DOI:10.13878/j.cnki.jnuist.2022.03.013

许历隆[1]　翟江涛[1]　林鹏[1]　崔永富[1]

一种基于改进深度残差收缩网络的恶意应用检测方法

摘要

恶意应用的快速增长给移动智能终端带来了巨大的安全威胁,实现恶意应用高精度检测对移动网络信息安全具有重要意义.本文提出一种基于改进深度残差收缩网络的恶意应用检测方法.首先将流量特征预处理成卷积神经网络输入,接着引入通道注意力机制和空间注意力机制,从通道和空间两个维度对样本特征进行加权.然后再引入深度残差收缩网络,自适应滤除样本冗余特征并通过恒等连接优化参数反向传播,减小模型训练和分类的难度,最终实现安卓恶意应用高精度识别.所提方法可避免手工提取特征,能实现高精度分类并且具有一定泛化能力.实验结果表明,所提方法在恶意应用的2分类、4分类和42分类中准确率分别为99.40%、99.95%和97.33%,与现有方法相比,具有较高的分类性能与泛化能力.

关键词

恶意应用;恶意家族;深度残差收缩网络;信息安全

中图分类号 TP393

文献标志码 A

收稿日期 2021-07-04

资助项目 国家自然科学基金(U1836104,61772281,61801073,61931004,62072250);南京信息工程大学人才启动基金(2020r061)

作者简介

许历隆,男,硕士生,研究领域为多媒体与信息安全.2837053804@qq.com

翟江涛(通信作者),男,博士,副教授,研究领域为多媒体与信息安全.jiangtaozhai@gmail.com

1 南京信息工程大学 电子与信息工程学院,南京,210044

0 引言

2020年第三季度,360安全大脑共截获移动端新增恶意应用样本约118.7万个,平均每天截获新增手机恶意应用样本约1.3万个[1].智能终端感染恶意应用快速增长的趋势给移动智能终端的用户隐私、财产安全等方面带来巨大危害.因此,移动终端恶意应用检测和分类成为网络安全研究的热点问题.

恶意应用的分类检测除了需给出待测样本是否具有恶意性之外,对其所属家族的判定同样具有重要意义.恶意应用的家族分类往往能够揭示恶意应用的恶意行为类别与执行目的[2],这也为恶意应用的危险程度等信息提供了重要的参考.另外,恶意应用家族分类的检测有利于快速跟踪恶意应用家族发展,以便对网络空间安全形势进行快速评估.因此,实现恶意家族的超多分类亦十分必要.

早期学者采用基于签名匹配的方式识别Android恶意软件.通过收集恶意应用签名构建数据库,将待检测的样本与数据库里的签名进行匹配,从而判断应用是否具有恶意性[3-5].但是这类方法需要对数据库进行实时的更新和补充,否则无法识别新出现的恶意应用.而且,开发者可通过简单修改应用代码在不影响语义的情况下逃避检测,因此该方法具有较大的局限性.随着机器学习的兴起,通过采用机器学习识别恶意应用的方案得到了广泛的研究.目前研究者主要通过提取恶意软件中的相关特征,训练机器学习分类器,并用满足训练条件的分类器对恶意样本进行识别.根据特征的不同,通常可分为静态分析和动态分析.

静态分析是利用反编译工具,提取APK文件中权限[6]、API调用[7]、网络地址[8]、关键代码字段[9]等特征.由于APK文件基本是固定的,因而从APK中提取的特征不容易发生改变.静态分析方法的优势在于数据易采集、检测效率高,但是其存在信息维度较少且难以解决代码混淆的问题.

动态分析是指捕获软件在运行中产生的行为特征分析并训练分类模型,如系统调用序列[10]、内存利用率[11]等.近些年,在动态特征中网络流量特征引起了学者们广泛关注.恶意应用往往通过获取移动终端的权限监视用户的浏览与输入信息,并通过网络将隐私信息泄露给攻击者.2012年,Sarma等[12]对超过15万个应用程序进行了研究,发现93%的恶意程序需要网络访问.同年,Zhou等[13]指出,其收集的

Android 恶意样本里,93%的软件通过网络与攻击者的 C&C 服务器连接接收指令.因此,流量交互是恶意应用产生恶意行为的重要一环,而通过分析流量特征检测恶意应用是可行方案.

Lashkari 等[14]从真实网络环境中捕获网络流量,公开一个新的数据集 CICAndMal2017.在此基础上,该团队提取 80 个流级流量特征并采用信息增益(IG)和基于相关特征选择(CFS)算法选取 9 类特征组成最佳特征集,通过随机林(RF)、K 近邻(KNN)和决策树(DT)算法训练模型实现恶意应用的快速检测和分类,但是所提方法精度不高.Noorbehbahani 等[15]在文献[14]工作的基础上评估 7 种分类器对勒索软件下的 10 类恶意家族的分类性能,其中随机森林分类器取得了最高的分类结果,10 分类精度达 85%.Taheri 等[16]提出两层框架的恶意应用检测算法.第 1 层框架 SBC 中,从 APK 文件中提取出 8 115 种权限与意图特征训练随机森林分类器并实现恶意应用的 2 分类.然后将识别出的恶意应用样本传进第 2 层框架 DMC 中.在第 2 层检测框架中,提取并结合 API 特征和流级流量特征训练随机森林分类器实现恶意应用类型多分类和恶意应用家族的超多分类.该方法在恶意应用 2 分类上取得了 95.3%的精度,恶意应用多分类上达到了 83.3%的精度,但是在恶意家族的分类上并不能取得理想精度.Abuthawabeh 等[17]认为相比流级特征,提取会话级特征可以充分捕获到通信双方流量数据交互的行为,并且有利于避免恶意软件使用端口随机化技术带来的干扰.于是提取了会话级流量特征,集成学习技术投票出最优特征并用以训练极端随机树和随机森林分类器,提高了恶意应用类型多分类和恶意家族超多分类的精度,最高分别达到 80.2% 和 67.21%.Chen 等[18]使用随机森林、K 近邻和决策树 3 个分类器实现 2 个分类任务:恶意-良性应用的 2 分类、恶意应用类型的 3 分类(选取的恶意应用类型为广告软件、勒索软件和恐吓软件).实验结果表明,随机森林分类器取得的分类效果最好.但是研究人员并没用使用完整的数据集并且缺少恶意家族分类,使得其方法泛化能力较弱.Arora 等[19]通过 IG 和卡方检验算法对 Android 恶意应用的网络流量特征进行优先排序,然后最小化网络流量特征,提高检测精度,减少训练和测试阶段的时间.通过实验发现 22 个特征中有 9 个特征可以满足更高的检测精度.同样,它可以减少 50%模型训练时间和 30%测试阶段的时间.

以上通过人工提取流量特征的传统机器学习方法需要大量专家经验,特征选取的种类和数目直接影响恶意应用检测的准确率.且在不同的分类任务中,往往需要研究人员有针对地提取不同特征来提高模型分类性能,这使得这类方法具有低泛化性与高复杂性.

为了克服上述困难,基于深度学习自动学习样本特征的方法得到了广泛关注.文献[20]提出一种基于深度学习的端到端的恶意流量分类方法:首先将流量数据映射成灰度图像样本,然后利用样本训练卷积神经网络(Convolutional Neural Networks,CNN)模型,最终实现恶意流量的检测.文献[21]设计了一个基于深度学习方法检测恶意软件的 DeepMAL 模型,通过从原始网络流和原始数据包中自动提取字节流特征,自主训练模型,实现对恶意应用的 4 分类检测.实验结果表明,与传统机器学习方法相比,DeepMAL 有效解决了传统方法依赖先验知识设计特征的问题,并能以更低的虚警率达到更高的检测精度.文献[22]提出一个双层的检测模型来实现恶意应用的多场景分类.第 1 层通过提取权限、组件信息、意图 3 种静态特征并基于全连接神经网络将应用分为良性和恶意应用,并将检测出的恶意应用样本传入第 2 层检测系统;第 2 层通过将原始流量数据转换成灰度图像,利用 CNN 与卷积自编码器(Convolutional Auto-Encoders,CAE)级联方法 CACNN 从灰度图像中自动提取特征并实现恶意应用类型的多分类和恶意家族的超多分类.该方法在恶意-良性应用 2 分类、恶意应用类型多分类 2 个场景中分别取得 99%与 98%的精度,但是在恶意家族超多分类中,精度只有 73%.基于深度学习的方法有效地解决了特征选取问题并且在检测精度上有所提高,但是针对恶意家族超多分类,上述深度学习的方法和传统机器学习方法往往都不能取得较高的分类精度.

在恶意应用大类下进行家族的多分类,由于子类样本之间区别更加细微,这使得模型需要更强的特征提取能力.人工提取特征较为依赖专家知识,而利用深度学习模型自动学习提取样本特征,有时并不能捕获样本中细粒度的特征.同时,恶意家族超多分类需要更多的特征种类和特征数目,其产生的冗余特征大大增加了传统机器学习与深度学习模型的分类难度.基于此,本文提出一种改进的残差收缩网

络方法,所提方法利用神经网络从原始流数据中学习特征表达,避免特征的人工设计带来的复杂性以及低泛化性.通过注意力机制与残差收缩网络,提取区分相似样本的细粒度特征,自适应滤除样本的噪声与冗余特征,进一步提升分类精度与在不同场景中的泛化性.本文贡献如下:

1)引入深度残差收缩网络和注意力机制,提出了一种新的端到端的样本检测方法;

2)所提模型通过抓取恶意家族样本中的细粒度特征并滤除冗余特征,显著提升了恶意家族超多分类的精度;

3)所提方法可以同时高精度识别出具有恶意行为的应用、恶意应用的类型、恶意家族种类,在不同的分类场景中具有较强的泛化能力.

1 本文方法

本文首先对原始流量数据集进行预处理,将流量数据映射成神经网络模型的输入.引入基于注意力机制改进的残差收缩网络模型,捕获样本细粒度特征,增加重要特征权重,自适应滤除每个流量样本中的噪声与冗余特征,进而提取有效特征并抑制对分类无用的特征.最后高精度地实现恶意-良性应用2分类、恶意应用类型多分类以及恶意家族超多分类.本文方法总体框图如图1所示.

1.1 流量预处理

在训练模型之前,必须将流量数据PCAP文件进行预处理,将它转化为模型可输入数据.本文的数据集处理包括流量切分、流量清洗以及生成灰度图像集等一系列操作流程.

1)加载并过滤PCAP文件:在网络流量检测阶段,HTTP协议是移动网络应用程序中首选的协议,而TCP和UDP是传输层最常见的协议,因此将TCP、UDP和HTTP作为关注的目标.通过Wireshark软件从原始流量文件中加载并过滤出含有相关协议的PCAP文件,以备下一步处理.

2)流量切分:使用USTC-TK2016[23]工具对已经过滤的PCAP文件进行切割,按照5元组(目的IP、源IP、目的端口、源端口和传输协议)进行分流.本文采取的流量单位是会话,即双向流数据.

3)流量清洗:清除没有应用层的会话和内容完全相同的会话.

4)统一长度:由于神经网络的输入要求统一的数据维度,因此需要对不同长度的会话文件进行统一长度.对会话数据进行裁剪,将所有会话数据修剪为1 521 B的文件以保证数据中至少包含一个数据包.截断超过1 521 B的PCAP文件,在字节数不满足1 521 B的文件后补上16进制的0.

5)划分样本:将统一长度后的PCAP文件按照9∶1比例划分成训练集与测试集.

6)生成包字节矩阵:以二进制读取每个固定长度的pcap会话文件,并将每8位二进制转换成十进制数,从而使得每个会话文件生成长度为1 521 B的十进制数组.接着将每个1 521 B的数组整形成39×39的包字节矩阵.

7)归一化处理:对矩阵数据归一化,消除数据量纲的影响,提升模型的收敛速度.

最后,将归一化后的包字节矩阵转换成灰度图像,并制成IDX3格式的灰度图像集.对标签进行独热编码处理,生成与灰度图像集对应的IDX1格式的编码集.预处理流程如图2所示.

王伟[24]通过分析流量可视化结果,发现不同种

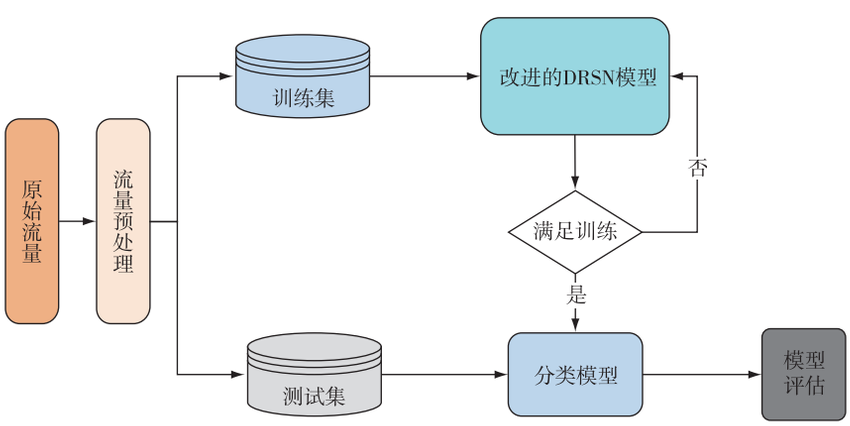


图1 本文方法总体框图

Fig. 1 Overall architecture of the proposed detection of malicious applications

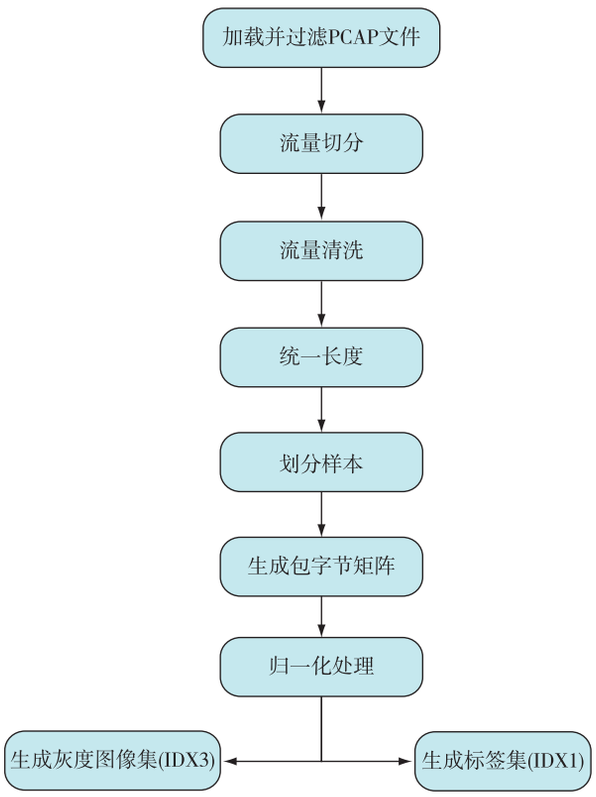


图 2　流量预处理流程

Fig. 2　Flow chart of the network traffic preprocessing

类流量之间的图片的区分度较为明显,认为使用图片分类的方法应该可以取得较好的效果.图 3、图 4 为从本文数据集中抽取的部分流量样本的可视化结果.图 3 为随机抽取的良性软件流量样本,图 4 为随机抽取的恶意软件流量样本.恶意软件流量样本分为 4 类,分别为广告软件、勒索软件、恐吓软件和短信恶意应用,并且从每一类恶意软件流量中随机抽取了 4 个恶意家族的流量样本.可以看出恶意流量与良性流量在肉眼上是可以区分的,而在同一恶意应用类型下,部分不同恶意家族的流量样本差异较小,纹理特征较为相似.因而模型需要提取样本间更加细粒度的特征来对恶意家族种类进行准确区分.

1.2　本文模型

本文模型框图如图 5 所示,使用卷积和注意力机制模块对输入的流量特征进行有效提取,提取的特征通过批归一化层(Batch Normalization,BN)后进一步通过 3 个残差收缩模块,自适应对每张特征图进行噪声的滤除并进一步提取有效特征,然后通过全局平均池化(Global Average Pooling,GAP)对提取出的抽象高维特征降维,大量缩小训练参数,避免过拟合,最后通过全连接层输出分类结果.

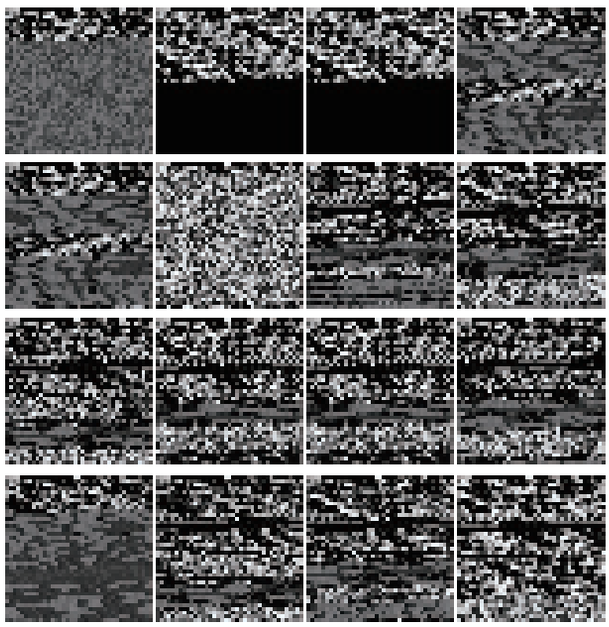
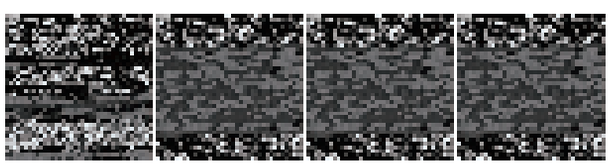
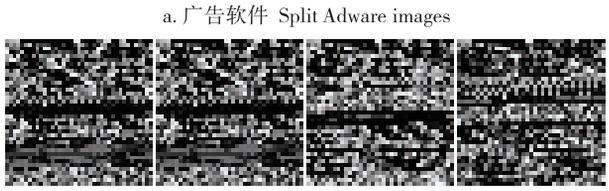


图 3　良性软件流量样本可视化结果

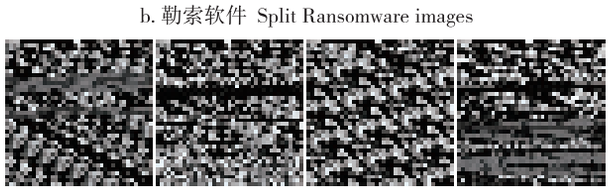
Fig. 3　Visualization results of benign software traffic samples



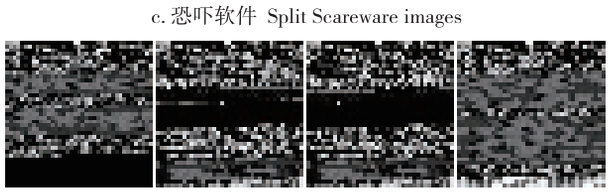
a. 广告软件 Split Adware images



b. 勒索软件 Split Ransomware images



c. 恐吓软件 Split Scareware images



d. 短信恶意应用 Split SMS images

图 4　恶意软件流量样本可视化结果

Fig. 4　Visualization results of malware traffic samples

1.2.1　残差收缩模块

深度残差收缩网络(Deep Residual Shrinkage Network,DRSN)是深度残差网络(Deep Residual Network,ResNet)的一种改进网络[25].引入该网络旨在

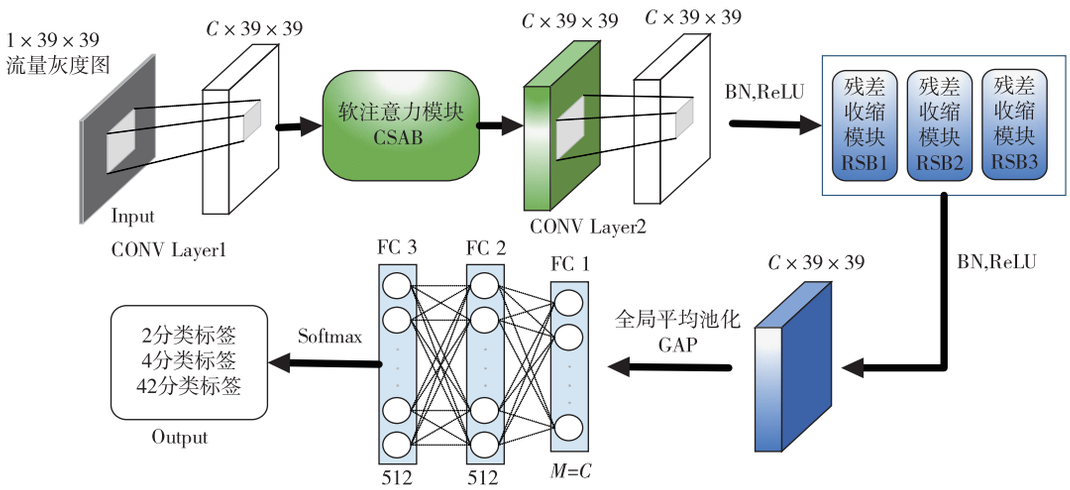


图 5　改进的深度残差收缩网络模型框图

Fig. 5　Block diagram of improved deep residual shrinkage network

加强深度神经网络从含噪声样本中提取有用特征的能力,剔除冗余特征,提升神经网络模型的分类准确率,并且通过残差网络的恒等映射,使反向传播更为方便,降低神经网络训练的难度并防止梯度爆炸.

软阈值化也是许多降噪算法的关键步骤,其将绝对值小于某个阈值的特征删除掉,将绝对值大于该阈值的特征朝着零的方向进行收缩.它可以通过以下公式来实现:

$$y=\begin{cases}x-\tau, & x>\tau,\\ 0, & -\tau\leqslant x\leqslant\tau,\\ x+\tau, & x<-\tau.\end{cases}\tag{1}$$

软阈值化的输出对于输入的导数为

$$\frac{\partial y}{\partial x}=\begin{cases}1, & x>\tau,\\ 0, & -\tau\leqslant x\leqslant\tau,\\ 1, & x<-\tau.\end{cases}\tag{2}$$

由式(2)可知,软阈值化的导数要么是 1,要么是 0.这个性质和 ReLU 激活函数是相同的.因此,软阈值化也能够减小深度学习算法遭遇梯度弥散和梯度爆炸的风险.深度残差收缩网络中嵌入的软阈值化模块是实现噪声数据剔除的关键部分.

图 6 为残差收缩模块(Residual Shrinkage Block,RSB).与普通残差模块不同,残差收缩模块嵌入了一个子网络来自适应生成阈值.在这个子网络中,首先对输入特征图的所有特征,求它们的绝对值,然后经

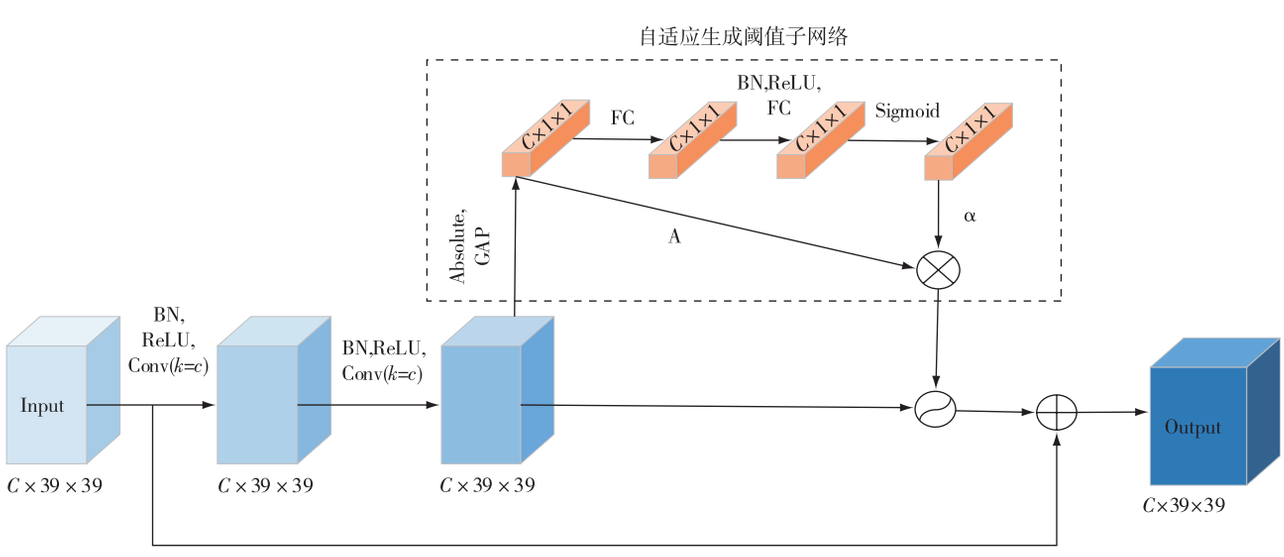


图 6　残差收缩模块 RSB

Fig. 6　Residual shrinkage block

过全局平均值池化,获得一个特征,记为 A.在另一条路径中,全局平均池化之后的特征图,被输入到一个小型的全连接网络.这个全连接网络以 Sigmoid 函数作为最后一层,将输出归一化到 0 和 1 之间,获得一个系数,记为 α.最终的阈值可以表示为 $\alpha \times A$.因此,阈值就是一个 0 和 1 之间的数字×特征图的绝对值的平均.这种方式,不仅保证了阈值为正,而且不会太大.而且,不同的样本就有了不同的阈值.因此,在一定程度上,可以理解成一种特殊的注意力机制:注意到与当前任务无关的特征,通过软阈值化,将它们置为零;或者说,注意到与当前任务有关的特征,将它们保留下来.最后,堆叠一定数量的基本模块以及卷积层、批标准化、激活函数、全局平均池化以及全连接输出层等,就得到完整的深度残差收缩网络.

1.2.2 注意力机制

深度学习中的注意力机制借鉴了人类的注意力思维方式,被广泛地应用在自然语言处理、图像分类及语音识别等各种不同类型的深度学习任务中,并取得了显著的成果.本文采用通道注意力机制与空间注意力机制串联的方式[26]构建软注意力模块(Channel and Spatial Attention Block,CSAB).软注意力模块框图如图 7 所示.输入特征先经过通道注意力机制,$W \times H \times C$ 的维度特征经过基于宽和高的全局平均池化和全局最大池化分别降维成 2 个 $1 \times 1 \times C$ 的特征向量.然后经过共享的多层感知机 MLP,并相加通过 Sigmoid 函数转换成 $1 \times 1 \times C$ 的权重特征向量,最后通过与输入特征相乘,结果即通道注意力机制模块的输出特征 MC.特征获取总体变换公式如下:

$$\mathrm{MC}(X)=\sigma(\mathrm{MLP}(\mathrm{MaxPool}(X))+\mathrm{MLP}(\mathrm{AvgPool}(X))), \tag{3}$$

式中,σ 为非线性激活函数 Sigmoid,MLP 为多层感知机,MaxPool 为最大池化,AvgPool 为平均池化.

将通道注意力机制的输出特征作为空间注意力模块的输入特征,分别在通道维度对其进行基于通道的全局平均池化和全局最大池化.将形成的特征图 Concat 后通过卷积层并经过 Sigmoid 变换,最后生成空间注意力模块特征 MS.总体变换公式如下:

$$\mathrm{MS}(X)=\sigma(f[\mathrm{AvgPool}(\mathrm{MC}(X));\mathrm{MaxPool}(\mathrm{MC}(X))]), \tag{4}$$

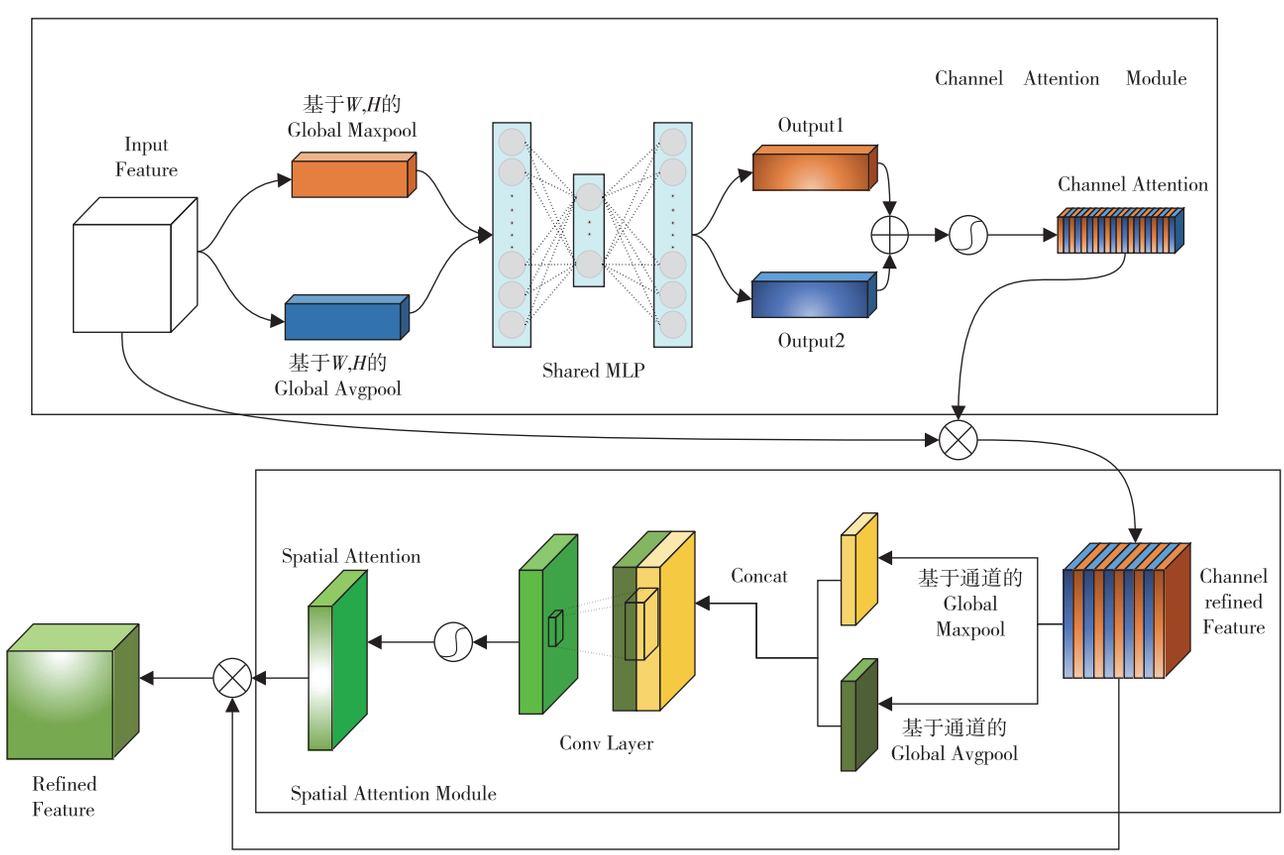


图 7 软注意力模块 CSAB

Fig. 7 Channel and spatial attention block

式中,MS 为最终得到的注意力矩阵,f 为卷积降维操作,σ 为非线性变换.

2 实验与结果分析

2.1 实验环境

本文所采用的实验环境,Windows10,系统处理器:Intel(R) Core(TM) i7-9700K CPU @ 3.60 GHz,RAM:16 GB,系统类型:64 位操作系统,基于 x64 的处理器,显卡:Nvidia GeForce RTX 2070.使用 Keras 深度学习库,Tensorflow 作为后端,利用 GPU 进行深度学习训练.同时,其他的第三方软件还有:Wireshark 、Python、Pycharm、Anaconda 等.

2.2 评价指标

为了公正地判断本实验方法的有效性,本文采用准确率(Accuracy,其量值记为 A)、召回率(Recall,其量值记为 R)、精确率(Precision,其量值记为 P)、F1 值(量值记为 F_1)作为本方法的评价指标,公式如下:

$$A=\frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FN}+\mathrm{FP}}, \tag{5}$$

$$P=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}, \tag{6}$$

$$R=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}, \tag{7}$$

$$F_1=\frac{2PR}{P+R}, \tag{8}$$

式中,TP(True Positive)是将正类预测为正类的数目,FP(False Positive)是将负类预测作为正类的数目,TN(True Negative)表示将负类预测为负类的数目,FN(False Negative)表示将正类预测成为负类的数目.

2.3 数据集简介

为了评估所提模型,本文使用了来自 CICAndMal2017[14] 数据集的 5 065 个良性应用程序和 4 354 个恶意应用样本.这些良性应用根据其受欢迎程度收集自 2015—2017 年发布的 Google play market,并根据 VirusTotal 的检测结果进行识别,只有被 VirusTotal 确定为良性的应用程序才包括在良性应用程序集中.最终,其中 5 065 个被保留为良性应用程序,4 354 个被保留为恶意应用应用程序.

所有的恶意应用有 4 类,它们是广告软件(Adware)、勒索软件(Ransomware)、恐吓软件(Scareware)和短信恶意应用(SMS Malware).并且每个类别有不同的恶意家族,例如广告软件下有 Dowgin、Ewind、Feiwo、Gooligan 等 10 个家族,勒索软件下有 Charger、Koler、Pletor、Ransombo 等 10 个恶意家族,恐吓软件和短信恶意应用下分别有 11 个恶意家族.4 类恶意应用共有 42 个恶意家族.

经过预处理后,良性样本与恶性样本数量分别为 369 211、437 555.本文对安卓恶意应用实现 3 个场景的分类.在 2 分类场景中,提取了原始数据集中所有的良性样本和恶意样本.在 4 分类和超多分类场景中,本文则提取数据集中所有的恶意样本,并分别划分成 4 类和 42 类.预处理后,不同分类场景的样本数目如表 1 所示.

表 1 数据集预处理后不同分类场景的样本数目

Table 1 Number of samples in different classification scenarios after data set preprocessing

分类场景	数据样本	样本数目
恶意-良性软件 2 分类	训练集	726 090
	测试集	80 676
恶意软件类型 4 分类	训练集	393 800
	测试集	43 755
恶意软件家族 42 分类	训练集	393 800
	测试集	43 755

2.4 对比实验

为验证本文方法的有效性,将本文方法用多个指标与文献[14,16-17]基于人工提取流量特征的传统机器学习方法与文献[22]和 CNN 基于流量特征的深度学习方法对比.

本文方法在 3 个分类场景中分类准确率(A)和损失函数(L)曲线如图 8—10 所示.在良性流量与恶意流量的 2 分类场景中,本文方法在训练样本中准确率达 99.61%,在测试样本中达到 99.40%.在恶意应用种类的 4 分类场景中,训练集与测试集样本准确率分别为 99.94%与 99.95%.在更复杂的恶意家族 42 分类场景中,训练集与测试集样本准确率分别为 99.31%与 97.33%.3 个分类场景中损失函数曲线在迭代 20 轮后逐渐收敛,准确率曲线也在迭代 10 轮后达到最大.本文方法在 3 个分类场景中准确率都能达到比较理想的水平,并且模型收敛迅速,在几轮迭代后模型便能达到较高的分类水平.

图 11—13 分别为与其他文献中方法进行对比的结果,进行比较的指标分别为 Precision(P)、Recall(R)、F1 值(F_1).文献[14]采用网络流级特征结合传统机器学习分类器实现 3 个场景的分类,文献

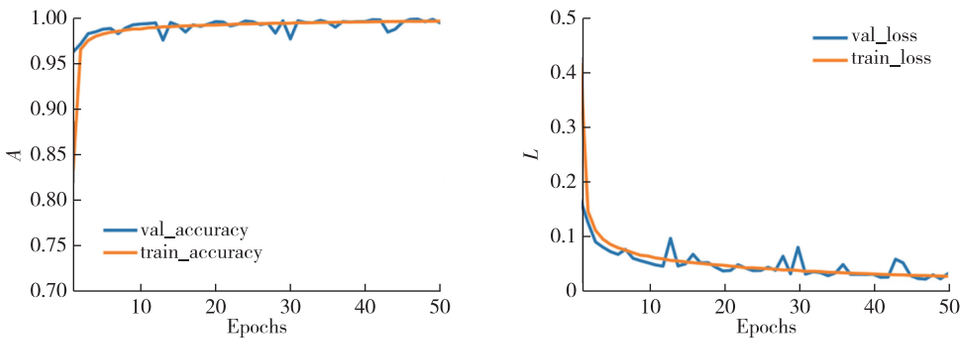


图 8　2 分类下的训练集和测试集的准确率与损失函数曲线

Fig. 8　Accuracy and loss curves of train and test under 2-classification

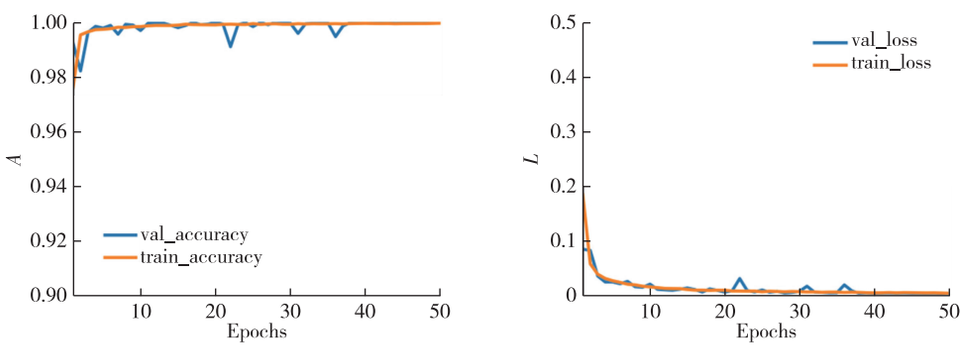


图 9　4 分类下的训练集和测试集的准确率与损失函数曲线

Fig. 9　Accuracy and loss curves of train and test under 4-classification

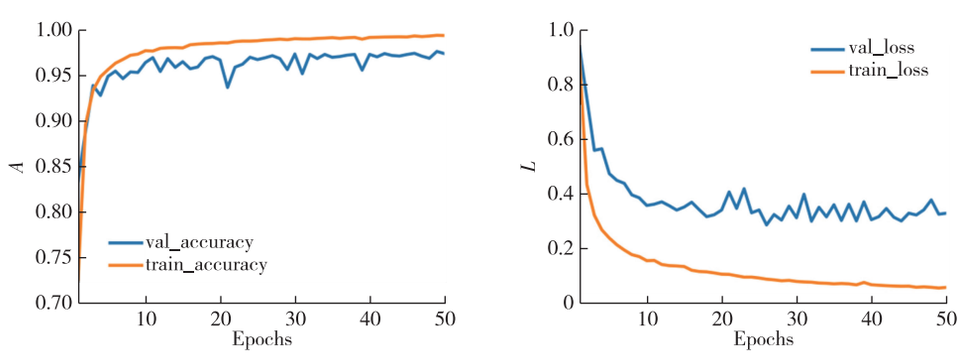


图 10　42 分类下的训练集和测试集的准确率与损失函数曲线

Fig. 10　Accuracy and loss curves of train and test under 42-classification

[16]提取了网络流量特征和 API 特征,文献[17]提取了会话级特征.以上 3 种方法,分别手工提取了流量样本的不同特征并结合传统分类器实现安卓恶意应用检测与分类.文献[22]通过深度学习 CACNN 模型结合流量特征实现 3 种分类任务,取得的准确率分别为 99.19%、97.3%、71.48%.另外,为验证本文方法是否比现有典型的卷积神经网络具有更出色的分类性能,搭建了 CNN 模型并测试.表 2 为本文方法较其他 5 种方法的对比结果.

综合 3 项指标可以看出,本文方法优于参与比较的 5 种方法.在 2 分类与 4 分类任务中,本文方法能取得高于 99%的分类准确率.在恶意家族的 42 超多分类任务中,本文方法精度、召回率、F1 值分别高达 96.04%、94.31%、95.17%,分类效果远高于目前现有方法.由于属于同一大类下的恶意家族样本特征较为相似并且种类较多,对其实现分类具有较高难度.因此其他方法在超多分类的准确率一直不能达到理想的水平.而本文方法可以聚焦样本间细粒度

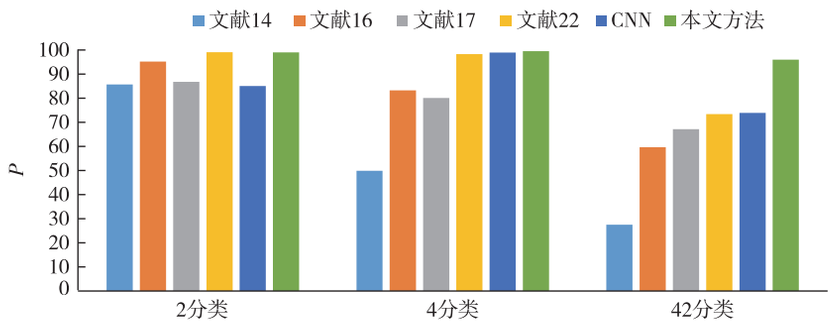


图 11　精确率对比

Fig. 11　Comparison of detection precision

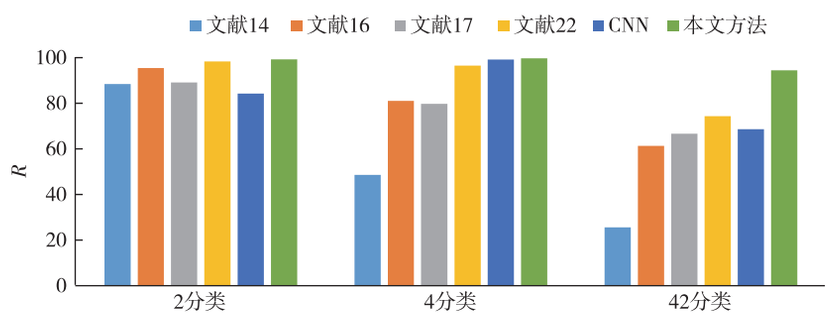


图 12　召回率对比

Fig. 12　Comparison of recall

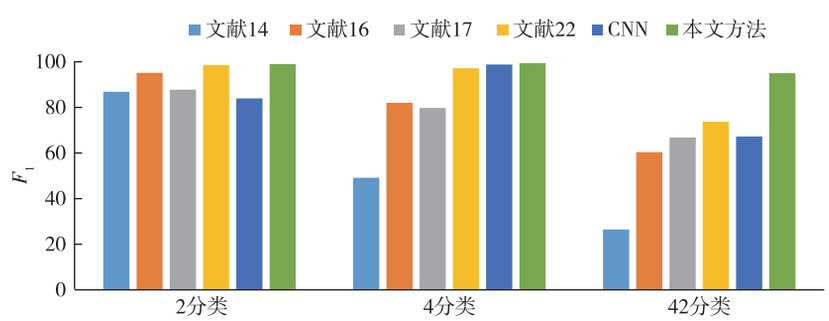


图 13　F1 值对比

Fig. 13　Comparison of F1

表 2　恶意应用分类结果对比

Table 2　Performance comparison of malicious application classifications

评价指标	分类任务	文献[14]	文献[16]	文献[17]	文献[22]	CNN	本文方法
P	2 分类	85.80	95.30	86.86	99.20	85.11	99.12
	4 分类	49.90	83.30	80.20	98.40	99.06	99.59
	42 分类	27.50	59.70	67.21	73.50	73.97	96.04
R	2 分类	88.30	95.30	89.00	98.20	84.13	99.12
	4 分类	48.50	81.00	79.64	96.40	99.04	99.59
	42 分类	25.50	61.20	66.59	74.20	68.55	94.31
F_1	2 分类	87.03	95.30	87.92	98.70	84.14	99.12
	4 分类	49.19	82.13	79.92	97.39	99.04	99.59
	42 分类	26.46	60.44	66.90	73.85	67.29	95.17

特征,提高对分类产生重要作用特征的权重,自适应滤除冗余特征,有效地提高超多分类任务的分类效果.相比现有方法,本文方法在恶意应用的3个分类场景中都具有优良的分类能力,因此,本文所提方法具有一定的泛化性.

3 结束语

本文提出了一种基于改进残差收缩网络的安卓恶意应用检测方法.所提方法通过预处理将流量数据映射成神经网络的输入,避免了人工提取特征的复杂性和繁琐性,实现端到端的自我学习.同时在网络架构的设计中,引入了注意力机制捕获样本间细粒度特征,又通过引入深度残差收缩网络,自适应滤除样本中大量冗余特征,减少大样本多分类任务给模型带来的分类难度,有效实现了安卓恶意-良性应用的2分类、安卓恶意应用类型的4分类以及恶意家族的42超多分类.3个场景下的准确率分别高达99.40%、99.95%和97.33%,与现有方法相比,具有较高的分类性能与泛化能力,并且在恶意家族超多分类任务中有较大的优势.下一步,将针对流量交互中探测器对数据包捕获存在丢失而导致检测精度降低的问题,研究相应的检测方法.

参考文献

References

[1] 360安全资讯.2020年第三季度中国手机安全状况报告[EB/OL].[2021-06-04].https://zt.360.cn/1101061855.php?dtid=1101061451&did=610689546

[2] 柯懂湘,潘丽敏,罗森林,等.基于随机森林算法的Android恶意行为识别与分类方法[J].浙江大学学报(工学版),2019,53(10):2013-2023
KE Dongxiang, PAN Limin, LUO Senlin, et al. Android malicious behavior recognition and classification method based on random forest algorithm[J].Journal of Zhejiang University (Engineering Science), 2019, 53(10): 2013-2023

[3] Zeng N Y, Wang Z D, Zineddin B, et al. Image-based quantitative analysis of gold immunochromatographic strip via cellular neural network approach[J].IEEE Transactions on Medical Imaging,2014,33(5):1129-1136

[4] 秦中元,王志远,吴伏宝,等.基于多级签名匹配算法的Android恶意应用检测[J].2016,33(3):891-895
QIN Zhongyuan, WANG Zhiyuan, WU Fubao, et al. Android malware detection based on multi-level signature matching[J].Application Research of Computers,2016,33(3):891-895

[5] Luo X, Zhou M C, Leung H, et al. An incremental-and-static-combined scheme for matrix-factorization-based collaborative filtering [J]. IEEE Transactions on Automation Science and Engineering, 2016, 13(1): 333-343

[6] Sun L C, Li Z Q, Yan Q B, et al. SigPID: significant permission identification for android malware detection[C]//2016 11th International Conference on Malicious and Unwanted Software (MALWARE). October 18-21, 2016, Fajardo, PR, USA. IEEE, 2016: 1-8

[7] Onwuzurike L, Mariconti E, Andriotis P, et al. MaMaDroid: detecting android malware by building markov chains of behavioral models (extended version)[J].ACM Transactions on Privacy and Security,2019,22(2):1-34

[8] Zhang L S, Niu Y, Wu X, et al. A3: automatic analysis of android malware [C] // Proceedings of the 1st International Workshop on Cloud Computing and Information Security. November 9-11, 2013, Shanghai, China. Paris, France: Atlantis Press, 2013: 89-93

[9] Sabhadiya S, Barad J, Gheewala J. Android malware detection using deep learning[C]//2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI). April 23-25, 2019, Tirunelveli, India. IEEE, 2019: 1254-1260

[10] Liang H L, Song Y, Xiao D. An end-to-end model for android malware detection[C]//2017 IEEE International Conference on Intelligence and Security Informatics (ISI). July 22-24, 2017, Beijing, China. IEEE, 2017: 140-142

[11] Amos B, Turner H, White J. Applying machine learning classifiers to dynamic android malware detection at scale [C]//2013 9th International Wireless Communications and Mobile Computing Conference (IWCMC). July 1-5, 2013, Sardinia, Italy. IEEE, 2013: 1666-1671

[12] Sarma B P, Li N H, Gates C, et al. Android permissions: a perspective combining risks and benefits [C] // Proceedings of the 17th ACM Symposium on Access Control Models and Technologies. June 20-22, 2012, Newark, New Jersey, USA. New York: ACM Press, 2012: 13-22

[13] Zhou Y J, Jiang X X. Dissecting android malware: characterization and evolution[C]//2012 IEEE Symposium on Security and Privacy. May 20-23, 2012, San Francisco, CA, USA. IEEE, 2012: 95-109

[14] Lashkari A H, Kadir A F A, Taheri L, et al. Toward developing a systematic approach to generate benchmark android malware datasets and classification[C]//2018 International Carnahan Conference on Security Technology (ICCST). October 22-25, 2018, Montreal, QC, Canada. IEEE, 2018: 1-7

[15] Noorbehbahani F, Rasouli F, Saberi M. Analysis of machine learning techniques for ransomware detection [C]//2019 16th International ISC (Iranian Society of Cryptology) Conference on Information Security and Cryptology (ISCISC). August 28-29, 2019, Mashhad, Iran. IEEE, 2019: 128-133

[16] Taheri L, Kadir A F A, Lashkari A H. Extensible android malware detection and family classification using network-flows and API-calls [C] // 2019 International Carnahan Conference on Security Technology (ICCST).

October 1-3,2019,Chennai,India.IEEE,2019:1-8
[17] Abuthawabeh M,Mahmoud K.Enhanced android malware detection and family classification,using conversation-level network traffic features[J].The International Arab Journal of Information Technology,2020,17(4A):607-614
[18] Chen R,Li Y Y,Fang W W.Android malware identification based on traffic analysis[C]//Artificial Intelligence and Security,2019:293-303
[19] Arora A,Peddoju S K.Minimizing network traffic features for android mobile malware detection[C]//Proceedings of the 18th International Conference on Distributed Computing and Networking.New York,NY,USA:ACM,2017:1-10
[20] Wang W,Zhu M,Wang J L,et al.End-to-end encrypted traffic classification with one-dimensional convolution neural networks[C]//2017 IEEE International Conference on Intelligence and Security Informatics(ISI).July 22-24,2017,Beijing,China.IEEE,2017:43-48
[21] Marín G,Caasas P,Capdehourat G.DeepMAL-deep learning models for malware traffic detection and classification[C]//Data Science:Analytics and Applications,2021:105-112
[22] Feng J Y,Shen L M,Chen Z,et al.A two-layer deep learning method for android malware detection using network traffic[J].IEEE Access,2020,8:125786-125796
[23] Wang W,Zhu M,Zeng X W,et al.Malware traffic classification using convolutional neural network for representation learning[C]//2017 International Conference on Information Networking(ICOIN).January 11-13,2017,Da Nang,Vietnam.IEEE,2017:712-717
[24] 王伟.基于深度学习的网络流量分类及异常检测方法研究[D].合肥:中国科学技术大学,2018
WANG Wei.Deep learning for network traffic classification and anomaly detection[D].Hefei:University of Science and Technology of China,2018
[25] Zhao M H,Zhong S S,Fu X Y,et al.Deep residual shrinkage networks for fault diagnosis[J].IEEE Transactions on Industrial Informatics,2020,16(7):4681-4690
[26] Woo S,Park J,Lee J Y,et al.CBAM:convolutional block attention module[M]//Computer Vision-ECCV 2018.Cham:Springer International Publishing,2018:3-19

Detection of malicious applications based on improved deep residual shrinkage network

XU Lilong[1] ZHAI Jiangtao[1] LIN Peng[1] CUI Yongfu[1]
1 School of Electronics & Information Engineering,Nanjing University of Information Science & Technology,Nanjing 210044

Abstract The rapid growth of malicious applications has posed a security threat to mobile intelligent terminals.It is of great significance to achieve high-precision detection of malicious applications for mobile network information security.Here,this paper proposes a method to detect malicious applications based on improved deep residual shrinkage network.First,the traffic features are preprocessed into convolutional neural network inputs,and then the channel attention mechanism and spatial attention mechanism are introduced to weight the sample features from the channel and spatial dimensions.Then,the deep residual shrinkage network is introduced to adaptively filter out the redundant features of the samples,and the parameters are back propagated through the identical connection optimization,so as to reduce the difficulty of model training and classification,and finally realize the high-precision identification of malicious android applications.The proposed method avoids manual feature extraction,achieves high-precision classification and has certain generalization ability.Experimental results show that the accuracy of the proposed method is 99.40%,99.95% and 97.33% in 2-classification,4-classification and 42-classification of malicious applications,respectively.Compared with the existing methods,the proposed method has better classification performance and generalization ability.

Key words malicious application;malicious families;deep residual shrinkage network;information security