

引用格式:张磊,殷梦婕,王建新,等. 基于随机森林的硬件木马检测方法 [J]. 微电子学与计算机, 2019, 36(2): 83-87.
[ZHANG L, YIU M J, WANG J X, et al. Hardware trojan detection method based on random forest [J]. Microeletronics & Computer, 2019, 36(2): 83-87.]

基于随机森林的硬件木马检测方法

张 磊,殷梦婕,王建新,董有恒,肖超恩,刘东阳,赵 成

(北京电子科技学院, 北京 100070)

摘 要: 针对 BP 神经网络和 SVM 这两种机器学习算法中存在参数选择困难和时间开销较大的问题, 本文提出了一种基于随机森林的硬件木马分类方法. 首先, 将硬件木马检测转化为二元分类问题, 对芯片的能量消耗进行多次采样, 再通过 PCA 对功耗曲线进行特征提取, 最后利用随机森林分类模型对特征向量进行分类, 达到检测硬件木马芯片的目的. 实验结果表明, 经 PCA 处理的相同硬件木马数据, 随机森林的判别准确率与 BP 神经网络相比提高了 9.13%, 与 SVM 方法相比判别准确率提高了 15.96%. 而相比其他两种方法, 时间开销也降低了 8 倍左右.

关键词: 侧信道分析; 硬件木马; 分类器; 随机森林

中图分类号: TP391.2

文献标识码: A

文章编号: 1000-7180(2019)02-0083-05

Hardware Trojan Detection Method Based on Random Forest

ZHANG Lei, YIN Meng-jie, WANG Jian-xin, DONG You-heng,

XIAO Chao-en, ZHAO Cheng

(Beijing Electronic Science & Technology Institute, Beijing 100070, China)

Abstract: Aiming at the problem of parameter selection and time overhead for BP neural network and SVM algorithms, this paper proposes a Hardware Trojan classification method based on Random Forests. Firstly, the Hardware Trojan detection problem is modeled as a binary classification problem and the power consumption of the chip is sampled several times. Then the characteristics of the power consumption curve are extracted by the PCA (principal component analysis). Finally, RF (Random Forests) classification model is used to classify the feature vectors in purpose of identifying Hardware Trojan chips. The experimental results show that, considering the same Hardware Trojan horse data processed by PCA, the discrimination accuracy of RF is improved by 9.13% compared with the BP neural network. Compared with the SVM (support vector machine) method, the discrimination accuracy is increased by 15.96%. Compared to the other two methods, the time cost of RF is reduced by about 8 times.

Key words: side-channel analysis; hardware trojan; classifier; random forests

1 引言

由于集成电路设计环节和制造过程的全球化, 集成电路更容易受到恶意更改和敌手攻击, 用户也更难控制集成电路供应链中的各个环节^[1-2]. 硬件木马可以达到操纵掩码改变设计网表, 恶意植入额外

逻辑电路的目的, 而不影响设计的主要功能. 在 RTL 设计、生产和代工阶段中, 各个环节芯片都有可能被植入硬件木马, 在军事系统、金融基础设施和运输安全等领域造成危害^[3-4]. 国际上硬件木马的研究发展迅速, 分类判别方法也越来越多, 其中包括失效分析、逻辑测试、木马激活以及侧信道分析等^[5].

失效分析方法耗时长、成本高,随着芯片集成度的提高,结构会变得越来越复杂,检测效率也大大降低.逻辑测试的方法所需的穷举测试十分耗时,特别是对诸如系统芯片等大规模集成电路的检测难度较大^[6].王力伟等人^[7]通过马氏距离判别法对信号进行统计分析,能够识别出在时域上无法区分的面积比为 2.72% 的硬件木马电路.薛明富等人^[8]首先利用时域约束估计器分离信号空间与噪声空间,然后通过空间投影与 PCA 凸显出信号特征,最终在较大的工艺偏差和测量噪声的情况下检测到微弱的木马信号.苏静等人^[9]通过对低秩信号上的主成分特征提取,运用欧式距离算法完成无木马芯片和硬件木马芯片的分类.

随着硬件木马检测技术的深入研究,基于马氏距离等分类判别方法难以进行有效检测.本文提出一种基于随机森林算法的硬件木马检测方法.

2 基于随机森林的硬件木马检测方法

随机森林是一种广泛应用于数据分类和非参数预测的机器学习算法^[10].在处理高维数据时具有泛化能力强,并行性良好,时间开销低的优势.在特定环境下对 FPGA 芯片进行多次侧信道信号采集,可以将其看做是在理想状态下对多个相同芯片进行单次采集.利用硬件木马检测模型,在特定环境下多次采集芯片运行功耗,并通过与模板功耗的对比,实现对硬件木马的检测.基于随机森林的硬件木马检测流程如图 1 所示.

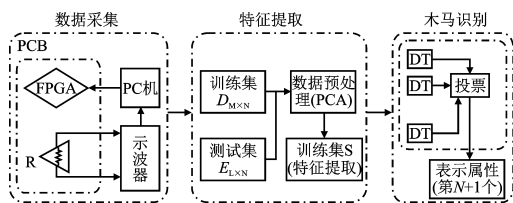


图 1 基于随机森林的硬件木马检测流程图

基于随机森林的硬件木马检测方法可分为三个步骤.

第一步是数据采集.以 Altera Cyclone IV E 最小系统板为核心,下载标准 AES 加密算法程序,并选取部分芯片加入木马模块.在芯片的接地引脚上连接一个大小合适的电阻,选用 Tektronix DPO 7104C 示波器测量出电阻的电压值,输出并显示在于数字示波器.示波器通过路由器与 PC 机连接,将

无木马芯片和木马芯片的功耗曲线传输到 PC 机并存储.

第二步是特征提取.在侧信道数据采集中,每条特征曲线代表一个训练样本,构成数据集的每一行;对每条特征曲线采样 N 次,表征为每个样本的 N 个属性,构成数据集的每一列.假设芯片共 $M+L$ 个,其中 M 个芯片信号用于训练,剩余 L 个芯片信号用于测试.最终得到原始训练集 $D_{M \times N}$ 和测试集 $E_{L \times N}$.再对数据集中每个样本设定标识属性,标记此样本代表的是干净芯片还是硬件木马芯片.将原始训练集 $D_{M \times N}$ 经过 PCA 处理,得到有效矩阵,即训练集 S .

第三步是木马识别.选定随机森林算法创建模型,根据训练样本的标识属性,进行训练和评估.根据训练集生成随机森林分类模型 C .考虑测试集 $E_{L \times N}$,根据分类模型 C 预测得出测试集中的第 $N+1$ 个标识属性;根据预测的标识属性,计算出随机森林模型正确识别的比例.

随机森林分类算法在实际应用中也存在一些问题^[11],如面对特征较少的数据,不能产生很好的分类结果;只能尝试不同的参数而无法控制模型内部的运行等.因此,本文在相同环境和条件下,采集标准芯片与待测芯片的工作功耗,构成训练集 $D_{M \times N}$ 和测试集 $E_{L \times N}$,选用随机森林算法进行有监督的模型训练,然后采用训练成功的随机森林模型 C 对测试集 $E_{L \times N}$ 进行检测,分离无木马芯片和硬件木马芯片.

3 随机森林分类模型

本文设计的随机森林分类模型主要由两部分构成:PCA 模块和随机森林模块.

3.1 PCA 模块

PCA 是一种对数据进行有效降维变换的方法,对于维度高、数据量大的数据表现良好^[12].它通过线性变换将数据从高维空间映射到低维空间,提取相关性较低的数据分量,用少量的数据分量表征原始高维数据,得到样本的主要特征^[13].PCA 算法实现流程如图 3 所示.

第一步分别计算原始训练矩阵 $D_{M \times N}$ 的均值和标准差,使用归一化方法,得到符合标准正态分布的功耗数据.

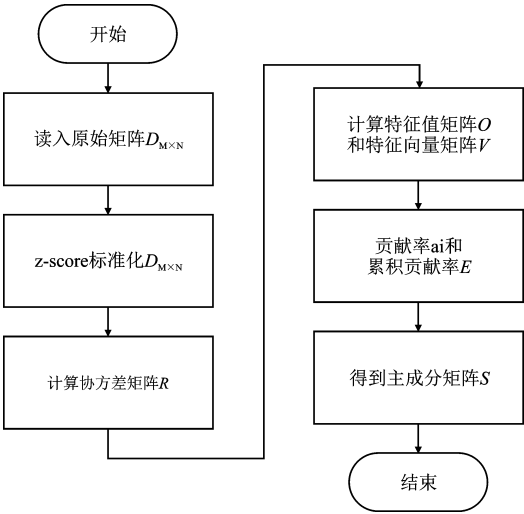


图 2 PCA 算法流程图

$$x_{ij} = \frac{x_{ij} - \bar{x}_j}{\sqrt{\text{var}(x_j)}} \quad (i = 1, 2, \dots, m; j = 1, 2, \dots, n) \quad (1)$$

式中, x_{ij} 和 $\text{var}(x_j)$ 为 $D_{M \times N}$ 中第 j 列的均差和标准差.

第二步求得协方差矩阵 R .

$$r_{ij} = \frac{1}{m-1} \sum_{i=1}^m x_{ii} x_{ij} \quad (i, j = 1, 2, \dots, n) \quad (2)$$

式中, r_{ij} 为协方差矩阵 R 中第 i 行第 j 列的值.

第三步由协方差矩阵 R 计算出特征值矩阵 O 和相应的特征向量矩阵 V , 最后将矩阵 O 中特征值向量按照从大到小的顺序排列, 获得矩阵 E .

$$E = (e_1, e_2, \dots, e_n) \quad (e_1 > e_2 > \dots > e_n) \quad (3)$$

第四步按照排列的排序, 计算各个向量的贡献率, 根据贡献率之和, 选取前 k 个向量构成投影变换矩阵 P .

$$\frac{\sum_{j=1}^k e_j}{\sum_{j=1}^n e_j} \geq 0.96 \quad (4)$$

$$P = (e_1, e_2, \dots, e_k) \quad (e_1 > e_2 > \dots > e_k) \quad (5)$$

第五步得到主成分矩阵 S .

$$S = D \times P \quad (6)$$

利用侧信道分析进行硬件木马检测, 实际上就是对多维数据进行分类, 选取 PCA 方法能够有效对大量功耗曲线进行特征提取.

3.2 随机森林模块

随机森林是一个构建并结合多个决策树的分类模型. 当随机森林处理测试集时, 实质上就是让每一棵决策树进行判断, 最终随机森林的判定结果由每棵树的输出标识的众数决定^[14]. 随机森林算法流程如图 3 所示.

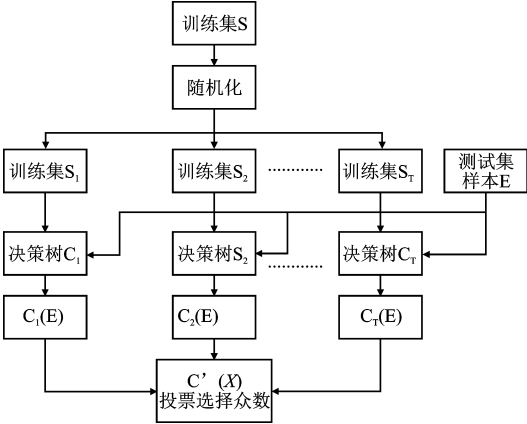


图 3 随机森林算法流程图

随机森林是集成学习中较典型的算法之一, 其中决策树的形成采用自助法重采样技术. 子集样本和属性抽取的随机性保证了随机森林的通用性和健壮性^[15]. 单棵决策树生成过程如下: 从训练集 S 中进行有放回的随机抽取, 重复 T 次, 得到一个新的训练子集, 根据这个训练子集, 得到一棵决策树. 假设有 Q 个属性, 则在这棵决策树的每个节点进行分裂时, 在从 Q 个属性中, 相同概率的随机抽取 q 个属性, 得到一个属性子集 R , 再依据基尼值最小、纯度最高的原则, 从属性子集 R 中选择一个最优属性 r 进行节点分裂生长, 之后再构造各个分支重复上述流程.

随机森林生成的具体步骤如下:

- (1) 利用自助法对训练集 S 进行随机有放回的抽取, 生成 T 个训练子集 $S_1, S_2, \dots, S_T$, 由此生成决策树 $C_1, C_2, \dots, C_T$.
- (2) 设每个决策树中属性为 Q 个, 从中随机抽取 q 个作为当前节点的属性子集, 通过计算每个属性的基尼指数, 选择基尼指数最小的属性, 通过这种方式对每棵树的每个非叶子节点进行分裂.
- (3) 每棵树都尽情分裂生长, 不进行任何裁剪.
- (4) 让每棵决策树对测试集 $E_{L \times N}$ 进行测试, 得到各自的分类标识 $C_1(E), C_2(E), \dots, C_T(E)$.
- (5) 采用投票法, 统计 T 组分类标识, 输出最多的标识就是测试集 $E_{M \times N}$ 的分类结果.

4 算法比较

在基于机器学习的硬件木马检测方法中,经常使用 BP 神经网络算法和 SVM 判别分析方法, BP 神经网络算法具有较强的非线性映射能力,通过构建前馈多级网络,从网络预测误差的负梯度方向,利用梯度校正法来修正权值和阈值,鉴别准确率较高,但也存在学习速度较慢,时间消耗过大,容易陷入过度拟合的缺陷. SVM 算法在求解小样本、非线性及高维数据判别中表现良好,但其惩罚参数 c 和核函数参数 g 的选择取决于测试经验,并且带有一定的随机性^[16].

本文选取了 BP 神经网络^[17]和 SVM^[18]方法为对比实验对象,在相同的实验条件下,对测试集进行分类判别,得到不同判别方法的鉴别结果,并比较各种算法在 PCA 前后的性能. 其中随机森林设定决策树数目为 50, SVM 使用径向基核函数,神经网络使用 3 层前馈网络,性能对比结果如表 1 所示.

表 1 与其它检测方法的性能对比表

是否降维	性能判别方法	检测准确率(%)	时间开销(s)
否	SVM	68.00	153.45
	BP 神经网络	76.51	21.40
	随机森林	90.27	232.21
是	SVM_PCA	79.05	9.31
	BP 神经网络_PCA	85.88	8.73
	随机森林_PCA	95.01	1.09

对于硬件木马检测,评估指标主要是:检测准确率及检测时间. 根据实验,可以看出,当不使用 PCA 降维时,BP 神经网络和 SVM 的检测准确率相似. 与随机森林算法相比,随机森林识别的准确率与 BP 神经网络高 13.76%. 与 SVM 方法相比提高了 22.27%. 然而由于数据的复杂性和算法的复杂度,随机森林和 SVM 方法时间开销较长,检测速度较低. 对数据进行 PCA 特征提取后,三种检测方法的检测准确率都得到了相应的提高,时间开销均大幅度降低. 准确率得到保证的前提下,PCA 弥补了随机森林时间开销上的缺陷,使得随机森林模型表现较好. 综合考虑检测准确率和时间开销,相较于其他两种分类方法,在二元分类硬件木马检测中,随机森林检测方法性能稳定,表现更佳.

5 结束语

针对基于 BP 神经网络和 SVM 算法的硬件木

马检测方法存在的问题,本文提出了一种基于随机森林算法的硬件木马检测方法. 首先,采集硬件木马芯片和无木马芯片的功耗曲线,然后通过 PCA 降维的方法提取原始数据,去除数据冗余,最后利用随机森林分类方法达到硬件木马的识别和分类.

(1)经过多次试验,综合错误样本个数与位置,进一步增强了随机森林分类模型的性能.

(2)考虑 PCA 处理的相同的硬件木马数据量,相较于 BP 神经网络和 SVM 方法相比,大幅度减少了计算所需的时间,均降低了 8 倍左右.

(3)随机森林的判别准确率与 BP 神经网络相比提高了 9.13%,与 SVM 方法相比,判别准确率提高了 15.96%.

(4)随机森林算法高效通用,灵活抗噪,在处理具体规则模糊的数据时优势明显.

参考文献:

[1] 赵志勋,倪林,李少青. 硬件木马电路功耗的检测方法[J]. 北京邮电大学学报,2015, 38(4):128-132.

[2] 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.

[3] 糜旗,胡麒,徐超,等. 硬件木马的分析与检测[J]. 计算机与现代化,2012(11):155-157.

[4] 陈华锋,瞿有甜,姜燕冰. 硬件木马电路检测技术及发展趋势[J]. 浙江大学学报(理学版),2014, 41(1):49-51.

[5] 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):6:1-23.

[6] Moein S, Subramnian J, Gulliver T A, et al. Classification of hardware trojan detection techniques[C]// Tenth International Conference on Computer Engineering & Systems. IEEE, 2016:23-24.

[7] 王力纬,贾鲲鹏,方文啸,等. 基于马氏距离的硬件木马检测方法[J]. 微电子学,2013, 43(6):817-820.

[8] 薛明富,胡爱群,刘威,等. 基于子空间域特征提取的硬件木马检测方法[J]. 东南大学学报(自然科学版),2014, 44(3):457-461.

[9] 苏静,赵毅强,何家骥,等. 旁路信号主成分分析的欧式距离硬件木马检测[J]. 微电子学与计算机,2015,32(1):1-4.

[10] 姚登举,杨静,詹晓娟. 基于随机森林的特征选择算法[J]. 吉林大学学报(工学版),2014, 44(1):

137-141.

[11] 张华伟,王明文,甘丽新. 基于随机森林的文本分类模型研究[J]. 山东大学学报(理学版), 2006, 41(3): 5-9.

[12] 林海明,杜子芳. 主成分分析综合评价应该注意的问题[J]. 统计研究,2013, 30(8):25-31.

[13] 王文波,张晓东,汪祥莉. 基于主成分分析的经验模态分解消噪方法[J]. 电子学报,2013, 41(7):1425-1430.

[14] Ristin M, Guillaumin M, Gall J, et al. Incremental Learning of Random Forests for Large-Scale Image Classification[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, 38(3):490-503.

[15] Zhang H R, Min F. Three-way recommender systems based on random forests[J]. Knowledge-Based Systems, 2016, 91(C):275-286.

[16] 李欣海. 随机森林模型在分类与回归分析中的应用[J]. 应用昆虫学报, 2013, 50(4):1190-1197.

[17] 董师师,黄哲学. 随机森林理论浅析[J]. 集成技术, 2013,2(1):1-7.

[18] Xu J, Ramos S. Hierarchical Adaptive Structural SVM for Domain Adaptation[J]. International Journal of Computer Vision, 2016, 119(2):1-20.

作者简介:

张磊男,(1979-),博士,副教授. 研究方向为信息安全.

殷梦婕(通讯作者)女,(1994-),硕士研究生. 研究方向为信息安全. E-mail:244241686i@qq.com.

王建新男,(1977-),博士,副教授. 研究方向为电磁防护.

董有恒男,(1995-),硕士研究生. 研究方向为信息安全.