侧信道多层感知器攻击中基于贝叶斯优化的超参数寻优

杨 欢[1] 吴 震[1*] 王 燚[1] 杜之波[1] 王 敏[1] 习 伟[2] 颜 伟[3]

[1](成都信息工程大学网络空间安全学院 四川 成都 610225)

[2](南方电网科学研究院有限公司 广东 广州 510080)

[3](振芯科技有限公司 四川 成都 640041)

摘 要 传统侧信道攻击利用加密设备泄露的物理信息来获取密钥,由于其需要大量人为干预,越来越多研究将机器学习算法运用到侧信道攻击中,其中神经网络攻击效果最好,而多层感知器又是神经网络的基础,其中超参数在很大程度上影响最终训练与攻击结果。为实现超参数自动寻优,将贝叶斯寻优的方法应用在侧信道多层感知器攻击中,并提出对离散值的处理方法,发展出能够结合超参数经验的侧信道多层感知器超参数寻优方法。实验对比了人工寻优与贝叶斯寻优两种算法用于侧信道多层感知器攻击中的效率,验证了多层感知器与侧信道攻击相结合及贝叶斯寻优的可行性和高效性。

关键词 侧信道攻击 多层感知器 贝叶斯寻优

中图分类号 TP3 **文献标志码** A DOI: 10.3969/j.issn.1000-386x.2021.05.052

HYPER – PARAMETERS OPTIMIZATION IN SIDE – CHANNEL ATTACK OF MULTILAYER PERCEPTRON BASED ON BYESIAN OPTIMIZATION

Yang Huan[1] Wu Zhen[1*] Wang Yi[1] Du Zhibo[1] Wang Min[1] Xi Wei[2] Yan Wei[3]

[1](*School of Cybersecurity, Chengdu University of Information Technology, Chengdu 610225, Sichuan, China*)

[2](*Research Institute of China Southern Power Grid, Guangzhou 510080, Guangdong, China*)

[3](*Vibration Core Technology Co., Ltd., Chengdu 640041, Sichuan, China*)

Abstract Traditional side-channel attack makes use of the physical information of encryption devices to get the key. This process requires a lot of manual intervention, so there is more and more research of machine learning algorithm is applied to the side channel attacks. Among the many algorithms the neural network attack is the best. The multi-layer perceptron is the basis of neural network, and the hyper-parameters of MLP greatly affect its results of final training and attack. To achieve automatic optimization of hyper-parameter, bayesian optimization method is applied to side channel attack of MLP, and the optimization method for discrete values is proposed, and we develop a optimization method of side channel attack of MLP for hyper-parameter which can combine hyper-parametric experience. The experiments were carried out to compare the efficiency of artificial optimization and Bayesian optimization in side channel attacks of MLP, and the result verified the feasibility and efficiency of combining MLP with side channel attacks and Bayesian optimization.

Keywords Side-channel attack Multilayer perceptron Bayesian optimization

---

收稿日期: 2019 – 08 – 22。**国家重点研发计划项目**(2018YFB0904900,2018YFB0904901);**国家科技重大专项基金项目**(2014ZX0103240);"**十三五**"**国家密码发展基金项目**(MMJJ20180224);**四川省重点研发项目**(2019YFG0096)。**杨欢,硕士,主研领域:侧信道攻击与神经网络。吴震,教授。王燚,教授。杜之波,高工。王敏,博士。习伟,硕士。颜伟,硕士。**

0 引 言

在密码学中,侧信道攻击常用于提取(寻找)加密设备的密钥。在密码算法实际运行时,硬件电路会泄漏出相关的电磁辐射、功耗等侧信道信息。侧信道攻击就是指利用这些泄露出来的侧信道信息,以及密码学、统计学原理相关知识,来分析和破译密钥信息[1-4]。通常侧信道攻击分为有学习攻击和无学习攻击两类,由于有学习攻击与机器学习技术原理类似,也包括训练与测试两个阶段,所以有学习攻击又被称为基于机器学习的攻击。有学习攻击是指利用一台可以被攻击人员所控制的、与攻击目标设备相似或相同的设备(称为实验设备[5]),来建立泄露信息的能耗模型(称为模板),这样就能对这一类型设备进行攻击。正是由于有学习攻击与机器学习原理类似,使得采用机器学习方法实现有学习攻击成为可能[6]。Backes 等[5]使用机器学习技术对打印机成功进行了声学侧通道攻击; Maghrebi 等[4]分别使用随机森林、支持向量机、深度网络等机器算法在 DPAContest 比赛数据中进行了侧信道模板分析,实验结果表明基于深度网络的模板攻击在匹配成功率方面要优于其他方法。

有学习攻击中的模板攻击[5]是目前众多攻击方法中研究得最深入也是最成功的一种,常用多元高斯分布表示模板。为提高模板攻击效率,目前的部分研究着眼于改进多元高斯分布中的协方差矩阵。这些研究分别尝试了单位协方差、共享协方差、池化协方差等方法,这些改进能在一定程度上降低计算成本、提高攻击效果[6-10]。其他提高模板攻击效率方法包括: 利用相位相关性消除构架模板中的数据干扰,以构建高质量的模板; 采用主成分分析(Principal Component Analysis,PCA)方法对能耗数据进行降维处理,该方法能有效降低协方差矩阵的计算复杂度并在一定程度上提高攻击效率[11-12]。

模板攻击另一种研究思路是采用机器学习中分类算法代替传统多元高斯分布模板,例如,利用多类支持向量机作为模板实施模板攻击[13-14],采用神经网络作为模板进行模板攻击等[15-16]。在传统模板攻击中,首先使用多元高斯分布对旁路信号轨迹特征进行刻画; 然后用极大似然方法对功率跟踪进行分类。

众多研究表明,基于神经网络的模板攻击是机器学习算法与侧信道攻击结合中的最优方案,而对于神经网络而言超参数是至关重要的因素,它往往决定了该神经网络学习性能与效率。对于某一个神经网络至少必须指定控制学习速度或底层模型容量等参数,比如学习率、神经网络层数等。因为不同模型在不同场景下需要不同的超参数,而且每个超参数的意义又各不相同,这就需要技术人员拥有丰富的调参经验以及投入大量工作时间。基于此,如果没有一个成熟的自动化的调参方案,算法开发人员也很难在有限时间和计算量情况下解决这一问题。在不同实际场景中对于众多训练数据集,很难事先知道什么方案是合适的,尽管是拥有丰富机器学习算法开发经验的人员也只有通过不断地尝试和更新迭代来找到合适的解决办法。因此,实现自动调参功能是十分必要的。时至今日,非参数学习研究正在帮助深度学习更加自动的优化模型中所需的超参数选择[17],比如常用的几种寻优方法: 网格寻优、随机寻优[18]以及贝叶斯寻优[19]。其中贝叶斯寻优已经被证明在许多具有挑战性的优化基准函数上优于其他全局优化算法。对于连续函数,贝叶斯优化的工作原理通常是假设未知函数是从高斯过程中采样的,并在进行观察时保持该函数的后验分布[20]。目前尚未有将贝叶斯优化方法运用于侧信道攻击中的研究。为了使得基于神经网络的模板攻击效果更佳,本文尝试将基于贝叶斯的超参数自动优化方法应用在侧信道攻击神经网络结构寻优中,帮助我们更好地借助神经网络方法实现侧信道攻击。

基于前人的研究成果,本文借鉴了基于神经网络侧信道攻击的模型,实现了将多层感知器用于模板攻击,而且首次将贝叶斯优化方法运用在基于多层感知器的模板攻击中,为多层感知器寻找合适的超参数使其用于模板攻击时攻击效果最佳; 其次,在原有的贝叶斯更新先验分布理论基础上发展出对离散值超参数的特殊处理方法,将超参数设置的经验结合到寻优算法中,更进一步地提高寻优效率。为与贝叶斯寻优算法作对比,本文还实现了人工寻优算法,分析了两种算法的优劣。

1 有学习攻击与多层感知器

1.1 侧信道攻击

1996 年,Kocher[20]提出了侧信道攻击概念,它是利用设备功耗等侧信道泄露的相关信息来提取密钥。侧信道攻击可以分为以下两类:

(1) 有学习攻击,比如模板攻击、随机攻击[21]或被称为基于机器学习的攻击。

(2) 无学习攻击,比如差分能量分析攻击[22]、相关性能量分析[23]或被称为相互信息分析。

想要实现有学习侧信道攻击,首先需要拥有两台

几乎相同的设备。其中一台接近于目标设备，而且技术人员应该对其具有一定的控制权，然后在上面运行一个具有固定密钥值 $k \in K$ 的加密操作，K 是所有可能的密钥值的集合。对于一个有学习能力的设备，应该对其输入和密钥具有充分的认识和控制权。在这种情况下，有学习的攻击分为以下两步执行:

(1) 分析阶段，同时使用从有学习能力的设备收集到的侧信道能迹与目标设备执行加密操作的能量泄露来分析所有可能的密钥值 $k \in K$。

(2) 攻击阶段，对从目标设备泄露收集到的能迹进行分析与分类，以恢复密钥值 k。

在分析阶段，为每个可能的键值 $k \in K$ 收集一组侧信道能迹。通常情况下，应该采用分而治之的策略。例如，$K = 0,1,\cdots,255$ 意味着收集 256 组能迹来执行分析。在攻击阶段，选择从目标设备收集最合适的侧信道能迹模型来揭示正确键值 k。有学习攻击被认为是最强大的侧信道攻击形式，因为攻击者能够在攻击前描述设备的侧信道泄露信息[24]。

1.2　模板攻击

模板攻击是针对泄露密钥相关信息的中间值，利用密码芯片的能量消耗与设备正在处理数据的相关性来实现攻击。模板攻击方法最早由 Chari 等[3]在 2002 年密码硬件与嵌入式系统国际会议上提出的，由于它能利用少量泄漏信息获取密钥，逐渐成为侧信道攻击研究的热点。传统模板攻击通常使用多元高斯分布对侧信道信息特征进行刻画，一般包含三个阶段: 特征点选取，主成分分析，模板匹配[23]。特征点就是对采集到的能量轨迹进行特征提取，选取最大相关点; 主成分分析是为了减少特征点个数，来提高攻击效率; 而模板匹配则是将构建好的模板与采集的能量集进行比较，然后根据极大似然判定准则选取匹配概率最大的模板即为最好模板，而该模板对应的密钥就是最可能的正确密钥。则正确密钥 $\hat{k}$ 为:

$$\hat{k} = \arg\max \hat{p}(T \mid O_k) \tag{1}$$

式中: T 表示能量迹集合; $\hat{p}$ 为匹配概率; O_k 表示的是猜测密钥。

1.3　多层感知器

由于模板实际上就是一个分类器，自然可以将其应用于分类神经网络实现攻击。同时神经网络是研究深度学习的基础，其中的深度神经网络技术已被应用于多个领域，如图像分类、语音识别等[24]。而多层感知器可以称得上是神经网络中一个里程碑的发展，它克服了单层感知器不能解决非线性可分的问题，为整个神经网络后续发展提供了可能。

多层感知器(MLP)也叫多层神经网络(DNN)，是神经网络的一种，由几个感知器单元组成。在经典三层神经网络结构中，继续添加层次。让原来的输出层变成中间层，新增加的层次成为新输出层。假设第一层为输入层，输入值为 $[x_1, x_2, x_3]$，偏置为 b_1、b_2。对于第 l 层，用 L_l 表示该层所有神经元，其输出为 y_l，其中第 j 个节点输出为 $y_l^{(j)}$，输入为 $u_l^{(j)}$，连接第 l 层与第 $(l-1)$ 层的权重矩阵为 $\boldsymbol{W}_l$，上一层(第 $l-1$ 层)第 i 个节点到第 l 层第 j 个节点权重为 $w_l^{(ji)}$。其结构如图 1 所示。

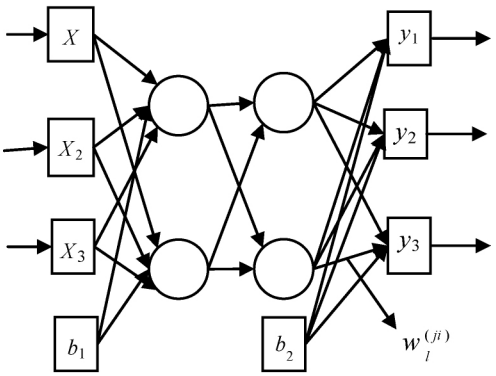

图 1　多层感知器

将 f 记为激活函数，网络中任意一层输出可以表示为:

$$\begin{cases} y_l^{(j)} = f(u_l^{(j)}) \\ u_l^{(j)} = \sum\limits_{i \in L_{i-1}} w_l^{(ji)} y_{l-1}^{(i)} + b_l^{(j)} \\ y_l = f(u_l) = f(\boldsymbol{W}_l y_{l-1} + b_l) \end{cases} \tag{2}$$

可以发现，隐含层每个神经元均是由输入特征 X 的线性组合构成。然而仅仅是线性组合，那么无论这个神经网络有多少层，结果都将与特征线性相关。于是在每个神经元计算结果之后，添加一个激活函数(Activation Function)，从而改变线性规则。常用的激活函数有修正线性函数(ReLU)、双曲正切函数(tanh)、Sigmod 函数等。

2　贝叶斯优化原理

2.1　超参数优化

超参数是指模型定义和训练中事先需要设置的参数。如神经网络层数、类型、层宽、激活函数、学习率、学习的批大小、L2 正则化项大小等。又例如机器学习算法中的支持向量机(Support Vector Machine，SVM)就有 gamma、kernel、ceof 等超参数需要调整。对于机器学习而言几乎所有算法都需要设置超参数。这些超

参数直接决定了模型性能训练是否能够成功或达到最优。然而针对不同模型找到最合适该模型的超参数组合(即超参数优化)确实是一项复杂的工作,如今越来越多调参工作使用了自动优化方法,这些方法旨在使用带有策略的启发式搜索在更短时间内找到最优超参数,除了少量初始设置之外,并不需要额外人工干预[16-17]。目前为止,已有多种较为成熟的自动调参算法,其中贝叶斯优化算法已被证明是目前最佳自动调参算法[18],本文正是通过运用这一算法对自动寻优进行研究。

2.2 贝叶斯优化算法

网格搜索和随机搜索在计算资源有限情况下推荐的结果不一定比建模工程师个人经验表现要好,而贝叶斯优化就是"很可能"比普通开发者或建模工程师调参能力更好的算法。

贝叶斯优化是贝叶斯回归的一种应用。贝叶斯回归是一种无参数函数回归方法。它利用高斯随机过程,根据观察到的函数输入和输出,使用贝叶斯定义,将假设的先验概率分布转换为后验分布。随着观察数据增加,函数的后验概率越来越准确。贝叶斯优化是根据拟合出的后验概率分布和一定策略,发现下一个寻优的位置。该策略称为 acquisition[19]。我们的目标是选择下一个观察位置,以便尽快发现该数据范围内的最大函数值。其策略需要考虑两个可能:一是开发(exploit),即在当前观察的最大值附近寻找,这是一种保守策略;二是探索(explore),即在方差最大的位置寻找,这是一种激进策略,方差最大地方可能带来意外的惊喜,而方差小的地方已经没有探索的价值了。根据这两种策略可以得到多种选择函数。典型的有 PI(Probability of Improvement)、EI(Expected Improvement)、GP-UCB(GP Upper Confidence Bound)等。

3 多层感知器的模板与贝叶斯优化

3.1 基于多层感知器的模板攻击

模板攻击是有学习攻击中研究得最深入也是有效率的一种攻击方式,而其中模板最基本形式是多元高斯分布。为进一步提高模板攻击效率,研究人员提出了多种改进方法。一种研究思路是采用机器学习中的部分算法代替传统多元高斯分布模板。例如,采用支持向量机作为模板进行模板攻击,采用神经网络作为模板进行攻击,将随机森林算法运用到模板攻击中等。本文就采用了结合多层感知器模板攻击方法。若攻击时使用 m 条攻击能迹,则密钥为:

$$K^* = \mathrm{argmax}_k \prod_{j=1}^{m} MLP(e_j \mid comb(x_j, k)) \quad (3)$$

式中:K^* 代表正确密钥;e_j 代表能量曲线;$comb(x,k)$ 代表无掩码中间组合值;$MLP(e_j \mid comb(x_j,k))$ 表示训练过程中学习到的分类器概率模型。在该攻击方式中,多层感知器训练结果起着关键作用。若通过学习得到模型本身分类效果就不好,那么攻击效果肯定也会不理想。

为成功实现基于多层感知器的模板攻击,本次实验使用的是无抖动防御的公开能迹集 ACAD。该能迹集是基于 AES 加掩码算法实现的一种电磁信号泄露数据,包含 60 000 条能迹数据,其中:50 000 条作为训练数据,10 000 条作为攻击数据[24]。采用模型训练的验证精度作为模型好坏的衡量指标。其中验证精度是机器学习中常用指标,反映了模型对数据的分类能力。模型训练验证精度定义为:

$$acc(D_{\mathrm{test}}) = \frac{\{e_i \in D_{\mathrm{test}}\} \mid K^* = \mathrm{argmax}_k Pr(K \mid e_i) \mid}{\mid D_{\mathrm{test}} \mid} \quad (4)$$

式中:D_{test} 表示验证集能迹;e_i 表示能迹;K^* 表示正确密钥;$Pr(K \mid e_i)$ 表示猜测密钥。简而言之,验证精度就是当猜测密钥与正确密钥相等时的能迹数与验证集能迹数之比。

有了数据集以及衡量指标作为前提,基于多层感知器的模板攻击具体实现步骤分为训练与攻击两个阶段,其中训练阶段主要步骤如下。

(1) 采用 PCA 对加掩 AES 算法能迹数据集进行降维处理。

(2) 针对降维后的能迹数据,建立相应多层感知器(MLP)模板。

(3) 利用数据集中的 50 000 条能迹数据对预测网络进行训练。

攻击阶段的主要步骤如下:

(1) 利用训练的多层感知器模板,对验证集能迹 D_{test} 进行分类。

(2) 根据分类结果计算正确密钥对应的能迹数与验证集能迹数之比也就是最大验证精度。

3.2 基于贝叶斯寻优先验分布的更新

贝叶斯优化与其他优化方法的不同之处在于,它首先为 $f(x)$ 构造一个概率模型(即先验分布),然后利用这个模型来决定下一步 x 取什么数值来计算函数值(更新先验分布)。其基本思想是利用以前对 $f(x)$ 评估所获得的所有信息,而不是简单地依赖于局部梯度和海森近似。

在使用贝叶斯优化时,首先通过样本点对高斯过

程进行估计与更新,然后通过选择函数来确定新的采样点。自然对贝叶斯寻优的研究重点就放在了高斯过程与选择函数上。

3.2.1　高斯过程

贝叶斯优化过程中使用高斯过程作为先验函数,它是一种强大、方便的先验分布函数,能有效拟合现实中的分布[20]。一个完整的高斯过程由且仅由均值函数 $m(x)$ 与协方差函数 $k(x,x')$ 确定,其中均值函数为一个向量,而协方差函数为矩阵。如此一来高斯过程 gp 就可以表示为:

$$f \sim gp(\boldsymbol{m},\boldsymbol{k}) \tag{5}$$

现假设有一组样本点 $D=\{(x_{1:t},y_{1:t})\}$,其协方差矩阵为:

$$\boldsymbol{K}=\begin{bmatrix} k(x_1,x_1) & k(x_1,x_2) & \cdots & k(x_1,x_t) \\ k(x_2,x_1) & k(x_2,x_2) & \cdots & k(x_2,x_t) \\ \vdots & \vdots & & \vdots \\ k(x_t,x_1) & k(x_t,x_2) & \cdots & k(x_t,x_t) \end{bmatrix} \tag{6}$$

一个新样本 x_{t+1} 加入会更新上述协方差矩阵 $\boldsymbol{K}$,假设有 $\boldsymbol{k}=[k(x_{t+1},x_1),k(x_{t+1},x_2),\cdots,k(x_{t+1},x_t)]$ 那么更新后的协方差可表示为:

$$\boldsymbol{K}=\begin{bmatrix} \boldsymbol{K} & \boldsymbol{k}^{\mathrm{T}} \\ \boldsymbol{k} & k(x_{t+1},x_{t+1}) \end{bmatrix} \tag{7}$$

有了更新后协方差矩阵就可以通过前 t 个样本估计出 f_{t+1} 的后验概率分布:

$$P(f_{t+1}\mid D_{1:t},x_{t+1})\} \sim N(\boldsymbol{u},\boldsymbol{\sigma}^2) \tag{8}$$

$$\boldsymbol{u}=\boldsymbol{k}^{\mathrm{T}}\boldsymbol{K}^{-1}f_{1:t} \tag{9}$$

$$\boldsymbol{\sigma}^2=k(x_{t+1},x_{t+1})-\boldsymbol{k}^{\mathrm{T}}\boldsymbol{K}^{-1}\boldsymbol{k} \tag{10}$$

上述公式的详细推导过程以及核函数与均值函数的选择详见文献[18]。

可以根据新加入的样本点对先验中高斯过程进行更新,使其能更好拟合现实情况。

3.2.2　选择函数

有了先验概率分布,接下来就需要通过选择函数确定用于更新先验的采样点。它是决定贝叶斯优化能否成功的另一重要因素。通过该采样点获得后验分布,使该分布更贴切模拟现实情况。虽然目前已有多种选择函数可供使用但它们主要思想都是平衡上面提到的两种策略即 explore 与 exploit。本文使用 EI 准则作为选择函数,经验证 EI 准则比 PI 表现得更好[19],而且与 GP_UCB 不同的是它自己不再需要确定额外参数。EI 准则 $a(\cdot)$ 如下:

$$a(x\mid D)=E_{y\sim f(x\mid D)}[\max(0,y-f_{\mathrm{best}})]=\begin{cases}\sigma(x)z\varphi(x)+\sigma(x)\Phi(z) & y>f_{\mathrm{best}}\\ 0 & \text{其他}\end{cases} \tag{11}$$

式中: f_{best} 为数据集 D 上的最大值; $z=\dfrac{f_{\mathrm{best}}-u(x)}{\sigma(x)}$; $E_{y\sim f(x\mid D)}$ 为期望函数; $\varphi(\cdot)$ 与 $\Phi(\cdot)$ 分别是高斯分布累计概率函数与概率密度函数。EI 准则最大优点就是能在 explore 与 exploit 两种策略间保持平衡,当 explore 时选择均值最大的点,exploit 时选择方差大的点。

4　实　验

为验证提出的基于多层感知器的攻击方法,本文对 AES 加掩数据集进行了攻击实验; 同时为了验证贝叶斯寻优算法,在攻击模型基础上使用了两种不同寻优方法: 人工寻优与贝叶斯寻优,并对这两种寻优方式进行了比较。为了进一步提高贝叶斯寻优算法效率,根据贝叶斯寻优理论知识分析了离散值超参数与连续值超参数的处理方式,并进行了改进。

4.1　实验环境

本文实验使用了 ASCAD 公开数据集作为实验数据,采用最大精度衡量优化算法好坏的指标。本文针对 MLP 神经网络所需超参数的搜索空间如表 1 所示。

表 1　多层感知器的超参数及其范围

超参数	类型	最小值	最大值
隐含层	int	1	3
神经元大小	enum	16	64
PCA	int	40	100
L2	enum	0.0	1.0
Activation Function	strenum	0	2

4.2　实验实施

4.2.1　离散值与连续值

(1) 离散超参数寻优。根据 3.2 节中描述的高斯过程中计算后验概率步骤,将离散型超参数看作枚举型(enum)处理。如需要推荐超参数中的 Layer_size 与 L2,它们的可能取值分别为 [16,32,48,64] 与 [0.0,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1.0],而后在处理这些数据时将每一个可能取值都当作单独一维,根据 sobol 序列函数产生的随机矩阵中最大的某一位置的原矩阵(存储的是需要优化超参数所有可能取值)值为本次推荐超参数。这样推荐参数的随机性较大,每个超参数成为推荐值概率都一样; 但由于新采样点是由 EI 准则所确定,若下一个采样点不存在该离散超参数取值范围内,就只能取最靠近计算出来采样点存在于取值范围内的值作为新的采样点,这样就不能

使更新后的后验分布更好模拟真实情况。

(2) 连续超参数寻优。连续超参数寻优与离散超参数处理类似,唯一不同是将连续超参数看作整型(int)处理。在配置文件中存的是它的取值范围,而不是具体值;在后面处理该类型数据时将每一种连续超参数视为一维,同样根据 sobol 序列函数产生的随机矩阵中最大的某一位置对应原矩阵中的值为本次推荐超参数。连续值超参数寻优较离散值超参数寻优的优点为:根据 EI 准则得到下一个采样点可以精确用于更新先验分布,能最大程度模拟真实情况中的分布。

由于离散型超参数不能取到两个数之间的值,本文借鉴了对连续超参数的处理方式对离散超参数进行优化:将一个离散超参数当且仅当作一维,通过一系列的处理,将应推荐超参数对应位置的值置为最大,从而得到推荐值。根据以上所述的两种处理方式,本文做了一个针对离散超参数寻优不同处理方式的对比实验,实验结果如图 2 所示。

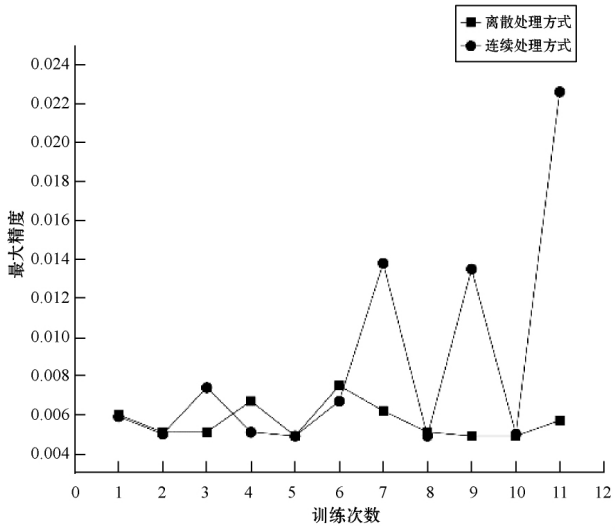

图 2 离散超参数寻优不同处理方式对结果的影响

由对比实验可知对离散值采取类似连续值处理方式要优于离散值原本处理方式。为进一步提高贝叶斯寻优效率,在后面实验过程中对离散值均采取了类似连续值的处理方式。

4.2.2 贝叶斯寻优与人工寻优

对比了采用贝叶斯优化算法进行自动寻优与人工寻优得到超参数所构建的多层感知器(MLP)网络结构运用于模板攻击中的两种结果。

(1) 贝叶斯自动寻优算法。针对上面需要寻优的 5 种超参数共进行了 5 轮贝叶斯自动寻优,将每轮寻优结果用于构建 MLP,将该 MLP 用于模板攻击,最终得到了每一轮攻击结果的最大精度,如图 3 所示。

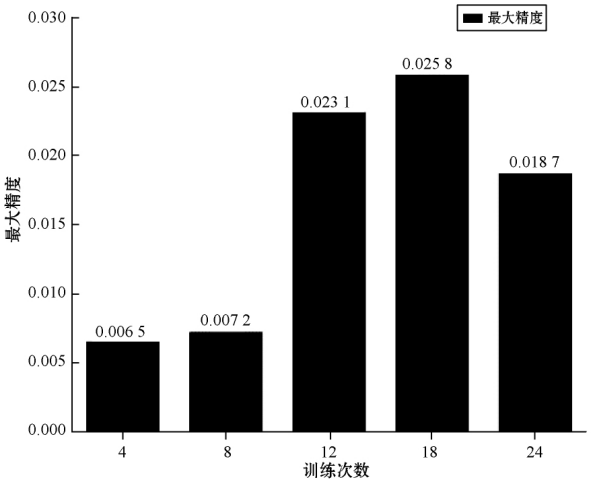

图 3 贝叶斯寻优算法在不同推荐次数下实验对比

在 5 轮推荐中,第 4 轮 18 次推荐次数使得攻击结果最大精度达到 0.025 8。其次,可以看出推荐次数与最大精度并不严格形成正比关系,但若是推荐次数过小,想要得到较好的结果则不太可能。

表 2 每一轮推荐的超参数结果

训练次数	隐含层	神经元大小	PCA	L2	Activation Fucnction
4	1	16	40	0.0	ReLU
8	1	16	40	0.0	ReLU
11	1	64	98	0.0	tanh
18	1	64	94	0.0	Sigmoid
24	1	64	41	1.0	Sigmoid

由图 3 和表 2 可知第四轮推荐结果最佳。当隐含层、神经元大小、PCA、L2 分别取 1、64、94、0.0 时,激活函数为 Sigmoid 构成的 MLP 用于模板攻击结果最佳。

(2) 人工寻优。在该部分中,实验对 MLP 需要的超参数进行了人工调整,得到了一个相对较好的结果。在此过程中针对 5 个超参数共进行了 40 次实验。实验采用了排列组合方式,即针对每个超参数可能取值都进行了测试。出于时间与人工成本考虑,实验过程中首先固定其他 4 个超参数取值再对剩下超参数所有可能取值进行测试,找到该超参数使得实验结果表现最佳的值。在接下来的实验中固定该超参数值,依次改变其他超参数,最终得到一个在所有实验中使得模板攻击在最大准确度与最小损失衡量下均表现为最佳的组合。对于上面的实验描述,得到实验结果如图 4 所示。

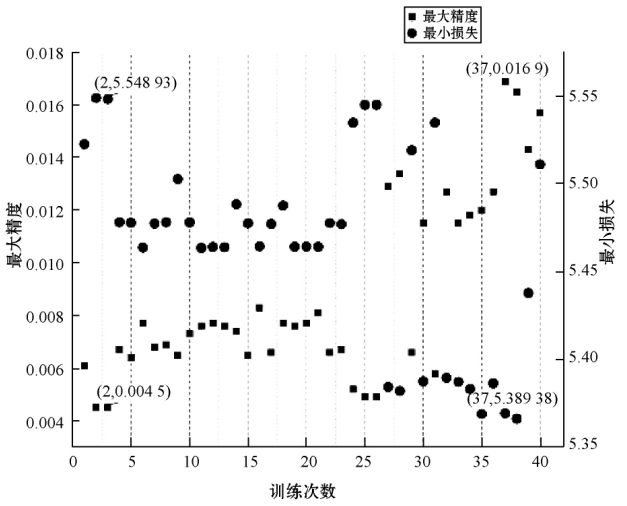

图 4　人工调参实验结果

在 40 次实验中表现最好为第 37 次实验，它最大精准度与最小损失分别为 0.016 9 与 5.369 38; 而表现最差为第 2 次实验，最大精度仅有 0.004 5，如表 3 所示。

表 3　人工调参中实验结果表现最好与最差超参数

训练次数	隐含层	神经元大小	PCA	L2	Activation Fucnction
37	1	48	42	0.0	ReLU
2	1	16	40	0.0	tanh

4.3　实验结果

经过上面两种不同寻优算法实验，得到了两种算法各自对应的最佳结果，现就两种算法得到结果进行对比并分析它们各自优缺点。将上面进行的 5 轮贝叶斯超参数优化实验中使得攻击结果表现得最好与最差的第 1 轮和第 4 轮分别从最大精度与最小损失以及所耗费的时间三个方面与使用人工寻优算法的攻击结果进行了对比，结果如图 5 所示。具体分析两种算法每次训练结果对实验的影响，结果如图 6 所示。

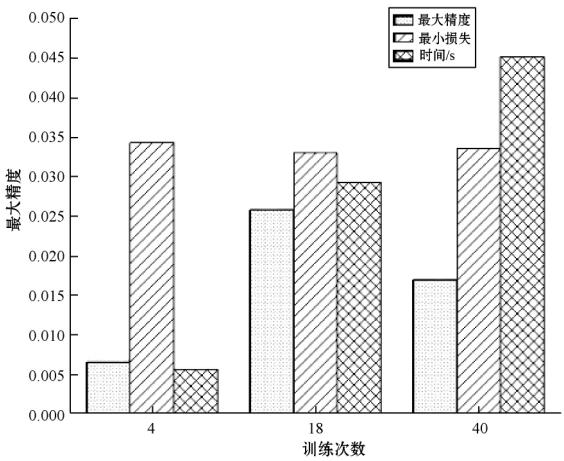

图 5　贝叶斯寻优算法与人工寻优的结果对比

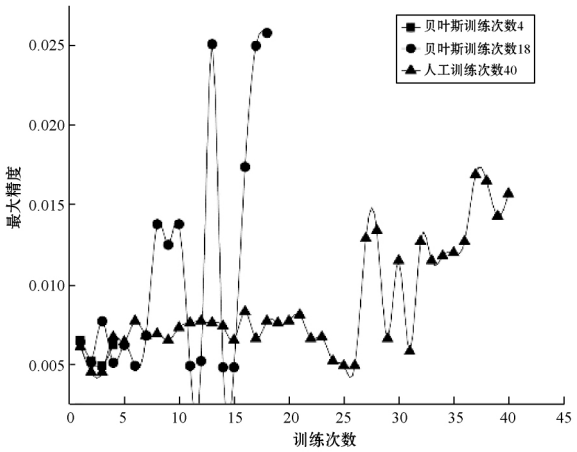

图 6　贝叶斯寻优与人工寻优每次寻优结果对比

可以看到，使用贝叶斯寻优算法时若推荐(训练) 次数过小会使得实验结果表现远不如训练次数多的人工寻优算法。可当贝叶斯寻优算法训练次数达到一定值时，不管是最大精度还是最小损失均优于人工寻优，且此时贝叶斯寻优算法推荐次数还不到人工寻优的一半，与之对应时间是使用人工寻优算法的二分之一。

4.4　实验分析

贝叶斯自动寻优算法与人工寻优算法各自的优缺点如下: 贝叶斯自动寻优算法不用人为干预，只要确定需要训练的超参数就可以得到该轮训练最佳结果; 达到相同实验结果贝叶斯自动调参算法效率更高。但贝叶斯算法存在偶然性，若是本轮训练次数少那么它的结果大概率会表现得很糟糕。由图 6 可以看到贝叶斯寻优明显存在冷启动问题，而人工寻优算法寻找组合更为全面，但是效率低下，需要大量重复劳动力。

5　结　语

本文基于现有的侧信道攻击神经网络调参技术，提出贝叶斯自动寻优算法，该方法利用高斯过程建立先验分布模型，通过选择函数确定采样点得到后验分布，继而根据后验更新先验分布。实验结果表明，贝叶斯寻优算法相对人工寻优具有高效性以及独立性，它能提供比其他寻优算法更好的学习效率，有利于减少训练次数，攻击质量取得了明显提高。在今后研究中，我们将致力于探索如何消除贝叶斯自动寻优算法冷启动的问题，以便于进一步提高贝叶斯寻优算法效率。

参　考　文　献

[1] 杜之波，吴震，王敏，等. 针对 SM4 轮输出的改进型选择明文功耗分析攻击 [J]. 通信学报，2015，36(10) : 85 - 91.

[2] 杜之波，吴震，王敏，等. 针对基于 SM3 的 HMAC 的能量分析攻击方法 [J]. 通信学报，2016，37(5) : 38 - 43.

[3] Chari S,Rao J R,Rohatgi P. Template attacks [C] //International Workshop on Cryptographic Hardware and Embedded Systems,2002.

[4] Maghrebi H,Portigliatti T,Prouff E. Breaking cryptographic implementations using deep learning techniques [C] //International Conference on Security,Privacy,and Applied Cryptography Engineering,2016:3 - 26.

[5] Backes M,Duermuth M,Gerling S,et al. Acoustic side-channel attacks on printers [C] //19th USENIX Conference on Security,2010.

[6] 刘飚,孙莹. 基于公共协方差矩阵的实用模板攻击 [J]. 计算机应用研究,2016(1):236 - 239.

[7] 崔琦,王思翔,段晓毅,等. 一种 AES 算法的快速模板攻击方法 [J]. 计算机应用研究,2017,34(6):1801 - 1804.

[8] Choudary O,Kuhn M G. Efficient template attacks [C] //International Conference on Smart Card Research and Advanced Applications,2014.

[9] Archambeau C,Peeters E,Standert X,et al. Template attacks in principal subspaces [M] //Cryptographic Hardware and Embedded Systems-CHES 2006. Springer,2006:1 - 14.

[10] 王红胜,徐子言,张阳,等. 基于模板的光辐射分析攻击 [J]. 计算机应用研究,2017,34(7):2151 - 2154.

[11] 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,2012.

[12] Heuser A,Zohner M. Intelligent machine homicide [M] //Constructive Side-Channel Analysis and Secure Design. Springer,2012:249 - 264.

[13] Martinasek Z,Zeman V. Innovative method of the power analysis [J]. Radioengineering,2013,22(2):586 - 594.

[14] Martinasek Z,Hajny J,Malina L. Optimization of power analysis using neural network [C] //International Conference on Smart Card Research and Advanced Applications,2013.

[15] Bergstra J,Bardenet R,Bengio Y,et al. Algorithms for hyper-parameter optimization [C] //International Conference on Neural Information Processing Systems,2011.

[16] Bergstra J,Bengio Y. Random search for hyper-parameter optimization [J]. Journal of Machine Learning Research,2012,13(1):281 - 305.

[17] Snoek J,Larochelle H,Adams R P. Practical bayesian optimization of machine learning algorithms [C] //25th International Conference on Neural Information Processing Systems,2012.

[18] Rasmussen C E,Williams C K I. Gaussian processes for machine learning [M]. MIT Press,2005.

[19] Schindler W,Lemke K,Paar C. A stochastic model for differential side channel cryptanalysis [C] //International Workshop on Cryptographic Hardware and Embedded Systems,2005:30 - 46.

[20] Kocher P C. Timing attacks on implementations of diffie-hellman,RSA,DSS,and other systems [M] //Advances in Cryptology-CRYPTO'96. Springer,1996:104 - 113.

[21] Kocher P,Jaffe J,Jun B. Introduction to differential power analysis and related attacks [C] //Annual International Cryptology Conference,1999:388 - 397.

[22] Benadjila R,Prouff E,Strullu R,et al. Study of deep learning techniques for side-channel analysis and introduction to ASCAD database [J]. Journal of Cryptographic Engineering,2020,10:163 - 188.

[23] 郭东昕,陈开颜,张阳,等. 针对密码芯片的模板攻击研究综述 [J]. 飞航导弹,2018(12):79 - 83.

[24] Lecun Y,Bengio Y,Hinton G. Deep learning [J]. Nature,2015,521(7553):43.

(上接第 287 页)

[20] West D B. Introduction to graph theory [M]. 2nd ed. Upper Saddle River: Prentice Hall,2001.

[21] Liu C Q,Li Y G,Li Z Y. A machining feature definition approach by using two-times unsupervised clustering based on historical data for process knowledge reuse [J]. Journal of Manufacturing Systems,2018,49:16 - 24.

[22] Getto R,Kuijper A,Fellner D W. Unsupervised 3D object retrieval with parameter-free hierarchical clustering [C] //Proceedings of the Computer Graphics International Conference. ACM,2017.

[23] 马露杰,黄正东,梁良,等. CAD 模型表面区域分割方法 [J]. 计算机辅助设计与图形学学报,2009,21(2):148 - 153.

[24] 张波. 基于三维模型智能工艺设计技术研究 [D]. 黑龙江:哈尔滨工业大学,2013.

[25] Mukherjee J,Mukhopadhyay J,Mitra P. A survey on image retrieval performance of different bag of visual words indexing techniques [C] //Proceedings of the 2014 IEEE Students' Technology Symposium. Los Alamitos: IEEE Computer Society Press,2014:99 - 104.

[26] 唐德权,吴绍兵,凌志刚. 一种新的图聚类算法研究 [J]. 计算机应用与软件,2014,31(6):18 - 21,58.

[27] 徐欣,舒阵宇,陈双敏,等. 基于决策图的三维模型无监督聚类算法 [J]. 宁波大学学报(理工版),2018,31(4):46 - 51.

[28] 吴烨,钟志农,熊伟,等. 一种高效的属性图聚类方法 [J]. 计算机学报,2013,36(8):1704 - 1713.

[29] 何家玉. 谱聚类算法的研究与应用 [D]. 安徽:安徽理工大学,2017.