

基于旁路分析的集成电路芯片硬件木马检测

赵崇征¹, 邓高明², 赵 强²

(1 北京橙时代信息技术有限公司, 北京 100080; 2 军械工程学院 计算机工程系, 河北 石家庄 050003)

摘 要: 针对密码芯片中硬件木马电路检测的困难性, 介绍了根据芯片旁路信息进行硬件木马检测的思想. 在形式化定义基于旁路分析的硬件木马检测问题的基础上, 分析了含硬件木马与不含硬件木马的密码芯片对应旁路信号在主成份分析结果上的差异, 并以此对 FPGA 实现的含硬件木马的 DES 密码原型芯片进行了检测实验, 实验结果表明了基于旁路信号主成份分析在密码芯片硬件木马检测中的效果.

关键词: 密码芯片; 硬件木马; 旁路分析; 主成份分析

中图分类号: TN918

文献标识码: A

文章编号: 1000-7180(2011)11-0005-05

Detecting Hardware Trojans in IC Chips with Side Channel Analysis

ZHAO Chong-zheng¹, DENG Gao-ming², ZHAO Qiang²

(1 Beijing Orange Time Information Technology Co. Ltd, Beijing 100080, China;

2 Dept. of Computer Engineering, Ordnance Engineering College, Shijiazhuang 050003, China)

Abstract: Aimed at the difficulty in detecting hardware Trojans inside IC chips, a novel idea about detecting Trojan circuits with side channel analysis is introduced. After formally defining the problem of detecting hardware Trojans in IC chips with side channel analysis, the differences between principle components analysis on side channel signals acquired from chips with Trojan circuits and without Trojan circuits were analyzed. Detection experiments with principle components analysis were taken out on a DES cipher chip implemented on a FPGA board. The results of experiments verified the effect of Hardware Trojans detection with principle components analysis on side channel signal.

Key words: cipher chips; hardware Trojan; side channel analysis; PCA

1 引言

伴随着半导体设计、制造的硬件外包全球化发展趋势, 近年来出现了一种针对集成电路芯片新型的硬件攻击方式, 称为“硬件木马 (Hardware Trojan)”^[1], 即在 IC 设计与制造过程中人为地制造一些非法电路, 留下“电子后门”, 为后续攻击打开方便之门. 密码芯片作为密码算法在 IC 芯片上的实现, 同样面临硬件木马的威胁, 比如密码芯片中的硬件木马可以实现以隐秘的通道泄露密钥, 在特定条件下篡改运算输出等功能.

硬件木马的检测困难性是其威胁密码芯片的重要原因. 现有一般的检测方法依赖电子显微扫描、电

压对照成像等复杂的故障分析技术^[2], 设备要求高, 代价昂贵, 而传统的基于 ATPG 的芯片检测方法对于激活条件非常独特的木马来说, 很难生成激活模式, 硬件木马的检测成为保证密码芯片安全性亟待解决的问题^[3]. 新型的基于旁路分析 (或称旁路攻击, Side Channel Analysis, SCA) 技术在用于硬件木马检测方面显示出较为乐观的应用前景. 旁路分析技术最初主要应用于密码分析学领域, 是通过监测采集 IC 密码芯片在工作时泄露的物理特征信息 (例如功耗、声波、运行时间、电磁辐射等), 利用数学分析 (重点是统计分析) 方法来获取 IC 密码芯片中的秘密信息^[4-6]. 基于旁路分析的密码芯片中硬件木马检测正是利用原始电路与木马电路之间旁路信号

收稿日期: 2011-01-06; 修回日期: 2011-02-08

基金项目: 国家自然科学基金项目 (60940019); 军械工程学院原始创新基金 (YSCX0903)

的差异来检测木马。

2 基于旁路分析的硬件木马检测问题描述

文中以实现数据加解密功能的密码芯片作为研究对象来分析硬件木马检测问题。要通过密码芯片的旁路信号检测其内部的硬件木马,首先要分析其旁路信号的组成,和基于旁路模板分析的密钥恢复类似,考虑一个密码芯片 C ,其上执行加密(解密)算法 $A(K)$ 。在 C 执行 $A(K)$ 的同时用测量设置 M 采集旁路信号(功耗、电磁辐射)轨迹,则旁路信号轨迹上时刻 t 的值($t; C; O(k); M$)可以建模成包含以下部分:①密码运算 $A(K)$ 相关的旁路信号均值部分 $l(t; A(K))$;②与运算相关的噪声 $n_l(t; A(K))$;③与芯片相关的噪声 $n_l(t; C)$;④测量噪声 $n_m(t; M)$ 。根据该模型,密码芯片的旁路泄露表述如下:

$$r(t; C; A(K); M) = l(t; A(K)) + n_l(t; A(K)) + n_l(t; C) + n_m(t; M) \quad (1)$$

而含有硬件木马的密码芯片对应的功耗轨迹会有由木马电路 T 导致的一个额外的部分 $\tau(t; C)$,上式则需改写为:

$$r(t; C; A(K); M) = l(t; A(K)) + n_l(t; A(K)) + n_l(t; C) + n_m(t; M) + \tau(t; C) \quad (2)$$

一般来说,测量噪声是每次测量都不同的随机噪声。由于功耗信号其它的任何一部分都与测量是无关的,因此随机噪声 $n_m(t; M)$ 可以通过对同一个密码芯片进行大样本测量求均值来消除,而且一般同一型号同一批芯片之间旁路泄露信号噪声规模只占信号的约0.75%,而测量噪声和运算噪声的规模一般都在5%以上^[7],因此在考虑对算法 $A(K)$ 的旁路信号特征进行建模时将这一部分可以忽略而不影响模型精度。上述两式中的噪声部分合并成一个,写成

$$r_C(t; C; A(K)) = l(t; A(K)) + n(t; A(K)) \quad (3)$$

$$r_T(t; C; A(K)) = l(t; A(K)) + n(t; A(K)) + \tau(t; C) \quad (4)$$

上式描述的是时刻 t 的旁路信号组成,在密码芯片运算过程中连续采集 m 个时刻的旁路信号组成一条旁路信号轨迹向量 $T = [r_0, r_1, \dots, r_{m-1}]$, T 上的每个点 $r_i (0 \leq i \leq m-1)$ 的组成都符合上面的描述。根据上述旁路信号组成,可将基于旁路模板的硬件木马检测问题定义如下:

定义1 基于旁路模板的硬件木马检测。给定已知不含硬件木马的密码芯片 C_G ,采集到 C_G 运行

密码运算 $A(K)$ 时对应的旁路信号轨迹向量 T_G ,给定另外一个同类型的未知密码芯片 C_X 以及对应的旁路信号轨迹向量 T_X ,问如何根据上述旁路信号向量确定芯片 C_X 中是否含有硬件木马。

换句话说,对未知密码芯片 C_X 有如下两个假设:

$$(1) H_G : T_X = T_G$$

$$(2) H_T : T_X = T_G + T_\tau$$

式中, T_τ 表示由木马电路产生的旁路信号轨迹。基于旁路模板分析的硬件木马检测问题就是要根据上述旁路信号确定哪个假设是正确的。

上述硬件木马检测问题可以看作是信号特征判别问题,需要做的就是刻画不含硬件木马密码芯片对应旁路信号的特征,以这种特征作为模板并核对待检验的未知芯片对应旁路信号中是否与模板存在显著差异。

由定义1可知,文中描述的基于旁路模板的密码芯片硬件木马检测需要一个基本假设,即假设存在一个或几个与待测芯片同型号并确定不含硬件木马的“纯净”芯片样本。作这样的假设是为了确保待测芯片与已知芯片对应旁路信号特征差异仅由硬件木马电路产生,而不会由芯片差异产生。实际操作中可以有許多途径可以满足这个假设,例如可以在有监督的情况下定制少量芯片作为这样的“纯净”芯片作为样本,或者在一批芯片中抽取少量芯片,采集完旁路信号之后进行解包,通过检查芯片实现与设计是否一致来确定“纯净”芯片样本。虽然上述确定“纯净”芯片样本的过程成本费用一般较高,但是对于快速检测数量较大的一批芯片的情形,定制和解包的成本费用分摊到每个待测芯片上也就降低到可以接受的程度了。

3 基于旁路信号主成份分析的硬件木马检测

根据上节的分析,基于旁路模板的硬件木马检测问题可以看作是信号中叠加的额外成分检测问题。为了更好地描述硬件木马电路对应的旁路信号与密码运算电路对应旁路信号的叠加情形,可以把由 m 个采样点组成的旁路信号轨迹 $t = [r_1, r_2, \dots, r_m]$ 看作 m 维空间中的一个向量 $\vec{t}$ 。这样含硬件木马的芯片对应的旁路信号向量 $\vec{t}_T$ 就可以看作是由密码运算电路对应的旁路信号向量 $\vec{t}_G$ 和木马电路对应的旁路信号向量 $\vec{t}_\tau$ 的叠加,由于木马电路不会和密码运算电路在同样的时刻作同样的运算操作,因

此向量和不会完全线性相关,即二者不会完全在一个方向上,图1给出了二维空间下向量叠加的示意图。

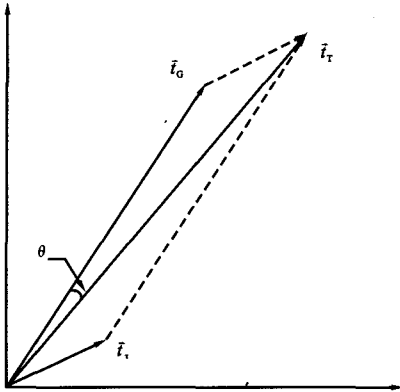


图1 二维空间下向量叠加的示意图

在已知“纯净”样本芯片对应的旁路信号向量模板 $\vec{t}_G$ 的条件下,可以通过待测芯片对应的旁路信号向量 $\vec{t}_X$ 与 $\vec{t}_G$ 之间的夹角来推断 $\vec{t}_X$ 中是否含有额外的成分:

$$\cos \theta = \frac{\langle \vec{t}_X, \vec{t}_G \rangle}{\|\vec{t}_X\| \cdot \|\vec{t}_G\|} \quad (5)$$

显然, $\cos \theta$ 越大表示 θ 越大,通过设定 $\cos \theta$ 的阈值,超过阈值则可认为 $\vec{t}_X$ 中含有超出 $\vec{t}_G$ 的额外成分。

但是,含硬件木马的芯片对应的旁路信号向量 $\vec{t}_T$ 中的 $\vec{t}_T$ 幅度很小(比 $\vec{t}_G$ 的幅度小3到4个数量级)的时候, θ 趋近于0, $\vec{t}_T$ 和 $\vec{t}_G$ 无法通过 $\cos \theta$ 来区分,这种情况下可以用 PCA 来区分这两个信号。

PCA 是统计学中分析数据的一种有效方法^[7],其目的是在数据空间中找一组向量尽可能地解释数据的方差,通过一个特殊的向量矩阵将数据从原来的高维空间映射到一个低维向量空间,在这个降维的过程中实际上实现了对向量的去相关分解。用 PCA 进行旁路信息数据降维与向量分解的步骤如下:

(1) 原始数据预处理

让芯片进行相同的运算 m 次并获取芯片工作时的旁路电磁信号,这样总共获取对应的 m 条电磁曲线,每条曲线有 n 个点,即 $t_{i1}, t_{i2}, \dots, t_{in}$ ($1 \leq i \leq m$)。接下来,电磁曲线样本组成了一个矩阵 $T_{m \times n}$,且对应的均值向量可以计算如下:

$$\begin{aligned} \bar{t} &= \langle \bar{t}_1, \bar{t}_2, \dots, \bar{t}_n \rangle \\ &= \langle \frac{1}{m} \sum_{i=1}^m t_{i1}, \frac{1}{m} \sum_{i=1}^m t_{i2}, \dots, \frac{1}{m} \sum_{i=1}^m t_{in} \rangle \end{aligned} \quad (6)$$

每一条曲线减去均值向量得到的偏差矩阵

$N_{m \times n}$,如式(5),其每一列 N_j 可以看作一组特征随机变量:

$$N_{m \times n} = [t_1 - \bar{t}, t_2 - \bar{t}, \dots, t_m - \bar{t}]^T \quad (7)$$

(2) 建立相关矩阵,求特征值与特征向量

根据偏差矩阵,求得每两个随机变量之间的相关系数 C_{ij} :

$$\begin{aligned} C_{ij} &= \frac{1}{m-1} N_i^T N_j \\ &= \frac{1}{m-1} \sum_{k=1}^m (t_{ki} - \bar{t}_i)^T (t_{kj} - \bar{t}_j) \end{aligned} \quad (8)$$

所有的相关系数构成一个相关系数矩阵 $C_{n \times n}$,按从大到小的顺序排列其特征值及对应的特征向量。

(3) 计算方差贡献率,选取主成分

计算主成分的贡献率(数值上等于特征值的贡献率),选择前面 k 个最大特征值,使这 k 个主成分的累计贡献率达到 85% 以上,即 $(\lambda_1 + \lambda_2 + \dots + \lambda_k) / (\lambda_1 + \lambda_2 + \dots + \lambda_n) \geq 85\%$ ($k \ll n$),这时对应的 k 个向量仍保留原信号的足够多的信息,用这 k 个主特征值对应的主特征向量构成 n 维状态空间中的 k 维特征子空间 $A_{n \times k}$ 。

(4) 建立旁路信号主成分特征空间

根据公式 $Y_{n \times n} = N_{m \times n} \times A_{n \times k}$ 计算出所需要的旁路信号特征主成分值,形成原始旁路信号向量在特征向量张成的空间中的正交分解,从而消除了旁路信号原始数据属性之间的相关性,得到了更能反映数据相关的旁路泄露信息本质的特征,这样用最大的有限个主成份即可以区分不含木马电路的芯片产生的电磁旁路信号和含木马电路的芯片产生的电磁旁路信号。

4 针对 FPGA 芯片上硬件木马原型的检测实验

4.1 FPGA 实现的调制辐射秘密泄露型硬件木马原型

Iowa State University 的 Alex Baumgarten 介绍了一种利用 AM 调制辐射方式泄露密码的木马电路设计思路^[8]。本节仿照文献^[8]中的思路,在 Xilinx Spartan 3E FPGA 开发板上实现 DES 密码芯片原型上,设计并实现了一个以 AM 调制辐射方式泄露芯片内部密钥的硬件木马原型。

硬件木马电路原理如图2所示。26 位计数器的高 3 位构成一个 8 状态的状态机,第一个状态是发出“嘀”声,第二个状态是间隔,第三个状态根据密钥

位是“1”还是“0”来确定是发出“嘀”声还是插入间隔,这里密钥位是由密钥块经移位寄存器得到的1位密钥,接下来的状态都是间隔状态以形成一个长间隔来分隔连续的密钥位.这样,如果密钥位是“0”,则 beeps 信号为“嘀”的一声,如果密钥位是“1”,则 beeps 信号为“嘀嘀”的两声.为了攻击者能用接收机听见声音,需要把 beeps 信号转为音频并调制到高频信号上,该木马电路用 26 位计数器的第 15 位输出形成 762Hz 的音频信号,用计数器的第 2 位输出(时钟频率为 50MHz)形成 12.5MHz 的载波信号,上述两个信号与 beeps 信号一起经与门之后形成载波为 12.5MHz 的调制信号,该信号输出到电路板的一个闲置引脚上发射出去,经调制接收机接收并解调后就能听见“嘀,嘀嘀,…”这样的声音,攻击者根据听到的声音能直接获取密钥的二进制序列.

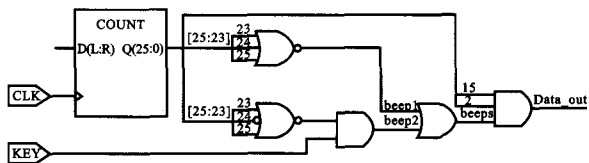


图 2 AM 调制辐射泄露密钥电路原理图

4.2 旁路信号主成份分析检测信息泄露型木马

为了利用旁路信号频谱模板检测密码芯片中的硬件木马电路,本节用图 3 所示旁路信号采集平台采集 DES 密码原型芯片对应的功耗轨迹. DES 密码系统使用 50MHz 的晶振频率,串联在 FPGA 供电回路中电阻 R 的电压用数字存储示波器 (Tektronix DPO4032, 被动探头 P6139A, 差分探头 TDP0500) 进行采集并通过 USB 传输到 PC 机存储,示波器的采集过程由 PC 机 (Pentium IV 4 2.8GHz, 160G HDD, 1G DDR RAM, USB 及 RS232 接口) 上用 LabWindow 编写的虚拟仪器控制平台实现自动控制,并且该虚拟仪器还通过 RS232 将明文发往 FPGA.

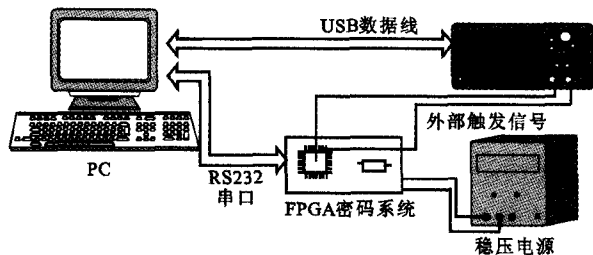


图 3 旁路信号采集平台原理图

用上述实验平台采集到不含硬件木马 FPGA DES 芯片对应旁路信号轨迹如图 4. 含硬件木马的

FPGA DES 芯片对应旁路信号轨迹如图 5. 二者难以看出区别.

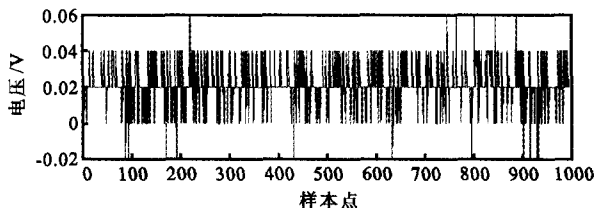


图 4 不含硬件木马的密码芯片旁路信号轨迹

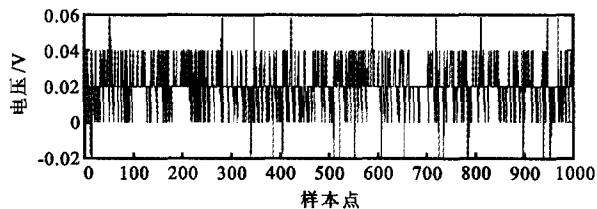


图 5 含硬件木马的密码芯片旁路信号轨迹

用第 3 节介绍的 PCA 方法对两组信号分别进行主成份分析,一般前 15 个主成分的贡献率已能达到 90% 以上. 对于 4.1 节描述的 AM 调制辐射泄露秘密信息型的硬件木马,从 20 个不含硬件木马和 20 个含硬件木马的芯片对应的电磁辐射旁路信号轨迹上各选三个点作图得到图 6 所示. 说明从原始电磁旁路信号上无法区分不含硬件木马的芯片和含硬件木马的芯片. 在完成主成份分析后,选取上述原始电磁旁路信号对应的最大的三个主成份作图得到图 7 所示. 说明不含硬件木马的芯片和含硬件木马的芯片通过主成份分析能够明显区分.

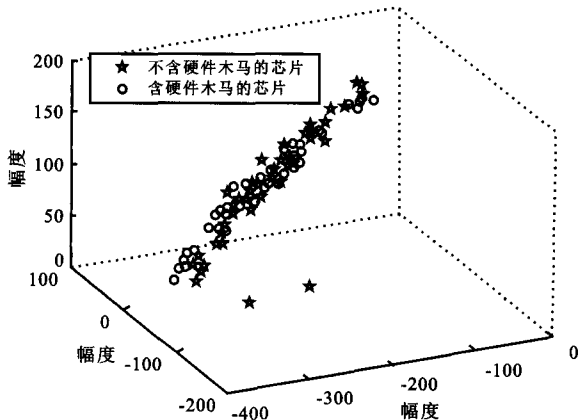


图 6 不含硬件木马和含硬件木马的密码芯片原始旁路信号对比

5 结束语

借鉴旁路密码分析的思想,文中根据含硬件木马和不含硬件木马的密码芯片对应的旁路信号之间差异,定义了基于旁路分析的密码芯片硬件木马检测问题,并介绍了基于旁路信号主成份分析的硬件

木马检测方法. 针对 FPGA 实现的 DES 密码芯片中硬件木马检测实验验证了该方法的有效性.

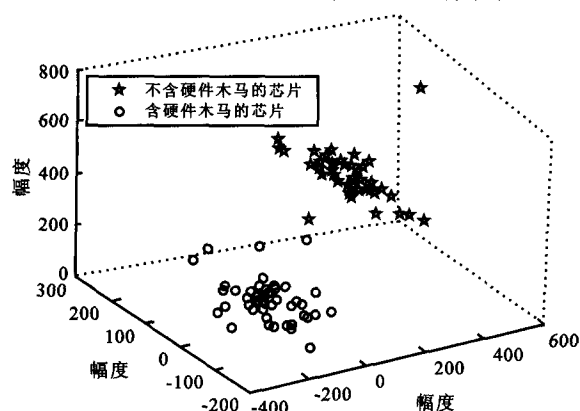


图7 不含硬件木马和含硬件木马的密码芯片旁路信号主成份分析结果对比

参考文献:

- [1] King S T, Tucek J, Cozzie A, et al. Designing and implementing malicious hardware[C]// Proceedings of the 1st USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET). San Francisco, 2008:1-8.
- [2] Soden J, Anderson R, Henderson C. IC failure analysis tools and techniques - magic, mystery, and science [C]// International Test Conference, Lecture Series II, Practical Aspects of IC Diagnosis and Failure Analysis;

A Walk through the Process. Washington, 1996:1-11.

- [3] Wang X, Tehranipour M, Plusquellic J. Detecting malicious inclusions in secure hardware: challenges and solutions[C]// 1st IEEE International Workshop on Hardware-Oriented Security and Trust (HOST). USA: California, 2008:15-19.
- [4] 邓高明, 赵强, 张鹏, 等. 针对密码芯片的电磁频域模板分析攻击[J]. 计算机学报, 2009, 32(4):602-610.
- [5] 张鹏, 邓高明, 赵强, 等. 基于 Cache 行为的旁路攻击[J]. 计算机工程, 2008, 34(21):20-22.
- [6] 邓高明, 赵强, 张鹏, 等. 基于 AES 算法的 Cache Hit 旁路攻击[J]. 计算机工程, 2008, 34(13):113-114.
- [7] 胡广书. 数字信号处理——理论、算法与实现[M]. 2版. 北京:清华大学出版社, 2003.
- [8] Alex Clark Baumgarten. Preventing integrated circuit piracy using reconfigurable logic barriers[D]. Ames, Iowa, Iowa State University, 2009.

作者简介:

赵崇征 男, (1977-), 博士研究生. 研究方向为银行卡防欺诈研究、基于旁路攻击理论的攻击与防御研究.

邓高明 男, (1983-), 博士研究生. 研究方向为电磁信息检测与主动防护技术等.

赵强 男, (1945-), 教授, 博士生导师. 研究方向为信息安全等.

(上接第4页)

- [6] 兰晓明, 颜峻, 马何平, 等. 一种宽频压控振荡器及高速双模预分频器的设计与实现[J]. 微电子学与计算机, 2010, 27(3): 45-48.
- [7] Ali M, Hegazi E. A multigigahertz multimodulus frequency divider in 90-nm CMOS[J]. IEEE Transaction on Circuits Systems II, Express Briefs, 2006, 53(12): 1333-1337.
- [8] Barale F, Sen P, Sarkar S, et al. Programmable frequency-divider for millimeter-wave PLL frequency synthesizers [C]// 38th European Microwave Conference. Amsterdam, 2008:460-463.

- [9] Kuo Y F, Wen R M. A dynamic-logic frequency divider for 5-GHz WLAN frequency synthesizer [C]// 17th International conference. Radioelektronika, 2007.

作者简介:

左红建 男, (1986-), 硕士研究生. 研究方向为数模混合集成电路设计.

郭阳 男, (1971-), 研究员. 研究方为超大规模集成电路设计与验证技术.

马卓 男, (1980-), 助理研究员. 研究方向为高性能模拟集成电路设计.