计算机研究与发展
Journal of Computer Research and Development
DOI:10.7544/issn1000-1239.2021.20200900
58(5): 1056-1064, 2021

基于深度学习的 SIMON32/64 安全性分析

王慧娇 丛 鹏 蒋 华 韦永壮
(广西可信软件重点实验室(桂林电子科技大学) 广西桂林 541004)
(whj@guet.edu.cn)

Security Analysis of SIMON32/64 Based on Deep Learning

Wang Huijiao, Cong Peng, Jiang Hua, and Wei Yongzhuang
(*Guangxi Key Laboratory of Trusted Software* (*Guilin University of Electronic Technology*), *Guilin*, *Guangxi* 541004)

Abstract With the rapid development of the Internet of Things, lightweight block cipher provides a solid foundation for the data security in various resource constrained environments. Currently, the security analysis of lightweight block ciphers tends to be more and more automated and intelligent. Applying deep learning to analyze the security of lightweight block ciphers appears to be a new research hotspot in this area. In this paper, the neural network technology is used to the security analysis of SIMON32/64, a lightweight block cipher algorithm released by the National Security Agency (NSA) in 2013. The feedforward neural network and the convolutional neural network are used to simulate the case of single input differential to multi output differential in multi differential cryptanalysis. Some deep learning distinguishers of 6-round (or even 9-round) reduced SIMON32/64 are designed, and both the advantages and disadvantages of the two neural network structures under different conditions are investigated. A candidate key sieving method for the 9-round reduced SIMON32/64 is also presented by extending the 7-round distinguisher of the feed-forward and the convolution neural networks, where one round forward and one round backward of this 7-round distinguisher are respectively considered. The experimental results show that 65535 candidate keys were dramatically reduced to 675 by only using 128 chosen plaintext pairs. Compared with the traditional differential distinguishers of reduced SIMON32/64, the new distinguishers combined with deep learning notably reduce both the time complexity and data complexity.

Key words SIMON32/64; deep learning; differential cryptanalysis; distinguisher; candidate key sieving

摘 要 轻量级分组密码的安全性分析越来越倾于向自动化和智能化的方向发展.目前基于深度学习对轻量级分组密码进行安全性分析正在成为一个全新的研究热点.针对由美国国家安全局在 2013 年发布的一款轻量级分组密码 SIMON 算法,将深度学习技术应用于 SIMON32/64 的安全性分析.分别采用前馈神经网络和卷积神经网络模拟多差分密码分析当中的单输入差分-多输出差分情形,设计了应用于

收稿日期:2020-11-06;**修回日期**:2020-12-23
基金项目:广西自然科学基金项目(2019GXNSFGA245004);广西科技重大专项项目(桂科 AA18118025);广西可信软件重点实验室项目(KX202056)
This work was supported by the Guangxi Natural Science Foundation (2019GXNSFGA245004), the Science and Technology Major Project of Guangxi (Guike AA18118025), and the Project of Guangxi Key Laboratory of Trusted Software (KX202056).
通信作者:韦永壮(walker_wyz@guet.edu.cn)

SIMON32/64 的 6～9 轮深度学习区分器,并比较了 2 种神经网络结构在不同条件下的优劣.通过对前馈神经网络和卷积神经网络的 7 轮深度学习区分器向前向后各扩展 1 轮,提出了针对 9 轮 SIMON32/64 的候选密钥筛选方法.实验结果证实:采用 128 个选择明文对,可以成功地将 65 535 个候选密钥筛选在 675 个以内.这说明基于深度学习的差分区分器相比传统差分区分器需要更少的时间复杂度和数据复杂度.

关键词 SIMON32/64;深度学习;差分密码分析;区分器;候选密钥筛选

中图法分类号 TP309.7

近年来,伴随无线传感器网络以及射频识别技术的发展和广泛应用,对资源受限的设备进行数据加密需要使用轻量级的密码算法.然而轻量级的密码算法追求低消耗与高效率,必然会导致安全性的降低,针对这一问题目前还没有公认的设计准则与安全标准,因此有必要对这些轻量级算法的安全性进行分析.

一个轻量级分组密码的分析关键在于构造一个有效的区分器,即根据密码算法的结构或其组件的特征将原密码算法和随机置换进行区分.差分区分器是迄今已知的攻击分组密码最有利的工具之一,由 Biham 等人[1]首次提出,该工具利用了分组密码算法在迭代过程中存在不平衡的差分统计量分布.轻量级分组密码在设计之初会充分考虑抵御差分密码分析和线性密码分析.如何发现未知的算法缺陷,设计新型区分器是一个新的研究方向.人工智能的飞速发展与深度学习技术的进步息息相关,在计算机视觉[2]、自然语言处理[3]、生物信息[4]等领域深度学习具有广泛的应用.深度学习是一种从数据当中发现复杂规律,并且利用规律对未来时刻、未知状况进行预测和判定的方法,因此深度学习在解决密码学问题上具有潜在的优势.

就密码学而言,深度学习主要应用于侧信道分析[5],对深度学习技术在经典密码分析的适用性上没有进行过多探讨.Abadi 等人[6]认为神经网络通常意味着并不擅长密码学,简单的感知机甚至无法对异或运算进行区分,而这恰恰是许多密码算法的基础.Rivest[7]则评价了深度学习与密码学之间的多种联系,提出了使用深度学习应用于密码分析中的一些可能的研究方向.Hu 等人[8]开发了一种前馈神经网络(feedforward neural network, FNN),该网络可以从 AES 密码的密文中发现明文,而无需使用密钥信息.Gohr[9]在 2019 年美密会上提出了基于深度学习的减轮 SPECK32/64 改进差分攻击,通过训练卷积残差神经网络用以区分固定明文差分加密的密文对和随机数据并以此构造区分器,证明了深度学习在对称密码上攻击的有效性和研究方向.Baksi 等人[10]延续 Gohr 的工作并借鉴文献[11]的思想利用神经网络构造多合一差分区分器应用于非马尔可夫密码 GIMLI. Yadav 等人[12]将差分区分器和深度学习技术结合,对经典差分区分器设计了基于深度学习的通用扩展,使区分器具有更多的轮数且需要更少的数据复杂度.Bellini 等人[13]采用多层感知机和卷积神经网络(convolutional neural network, CNN)来构造减轮 TEA 和 RAIDEN 算法的神经网络区分器,并提出了传统区分器无法应用的场合以及该方法的局限性.Jain 等人[14]对轻量级密码算法 PRESENT 构造 3～6 的神经网络区分器,该区分器能够将密码数据和随机数据以极高的概率区分开,进一步拓展了深度学习在分组密码的工作.So[15]尝试利用深度学习对简化版本的 DES,SPECK 和 SIMON 在基于密钥空间受限下的全轮攻击,成功发现了明/密文对和密钥之间的线性近似,但在密钥空间不受限制的条件下,该方法并不适用.

2013 年轻量级分组密码 SIMON[16]一经发布,便受到了广泛的关注.最初由 Abed 等人[17]采用传统差分分析获得了 SIMON 算法在各种分组长度下的高概率差分特征.Biryukov 等人[18]在 FSE'14 对自动化搜索算法寻找差分路径进行改进,提出了基于 ARX(addition rotation exclusive-or)密码的阈值搜索技术.该方法能够更有效地利用密码算法的强差分效应,进一步改善了 SIMON 类算法的最佳差分.Kölbl 等人[19]在 2015 年美密会上提出关于 SIMON 系列密码的差分和线性分布特征,采用自动化搜索技术 SAT/STM 寻找最佳的差分和线性特征,并研究了在不同轮数下 SIMON 算法抵抗攻击的能力.但是这类方法在构造区分器的过程中数据复杂度和时间复杂度相对较高.如何针对 SIMON 算法构建更加智能化、便捷化的深度学习区分器有待进一步去解决.

本文借鉴差分密码分析的攻击思想,通过深度学习发现轻量级分组密码在迭代过程中存在的差分不均匀性,提出了基于深度学习的减轮 SIMON32/64 的区分器模型,该区分器可以将密码算法和随机数据以极高的概率区分开.这是对 SIMON 算法安全性分析的全新思路,相较于传统的区分器具有更强的通用性和可实现性.该方案采用 FNN 和 CNN 深度学习算法来构造区分器,讨论了在相同环境下 2 种模型各自具有的优势.进一步通过对 FNN 和 CNN 区分器进行结合构造混合区分器,在保持模型的相对精度下,增加个体间的差异,提高区分器的泛化能力.通过实验结果证明,混合区分器在进行侯选密钥的筛选任务时能够以相对较高的概率将真实子密钥的范围确定在一个三者最小的集合中.与文献[18-19]相比,本文提出的深度学习区分器具有较好的性能,详细的 9 轮攻击复杂度对比如表 1 所示:

Table 1 Comparison of 9 Rounds of Attack Complexity for SIMON32/64

表 1 SIMON32/64 的 9 轮攻击复杂度对比

构造区分器算法	数据复杂度	时间复杂度	存储复杂度
文献[18]算法	2^{26}	$2^{42.000}$	$2^{26.000}$
文献[19]算法	2^{18}	$2^{34.000}$	$2^{18.000}$
FNN	$2^{7}+1$	$2^{23.000}$	$2^{10.115}$
CNN	$2^{7}+1$	$2^{23.000}$	$2^{9.760}$
FNN+CNN	$2^{7}+1$	$2^{23.358}$	$2^{9.399}$

1 预备知识

1.1 SIMON 算法

轻量级分组密码 SIMON[16] 基于平衡 Feistel 结构,其设计上只使用了简单的与运算、异或运算、移位运算,特别适用于物联网设备当中.其可选择的分组长度为 $2n$,其中,n 取值为 16,24,32,48,64,主密钥长度为 $n\times m$,其中,m 取值为 2,3,4,设计者基于不同 n 和 m 以及迭代轮数给出了 SIMON 算法的多个版本.其非线性变换为

$$F(x)=(x\ll 1)\ \&\ (x\ll 8)\oplus(x\ll 2). \tag{1}$$

假设其输入为(L_i,R_i),子密钥为 k_i,经过 1 轮加密后其输出为

$$\begin{aligned} L_{i+1}&=F(L_i)\oplus R_i\oplus k_i,\\ R_{i+1}&=L_i. \end{aligned} \tag{2}$$

图 1 为 SIMON 的 1 轮迭代变换结构,其中$\ll$为左移运算符、& 为与运算符、$\oplus$为异或运算符.

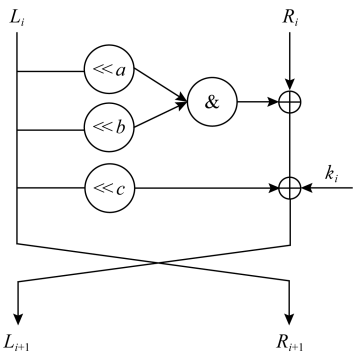


Fig. 1 Process of SIMON single round encryption

图 1 SIMON 单轮加密过程

1.2 前馈神经网络与卷积神经网络

FNN 和 CNN 在设计上具有相似性同时也存在差异,因此对数据集进行处理的过程中会有不同的敏感偏好.

一个 FNN 区分器由输入层、隐藏层和输出层构成,如图 2 所示,其中 C_i^j 表示第 i 个密文的第 j 个比特位.通过增加隐藏层的数量,FNN 能够很好地学习一些非线性不可分的任务,但相对而言一个深层的 FNN 容易出现过拟合现象.

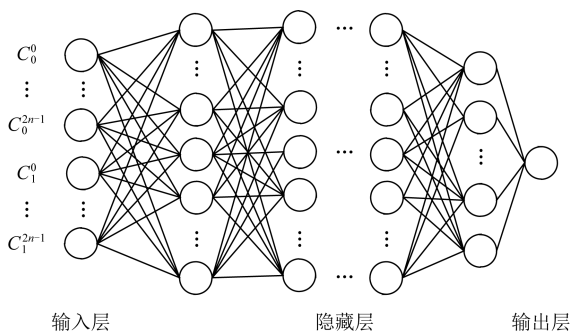


Fig. 2 Model of feedforward neural network distinguisher

图 2 前馈神经网络区分器模型

CNN 是一种特殊结构的 FNN,能够接受矩阵作为输入,具有跨空间(图像等)或跨时间(音频信号等)的重复神经元块(卷积核),这种特殊的设计结构使得 CNN 具有部分平移不变性.图 3 给出了 CNN

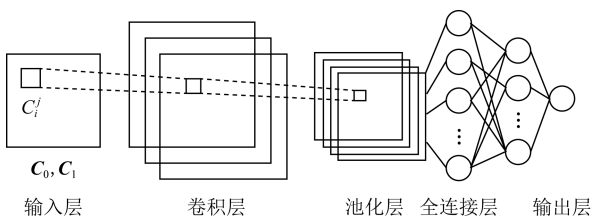


Fig. 3 Model of convolutional neural network distinguisher

图 3 卷积神经网络区分器模型

区分器的模型,其中输入层可由密文对($\boldsymbol{C}_0$,$\boldsymbol{C}_1$)灵活设置其矩阵的尺寸.

以截断差分区分器为例:算法只考虑差分的一部分性质,比如,差分落在某个集合中以及差分的某个比特位为 0.对于不同的数据集这些差分集合以及其中的某比特可以落在空间中的任意位置,假如要像 FNN 一样在每个空间位置学习独立的权重,训练数据需要多出几个数量级.简单地说,对于没有跨空间重复权重的 FNN,连接到输入矩阵左上方的输入神经元组不得不独立于连接到输入矩阵右下方的输入神经元组来学习表示某个差分集合或某个比特位.因此需要足够多的训练样本,才能够让 FNN 学习到某些差分集合在各个可能位置.

2 基于深度学习的区分器设计

本节采用 1.2 节中的 2 种深度学习模型来构建深度学习区分器,给出了深度学习区分器的构建方法和实现过程,以及如何利用该区分器进行候选密钥筛选.

2.1 区分器的设计

轻量级分组密码在迭代轮数低的情况下,明文中的统计规律和结构特征没有完全隐藏在密文当中,本文期望通过深度学习发现这种隐含的特征,从而将密文对和随机数据区分开,进而筛选出具有高概率差分特点并且隐含了密钥信息的密文对.对此本文选取了 2 种网络结构模型:FNN 和 CNN,通过灵活调节超参数,使其达到最优的性能.一方面是为了验证 2 种神经网络结构对于加密数据的敏感度;另一方面通过结合 2 种神经网络区分器进行候选密钥筛选,得到更好的结果.

深度学习技术能够揭示数据中的隐藏结构而无需显式标注其特征,借鉴差分密码分析中寻找多条差分特征的思想将其转化为深度学习的分类问题,采用 2 阶段策略构造深度学习区分器:

1) 离线阶段

离线阶段是对神经网络进行 2 分类学习,训练集与验证集由密文对和随机数据组成.密文对来自于固定明文差分通过不同密钥加密 r 轮生成.

2) 在线阶段

在线阶段是通过神经网络进行预测.测试集来自于固定明文差分采用统一密钥进行加密 r 轮生成的密文对.确定区分器阈值 δ,当测试样本预测值大于 δ 时将其归类为正例,否则归类为负例.

用于构造区分器数据集的过程定义为:假设一个长度为 $2n$ 比特的分组,经过 r 轮迭代得到分组密码算法 E.存在 2 个明文向量($\boldsymbol{P}_0$,$\boldsymbol{P}_1$)和 1 个密钥向量 $\boldsymbol{K}$,其 $\|\boldsymbol{P}_0\|$,$\|\boldsymbol{P}_1\|$,$\|\boldsymbol{K}\|$ 都为 l,且 $P_0^i \oplus P_1^i = \Delta_x$,对应的密文向量为 $\boldsymbol{C}_0 = E(\boldsymbol{P}_0, k)$ 和 $\boldsymbol{C}_1 = E(\boldsymbol{P}_1, k)$,其中当密文来自于训练集时 k 是 $\boldsymbol{K}$ 生成的轮密钥,当密文来自测试集时 k 是同一个密钥下的轮密钥,由此生成 l 个密文对(C_0^i,C_1^i),其中 $i \in \{1,2,\cdots,l\}$.算法 1,2 分别给出了深度学习区分器 2 阶段的构造过程.

算法 1. 离线阶段.

输入:样本数量 n、明文差分 Δ_x;

输出:训练成功的网络模型 $model$.

① $TD \leftarrow \varnothing$; /* 初始训练集为空 */

② $\boldsymbol{P}_0, \boldsymbol{K} \leftarrow Random$; $\boldsymbol{P}_1 \leftarrow \boldsymbol{P}_0 \oplus \Delta_x$;

③ $\boldsymbol{C}_0 \leftarrow encrypt(\boldsymbol{P}_0, k)$; $\boldsymbol{C}_1 \leftarrow encrypt(\boldsymbol{P}_1, k)$;

④ for $i \in \{1, n\}$ do /* 设置样本集标签 */

⑤ if $P_0^i \& 1 = 0$ then

⑥ $C^i \leftarrow Random$;

⑦ $Y^i \leftarrow 0$;

⑧ else

⑨ $Y^i \leftarrow 1$;

⑩ end if

⑪ end for

⑫ $TD \leftarrow (\boldsymbol{X}(\boldsymbol{C}_0, \boldsymbol{C}_1), \boldsymbol{Y})$;

⑬ $model \leftarrow trainNetwork(TD)$.

算法 2. 在线阶段.

输入:样本数量 n、明文差分 Δ_x、阈值 δ、模型 $model$;

输出:隐藏密钥信息且具有高概率差分密文对.

① $TD' \leftarrow \varnothing$; /* 初始测试集为空 */

② $\boldsymbol{P}_0 \leftarrow Random$; $\boldsymbol{P}_1 \leftarrow \boldsymbol{P}_0 \oplus \Delta_x$;

③ $\boldsymbol{C}_0 \leftarrow encrypt(\boldsymbol{P}_0, k)$; /* 唯一密钥加密 */

④ $\boldsymbol{C}_1 \leftarrow encrypt(\boldsymbol{P}_1, k)$;

⑤ $TD' \leftarrow \boldsymbol{X}(\boldsymbol{C}_0, \boldsymbol{C}_1)$;

⑥ for $i \in \{1, n\}$ do

⑦ if $predict(TD_i') > \delta$ then /* $model$ 预测 */

⑧ $save(TD_i')$;

⑨ else

⑩ $drop(TD_i')$;

⑪ end if

⑫ end for

2.2 神经网络的设置

为了保留密文对的统计信息以及神经网络模型能够顺利收敛到一个局部最优解,分别对输入输出格式、激活函数、部分超参数进行设置.

1) 输入输出格式:输入层由密文对$(\boldsymbol{C}_0,\boldsymbol{C}_1)$组成,其中$C_i^j$中$i\in\{0,1\}$,$j\in\{0,1,\cdots,2n-1\}$.FNN只能接收1维向量,故需要将密文对进行顺序排列,输入层的神经元数量为$2(n-1)$.CNN可以选择矩阵进行输入,本文灵活地构造SIMON32/64的2种输入数据格式2×32和4×16,分别代表密文对$(\boldsymbol{C}_0,\boldsymbol{C}_1)$的顺序排列和面向字节的结构.

2) 基于不同的数据集和应用场景神经网络在训练过程中采取不同的非线性激活函数:*sigmoid*函数是1个*logistic*函数,表示为不管输入是什么得到的输出被约束在0~1之间.本文中的神经网络模型只有1个输出神经元,因此将其应用于输出层,其函数表达式为

$$sigmoid(x)=\frac{1}{1+e^{-x}}. \tag{3}$$

*ReLU*函数应用于轮数比较低的区分器训练当中,减少了梯度消失的可能性,其函数表达式为

$$ReLU(x)=\max(0,x). \tag{4}$$

当对轮数比较高的区分器进行训练时,由于数据集当中的密文对和随机数据没有出现明显的差异性,因此往往在训练过程中产生梯度爆炸的情况,训练模型存在向2个方向倾斜的可能,基于此本文选择了tanh函数,其函数表达式为

$$\tanh(x)=\frac{\sinh(x)}{\cosh(x)}=\frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}. \tag{5}$$

3) 超参数设置.损失函数设置为其均方误差

$$L(y;f(x))=\frac{1}{n}\sum_{i=1}^{n}(y_i-f(x_i))^2. \tag{6}$$

同时为了防止出现过拟合现象,对均方误差采用$L2$范数最小化结构风险.每批次(batch)训练数据为2^{13},优化器采用Adam算法,学习率随训练轮数不断下降,经过20个回合(epoch)后不再降低,所以,算法中训练周期设置为20个epoch.其间通过回调函数触发ModelCheckPoint方法保存最佳的学习模型.

2.3 候选密钥筛选方案

采用上述算法获得了一个$r-1$轮深度学习区分器,对一个r轮加密算法,其攻击步骤为:

1) 均匀随机的选取明文$\boldsymbol{P}_0$,令$\boldsymbol{P}_1=\boldsymbol{P}_0\oplus\boldsymbol{\Delta}_x$,在同一密钥$k(k\in k_1,k_2\cdots,k_r)$下加密,获得相应密文对$(\boldsymbol{C}_0,\boldsymbol{C}_1)$.

2) 对$(\boldsymbol{C}_0,\boldsymbol{C}_1)$利用其$r$轮中所有存在的候选密钥进行1轮"解密",得到其$r-1$轮密文对$(\boldsymbol{\sigma}_0,\boldsymbol{\sigma}_1)$.

3) 将$(\boldsymbol{\sigma}_0,\boldsymbol{\sigma}_1)$分别通过由FNN和CNN构造的$r-1$轮深度学习区分器获得其对应输出$\delta_{\mathrm{FNN}}^k$,$\delta_{\mathrm{CNN}}^k$,计算输出的加权平均值为

$$\delta^k=\frac{\alpha\times\delta_{\mathrm{FNN}}^k+\beta\times\delta_{\mathrm{CNN}}^k}{\alpha+\beta}, \tag{7}$$

其中,α,β由FNN和CNN区分器的准确率分配.

4) 进一步统计每一个候选密钥来自于解密后的得分

$$v_k:=\sum_{i=1}^{n}\ln\frac{\delta_i^k}{1-\delta_i^k}, \tag{8}$$

通过v_k对所有候选密钥进行排名.算法3对步骤1)~4)进行了阐述.

算法3. 候选密钥筛选.

输入:选择明文差分加密r轮的密文对$(\boldsymbol{C}_0,\boldsymbol{C}_1)$、$r-1$轮深度学习区分器$model_{\mathrm{FNN}}$,$model_{\mathrm{CNN}}$;

输出:候选密钥排名.

① for $i\in\{1,n\}$ do /* n为候选子密钥数量 */

② $\boldsymbol{\sigma}_0\leftarrow decryptOneRound(\boldsymbol{C}_0,k^i)$;

③ $\boldsymbol{\sigma}_1\leftarrow decryptOneRound(\boldsymbol{C}_1,k^i)$;

④ $\delta_{\mathrm{FNN}}^k\leftarrow model_{\mathrm{FNN}}(\boldsymbol{\sigma}_0,\boldsymbol{\sigma}_1)$;

⑤ $\delta_{\mathrm{CNN}}^k\leftarrow model_{\mathrm{CNN}}(\boldsymbol{\sigma}_0,\boldsymbol{\sigma}_1)$;

⑥ $\delta^k\leftarrow weightedAverage(\delta_{\mathrm{FNN}}^k,\delta_{\mathrm{CNN}}^k)$;

⑦ $v_k\leftarrow finalGrade(\delta^k)$;

⑧ end for

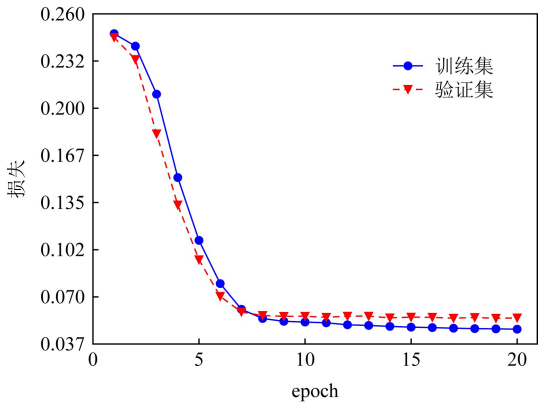
⑨ $descendingSort(v_k)$. /* 降序排列 */

3 实验结果与分析

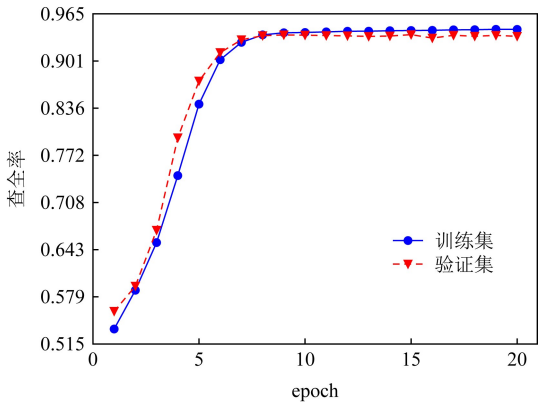
本节主要给出了深度学习区分器在训练阶段和测试阶段的实验结果,并展示了通过深度学习区分器在对SIMON32/64的9轮攻击中呈现的效果.所有实验平台其硬件环境为处理器:Intel® Core™ i7-7700,内存:8.00 GB.深度学习库:后端Tensorflow,前端Keras,采用CPU进行计算实现.

3.1 离线阶段实验

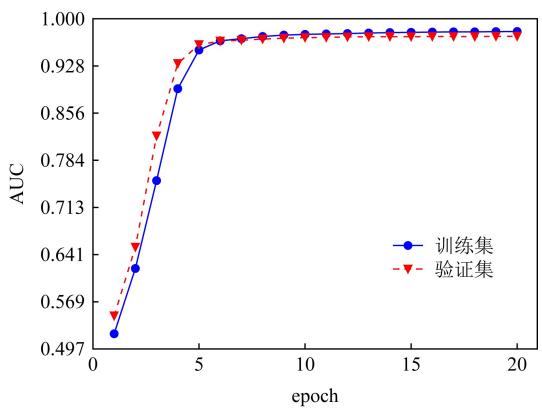
选择明文差分$\Delta_x=0x0040/0000$,通过第2节的算法1生成2^{20}个训练集样本,2^{17}个验证集样本,正负样本的数量各占其中的1/2.图4给出了基于6轮FNN在20个回合的训练情况:



(a) 损失曲线



(b) 查全率曲线



(c) AUC曲线

Fig. 4 Performance measurement of SIMON 6-round neural network differentiator

图 4 SIMON 的 6 轮神经网络区分器的性能度量

损失曲线能够反映模型的学习率变化和函数的收敛情况,在图 4(a)中,由于初始时学习率设置较大,损失值下降很快.随着迭代学习率不断得到调整,模型的训练精度增加,损失曲线也逐渐趋于稳定.其后面 epoch 验证集的曲线变化有向低性能过渡的趋势,反映出模型存在轻微的过拟合现象.查全率表示被预测正确的正例样本占总正例样本的比例,衡量了区分器对正例的识别能力.其 20 个 epoch 的查全率变化如图 4(b)所示,反映该模型能够比较全面地选择出更多的密文对,从而能够寻找出更多的高概率差分.AUC(area under curve)反映了模型对样本的排序能力,它能够直观地展示一个模型的好坏.图 4(c)展示了在第 6 个 epoch 后 AUC 的值开始接近于 1,说明采用 Δ_x = 0x0040/0000 为初始差分经过 6 轮迭代后生成的密文对与随机数据之间存在较大的差别,神经网络能够轻易地提取到这种数据间差异.综上说明这是一个近乎完美的 6 轮区分器.

表 2 列出了 FNN 和 CNN 在 6～9 轮区分器的学习参数、训练时间和准确率.可以看出 CNN 能够更好的收敛到一个局部最小值,更容易被优化.但相应的由于 FNN 结构简单,训练时间将大大降低,准确率上略逊于 CNN.

Table 2 Comparison of SIMON32/64 6～9 Rounds FNN and CNN

表 2 SIMON32/64 的 6～9 轮 FNN 与 CNN 对比

区分器模型	训练参数个数	训练时间/s	准确率/%
FNN-6	11 489	71	96.76
CNN-6	77 025	3 211	99.83
FNN-7	11 489	69	78.06
CNN-7	83 489	5 362	93.60
FNN-8	11 489	71	62.34
CNN-8	83 489	5 667	67.36
FNN-9	11 489	70	53.03
CNN-9	83 489	5 240	56.70

图 5 展示了 6～9 轮区分器训练过程中验证集准确率的变化,表明深度学习区分器对低轮的 SIMON 算法轻松的学习到了加密数据和随机数据的区分,但随着其迭代轮数的增加,准确率会不断降低.其本质原因在于根据密码算法的混淆和扩散原则,其迭代轮数越高,明密文之间的统计信息越弱,相应的正负样本相似性很高,使深度学习难以进行有效的特征选择.为了使 8,9 轮的区分器有更强的泛化能力,通过增加训练集和验证集的数据量、延长训练 epoch,2 种神经网络模型都获得了不同程度的提升,CNN 相比于 FNN 提升效果更加明显.

3.2 在线阶段实验

对 2 种神经网络模型 6～9 轮进行测试,采用选择明文差分为 Δ_x = 0x0040/0000,定义阈值 δ > 0.5.

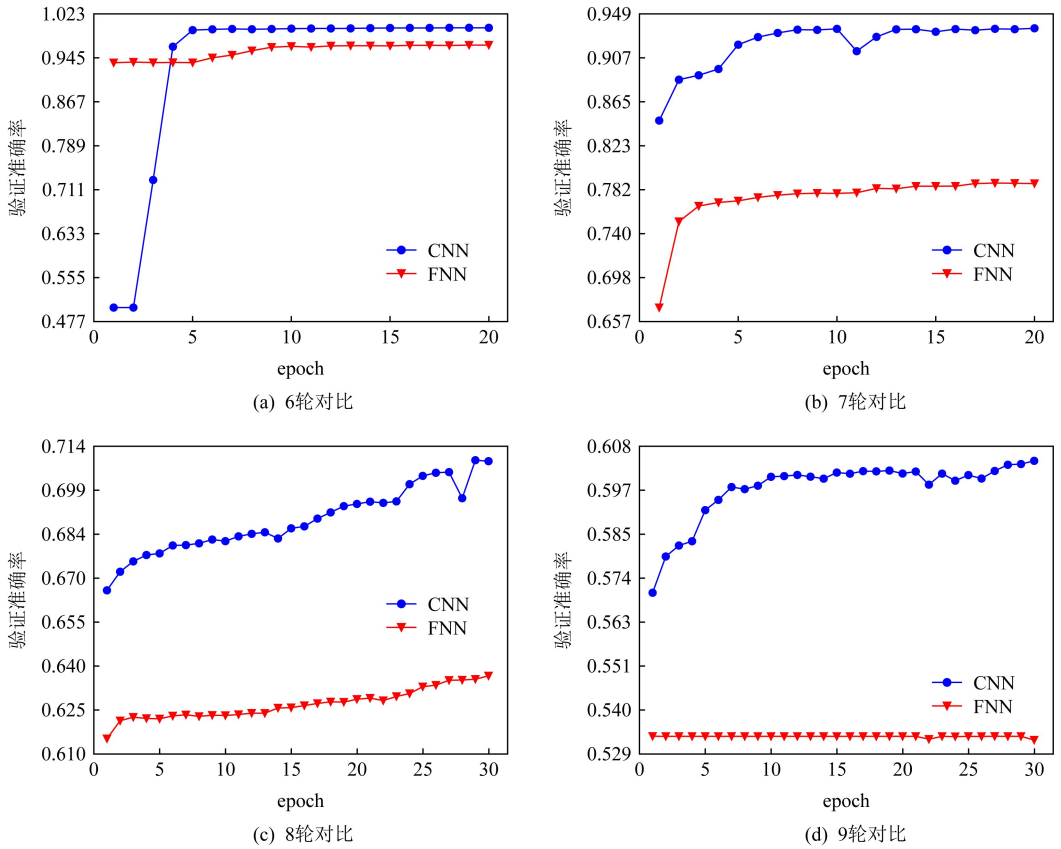


Fig. 5 Comparison of the accuracy in verification set of SIMON 6～9 rounds FNN and CNN

图 5 SIMON 的 6～9 轮 FNN 与 CNN 验证集准确率对比

为了防止可能存在的密钥依赖,每一轮采用 1 000 个密文对作为数据集,循环 1 000 次,总计 100 万个数据集.表 3 给出了 2 种神经网络模型 6～9 轮的预测结果,可以看出 2 种模型在 6,7 轮都具有可信的准确率,同时证明该模型具有良好的泛化能力.在 8,9 轮中,CNN 在预测上具有一定的参考意义,基于 9 轮的 FNN 几乎没有学到任何东西.图 6 给出了 6 轮 CNN 区分器对 300 个随机正样本的预测分布.可以看出我们的 6 轮 CNN 区分器在预测值的分布上具有较高的可信度,能够容易地找到具有高概率差分的密文对.

Table 3 Test Set Accuracy of Deep Learning Distinguisher

表 3 深度学习区分器测试集准度

迭代轮数	准确率	
	FNN	CNN
6	$0.997\pm9.71.10^{-4}$	$0.999\pm7.74.10^{-4}$
7	$0.815\pm2.32.10^{4}$	$0.983\pm3.56.10^{-4}$
8	$0.540\pm2.81.10^{-4}$	$0.602\pm9.98.10^{-4}$
9	$0.480\pm9.35.10^{-4}$	$0.564\pm3.74.10^{-4}$

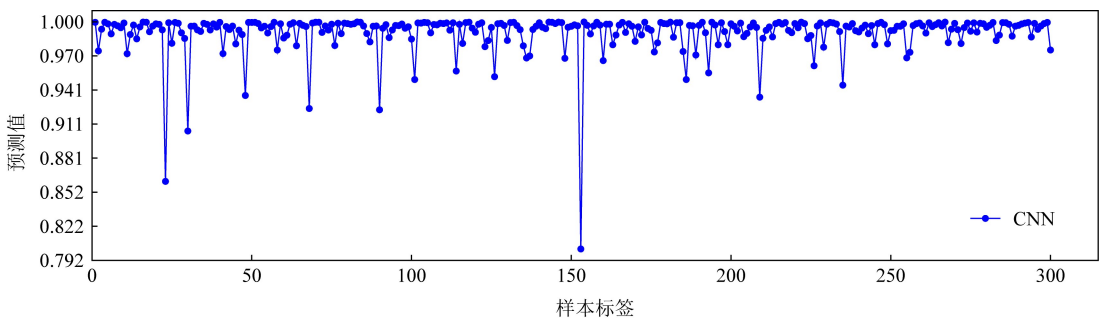


Fig. 6 Prediction results of SIMON 6 rounds of convolutional neural network

图 6 SIMON 的 6 轮卷积神经网络预测结果

3.3 候选密钥筛选结果

训练集包含 2^{20} 个样本,验证集包含 2^{17} 个样本,每批次采用 2^{13} 样本,训练 epoch 为 20 轮.模型初始学习率为 0.02,每轮递减 5E−4,经过 20 个 epoch 后学习率为 0.01.最终得到了基于 FNN 和 CNN 的最佳学习模型.故其每个区分器的离线复杂度为 $2^{20}+2^{17}=2^{20.170}$(加密次数).

利用算法 1 构造了一个 7 轮的深度学习区分器.采用 128 个输入差分为 $\Delta_x=$0x0040/0000 的已知明文对,基于 SIMON32/64 的自身性质,对其无损失的向上扩展 1 轮,最后通过真实密钥加密至 9 轮.

利用算法 3 对 9 轮 SIMON32/64 的 65,535 个候选密钥进行筛选.采用 2^7 个选择明密文对,该攻击的时间复杂度为 $2^{20.170}+2^7\times2^{16}=2^{23.000}$(加密次数),混合区分器的攻击时间复杂度为 $2^{20.170}\times2+2^7\times2^{16}=2^{23.358}$(加密次数).

通过对 FNN 和 CNN 以及两者结合的混合区分器进行实验,表 4 给出了以真实密钥为分界线,高于真实密钥成绩的候选密钥数量.筛选出的候选密钥数量越少,表明所设计的区分器越有效.候选密钥成绩的平均取值范围在−537.107～−500.221 之间.

Table 4 The Number of Candidate Keys and the Actual Key Scores

表 4 候选密钥数量和真实密钥成绩

区分器	真实密钥成绩	候选密钥数量	剩余候选密钥数量
FNN	−506.564±94	$2^{16.000}$	$2^{10.115}$
CNN	−503.927±66	$2^{16.000}$	$2^{9.760}$
FNN+CNN	−503.133±87	$2^{16.000}$	$2^{9.399}$

3.4 深度学习区分器与传统区分器的比较

采用 STA/STM 方法 Kölbl 等人[19]给出了减轮 SIMON32/64 的 16 轮差分轨迹,其中 6～9 轮的差分轨迹的转移概率分别为 2^{-12},2^{-14},2^{-18},2^{-20},相应所需的数据复杂度和存储复杂度至少分别需要 2^{12},2^{14},2^{18},2^{20}.采用深度学习构造的区分器在 9 轮攻击中数据复杂度仅为 2^7+1,FNN 和 CNN 以及两者结合的混合区分器存储复杂度分别为 $2^{10.115}$,$2^{9.760}$ 和 $2^{9.399}$.表 1 对文献[18-19]和深度学习区分器所需的数据复杂度、时间复杂度和存储复杂度进行了总结,可以反映出深度学习区分器在这些度量标准上都具有明显的优势,采用 FNN 和 CNN 结合的候选密钥筛选方案能够将存储复杂度有效降低.

4 总结与展望

本文分别基于 FNN 和 CNN 2 种深度学习模型对减轮 SIMON32/64 的安全性进行了分析.实验结果发现:CNN 构造的区分器在准确率上要优于 FNN;对于 6,7 轮的 SIMON32/64 该区分器都能够以很高的概率将密文对和随机数据区分开;在利用 FNN 和 CNN 区分器共同参与的候选密钥筛选策略中,不但成功缩减了候选密钥可能的范围,而且有效降低了攻击复杂度、数据复杂度及时间复杂度.这也进一步说明了通过结合神经网络和差分对一些轻量级分组密码进行安全性分析是一个可行的手段.另一方面,能否利用机器学习中的其他算法来构建分组密码的新型区分器,有待进一步研究.

参 考 文 献

[1] Biham E, Shamir A. Differential cryptanalysis of DES-like cryptosystems [J]. Journal of Cryptology, 1991, 4(1): 3-72

[2] Li Zhixin, Wei Haiyang, Huang Feicheng, et al. Combine visual features and scene semantics for image captioning [J]. Chinese Journal of Computers, 2020, 43(9): 1624-1640 (in Chinese)
(李志欣, 魏海洋, 黄飞成, 等. 结合视觉特征和场景语义的图像描述生成[J]. 计算机学报, 2020, 43(9): 1624-1640)

[3] Liu Ye, Huang Jinxiao, Ma Yutao. An automatic method using hybrid neural networks and attention mechanism for software bug triaging [J]. Journal of Computer Research and Development, 2020, 57(3): 461-473 (in Chinese)
(刘烨, 黄金筱, 马于涛. 基于混合神经网络和注意力机制的软件缺陷自动分派方法[J]. 计算机研究与发展, 2020, 57(3): 461-473)

[4] Esteva A, Robicquet A, Ramsundar B, et al. A guide to deep learning in healthcare [J]. Nature Medicine, 2019, 25(1): 24-29

[5] Das D, Golder A, Danial J, et al. X-DeepSCA: Cross-device deep learning side channel attack [C] //Proc of the 56th Annual Design Automation Conf. Piscataway, NJ: IEEE, 2019: 1-6

[6] Abadi M, Andersen D G. Learning to protect communications with adversarial neural cryptography [J]. arXiv preprint, arXiv:1610.06918, 2016

[7] Rivest R L. Cryptography and machine learning [C] //Proc of the Int Conf on the Theory and Application of Cryptology. Berlin: Springer, 1991: 427-439

[8] Hu Xinyi, Zhao Yaqun. Research on plaintext restoration of AES based on neural network [J/OL]. Security and Communication Networks, 2018 [2019-11-13]. https://doi.org/10.1155/2018/6868506

[9] Gohr A. Improving attacks on round-reduced SPECK32/64 using deep learning [C] //Proc of the Annual Int Cryptology Conf. Berlin: Springer, 2019: 150-179

[10] Baksi A, Breier J, Dong Xiaoyang. Machine learning assisted differential distinguishers for lightweight ciphers [DB/OL]. IACR Cryptology ePrint Archive. (2020-11-10) [2020-11-16]. https://eprint.iacr.org/2020/571.pdf

[11] Albrecht M R, Leander G. An all-in-one approach to differential cryptanalysis for small block ciphers [G] //LNCS 7707: Proc of the Int Conf on Selected Areas in Cryptography. Berlin: Springer, 2012: 1-15

[12] Yadav T, Kumar M. Differential-ML distinguisher: Machine learning based generic extension for differential cryptanalysis [DB/OL]. IACR Cryptology ePrint Archive. (2020-07-21) [2020-08-25]. https://eprint.iacr.org/2020/913.pdf

[13] Bellini E, Rossi M. Performance comparison between deep learning-based and conventional cryptographic distinguishers [DB/OL]. IACR Cryptology ePrint Archive. (2020-08-13) [2020-09-10]. https://eprint.iacr.org/2020/953.pdf

[14] Jain A, Kohli V, Mishra G. Deep learning based differential distinguisher for lightweight cipher PRESENT [DB/OL]. IACR Cryptology ePrint Archive. (2020-07-30) [2020-08-25]. https://eprint.iacr.org/2020/846.pdf

[15] So J. Deep learning-based cryptanalysis of lightweight block ciphers [OL]. 2020 [2020-08-20]. https://doi.org/10.1155/2020/3701067

[16] Beaulieu R, Shors D, Smith J, et al. The SIMON and SPECK families of lightweight block ciphers [DB/OL]. IACR Cryptology ePrint Archive. (2013-06-19) [2019-10-01]. http://eprint.iacr.org/2013/404.pdf

[17] Abed F, List E, Lucks S, et al. Differential cryptanalysis of round-reduced Simon and Speck [G] //LNCS 8540: Proc of the Int Workshop on Fast Software Encryption. Berlin: Springer, 2014: 525-545

[18] Biryukov A, Roy A, Velichkov V. Differential analysis of block ciphers SIMON and SPECK [G] //LNCS 8540: Proc of the Int Workshop on Fast Software Encryption. Berlin: Springer, 2015: 546-570

[19] Kölbl S, Leander G, Tiessen T. Observations on the SIMON block cipher family [C] //Proc of the Annual Cryptology Conf. Berlin: Springer, 2015: 161-185

Wang Huijiao, born in 1976. PhD, associate professor, master supervisor. Member of CCF. Her main research interests include cryptography and information security.

王慧娇, 1976 年生. 博士, 副教授, 硕士生导师, CCF 会员. 主要研究方向为密码学和信息安全.

Cong Peng, born in 1994. Master candidate. His main research interests include machine learning and information security.

丛 鹏, 1994 年生. 硕士研究生. 主要研究方向为机器学习与信息安全.

Jiang Hua, born in 1963. PhD, professor, master supervisor. His main research interests include information security and data mining.

蒋 华, 1963 年生. 博士, 教授, 硕士生导师. 主要研究方向为信息安全和数据挖掘.

Wei Yongzhuang, born in 1976. PhD, professor, PhD supervisor. His main research interests include cryptography and information security.

韦永壮, 1976 年生. 博士, 教授, 博士生导师. 主要研究方向为密码学和信息安全.