[7] Jeroen Delvaux, Ingrid Verbauwhede: Side channel modeling attacks on 65nm arbiter PUFs exploiting CMOS device noise. HOST 2013.

[8] Jeroen Delvaux, Ingrid Verbauwhede: Fault Injection Modeling Attacks on 65nm Arbiter and RO Sum PUFs via Environmental Changes. IACR Cryptology ePrint Archive, Report 2013/619.

[9] Gabriel Hospodar, Roel Maes, Ingrid Verbauwhede: Machine learning attacks on 65nm Arbiter PUFs: Accurate modeling poses strict bounds on usability. WIFS 2012: 37-42

[10] Daihyun Lim: Extracting Secret Keys from Integrated Circuits. MSc Thesis, MIT,2004.

[11] P. C. Kocher: Timing attacks on implementations of Diffie-Hellman, RSA, DSS, and other systems, in Proc. CRYPTO',96, vol. LNCS 1109, pp.104-113 1996

[12] Dominik Merli, Dieter Schuster, Frederic Stumpf, Georg Sigl: Semi- invasive EM attack on FPGA RO PUFs and countermeasures. ACM Workshop on Embedded Systems Security (WESS'11), 2011.

[13] R. Pappu, B. Recht, J. Taylor, and N. Gershenfeld. Physical one-way functions. Science, 297(5589):2026, 2002.

[14] Ulrich Ru¨hrmair, Srinivas Devadas, Farinaz Koushanfar: Security based on Physical Unclonability and Disorder. In M. Tehranipoor and C. Wang (Editors): “Introduction to Hardware Security and Trust”. Springer, 2011.

[15] Ulrich Ru¨hrmair, Christian Jaeger, Michael Algasiner: An Attack on PUF-based Session Key Exchange, and a Hardware-based Counter- measure: Erasable PUFs. Financial Cryptography and Data Security 2011.

[16] Ulrich Ru¨hrmair, Frank Sehnke, Jan So¨lter, Gideon Dror, Srinivas Devadas, Ju¨rgen Schmidhuber: Modeling Attacks on Physical Unclon- able Functions. ACM Conference on Computer and Communications Security, 2010.

[17] Ulrich Ru¨hrmair, Jan So¨lter, Frank Sehnke, Xiaolin Xu, Ahmed Mahmoud, Vera Stoyanova, Gideon Dror, Ju¨rgen Schmidhuber, Wayne Burleson, Srinivas Devadas: PUF Modeling Attacks on Simulated and Silicon Data. IEEE Transactions on Information Forensics and Security (IEEE T-IFS), 2013.

[18] G. Edward Suh and Srinivas Devadas: Physical unclonable functions for device authentication and secret key generation. In Proceedings of the 44th Design Automation Conference, pages 9-14. IEEE, 2007.

[19] Blaise Gassend, Daihyun Lim, Dwaine Clarke, Marten v. Dijk, Srinivas Devadas: Identification and authentication of integrated circuits Con- currency and Computation: Practice & Experience, pp. 1077 - 1098, Volume 16, Issue 11, September 2004.

[20] Ahmed Mahmoud, Ulrich Ru¨hrmair, Mehrdad Majzoobi, Farinaz K- oushanfar: Combined Modeling and Side Channel Attacks on Strong PUFs. IACR Cryptology ePrint Archive 2013: 632 (2013)

[21] Ulrich Ru¨hrmair, Xiaolin Xu, Jan So¨lter, Ahmed Mahmoud, Fari- naz Koushanfar, Wayne Burleson: Power and Timing Side Channels for PUFs and their Efficient Exploitation. IACR Cryptology ePrint Archive 2013: 851 (2013)

[22] D. J. Bernstein. Cache-timing attacks on AES. Technical Report, 37 pages, April 2005. Available at: http://cr.yp.to/antiforgery/cachetiming- 20050414.pdf

[23] David Brumley , Dan Boneh, Remote timing attacks are practical, Proceedings of the 12th conference on USENIX Security Symposium, p.1-1, August 04-08, 2003, Washington, DC

[24] D. Merli, D. Schuster, F. Stumpf, and G. Sigl. Side-channel analysis of pufs and fuzzy extractors. In 4th International Conference on Trust and Trustworthy Computing (TRUST2011), Pittsburgh, PA, USA, June 2011. Springer.

[25] Daihyun Lim: Extracting Secret Keys from Integrated Circuits. MSc Thesis, MIT, 2004.

[26] E. Simpson and P. Schaumont: Offline Hardware/Software Authenti- cation for Reconfigurable Platforms. In L. Goubin and M. Matsui, editors, Cryptographic Hardware and Embedded Systems - CHES 2006, volume 4249 of LNCS, pages 311-323. Springer, October 10-13, 2006.

[27] M. Balliu and I. Mastroeni: A weakest precondition approach to active attacks analysis. In PLAS 09: Proc. of the ACM SIGPLAN Fourth Workshop on Programming Languages and Analysis for Security, pages 5971. ACM, 2009.

[28] Francesco Regazzoni, Thomas Eisenbarth, Johann Groscha¨dl, Luca Breveglieri, Paolo Ienne, Israel Koren, and Christof Paar: Power At- tacks Resistance of Cryptographic S-Boxes with Added Error Detection Circuits. In DFT, pages 508-516. IEEE Computer Society, September 26-28 2007. Rome, Italy.

[29] C.M. Bishop et al. Pattern recognition and machine learning. Springer New York:, 2006.

[30] Georg T. Becker, Raghavan Kumar:Active and Passive Side-Channel Attacks on Delay Based PUF Designs. IACR, 2014.