基于集成学习的功耗分析研究

刘 飚[1]，潘 扬[1*]，许盛伟[1]，李佳丽[2]，封化民[1,2]

(1.北京电子科技学院管理系 北京 丰台区 100070；2. 西安电子科技大学计算机科学与技术学院 西安 710071)

【摘要】针对单模型分类算法在训练样本数量较少时成功率偏低的问题，提出一种集成学习算法，并在DPA_Contest_V4数据集上进行实验。首先使用传统方法破解循环掩码，再使用SVM(support vector machine)、随机森林和k近邻(k-nearest neighbor, kNN)等分类算法进行训练和预测，最后将这些模型的结果集成。实验结果表明，集成模型优于单一模型，尤其当训练集中的能量迹数目较少时集成模型的成功率比单一模型高10%左右。

关 键 词 集成学习； k近邻； 功耗分析； 随机森林； 支持向量机

中图分类号 TP309 **文献标志码** A doi:10.3969/j.issn.1001-0548.2019.02.015

Research of Power Analysis Based on Ensemble Model

LIU Biao[1], PAN Yang[1*], XU Sheng-wei[1], LI Jia-li[2], and FENG Hua-min[1,2]

(1. Department of Management, Beijing Electronic Science and Technology Institution Fengtai Beijing 100070;
2. School of Computer Science and Technology, Xidian University Xi'an 710071)

Abstract Aiming at the problem that the single model classification algorithm has a low success rate when the number of training samples is low, an ensemble learning algorithm is presented in this paper. The experiment was conducted by applying DPA_Contest_V4 dataset. First the traditional method is used to break the mask, and then SVM, RF and kNN classification algorithms are applied to train and predict. Finally, the results of these models are combined as an ensemble model. The experimental results show that the integrated model is superior to the single model, and the success rate of the ensemble model can be about 10% higher than that of the single model especially when the number of training samples is low.

Key words ensemble learning; kNN; power analysis; random forest(RF); SVM

在对密码设备进行功耗分析攻击时，攻击者需要建立密钥或者与密钥关联的数据值与被攻击设备的功耗相关性模型，借此通过对功耗的分析破解出敏感信息。因为不同的密钥值会在能量迹的特定位置反映出不同的特征，因此可以把猜测密钥转化为一个分类问题[1]。

模板攻击本质上是一种分类方法，但是这种方法对能量迹的数目要求比较高[2]，实际攻击中往往难以达成。为解决这个问题，文献[3]将SVM等机器学习算法应用到侧信道攻击中，结果表明可以降低对能量迹的数目和质量的要求。

考虑到单模型分类算法在能量迹数目较少的情况下准确率会大幅下降，同时为进一步提高分类的准确率，本文将集成学习的方法首次应用在功耗分析领域，该方法综合了多种攻击方法，使得在能量迹较少的情况下，分类器的准确率可以获得显著的提升，同时在能量迹足够的情况下准确率也可以获得一定的提升。

1 集成学习

集成学习通过结合多个学习器来完成学习任务。通常先产生一组学习器，称之为个体学习器，再通过某种策略将这些个体学习器进行结合。

	样本1	样本2	样本3
模型1	√	√	×
模型2	×	√	√
模型3	√	×	√
集成	√	√	√

图1 投票法提高准确率

通常情况下，将性能相当的多个学习器相集成会获得一个性能比个体学习器更好的学习器，如图1

收稿日期：2017－09－08；修回日期：2017－11－01
基金项目：国家重点研发计划(2018YFB0803600)；中央高校基本科研业务费专项资金(328201507)
作者简介：刘飚(1981－)，男，博士，主要从事侧信道攻击与机器学习方面的研究.
通信作者：潘扬，E-mail: 136639355@qq.com

所示，虽然3个个体学习器都只有66%的正确率，但用投票法集成后正确率就提高到了100%。

常用的集成算法有Bagging，Boosting，此外还有一些较为简单的结合策略。例如：1) 平均法，对多个个体学习器的输出结果取平均或加权平均得出最终结果；2) 投票法，在分类任务中，对多个个体学习器的结果进行投票，当某一分类结果最多时，采用该分类结果；3) 学习法，通过一个学习器来结合，该学习器称之为次级学习器，个体学习器称之为初级学习器，训练时用初级学习器的输出来训练次级学习器，考虑到过拟合的问题，初级学习器的输出结果应采用交叉验证，即使用训练初级学习器时未使用的样本来产生结果[4]。

本文采用两种集成策略，一是线性加权集成，二是投票集成。由于中间值汉明重量的大小与能量迹上相应点的电压值大小存在线性关系[5]，因此本文使用支持向量机、随机森林和kNN时均对汉明重量进行回归预测，而不是分类预测，输出结果的时候再将回归值转化为具体值。

在进行基于线性加权的集成学习(weighted averaging ensemble learning, WAEL)的过程中，首先分别训练出多个个体学习器，再使用这些个体学习器的训练结果依照最小二乘法计算每个学习器的权重，具体步骤见本文2.3节。

在进行基于投票法的集成学习(voting ensemble learning, VEL)的过程中，首先分别训练出多个个体学习器，然后对这些学习器的预测结果进行投票，具体步骤见本文2.4节。

通常，为使集成学习具有较好的结果，待集成的个体学习器应当具有大致相当且不算太坏的性能，同时个体学习器之间应当有较大的不同，即所谓和而不同，综合考虑以上因素，本文选择支持向量机、随机森林和k近邻3种算法作为个体学习器来构成集成模型。

1.1 支持向量机(SVM)

SVM是一种建立在统计学习理论和结构风险最小原理基础上的机器学习算法，可用于处理分类和回归问题。

SVM的基本思想是在样本空间中划分一个超平面，将不同的样本分开。超平面可由线性方程 $\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{x}+b=0$ 表示，$\boldsymbol{\omega}$ 表示法向量，b 表示位移项，如图2所示。

优化的目标是找出间隔 γ 为最大时的超平面。通常样本在低维空间并非线性可分，因此，借助核函数将样本从原始空间映射到更高维的空间中，使样本在高维空间表现出线性可分的特性，通过求解下式中的 α 即可求出最终的模型。

$$\max_{\alpha}\sum_{i=1}^{m}\alpha_i-\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_i\alpha_j y_i y_j \kappa(x_i,x_j)$$

$$\text{s.t.}\sum_{i=1}^{m}\alpha_i y_i=0 \quad \alpha_i\geqslant 0, i=1,2,\cdots,m$$

式中，α 为拉格朗日乘子；$\kappa()$ 为核函数[4]。

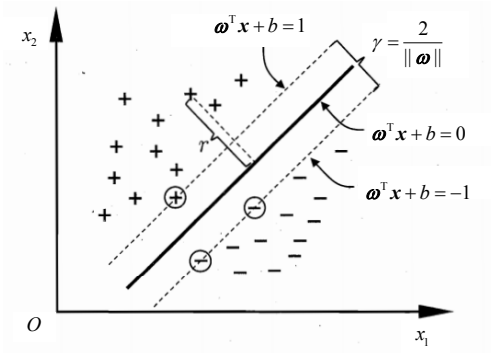

图2 支持向量机

1.2 随机森林(RF)

随机森林是一种以决策树为基学习器构建Bagging的集成学习算法。

1.2.1 决策树

决策树是一种基于树结构进行决策的分类算法，在分类的时候，通常会根据多个特征进行一系列的判断。

如图3所示，根节点为整个数据集，每个子节点则代表数据集的一部分，叶子节点则代表一部分最终的分类结果。该算法的最终目的是生成一个泛化能力强的决策树，可对未知数据保持良好的预测能力[4]。

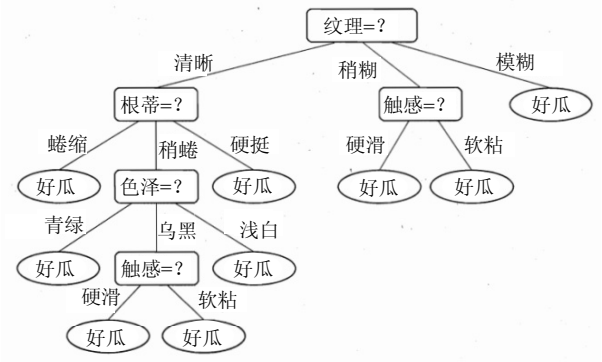

图3 决策树

1.2.2 Bagging

Bagging是一种并行式学习算法。这种算法使用自助采样法从原数据集中采样，即对原数据集进行

m 次有放回的采样，将获得的 m 个样本作为采样集，重复上述操作 T 次即可获得 T 个含 m 个样本的采样集，然后对每个采样集单独训练一个模型，最终将这些模型通过投票法或平均法进行结合，得出最终预测结果[4]。

随机森林在Bagging的基础之上，还在决策树的训练中引入了随机属性选择，即对于每个决策树节点，在该节点的属性集合中随机选出一个包含k个属性的子集，再从中选出一个最优属性进行划分，而k则是为控制随机性而引入的参数。这样就可以使不同的决策树之间不仅有样本的差异，而且还有属性的差异，使最终结果有进一步的提升[4]。

1.3　k近邻算法(kNN)

k近邻算法是一种无需训练的分类算法，在预测新样本时，将该样本与训练集中的样本进行比较，然后提取训练集中与待预测样本的特征最相似的样本的标签作为预测结果。通常选取训练集中前k个最相似的数据中出现次数最多的分类(分类问题)或标签的平均值(回归问题)作为对新样本的预测结果，本文在计算样本之间的相似度时采用欧氏距离公式：$d=\sqrt{\sum_{i=0}^{n}(xA_i-xB_i)^2}$，其中$d$为样本距离，$n$为样本的特征数。

2　基于集成学习的功耗分析

本文采用DPA_Contest_V4数据集[6]，该数据集采集自带有循环掩码的AES-256加密算法。攻击过程中，首先运用SVM破解掩码[7-8]，再对S盒输出值的汉明重量分别用SVM、RF和kNN进行训练并预测；最后将这3个单模型的结果通过线性加权或投票法集成，得出最终结果。

2.1　能量迹特征提取

研究表明，掩码偏移量或中间值的汉明重量仅与能量迹中很小一部分的点有较高的相关性，因此，为提高计算的准确性并降低计算量，需要对能量迹进行特征提取。提取特征的方法有主成分分析法[9]、相关系数法[10]等。本文采用相关系数法来进行特征提取，具体步骤如下。

选取所有能量迹上某一时刻的电压值 $v_{i,j}(i\in[1,N],j\in[1,T])$ 和所有能量迹的偏移量 $\mathrm{offset}_i(i\in[1,N])$ 或中间值的汉明重量，其中 N 为用于提取特征的能量迹总数，T 为数据集中能量迹上的采样点的总数，根据皮尔逊相关系数公式：

$$\rho=\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{D}(X)}\sqrt{\mathrm{D}(Y)}}=\frac{\mathrm{E}((X-\mathrm{E}(X))(Y-\mathrm{E}(Y)))}{\sqrt{\mathrm{D}(X)}\sqrt{\mathrm{D}(Y)}}$$

计算两个变量的相关系数[10]。再对所有时刻的相关系数的绝对值进行排序，根据相关系数最大的 m 个点所对应的时刻在能量迹上进行采样。

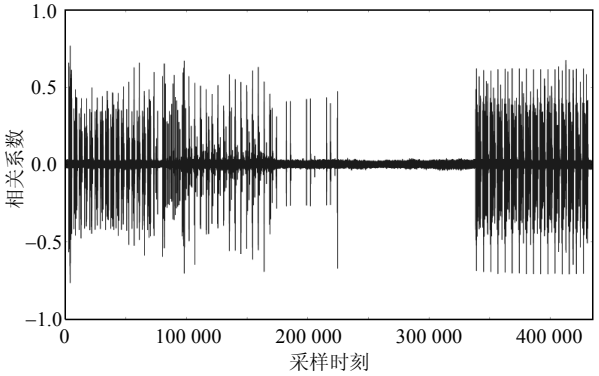

图4　掩码偏移量与能量迹上的点的相关系数

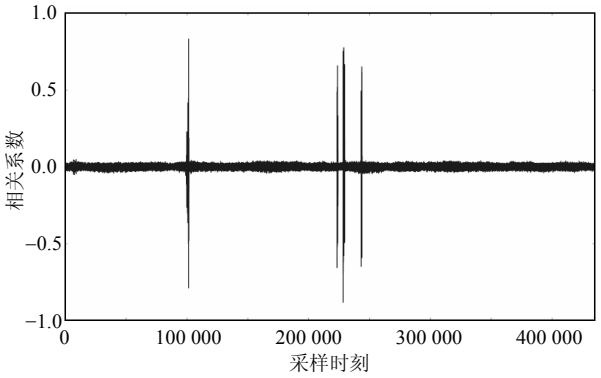

图5　第一个S盒的输出值的汉明重量与能量迹上的点的相关系数

如图4和图5所示，相关系数在某些时刻出现明显尖峰，表明在这些时刻发生了与掩码偏移量或中间值相关性较大的运算，比如明文盲化或S盒运算，本文的目标就是在出现尖峰的时刻提取特征。其中，掩码破解阶段依照图4提取特征，集成学习阶段依照图5提取特征。

2.2　基于SVM的掩码破解

由于数据集中的能量迹仅包括AES-256的第一轮加密，所以只考虑第一轮加密过程中的掩码S盒。

掩码运算步骤：

1) 选择16个8位数作为基本掩码，该数据集选择的这16个8位数 $M_i(i\in[0,15])$ 为{0x00; 0x0f; 0x36; 0x39; 0x53; 0x5c; 0x65; 0x6a;0x95; 0x9a; 0xa3; 0xac; 0xc6; 0xc9; 0xf0; 0xff}

2) 生成16个掩码S盒，满足 $\mathrm{MaskedSubBytes}_i(X_i)=\mathrm{SubBytes}(X_i\oplus M_i)\oplus M_{i+1}$，其中 X_i 表示某一位明文字节，$i\in[0,15]$。

3) 计算第一轮掩码S盒的输出：

$$\text{for } i \in [0,15] \quad \text{do}$$

$$X_i = X_i \oplus M_{\text{offset}}(\text{offset} \in [0,15])$$

$$\text{for} \quad i = [0,15]$$

$$X_i = \text{MaskedSubBytes}_i(X_i)$$ [6]

每一条能量迹的掩码偏移量offset都是从$[0,15]$中随机选取的，改变偏移量的值会使能量迹中部分点发生相应的变化，因此可将offset的破解看作是一个分类问题。

在训练模型之前同样需要对能量迹进行特征提取，具体方法如2.1节所述。

用提取特征后的数据集$\text{dataset}_{i,j}(i \in [1,N], j \in [1,m])$和offset值训练SVM模型，使用训练后生成的分类模型对offset值进行预测[7-8]，在DPA_Contest_V4数据集下，当m取100，N取1 000时成功率即可达到99.4%。

2.3 基于线性加权的集成学习WAEL

为提高各个单模型权重的准确度，本文的模型权重由基于最小二乘法的线性回归模型计算出，具体步骤如下。

按照2.1节的方法提取出的特征点作为特征，中间值字节的汉明重量作为标签分别训练一个SVM模型和一个RF模型，并用这两个模型分别对测试集进行预测，生成两个单模型的预测结果$\text{result}_{\text{svm}}$和$\text{result}_{\text{rf}}$ [7]。为防止过拟合，需要对训练集进行交叉预测，得到两组与训练集等长的结果$\text{cross_result}_{\text{svm}}$和$\text{cross_result}_{\text{rf}}$。接下来$\text{cross_result}_{\text{svm}}$和$\text{cross_result}_{\text{rf}}$作为特征，中间值字节的汉明重量作为标签训练线性回归模型；最后用$\text{result}_{\text{svm}}$和$\text{result}_{rf}$作为测试集，用训练好的线性回归模型进行预测，预测结果即为SVM与RF模型集成的结果，基本步骤如图6所示。本文还进行了将3个模型线性加权的实验，步骤与两个模型的类似。

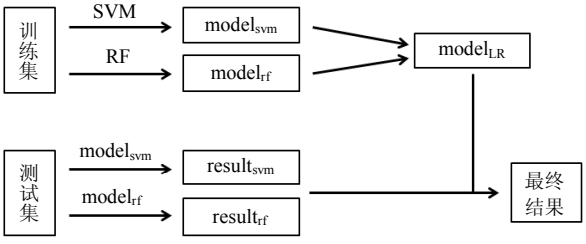

图6 基于线性加权的集成学习

2.4 基于投票法的集成学习VEL

本节使用和2.3节中相同的特征和标签分别训练一个SVM模型、一个RF模型和一个kNN模型，并用这3个模型分别对测试集进行预测，生成3个单模型的预测结果$\text{result}_{\text{svm}}$、$\text{result}_{\text{rf}}$和$\text{result}_{\text{kNN}}$，对这3个预测结果进行投票，最后将投票结果作为集成的结果输出，基本步骤如图7所示。

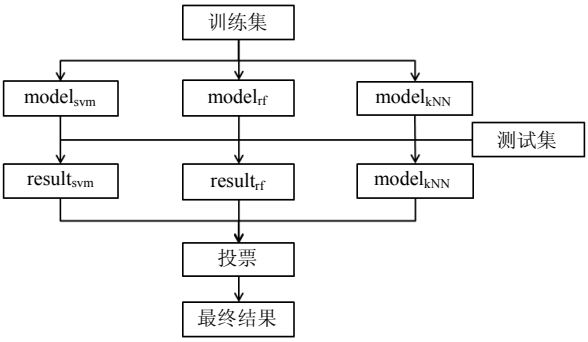

图7 基于投票法的集成学习

3 实验结果与分析

DPA_Contest_V4数据集是由法国电信提供的公开数据集，包含100 000条能量迹，攻击的对象是一个实现了带有循环掩码的AES-256的ATMega 163单片机。其中所有的能量迹采用的密钥、明文和偏移量offset都是已知的，每条能量迹有435 002个采样点，包含AES-256加密算法的第一轮和第二轮开始部分[6]。

本实验中取相关系数最大的100个点作为特征，掩码S盒输出字节的汉明重量作为标签，分别选取50，100，200，500，800，1 000，2 000，3 000，4 000，5000条能量迹作为训练集，另选10 000条能量迹作为测试集。SVM和RF的线性加权集成(WAEL)的预测成功率记录如表1和图8所示。

表1 SVM，RF与集成算法的成功率对比

能量迹数/条	模型		
	SVM	RF	WAEL
50	0.675	0.669	0.755
100	0.736	0.747	0.826
200	0.815	0.832	0.873
500	0.847	0.853	0.882
800	0.862	0.874	0.907
1 000	0.871	0.882	0.913
2 000	0.879	0.893	0.918
3 000	0.887	0.902	0.927
4 000	0.892	0.910	0.933
5 000	0.904	0.923	0.941

从表1和图8中可以看出，随着训练集能量迹数目的增加，预测的成功率在逐渐提高，同时集成模型的成功率始终比单个模型的成功率高。另外需要特别注意的是当训练集的能量迹数目越少时，集成后准确率提升得就越显著。

表2和图9记录了SVM、RF和kNN的投票集成(VEL)和线性加权集成(WAEL)的结果，训练集与测试集的选取方法同上，由于样本是随机选取的，因此结果与表1略有不同。

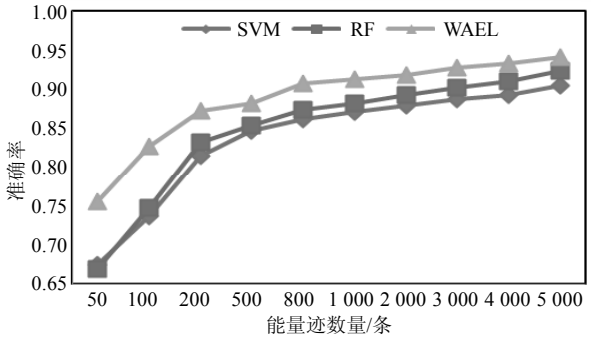

图8　SVM、RF与集成算法的成功率对比，数据源自表1

表2　SVM，RF，kNN与集成算法的成功率对比

能量迹数/条	模型				
	SVM	RF	kNN	VEL	WAEL
50	0.698	0.680	0.654	0.754	0.804
100	0.739	0.753	0.748	0.821	0.855
200	0.802	0.824	0.792	0.866	0.883
500	0.842	0.858	0.824	0.887	0.908
800	0.851	0.867	0.829	0.904	0.916
1 000	0.878	0.875	0.851	0.910	0.924
2 000	0.882	0.893	0.871	0.918	0.930
3 000	0.891	0.908	0.879	0.922	0.937
4 000	0.901	0.914	0.893	0.931	0.943
5 000	0.903	0.923	0.898	0.936	0.952

由表2和图9可知，随着训练集能量迹数目的增加，所有模型的预测成功率都在逐渐提高，SVM、RF和kNN这3个模型的成功率基本相当，而三者集成后的模型的正确率则显著高于单个模型。同时，和表1一样，当训练集中能量迹数目较少时，集成模型的提升更加显著。

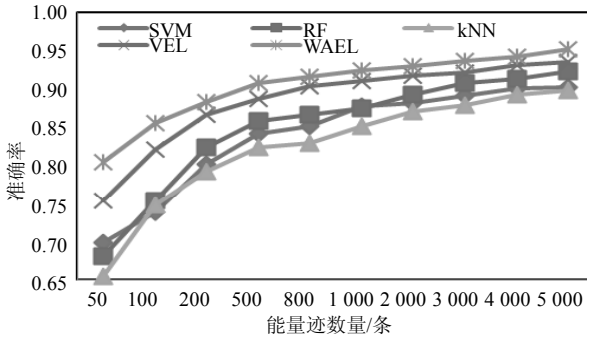

图9　SVM，RF，kNN与两种集成算法的成功率对比，数据源自表2

对比表2中的两种的集成方法，发现线性加权集成(WAEL)的成功率始终高于投票集成(VEL)。对比表1和表2中的WAEL的结果，发现在增加了kNN模型之后，线性加权集成(WAEL)的结果有了进一步的提升。

本文还对特征数不同的情况进行了实验。取1 000条能量迹，分别选取相关系数最大的50，100，150，200，250，300，350，400，450，500，550，600个点作为特征训练模型，各模型的准确率如图10所示。从图中可以看出，在特征数较少时集成模型的提升更为明显。

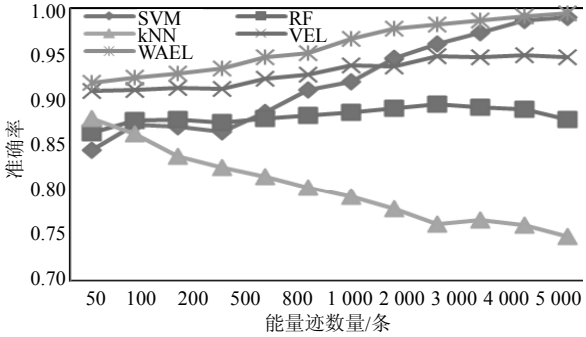

图10　SVM，RF，kNN与两种集成算法在不同特征数时的成功率对比

4　结 束 语

本文首次将多个分类模型的集成应用在了功耗分析领域，从实验中可以看出，无论是投票法(VEL)还是线性加权法(WAEL)都可以获得比单模型更高的成功率，尤其在训练样本数目较少的情况下线性加权法(WAEL)能够使模型的成功率获得超过10%的提升。

本文的实验表明，无论是线性加权集成还是投票集成，都可以在训练样本数量比较少(小于500条)的情况下使准确率有显著的提升，同时线性加权集成优于投票集成，三模型集成优于两模型集成。在训练样本较少的情况下，半监督学习也是一种有效的应对方法，因此在下一步的工作中，将探讨将半监督学习应用到功耗分析领域的具体方法，以期能在有标签能量迹较少的情况下进一步提高模型的准确率。

参 考 文 献

[1] HOSPODAR G, GIERLICHS B, MULDER E D, et al. Machine learning in side-channel analysis: a first study[J]. Journal of Cryptographic Engineering, 2011, 1(4): 293-302.

[2] MARKOWITCH O, LERMAN L, BONTEMPI G. Side channel attack: an approach based on machine learning[C]// International Workshop on Constructive Side-Channel Analysis and Security Design. [S.l.]: Springer-verlag, 2011: 29-41.

[3] LERMAN L, POUSSIER R, BONTEMPI G, et al. Template

attacks vs. machine learning revisited (and the curse of dimensionality in side-channel analysis)[C]//Constructive Side-Channel Analysis and Secure Design, International Workshop, Cosade 2015. Berlin, Germany: [s.n.], 2015: 20-33.

[4] 周志华. 机器学习[M]. 北京: 清华大学出版社，2016.
ZHOU Zhi-hua. Machine learning[M]. Beijing: Tsinghua University Press, 2016.

[5] BARTKEWITZ T, LEMKE-RUST K. Efficient template attacks based on probabilistic multi-class support vector machines[C]//International Conference on Smart Card Research and Advanced Applications. [S.l.]: Springer-Verlag, 2012: 263-276.

[6] TELECOM ParisTech SEN research group. DPA Contest (4th edition) [EB/OL]. [2017-5-12]. http://www. dpacontest. org/V4/.

[7] LERMAN L, MEDEIROS S F, BONTEMPI G, et al. A machine learning approach against a masked AES[M]// Smart Card Research and Advanced Applications. [S.l.]: Springer, 2013: 61-75.

[8] ZENG Z, GU D, LIU J, et al. An Improved side-channel attack based on support vector machine[C]//10th International Conference on Computational Intelligence and Security. [S.l.]: IEEE, 2014: 676-680.

[9] 邓高明, 张鹏, 赵强, 等. 基于PCA和SVM的电磁模板分析攻击[J]. 计算机测量与控制, 2009, 17(9): 1837-1839.
DENG Gao-ming, ZHANG Peng, ZHAO Qiang, et al. Electromagnetic template analysis with PCA and SVM[J]. Computer Measurement & Control, 2009, 17(9): 1837-1839.

[10] RECHBERGER C, OSWALD E. Practical template attacks[J]. Lecture Notes in Computer Science, 2005, 3325: 440-456.

编 辑 税 红