SYNOPSYS®
Silicon to Software™

# 真正安全系统的真随机数发生器

2015年12月

作者: 美国新思公司
现场应用工程师
**David A. Jones**

## 简介

大多数归类于“物联网”广阔范畴下的计算机、电子消费品和设备都需要达到令人惊讶程度的密码安全。即使是诸如PlayStation或Xbox游戏机之类的玩具和游戏，也具备非常复杂的各种设计预置安全特性。然而令人遗憾的是: 当密码或登录凭证遭到泄露，这些特性往往受到阻挠或被破解，从而导致经济损失(如: 失去音乐版税)、个人数据泄露(如: 医疗记录、家庭地址、或银行信息)或后续安全风险。

因此，越来越多的人开始关注电子产品设计师，他们中的许多人因装备不良而无法实现这些特性。就其性质而言，密码安全是复杂的且(大多数情况下)完全违反直觉的。当被问及实施标准安全特性时，即使是经验丰富的程序员、硬件设计师和工程管理人员也会发现自己有些不知所措，更不要说高端或关键任务设备所需的最先进密码安全。

随机数是大多数安全系统的核心。然而，从实践和效能方面来讲，生成随机数的方法却差别很大(对于实践而言尚好，但对效能而言则不很理想)。久而久之，许多流行的随机“噪声”生成算法和电路已被证明存在缺陷。随机数看起来似乎是，事实也是，可预见的以及/或者可被发现的。这样的缺陷会破坏安全系统的基础，导致弱点、后门或攻击媒介。更为糟糕的是，这样的弱点通常对产品设计师来讲并不是显而易见的，这就会让他们产生一种虚假的安全感。多年来，设计师可能察觉不到这些漏洞（如有）的存在。同时，所述的系统可能已经遭受反复攻击和利用。

本文将基于各种现实熵的不同源（即所谓的真随机数发生器--TRNGs）以及相关的攻击技术（包括物理、统计和电子法），对产生随机数的现有方法进行探讨。继而生成不易受攻击（更真实随机）数的新方法，以及采用标准数字电路和/或软件这样做的方法。最后，通过检查声称随机性的可验证性，以避免落入提供了随机性外表而未获实际效益的陷阱。

虽然加密算法、“黑帽”攻击、电子干扰以及安全性的种种其他方面不在本文的研究范围，但它将侧重于为产品设计和产品管理团队提供有关安全性基础知识的有用和实用知识。这些基本原则的关键是“信任链”概念。

## 安全性: 群体努力

人们通常用链条来比喻安全。“信任链”揭示了某一整体安全策略中若干完全不同且各自独立方面的相互联系，它们其中的每个要素均依赖于其他要素。如果系统安全性的任何一个方面受到损害，则整个系统就会受到损害。正如老话中所说的那样: 信任链的强度取决于其最薄弱的环节。

链条中的其中一环是密码技术。 密码、敏感数据、商业机密、专有固件、内存内容、受版权保护的材料均通常被加密处理。的确，设计团队有时会将加密与安全混为一谈， 并可能互换“加密”和“安全”的概念， 而事实上，前者只是后者的一个组成部分而已。

按照定义，密码系统是需要保密的。 这种保密性构成了系统密码安全特性的基础。如果秘密被泄露，那么密码安全就会遭到破坏，整个系统是脆弱的。 保密成为建立和维持总体安全的重中之重。

不管属于何种密码保密，都可被描述为安全链的一个基本环节。 它构成了所有后续加密元素，乃至整个系统总体安全的基础。

为确保有效，该等秘密必须满足四个条件：

1. 其必须是不可预知的
2. 其必须被均匀分布(即：任何数均有平等出现的机会)
3. 其必须统计独立 (即：与以前生成的随机数不存在明显关系)
4. 其必须被保密。

简而言之，它们必须看起来是完全随机的。

随机性比许多(或许大多数)开发人员所意识到的情形更难以实现。从人类的主观意识来讲，“随机”可以描述无法被普通观察者立即辨识的任何模式。在赌场中掷骰子是随机的， 它甚至满足了前三个标准(不可预测性、均匀分布和从一个到另一个的独立性)。然而，从1到6的数字中随机选择一个数字只会给电子安全系统奠定蹩脚的基础， 其原因在于：在第一次尝试中，仅有六分之一的猜对结果。此外，它也不是一个秘密。经过几次尝试后，安全性很可能会遭到破坏。

## NIST标准剔除臆测

“很难猜测”不是一个充足的随机性检验。对谁来说很难？一位数学家？一台电脑？还是一位拥有近乎无限技术、财源且受到强大成功激励的专门黑客？展示真正的随机性与简单忽悠一屋子工程师之间，存在着很大的差异。

为此，在美国国家标准和技术研究所（NIST）开发的NIST SP 800 - 22和SP 800 - 90（A至C的部分）中，定义了，被认为具有足够的随机性以用于加密应用之前，随机数发生器（RNG[1]）必须满足的统计分析标准。这些标准足以用来清除看似有效但实际上可能会存在破坏系统安全性统计缺陷的随机发生器。

服从于NIST SP 800 - 90的RNG的被推荐架构，旨在采用某一未知种子值来作为一个具有密码品质的伪随机数发生器(PRNG)的种子，并使用PRNG一段时间，或生产一些随机数据。PRNG将被补种，并再次暂时使用一段时间，等等。PRNG的种子应为一个来自于某一“熵源”的保密、随机输入。高质量的TRNG是PRNG的首选输入设备。经适当建构的随机数发生器（TRNG）从一些随机过程中收获熵(如：晶体管中电流产生的噪音，或放射性衰变事件之间的时间)，继而支配熵信号，以消除偏离，并白化输出结果顺序的光谱。当然， 在设计不受外来者（试图扰乱操作）影响的某一适当随机数发生器（TRNG）电路时，有必要控制诸如工作温度、老化、易受电子噪声和扰动等因素。

NIST还发布了联邦信息处理标准(FIPS)140 - 2，它描述了用于验证、测试和监控此等系统的一种客观方式。这些标准免去了在对某一随机数发生器质量进行评估时的臆测性。

这些标准没有描述如何创建一个随机数发生器（TRNG）;它们仅仅提及了如何验证其是否工作。这些实施细节给设计师发挥创造力留出了空间，因此，允许存在许多替代方法。然而，在所有情况下， 随机数发生器（TRNG）必须满足上述四个条件： 必须具备不可预测性、均匀性、独立性和不可发现性。

## 攻击媒介

有很多方法可危及或攻击一个安全系统。许多方法完全是违反直觉的，几乎有资格媲美爱因斯坦的“远距离的量子纠缠。”例如：仅通过观察电脑发出的射频能量，就可以从电脑中抽取用作密匙的随机数。甚至在没有触碰机器的情况下，黑客就能够从稀薄的空气中，正确揭示用于加密和解密的密钥。整个过程在目标机上不会留下任何痕迹，当然，除了标准台架试验设备之外，其他都很容易被复制。

遭受临近攻击也是可能的。故意在微处理器或SOC的I/O引脚上施加小故障就可以导致芯片锁定或偏离其预定的路径，从而使其容易受到示波器探头、逻辑分析仪和恶意软件的影响。这种攻击也可以破坏被转移的数据值（无论在芯片外或是芯片本身），或改变特权级，降低内置的硬件防御。

密封的芯片可被解封，暴露其存储单元、门配置、引脚、导线，并在某些情况下，会暴露非易失性存储器中被编程的秘密关键码值。只读存储器可被转出，总线可被监控。基于简单功耗分析（SPA）或差分功率分析（DPA）攻击中只不过是系统的超限时间电流消耗，一些攻击就可以依靠监控电源，梳理出电路行为。故障植入式攻击可向某一系统馈给一系列精心设计的值，以破坏电路的正确操作。在一种被称为定时分析的技术中，个别操作的执行时间中存在的数据依赖性被用来确定机密数据（如加密密钥）的秘密数据值。这些攻击连同许多其他看似不可能的攻击，均已被记录在案、并被证明和重复。一名专门的黑客在破解一个所谓的安全系统方面，几乎不花费任何努力。这些攻击依赖于许多系统中存在的“可发现性”弱点。即使原来的随机数可能是未知的，也可以通过观察一些看似无关的现象，从而间接地发现这些随机数。

---

[1] NIST将这些称为随机位发生器， RNG则是较常见的俗称。

## 可观察性与理解

最后，还有另一类攻击媒介：操纵。值得注意的是：操纵攻击无需了解安全特性的工作方式。只需干扰它们的功能，经常通过采用一些技术含量非常低的方式。例如：像用廉价吹风机加热芯片这样简单的事情就可能会使芯片的性能表现发生变化，这也许是因为在其一些电路中，传播延迟发生了改变。这种低技术含量的攻击可能不会揭示任何有关系统的加密方法(虽然它也可能会这么做)，但它没必要去做。只要它改变了系统的行为，就可以被用作武器。

许多这样的攻击属于令人胆怯的“离散熵的贝叶斯估计”。只要获得足够的样本，黑客就可以推断出某一系统的运作方式（即使他从未完全理解系统的工作方式）。这是一种“船货崇拜”现象的现代翻版。如果我这样做，那就发生了，即使我不知道其原因。因为大多数电脑运行非常快， 所以，在合理的时间内，不难积累和记录数以千计、成千上万、甚至上百万的相关数据样本。通过采用那种类型的数据进行分析，黑客可以推断系统的工作方式—或者即使不是那样，至少也知道如何去影响系统的行为。

聊以自慰的是，设计师们可以通过使数据收集费用过高，而使他们的系统获得有效的安全（即使这些系统并不安全）。贝叶斯分析取决于积累可用于分析的大型样本数据库，以确定相似性和统计上的显著差异，所以，关键在于减少或消除使这些相关性有意义的机会。请注意：通常情况下，不允许减少加密事件本身（存在产生潜在有害样本的电路运作）的数量，因此，掩盖他们的行为—保守秘密—是最为重要的。如果做一次彻底分析需要五年，以积累足够的样本，大多数黑客将会在发现有用东西之前离开。

同样地，如果解封芯片以及在电子显微镜下扫描复杂的ASIC层需花费数千美元，那么，很多黑客将不会在意微薄的经济回报。安全锁的大小应该与财宝的规模相称。就个人而论，能够提供访问安全内容的电视机顶盒可能在经济上缺乏吸引力，但却给能够破解数百万机顶盒密钥的黑客们提供了可观的奖励。阻挠批量生产媒体播放器单一模型的数字版权管理(DRM)可能提供配得上耗时逆向工程的重要动力，如果同一黑客致力于研究所有单元。所以技术战略的一部分应在于减少激励。最重要的是：破解一个设备的加密密钥不应该有助于对第二台设备的破解，等等。

## 随机性的悖论

让人觉得不方便的是：安全系统通常需要频繁生成随机数， 而这与希望最大程度减少可观测事件（揭示可衡量侧信道数据）背道而驰。许多系统连续不断地加密和解密数据。甚至一个直接的的VPN连接都会加密和解密在两个方向流动数据的每个字节， 仅在数分钟之内就会产生数十亿次潜在可观察到的安全操作。

为强化密钥的安全性，系统应在每次需要生成密钥时，理想地使用不同的随机数“种子”。由于从不重复使用相同的种子，该系统降低了所有最重要密钥的暴露和寿命，从而减少了暴露给黑客的“攻击面”。不管怎样，应频繁地重做种子，以限制种子的曝光量和基于其给定值所生成的数据量。

那么，该如何去调和这两个要求呢?一方面，应该生成随机数并频繁更换；另一方面应对这样的活动保密。一个安全系统将理想地拥有完全不可见观测的随机数发生器（TRNG），可发出任何非计划的侧信道数据(射频发射、功率波动等)，同时生成真随机数。

## 从可预测性推导随机性

从数字电子系统中产生随机数是非常困难的。计算机是可预测的。事实上，决定论是界定数字电子电路和二进制系统的基本特征之一。从根本上讲，模拟电压和电流信号被转换为二进制数字信号，精确地消除来自模拟信号的不确定性。在这些电路中产生真正的随机性，而不仅仅是显而易见的随机性，是非常复杂的。尽管许多方法都试过了，但它们中的有不少方法均已被证明具有可被利用的漏洞。

许多早期的随机数据生成器(RNGs)依赖基于时间的熵而产生随机数。也就是说，计算机使用了当前时间（包括秒、分、十分之一、百分之一，也许数以千计的一秒），并在数学上结合这些数字，表面上随机的数。由于结果是通过始于随机种子的确定性的过程而生成的，因此，它可被适当地归类为一个伪随机数生成器（PRNG）。像所有的数学函数一样，输出完全取决于输入。考虑到相同的输入(即种子 、时间)，伪随机数生成器（PRNG）将产生相同的输出。虽然结果对漫不经心的观察者来讲是随机的，但它简直像时刻一样是可预测的。

请注意：它跟其余系统安全功能如何先进没关系。以时间为种子的伪随机数生成器（PRNG）是众所周知的薄弱环节。同样地，对于其他的数字源（作为种子）也应持怀疑态度，如磁盘或者网络控制器的中断率，特别是在攻击者可以影响它们的情况下。因为种子是可预测的（在这种情况下，其熵的来源），加密措施在独立性测试中失效，所以整个系统会受到损害。

## 利用技术中存在的缺陷

或许一个更好的熵源可以解决这个问题。许多随机数生成器（RNG）依赖半导体电路物理特性中的微小差异，作为熵源。例如：芯片某一部分的电容，或给定布线的电阻，或未初始化存储单元的随机内容。复合半导体中存在着很多不可测知的物理怪癖，可被利用，作为熵源。尽管半导体是大规模生产并被复制到难以置信的公差，但有些差异依然存在。这无疑是一个真正的随机起点。

这种方法有其优点，也有其缺点。批量生产零件中的物理差异确实存在，但它们得靠运气。也就是说，它们不是预置在设备之中;它们仅作为生产过程的一种自然的副作用。电路设计人员可能无法依赖于那里存在的各种“缺陷”， 或这些缺陷将会一种有用的方式出现“故障”。这种物理熵得看运气。

这样的缺陷也可能在可观测性测试中失败。由于制造差异是电路的物理性质，故可采用显微镜进行视觉扫描或用仪器进行测量。这很容易理解：如果RNG本身可检测出电路缺陷， 那么，则用其他电路也可检测出来，这一点是毋庸置疑的。

在实践层面，基于物理怪癖的熵很难复制。这是蓄意的，但它却给大容量设备的创设者们带来了一些问题(也就是说，他们中的大多数)。物理缺陷非常特定于个体设备、半导体过程节点、制造技术、硅供应商、制造过程的化学作用和其他因素。这个过程在从加工车间到加工车间或者从设计到设计过程中是不可扩展、不可移植的，其原因在于其目的不在于此。

## 对策

为阻止基于可观察到现象的侧信道攻击， 随机数发生器（TRNG）最好进行“无形”操作， 不辐射与内部活动相关的射频能量，不影响主机设备的电源需求，不暴露输入输出脚的未加密(甚至加密)数据，甚至设备内部的“引脚”。不幸的是，如此完美的需求模型与物理性质背道而驰。所有的电子设备都会产生可观察到的现象，如果不这样做的话，它们将会变得毫无用处。

实际的回应是屏蔽随机数发生器（TRNG）的行为。这通常被称为“白化”电路，类似于产生白噪声，以屏蔽机械系统的声音。适当的白化对策可有效庇护随机数发生器（TRNG）以避开基于可观察到现象的侧信道攻击，如上述所提及的。从外面看，无法获知随机数发生器（TRNG）的运行情况、其正在处理何种数据，甚至其是否正在运行。对于所有的意图和目的，其行为均采用完全相同的方式，与密钥长度、熵源或活动水平无关。

为了进一步阻止基于可观测性的攻击，随机数发生器（TRNG）应不对时间、热、电压和其他输入现象敏感。作为数字电路，其传播延迟不应与密钥长度或加密复杂性相关。换句话说，其应精确地花费X周期或某个数量的时间，以产生一个新的随机数，即便在不同输入参数的情况下，亦应如此。这是难以实现的，其原因在于：大多数设计师青睐于能够容纳各种加密算法，或可以与不同密钥长度一起工作（128位密钥、1024位密钥等）的随机数发生器（TRNG）。

显然，任何随机数发生器（TRNG）应依靠真正随机熵源（而非时基或用户提供密钥），并通过预见性测试。它可通过服从NIST 和FIPS，而非电路分析（取决于人类预想、偏见和彻底的臆测），加以验证。其主要部分为：该电路或其周围的子系统必须监视其正确行为，并不断测试电路的操作，如果检测到随机性偏差时，能发出错误信号。

### 新思公司的DesignWare真随机数发生器

新思公司以可综合的IP形式，开发出其自己的真随机数发生器。换言之，电路中不存在必须针对每个工艺节点、芯片设计、晶圆厂或供应商进行人工调整的硬宏元。相反，它是一个易于综合的IP硬件设计语言块，可被集成到任何数字电路中。

- DesignWare®真随机数发生器(TRNG) —在NIST命名法中，被分类为 “带电、受制约的数字化噪声源”， 该IP结合了一个带噪音源的白化电路，可被用于播种一个随机数流，以及提供一个不间断的熵源。
- DesignWare真随机数发生器(TRNG) NIST SP 800-90c兼容 —在NIST命名法中，被分类为 “带电、增强型非确定性随机位发生器”， 该 IP结合了一个NIST SP 800-90b认可的带噪音源的调理电路，以及一个NIST SP 800-90a认可的确定性随机数产生器(DRBG)。

## 结论与观察

- 安全和隐私是关键的设计考量，即便对于简单的消费设备而言，亦是如此。这两种特性均依赖于加密。
- 各种加密功能实现成功是非常困难的，其原因在于：它们往往是违反直觉的，并依赖于诸如射频辐射、贝叶斯分析、数值独立性以及复杂的数论。这些以及其他因素通常会把刚入行的开发者误导进入实施薄弱的安全系统（仅能起到掩饰自身无效性的作用）。
- 许多系统和应用程序的安全强度依赖于熵源的质量，如：真随机数发生器（TRNG）不能被任何数量的物理、电子或统计攻击所阻止、观察或利用。
- 业界已制订出国际标准，以采用可验证和统计严谨的态度，验证了某一真随机数发生器（TRNG）的真随机性质。
- 新思公司已开发出通过NIST FIPS规范、且适用于任何数字半导体器件的真随机数发生器（TRNG），而不考虑过程节点。


新思公司 • 690 East Middlefield Road • Mountain View, CA 94043 • www.synopsys.com

01/17/16.TT_CS6806_TRNG_WP_Chinese.