计算机科学
COMPUTER SCIENCE

基于卷积神经网络的旁路密码分析综述

刘林云，陈开颜，李雄伟，张阳，谢方方

引用本文

刘林云，陈开颜，李雄伟，张阳，谢方方. 基于卷积神经网络的旁路密码分析综述[J]. 计算机科学, 2022, 49(5): 296-302.

LIU Lin-yun, CHEN Kai-yan, LI Xiong-wei, ZHANG Yang, XIE Fang-fang. Overview of Side Channel Analysis Based on Convolutional Neural Network[J]. Computer Science, 2022, 49(5): 296-302.

计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.210300286

基于卷积神经网络的旁路密码分析综述

刘林云　陈开颜　李雄伟　张　阳　谢方方
陆军工程大学石家庄校区装备模拟训练中心　石家庄 050003
(llyun324@163.com)

摘　要　旁路建模分析方法可以有效攻击密码实现,其中基于卷积神经网络的旁路密码分析方法(CNNSCA)可以高效地进行密码攻击,甚至能够攻击有防护的加密算法设备。针对现阶段旁路密码分析建模方法的研究现状,对比分析了几种CNNSCA的模型特点和性能差异,并针对典型CNN模型结构以及旁路信号公共数据集ASCAD,通过模型对比及实验结果分析不同的CNN网络建模方法的效果,进而分析影响CNNSCA方法的性能因素、基于卷积神经网络的旁路建模方法的优势。由分析可知,基于VGG变体的CNNSCA在攻击各种情况的目标数据集时泛化性、鲁棒性表现最好,但使用的CNN模型训练程度及超参数设置是否最适用于SCA场景并未得到验证。今后研究者可通过调整CNN模型的各种超参数,使用数据增强技术,结合Imagenet大赛中优秀CNN网络等手段,来提升CNNSCA的分类准确率和破密性能,探索最适用于SCA场景的CNN模型是未来的发展趋势。

关键词:旁路分析;建模方法;卷积神经网络;超参数;性能评估
中图法分类号　TP309.7

Overview of Side Channel Analysis Based on Convolutional Neural Network

LIU Lin-yun,CHEN Kai-yan,LI Xiong-wei,ZHANG Yang and XIE Fang-fang
Center of Equipment Simulation Training,Shijiazhuang Campus of the Army Engineering University,Shijiazhuang 050003,China

Abstract　The profiled side-channel analysis method can effectively attack the implementation of cryptographic,and the side-channel cryptanalysis method based on convolutional neural network (CNNSCA) can efficiently carry out cryptographic attacks,and even can attack the implementation of protected encryption algorithms. In view of the current research status of side-channel cryptanalysis profiling methods,this paper compares and analyzes the characteristics and performance differences of several CNNSCA models,and focuses on the typical CNN model structure and side-channel signal public data set ASCAD. Through model comparison and experimental results,it compares and analyzes the effects of different CNN network modeling methods,and then analyzes the performance factors that affect the CNNSCA method and the advantages of the side-channel profiling method based on convolutional neural networks. Research and analysis show that CNNSCA based on VGG variants performs best in generalization and robustness when attacking target data sets in various situations,but whether the training level of the used CNN model and the hyperparameter settings are most suitable for SCA scenarios have not been verified. In the future,researchers can improve the classification accuracy and decryption performance of CNNSCA by adjusting various hyperparameters of the CNN model,use data enhancement techniques and combine the excellent CNN network in the Imagenet competition to explore the most suitable CNN model for SCA scenarios,which is a development trend.

Keywords　Side-channel analysis,Profiling method,Convolutional neural network,Hyperparameter,Performance evaluation

1　引言

旁路分析(Side Channel Analysis,SCA)[1]指绕过对加密算法的烦琐分析,利用密码算法的硬件实现在运算中泄露的信息,如执行时间、功耗、电磁辐射等,并结合统计理论快速地破解密码系统。这类新发现的物理泄漏信息被学者称为旁路信息(Side-channel Information),与之对应的攻击方法被称为旁路攻击(Side-channel Attack)。

旁路密码分析方法(SCA)分为建模方法和非建模方法。建模类方法包含模板攻击[5](Template Attack,TA)、基于多层感知器的旁路密码攻击(Side-channel Attack Based on Multi-layer Perceptron,MLPSCA,)以及基于卷积神经网络的旁路密码攻击(Side-channel Attack Based on Convolutional Neural Network,CNNSCA)。非建模类方法包括差分能量

到稿日期:2021-03-29　返修日期:2021-07-21
基金项目:国家自然科学基金(51377170,61602505)
This work was supported by the National Natural Science Foundation of China(51377170,61602505).
通信作者:陈开颜(chen_wu2013@163.com)

攻击[2]（Differential Power Attack，DPA）、相关系数攻击[3]（Correlation Power Attack，CPA）以及互信息攻击[4]（Mutual Information Attack，MIA）。虽然非建模类方法攻击方式简单直接，但旁路信号微弱或环境噪声过大会造成攻击失效，而建模类方法能有效分析旁路信号特征，预先获得攻击设备的加密知识，因此更易破获秘钥。

早期，从信息论角度看，如果有大量的旁路泄露信号（以下简称能量迹），那么传统建模方法中破密效果最好的是TA[5-9]。但随后研究人员发现，TA处理高维度旁路信号时存在统计困难，也无法攻击带防护加密实现。中期，随着计算机硬件性能的提升和人工智能领域机器学习的兴起，受监督机器学习算法在其他领域中能有效分析类似能量迹的一维数据，一些研究人员开始提出基于机器学习的旁路密码分析（Side-channel Analysis Based on Machine Learning，MLSCA）[10-15]。这些方法主要针对带加密算法AES（Advanced Encryption Standard）的设备，自此传统建模方法向结合机器学习算法的新型建模方法转变。新型建模方法MLP-SCA在攻击性能上超越了传统建模方法[8,16]，弥补了模板攻击不能处理高维度旁路信号的缺陷，但在攻击带防护的加密实现时也会失去效力。如今，随着机器学习的发展，在图像分类、目标识别上表现优异的深度学习技术[17]开始盛行，已有研究表明，深度学习下的卷积神经网络算法应用在旁路分析上具有较好的破密性能[18-22]，而且CNNSCA能有效攻击带防护的加密实现。虽然CNNSCA弥补了以往建模方法的缺陷，破密性能有了提升，但并没有提出最适用于旁路密码分析的卷积网络模型，而且模型本身的训练准确率都不高，因此CNNSCA的模型学习能力和破密性能并未达到最优，仍有提升空间。

目前，在旁路分析领域，不同CNNSCA方法的汇总分析较少，或者不够完善，本文分析了现有几种CNNSCA方法的破密性能的差异和特点，并通过实验论证了CNNSCA在攻击新目标与原文献目标时性能存在差异，但调整CNNSCA模型超参数并重新训练模型可以使其攻击新的目标。本文还论证了CNNSCA在攻击添加了高斯噪声的无防护能量迹时其破密性能更好。实验还发现，基于Alexnet的CNNSCA在攻击带防护的加密设备时不失为一种新的有效模型。

2 基于卷积神经网络的旁路密码分析

2.1 旁路泄露公开数据集简介

最新公布的ASCAD数据库[19]采集的目标是带一阶掩码防护的AES-128实现，即8位AVR微控制器（ATmega8515），其中能量迹是由采集的电磁辐射转换的数据信号。敌手针对AES第一轮加密的第三个S盒输出采集信号，并针对第一个AES密钥字节发起攻击，该数据库遵循MNIST数据库规则，共提供4个数据集，每个数据集包含60000条能量迹，其中50000条能量迹用于分析/训练，余下的10000条能量迹用于测试/攻击。前3个ASCAD数据集分别代表设置3种不同随机时延防护对策的加密实现泄露，分别用信号偏移$desync=0$，$desync=50$，$desync=100$来表示带掩码和时延这两种策略的数据集。在前3类数据集中，所有能量迹都包含700个特征点，这些特征点是在包含100000个特征点的原始能量迹中选取的，选取依据是信号尖峰最大的位置。在掩码已知的情况下，数据集的信噪比最大值可达到0.8，而在掩码未知的情况下几乎为0。最后一个ASCAD数据集存放的是原始能量迹。

2.2 CNN

卷积神经网络（Convolutional Neural Network，CNN）是人工智能最成功的算法之一，是一种新结构的多层神经网络[23]。CNN的设计受到了视神经感受野研究[24-25]的启发，其核心部件卷积核就是局部感受野的结构体现。它属于反向传播训练的深度网络，利用数据二维空间关系减少需要学习的参数数目，在一定程度上提升了BP算法的训练性能。CNN与MLP（Multilayer Perceptron）的主要区别在于其增加了卷积块结构。在卷积块中，输入数据的一小部分作为网络结构的原始输入，数据信息在网络中逐层向前传递，每层通过若干卷积核对输入的数据进行特征提取。卷积神经网络已在计算机视觉、自然语言处理等领域成功应用[26-27]，尤其在计算机视觉领域的Imagenet大规模视觉识别挑战赛（ILSVRC）[28]上大放异彩。

2.3 核心算法及网络结构

2.3.1 网络结构

结合旁路密码分析场景，应用到旁路攻击的CNN主要有6种逐层堆叠的网络层。

（1）卷积层（Convolutional Layers，CONV）。该层属于线性层，层与层之间不完全连接，能避免全连接网络的两个缺陷：训练权值需要巨大的计算量和模型过拟合。同一层中的同一个卷积核（Filters）权值共享，可让卷积层提取不变位移特征，同时减少参数。卷积层也可以使用多个卷积核，每个卷积核从输入向量中提取不同的抽象特征，这些抽象特征在附加维度（所谓的深度）上并排排列，使得CNN能够抵抗时域变形的向量特征[29-30]。卷积层通常需要设置填充（Padding）方式，一种是不填充（Valid Padding），使得卷积后的特征向量维度小于原始向量；另一种是相同填充（Same Padding），使得卷积后的特征向量维度与原始向量相同。

（2）批量归一化层[31]（Batch Normalization Layers，BN）。该层的作用是减少协变量在训练和预测两个阶段发生偏移，有利于网络模型使用更高的学习率[32]。

（3）激活层（Activation Layers，ACT）。该层属于非线性层，由单个实函数组成，该函数作用于输入向量的每个坐标。目前在深度学习中首选ReLU函数。

（4）池化层（Pooling Layers，POOL）。该层是非线性层，使用池化窗口在输入向量上滑动，提取显著特征点，以减小特征维度。池化层不存在权值，不会造成输入信号的变形。

（5）全连接层（Fully-Connected Layers，简称FC）。其层与层之间的神经元完全连接，这些层需要训练大量权值。该层用仿射函数表示为：D维$\boldsymbol{x}$向量为输入，$\boldsymbol{Ax}+\boldsymbol{B}$为输出。其中$\boldsymbol{A}\in R_{C\times D}$是权值矩阵、$\boldsymbol{B}\in R_C$是偏差向量，这些权值和偏差是FC层的训练参数。

（6）softmax层（Softmax Layer，SOFT）。在多分类任务中通常使用softmax作为输出层的激活函数，这里用softmax

代表输出层。该层对输入进行分类，得到各个标签的预测值，取最大值对应的标签作为全局分类结果。

CNN 的卷积块由 CONV 层、BN 层、ACT 层构成，在该块之后通常会添加一个 POOL 层以减少特征维度，组成的新卷积块在网络模型中重复 n 次，直至获得合理大小的输出。然后，引入 n 个 FC 层，在最后一个 FC 层使用 softmax 函数，最后输出分类预测结果。旁路密码攻击的卷积网络结构如图 1 所示。

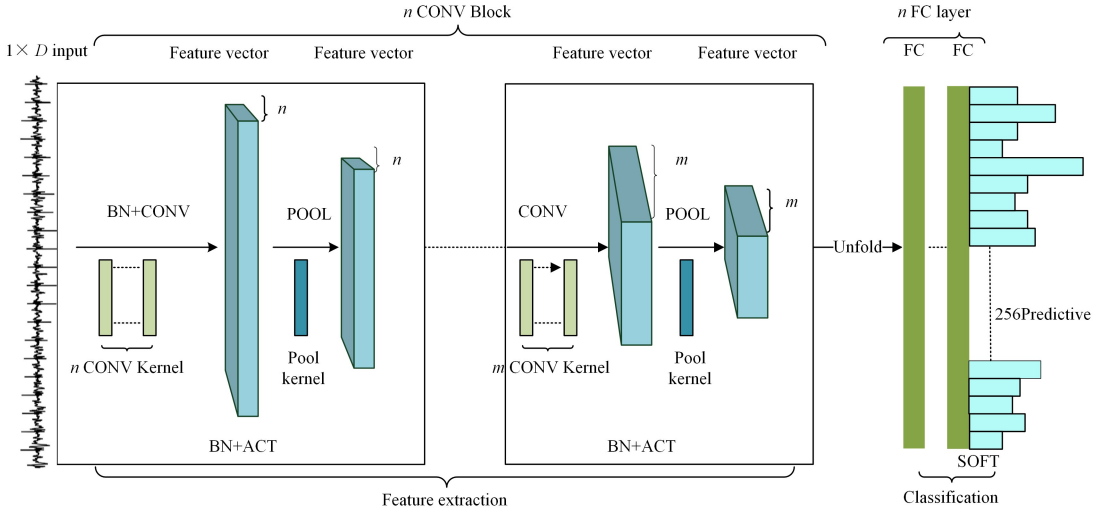


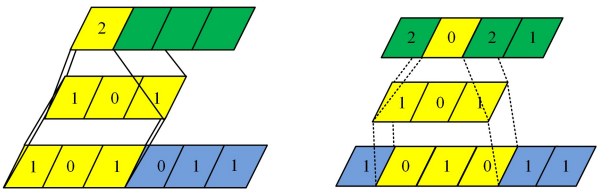
图 1 旁路攻击场景下的卷积网络结构

Fig. 1 Convolutional network structure in side-channel attack scenarios

2.3.2 核心算法

(1)卷积计算

将卷积核在一维向量上滑动，每次移动的步数称为步长，每次滑动时进行卷积计算得到一个数值，一轮计算完成后得到一个表示向量特征的特征向量。数值运算规则是将一个一维卷积核与一个一维向量对应位置的数值相乘，然后再求和。例如，有一个 1×3 的卷积核，对 1×6 的一维向量求卷积，步长为 1，计算过程如图 2 所示。



(a)卷积核对应向量窗口滑动第一次运算　(b)卷积核对应向量窗口滑动第二次运算

图 2 卷积计算过程

Fig. 2 Convolution calculation process

图 2(a)中，卷积核从输入向量左侧开始滑动，第一步数值运算为 1×1+0×0+1×1=2，得到新特征向量的第一个数值 2；接着，卷积核向右滑动一步，继续数值运算 1×0+0×1+1×0=0，得到新特征向量的第二个数值 0，如图 2(b)所示。重复此过程，直至卷积核滑动到输入向量的最右边，卷积计算完成。

(2)池化计算

池化有最大池化(Max-pooling)、平均池化(Mean-pooling)和随机池化(Stochastic Pooling)3 种方式。最大池化是提取池化窗口内数值的最大值，平均池化是提取池化窗口内数值的平均值，随机池化则是随机提取池化窗口内的数值。

(3)softmax 函数

该函数对输出值进行归一化操作，把所有输出值都转化为概率，概率之和为 1，softmax 的计算式为：

$$softmax(x_i)=\frac{\exp(x_i)}{\Sigma_j \exp(x_j)} \tag{1}$$

其中，x_i 是 softmax 层第 i 个神经元的输入，x_j 是 softmax 层所有神经元的输入，Σ_j 是对 x_j 的计算求和。函数结果作为第 i 个神经元标签的拟合概率。

(4)权值调整原理

使用代价函数与梯度下降算法[33]，网络模型每训练一次，权值往误差减小的方向自动调整一次，这样重复训练调参，直到所有迭代结束，权值调整完成。

(5)破秘性能评估

通常安全员在评估 CNNSCA 破密性能时会考虑两个方面的指标：1)建模时神经网络模型的训练准确率 Acc[34]；2)攻击阶段获取密钥的安全指标猜测熵[35-36]。猜测熵用于衡量密钥破解效率，每次攻击通过一个秩函数输出来猜测密钥的排名位置。在 n 次攻击中，破密方法性能越好，效率越高，排名就越快收敛于首位零。

3 现有 CNNSCA 的模型特点及实验分析

3.1 现有 CNNSCA 的网络模型特点

在文献[37]中，CNN 已被证明是时序数据的强大分类器(如音频)，而旁路信号与一维时序信号类似，且共享大量特征，这使得旁路分析使用 CNN 算法受到了很大启发，文献[8,16]鼓励研究人员在旁路分析中使用 CNN。文献[16]发现，基于深度神经网络的旁路分析在特征提取过程中使用 CNN 算法比以往的经典机器学习算法 AE(自动编码器)效率更高。近年来，已有一些将 CNN 算法成功应用到旁路分析的文献。

Maghrebi 等[16]首次将 CNN 用于旁路密码分析，并分别针对无防护、带一阶掩码防护的 AES 加密设备泄露能量迹实施攻击。他们采取猜测熵指标来衡量破秘性能，将 CNNSCA 与 MLPSCA 攻击进行对比，结果显示，CNNSCA 对上述两类

能量迹均可实现破密，且攻击性能优于 MLPSCA。

Cagli 等[38]针对带抖动防护对策的 AES 加密设备进行 CNNSCA 攻击，他们使用 CNN 神经网络的数据增强技术[39]来消除这种防护对策对旁路攻击带来的影响。该 CNNSCA 方法的 CNN 有 10 层网络结构，选取汉明重量泄露模型。实验结果证明，CNNSCA 的鲁棒性较好，能攻击带抖动防护策略的 AES 加密实现。

Benadjila 等[19]创建了 ASCAD 公共数据集，本文 2.1 节对此进行了详细阐述。他们针对带一阶掩码防护或叠加抖动防护策略的 AES 加密实现进行 CNNSCA 攻击，并分别利用 ImageNet 竞赛中表现突出的 VGG16[40]，Inception-v3[41]以及 ResNet-50[42]3 种 CNN 网络对 ASCAD 能量迹进行旁路攻击。通过实验发现，基于 VGG16 的 CNNSCA 的破秘性能明显优于另外两种网络。在此基础上，Benadjila 等对 VGG16 的原结构进行了改造，以提升该 CNNSCA 的破密性能，改造后的网络称为 VGG16 变体。接下来用 VGG16 变体进行旁路分析实验，结果发现，该 CNNSCA 对 ASCAD 数据集的攻击效果提升较大，且 CNNSCA 能攻击带防护的 AES 加密实现。

Masure 等[43]提出了使用敏感度分析来选择能量迹的特征点。他们分别用 CNN 梯度可视化技术和传统信噪比技术，将选择的特征点进行模板攻击。结果表明，当使用基于 CNN 的梯度可视化技术选择特征点时，模板攻击破密效果更好，这说明使用 CNN 进行特征提取更具优越性。

Carbone 等[44]针对 RSA 实现采集了不同类型的泄露数据集，利用 CNN 算法攻击这些数据集时，显示了 CNN 高潜力的破密效率，但其设计的 CNNSCA 高度依赖原目标加密设备和测量活动，泛化性极低。

Kim 等[45]提出了在能量迹添加人工噪声的方法，并研究了 CNNSCA 攻击带噪声能量迹时的性能。他们同样以 VGG16 基网络搭建了 11 层 VGG16 变体，对 ASCAD 数据集添加人工噪声后再实施攻击。实验结果显示，在能量迹中添加适量噪声不但不会影响 CNN 的分类效果，反而提升了 CNN 的分类精度，密钥攻击效果同样也得到了提升。

Chen 等[46]提出了一种优化的 CNNSCA，该模型使用了一种新的卷积层 SincNet，SincNet 层只需要训练卷积核的高低两个核参数，相比传统的卷积层，它减少了需要学习的参数量，但没有影响模型的破密效果。

Perin 等[47]解决了如何确定正确的训练迭代次数这个问题，以阻止 DLSCA 模型过度训练或训练不足，并证明了可以测量传送到输出层的信息量——互信息，将其作为参考度量，以确定网络模型的最佳泛化迭代次数。

Guo 等[48-49]通过调整超参数来改变神经网络结构，进而将 Alexnet 与 VGGNet 卷积神经网络应用到旁路分析中，并结合建模类旁路分析原理，成功实现了破密，但训练耗时多且训练参数量较大。现有 CNNSCA 方法中，收集到的 CNN 网络及其参数总结如表 1 所列。

表 1 现有 CNNSCA 方法中的 CNN 网络及其参数

Table 1 CNN network and parameters in existing CNNSCA method

Author	CNN	block	Convolutionallayer	Kernelsize	BN+Activation function	Pooling layer	dense layer	Learning rate	Batch size	epoch
Maghrebi et al	LeNet5	—	—	—	—	—	—	—	—	—
Cagli et al	—	4	4 layer	—	ReLU, softmax	max	1 layer	—	—	120
Benadjila et al	VGG16 Variants	5	5 layer {64,128,256, 512,512}, same	All layer 1×11	ReLU, softmax	All layer Mean, valid(2,2)	3layer {4096, 4096,256}	10^{-5}	200	75
Kim et al	VGG16 Variants	6	8layer {8,16,32,(64,64), 128,(256,256)}, valid	All layer 1×3	ReLU, softmax	All layer max, valid(2,2)	2 layer {256,256}	10^{-4}	256	75
Guo et al	Alexnet Variants	4	5 layer {96,256,(384, 384),256}, same	Respectively are 1×11,1×5,1×3, 1×3,1×3	ReLU, softmax	1,2,5 layer max, valid(2,2),(2,2), (3,3)	3 layer {4096, 4096,256}	10^{-2}	10	20
Guo et al	VGGNet Variants	5	13 layer {(64,64),(128, 128),(256,256,256), (512,512,512), (512,512,512) }, same	All layer 1×2	ReLU, softmax	2,4,7,10,13 layer respectively are max(2,2)	3 layer{4096, 4096,256}	10^{-6}	1800 (150)	5000 (75)

3.2 现有 CNNSCA 实验对比分析

本文实验涉及的算法均使用 Python 语言编程，并使用深度学习架构 Keras 库[50]（版本 2.4.3）或直接使用 GPU 版 Tensorflow 库[51]（版本 2.2.0）。实验在配备 16GB RAM 和 8GB GPU(Nvidia GF RTX 2060)的普通计算机上进行。

以下 5 个实验均使用本文 2.1 节中的 ASCAD 公共数据集，并使用本文 2.3.2 节中的猜测熵评估方法进行破密性能评估。实验主要复现文献[19,45,48-49]中破密性能较好的 4 个 CNN 结构，分别是国外 Benadjila 等和 Kim 等的两个 CNN 结构，国内 Guo 等的两个 CNN 结构，详细参数如表 1 所列。由于实验用计算机显卡和 CPU 的限制，将 Guo 等基于 VGG 的 CNNSCA 下 CNN 训练参数迭代次数、批次更改为 75 和 150。

3.2.1 使用原模型参数

实验 1 使用 Benadjila 等、Kim 等和 Guo 等的 4 类 CNNSCA 方法，几乎保持原 CNN 所有参数不变，都攻击已知掩码的 ASCAD 数据集，该数据集代表无防护且高信噪比的旁路信号。其中 Benadjila 等、Kim 等的原文献也是攻击 ASCAD 数据集。而 Guo 等的原文献是攻击自采数据集，基于 Alex 的 CNNSCA 破密成功率为 0.611[48]，基于 VGG 的 CNNSCA 破密成功率为 0.923[49]，实验结果如图 3 所示。由图 3 中可以看出，Benadjila 等和 Kim 等应用基于 VGG 变体的 CNNSCA 猜测熵收敛最快。Guo 等的两种方法中，基于 VGG 变体的 CNNSCA 在攻击与原文献不同的目标——ASCAD 时猜测熵收敛较慢，而基于 Alex 变体的 CNNSCA 却没有收敛。这些 VGG 变体使用不同的参数，代表不同的 CNN

网络。这说明基于 VGG 的 CNNSCA 泛化性较好。

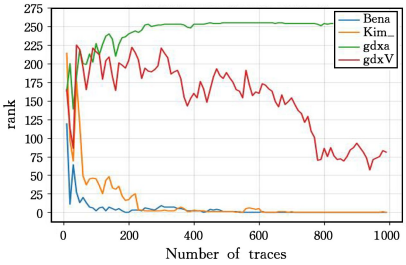


图 3 Benadjila 等、Kim 等和 Guo 等的 4 类 CNNSCA 猜测熵(原文献参数)(电子版为彩图)

Fig. 3 Four types of CNNSCA guess entropy of Benadjila, Kim and Guo(original document parameters)

3.2.2 改变超参数

实验 2 在实验 1 的基础上将 Guo 等基于 Alex 的 CNNSCA 训练迭代次数由 20 改为 75,结果如图 4 所示。

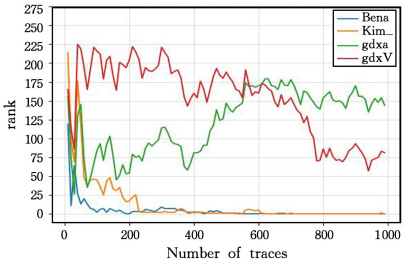


图 4 Benadjila 等、Kim 等、Guo 等的 4 类 CNNSCA 猜测熵(迭代次数都为 75)(电子版为彩图)

Fig. 4 Four types of CNNSCA guess entropy of Benadjila, Kim and Guo(all epochs are 75)

从图 4 可知,该基于 Alex 的 CNNSCA 猜测熵相比图 3 有所收敛,说明通过调整超参数迭代次数,CNNSCA 方法可实现对新目标的攻击。

3.2.3 在能量迹添加高斯噪声

实验 3 在实验 2 的基础上为上述实验数据集添加高斯噪声。在一维旁路信号中添加高斯噪声的措施,类似于二维图像数据在进行模型训练之前的数据增强手段[17]。这里添加高斯噪声的方法是:在实验 2 使用的原数据集的基础上任选三分之一的能量迹,然后在这三分之一能量迹上逐条添加高斯分量,用于模拟采集信号时的高斯噪声[1],实验结果如图 5 所示。

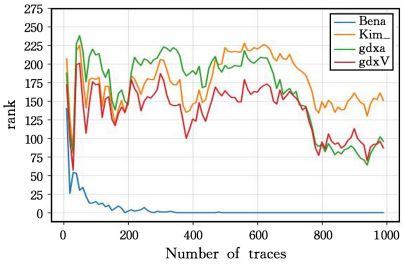


图 5 Benadjila 等、Kim 等、Guo 等的 4 类 CNNSCA 猜测熵(迭代数都为 75,高斯噪声)(电子版为彩图)

Fig. 5 Four types of CNNSCA guess entropy of Benadjila, Kimand Guo (all epochs are 75, Gaussian noise)

由图 5 的对比显示可知,其中 Kim 等的 CNNSCA 猜测熵结果最差,Guo 等的 CNNSCA 猜测熵进一步收敛,Benadjila 等的 CNNSCA 猜测熵收敛速度翻了一倍,在未加高斯噪声前其猜测熵排名为 650 左右(见图 3),添加之后为 270 左右(见图 5)。这说明对旁路信号使用数据增强手段也会提升原 CNNSCA 模型的性能。

3.2.4 攻击带防护的能量迹

实验 4 在实验 2 的基础上,改换信号偏移 desync=100 的 ASCAD 数据集攻击,代表攻击有时延防护对策的加密设备,实验结果如图 6 所示。

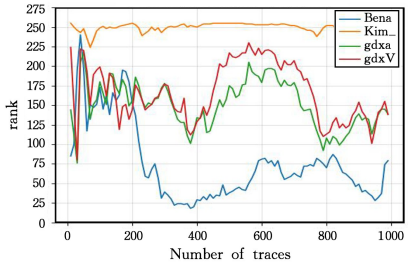


图 6 Benadjila 等、Kim 等、Guo 等的 4 类 CNNSCA 猜测熵(迭代数都为 75,时延 100)(电子版为彩图)

Fig. 6 Four types of CNNSCA guess entropy of Benadjila, Kim and Guo(all epochs are 75, desynchronization 100)

由图 6 可知,Kim 等的 CNNSCA 猜测熵没有收敛,破密完全失效。Benadjila 等、Guo 等的 CNNSCA 猜测熵依然有收敛的趋势,而且 Guo 等基于 Alex 变体的 CNNSCA 猜测熵收敛情况与实验 2 中攻击无防护数据集相比(见图 4),绿色曲线反而向下收敛。

3.2.5 在带防护的能量迹中添加高斯噪声

实验 5 在实验 4 的基础上,在信号偏移 desync=100 的 ASCAD 数据集上添加高斯噪声,加噪方法已在第 3.2.3 节中详细阐述,实验结果如图 7 所示。从图 7 可以看出,Kim 等的 CNNSCA 猜测熵与实验 4 相比往下收敛,Guo 等的两种 CNNSCA 猜测熵变化不大,Benadjila 等的 CNNSCA 猜测熵对比实验 4 收敛效果反而变差。这说明攻击带时延防护的数据集时,使用添加高斯噪声这种数据增强手段可改善 kim 等的 CNNSCA 性能,但对其他方法无效。

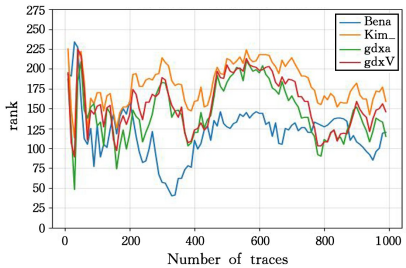


图 7 Benadjila 等、Kim 等、Guo 等的 4 类 CNNSCA 猜测熵(迭代数都为 75,时延 100,高斯噪声)(电子版为彩图)

Fig. 7 Four types of CNNSCA guess entropy of Benadjila, Kim and Guo (all epochs are 75, desynchronization 100, Gaussian noise)

本文分别在实验 1、实验 3 和实验 4 中记录了 Benadjila 等基于 VGG 的 CNNSCA 模型训练准确率(accuracy),结果如图 8 所示。显然这些准确率都不高,说明该 CNN 网络在

SCA 场景下的分类准确率仍有提升空间。

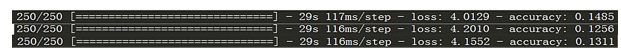


图 8 Benadjila 等基于 VGG 的 CNNSCA 网络在实验 1、实验 3 和实验 4 中的训练准确率

Fig. 8 Training accuracy of Benadjila's VGG-based CNNSCA network in experiments 1, 3 and 4

结束语 本文介绍了目前已经破密成功的几种 CNNSCA 方法，并复现了几类提供详细参数的 CNNSCA 模型，同时通过实验分析了这些新型建模方法的性能差异和特点。通过研究发现，CNNSCA 拥有良好的高维数据处理能力以及特征提取能力，其对旁路信号中的噪声、抖动、非对齐等信号具有强鲁棒性。其中，基于 VGG 变体的 CNNSCA 的破密性能，目前已超过其他 CNNSCA，但并未达到最优。本文通过实验证明，可以利用 3 种手段来提升 CNNSCA 的分类准确率和破密性能，即优化 CNN 模型的各种超参数、使用数据增强技术、使用 Imagenet 大赛中的其他优秀 CNN 算法。因此，继续探索最适用于 SCA 应用场景的深度学习 CNN 模型，设计或优化出性能强大的 CNNSCA 分析器都具有很大的研究空间。这对实现高效的旁路密码攻击有重要意义，并且对信息安全和加密防护也有一定的学术意义。

参 考 文 献

[1] MANGARD S, OSWALD E, POPP T. Energy analysis attack [M]. Beijing: Science Press, 2010.

[2] KOCHER P, JAFFE J, JUN B. Differential power analysis[C]// Annual International Cryptology Conference. Berlin: Springer, 1999: 388-397.

[3] BRIER E, CLAVIER C, OLIVIER F. Correlation power analysis with a leakage model[C]// International Workshop on Cryptographic Hardware and Embedded Systems. Berlin: Springer, 2004: 16-29.

[4] GIERLICHS B, BATINA L, TUYLS P, et al. Mutual information analysis[C]// International Workshop on Cryptographic Hardware and Embedded Systems. Berlin: Springer, 2008: 426-442.

[5] CHARI S, RAO J R, ROHATGI P. Template attacks[C]// International Workshop on Cryptographic Hardware and Embedded Systems. Berlin: Springer, 2002: 13-28.

[6] LERMAN L, BONTEMPI G, MARKOWITCH O. Power analysis attack: An approachbased on machine learning[J]. International Journal of Applied Cryptography: IJACT, 2014, 3(2): 97-115.

[7] PICEK S, HEUSER A, GUILLEY S. Template attack versus Bayes classifier[J]. Journal of Cryptographic Engineering, 2017, 7(4): 1-9.

[8] CAGLI E, DUMAS C, PROUFF E. Convolutional Neural Networks with Data Augmentation Against Jitter-Based Countermeasures Profiling Attacks Without Preprocessing[C]// Cryptographic Hardware and Embedded Systems CHES 2017 19th International Conference. Taipei, Taiwan, 2017: 45-68.

[9] CHOUDARY O, KUHN M G. Efficient template attacks[C]// International Conference on Smart Card Research and Advanced Applications. Springer, 2013: 253-270.

[10] 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-6th International Workshop, COSADE 2015. Berlin, Germany, 2015: 20-33.

[11] LERMAN L, BONTEMPI G, MARKOWITCH O. A machine learning approach against a masked AES-Reaching the limit of side-channel attacks with a learning model[J]. Jounal of Cryptographic Engineering, 2015, 5(2): 123-139.

[12] PICEK S, HEUSER A, JOVIC A, et al. Climbing down the hierarchy: Hierarchical classification for machine learning side-channel attacks[C]// 9th International Conference on Cryptology in Africa. Springer, 2017: 61-78.

[13] HEUSER A, ZOHNER M. Intelligent Machine Homicide Breaking Cryptographic Devices Using Support Vector Machines [C]// COSADE. Springer, 2012: 249-264.

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

[15] PICEK S, HEUSER A, JOVIC A, et al. Side-channel analysis and machine learning: A practical perspective[C]// 2017 International Joint Conference on Neural Networks, IJCNN 2017. Anchorage, AK, USA, 2017: 4095-4102.

[16] MAGHREBI H, PORTIGLIATTI T, PROUFF E. Breaking cryptographic implementations using deep learning techniques[C]// 6th International Conference on Security, Privacy, and Applied Cryptography Engineering (SPACE 2016). Hyderabad, India, 2016: 3-26.

[17] BENGIO Y, GOODFELLOW I, COURVILLE A. Deep learning [M]. MIT press, 2017: 170-200.

[18] PICEK S, SAMIOTIS, I P, HEUSER A, et al. On the performance of convolutional neural networks for side-channel analysis [OL]. https://eprint.iacr.org/2018/004.

[19] BENADJILA R, PROUFF E, STRULLU R, et al. Deep learning for side-channel analysis and introduction to ASCAD database [J]. Journal of Cryptographic Engineering, 2019, 10.

[20] HEUSER A, PICEK S, GUILLEY S, et al. Lightweight ciphers and their side-channel resilience[J]. IEEE Transactions on Computers, 2017, 69(10): 1434-1448.

[21] HUANG J, WANG Y. Experimental Research on Convolutional Neural Network Structure Suitable for Side Channel Analysis [J]. Journal of Chengdu University of Information Technology, 2019(5): 449-456.

[22] MAGHREBI H. Deep learning based side channel attacks in practice [J/OL]. IACR Cryptol. ePrint Arch., 2019: 578. https://eprint.iacr.org/2019/578.

[23] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 4700-4708.

[24] HUBEL D H, WIESEL T N. Receptive Fields And Functional Aechitecture of Monkey Striate Cortex[J]. The Journal of Phy-

siology, 1968, 195(1): 215-243.
[25] LECUN Y, BENGIO Y. Convolutional networks for images, speech, and time series[M]// The Handbook of Brain Theory and Neural Networks. MIT Press, 1998: 255-258.
[26] SHI H, YANG Q, LIU S H, et al. Research on information extraction of power grid failure plans based on deep learning[J]. Computer Science, 2020, 47(S2): 62-66.
[27] YIN W, KANN K, YU M, et al. Comparative Study of CNN and RNN for Natural Language Processing[J]. arXIV: 1702.01923, 2017.
[28] RUSSAKOVSKY O, DENG J, SU H, et al. Imagenet large scale visual recognition challenge[J]. International Journal of Computer Vision, 2015, 115(3): 211-252.
[29] GILMORE R, HANLEY N, O'NEILL M. Neural network based attack on a masked implementation of AES[C]// 2015 IEEE International Symposium on Hardware Oriented Security and Trust (HOST). 2015: 106-111.
[30] ZOTKIN Y, OLIVIER F, BOURBAO E. Deep Learning vs Template Attacks in front of fundamental targets: experimental study[J/OL]. IACR. https://xs.dailyheadlines.cc/scholar?q=Deep+Learning+vs+Template+Attacks+in+front+of+fundamental+targets%3A+experimental+study
[31] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift[J]. arXiv: 1502.03167, 2015.
[32] GOODFELLOW I J, BENGIO Y, COURVILLE A C. Deep Learning[M]//Adaptive Computation and Machine Learning. Cambridge: MIT Press, 2016.
[33] HAN L Q, KANG Q. Artificial Neural Network Theory, Design and Application——Nerve Cells, Neural Networks and Neural System[J]. Journal of Beijing Technology and Business University (Natural Science Edition), 2005, 23(1): 52-52.
[34] HAWKINS D M. The problem of overfitting [J]. Journal of Chemical Information and Computer Sciences, 2004, 44(1): 1-12.
[35] STANDAERT F X, MALKIN T G, YUNG M. A unified framework for the analysis of side-channel key recovery attacks[C]// Annual International Conference on The Theory and Applications of Cryptographic Techniques. Berlin: Springer, 2009: 443-461.
[36] MASURE L, DUMAS C, PROUFF E. A comprehensive study of deep learning for side-channel analysis[J]. IACR Transactions on Cryptographic Hardware and Embedded Systems, 2020 (1): 348-375.
[37] OORD A V D, DIELEMAN S, ZEN H, et al. Wavenet: A generative model for raw audio[J]. arXiv: 1609.03499, 2016.
[38] CAGLI E, DUMAS C, PROUFF E. Convolutional neural networks with data augmentation against jitter-based countermeasures[C]// International Conference on Cryptographic Hardware and Embedded Systems. Cham: Springer, 2017: 45-68.
[39] WONG S C, GATT A, STAMATESCU V, et al. Understanding data augmentation for classification: when to warp? [C]// International Conference on Digital Image Computing: Techniques and Applications (DICTA). IEEE, 2016: 1-6.
[40] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large scale image recognition[J]. arXiv: 1409.1556, 2014.
[41] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015: 1-9.
[42] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 770-778.
[43] MASURE L, DUMAS C, PROUFF E. Gradient visualization for general characterization in profiling attacks[C]// International Workshop on Constructive Side-Channel Analysis and Secure Design. Cham: Springer, 2019: 145-167.
[44] CARBONE M, CONIN V, CORNÉLIE M A, et al. Deep learning to evaluate secure RSA implementations[J]. IACR Transactions on Cryptographic Hardware and Embedded Systems, 2019 (2): 132-161.
[45] KIM J, PICEK S, HEUSER A, et al. Make some noise. unleashing the power of convolutional neural networks for profiled side-channel analysis[J]. IACR Transactions on Cryptographic Hardware and Embedded Systems, 2019(3): 148-179.
[46] CHEN P, WANG P, DONG G F, et al. Side channel attack based on SincNet[J/OL]. Journal of Cryptography. http://kns.cnki.net/kcms/detail/10.1195.TN.20200520.1652.002.html.
[47] PERIN G, BUHAN I, PICEK S. Learning when to stop: a mutual information approach to fight overfitting in profiled side-channel analysis [C]// International Workshop on Constructive Side-Channel Analysis and Secure Design. Springer, Cham, 2021: 53-81.
[48] GUO D X, CHEN K Y, ZHANG Y, et al. A new method for attacking encrypted chip templates based on Alexnet convolutional neural network[J]. Computer Measurement and Control, 2018, 26(10): 246-249, 254.
[49] GUO D X, CHEN K Y, ZHANG Y, et al. A new method of attacking encrypted chip templates based on VGGNet convolutional neural network [J]. Computer Application Research, 2019, 36(9): 2809-2812, 2855.
[50] GULLI A, PAL S. Deep learning with Keras[M]. Packt Publishing Ltd, 2017.
[51] ABADI M, AGARWAL A, BARHAM P, et al. Tensor Flow: Large-scale machine learning on heterogeneous systems[OL]. https://www.tensorflow.org/. Software available from tensorflow.org.

LIU Lin-yun, born in 1988, postgraduate. Her main research interests include side-channel attack and so on.

CHEN Kai-yan, born in 1970, Ph.D, associate professor. Her main research interests include cryptography and so on.

(责任编辑:柯颖)