基于灰度图像转化的时间型隐蔽信道检测方法

陈虹吕[1]，王诗蕊[2]，李　峰[1]，王俊峰[1]
(1. 四川大学计算机学院(软件学院)，成都 610065；2. 国家工业信息安全发展研究中心，北京 100040)

摘　要：时间型网络隐蔽信道是一种隐蔽性极高的信息泄露方式. 其作为 APT 攻击的主要通信手段，对网络安全产生了极大威胁. 目前针对隐蔽信道的检测方法通用性不足、误检率高，且人工提取流量特征耗时耗力. 本文提出了一种基于灰度图像转化的检测方法. 该方法将报文到达时间间隔归一化，转换成像素值，再将其转为灰度图像，由此把一维序列分类问题转成二维图像分类问题. 本文使用卷积神经网络自动获取图像特征，并利用卷积块注意力模块，从空间与通道两个维度进行特征自适应优化. 本文用合法流量和隐蔽信道流量组成的数据集训练网络，所得到的二分类模型用于判别被检测流量是否为时间型隐蔽信道流量. 最后将提出的方法与现有的 4 种检测方法做对比. 实验结果表明，本文方法具有更高的精确率和召回率，所得模型的通用性更好且误检率更低.

关键词：时间型隐蔽信道；灰度图像；卷积神经网络；注意力
中图分类号：TP393　　**文献标识码**：A　　DOI：10.19907/j.0490-6756.2023.033003

Covert timing channel detection method based on grayscale image transformation

CHEN Hong-Lü[1], *WANG Shi-Rui*[2], *LI Feng*[1], *WANG Jun-Feng*[1]
(1. College of Computer Science (College of Software), Sichuan University, Chengdu 610065, China;
2. National Industrial Information Security Development Research Center, Beijing 100040, China)

Abstract: Network covert timing channel is a highly concealed method of information leakage. As the main communication method of APT attack, it poses a great threat to network security. The current detection methods for covert channel are insufficient generality, have high false detection rate, and manual extraction of features is time-consuming. This paper proposes a detection method based on grayscale image transformation. This method normalizes the inter-arrival time sequence of packets, and converts them into pixel values, and then converts into a grayscale image, thereby transforming a problem of one-dimensional sequence classification into a problem of two-dimensional image classification. The authors use the Convolutional Neural Networks to automatically acquire image features, and use the Convolutional Block Attention Module to optimize the feature adaptively from two dimensions of space and channel. The authors train the network with the data set composed of legitimate traffic and covert channel traffic, and the obtained binary classifier can be used to judge whether the detected traffic is covert timing channel traffic. Finally, the proposed approach is compared with the four existing detection methods. The experimental results show that the proposed method has higher precision and recall rate, and the pro-

收稿日期：2022-06-24
基金项目：国家重点研发计划(2019QY1400)；国家自然科学基金(U2133208)；四川省青年科技创新研究团队(2022JDTD0014)
作者简介：陈虹吕(1997－)，女，重庆人，硕士研究生，研究方向为网络空间安全. E-mail：13088021029@163.com
通讯作者：王俊峰. E-mail：wangjf@scu.edu.cn

posed model has better generality and lower false detection rate.

Keywords: Covert timing channel; Grayscale image; Convolutional neural network; Attention

1 引　言

随着互联网技术不断发展,信息安全问题变得日益严峻.为规避信息泄露事故的发生,网络入侵检测也更加严格.因此,如何绕过检测系统盗取信息成为了攻击者的研究热点.目前存在一种名为隐蔽信道的信息盗窃方式,它凭借极高的隐蔽性令许多入侵检测系统束手无策[1].隐蔽信道这一概念在1973 年由 Lampson 等首次提出[2],它利用协议的规则漏洞,通过现有系统资源来传输信息,而这些资源本不该用于传输数据.隐蔽信道对于目前的网络安全机制不可见,它的高隐蔽性给网络环境带来了极大的安全隐患.因此,对隐蔽信道检测方法的研究具有重要意义.

隐蔽信道大多数时候被用于非法目的.例如,高级可持续性威胁(APT)攻击中,隐蔽信道常常作为信息回传的工具;一些 APT 组织会利用隐蔽信道潜伏在网络中搜集政府的隐私信息[3];计算机病毒可以通过隐蔽信道在不被发现的情况下自行传播[4].2010 年发生了一场名为"极光行动"的攻击,攻击者利用 IE 漏洞和基于 SSL 协议的隐蔽信道,获取到谷歌的账号密码等信息,导致谷歌网络被渗入数月,大量机密数据被窃取[5].

网络隐蔽信道作为信息安全中一个重要研究领域越来越受到重视[6].我国的《计算机信息系统安全保护等级划分准则》(GBl7859-1999)[7]以及国际标准化组织 ISO 发布的《信息技术安全评估通用准则》(ISO/IEC 15408)[8]对隐蔽信道分析提出了明确规定,要求高安全等级的信息系统必须进行隐蔽信道分析.因此,对隐蔽信道检测方法的研究具有着现实的需求和意义.

网络隐蔽信道种类繁多,根据它们隐藏信息的载体差异可分为两类:时间型隐蔽信道和存储型隐蔽信道[9].存储型隐蔽信道,顾名思义,是以通信双方的共享资源为载体传递隐蔽信息,共享资源主要包括协议的首部字段,如:IP 协议的服务类型(TOS)[10]、标志位(flag)[11]、生存时间(TTL)[12]字段、TCP 协议的序列号(ISN)字段[13]和 SSH 协议的 MAC 字段[14]等.存储型隐蔽信道受到网络协议的严格约束,不能偏离特定行为.针对这一类型,通常只需要监视数据包的特定字段,判断其取值范围和分布是否存在异常即可.而时间型隐蔽信道则利用报文的时间特性来传递隐蔽信息,如:报文的发送时间、到达时间间隔等.它只改变报文的时间特性,并不违反网络安全策略,与存储型信道相比,有着更大的随机性和更强的隐蔽性,因此对该类型的信道检测更具挑战.针对时间型隐蔽信道的检测成为了近年来相关领域的热门研究课题.目前有学者针对时间型隐蔽信道提出了不同的检测方法,但经调研,这些方法只在特定类型的信道上有较好的效果,且对合法流量的误报率较高.

针对上述问题,本文提出了一种新颖的基于灰度图像转化的时间型隐蔽信道检测方法,该方法先将报文到达接收方的时间间隔归一化,再转换为灰度图像.在此基础上利用二维卷积神经网络(2D-CNN)和卷积块注意力模块(CBAM)训练二分类模型,以达到更好的检测效果.本文的主要创新和贡献如下:(1) 提出了基于灰度图像转化的时间型网络隐蔽信道检测方法,将流量可视化,把一维序列分类问题转换为二维图像分类问题,使用图像分类技术检测隐蔽信道.(2) 提出了将卷积神经网络和 CBAM 注意力模块相结合的自动检测方法,从图像空间维度和网络通道维度分别生成注意力特征图并加以融合,该方法在检测性能上优于手工提取特征的方法.(3) 本文仿真了四种经典的时间型隐蔽信道并收集了隐蔽信道流量,结合真实环境下的合法流量进行大量的实验.实验结果表明,本文提出的方法可以实现隐蔽信道的自动检测,该方法具有更好的通用性,较其他方法在精确率、召回率上均有一定提升.

2 相关工作

网络隐蔽信道按照传输数据的网络资源类型可分为两类:存储型和时间型.由于本文的研究主要针对时间型隐蔽信道,因此对存储型不再展开介绍.下面我们将对经典的时间型隐蔽信道的设计、现有的检测方法及图像处理的应用展开介绍.

2.1 时间型隐蔽信道

在本文中,我们主要关注 4 种经典的时间型隐蔽信道:BER、GAS、CAB 和 ZAN.我们采用了在文献中首次提出这种隐蔽信道设计的作者名字的前三个字母大写来命名.本小节将详细介绍这 4 种

隐蔽信道的设计.

(1) BER. 对于 BER[15] 隐蔽信道，我们首先设定两个值：t_0 和 t_1. 把需要发送的隐蔽信息编码为二进制的形式，当需要发送 0 时，就将报文时间间隔设置为 t_0；当需要发送 1 时，就将时间间隔设置为 t_1. 接收方收到报文后，可以根据提前约定好的规则解码二进制数据，从而提取隐蔽信息.

(2) GAS. Gasior 等人[16] 提出了一种新的嵌入隐蔽信息的方法. 首先需要设定一个基准值 th：当需要发送 0 时，在 th 上减去一个随机值 Δt，作为报文时间间隔；当需要发送 1 时，则在 th 的基础上加上 Δt. 接收方通过比较真实的报文到达时间间隔和 th 的值来解码信息.

(3) CAB. Cabuk[17] 提出了时间重放隐蔽信道. 该信道在作者早期提出的模型上做了改进. 为了规避基于统计方法的检测，该信道选取了合法流量的报文时间间隔，将间隔值排序后，选取一个阈值，根据阈值将时间间隔分为两个不相交的集合 T_0 和 T_1. 从间隔值较小的集合 T_0 中无重复地随机选取一个值代表 0；相应地，从集合 T_1 中随机选取一个值代表 1.

(4) ZAN. Zander 等人[18] 提出了 ZAN 信道，它源自一种将隐蔽信息编码进 IP 报文 TTL 字段的调制技术. 作者将初始间隔时间设为 t_0. 当发送数据 0 时，保持间隔时间不变，即 $t_k = t_{k-1}$；当发送数据 1 时，在上一个间隔时间 t_{k-1} 的基础上加上或减去一个随机值 ε，即 $t_k = t_{k-1} \pm \varepsilon$.

2.2　检测方法介绍

经典的时间型隐蔽信道检测方法可分为三类：基于统计的方法、基于熵的方法和基于机器学习的方法，下面我们将对这些方法展开介绍.

2.2.1　*基于统计的方法*　Cabuk 等人[19] 提出了一个基于规律性度量的方法. 将采样的间隔时间分为 N 段，得到每一段的标准差，表示为 $\sigma_1, \sigma_2, \cdots, \sigma_N$，然后计算相邻 σ_i 相对差的标准差，表示为

$$Regularity = STDEV\left(\frac{|\sigma_i - \sigma_j|}{\sigma_i}, i<j, \forall i,j\right).$$

Cabuk 等人[19] 在同一文献中提出了第二种检测方法，名为 ε-*similarity* 测试. 它首先按升序对报文的时间间隔进行排序，排序后的间隔表示为 $a_1, a_2, \cdots, a_n$，然后计算相邻间隔时间的相对差 $\frac{|a_i - a_{i+1}|}{a_i}$. ε-*similarity* 即为相对差小于 ε 的百分比. Yang 等人[20] 提出可以使用 Kolmogorov-Smirnov(K-S)测试，度量待测样本与预期合法流量样本的经验分布函数差距，来将隐蔽信道和合法流量区分开来. 距离较大则表明可能存在隐蔽信道. Rezaei 等人[21] 提出了使用 Spearman-Rho 测试、Mann-Whitney-Wilcoxon 秩和检验、Wilcoxon 秩序检验三种非参数统计测试来检验是否存在隐蔽信道流量.

2.2.2　*基于熵的方法*　Gianvecchio 等人[22] 提出使用熵和修正条件熵来检测隐蔽信道. 熵值较高意味着随机性较大；而熵值低意味着数据分布可能存在固定模式，则样本为隐蔽信道流量的可能性更高. 为了同时评估待测数据的分布特性和数值特性，张宇飞等人[23] 提出了使用差分信息熵来检测隐蔽信道的方法，该方法对数据的变化具有更高的敏感度.

2.2.3　*基于机器学习的方法*　Shresth 等人[24] 提出使用 SVM 的方法检测隐蔽信道，他们使用报文到达时间间隔的 4 个统计特征作为流量指纹：K-S 值、regularity 分数、熵和修正条件熵. 通过四个指纹训练模型来检测隐蔽信道. Fu 等人[25] 提出了基于随机森林的检测方法，方法中用到了 8 个统计特征. Darwish[26] 提出了一种基于层次特征提取的方法，根据提前制定的规则对每一层特征进行剪枝处理，结合深度神经网络训练检测模型，该方法在检测准确率上有较大的提升. Han 等人[27] 提出了一种基于时间间隔和有效载荷长度分析的检测方法，该方法提取了与时间和载荷相关的 20 个统计量，并最终验证了使用 KNN 算法的模型相比其他算法具有更好的检测性能. Wu 等人[28] 提出将时间间隔离散化后转换为符号序列，通过计算状态转换概率矩阵来检测隐蔽信道.

现有方法虽然对部分隐蔽信道有一定的检测效果，但依然存在以下缺点：(1) 大多数方法通过手动提取特征，序列中的局部显著性特征难以得到关注并有效地表达. 且手动选择特征过度依赖专家知识，耗时耗力，特征的好坏直接影响模型的检测性能，特征的适应性和表征性不足会导致模型的鲁棒性降低. (2) 模型通用性和泛化能力不足，只能对部分隐蔽信道有较好的效果，无法可靠地检测出多种类型的信道. (3) 现有方法对合法流量的误报率较高.

2.3　图像技术在网络安全领域的应用

近年来，图像技术被广泛应用于其他领域，包括网络安全领域. 众多研究表明，图像技术在处理

分类问题上有较高的准确性.

Nataraj 等人[29]使用图像处理技术将二进制代码样本结构转换为二维灰度图像,然后从这些图像中提取各种特征,以训练机器学习分类器来检测恶意软件.该分类器在检测多种类型的恶意软件时达到了 97.18%的准确率.2017 年,Chen 等人[30]鉴于传统 IP 流量分类严重依赖于手工特征提取,提出了将序列转化为图像后利用卷积神经网络进行分类的新方法,并用实验验证了该方法的有效性.2018 年,Taheri 等人[31]提出将合法流量和僵尸网络流量数据转换为图像,再利用 DenseNet 训练分类器来检测僵尸网络的方法.2020 年,He 等人[32]提出了将会话的前几个非零有效载荷转换为灰度图像,并使用卷积神经网络对转换后的灰度图像进行分类,来分类加密后的网络流量. Al-Eidi[33]提出了将时间型隐蔽信道流量转化为彩色图像,手动提取图像的 8 个常见特征,用以训练分类器,以克服统计方法检测精度低的缺点.

鉴于现有隐蔽信道检测方法的缺点,以及图像处理技术在解决网络安全领域问题时较为有效,本文提出了一种基于灰度图像转化的时间型隐蔽信道检测方法.该方法将报文时间间隔序列转换为灰度图像,结合 CNN 和 CBAM 注意力模块训练分类器,实现了图像特征的自动提取.该方法不再依赖专家知识,节省了人力,且提高了模型的通用性和检测性能,降低了误检率.

3　检测方法

3.1　方法概述

如图 1,本文的系统架构一共由 4 个部分组成:数据生成、数据预处理、模型训练和模型测试.首先,我们搭建了一个包含多个主机的通信环境,用于构建多种隐蔽信道.两个主机之间通信,生成隐蔽信道流量,用 wireshark 捕获流量后保存为 pcap 文件.合法流量取自 MAWI 工作组发布的骨干网络日流量.随后将这些流量做预处理,根据<源 IP 地址,目的 IP 地址,源端口,目的端口,协议类型>五元组提取出流,得到每条流的报文时间间隔,归一化后转换为灰度图像.再用 CNN 网络结合 CBAM 注意力模块训练检测模型.考虑到现实场景中会遇到未知类型的隐蔽信道,因此我们将该检测问题定为二分类问题.

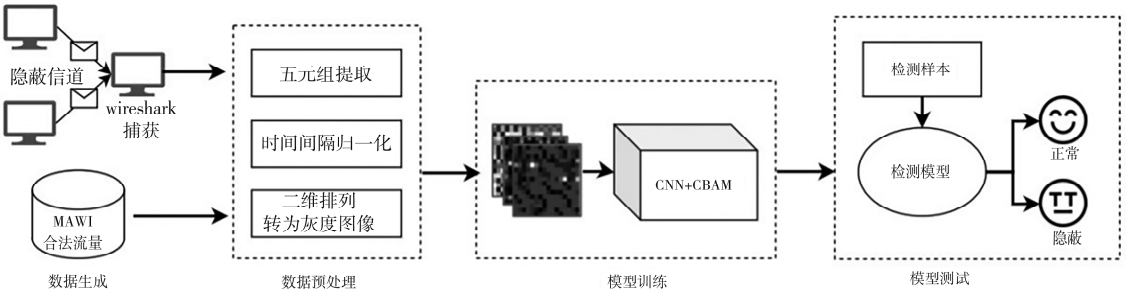


图 1　基于灰度图像转化的时间型隐蔽信道检测方法框架

Fig. 1　Covert timing channel detection method framework based on grayscale image transformation

3.2　数据预处理

由于时间型隐蔽信道的信息主要藏匿于报文时间间隔中,因此本文对流量的主要关注点也是报文的时间间隔.原始数据是 pcap 文件,将文件中的报文按照五元组<源 IP 地址,目的 IP 地址,源端口,目的端口,协议类型>分流后,提取报文的到达时间间隔,每 401 个报文作为一个组,则一组应包含 400 个时间间隔值.然后我们将 400 个值归一化.归一化是将数值的绝对值变成相对关系,将数据限制在所需要的范围内.归一化过程使用了最大最小缩放法,要将原始数据归一化到区间$[a,b]$,计算方法如式(1)和 式(2)所示.其中X_i、X_i'为序列中第 i 个数进行归一变换前后的值,$X_{\max}$、$X_{\min}$分别为该序列中的最大值和最小值. 该缩放为可逆变换,如式(3)所示.

$$X_i' = a + k(X_i - X_{\min}) \tag{1}$$

$$k = \frac{b-a}{X_{\max} - X_{\min}} \tag{2}$$

$$X_i = \frac{X_i' - a}{k} + X_{\min} \tag{3}$$

由于要将时间间隔转为灰度图像,所以我们选择的范围是$[0,255]$.接下来将处理后的每组数据二维排列,从左到右、从上到下地填充一个 20×20 的矩阵.矩阵中的每个位置都代表灰度图像中的一个像素点.每张图像都可以在展平后通过式(3)的逆变换变回原始数据. 由此可见该预处理为无损变换.

我们采用了 Python 的 ImageIO 库写入图像数据，图 2 是各类隐蔽信道流量和合法流量生成的灰度图示例.

图 2　转换后的流量灰度图
Fig. 2　Transformed traffic grayscale images

3.3　模型构建

在模型训练中，我们采用卷积神经网络(Convolutional Neural Networks，CNN)和卷积块注意力模块(Convolutional Block Attention Module，CBAM)相结合的网络结构，如图 3.

CNN 是一种被广泛应用于图像识别领域的神经网络，它可以实现图像特征的自动提取. CNN 由卷积层、池化层、激活函数和全连接层组成. 通过卷积和池化，网络提取不同层次的图像特征并进行压缩，最终可学到高层次的特征. CNN 模块输出的特征图为 F，$F\in\mathbb{R}^{C\times H\times W}$. 其中 C 为通道数，H 和 W 分别为特征图的高度和宽度.

CBAM 是一种混合注意力机制[34]，它结合了通道注意力M_c和空间注意力M_s. 每个注意力模块的输出都与输入做⊗运算，运算结果作为下一模块的输入，如式(4)、式(5)所示. ⊗表示哈达玛积运算，将输出的特征图维度与输入保持一致.

$$F'=M_c(F)\otimes F \tag{4}$$

$$F''=M_s(F')\otimes F' \tag{5}$$

通道注意力模块的处理过程如式(6)所示. CNN 模块的输出为 F，将作为通道注意力模块的输入，使用全局平均池化和最大池化，将特征图在空间维度进行压缩，传入共享网络 MLP，最后经激活函数 Sigmoid 激活，产生通道注意力权重$M_c(F)$，$M_c(F)\in\mathbb{R}^{C\times1\times1}$. 其中 σ 为激活函数 Sigmoid.

$$M_c(F)=\sigma\begin{pmatrix}MLP(\mathrm{AvgPool}(F))\\+MLP(\mathrm{MaxPool}(F))\end{pmatrix} \tag{6}$$

与通道注意力不同的是，空间注意力模块关注特征的空间关系，分别用平均池化和最大池化压缩特征图的通道维度，将池化的结果拼接后传入卷积层生成空间注意力权重$M_s(F')$，卷积核为 7×7 大小，如式(7).

$$M_s(F')=\sigma\begin{pmatrix}f^{7\times7}([\mathrm{AvgPool}(F');\\\mathrm{MaxPool}(F')])\end{pmatrix} \tag{7}$$

两个注意力模块在通道和空间两个维度互为补充，更好地计算出注意力特征图，有益于提升模型的检测性能. 这一点我们将在下一节的实验中加以验证.

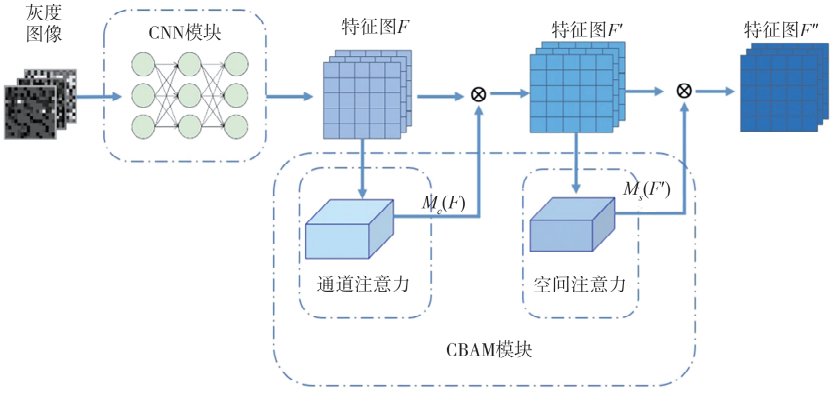


图 3　基于 CNN 和 CBAM 的网络模型图
Fig. 3　Network model diagram based on CNN and CBAM

4　实验结果

4.1　数据集

为了训练分类模型，我们需要收集隐蔽信道流量和合法流量. 其中，合法流量数据我们选用了 MAWI 工作组在 2021 年 10 月 1 日发布的日流量中的 TCP 流量[33]. 对于隐蔽信道流量，我们选用了 2.1 小节中描述的四种经典信道 BER[14]、GAS[15]、CAB[16]、ZAN[17]，并根据它们在原论文中的设计进行仿真. 我们在局域网中部署了 3 台主机，借助主机间建立的 TCP 连接实现隐蔽通信.

对于 BER 信道，我们设置$t_0=40$ ms，$t_1=80$ ms. 对于 GAS 信道，我们设基准值 $th=60$ ms，随机值 Δt 范围为 10～40 ms. 因此发送信息 0 和 1

的取值分别为 $th-\Delta t$ 和 $th+\Delta t$. 对于 CAB 信道，我们提取合法流量中的 TCP 流，按升序进行排列，并将其分为两个集合 T_0 和 T_1，其中 T_0 和 T_1 具有相同的元素个数. 对于 ZAN，我们设初始值 $t_0=40$ ms，随机值 ε 的范围为 30～60 ms. 当发送隐蔽信息 1 时，我们让加和减操作交替进行，即若 $t_{k-1}=t_{k-2}+\varepsilon$，则 $t_k=t_{k-1}-\varepsilon$. 由于 ε 是一个随机值，为了避免 t_k 过大或者过小，我们约定：若 $t_{k-1}>80$ ms，则 $t_k=t_{k-1}-\varepsilon$；若 $t_{k-1}-\varepsilon<10$ ms，则 $t_k=t_{k-1}+\varepsilon$. 这是因为当间隔时间太大，很可能会呈现较为明显的异常，这将导致隐蔽信道更容易被检测出来；若间隔时间太小，将更容易受到网络抖动因素的影响而造成信息解码错误，因此我们做了间隔范围的控制.

生成原始数据后，我们按照五元组划分流，然后保存每一条单向流的时间间隔. 我们一共收集了 680 000 个合法时间间隔，然后按照每组 400 个间隔进行切分，共分为 1700 组，生成 1700 张 20×20 的灰度图像样本. 其中 1600 个样本被用作训练，100 个被用作测试. 对于每类隐蔽信道，我们收集了 200 000 个时间间隔，仍然按照每组 400 个间隔进行分组，得到 500 组数据，生成 500 张图象样本. 其中 400 张用于训练，100 张用于测试. 我们让所有隐蔽信道的训练样本同时参与训练，然后对每类隐蔽信道分别进行检测，这样可以保证训练集和测试集中的合法流量与隐蔽信道流量的样本比例为 1∶1.

4.2　性能评估

我们用基于 CNN 和 CBAM 注意力模块训练得到的模型对每一类隐蔽信道分别进行了检测实验，本小节用混淆矩阵来评估模型的检测性能.

(1) BER. 表 1 展示了我们的模型检测 BER 信道的混淆矩阵，合法流量和隐蔽流量的检测率分别为 90%和 98%. 据分析，在 BER 信道上的高检测率是由于它的编码规则将与合法流量的间隔分布呈现较为显著的差异. BER 信道的时间间隔分布主要取决于通信双方设定的两个固定值以及隐蔽信息的二进制编码. 网络抖动会对间隔值的分布产生一定影响，但间隔时间仍主要分布在 t_0 和 t_1 附近.

表 1　针对 BER 分类的混淆矩阵

Tab. 1　Confusion matrix of BER classification

流量类型	预测为合法流量	预测为隐蔽信道流量
实际的合法流量	90	10
实际的隐蔽信道流量	2	98

(2) GAS. 分类模型在 GAS 信道流量上的实验结果如表 2 所示. 隐蔽信道的检测率为 96%，相对 BER 略有下降，这是因为 GAS 信道中的 Δt 为间隔时间赋予了一定的随机性. 但不可否认的是，GAS 的编码规则依然使得隐蔽流量的分布具有规律性，呈现出与合法流量不同的特征.

表 2　针对 GAS 分类的混淆矩阵

Tab. 2　Confusion matrix of GAS classification

流量类型	预测为合法流量	预测为隐蔽信道流量
实际的合法流量	93	7
实际的隐蔽信道流量	4	96

(3) CAB. 如表 3，我们的模型在 CAB 信道上的检测率为 91%，这比其他隐蔽信道上的检测率都低，但也在我们意料之中. 因为 CAB 的间隔值选取自合法流量，这导致该信道的数值特性与合法流量是相似的. 在两个不相交集合中无重复的取值会让间隔的随机性较大，相对其他隐蔽信道来说更难检测.

表 3　针对 CAB 分类的混淆矩阵

Tab. 3　Confusion matrix of CAB classification

流量类型	预测为合法流量	预测为隐蔽信道流量
实际的合法流量	90	10
实际的隐蔽信道流量	9	91

(4) ZAN. 表 4 展示了模型对 ZAN 信道的检测结果，结果显示模型对隐蔽信道和合法流量的检测率分别为 2%和 11%，这表明我们的模型可以检测出绝大部分 ZAN.

表 4　针对 ZAN 分类的混淆矩阵

Tab. 4　Confusion matrix of ZAN classification

流量类型	预测为合法流量	预测为隐蔽信道流量
实际的合法流量	89	11
实际的隐蔽信道流量	2	98

4.3　性能对比

为了验证我们提出的模型在检测性能上的优越性，我们选取了四种对比方法，分别是：经典的 Regularity 测试、基于 SVM 的模型、基于随机森林(RF)的模型和基于层次特征提取(Hier)的方法. 为保公正，我们复现了这四种方法并在同一批数据上进行了实验. 然后选取精度(Precision)、召回率(Recall)、$F1$ 分数($F1$-score)和误报率 RFP(Rate

of False Positive，RFP)作为评价指标.

Precision 也叫查准率，即被预测为隐蔽信道的样本中真正为隐蔽信道的样本比例，见式(8). 其中 TP 表示模型预测正确且类别为隐蔽信道的样本数，*FP* 表示模型预测正确且类别为合法流量的样本数. Precision 越高，证明模型预测为隐蔽信道的出错率越低.

$$Percision=\frac{TP}{(TP+FP)} \tag{8}$$

Recall 也叫查全率，即所有隐蔽信道样本中被预测出来为隐蔽信道的样本占比，如式(9). 为了系统安全考虑，我们希望尽可能多地检测出所有隐蔽信道. *Recall* 越高表明该方法越能够保证系统的安全.

$$Recall=\frac{TP}{(TP+FN)} \tag{9}$$

式(10)为 *F*1-score 的计算方式. *P* 代表精度(Precision)；*R* 代表召回率(Recall). 这两个指标是互相制约的关系，一味追求更高的精度会以牺牲召回率为代价，反之亦然. *F*1-score 综合考虑了两者，是评价一个模型好坏的关键指标. *F*1-score 越高，模型的性能越好.

$$F1\text{-}score=\frac{2\times P\times R}{(P+R)} \tag{10}$$

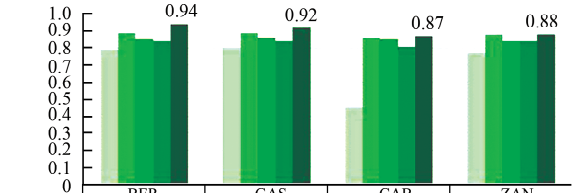
已知精度 *Precision*，可以计算误报率 *RFP*，如式(11). *RFP* 越低，模型的性能越好.

$$RFP=1-Precision \tag{11}$$

图 4～图 6 分别是 Precision、Recall、*F*1-score 三个评价指标下的对比实验结果. 其中不同深浅的颜色代表不同的方法，本文提出的方法在图中每一簇的最后一列，不同簇代表不同种类的隐蔽信道. 从图 4 中我们可以看出，在 Precision 指标上，我们的方法在四种隐蔽信道上的实验结果分别到达了 94%、92%、87% 和 88%. 除了 ZAN 信道上的 SVM 方法与我们的方法结果持平外，其余方法在该指标上的实验结果均低于我们提出的方法. 另外，在 *Recall* 和 *F*1-score 两个指标上，我们的方法也明显优于其他方法. 图 7 展示了我们的方法和对比方法的误报率情况. 从图 7 可以看到，我们的模型除了在 ZAN 信道上与 SVM 方法的误报率持平在 12%，其他情况下的误报率都低于对比方法.

以上实验结果表明，我们的模型性能优于其他对比方法，该模型具有通用性，可以检测出绝大部分的隐蔽信道，且误报率更低.

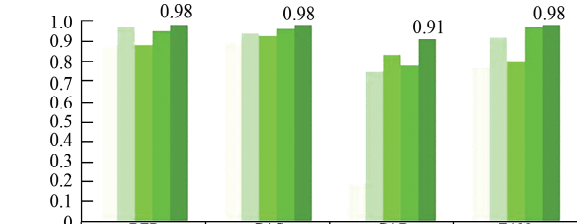
其次，本小节的 4 种对比方法均需手动提取流量特征. 这需要人工对每种隐蔽信道的报文时间间隔序列进行分析，总结多类隐蔽信道的共性，并选取最具表征能力的特征集，此过程将消耗较多时间. 而本文提出的基于图像的特征自动提取方法，将特征提取的过程交予 CNN 和 CBAM 注意力机制相结合的框架完成，这将节省约 40% 的人力成本.



	BER	GAS	CAB	ZAN
Regularity	0.79	0.8	0.45	0.77
SVM	0.89	0.89	0.86	0.88
RF	0.85	0.86	0.85	0.84
Hier	0.84	0.84	0.81	0.84
Proposed	0.94	0.92	0.87	0.88

图 4　隐蔽信道检测的精确率

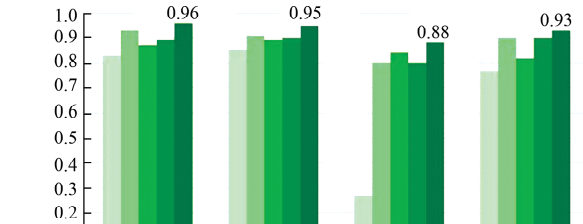
Fig. 4　Precision of covert channel detection



	BER	GAS	CAB	ZAN
Regularity	0.88	0.9	0.19	0.77
SVM	0.97	0.94	0.75	0.92
RF	0.88	0.93	0.83	0.8
Hier	0.95	0.96	0.78	0.97
Proposed	0.98	0.98	0.91	0.98

图 5　隐蔽信道检测的召回率

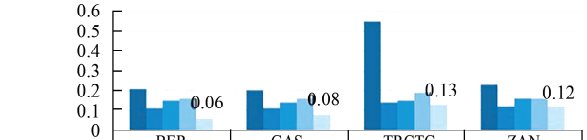
Fig. 5　Recall of covert channel detection



	BER	GAS	CAB	ZAN
Regularity	0.83	0.85	0.27	0.77
SVM	0.93	0.91	0.8	0.9
RF	0.87	0.89	0.84	0.82
Hier	0.89	0.9	0.8	0.9
Proposed	0.96	0.95	0.88	0.83

图 6　隐蔽信道检测的 *F*1-分数

Fig. 6　*F*1-score of covert channel detection



	BER	GAS	TRCTC	ZAN
Regularity	0.21	0.2	0.55	0.23
SVM	0.11	0.11	0.14	0.12
RF	0.15	0.14	0.15	0.16
Hier	0.16	0.16	0.19	0.16
Proposed	0.06	0.08	0.13	0.12

图 7　隐蔽信道检测的误检率

Fig. 7　Rate of false positive of covert channel detection

4.4　消融实验

为了评估基于灰度图像转化的模型对于隐蔽信道检测的有效性,我们将基线方法和基于灰度图像转化的分类方法做了比较. 基线定义为将一维卷积神经网络直接用于时间间隔序列. 我们选取了正确率(*Accuracy*)、召回率(*Recall*)、*F*1 分数(*F*1-score)三个指标来衡量模型的性能. 式(12)给出了 Accuracy 的计算公式:

$$Accuracy=\frac{TP+TN}{P+N} \tag{12}$$

其中,TP 和 TN 分别表示模型预测正确且样本类别为隐蔽信道和合法流量的个数;P 和 N 分别表示类别为隐蔽信道和合法流量的样本总个数. 有关 *Recall* 和 *F*1-score 的计算方式,上一节的式(9)、式(10)已经给出. 为消除实验的偶然性,我们进行了 10 次随机实验,并计算实验结果的平均值(%)和标准差,如表 5～表 7,表中的数字是指标平均值,括号中为标准差.

表 5　和基线方法的正确率对比结果

Tab. 5　Accuracy comparison with baseline method

Method	*Aeeuracy*			
	BER/%	GAS/%	CAB/%	ZAN/%
Baseline(1-D CNN)	94.1(0.016)	93.1(0.011)	85.5(0.026)	90.9(0.049)
2-D CNN	94.6(0.027)	93.5(0.018)	87.5(0.029)	91.8(0.012)
2-D CNN+CBAM	95.8(0.019)	94.6(0.018)	88.3(0.22)	92.3(0.007)

表 6　和基线方法的召回率对比结果

Tab. 6　Recall comparison with baseline method

Method	*Recall*			
	BER/%	GAS/%	CAB/%	ZAN/%
Baseline(1-D CNN)	94.5(0.007)	97.0(1.18e-16)	77.4(0.053)	88.7(0.107)
2-D CNN	96.6(0.037)	97.4(0.010)	87.3(0.065)	96.5(0.031)
2-D CNN+CBAM	98.3(0.036)	97.6(0.008)	99.9(0.022)	97.6(0.011)

表 7　和基线方法的 *F*1 分数对比结果

Tab. 7　*F*1-score comparison with baseline method

Method	*F*1-score			
	BER/%	GAS/%	CAB/%	ZAN/%
Baseline(1-D CNN)	94.1(0.015)	93.6(0.012)	84.1(0.032)	90.6(0.058)
2-D CNN	94.7(0.027)	93.6(0.014)	87.5(0.036)	92.1(0.011)
2-D CNN+CBAM	96.0(0.018)	95.0(0.018)	88.4(0.019)	92.7(0.007)

如表 5～表 7 的前两行所示,与基线模型相比,基于灰度图像转化的分类模型在 *Accuracy*、*Recall*、*F*1-score 三个指标上均有所提升. 效果最为显著的是,针对 CAB 类隐蔽信道的检测 *Recall* 从 77.4%提升到了 87.3%. 结果表明,将流量转为灰度图像可以有效地提升模型的检测性能.

为定量评估 CBAM 注意力模块在检测模型中的贡献,我们比较了单独使用二维卷积神经网络的模型和加入 CBAM 注意力模块后的模型,如表 5～表 7 的后两行所示. 加入 CBAM 注意力模块后,模

型对各隐蔽信道的检测效果均有不同程度的提升，针对最难检测的 CAB 信道，模型的 *Accuracy*、*Recall*、F_1-score 分别提升了 0.8%、3.6%和 0.9%. 根据定量比较的结果，我们可以合理推断：CBAM 注意力模块的加入可以有效提升模型的检测性能.

5　结　论

本文分析了现有的时间型隐蔽信道检测方法的不足，提出了一种将流量转换成灰度图像的处理方法. 该方法结合了 CNN 和 CBAM 注意力模块，解决了现有方法检测通用性不足、误检率高的问题. 我们的方法在与现有的四类方法比较中，表现出了更好的检测性能. 但由于现实场景中会出现更多未知的隐蔽信道，且训练数据可能不充分，因此，如何识别出未知的隐蔽信道，且在数据不充分的情况下让模型依然保持较好的检测性能，是我们未来要探索的方向.

参考文献：

[1] Zseby T, Vázquez F I, Bernhardt V, *et al*. A network steganography lab on detecting TCP/IP covert channels [J]. IEEE T Educ, 2016, 59: 224.

[2] Lampson, Butler W. A note on the confinement problem [J]. Commun ACM, 1973, 16: 613.

[3] 王晓琪，李强，闫广华，等. 高级持续性威胁中隐蔽可疑 DNS 行为的检测[J]. 计算机研究与发展，2017，54：2334.

[4] 张红，申滨，张燕，等. 基于卷积神经网络的宽带合作频谱感知[J]. 重庆邮电大学学报：自然科学版，2022，34：766.

[5] Zander S, Armitage G, Branch P. A survey of covert channels and countermeasures in computer network protocols [J]. IEEE Commun Surv Tut, 2007, 9: 44.

[6] 巫祺炜. 高级持续性威胁(APT)隐蔽信道的检测研究[D]. 上海：上海交通大学，2016.

[7] 国家质量技术监督局. 计算机信息系统安全保护等级划分准则：GB 17859-1999 [S]. 北京：中国标准出版社，1999.

[8] Jones R H, Goodrich J K, Sabiston D C. Information technology-security techniques-evaluation criteria for IT security-part1: introduction and general model [J]. Inf Technol Stand, 2009, 15: 598.

[9] 李彦峰，丁丽萍，吴敬征，等. 网络隐蔽信道关键技术研究综述[J]. 软件学报，2019，30：2470.

[10] Handel T, Sandford M. Hiding data in the OSI network model [C]//International Workshop on Information Hiding. Berlin: Springer, 1996: 23.

[11] Kundur D, Ahsan K. Practical Internet steganography: data hiding in IP [J]. Secur Inform Syst, 2003 (2): 10.

[12] Zander S, Armitage G J, Branch P A. Covert channels in the IP time to live field [C]//Australian Telecommunication Networks and Application Conference (ATNAC). Melbourne: University of Melbourne, 2006.

[13] Mileva A, Panajotov B. Covert channels in TCP/IP protocol stack-extended version [J]. Open Comput. Sci, 2014, 4: 45.

[14] Lucena N B, Pease J, Yadollahpour P, *et al*. Syntax and semantics-preserving application-layer protocol steganography [C]//International Workshop on Information Hiding. Berlin: Springer, 2004: 164-179.

[15] Berk V, Giani A, Cybenko G. Detection of covert channel encoding in network packet delays: TR2005536 [R]. New Hampshire: Department of Computer Science, Dartmouth College, 2005: 1.

[16] Gasior W, Li Y. Network covert channels on the android platform [C]//Proceedings of the Seventh Annual Workshop on Cyber Security and Information Intelligence Research. [S. l.]: ACM, 2011.

[17] Cabuk S. Network covert channels: Design, analysis, detection, and elimination [D]. Hammond: Purdue University, 2006.

[18] Zander S, Armitage G, Branch P. An empirical evaluation of IP time to live covert channels [C]//Proceedings of the 2007 15th IEEE International Conference on Networks. [S. l]: IEEE, 2007: 42.

[19] Cabuk S, Brodley C E, Shields C. IP covert timing channels: design and detection [C]//Proceedings of the 11th ACM Conference on Computer and Communications Security. [S. l.: S. n.], 2004: 178.

[20] Yang P, Zhao H, Bao Z. A probability-model-based approach to detect covert timing channel [C]//Proceedings of the 2015 IEEE International Conference on Information and Automation. Lijiang: IEEE, 2015: 1043.

[21] Rezaei F, Hempel M, Sharif H. Towards a reliable detection of covert timing channels over real-time network traffic [J]. IEEE T Depend Secure Comput, 2017, 14: 249.

[22] Gianvecchio S, Wang H. An entropy-based approach to detecting covert timing channels [J].

IEEE Trans Dependable Secure Comput, 2010, 8: 785.

[23] 张宇飞, 沈瑶, 杨威, 等. 差分信息熵的网络时序型隐蔽信道检测[J]. 软件学报, 2019, 30: 1.

[24] Shrestha P L, Hempel M, Rezaei F, *et al*. A support vector machine-based framework for detection of covert timing channels [J]. IEEE Trans Dependable Secure Comput, 2015, 13: 274.

[25] Li Q, Zhang P, Chen Z, *et al*. Covert timing channel detection method based on random forest algorithm[C]//Proceedings of the 17th International Conference on Communication Technology (ICCT). Chengdu: IEEE, 2017: 165.

[26] Darwish O, Al-Fuqaha A, Brahim G B, *et al*. Statistical hierarchical analysis with deep neural network-based framework for covert timing channels detection [J]. Appl Soft Comput, 2019, 82: 105546.

[27] Han J, Huang C, Shi F, *et al*. Covert timing channel detection method based on time interval and payload length analysis [J]. Comput Secur, 2020, 97: 101952.

[28] Wu S, Chen Y, Tian H, *et al*. Detection of covert timing channel based on time series symbolization [J]. IEEE Open J Commun Soc, 2021, 2: 2372.

[29] Nataraj L, Karthikeyan S, Jacob G, *et al*. Malware images: visualization and automatic classification [C]//Proceedings of the 8th International Symposium on Visualization for Cyber Security. [S. l.: s. n.], 2011: 1.

[30] Chen Z, Ke H, Jian L, *et al*. Seq2Img: A sequence-to-image based approach towards IP traffic classification using convolutional neural networks [C]//Proceedings of the International Conference on Big Data. Osaka: IEEE, 2017.

[31] Taheri S, Salem M, Yuan J S. Leveraging image representation of network traffic data and transfer learning in botnet detection [J]. Big Data Cognit Comput, 2018, 2: 37.

[32] He Y, Li W. Image-based encrypted traffic classification with convolution neural networks [C]//Proceedings of the 2020 IEEE 5th International Conference on Data Science in Cyberspace (DSC). Hong Kong: IEEE, 2020: 271.

[33] Al-Eidi S, Darwish O, Chen Y, *et al*. SnapCatch: automatic detection of covert timing channels using image processing and machine learning [J]. IEEE Access, 2020, 9: 177.

[34] Woo S, Park J, Lee J Y, *et al*. Cbam: convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision (ECCV). Munich: [s. n.], 2018: 3.

引用本文格式:

中　文: 陈虹吕, 王诗蕊, 李峰, 等. 基于灰度图像转化的时间型隐蔽信道检测方法[J]. 四川大学学报: 自然科学版, 2023, 60: 033003.

英　文: Chen H L, Wang S R, Li F, *et al*. Covert timing channel detection method based on grayscale image transformation [J]. J Sichuan Univ: Nat Sci Ed, 2023, 60: 033003.