硬件木马检测与防范研究综述

许强[1]，蒋兴浩[1]，姚立红[1]，张志强[2]，张诚[2]
（1. 上海交通大学电子信息与电气工程学院，上海 200240；
2. 上海机电工程研究所，上海 201109）

摘　要：集成电路芯片的设计和制造是当今电子化产业中关键的组成部分，第三方技术服务的日益普及可能导致电路芯片在制造过程中被植入硬件木马，这对电子设备在安全性上带来了很大的挑战。在简单介绍硬件木马的概念的基础上，分析了硬件木马的组成与特点，然后重点探讨了现阶段硬件木马的检测和防范技术，最后对硬件木马的发展趋势进行了总结。

关键词：集成电路芯片；硬件木马；检测；防范
中图分类号：TN972
文献标识码：A
doi: 10.11959/j.issn.2096-109x.2017.00160

Overview of the detection and prevention study of hardware Trojans

XU Qiang[1], JIANG Xing-hao[1], YAO Li-hong[1], ZHANG Zhi-qiang[2], ZHANG Cheng[2]
（1. School of Electronic Information and Electrical Engineering, Shanghai Jiaotong University, Shanghai 200240, China;
2. Shanghai Electro-Mechanical Engineering Institute, Shanghai 201109, China）

Abstract: The design and manufacture of integrated circuit chips are now a key component of the electronic industry. The increasing popularity of third-party technology services may lead to the implantation of hardware chips in the manufacturing process, which brings great challenge to the security of electronic devices. After introducing the concept of hardware Trojans, the characteristics and its forms were briefly analyzed. Then the detection technologies, and the prevention strategies of the hardware Trojans were discussed. At last, the development trend of the hardware Trojans was summarized.

Key words: integrated circuit chips, hardware Trojans, detection, prevention

1　引言

随着专用集成芯片（ASIC，application specific integrated circuits）设计以及 IP 重用技术的不断进步，集成电路设计制造技术得以迅猛发展，芯片功能水平也得到极大的提升，其在民生生活、军事攻防、医疗卫生等各个领域应用日益广泛。集成电路芯片从设计到最终成型的制造主要包括应用系统设计、系统开发、原型验证、代工厂制版、圆片制造测试、封装成型等工艺，在经济全球化的条件下，芯片、半导体设计与制造方式也逐步全球化，同一块芯片要完成不同的工艺或许要经过多方安全性未知的制造商来完成，在这些环节中，芯片电路可能被竞争者植入硬件木马，以达到芯片遭到恶意篡改和攻击控制的目的，这给集成电路芯片的安全性带来了极大的挑战。

收稿日期：2017-01-24；**修回日期**：2017-02-27。**通信作者**：蒋兴浩，xhjiang@sjtu.edu.cn
基金项目：航天先进技术联合基金资助项目
Foundation Item: The Aerospace Advanced Technology Jointly Funded Project

集成电路如果感染硬件木马，会导致功能和规格的变化，造成敏感信息的泄露，甚至造成系统的瘫痪，这很有可能威胁到国家军事系统、交通、金融、医疗等方面的安全。现阶段，很多研究都指出硬件木马对原始电路的危害性，Yang 等[1]验证了无论是在模拟电路还是数字电路中，远程攻击者都能通过规模很小的木马电路对电路进行控制。文献[2,3]的研究也表明硬件木马能够实现对专用集成电路、微处理器、微控制器、网络处理器、数字信号处理器等硬件的修改以及对 FPGA 比特流的修改。硬件木马设计的隐蔽性是其与现阶段制造缺陷最大的不同点，制造缺陷是随机生成并且非刻意的，它含有特定的缺陷模型，如延迟缺陷模型等。对硬件木马而言，很难生成一种对所有硬件木马都适配的模型，且制造缺陷完全是制造过程中产生的，而硬件木马可以在集成电路设计、制造、大规模复制期间产生，因此硬件木马问题更具复杂性。

随着半导体工艺的不断发展，电路逐渐高集成度，尺寸纳米级别化，在芯片中植入的硬件木马只有在特定条件下才可以被激活，其余时间对原始的电路功能不影响，因此硬件木马被检测出来的难度非常大。硬件木马是一类实体电路，在原电路中一旦被植入便会长期存在，只有通过合理的检测手段对电路进行木马检测才能使风险最小化，近年来，关于硬件木马方面的研究越来越多地引起国内外学者的关注，对硬件木马检测及防范技术的研究已经到了刻不容缓的地步。有效的硬件木马检测方法以及防范策略对保证芯片安全、稳定电路系统正常运行至关重要。

2　硬件木马的概念

2.1　硬件木马的定义及产生

硬件木马是指在电路设计阶段对电路进行恶意、蓄意更改，使电路系统在某种特定触发条件下产生一些不希望出现的响应[3]。硬件木马的生成一般是由于硬件制造商为了缩短生产周期、降低制造成本，在芯片制造生产过程中使用了不可信的设计工具或依赖不可信的制造商，如图 1 所示，这些不可信的元素细分如下。

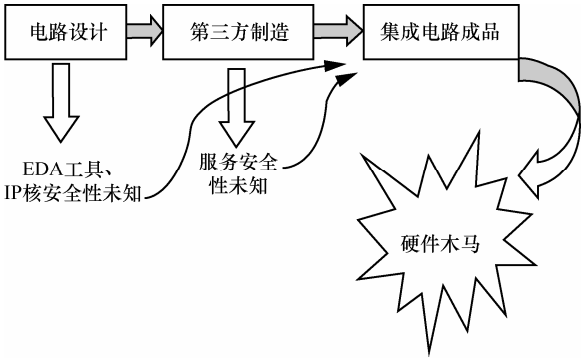


图 1　硬件木马的产生

1) 不可信的 IP 核、SOC 开发商、外包商。电子线路的高度集成使集成芯片开发商无法自行设计所有的 IP 核，因此，需要利用第三方 IP 核进行集成电路的开发制作，第三方 IP 核能够携带硬件木马，从而造成集成电路遭遇硬件木马攻击。同理，SOC 开发商、外包商也有可能安插硬件木马到集成电路中[4]。

2) 不可信的成品部件。如今越来越多的公司、机构利用成品部件来完成系统的功能，成品部件往往因其能够实现某一特定功能而在电子市场泛滥。当然，这也是木马设计者的可趁之机，在成品部件上植入硬件木马，进而侵占整个系统。

3) 不可信的系统集成。许多半导体公司都能提供面向应用的集成电路设计以满足不同客户的需求，客户可以规定他们 SOC 设计所需的特定的 IP 核，芯片被制造、测试、包装后送至客户方，一些公司拥有自己的制造产业以及设计团队，他们同时提供芯片设计方面特殊的制造服务[4]。

2.2　硬件木马的组成

硬件木马在集成电路的设计过程中，为了实现在特定触发条件下改变系统运行状态，影响系统正常运作，泄露系统信息以及使系统瘫痪的目的，对原有电路系统进行人为的恶意改动[5]。硬件木马由触发模块和负荷模块组成[3]：触发模块能够激活木马电路；负荷模块是木马激活后所攻击的电路模块。硬件木马的触发条件往往是基于一个小概率事件来实现的，这种小概率事件通常是一个很少发生的信号或信号的组合。

图 2 是常见的硬件木马基础模型，正常情况下硬件木马没有触发，系统在输入信号的激励后经宿主电路信号处理模块后输出正常信号。在特

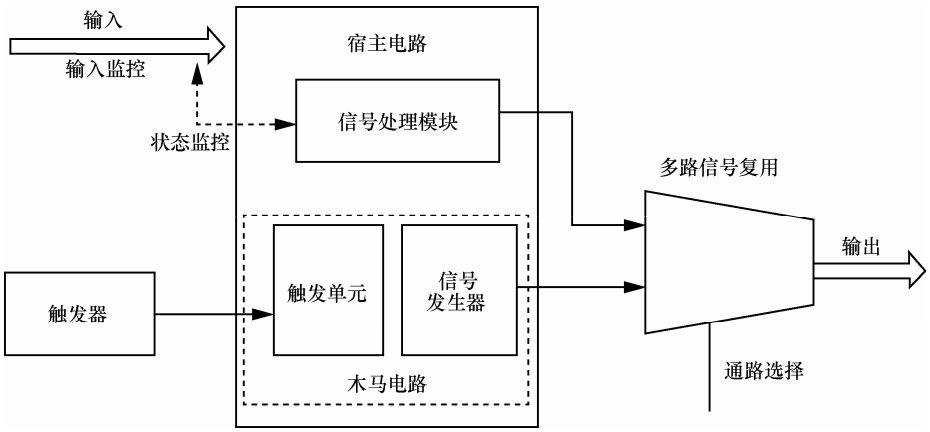


图 2　硬件木马模型

定的条件下，触发单元检测到特定的输入，触发器激活，硬件木马被触发，在多路复用器的作用下，输入信号不经过标准信号处理模块，而是在木马功能模块处理后从输出端口发出恶意信号。

2.3　硬件木马的分类方法

目前，已有很多研究都对硬件木马做了详细的分类，本文综合现阶段已有的分类方法并结合硬件木马的各方面特点对硬件木马进行了如图 3 所示的分类。

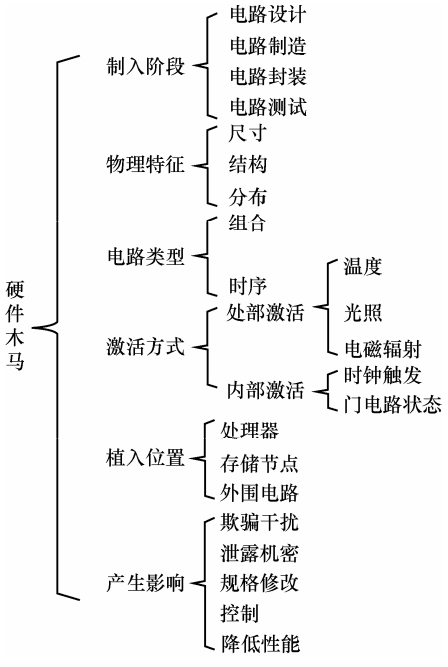


图 3　硬件木马分类方法

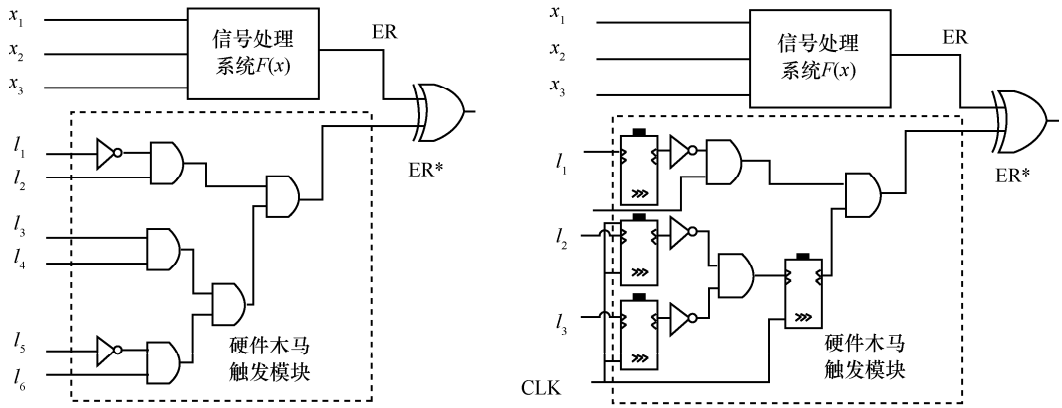
硬件木马的植入可能处于不同的时期，可以在电路的设计、制造、封装和测试等各阶段来植入，因此可以根据其植入方式进行类别划分[4]。

此外，硬件木马在物理层面各不一样，木马物理特征可以分为尺寸、分布、结构这 3 类，其中，尺寸是根据木马在电路中添加、删除或损坏芯片的数量来划分；分布是指木马电路在宿主电路分布的情况，可以是物理布局的紧致程度，也可以是在原始电路上布局的位置。硬件木马在植入过程中，电路会进行二次布局，这会导致芯片的规格和结构发生变化，因此这些物理特征也可以作为木马的分类标准。

Karri 等[3]按硬件木马的 5 种特性（插入相位、抽象层次、触发机制、影响及位置）对其进行分类；根据信号类型也可将硬件木马分为数字型和模拟型；根据触发方式对信号的依赖情况可分为外部触发木马和内部触发木马，外部触发木马仅依赖于外来信号的干预[2]。

Banga 等[6]将常规的硬件木马根据电路类型分为组合型和时序型，组合型硬件木马的触发模块不包含寄存器电路，一般由门电路构成，如图 4(a)所示；而时序型硬件木马的触发模块包含寄存器，触发条件受时钟和输入信号控制，如图 4(b)所示，当特定信号到来时，硬件木马被触发，导致芯片输出由 ER 变为 ER*。

Bhunia 等[7]在木马触发和负载的基础上，将其分为数字型木马和模拟型木马，数字型的木马包含组合型和时序型 2 类，模拟型木马通过传感器感应外界条件变化来触发。Zhang 等[2]根据硬件木马对电路功能产生的影响将木马分为故障型和寄生型 2 类，故障型硬件木马主要通过改变原始电路影响系统的正常功能，寄生型硬件木马



(a) 组合型硬件木马　(b) 时序型硬件木马

图 4　组合型和时序型硬件木马

一直隐藏在原始的宿主电路中，它在特定情况下才能被激活，未激活时不影响原始电路的政策功能。研究几乎都是按照硬件木马的物理特性、触发方式、响应模式以及对原电路的影响进行分类的[3,5]。

除了常见的按电路类型、激活方式分类外，硬件木马的植入位置以及危害都可以作为其分类的标准，不同的硬件木马为了达到其特有的功能可能被植入芯片的不同位置，可以是在处理器、存储节点、外围电路中。理解不同类型的硬件木马工作机制有利于今后检测和防范木马技术的实施。

2.4　硬件木马的特点

自从阿格拉沃尔于 2007 年首次发表关于硬件木马的文章后，该方面的研究取得了很大的进展。很多研究都表明攻击者为了实现对电路系统不同方式的攻击[8~10]，多种木马电路被设计出来。在设计阶段，设计者往往会考虑如何保证硬件木马的功能散化、低激活、高可靠以及隐蔽性，以此保证攻击能够实施并且不易被察觉。

通过分析硬件木马设计思路及其产生的影响可以发现，硬件木马同软件木马一样具备以下特点。

1) 隐蔽性强：隐蔽性体现在其植入芯片的位置不易被发现，自身逻辑器件很小，并且木马触发的概率极低。

2) 设计灵活：硬件木马触发方式以及植入方法的多样化也使其设计方式很灵活。

3) 破坏力大：一旦木马被激活，能够对原系统造成控制、窃取资源、瘫痪的危害。

4) 设计要求高：保证木马芯片的功能散化、低激活、高可靠性以及隐蔽性。

5) 防护检测难度大：没有统一的检测模型，现有的检测方法耗时耗力，效率低。

3　硬件木马的攻击模式

硬件木马的工作方式灵活多样，它通常通过监测原始电路的输入、外界条件刺激、系统总线内容来触发攻击，如果电路的状态不满足其触发条件，则电路正常运行输出；当触发条件到来时，如触发逻辑电路接收到特定信号或频率传感器探测到特定频率的信号等，硬件木马电路被激活并实现特定功能，如欺骗干扰、控制、拒绝服务、密码泄露、物理摧毁等。硬件木马的工作原理如图 5 所示。

目前，关于硬件木马攻击方式的研究没有统一的模式，根据各个应用环境的不同，相应的模式也不一样[11]。由于系统芯片的制作流程可以被划分为 IP 核设计、系统芯片设计、成品制造这 3 个流程，每一个流程对应的开发设计方都有可能实现木马的植入，木马的攻击模式也相应的不同。下文将详细介绍目前已有的木马攻击模式，通过分析这些模式的组成特点，为新模式的提出提供思路。

3.1　基于旁路型的攻击模式

旁路型的攻击模式是近年来提出的一种新型攻击手段，不同于其他硬件木马的原理，它利用密码算法分析目标芯片的功耗信息、延时，以及其他物理信息来破解密钥。近年来，关于旁路型

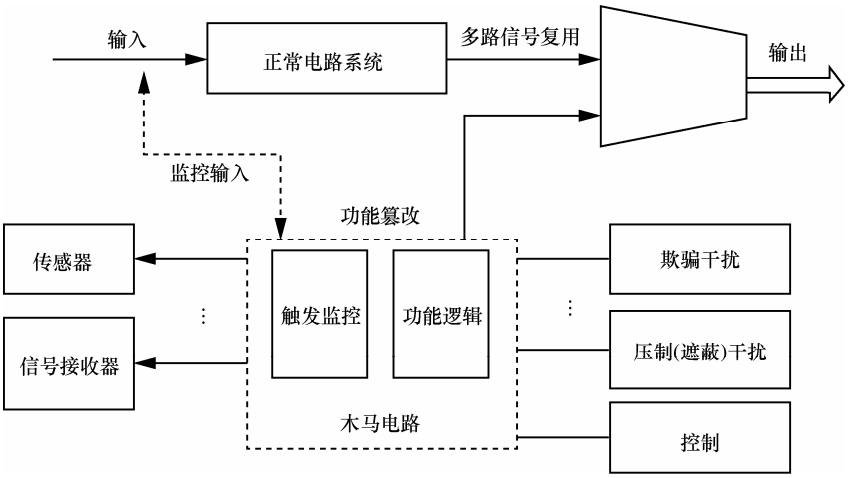


图 5　硬件木马工作原理

攻击模式的研究也层出不穷。

电路的电磁辐射、功耗延时作为电路的固有特征，成为旁路型木马设计的重要依据。谢海等[12]设计了一种电磁泄露型硬件木马，该木马电路能通过电磁发射的方式泄露密码芯片的密钥，通过搭建基于现场可编程门电路 FPGA 的 AES（advanced encryption standard）加密电路测试平台，利用旁路分析的方法，对此结构类型的硬件木马电路进行检测与分析，发现该硬件木马电路能在使用者毫不知情的情况下成功获取 128 bit 的 AES 加密密钥。Lin 等[9]提出一种基于 FPGA 芯片的功率旁路型硬件木马模式，通过监控功率信息，触发木马，从而达到泄露芯片隐密信息的目的，该攻击模式主要针对功耗信息及差分分析技术来逐位检测获取密钥，邹程等[13]在此基础上对其进行了改进，利用 FPGA 芯片运行时的功率旁路泄露，有意形成功率旁路通道以达到泄露电路私密信息的目的。

除了利用电路功率、延时特征，一些其他的特性（如 LED 闪烁频率等）也可作为旁路型木马设计的思路，此类木马攻击也是旁路型攻击中的一种重要模式。Baumgaten 等[8]提出一种利用 LED 闪光频率作为触发信号的木马电路，它需要一块特殊的旁路电路来监控 LED 灯的闪光频率，灯闪动频率一般位于人眼无法察觉的 1~2 kHz，并且监控电路的位置应尽可能布置在离 LED 灯较近的位置，从而方便其获取 LED 灯的信息。Jin 等[10]利用无线信道特征（如幅度、频率等）在数字模拟混合无线密码集成电路中实现了无线旁路硬件木马的植入。Zhang 等[2]针对集成芯片制造过程中可能被嵌入恶意硬件木马的问题，利用扩频通信的原理，对芯片中的密钥或密码数据进行扩频调制，从而得到器件噪声级别以下的电磁泄露旁路，然后利用相关性来提取芯片中的秘密信息，对 AES 电路进行木马攻击实验。同样地，针对 AES 加密电路，吴志凯等[14]提出了一种基于少态的硬件木马植入方法，通过利用在门级、网表级各节点翻转次数统计结果的少态实现硬件木马的触发逻辑，完成低面积开销、低旁路功耗增加的硬件木马植入，大大减小了硬件木马被检测的可能性。

3.2　基于时间有限状态机触发的攻击模式

有限状态机是指系统的输出仅取决于当前和过去输入部分的时序逻辑电路，是一种包含有限个状态以及在这些状态之间转移和动作等行为的数学模型。与普通组合型木马不同，基于时间有限状态机触发的木马并非在木马接收到某一单独输入序列就被激活，而是考虑时间因素在内，即除了各输入在阈值上存在范围限制，在时间上应该符合某种顺序关系，使触发逻辑空间增大，降低触发概率，保证硬件木马的隐蔽性[5]。

王晓晗等[15]利用线性反馈移位寄存器生成的最大周期递归序列作为木马的激活序列，以在密码芯片中注入故障作为攻击手段，设计了一种规模可控的硬件木马电路。在 FPGA 芯片上实现 AES 加密电路中植入木马。线性反馈移位寄存器的本质是一种有限状态机，利用二元 n 级寄存器

（每个寄存器为线性反馈移位寄存器的一级，只有 0 和 1 这 2 种状态）存储当前时刻有限状态机的状态，并通过反馈、移位等操作生成下一时刻的状态。李蕾等[16]提出一种基于有限状态机的硬件木马攻击方法，有限状态机作为木马的触发模块，通过监测输入信号和目标电路内部节点决定下一状态，负荷模块包含 1 个两输入与门电路，2 个输入分别来自触发器和目标电路，通过ISCAS89基准电路中的 S349 作为目标电路，对电路功能和延时信息进行仿真，结果显示，这种攻击模式能够提升木马的激活难度，并且可以有效隐藏延时信息。

3.3　基于含攻击后门的处理器的攻击模式

硬件安全专家 Damien Zammit 指出，在新款的 Intel x86 处理器内部藏有隐蔽子系统 ME，它作为一个单独的处理器在 Intel 处理器内运行，计算机使用者一般无法禁用，并且无人可以查看它封闭的专有代码。文献[17]发现，为处理器保留隐藏后门的情况确实存在，单个组件隐藏在数亿的器件之间，芯片设计者无法看见，也不能确定它是否是芯片制造商添加的，这成为硬件木马寄宿的良好场所。常见的硬件木马都是针对具体硬件或防御体系设计的，基于含攻击后门的处理器的攻击模式也应运而生。文献[4]提出了一种含攻击后门且支持多方式攻击的处理器，文中提到的硬件木马并非针对某一硬件来插入，而是设计了 2 种基于硬件的攻击方式进行多种攻击。

随着硬件电路设计的多样化，电路的功能、结构以及应用场景变得越来越丰富，不同类型电路对应的电路特征也各不相同，通过分析现阶段不同的硬件木马攻击模式，能够把握硬件木马的应用特点以及规划未来的研究方向。

4　硬件木马检测

硬件木马的出现给集成电路安全带来了极大的挑战，硬件木马种类多样，设计方法和功能都没有特定的限定，且其植入方式不尽相同，因此，硬件木马的检测十分困难，如何检测和防范硬件木马成为一个很关键的问题。现阶段的研究中，很多木马检测技术被提出，从最早的基于失效分析的检测方法到如今基于旁路分析的方法，不同的检测技术有其各自的优势与特点，对不同的检测方法进行分析能够把握检测技术的发展方向和思路，对保证集成芯片的安全意义重大。

4.1　检测方法分析

硬件木马的植入方式有：直接存于第三方 IP 核、插入在空余空间内、存于 RTL 级代码中，不同的植入方式功能各不相同，因此对硬件木马的检测难度很大。现有的硬件木马检测技术主要有逆向工程、物理检测、基于反向解剖检验、旁路分析检测和功能测试。目前，很多学者对硬件木马检测技术根据其检测实施阶段和方法的不同进行类别划分。倪林等[11]根据这些检测方法的特点，将其科学地分为电信号检测和非电信号检测两大类。Xiao 等[4]根据检测对象的呈现方式将这些检测技术分为晶片木马检测及硅前木马检测两类。本文按照不同的检测手段，将检测技术分为对原电路有破坏和无破坏两类，如图 6 所示。

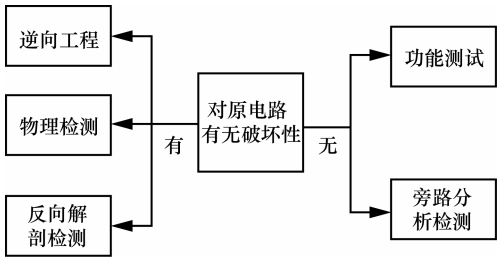


图 6　硬件木马检测技术分类

4.1.1　破坏性检测

现今，利用光学方法检测元器件异常的技术已经成为可能，为了对芯片进行硬件木马检测，必须对芯片进行逐层拆分，并且只有去除上层模块才能到达底层模块，进而对底层进行线路分析，然而经过拆分后，会导致芯片的损坏。对芯片的破坏性检测能够有效地检测出芯片是否被硬件木马感染，但这一结果只对被测芯片成立，不能确定未测芯片是否被感染。

1) 逆向工程

现阶段的很多硬件木马检测技术往往需要事先保证金片的存在，利用金片来对安全性未知的芯片进行对比分析，而发现金片的过程往往需要进行逆向工程的处理，集成芯片的逆向工程是分析芯片内部构造、连接的重要一步，通过逆向工程，工程师能够确定芯片是如何设计和运行的，通过分析芯

片结构能够更加有效地对硬件木马进行检测。

逆向工程一般包括以下几步。①解封装：去除芯片外部晶粒。②结构平整化：为了保证芯片表面平整，利用化学方法一次性去除每一层的晶粒。③成像：利用电子扫描显微镜对每一裸层拍摄上千张高分辨率图像，所有图像将利用特殊软件工具进行拼接，最终形成对芯片裸层的完整呈现[9]。对于不同的裸层，层与层之间必须对齐，并保证上下间隙以及接触点的位置要精准。④注释：系统中包括连接线、晶体管在内的所有器件都必须利用图像识别软件或人工进行标注。⑤原理图的搭建、组织和分析：利用带标注的图像以及数据表在内的共用信息形成层次化原理图。以上步骤时间代价高且易出错，前 3 步本质上是对集成芯片结构图像进行提取，后 2 步在电路设计方面度量同样也具有很大的挑战[5]。

Xiao 等[4]研究发现，现有的木马检测技术利用逆向工程来解包芯片，逆向工程能够百分之百保证检测到芯片中的恶意篡改，但它的时间代价太高，检测复杂度一般的芯片要花费几个星期到几个月。Bao 等[5]为了提升逆向工程效率，减少寻找安全芯片的时间代价，提出了一种新颖、具顽健性的逆向工程方法来验证芯片没有被植入木马。该文采用机器学习方法，利用解封装、减层级、成像 3 步来获得集成电路每一层的原始图像，然后对图像进行无重复的分块，并对每一层次的块进行特征提取，随后对分类器进行训练，获得每一层的决策边界，训练完成后，利用 SVM 每一层的决策边界将芯片每一层的块进行分类，最后对分类块的芯片定义标签，如图 7 所示。

通过利用现有工具在公共电路上测试的结果显示，对于不同的模型，该方法能够高效、高精度地检测出硬件木马，与传统逆向工程需要人工参与的方法相比，该方法能够实现自动化检测并且在计算和存储资源上更加有效。

2) 物理检测

物理检测的原理很简单，它利用电子仪器（如显微镜、电路分析仪等）观察电路裸芯片与原电路是否存在差异。芯片裸化的程序十分复杂，需要经过机械打磨、拆分等工艺，是一种具有破坏性的检测方法，Bhasin 等[18]利用光学器件检测芯片是否被制造商植入硬件木马，通过显微镜光学成像设备获取硅片图像，并与 GDSII 布局数据库视图中的原始图片进行对比，同时运用了形状识别、自动识别、相似度判别工具，结果显示金属层面的单一更改能够很容易被检测。物理检测方法比较适用于结构相对简单的芯片，对于结构复杂的芯片，这种检测方法耗时较长，检测的成本很大，且当芯片的数量较多时，这种检测方法显得有些捉襟见肘，当然，物理检测方法能够得到金片，为其他检测方法提供对比母片。

3) 基于反向解剖检测

目前，大部分的芯片都是基于 CMOS 结构标准单元门电路制造而成，这让反向解剖检测硬件木马的存在成为可能[19]。反向解剖的检测技术必须先把芯片的电路硅片彻底暴露出来；然后利用显微镜成像设备对硅片表面进行图像采集，获取各块的细节图像，再利用特有的逆向分析工具把各区域照片组合成为完整的芯片图像；最后，利用软件把原来的芯片结构样图与组合后的样式图进行同规格的切割，再把切割之后的图像进行对比检查。

利用反向解剖技术可以检测除工艺参数类硬件木马外所有版图级的变化。然而，反向解剖技术是一种带破坏性检测技术，对硬件电路具有破坏性，并且芯片逆向分析验证工程量很大，成本和时间代价太高[19]，针对这些不足，在检测硬件木马时可以对版图中最有可能存在木马的部分进行分析检验，从而提高检测效率。

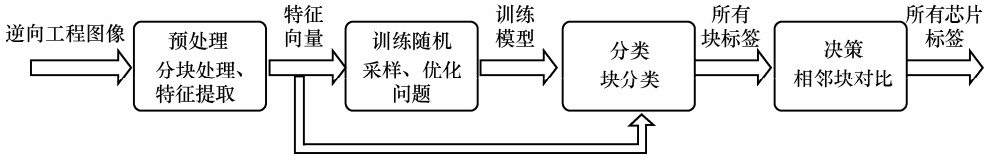


图 7　利用机器学习的芯片分类算法流程

4.1.2 非破坏性检测

1) 旁路分析检测

旁路分析检测是一种常见的检测硬件木马是否存在的方法[19]，电路系统在被植入硬件木马电路后，原有电路的结构和线路会发生一定程度的变化，并且会导致功耗增加、响应延时等旁路特征，为此，可以通过旁路分析的方法判断电路系统中是否被硬件木马感染。

目前，很多研究[20~23]都提出了旁路分析检测的实施方略和优化策略。Agrawal 等[17]提出一种基于功率信号的 IC 指纹法的硬件木马检测技术，通过测试集成电路的功耗得到功耗特性曲线，该曲线用 IC 指纹来表示，然后用该 IC 指纹认证其他待测的芯片。在获取原始集成电路指纹时，需要对原始集成电路进行多次测试，并且分析对应输出的功率特性，进而生成指纹。在功耗的测量过程中，需要考虑环境噪声功耗以及设备工艺的噪声，在环境噪声较大时，木马噪声容易被覆盖，因此，采用 KL 分析环境噪声特征来确定芯片是否被植入木马。Exurville 等[24]利用一种时钟短时脉冲波干扰注入的方法来观察集成电路芯片内部的延时，以此来检测是否存在木马。李雄伟等[25]通过对 CMOS 电路特性的影响，提出了利用电路平均动态电流与最大工作频率旁路分析的方法来检测硬件木马，此方法能够克服噪声的干扰成功检测出硬件木马，然而实验采用的蒙特卡洛法在实验效率及时间代价上略有不足。Ngo 等[26]提出了一种基于电磁辐射的硬件木马检测，该旁路特征由于其良好的时空分辨率被优先选择来分析，该文搭建特征矩阵来测量电路运行信号状态变化，以此来检测硬件木马位置和大小，同时，给出硬件木马的检测概率，这种方法在 Virtex-5 FPGAs 上的 AED-128 加密核上被验证，然而该文并没有给出一个精细的衡量准则来判定参数变化对检测概率的影响。

旁路分析检测是一种非破坏性检测，也是硬件木马检测中的关键检测方法，然而，旁路分析方法存在一个普遍的问题，即系统中存在系统噪声和设备特定信号偏差，对旁路信号去噪也会削弱原始信号，因此如何分辨各种异常信号来检测硬件木马成为一个急需解决的问题[27,28]。

2) 功能测试

在特定的触发信号下，硬件木马能够使系统输出发生改变，干扰正常的输出信号，这使利用功能测试的方法对硬件木马进行检测成为可能。功能测试技术一开始是根据 VLSI 中故障测试的 ATPG 测试技术，利用自动测试工具，在电路芯片的输出端提供激励信号，检测芯片的输出端与正常的输出逻辑是否一致，若产生不一致的逻辑输出，则可以证明硬件木马的存在，以此来断定芯片是否被木马污染[29]。

功能测试也在一些缺点，如木马激活的概率极低，现代电子设备通常有多个 I/O 口，遍历所有输入时间代价很高，此外，有些硬件木马的触发依赖于设备内部状态，单纯靠遍历外部输入来验证硬件木马是否存在变得不可能。硬件木马在设计时一般会选择低活性状态作为其触发的条件，正常情况下硬件木马很难被激活，普通故障覆盖率模型很难进行有效的检测。然而，通过故障模型测试算法，可以生成极精准的测试向量集，如启发式硬件木马测试码生成算法 MERO[30]。该算法可以对芯片状态穷举测试，进而判断出是否存在因木马而导致的错误输出。该技术能够借助测试平台实现自动化检测，智能化程度较高，能够支持其他方法进行检测，难点在于如何生成合适的测试码，且算法的复杂度较高。冯秋丽等[31]提出一种基于节点活性的硬件木马检测方法，通过对目标电路施加随机测试向量并统计电路节点的翻转概率，选取节点翻转概率临界阈值，筛选出目标电路中的低活性节点作为木马节点；然后，利用格雷码编码电路生成候选测试向量，采用二分法对候选测试向量进行分组，并统计每组测试向量下木马节点的翻转概率，挑选出使木马节点翻转概率最大的测试向量作为生成的硬件木马测试向量；最后，结合多参数旁路检测方法实现对硬件木马的检测。

4.2 检测方法对比

上述几种硬件木马检测方法实现方式各不相同，各有各的特点，通过对比这些方法的优缺点（如表 1 所示），能够使硬件木马检测技术优势互补，提升木马检测精度和效率。

表 1　　**硬件木马检测方法对比**

检测方法	优势	缺陷
逆向工程	能够获得金片，精确度较高（不含木马的芯片）	耗时、流程复杂、对原电路有破坏性
物理检测	原理简单、能够获得金片	不适用于结构复杂的芯片，对原电路有破坏性
反向解剖	检验方式彻底，正检率高	采集芯片版图工艺复杂，时间代价高，对原电路有破坏性
旁路分析	对大型木马电路有效，测试方法较为简单，对原电路无破坏	受工艺噪声影响，成熟性有待提高
功能测试	对小型木马电路有效，不受工艺噪声影响，易于自动化实现，对原电路无破坏	测试方式较死板，时间代价大

总地来说，木马检测方法中逆向工程、物理检测、反向解剖及功能测试趋于解决小型硬件木马电路的检测，而旁路分析由于其独特的检测手段，能够适用于如今较为普遍的大型集成电路。由于现在还没有统一的木马模型和衡量标准，各种检测技术的优劣很难进行盖棺定论，各有各的优势。除了上述几种检测方法外，还有一些硬件木马检测手段，如专门性检测技术，即在设计阶段，专门设计防范硬件木马的举措，以降低硬件木马的检测效率。目前，硬件木马检测技术发展所面临的最大挑战是缺少实际的木马样本供研究人员分析考察，没有样本就没有数据，进而难以采用科学的方法进行硬件木马检测研究。

5　硬件木马防范

从 3.1 节中可以发现，木马检测对电路可靠性有一定程度的保证，然而现有的硬件木马检测技术对低触发概率的硬件木马检测难度仍然很大，为了进一步增强电路安全性，应当在电路设计的同时兼顾木马抗植入技术，即集成电路硬件木马防范技术，在集成电路设计和制造过程中针对硬件木马问题设置信任门槛机制有助于更有效地抵抗硬件木马的攻击。为此，现阶段很多学者提出授信设计（DFT，design-for-trust）、授信分块制造（SMFT，split-manufacturing-for-trust）以及实时监控与电路增强技术（RMCET，real-time monitoring and circuit enhancement technology）的概念以提高集成芯片电路的安全性。

5.1　授信设计

授信设计根据其实现原理可以划分为 3 个层次的内容。

5.1.1　电路模糊和后版图填充技术

电路模糊技术即对电路功能或结构进行模糊化[32]，通过对电路进行模糊处理，使硬件木马植入者难以发现原始电路的结构层次，为硬件木马植入增加难度。后版图填充技术是填充电路中剩余空间以减少硬件木马植入所需空间的一种技术。Xiao 等[33]提出一种 BISA（built in self-authentication）技术来填充电路芯片未被利用的空间，使硬件木马植入成为不可能。

5.1.2　反逆向工程技术

为了植入硬件木马，植入者往往需要熟悉集成电路的各个功能模块，对于不熟悉集成电路各模块功能的木马植入，往往需要对电路进行逆向转化工程来识别电路[4]。利用逻辑混淆、伪装、功能填充单元能够有效阻止硬件木马的植入。逻辑混淆通过在原始设计中插入内置锁来隐藏真正的功能，内置锁对外界来说是透明的，在运用到特定的钥匙时，真正的功能才会运行。对于组合逻辑电路，异或门及与或非门电路将被引入[34]。Liu 等[35]提出插入可重构的逻辑单元来实现逻辑混淆，当可重构电路被制造商和最终使用者正确编译后，设计电路正常运行。伪装即通过在不同层之间增加虚拟、虚假连接来制造不可区分的布局[36]，这种技术能够阻止攻击者提取合适的网络表，从而保护原始电路免受木马的植入。Bi 等[37]利用虚拟连接的方法，在两极控制 SINWFET（silicon nanowire FET）的基础上设计出一系列伪装胞元，从而实现伪装功能。

5.1.3　可信度计算

运行监控和可信计算的区别在于可信计算往往忽略木马攻击的因素，保证不可信单元进行可

信度的计算是芯片防护的有效手段。运行状态下的木马检测和恢复作为芯片防范的最后一道防线是至关重要的，现阶段一些研究在多核处理器中利用分布式软件调度协议来完成可信度的计算。Keren[38]提出的并发错误检测技术能够检测出因木马产生的恶意输出。此外，Reece 等[39]提出运用第三方 IP 供应商不同的参数设置来阻止木马感染。IP 核硬件木马难度很大，通过绕过检测的问题，使木马不能激活，同样可以达到安全的目的。王龙等[40]提出一种输入序列检测器的设计方案，首先对输入序列进行相关性处理，把符合相关性的序列通过 IP 核传输进逻辑芯片，在 IP 核与芯片间再做解密处理输入芯片。硬件木马激活逻辑因为不能得到所需的激活序列而无法激活，从而达到硬件木马防护的目的。

5.2 授信分块制造

授信分块制造作为一种利用最顶尖半导体制造商最小化集成电路风险的制造技术，在近期被提出，它将集成电路的制造划分为前序工艺和后道工艺 2 个部分，且这 2 个部分分别给不同制造商制造。前序工艺由不可信的制造商完成，然后将芯片运输到可信的后道工艺制造商制造。不可信制造商由于不能进入后道工艺层，因此无法选择合适的电路区域植入木马。截至目前，分块制造仍处于 2D 到 3D 集成领域[41]，2.5D 集成技术首次将一次电路设计划分为 2 个部分，非信任商生产制造后在芯片与封装载板上插入硅中介来进行连接，因此，一部分在信任制造商中的连接将被隐藏在中介层。而在 3D 集成领域，集成电路被划分为 2 层，且由不同制造商制造，一层叠放在另一层之上，上层通过 TSV 的连接器连接，考虑 3D 制造的障碍，2D 及 2.5D 切分制造技术在如今较易现实。

现阶段关于授信分块制造的相关研究很多，Imeson 等[41]提出一种 3D 集成电路技术对抗硬件木马带来的安全威胁，在对较高层次进行分块时，利用 *k*-安全标准来选择合适的线路，提升其信任等级。然而，提升大量的线路将引起时耗和功耗的过载，进而导致芯片运行异常。Vaidyanathan 等[42]通过比较芯片的性能验证了经过 M1 分块制造后的可行性。Xiao 等[43]利用模糊 BISA 技术在原始电路中插入虚拟电路，它能在后续分块制造中实现模糊设计，进而加大木马植入的难度。

5.3 实时监控与电路增强技术

实时监控技术通过在芯片中植入监控电路对其进行状态检测与控制，一旦电路出现异常，则会采取关闭信息输入输出通道等安全措施，以此保证电路安全，防止木马危害继续扩大。Bhunia 等[44]提出一种在片上系统芯片上植入监控电路，不仅可以检测电路异常状态，还能对电路内存的非正常使用进行监测，这大大保证了电路的安全性。

电路增强技术主要针对硬件木马检测技术进行改进，如功能测试技术，由于设计过程中大量的低控制率和低可度量网络阻止了激发木马的可能性，Salmani 等[45]为了增加节点的控制率和可度量特性，在电路中插入额外的测试节点。对于旁路分析检测方法，为了提高旁路检测方法的敏感性，有许多方法被提出，Salmani 等[46]利用单元扫描记录技术实现最小化局部旁路背景信号，许多新构建的模型和传感器的提出使良性旁路分析技术相比传统检测方式有更高的测试敏感度，对集成电路关键计算单元的运行监测技术能有效提升硬件木马对抗的可信程度。Jin 等[10]设计了一种模拟神经网络的芯片，它能够通过测量传感器获取数据并进行训练，从而区分可信电路和不可信电路。石朝阳等[47]提出一种基于密钥的电路增强技术，通过在电路中增加初始序列（密钥）、迷惑电路和冗余电路，隐藏有正确功能的原始电路，以预防在设计及后续环节中可能被植入的硬件木马，优化的预防电路在没有太多额外电路资源开销的情况下能有效保护电路不被硬件木马破坏，且不影响正常的功能。

实时监控技术能够控制木马的传播[48]，减少硬件木马触发后给整个系统带来的危害[45]，然而它并不能从根源上阻止木马的植入，电路增强技术通过在原始电路上增加额外的增强电路，能够加大电路抵抗木马植入的能力。

5.4 防范技术对比

授信设计、授信分块制造、实时监控与电路增强技术作为常见的硬件木马防范技术有其各自的优势和缺陷，如表 2 所示。

表 2　　硬件木马防范技术对比

防范技术	优势	缺陷
授信设计	设计灵活，较易实现	工程量大，导致额外时延和功耗
授信分块制造	技术含量高，风险最小化，不改变原来电路构造	难度较大，工艺复杂，制造周期长，涉及多方向加工协作
实时监控与电路和增强	能控制危害伟播，设计方式灵活，能够把握电路实时状态	导致额外时延和功耗，制造成本大

对授信设计而言，其设计方式灵活多样，然而它很大程度需要依靠在原始电路中增加额外的元器件来影响硬件木马的植入，当电路规模增加时，信号处理的复杂度及功耗、延时相应增加，这也是电路设计者最关注的问题。因此，授信设计技术在拥有上百万门器件的大规模集成电路上仍然很难运用。授信分块制造不改变原始电路的构造，但其复杂的工艺设计以及多方面的加工协作处理，往往使制造周期较长，而实时监控与电路增强通过监测电路的实时状态能够很好地把握电路运行情况，在硬件木马触发时能够控制危害的传播，增强电路在设计方式上灵活多样，但该技术制造成本较大，且会导致额外的时延和功耗。

对比这些技术的优缺点能够使硬件设计开发者及硬件木马研究者对木马的防范技术有较全面的认知，能够促进硬件木马防范技术的优势互补，提升木马防范技术的全面性。

6　结束语

在经济全球化的今天，大规模芯片制造往往涉及多方资源，在激烈的市场竞争下，芯片的安全性问题逐渐引起了人们的重视。随着电路制造工艺的不断进步，集成度的不断提高促使硬件木马植入变得更加灵活、简便、隐蔽，因此，对硬件木马特性和检测防范方法进行研究显得极为重要。本文对硬件木马的产生、组成特点以及攻击模式进行了详细描述，并针对现阶段研究成果重点分析了硬件木马检测和防范技术，根据检测、防范技术的特点对其进行合理分类，并对比了每种检测技术的优缺点。

如何保证芯片的安全是开发设计人员必须重视的问题，硬件木马检测及防范尚没有形成成熟的行业技术，更多的研究还处于实验阶段，如何加强木马检测技术以及如何设计和搭建高集成度仿真环境有待进一步研究。

参考文献：

[1] YANG K Y, HICKS M, DONG Q, et al. A2: analog malicious hardware[C]//2016 IEEE Symposium on Security and Privacy (SP). 2016:18-37.

[2] ZHANG J, YUAN F, WEI L, et a1. VeriTrust：verification for hardware trust[C]//The Computer-aided Design of Integrated Circuits and Systems. 2013:1148-1161.

[3] KARRI R, RAJENDRAN J, ROSENFELD K. Trustworthy hardware: identifying and classifying hardware Trojans[J].Journal of Computer, 2010, 43(10): 39-46.

[4] XIAO K, FORTE D, JIN Y, et al. Hardware Trojans: lessons learned after one decade of research[J]. ACM Transactions on Design Automation of Electronic Systems, 2016, 22(1):1-23.

[5] BAO C, FORTE D, SRIVASTAVA A. On application of one-class SVM to reverse engineering-based hardware Trojan detection[C]//The International Symposium on Quality Electronic Design. 2014:47-54.

[6] BANGA M, HSIAO M S. A region based approach for the identification of hardware Trojans[C]//IEEE Workshop on Hardware Oriented Security and Trust-HOST. 2008:40-47.

[7] BHUNIA S, HSIAO M S, BANGA M, et al. Hardware Trojan attacks: threat analysis and counter measures[J]. Proceedings of the IEEE, 2014, 102(8):1229-1247.

[8] BAUMGATEN A, STEFFEN M, CLAUSMAN M, et al. A case study in hardware Trojan design and implementation[J]. International Journal of Information Security, 2011, 10(1): 1-14.

[9] LIN L, KASPEr M, GUNEYSU T, et al. Trojan side-channel: light weight hardware trojans through side-channel engineering[C]//The 11th International Workshop on Cryptographic Hardware and Embedded Systems (CHES 2009) . 2009: 382-395.

[10] JIN Y, MAKRIS Y. Hardware Trojans in wireless cryptographic ICs[J]. IEEE Design & Test, 2010, 27(1): 26-35.

[11] 倪林, 李少青, 马瑞聪, 等. 硬件木马检测与防护[J]. 数字通信, 2014, 41(1): 59-63.
NI L, LI S Q, MA R C, et al. Hardware Trojan detection and protection[J].Digital Communications, 2014, 41(1): 59-63.

[12] 谢海, 恩云飞, 王力纬. 电磁泄露型硬件木马设计与检测[J]. 广

东工业大学学报, 2013(4):70-73.
XIE H, EN Y F, WANG L W. Design and detection of hardware Trojan based on electromagnetic leakage[J]. Journal of Guangdong University of Technology, 2013(4):70-73.

[13] 邹程, 张鹏, 邓高明, 等. 基于功率旁路泄露的硬件木马设计[J]. 计算机工程, 2011, 37(11):135-137.
ZOU C, ZHANG P, DENG G M, et al. Design of hardware Trojan based on power side-channel exposure[J]. Computer Engineering, 2011, 37(11):135-137.

[14] 吴志凯, 魏佩, 陈吉华, 等. 一种基于少态触发的硬件木马设计与实现[C]//微处理器技术论坛. 2014.
WU Z K, WEI P, CHEN J H, et al. Design and implementation based on less state trigger hardware Trojans[C//Microprocessor Technology BBS. 2014.

[15] 王晓晗, 李雄伟, 张阳, 等. 一种基于故障注入的硬件木马设计[J]. 军械工程学院学报, 2015(5): 57-61.
WANG X H, LI X W, ZHANG Y, et al. Hardware Trojan design based on fault injection[J]. Journal of Ordnance Engineering College,2015(5):57-61.

[16] 李蕾, 尚子靖, 冯建华, 等. 基于有限状态机的硬件木马设计和插入[J]. 北京大学学报:自然科学版, 2013, 49(6): 1105-1110.
LI L, SHANG Z J, FENG J H, et al. Design and insertion of hardware Trojan based on finite state machine[J]. Acta Scientiarum Naturalium Universitatis Pekinensis, 2013, 49(6):1105-1110.

[17] AGRAWAL D, BAKTIR S, KARAKOYUNLU D, et al. Trojan detection using IC fingerprinting[C]///The 2007 IEEE Symposium Security and Privacy. 2007: 296-310.

[18] BHASIN S, DANGER J L, GUILLEY S, et al. Hardware Trojan horses in cryptographic IP Cores[C]//Fault Diagnosis and Tolerance in Cryptography. 2013:15-29.

[19] 刘华锋, 罗宏伟, 王力纬. 硬件木马综述[J].微电子学, 2011, 41(5): 709-713.
LIU H F, LUO H W, WANG L W, et al. Survey on hardware Trojan horse[J]. Microelectronics, 2011, 41(5):709-713.

[20] POTKONJAK M, NAHAPETIAN A, NELSON M, et al. Hardware Trojan horse detection using gate-level characterization[C]//Design Automation Conference (DAC '09). 2009:688-693.

[21] NARASIMHAN S, DU D, CHAKRABORTY R S, et al. Multiple-parameter side-channel analysis: a non-invasive hardware trojan detection approach[C]//IEEE Workshop on Hardware-Oriented Security and Trust-HOST. 2010:13-18.

[22] BANGA M, HSIAO M S. VITAMIN: voltage inversion technique to ascertain malicious insertions in ICs[C]//IEEE Workshop on-Hardware-Oriented Security and Trust-HOST. 2009:104-107.

[23] JIN Y, MAKRIS Y. Hardware Trojan detection using path delay fingerprint[C]//IEEE Workshop on Hardware-Oriented Securityand Trust- HOST, 2008:51-57.

[24] EXURVILLE I, FOURNIER J, DUTERTRE J M, et al. Practical measurements of data path delays for IP authentication and integrity verification[C]//IEEE International Workshop on Reconfigurable and Communication-Centric Systems-on-Chip-ReCoSoC. 2013: 1-6.

[25] 李雄伟, 王晓晗, 张阳, 等. 基于多旁路综合分析的硬件木马检测方法[J]. 计算机仿真, 2015, 32(3):216-219.
LI X W, WANG X H, ZHANG Y, et al. Hardware Trojan detection method based on multiple side-channels analysis[J]. Computer Simulation, 2015, 32(3):216-219.

[26] NGO X T, NAJM Z, GUILLEY S, et al. Method taking into account process dispersion to detect hardware Trojan horse by side-channel[C]//Security Proofs for Embedded Systems−PROOFS. 2014.

[27] RAD R, PLUSQUELLIC J, TEHRANIPOOR M. Sensitivity analysis to hardware Trojans using power supply transient signals[C]// IEEE Workshop on Hardware-Oriented Security and Trust-HOST. 2008:3-7.

[28] RAD R, WANG X, TEHRANIPOOR M, et al. Power supply signal calibration techniques for improving detection resolutionto hardware Trojans[C]//IEEE International Conference on Computer-Aided Design-ICCAD. 2008:632-639.

[29] SUI Q. Hardware Trojan detection based on side channel signal analysis[D]. Changsha: National University of Defense Technology, 2012.

[30] CHAKRABORTY R S, WOLFF F, PAUL S, et al. ME-RO: a statistical approach for hardware Trojan detection[C]//Lecture Notes in Computer Science. 2009: 396-410.

[31] 冯秋丽. 基于节点活性的硬件木马检测方法研究[D]. 广州：广东工业大学. 2016.
FENG Q L. The research of hardware Trojan detection method based on nodes activity[D]. Guangzhou: Guangdong University of Technology. 2016.

[32] CHAKRABONY R S, BHUNIA S. Security against hardware Trojan attacks using key—based design obfuscation[J]. Journal of Electronic Testing, 2011,27(6):767-785.

[33] XIAO K, TEHRANIPOOR M. BISA: built-in self-authentication for preventing hardware Trojan insertion[C]//IEEE Intemational Workshop on Hardware-oriented Security and Trust. 2013:45-50.

[34] ROY J A, KOUSHANFAR F, MARKOV I L. EPIC: ending piracy of integrated circuits[C]//Design, Automation & Test in Europe. 2008:1069-1074.

[35] LIU B, WANG B. Embedded reconfigurable logic for ASIC design obfuscation against supply chain attacks[C]//The Design, Automation and Test in Europe Conference and Exhibition (DATE'14). 2014:1-6.

[36] RAJENDRAN J, ZHANG H, SINANOGLU O, et al. High-level synthesis for security and trust[C]//The 2013 IEEE 19th International on-Line Testing Symposium (IOLTS'13). 2013:232-233.

[37] BI Y, GAILLARDON P E, HU X S, et al. Leveraging emerging technology for hardware security - case study on silicon nanowire FETs and graphene SymFETs[C]// Test Symposium. 2014:342-347.

[38] KEREN I L, KARPOVSKY M. Duplication based one-to-many coding for Trojan HW detection[C]//The 2010 IEEE 25th International Symposium on Defect and Fault Tolerance in VLSI Systems (DFT'10). 2010:160-166.

[39] REECE T, LIMBRICK D B, ROBINSON W H. Design comparison to identify malicious hardware in external intellectual property[C]//IEEE International Conference on Trust, Security and Privacy in Computing and Communications. 2011:639-646.

[40] 王龙, 陈吉华, 李少青, 等. 基于第三方 IP 核硬件木马防护性设计[C]//微处理器技术论坛. 2015.
WANG L, CHENJ H, LI S Q, et a1. Based on the third party IP core hardware Trojan defensive design[C]//Microprocessor Technology BBS. 2015.

[41] IMESON F, EMTENAN A, GARG S, et al. Securing computer hardware using 3D integrated circuit (IC) technology and split manufacturing for obfuscation[C]//The Usenix Security Symposium. 2013.

[42] VAIDYANATHAN K, DAS B P, SUMBUL E, et al. Building trusted ICs using split fabrication[C]//IEEE International Symposium on Hardware-Oriented Security and Trust. 2014:1-6.

[43] XIAO K, FORTE D, TEHRANIPOOR M M. Efficient and secure split manufacturing via obfuscated built-in self-authentication[C]// The 2015 IEEE International Symposium on Hardware Oriented Security and Trust (HOST'15). 2015:14-19.

[44] BHUNIA S, ABRAMOVICI M, AGRAWAL D, et a1.Protection against hardware Trojan attacks: towards a comprehensive solution[J]. IEEE Design & Test, 2013(3): 6-17.

[45] SALMANI H, TEHRANIPOOR M, PLUSQUELLIC J. A novel technique for improving hardware Trojan detection and reducing trojan activation time[C]//IEEE Transactions on Very Large Scale Integration (VLSI) Systems. 2012:112-125.

[46] SALMANI H,TEHRANIPOOR M, PLUSQUELLIC J. A layout-aware approach for improving localized switching to detect hardware Trojans in integrated circuits[C]//2010 IEEE International Workshop on Information Forensics and Security(WIFS 2010). 2010.

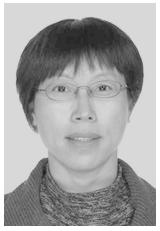
[47] 石朝阳, 邹雪城, 明瑞华, 等. 一种基于密钥的硬件木马预防方法研究[J]. 现代电子技术, 2016, 39(20).
SHI Z Y, ZOU X C, MING R H, et a1. Hardware Trojan prevention method based on secret key[J]. Modern Electronics Technique, 2016, 39(20).

[48] 曾辰熙, 吴泉源, 李爱平, 等. 基于模糊层次分析的木马攻击效果评估技术研究[J]. 网络与信息安全学报, 2016, 2(7): 49-58.
ZENG C X, WU Q Y, LI A P, et a1. Research on FAHP based Trojan attack effect evaluation[J]. Chinese Journal of Network and Information Security, 2016, 2(7): 49-58.

作者简介：

许强（1992-），男，江西赣州人，上海交通大学博士生，主要研究方向为信息安全。

蒋兴浩（1976-），男，河南邓州人，博士，上海交通大学教授，主要研究方向为网络空间安全、信息处理。



姚立红（1974-），女，江苏建湖人，博士，上海交通大学高级工程师，主要研究方向为系统安全、网络访问控制。

张志强（1978-），男，山东高青人，上海机电工程研究所高级工程师，主要研究方向为信息对抗。

张诚（1985-），女，江苏徐州人，上海机电工程研究所工程师，主要研究方向为信息对抗。