(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 113158181 B
(45)授权公告日 2022.04.05

(21)申请号 202110406620.5
(22)申请日 2021.04.15
(65)同一申请的已公布的文献号
申请公布号 CN 113158181 A
(43)申请公布日 2021.07.23
(73)专利权人 上海交通大学
地址 200240 上海市闵行区东川路800号
专利权人 国网宁夏电力有限公司电力科学研究院
(72)发明人 陆相君 谷大武 陆海宁 沙伟燕 张佩
(74)专利代理机构 上海交达专利事务所 31201
代理人 王毓理 王锡麟
(51)Int.Cl.
G06F 21/55(2013.01)
G06K 9/62(2022.01)
G06N 3/04(2006.01)
G06N 3/08(2006.01)

(56)对比文件
CN 111565189 A,2020.08.21
CN 111985411 A,2020.11.24
CN 112615714 A,2021.04.06
CN 112367396 A,2021.02.12
CN 111680787 A,2020.09.18
CN 110933031 A,2020.03.27
CN 112464248 A,2021.03.09
US 2019130101 A1,2019.05.02
IN 201721015989 A,2017.06.16
谷大武 等.“密码系统的侧信道分析:进展与问题”.《西安电子科技大学学报》.2021,第48卷(第1期),14-49.
黄洁 等.“适用于侧信道分析的卷积神经网络结构的实验研究”.《成都信息工程大学学报》.2019,第34卷(第5期),449-456.

审查员 甄红欣

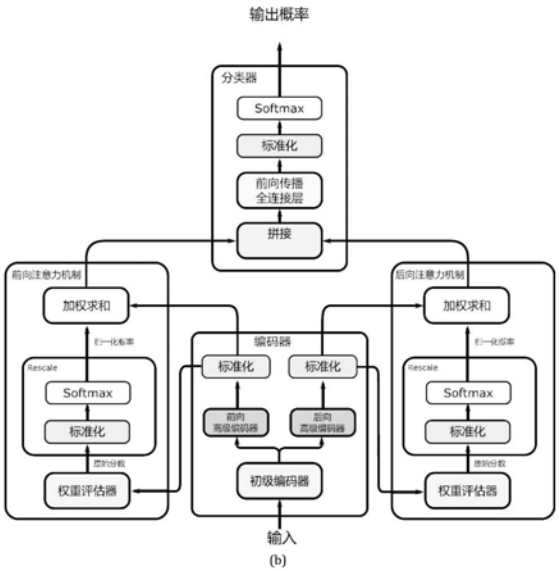
权利要求书2页 说明书4页 附图5页

(54)发明名称

使用神经网络对原始侧信道数据进行端到端攻击的方法

(57)摘要

一种使用神经网络对原始侧信道数据进行端到端攻击的方法,通过构建神经网络并使用任意原始侧信道数据对神经网络进行训练,并从中随机不超过其中50%的数据作为验证集,训练过程中使用交叉熵作为损失函数,当验证数据集上的损失函数数值开始上升时,停止训练;使用训练后的神经网络在攻击数据集上开展攻击,即将攻击数据集输入网络,网络将返回每条侧信道信息曲线的分类概率,根据多条侧信道信息分类的概率值,利用最大似然估计,求得一组侧信道信息背后密码算法运行时的密钥值;本发明能够自动地在原始侧信道曲线上发现并组合掩码和中间值泄露,并利用组合后的信息直接进行侧信道攻击,有效解决掩码防护下进行侧信道攻击困难的问题。



(b)

CN 113158181 B

1.一种使用神经网络对原始侧信道数据进行端到端攻击的方法,其特征在于,包括:

步骤1,构建包括初级编码器、高级编码器、至少一个注意力机制单元以及分类器的神经网络用于后续的训练和攻击步骤;

步骤2,使用任意原始侧信道数据对步骤1搭建的神经网络进行训练,并从中随机不超过其中50%的数据作为验证集,训练过程中使用交叉熵作为损失函数,当验证数据集上的损失函数数值开始上升时,停止训练;

步骤3,使用训练后的神经网络在攻击数据集上开展攻击,即将攻击数据集输入网络,网络将返回每条侧信道信息曲线的分类概率,根据多条侧信道信息分类的概率值,利用最大似然估计,求得一组侧信道信息背后密码算法运行时的密钥值,具体包括:

步骤3.1,使用神经网络中的初级编码器对原始侧信道数据进行细粒度特征提取,同时实现对原始侧信道数据的维度压缩;

步骤3.2,使用神经网络中的高级编码器,对步骤3.1得到的细粒度特征进行组合,从而达到结合掩码和被掩中间值泄露信息的目的,以实现最终的端到端攻击;

步骤3.3,使用注意力机制单元计算高级编码器输出的组合后的特征之间的权重并求和,对最终得到的特征向量使用分类器进行分类,得到该侧信道信息属于不同中间值类别的概率;

步骤3.4,使用一层全连接层和softmax为加权求和后的特征向量分类;

所述的攻击数据集是指:在训练过程中未使用的,用于实际攻击密码算法的数据集,在机器学习的分类问题背景下通常被称为测试集,用于测试匹配率,在侧信道背景下,因为测试集的分类概率可直接用于侧信道攻击,即攻击数据集;

两个LSTM网络结构按照数据通道维度结合或按照时间维度结合,其中:按照数据通道维度结合时,中间特征向量的通道数量翻倍,时间步数量不变且两个LSTM网络结构共享一个批标准化操作;按照时间维度结合时,中间特征向量的通道数量不变,时间步翻倍且两个LSTM网络结构分别具有独立的批标准化操作;

所述的至少一个注意力机制单元是指:当两个LSTM网络结构按照数据通道维度结合时,一个注意力机制单元的输入端与两个LSTM网络结构结合后的输出端相连;当两个LSTM网络结构按照时间维度结合时,两个相互独立的注意力机制单元的输入端分别与两个LSTM网络结构的输出端相连,两个方向不同的注意力机制单元相互配合,从不同方向确定侧信道信息的主要泄露区间,帮助上层LSTM缩小实际训练中的学习序列长度。

2.根据权利要求1所述的使用神经网络对原始侧信道数据进行端到端攻击的方法,其特征是,所述的初级编码器包括:局部连接网络和卷积网络,其中:局部连接网络由一层局部连接层和整形(Reshape)层组成,卷积网络包括至少一层卷积层和最大池化层。

3.根据权利要求1所述的使用神经网络对原始侧信道数据进行端到端攻击的方法,其特征是,所述的高级编码器包括:两个分别与初级编码器的输出端相连的长短期记忆(LSTM)网络结构,该两个LSTM网络结构分别按照时间顺序从前到后和从后到前遍历数据。

4.根据权利要求1所述的使用神经网络对原始侧信道数据进行端到端攻击的方法,其特征是,所述的注意力机制单元包括单个神经元和softmax激活函数,其中:单个神经元按照相同标准给所有时间步的重要性打分,softmax激活函数将分数映射成概率,该概率可以进一步控制不同时间步之间的权重,以权重形式作用于注意力单元的输入,并给出加权求

和的结果向量，从而帮助高级编码器中的LSTM在大量的时间步中筛选重要时间步，并通过不同权重控制训练过程中的梯度方向，起到了软性的时间步截断作用。

5.根据权利要求1所述的使用神经网络对原始侧信道数据进行端到端攻击的方法，其特征是，所述的原始侧信道数据是指：未经过特征点选择和对齐处理的侧信道数据。

6.根据权利要求1所述的使用神经网络对原始侧信道数据进行端到端攻击的方法，其特征是，当原始侧信道信息未对齐或无法对齐，通过替换不同结构的初级编码器，对未对齐的原始侧信道数据进行细粒度特征提取。

7.根据权利要求2所述的使用神经网络对原始侧信道数据进行端到端攻击的方法，其特征是，所述的局部连接网络使用局部权重对局部侧信道信息进行点乘处理，卷积网络使用共享权重对全局侧信道信息进行点乘处理，所述两个LSTM网络都可提取到细粒度的侧信道特征。

8.根据权利要求1所述的使用神经网络对原始侧信道数据进行端到端攻击的方法，其特征是，所述的高级编码器中LSTM的数据流控制门可根据训练数据的不同对自身的权重向量进行自动化学习，并依不同门控逻辑中的不同权重值对数据流和内部存储单元进行输入、输出、记忆和遗忘操作。

使用神经网络对原始侧信道数据进行端到端攻击的方法

技术领域

[0001] 本发明涉及的是一种信息安全领域的技术，具体是一种使用神经网络对原始侧信道数据进行端到端攻击的方法，即直接对掩码防护下未对齐的高维原始侧信道数据进行侧信道攻击。

背景技术

[0002] 侧信道分析主要是基于物理特征的分析技术，包括功耗分析，电磁分析，错误分析，时间分析等等，其中功耗分析是指通过分析密码运算过程中呈现的电流/电压变化得出功耗的变化，进而将功耗与密钥信息联系起来，最终获取密钥信息。其进一步包括简单功耗分析(SPA)和差分功耗分析(DPA)：SPA是指根据功耗曲线上所呈现的特殊特征来推测密钥信息，DPA利用的是操作数的变化所引起的微小的功耗变化，需要通过对大量功耗曲线进行统计分析才能得出密钥信息。电磁分析与功耗分析类似，只是获取曲线的方式有别。错误分析是利用错误结果进行分析得出密钥信息的分析技术。时间分析是指有的算法运行时间会因密钥的不同而不同，因而可以通过运行时间来推测密钥。

[0003] 现有的随机掩码防护方法(masking)采用随机数掩盖运算过程中的真实数据，以此阻止攻击者找到中间值和侧信道信息的直接相关关系。此时随机掩码和中间值依然会在侧信道信息中泄露，但由于掩码的随机性，攻击者无法直接定位它们的位置，致使攻击代价大幅提高。由于所述掩码的存在，在实际分析中进行直接的特征点选择是不可行的，而结合不同位置功耗点的高阶攻击会因为原始曲线过长而消耗大量存储资源(随掩码阶数增长呈指数增长)，也不可行。

发明内容

[0004] 本发明针对现有技术存在的上述不足，提出一种使用神经网络对原始侧信道数据进行端到端攻击的方法，使用神经网络技术直接对掩码防护下未对齐高维原始侧信道信息进行攻击，能够自动地在原始侧信道曲线上(无需预先时序对齐或降维)发现并组合掩码和中间值泄露，并利用组合后的信息直接进行侧信道攻击，有效解决掩码防护下进行侧信道攻击困难的问题。

[0005] 本发明是通过以下技术方案实现的：

[0006] 本发明涉及一种使用神经网络对原始侧信道数据进行端到端攻击的方法，包括：

[0007] 步骤1，构建包括初级编码器、高级编码器、至少一个注意力机制单元以及分类器的神经网络用于后续的训练和攻击步骤。

[0008] 所述的初级编码器包括：局部连接网络和卷积网络，其中：局部连接网络由一层局部连接层和整形(Reshape)层组成，卷积网络包括至少一层卷积层和最大池化层。

[0009] 所述的高级编码器包括：两个分别与初级编码器的输出端相连的长短期记忆(LSTM)网络结构，该两个LSTM网络结构分别按照时间顺序从前到后和从后到前遍历数据。

[0010] 所述的两个LSTM网络结构按照数据通道维度结合或按照时间维度结合，其中：按

照数据通道维度结合时,中间特征向量的通道数量翻倍,时间步数量不变且两个LSTM网络结构共享一个批标准化(batch normalization)操作;按照时间维度结合时,中间特征向量的通道数量不变,时间步翻倍且两个LSTM网络结构分别具有独立的批标准化操作。

[0011] 所述的至少一个注意力机制单元是指:当两个LSTM网络结构按照数据通道维度结合时,一个注意力机制单元的输入端与两个LSTM网络结构结合后的输出端相连;当两个LSTM网络结构按照时间维度结合时,两个相互独立的注意力机制单元的输入端分别与两个LSTM网络结构(FWAttention和BWAttention)的输出端相连,两个方向不同的注意力机制单元相互配合,从不同方向确定侧信道信息的主要泄露区间,帮助上层LSTM缩小实际训练中的学习序列长度。

[0012] 所述的注意力机制单元包括单个神经元和softmax激活函数,其中:单个神经元按照相同标准给所有时间步的重要性打分,softmax激活函数将分数映射成概率,该概率可以进一步控制不同时间步之间的权重,以权重形式作用于注意力单元的输入,并给出加权求和的结果向量,从而帮助高级编码器中的LSTM在大量的时间步中筛选重要时间步,并通过不同权重控制训练过程中的梯度方向,起到了软性的时间步截断作用。

[0013] 所述的分类器为使用softmax作为激活函数的全连接层。

[0014] 步骤2,使用任意原始侧信道数据对步骤1搭建的神经网络进行训练,并从中随机不超过其中50%的数据作为验证集,训练过程中使用交叉熵作为损失函数,当验证数据集上的损失函数数值开始上升时,停止训练。

[0015] 所述的原始侧信道数据是指:未经过特征点选择和对齐处理的侧信道数据。

[0016] 步骤3,使用训练后的神经网络在攻击数据集上开展攻击,即将攻击数据集输入网络,网络将返回每条侧信道信息曲线的分类概率,根据多条侧信道信息分类的概率值,利用最大似然估计,求得一组侧信道信息背后密码算法运行时的密钥值。

[0017] 所述的攻击数据集是指:在训练过程中未使用的,用于实际攻击密码算法的数据集,在机器学习的分类问题背景下通常被称为测试集,用于测试匹配率,在侧信道背景下,因为测试集的分类概率可直接用于侧信道攻击,也被称为攻击数据集。

[0018] 所述的步骤3具体包括:

[0019] 步骤3.1,使用神经网络中的初级编码器对原始侧信道数据进行细粒度特征提取,同时实现对原始侧信道数据的维度压缩。

[0020] 优选地,当原始侧信道信息未对齐或无法对齐,通过替换不同结构的初级编码器,对未对齐的原始侧信道数据进行细粒度特征提取。

[0021] 所述的局部连接网络使用局部权重对局部侧信道信息进行点乘处理,卷积网络使用共享权重对全局侧信道信息进行点乘处理,所述两种网络都可提取到细粒度(一个或几个时钟周期提取一个特征向量)的侧信道特征。

[0022] 步骤3.2,使用神经网络中的高级编码器,对步骤3.1得到的细粒度特征进行组合,从而达到结合掩码和被掩中间值泄露信息的目的,以实现最终的端到端攻击。

[0023] 所述的高级编码器中LSTM的数据流控制门可根据训练数据的不同对自身的权重向量进行自动化学习,并依不同门控逻辑中的不同权重值对数据流和内部存储单元进行输入、输出、记忆和遗忘操作。

[0024] 步骤3.3,使用注意力机制单元计算高级编码器输出的组合后的特征之间的权重

并求和,对最终得到的特征向量使用分类器进行分类,得到该条侧信道信息属于不同中间值类别的概率。

[0025] 步骤3.4,使用一层全连接层和softmax为加权求和后的特征向量分类。

[0026] 技术效果

[0027] 与现有技术相比,本发明直接使用原始侧信道信息建模,可实施切实可行的端到端攻击,可以省略对有掩码防护的实现进行侧信道攻击时的特征点选择过程。

附图说明

[0028] 图1为实施例中两种网络构型的抽象结构实例;

[0029] 图2为本发明初级编码器中局部连接层示意图;

[0030] 图3为本发明初级编码器中卷积层示意图;

[0031] 图4为使用神经网络对原始侧信道数据进行端到端攻击的方法示意图;

[0032] 图5-图7为实施例效果示意图。

具体实施方式

[0033] 本实施例针对开源数据集ASCAD进行分析,ASCAD数据集中单个时钟周期长度约为52个时间点,本实施例涉及一种用于对原始侧信道数据进行端到端攻击的神经网络,包括:初级编码器、高级编码器、注意力机制单元以及分类器。

[0034] 所述初级编码器中的局部连接网络,该局部连接网络包括:局部连接层和整形(Reshape)层。

[0035] 所述的局部连接层中的过滤器(filter)大小是原始曲线中一个时钟周期长度的整数倍,通常取一到两个时钟周期,步进(stride)可以整除过滤器的长度,通常取过滤器的长度的一半。

[0036] 所述的整形层的整形参数为:(-1,int(f/s)),其中:f为过滤器长度,s为步长,int为取整。

[0037] 所述初级编码器中的卷积网络,该卷积网络包括:若干卷积层和池化层,其中:第一层卷积层的卷积核长度为应用曲线集的时钟周期长度步进为1,其他卷积层卷积核长度皆为3,步进皆为1;池化层中池化长度皆为2,步进皆为2,使用最大池化。

[0038] 所述的卷积层在每次池化层后通道数量翻倍。

[0039] 所述的高级编码器采用的长短期记忆结构(LSTM),分别从正向和反向遍历所有的初级编码器输出,并依据侧信道信息复杂程度不同,使用不同的组合方式进行结合(按时间维度或按数据通道维度),LSTM中的单元数为128或256,激活函数为tanh,循环激活函数为sigmoid。

[0040] 所述两个方向不同的LSTM网络结构,其各自拥有独立的注意力机制,使得注意力机制有方向性,两个方向不同的注意力机制可以相互配合确定侧信道信息泄露的主要区间,帮助上层LSTM缩小实际训练中的学习序列长度。

[0041] 所述的注意力机制单元直接作用于高级编码器的输出之上,使用单神经元的网络结构,按照统一标准评判每一个时间步数据的重要程度,该单神经元的网络结构的输出将输入softmax激活函数,最终得到求和为1的一组概率值,并使用该组概率值,将高级编码器

所有时间步的输出进行加权求和，具体为：$a' = \mathrm{BatchNorm}(v^T H)$，$a = \mathrm{softmax}(a')$，$r = Ha^T$，其中：H为高级编码器输出，v为单神经元中的可训练权重向量，a′为加权分数，a为注意力概率向量，r为加权求和后的特征向量。

[0042] 本发明中的注意力机制额外加入了批标准化(batch normalization)操作，该操作会在统一标准的加权分数a′之上，再对每个不同的时间步引入缩放和偏置自由度，当时间步数量较多时，可以有效加速注意力机制自身的收敛速度。

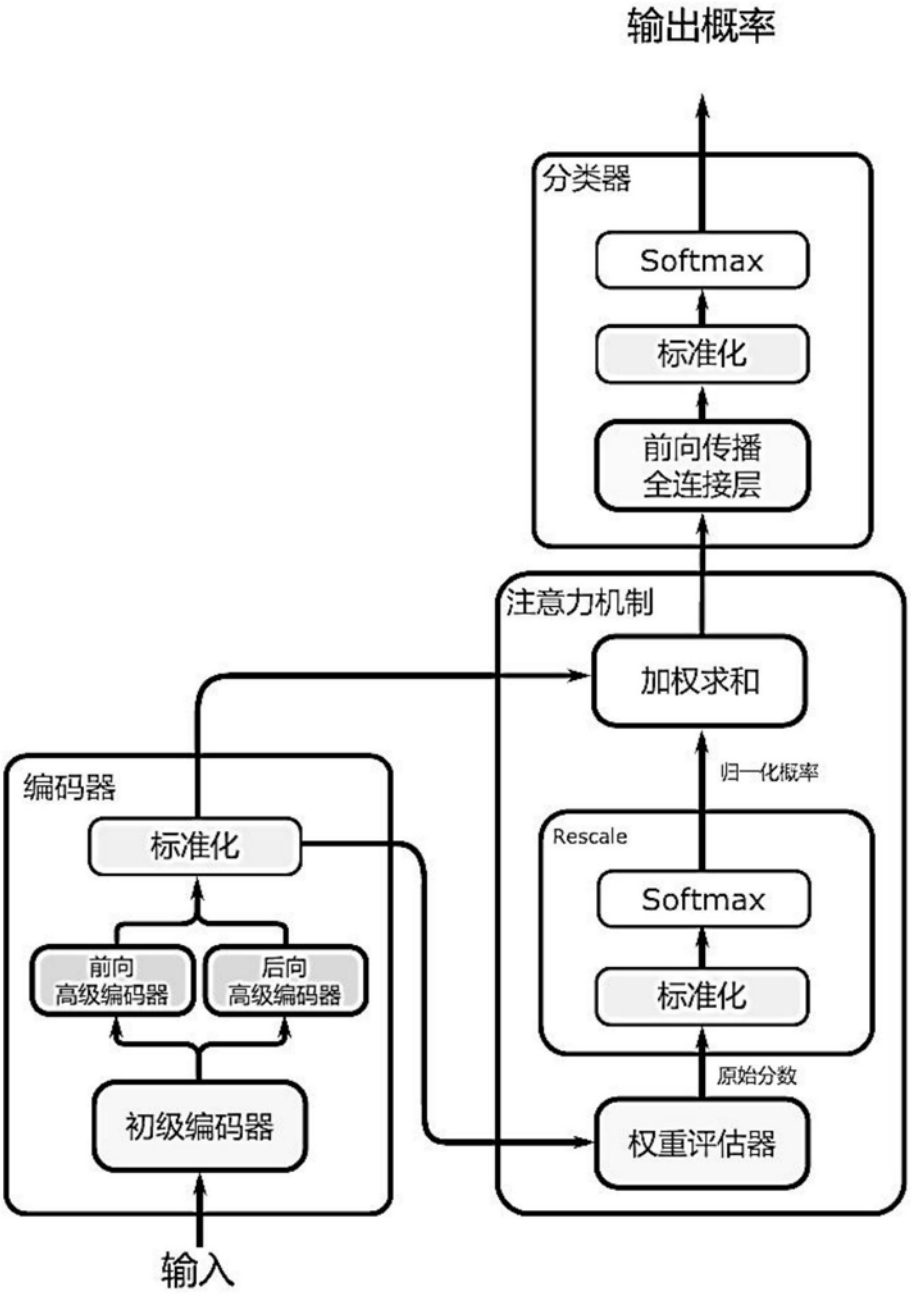
[0043] 经过具体实际实验，在Ubuntu 20.04，Python3.6，Keras2.2.4，Tensorflow 1.13.1的环境设置下，以Keras库中的默认随机初始化网络参数构建网络，使用批大小为8，学习率为0.0001，优化函数为Adam开始训练，以公开数据集ASCAD为攻击目标，能够得到以下实验数据。

[0044] 如图5所示，本发明构建的神经网络，在攻击对齐后的ASCAD数据集时，可实现7条攻击曲线恢复正确的密钥(猜测熵将至0)。

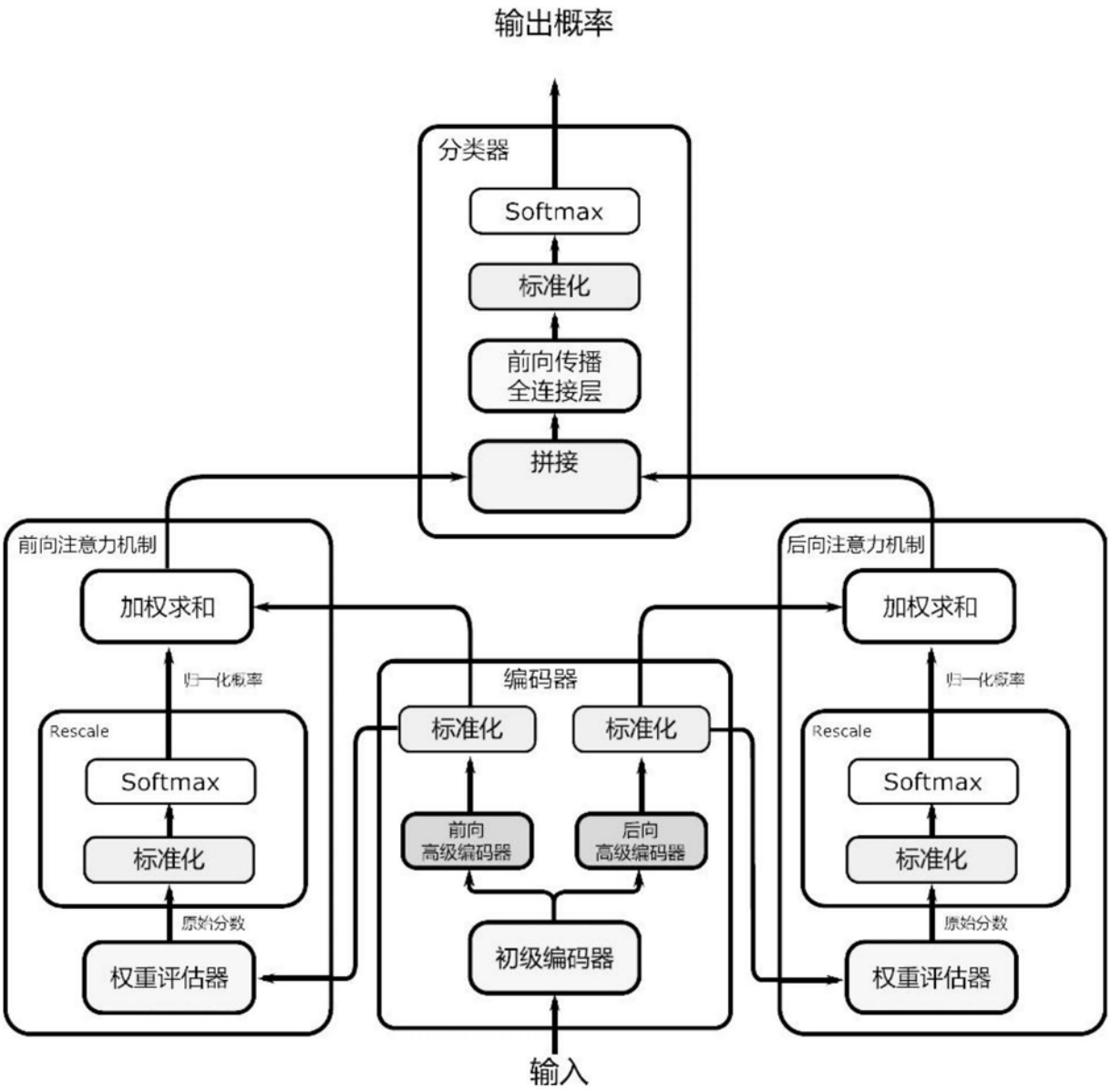
[0045] 如图6所示，本发明构建的神经网络，在攻击未对齐的ASCAD数据集时(进行了额外的随机平移作为数据增强，数据增强移动区间长度为80个时间点)，可实现20条攻击曲线恢复正确的密钥(猜测熵将至0)。

[0046] 如图7所示，为本发明的攻击结果和现有ZBHV20技术(Gabriel Zaid,Lilian Bossuet,AmauryHabrard,andAlexandreVenelli.Methodologyforefficient CNN architectures in profilingattacks.IACRTrans.Cryptogr.Hardw.Embed.Syst.,2020(1):1-36,2020.)的比较，相比之下现有技术需要选取700个特征点，本发明方法直接攻击原始曲线共10万点。

[0047] 上述具体实施可由本领域技术人员在不背离本发明原理和宗旨的前提下以不同的方式对其进行局部调整，本发明的保护范围以权利要求书为准且不由上述具体实施所限，在其范围内的各个实现方案均受本发明之约束。


输出概率
分类器
Softmax
标准化
前向传播
全连接层
注意力机制
加权求和
归一化概率
Rescale
Softmax
标准化
原始分数
权重评估器
编码器
标准化
前向
高级编码器
后向
高级编码器
初级编码器
输入

(a)


输出概率
分类器
Softmax
标准化
前向传播
全连接层
拼接
前向注意力机制
加权求和
归一化概率
Rescale
Softmax
标准化
原始分数
权重评估器
编码器
标准化
标准化
前向
高级编码器
后向
高级编码器
初级编码器
输入
后向注意力机制
加权求和
归一化概率
Rescale
Softmax
标准化
原始分数
权重评估器

(b)

图1

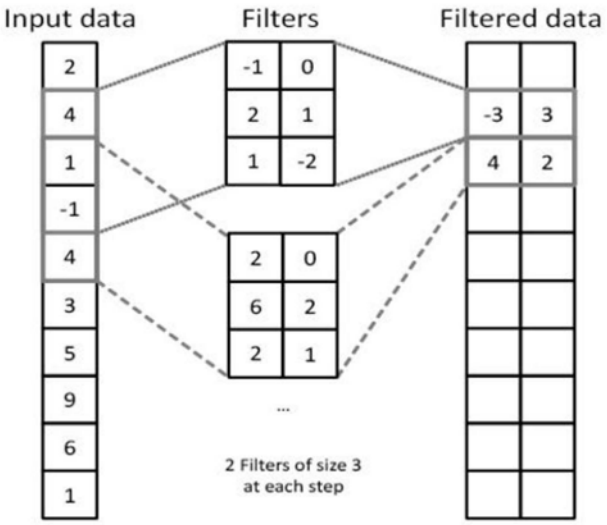

Input data
Filters
Filtered data
2
4
1
-1
4
3
5
9
6
1
-1 0
2 1
1 -2
2 0
6 2
2 1
...
2 Filters of size 3
at each step
-3 3
4 2

图2

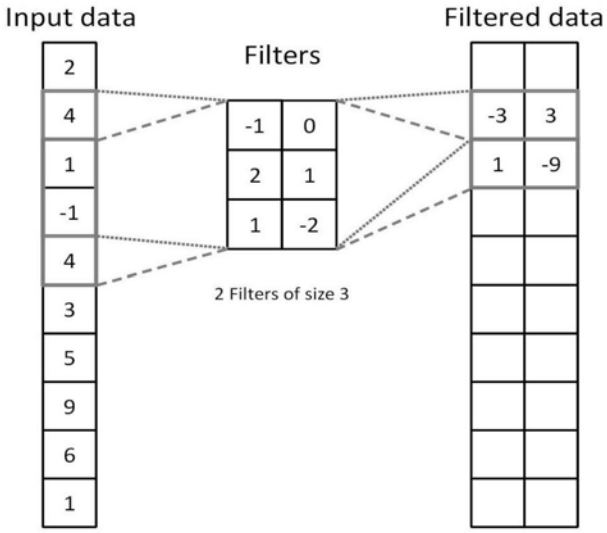

Input data
Filters
Filtered data
2
4
1
-1
4
3
5
9
6
1
-1 0
2 1
1 -2
2 Filters of size 3
-3 3
1 -9

图3

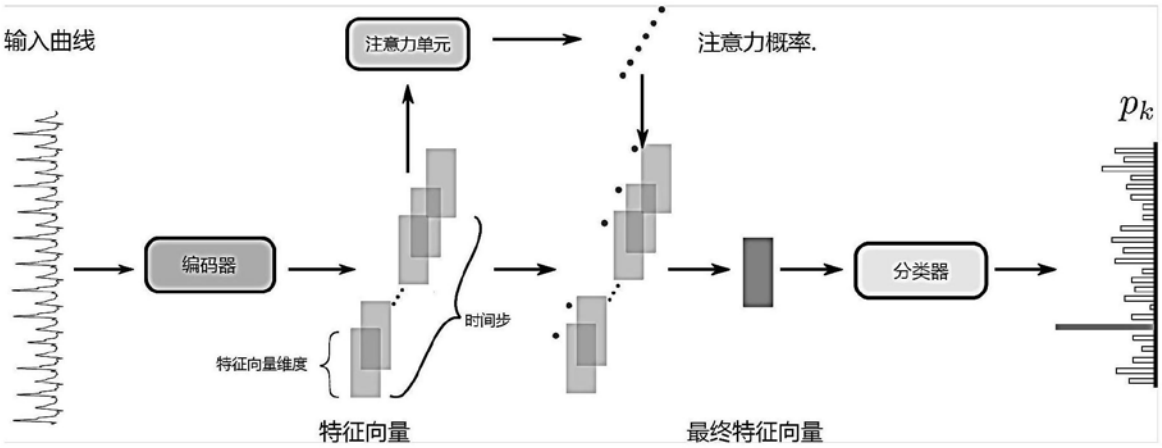

输入曲线
注意力单元
注意力概率.
p_k
编码器
分类器
时间步
特征向量维度
特征向量
最终特征向量

图4

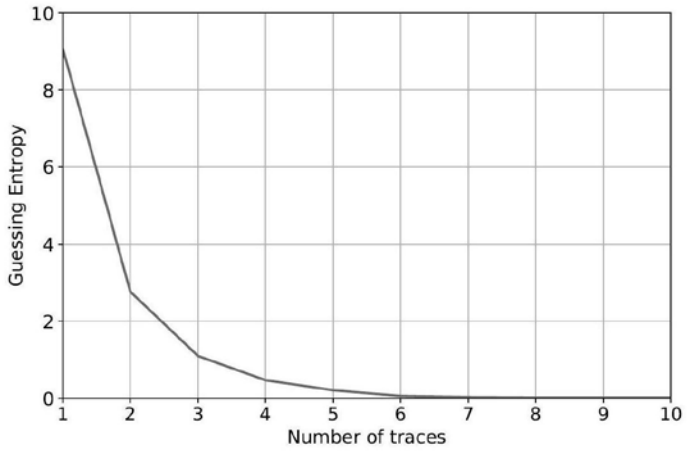

Guessing Entropy
Number of traces
0
2
4
6
8
10
1
2
3
4
5
6
7
8
9
10

图5

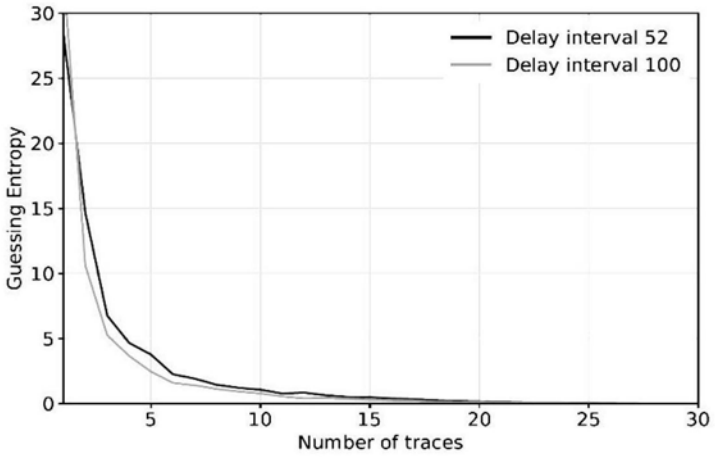

Delay interval 52
Delay interval 100
Guessing Entropy
Number of traces
0
5
10
15
20
25
30
5
10
15
20
25
30

图6

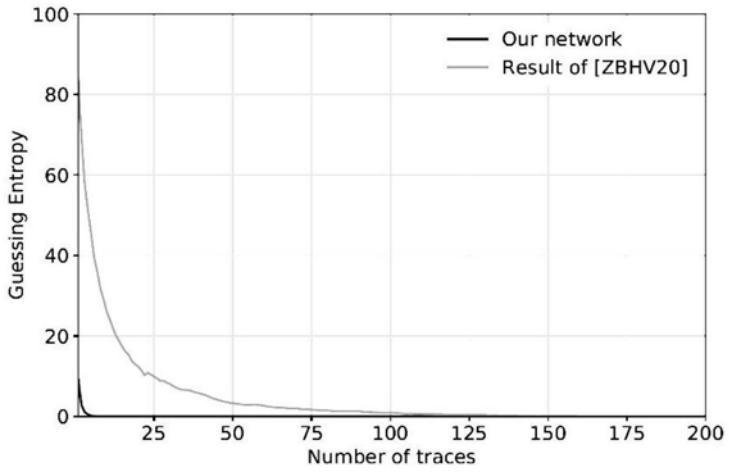

Our network
Result of [ZBHV20]
Guessing Entropy
Number of traces
0
20
40
60
80
100
25
50
75
100
125
150
175
200

图7