文章编号:1671-4598(2016)01-0196-03 DOI:10.16526/j.cnki.11—4762/tp.2016.01.054 中图分类号:TN918 文献标识码:A

基于核主成分分析的硬件木马检测方法研究

王晓晗,李雄伟,张 阳,陈开颜,徐 璐
(军械工程学院 信息工程系,石家庄 050003)

摘要:针对旁路信号样本在高维空间中的分布,提出了一种基于核主成分分析的硬件木马检测方法,该方法能够找出旁路信号样本分布中的非线性规律,将高维的旁路信号映射到低维子空间同时更精确地反映旁路信号样本的分布特性,从而发现由木马引起的非线性特征差异;针对AES加密电路植入约占电路3%的组合型木马并进行检测,实验结果表明,该方法能够有效分辨基准电路与含木马电路之间旁路信号的非线性特征差异,实现木马的检测,并取得比K—L变换更好的检测效果。

关键词:集成电路;硬件木马检测;旁路分析;核主成分分析

Hardware Trojan Detection Method Based on Kernel Principal Component Analysis

Wang Xiaohan, Li Xiongwei, Zhang Yang, Chen Kaiyan, Xu Lu
(Department of Information Engineering, Ordnance Engineering College, Shijiazhuang 050003, China)

Abstract: Aimed at the distribution of side-channel signal sample in high-dimensional space, a Hardware Trojan detection method based on kernel principal component analysis is proposed. This method can identify the nonlinear law in distribution of side-channel signal sample, project high-dimensional side-channel signal onto low-dimensional subspace, and more accurately reflect the distribution characteristics of side-channel signal sample, so that non-linear characteristics of differences caused by Trojan are found. In against the AES encryption circuit, the combined Trojan circuit about 3% is implemented, and detection experiments are performed. Experimental results show that, the method can effectively distinguish the difference about nonlinear characteristics between side-channel signal of the reference chip and side-channel signal of chip with Trojan, achieve the detection about Trojan, and get a better detection result than the K-L Transform.

Keywords: integrated circuit; hardware trojan detection; side-channel analysis; kernel principal component analysis

0 引言

硬件木马(hardware trojan)[1-2]是威胁集成电路(IC)安全的一种新型硬件攻击方式,其本质是对IC芯片原始设计的修改或者是植入的具有恶意功能的冗余电路,能够在特定条件下实现破坏性功能或者泄露芯片内部的秘密信息。硬件木马的检测十分困难,传统的功能测试和逻辑测试无法识别由木马激活而引起的额外功能,为了高效地认证芯片保证芯片免受硬件木马的侵袭,研究人员围绕硬件木马的预防与检测展开研究,并取得许多研究成果。其中木马的预防主要是修改原始电路设计,使木马更难植入或者更容易被检测,例如内建自检测技术(BIST)[3]、环形振荡器[4]等,这类方法能够起到预防木马的功用,但是增加了电路冗余,影响了集成电路芯片的性能。木马的检测主要是比对待测电路与参考基准之间的特征差异(由木马引起)判断待测电路中是否含有木马。其中,旁路分析技术是当下应用较广的检测方法[5],如Wang等人[6]对功耗旁路信号求积分,根据电路在一段时间内工作电流的累计差异来检测木马;Agrawal等人[7]采用K—L变换处理功耗旁路信号,投影到噪声的特征空间,建立“指纹”并进行比对检测。

通过实验验证发现,文献[7]中的K—L变换对于高维、分布复杂的旁路信号而言,检测效果并不理想。因此本文提出一种新的信号分析方法,通过核主成分分析寻找投影子空间,在提取特征时致力于发现旁路信号的分布规律,从而发现含木马IC与“金片”之间的旁路信号差异,并有效的检测出木马。

1 基于核主成分分析的检测方法

1.1 功耗旁路信号的统计分布特性

功耗旁路信号是基于旁路分析的硬件木马检测中应用较多的一种旁路信号,主要是在同等条件下测量“金片”与待测IC的工作电流,通过比较电流的差异来判断待测IC中是否含有木马。然而由于电路运行时产生的功耗旁路信号十分微弱,由实验设备和逻辑运算造成的噪声对功耗旁路信号的影响很大,使得同样条件下对同一集成电路芯片测量得到的一组旁路信号不尽相同,分布十分复杂,极大地干扰了硬件木马的检测,因此,有必要探究噪声影响下“金片”的功耗旁路信号分布。

用于检测的功耗旁路信号,本质上是离散多元随机变量($X(t_1)$, $X(t_2)$, ……, $X(t_n)$),对于“金片”的功耗旁路

收稿日期:2015-06-30; **修回日期**:2015-08-27。
基金项目:国家自然科学基金(61271152,51377170);河北省自然科学基金(F2012506008)。
作者简介:王晓晗(1992-),男,河北衡水人,硕士研究生,主要从事信息安全方向的研究。
李雄伟(1975-),男,河北定州人,副教授,博士,硕士生导师,主要从事信息安全方向的研究。

信号而言，在某一时刻其功耗组成为：(1) 主电路电流 (P_e)；(2) 全噪声 (P_n)[8]。其中，主电路电流是电路正常工作时产生的电流，若测试向量相同，电路执行相同的操作、处理相同的数据，则主电路电流相同。全噪声与实验设备和在某时刻并行执行的运算有关，其噪声波动范围随并行执行计算的增多而增大，但总是服从正态分布[8]，即 $P_n \sim N(0, \sigma^2)$，其中（可变。若将一条离散功耗旁路信号看作高维空间（n 维）中的一个样本点，则噪声的存在，使得高维空间中一组旁路信号样本在均值样本周围散布。由于某一时刻的噪声分布都服从正态分布，相邻时刻的噪声分布之间存在一定的相关性，且不同时刻噪声波动范围（方差）不确定，使得同样条件下测得的一组“金片”的功耗旁路信号样本的分布为超椭球（非线性）。

文献 [7] 中的 K－L 变换，本质上是根据旁路信号样本在高维空间中的分布来寻找一组正交的线性投影方向，使该方向上的投影尽可能好的反映“金片”功耗旁路信号样本的分布信息，根据投影方向上“金片”旁路信号样本投影与待测旁路信号样本投影的差异判断是否含有木马。然而该方法比较适合高维空间中沿一定方向分布的样本数据，不足以精确地刻画分布为超椭球的“金片”功耗旁路信号样本分布中的非线性规律，寻找合适的非线性（曲线）投影方向，可以更好的表述功耗旁路信号样本的分布特性，从而发现由木马引起的功耗旁路信号样本分布之间的差异。

1.2　理论分析

核主成分分析（Kernel PCA，KPCA）是由主成分分析发展的一种非线性变换方法[9]，是核函数方法中的一个特例，其基本思想是利用支持向量机中的核函数技巧来实现非线性变换，在变换后的特征空间中提取反映样本分布的主要特征，从而实现非线性特征提取。该方法与一般非线性方法相比具有很多优点：1）避免维数灾难。非线性变换能增加样本的特征维数，非线性程度越强，维数越高，呈指数增长趋势，而核函数方法的求解空间维数仅为样本总数，成功的克服了维数灾难。2）减少计算量。核函数方法成功的将非线性变换转换成线性问题，使得问题的求解只需采用线性手段，简化了问题的求解过程。3）保证推广能力，非线性变换使得样本维数增高，但没改变样本个数，降低了解的可信性（空间维数越高，样本越分散，越容易求解），而核函数方法的求解限制了空间维数，使得解仍具有一定的推广能力。

对于空间中的 n 个样本 X_i，若非线性映射为 φ，则样本在高维空间中的投影可以表述为 $\varphi(x)$，根据再生核理论，高维空间中样本投影的协方差矩阵为

$$\boldsymbol{K} = \frac{1}{n}\{K_{ij}\}_{n\times n}, i, j = 1\cdots n \tag{1}$$

其中：$K_{ij} = \varphi(x_i) \quad \varphi(x_j) = k(x_i, x_j)$，协方差矩阵 $\boldsymbol{K}$ 的前 d 个特征值对应特征向量 a^d 即为高维空间中的 d 个主成分方向（非线性），则样本在非线性主成分方向 a^d 上的投影为

$$z^d(x) = (a^d \cdot \varphi(x)) = \sum_{i=1}^{n} a_i{}^d k(x_i, x) \tag{2}$$

1.3　检测方案

基于功耗旁路信号的硬件木马检测一般是在同等条件下运行“金片”与待测 IC，并测量其工作电流，采用信号处理方法提取特征，通过比较特征的差异来判断待测 IC 中是否含有木马。根据该检测流程，下面给出基于核主成分分析的详细检测方案。

Step1：旁路信号的采集。对“金片”施加固定的测试向量并测量电路在该测试向量下的功耗旁路信号，得到一组无木马的功耗数据集 $B = \{b(i, j) \mid i=1, 2, \cdots, n; j=1, 2, \cdots, m\}$，其中，$n$ 表示功耗旁路信号的数量，m 表示每条功耗轨迹的采样长度。类似的，以同样的方法得到待测 IC 的功耗数据集 $T = \{t(i, j) \mid i=1, 2, \cdots, n; j=1, 2, \cdots, m\}$，通过比较两组数据集来判断待测芯片中是否含有木马。

Step2：计算非线性投影方向。同时处理两组数据集中的旁路信号样本，利用公式 (1) 计算高维空间中样本投影的总离散度矩阵 $\boldsymbol{K}$，随后计算矩阵 $\boldsymbol{K}$ 的一组特征向量 a^d（非线性投影方向），并按照对应特征值从大到小的顺序排列。

Step3：将两组数据集分 $\boldsymbol{B}$ 和 $\boldsymbol{T}$ 分别投影到每个非线性投影方向上，并获得投影数据集 B_d' 和 T_d'，分别统计每个非线性投影方向上两组样本投影的分布，根据“金片”的样本投影分布定义阈值，此处将“金片”的样本投影的 $\mu \pm 3\sigma$ 定为阈值，若在某个投影方向上待测 IC 的样本投影超出阈值的比例大于 1/2，则判定待测 IC 中含有木马，反之则不含木马。

2　硬件木马检测实验

2.1　实验配置

为了验证方法的有效性，搭建了基于 FPGA 的硬件木马检测平台，其原理如图 1 所示。平台选用 Xilinx 公司的 XC5VLX30 FPGA 芯片来模拟集成电路芯片，通过修改 FPGA 比特流来获取基准加密电路和木马电路，为了减少 FPGA 布局布线对检测效果的影响，采用 ISE 软件中的增量编译技术生成电路。平台采用 Tektronix DPO4032 示波器（带宽为 350 MHz）采集 FPGA 芯片正常工作时产生的功耗旁路信号，在计算机（CPU 为 I5－4570，主频为 3.20 GHz，内存为 4 G）上采用 LabVIEW 编写的虚拟仪器控制平台控制旁路信号的采集，并用 Matlab R2009 处理采集到的功耗旁路信号，整个检测过程都采用 USB 数据线传输数据。

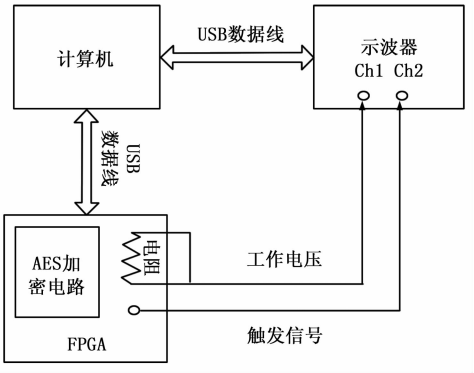

图 1　实验平台原理图

在该 FPGA 芯片中植入高级加密标准（advanced encryption standard，AES－128）电路，并针对该电路设计了组合型硬件木马，木马电路约占 AES 加密电路的 3%，其电路结构图如图 2 所示。其中 a1 至 an 代表加密过程中分组明文和密钥

的某些数据位，经过一定规则的逻辑运算（触发）使负载工作，当输入满足一定条件时修改加密结果 s1，使加密发生故障。这种木马结构简单，易于实现，通过修改运算规则更改木马规模，功能不发生变化，且木马始终处于运行状态，便于检测，不需要特定的测试向量来激活木马。

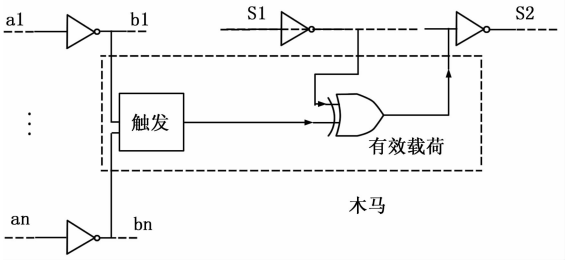

图 2 木马电路示意图

采集时，固定一组加密明文和密钥，对每个测量电路重复采集 1 000 条功耗旁路信号，每条旁路信号采样长度为 1 000（采样频率为 2.50 GSa/s，采样时长为 0.4），为了削弱噪声的影响，每条旁路信号样本为重复 16 次采样求平均得到的结果。

2.2 实验结果

采用 K－L 变换的实验结果如图 3 所示。其中，横坐标表示按照特征值大小排序的特征向量，纵坐标为每个特征向量上的投影值，红色部分为“金片”的投影分布，绿色部分为含木马 IC 的投影分布，两条黑线为根据“金片”的样本投影确定的 $\mu\pm3\sigma$ 检测边界，经统计，K－L 变换在第二个投影方向上超出检测边界的样本投影数最多仅有 390 个，不能检测出木马。

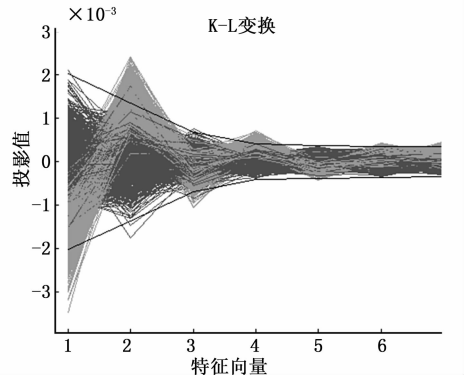

图 3 基于 K－L 变换的木马检测

采用核主成分分析的实验结果如图 4 所示，此处选取的核函数为径向基（RBF）核函数 $k(x,x')=\exp(-\frac{\|x-x'\|^2}{\sigma^2})$，当参数 σ^2 设为 0.000 01 时，该方法达到最优的检测效果，此时超出边界的样本数达 791 个（第二个特征向量处），能够检测出木马。

3 结束语

基于核主成分分析的硬件木马检测能够更为有效的刻画功耗旁路信号的分布特性，提取样本分布中的非线性分布规律，

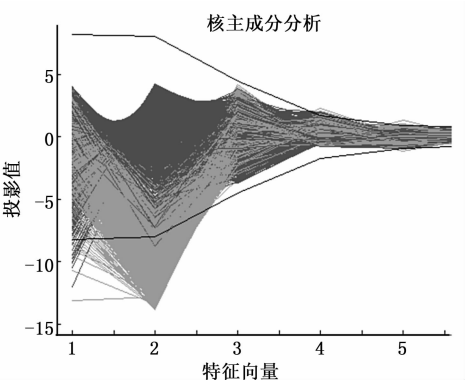

图 4 基于核主成分分析的木马检测

取得了比 K－L 变换更好的检测效果，为硬件木马的检测提供新的思路。然而该方法仍然有一些不足，一方面本文选用常用核函数进行计算，其效果对于木马检测未必最优；另一方面，该方法的计算代价会随迭代次数的增加而成倍增长，不利于木马的快速检测。针对这两点不足，在下一步的工作中将选取其他核函数或者采用多个核函数进行计算，以期进一步提高木马检测精度和稳定性。同时，还寻找启发式寻优算法，寻找自适应的最优参数 σ^2，减少迭代次数，提高木马检测效率。

参考文献：

[1] Tehranipoor M, Koushanfar F. A Survey of hardware Trojan Taxonomy and Detection [J]. IEEE Design & Test of Computers, 2010, 27 (1): 10－25.

[2] 李雄伟，徐 徐，张 阳，等. 基于简单电磁分析的硬件木马设计 [J]. 计算机测量与控制，2013，21 (12)：3396－3398.

[3] Chakraborty R, Paul S, Bhunia S. On－demand Transparency for Improving Hardware Trojan Detectability [A]. in Proceedings of IEEE International Workshop on Hardware－ Oriented Security and Trust (HOST' 2008) [C] . 2008: 48－50.

[4] Tehranipoor M, Salmani H, Zhang X, et al. Trustworthy Hardware: Trojan Detection and Design－for－Trust Challenges [J]. the IEEE Computer Society, 2011, 44 (7): 66－74.

[5] Chakraborty R, Narasimhan S, Bhunia S. Hardware Trojan: Threats and emerging solutions [A] . in proceedings of the IEEE International Workshop on High Level Design Validation and Test Workshop [C] . 2009: 166－171.

[6] Wang X, Salmani H, Tehranipoor M, et al. Hardware Trojan Detection and Isolation Using Current Integration and Localized Current Analysis [A]. in Proceedings of the International Symposium on Fault and Defect Tolerance in VLSI Systems (DFT' 2008) [C] . 2008: 87－95.

[7] Agrawal D, Baktir S, Karakoyunlu D, et al. Trojan Detection Using IC Fingerprinting [A], in Proceedings of the Symposium on Security and Privacy (SP' 2007) [C] . 2007: 296－310.

[8] S. Mangard, E. Oswald, T. Popp 著. 冯登国 等译. 能量分析攻击 [M]. 北京：科学出版社，2010.

[9] Kirby M, Sirovich L. Application of the K－L procedure for the characterization of human faces [J]. IEEE Transaction on Pattern Analysis Machine Intelligence, 1990, 12 (1): 103－108.