计算机科学
COMPUTER SCIENCE

基于会话统计编码器的恶意加密流量检测方法研究

巩思越, 刘辉, 王宝会

引用本文

巩思越, 刘辉, 王宝会. 基于会话统计编码器的恶意加密流量检测方法研究[J]. 计算机科学, 2024, 51(11): 340-346.

GONG Siyue, LIU Hui, WANG Baohui. Malicious Encrypted Traffic Detection Method Based on Conversation Statistical Encoder Model [J]. Computer Science, 2024, 51(11): 340-346.

计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.231000121

基于会话统计编码器的恶意加密流量检测方法研究

巩思越　刘　辉　王宝会
北京航空航天大学软件学院　北京 100000
(gongsy@buaa.edu.cn)

摘　要　随着网络技术的发展和广泛应用,加密流量已成为保护用户隐私的关键技术。但同时,恶意软件和攻击者也利用加密流量来隐藏其行为,规避传统的网络入侵检测系统。现有的恶意加密流量检测方法存在一些问题,如基于统计特征的方法需要依赖专家经验进行特征提取,且不同协议的特征无法通用;基于原始输入的深度学习方法存在信息不完整和字段填充等数据问题,对加密流量交互行为的语义表征不足。为解决上述问题,提出了一种名为会话统计编码器模型(Conversation Statistic Encoder Model,CSEM)的方法。与传统的将字节流输入深度神经网络的模式不同,该方法借鉴了 transformer-encoder 模型,引入了一种新的流量包特征解析方式。所提方法能够针对每个流量包构建出固定长度的向量表示,并且无需进行零填充,同时避免了特征提取过程对具体加密协议的依赖,构建了一个混合深度神经网络,为恶意加密流量检测提供了一种新的思路。在 DataCon 和自建数据集上对所提模型进行了验证,其在 DataCon 公开数据集上的召回率达到了0.9911,精确率达到了0.9407,F1 值达到了 0.9652(相比随机森林模型 F1 值提升了 9%),几项指标均达到了目前的最佳水平。

关键词:会话;加密流量检测;编码器
中图分类号　TP312

Malicious Encrypted Traffic Detection Method Based on Conversation Statistical Encoder Model

GONG Siyue,LIU Hui and WANG Baohui
College of Software,Beihang University,Beijing 100000,China

Abstract　Abstract:With the development and widespread application of network technology,encrypted traffic has become a key technology for protecting user privacy. However,malware and attackers also use encrypted traffic to hide their behaviors and evade traditional network intrusion detection systems. Existing malicious encrypted traffic detection methods have some problems. Statistics-based methods rely on expert experience for feature extraction,and features of different protocols cannot be generalized. Deep learning methods based on raw inputs have incomplete information and field padding data issues,leading to insufficient semantic representation of encrypted traffic interactions. To solve the above problems,this paper proposes a method called "conversation statistic encoder model(CSEM)". The method draws on the transformer encoder model and introduces a new traffic packet feature parsing method,and it is different from the traditional mode of inputting byte streams into deep neural networks. The proposed method can construct fixed-length vector representations for each traffic packet without padding zeros,while avoiding dependence on specific encrypted protocols in the feature extraction process. A hybrid deep neural network is constructed to provide a new idea for malicious encrypted traffic detection. The proposed method is verified on the DataCon dataset and self-built dataset,and the experimental results on Datacon dataset show a recall of 0.9911,precision of 0.9407,and F1 score of 0.9652,reaching the current best level,and the F1 score is 9% higher than that of the random forest model.

Keywords　Conversation,Encrypted traffic detection,Encoder

1　引言

根据 2021 年上半年我国互联网网络安全监测数据分析报告[1],恶意程序样本约 2307 万个,日均传播约 582 万次,涉及恶意程序家族约 20.8 万个。国内感染计算机恶意程序的主机约 446 万台,同比增长 46.8%。基于思科 2020 年的 ESG 白皮书 Network Traffic Analysis(NTA):A Cybersecurity 'Quick Win'[2],目前有高达 80%的网络流量采用加密方式,其中 63%的威胁流量和 76%的关键/高风险威胁是在加密流量中被发现的,并且越来越多的高危问题是由高级持续性威胁(Advanced Persistent Threat,APT)引起的。APT 攻击通常由高度有组织的团队发起,使用加密流量进行通信,利用复杂定制的攻击手段,长时间秘密地渗透目标网络,从而达到数据窃取或系统破坏的目的。

传统检测方法主要采用基于载荷的深度包检测(Deep Packet Inspection,DPI)[3]。该方法需要解密加密流量,消耗

到稿日期:2023-10-18　返修日期:2024-03-07
通信作者:王宝会(wangbh@buaa.edu.cn)

大量资源，需要花费大量人力来维护规则库，无法自动更新，且引发数据隐私担忧。为规避对载荷进行解密，学术界提出了基于统计特征和基于原始输入的方法。基于统计特征是通过提取流量的统计特征进行识别，然而这类方法的特征提取仰赖于专家经验，缺乏对流量交互行为语义的挖掘，需要专家持续的分析更新。基于原始输入的方法则存在直接使用灰度图像表征信息密度低，加密信息不可解密；流量包长度变化大，会话包个数变化大，直接截断与补零会导致输入信息缺失或混淆等问题。

针对上述问题，本文创新性地结合加密流量会话过程特征和载荷统计特征，并通过 Transformer-Encoder 架构将其应用于恶意加密流量检测。该方法主要解决以下 3 个问题：

1）应用层加密信息难以表征，采用二进制表征信息密度低，且加密信息的不可解性为共识。

2）流量包长度和会话包个数变化大，需要合理构造输入特征。

3）加密流量攻击形式多样，许多研究仍基于握手特征判定；对非 SSL 加密协议支持有限，需训练多种模型进行判定。

本文提出了一种基于会话统计编码器模型的恶意加密流量检测方法。该方法基于 Transformer-Encoder 模型，使用统计特征和包基础特征将每个流量包描述为固定长度的向量，将每个向量视为一个 token，将一组会话视为一个序列，使用 encoder 模型训练，对流量会话进行特征表征，结合会话整体统计特征，构建混合深度神经网络，对恶意加密流量进行分类。

2 相关研究

使用机器学习或深度学习对恶意加密流量进行检测，根据模型输入信息的不同，一般来说有 3 种方式，如图 1 所示，分别为基于统计特征、基于原始输入和基于混合输入 3 种[3]。

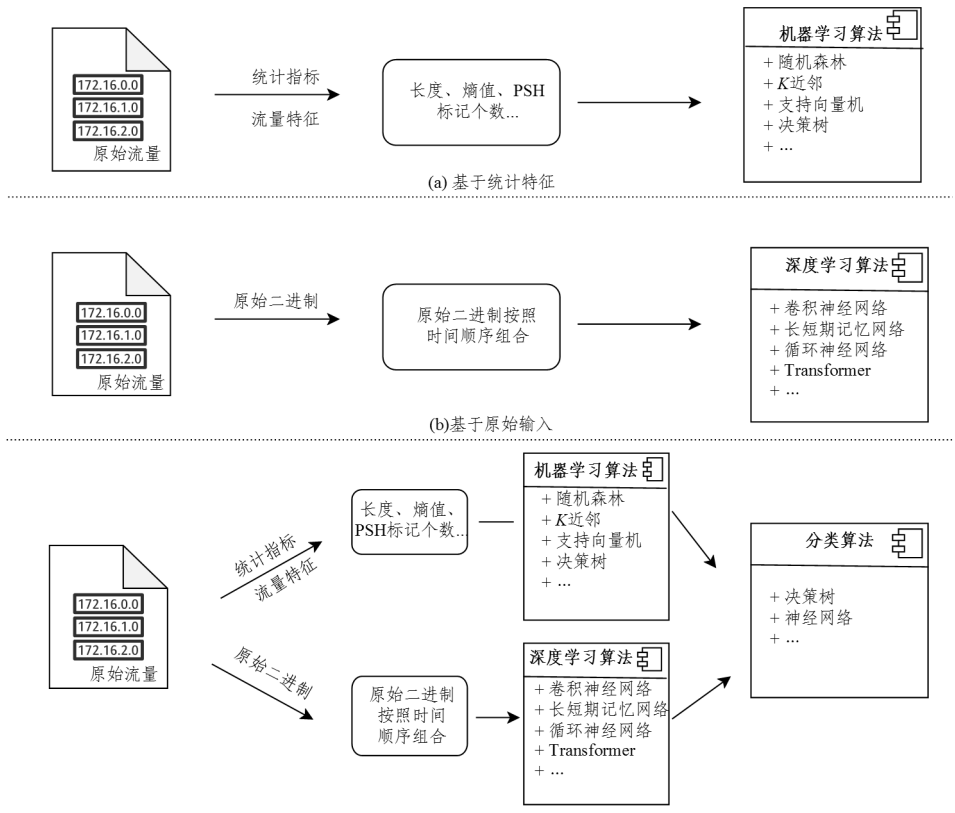


图 1 恶意加密流量常用检测方法

Fig. 1 Common methods for detecting malicious encrypted traffic

2.1 基于统计特征

在基于统计特征进行恶意加密流量识别领域，通常分为两种方法：一种以包级统计特征为输入进行识别，另一种以会话级统计特征为输入进行识别。在已有研究中，许多方法被提出以提高识别效果。例如，Fang 等[4]提出了一种基于随机森林的恶意加密流量识别方法，该方法结合了数据包信息、时间、TCP 标志字段和应用层有效载荷信息的特征；Khraisat 等[5]提出了基于 C5 决策树的 IDS 流量检测系统；Li 等[6]通过一系列机器学习策略提出了攻击检测算法，其中 SVM 用于构建分类器，并在 KDD Cup 数据集上取得了高达 98.624 9%的准确率；Lin 等[7]提出了聚类中心和近邻(CANN)方法；Ashkari 等[8-9]对流量的时序特征进行了广泛统计，总结出 86 维特征，研究发现随机森林在效果上表现最佳。

2.2 基于原始输入

鉴于深度神经网络在图像分类、文本翻译等任务上取得了显著进展，网络安全领域正积极探索将深度神经网络应用于网络流量攻击检测。学术界已开始研究利用深度神经网络进行特征提取并利用混合模型进行分类的算法，常用的网络架构包括 CNN，LSTM，Transformer 和 ResNet 等。

卷积神经网络方面，Wang 等[10]提出了一种基于一维卷积神经网络的端到端加密流量分类方法，将数据处理过程集中在一个端到端框架中，然后通过学习特征进行分类。Bazuhair 等[11]提出基于二维 CNN 的方法，提取前 784 字节的负载内容，并将其转化为 28×28 的图片，通过 LeNet-5 进行学习。Cheng 等[12]则提出了 RTETC 方法，利用前 3 个包的嵌入表示，采用 Multi-head Attention 和 1D CNN，提取流量包内部和流量包之间的交互。在 LSTM 方面，Zou 等[13]使用负载大小、时间间隔等作为包的表示，通过 LSTM 挖掘时序信息。在 Transformer 架构中，Lin 等[14]提出了 ET-BERT 流量表征模型，使用双字节表征法对流量包的 16 位进行编码。Zeng 等[15]使用 CNN，LSTM 和 SAE 模型，自动挖掘特征进行检测。

2.3 基于混合输入

基于混合输入的神经网络一般是网络结构中同时输入原始流量二进制流和会话统计特征。在混合输入的深度神经网络研究中，Bader 等[16]提出了一种方法，将原始负载与前 32 个包的行为序列作为输入，这些输入分别经过 1D CNN 和 GRU 模型处理；此外，基于流量方向与是否握手等特征，生成流量画像作为第三通道；通过 2D CNN 进行特征提取，再与前两个通道的输出进行融合。Gu 等[17]提出了一种基于多粒度表征学习的加密恶意流量检测方法。在字段级粒度上，他们利用词向量进行局部行为建模，并采用 Multi-head Attention 来计算字段间的交互作用，接着通过 BiLSTM 获得报文级语义信息。在包级粒度上，他们基于时空信息进行全局行为建模，提取包的时空状态，并运用 LSTM 模型获取流级语义。研究中将这两个粒度下的局部行为语义和全局行为语义融合，形成加密流量的综合表征。Wei 等[18]提出了 HNNIM (Hybrid Neural Network Identification Model)模型，用于识别与分类任务。如图 2 所示，特征提取分为两部分：一部分由深度神经网络自动挖掘；另一部分由专家根据经验挑选，随后再经由深度神经网络进一步筛选。

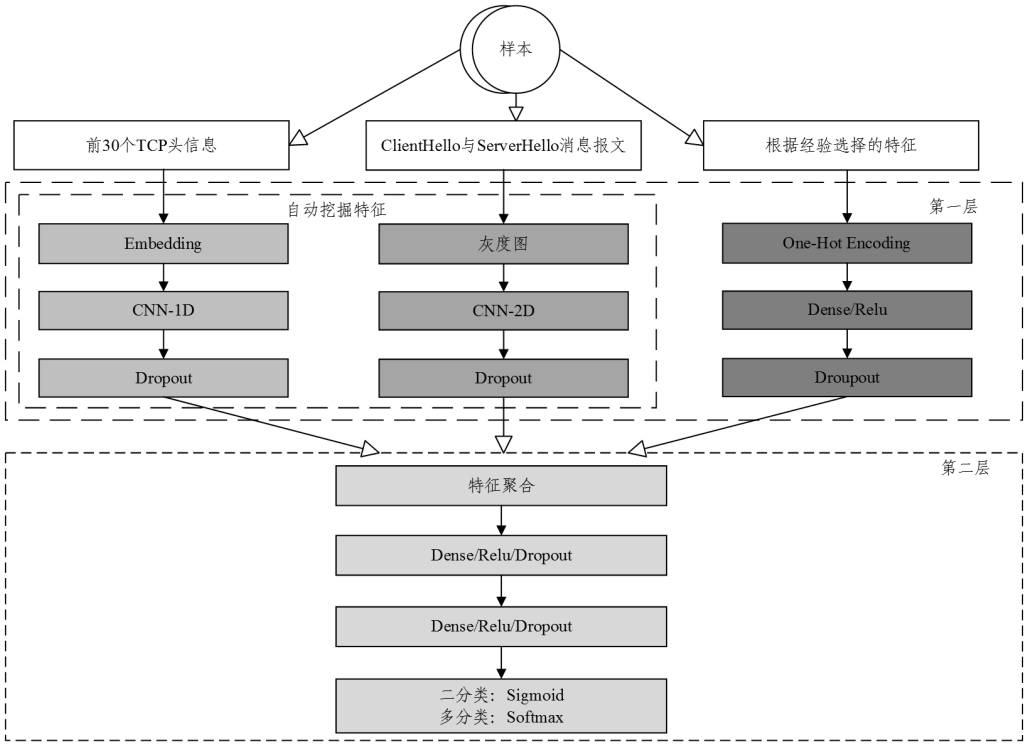


图 2 HNNIM 模型结构图

Fig. 2 Structure of HNNIM model

以上研究表明，机器学习和深度学习在检测恶意加密流量方面备受关注，特征提取方法已经发生显著变化。机器学习主要从网络会话中提取统计或协议专属特征；而深度学习则以字节为单位进行分析，将字节作为原始输入。混合神经网络将这两种方法结合，以更全面的方式分析流量数据。深度学习解决了加密流量特征提取的难题，如不准确和不完整，它通过引入网络安全领域知识提高了加密流量分类的准确性。然而，由于网络流量的属性，深度神经网络在字节截取和填充方面仍依赖经验。此外，在进行零填充时难以与非加密流量的零填充区分开，这些填充数据不仅浪费计算资源，还引入干扰。为了提升效果，一些研究专注于特定协议，提取其独特字段，但这种方法未充分发挥深度神经网络的优势。

因此，本文旨在基于统计特征构建 Transformer 编码器模型，利用包统计信息等特征，结合会话统计特征，无需依赖人工分析协议特征，从而提升加密流量的特征表征能力以及对恶意加密流量的识别能力。

3 建模方法

CSEM 模型整体流程如图 3 所示，分为 3 个阶段：数据预处理阶段、会话特征表征阶段、混合特征拼接 3 个阶段。该方法基于 Transformer-Encoder 模型[19]，使用载荷统计特征和包基础特征将每个流量包描述为固定长度的向量，将每个向量视为一个 token，将一组会话视为一个序列，使用 Encoder 模型训练，对流量会话进行特征表征，

同时在混合神经网络中结合会话整体统计特征，构建混合深度神经网络，对恶意加密流量进行分类。

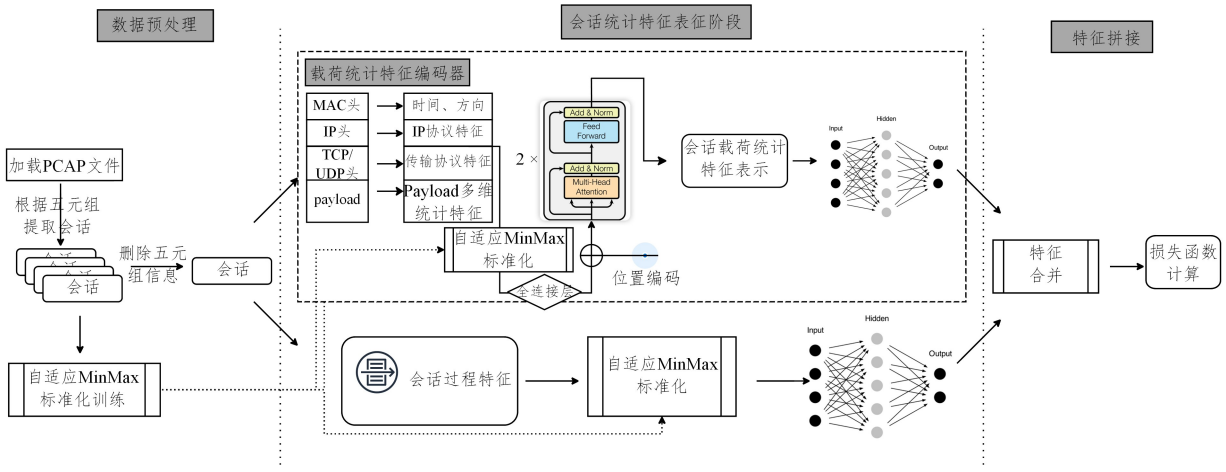


图 3 CSEM 恶意加密流量检测整体框架

Fig. 3 Framework of CSEM malicious encrypted traffic detection

3.1 数据预处理阶段

使用 Python 的 Scapy 库读取 PCAP 文件，通过五元组信息（源 IP 地址、目标 IP 地址、源端口号、目标端口号和传输协议）确定会话划分。将训练数据和测试数据按 7∶3 的比例划分为独立会话，并移除对应的源 IP 地址、目标 IP 地址、源端口号、目标端口号信息。从训练集中抽取十分之一的样本，采用本文提出的自适应 MinMax 标准化方法对会话过程特征和载荷统计特征进行标准化。

传统标准化方式一般有最大最小值标准化和正态标准化两种。针对统计特征和计数特征这种差异较小的特征，最大最小值标准化适用，但某些特征的极值之间差异较大，直接应用会将大量数据映射到极小的范围，不利于表征。为此，本文提出自适应 MinMax 标准化方法。具体而言，针对需要处理的特征，将其在数据集上的 95 分位数和 0 分位数设为最大值和最小值。若两者差异小于 1×10^6，则将数据集中大于或等于 95 分位数的值映射为该 95 分位数，然后进行最大最小值转换（如式(1)所示）。若 95 分位数和 0 分位数差异超过 1×10^6，则说明特征分布广泛，需先进行对数转换，再进行标准化。如果 $X_{\min}$ 小于 3，就令 $X_{\min^*}=X_{\min}+3$，以确保 ln 函数在定义域内，如式(2)所示。

$$X_{sc}=\frac{X_i-X_{min}}{X_{\max}-X_{\min}} \tag{1}$$

$$X_{sc}=\frac{\ln(X_i)-\ln(X_{\min^*})}{\ln(X_{\max})-\ln(X_{\min^*})} \tag{2}$$

3.2 会话特征表征阶段

本文采用混合神经网络，使用会话过程特征和载荷统计特征编码器对同一样本进行表征。会话过程特征表征借鉴 CICFlowmeter 思路[8]，根据会话过程长度、包个数、载荷长度、标记位个数、会话载荷熵等特征，统计会话的一般特征。

载荷统计特征编码器将 TCP/IP 协议体系中的网络层和传输层特征按协议定义拆分表示，结合应用层载荷统计特征，描述包携带信息。网络层、传输层根据协议定义，提取其中的特征，共计 25 维；TCP/UDP 层的载荷参考了中华人民共和国密码行业标准 GM/T0005-2012 随机性检测规范[20]与美国国家标准与技术研究院 NIST SP 800-22 随机数统计测试套件[21]，计算生成 19 维向量，其中包括长度、熵、平滑熵[22]、单比特频数统计、块内频数统计、扑克统计、重叠子序列统计、游程总数统计（两个统计量）、游程分布统计、块内最大“1”游程统计、二元推导统计、自相关统计、矩阵秩统计、累加和统计、近似熵统计、线性复杂度统计、Maurer 通用统计、离散傅里叶统计。对于这 19 个统计指标，每个统计量记为函数 F_i，将所有特征值拼接，从而构建 44 维特征向量。编码结构类似 NLP 中的序列预测，将每个包视为 token，多个 token 组成序列，输入编码器模型中，编码器模型经过训练，输出嵌入向量，具体过程见式(3)－式(8)。

$$\boldsymbol{X}_i=F_i(payload) \tag{3}$$

$$STAT(payload)=(\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3\cdots\boldsymbol{X}_{19}) \tag{4}$$

$$d(packet)=concat([\boldsymbol{direct},\boldsymbol{time},\boldsymbol{features},\boldsymbol{STAT}(\boldsymbol{payload})]) \tag{5}$$

$$\boldsymbol{head}_{ai}=\sum_{j=1}^{n}\mathrm{softmax}\left(\frac{Q_{aj}K_{ai}^{\mathrm{T}}}{\sqrt{d_k}}\right)V_{aj} \tag{6}$$

$$z_i=concat(\boldsymbol{head}\,1i\parallel\boldsymbol{head}\,2i\parallel\boldsymbol{head}\,3i)\cdot\boldsymbol{W}^o \tag{7}$$

$$out=LN(FFN(LN(\boldsymbol{x}_i+\boldsymbol{z}_i))+\boldsymbol{z}_i) \tag{8}$$

其中，Q,K,V，分别是 $d(packet)$经过不同的线性变换得到的。

将句子输入模型之前，已有的包序列构成了一个完整的向量。与自然语言处理领域不同，这里不需要使用 tokenizer 进行编码。相反，可以直接使用一个全连接层，将具有 44 维特征的向量转换为与 Encoder 隐藏层维度相匹配的向量，然后将其输入 encoder 模块，如图 3 所示。在 Transformer-Encoder 架构中，token 的顺序会对模型输出产生影响，因此位置编码包括两个方面：一个是入方向和出方向的编码，在每个网络包的首个特征位置上进行表示，使用 0 和 1 进行区分；另一个是网络流中包的先后顺序，我们继续使用 Transformer 中的位置嵌入，详见式(9)、式(10)。

$$PE_{(pos,2i)}=\sin(pos/10\,000^{2i/d_{\mathrm{model}}}) \tag{9}$$

$$PE_{(pos,2i+1)}=\cos(pos/10\,000^{2i/d_{\mathrm{model}}}) \tag{10}$$

模型结构有效地避免了加密流量会话在包级别使用零填充，而会话级别的零填充由于采用了 mask-attention 模块，模型不要求每个输入会话的长度都是固定的。在同一个 batch 内，只需规定一个最大长度，对于较短的会话，可以填充零向量

替代流量包，以达到最大长度。一旦 mask 字段设置完毕，注意力掩码结构将自动忽略这些零填充。

3.3 混合特征拼接阶段

得到了会话过程特征表征与载荷统计特征编码器表征之后，使用 projection 将双方的输出映射到同一个维度下。经本文测试，最终两种特征都是被映射到 2 维特征下，使用可学习参数 p 将其相加，最终 $out = p * out_1 + (1-p) * out_2$。然后使用 Cross-entropy 损失函数计算损失，进行模型训练。其中，projection 在本文中为一个含有 2 000 个神经元隐藏层的单层神经网络。

3.4 可解释性分析

本文的创新点之一就是将载荷统计特征输入模型，代替了直接将载荷二进制输入模型中。针对该创新点，本节论证使用载荷统计特征的可行性。

首先，现代加密协议加密的信息不可被解密目前还是共识。为了验证使用多种维度的统计量对加密流量的载荷信息进行表示的方案是否可行，设计如下实验：随机生成 XSS 攻击代码、中英文对话代码、音乐文件二进制切片、图片文件二进制切片等共计 10 200 条数据，使用 AES，DES，3DES，RSA 4 种算法将其转换为二进制，其中 AES，DES 和 3DES 是使用 openPGP 协议，密钥为随机生成。分别验证：1）统计指标之间是否具有相关性；2）统计指标是否能够区分不同加密协议。

使用 AES 加密算法加密的数据各个统计指标之间的相关性绝对值热力图如图 4 所示，可以发现大部分统计指标特征之间的相关性都低于 0.4。对比使用 AES 和 RSA 协议加密的数据的 4 种统计特征值的分布（如图 5 所示）可以发现，相同统计指标在不同加密协议下，统计值分布不同，有倾向性。总的来说，相同统计指标在不同加密协议下分布不同，相同协议在不同统计指标下无明显相关性，因此可以说使用本文所用的 19 种统计量对加密流量载荷部分进行描述是可行的。

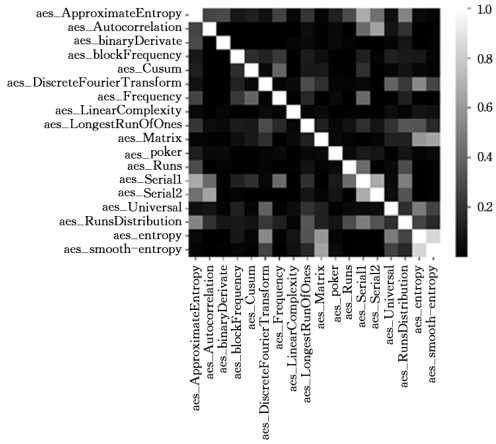
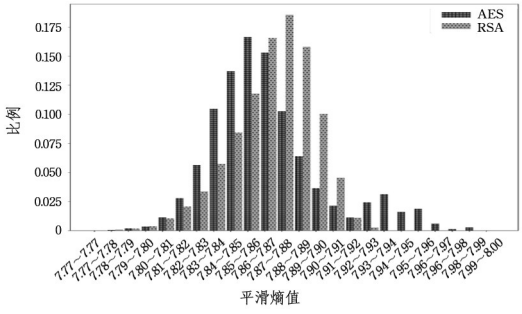
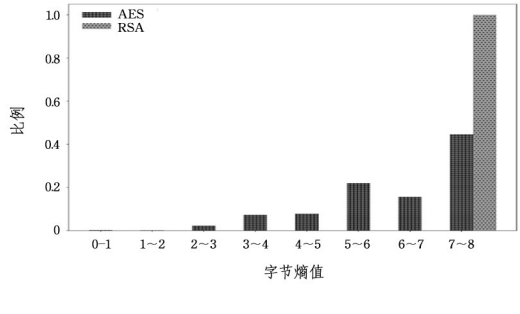


图 4 AES 统计指标相关性热力图

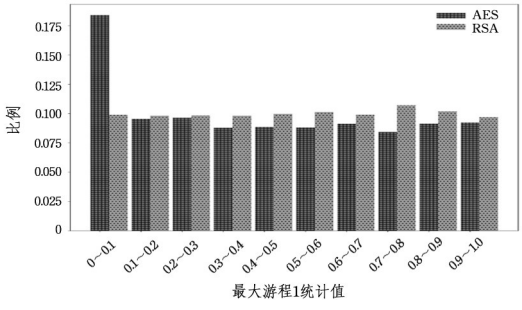
Fig. 4 Correlation heat map of AES statistical indicators



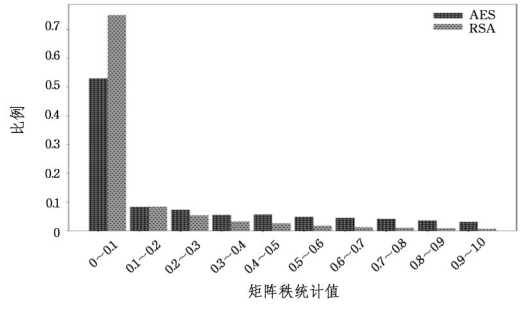
(a)对比 AES 与 RSA 平滑熵分布



(b)对比 AES 与 RSA 字节熵分布



(c)对比 AES 与 RSA 最大游程 1 统计值分布



(d)对比 AES 与 RSA 矩阵秩统计值分布

图 5 AES 与 RSA 加密数据统计特征分布对比

Fig. 5 Comparison of statistical feature distribution of AES and RSA encrypted data

4 实验设计

首先使用完整模型进行对比实验，与当前最先进的模型进行召回率、精确率、F1 值的对比，来验证模型的有效性。其次会进行消融实验，验证本文增加的载荷统计特征编码器的有效性和对传统统计特征模型的提升程度。程序运行在 CentOS 7 系统中，CPU 为 Intel(R) Xeon(R) Gold 6133 CPU@ 2.50GHz，GPU 为 NVIDIA GeForce RTX 4090，内存128GB，PyTorch 版本为 2.0.0+cu117。

4.1 数据集选择

本文主要使用的数据集有两个，一个是 DataCon2020-加密恶意流量数据集，其由 2020 年 2 月—6 月奇安信技术研究院天穹沙箱运行产生的流量筛选生成，流量内容为 443 端口产生的 TLS/SSL 数据包，将其拆分为会话，共计 94 256 条恶意加密流量会话，314 98 条正常加密流量会话[23]。另一个是本文生成的恶意加密流量数据集，正常流量包含访问网站、

微信聊天和 SSH 链接远程服务器等操作，协议为 SSH，HTTPS，SSL/TLS 的正常加密流量，共计 17 465 条；恶意流量包括使用菜刀、一句话木马的 WebShell 后门，CobaltStrike 模拟攻击内网机，以及通过 VMess 和 ShadowSockets 代理的恶意加密流量，共计 18 764 条。删除目的/源 IP、端口 4 个字段，使得模型无法根据 IP 或端口这种无法泛化的特征进行判断，从而能够真实地分析网络会话通过编码器生成的特征。

4.2 评价指标

本文主要以恶意样本的精确率、召回率和 F1 值作为评价指标。检测目的是区分恶意加密流量，这属于典型的二分类问题。使用恶意样本的精确率（见式(11)）和召回率（见式(12)）来衡量模型是否能够胜任实际工作。F1 值（见式(13)）作为召回率和精确率的平衡，显示预测的综合效果。

$$Precision=\frac{TP}{TP+FP} \tag{11}$$

$$Recall=\frac{TP}{TP+FN} \tag{12}$$

$$F1=2\times\frac{Precision*Recall}{Precision+Recall} \tag{13}$$

4.3 对比实验

在对比实验中，由于网络协议在不断进步，如果模型选择对加密协议的明文或某些特征字段进行指纹匹配，有可能导致其被误导，因此不应该匹配有关 IP 和端口的信息。在 DataCon数据集和自建数据集上，去除掉会话源和目的 IP、端口信息后进行训练预测，实验如表 1 所列。

表 1 对比实验

Table 1 Comparative experiments

model	DataCon			自建数据集		
	Precision	*Recall*	*F*1	*Precision*	*Recall*	*F*1
RF[4]	0.7846	0.9933	0.8767	0.7923	0.9792	0.8759
C5[5]	0.7447	0.9648	0.8406	0.7562	0.9602	0.8461
2D CNN[11]	0.8365	0.9883	0.9061	0.8715	0.9301	0.8998
ID CNN[10]	0.8485	0.9868	0.9124	—	—	—
RTETC[12]	0.8307	0.9917	0.9041	—	—	—
LSTM[13]	0.8149	0.9841	0.8916	0.8421	0.9371	0.8871
MaIDIST[16]	0.8600	0.9893	0.9201	—	—	—
ET-Bert[14]	0.9306	0.9833	0.9562	0.9388	0.9689	0.9536
CSEM	**0.9407**	**0.9911**	**0.9652**	**0.9425**	**0.9894**	**0.9654**

DataCon 数据集中的黑白流量均是加密流量。从表中可以发现，仅基于统计特征输入的模型效果一般会劣于基于原始输入的模型效果，如 CNN 和 LSTM 等；加入了深度学习算法的模型能力一般都有所提升。这主要是因为基于统计输入的模型往往只能关注到比较明显的部分，部分模型的召回率极高但是精确率较差，导致 F1 值较低。在不增加特定协议数据集的检测方法中，CSEM 效果最好，泛化性更强；ET-Bert 作为类似架构的模型，效果劣于 CSEM。猜测有两方面的原因，其一是该模型对会话预测时仅使用前 5 个流量包的原始输入的编码映射，会话特征表征不完全；其二是自定义加密流量载荷字节数较多，而 ET-Bert 对流量包载荷的输入进行了截断，长度仅为 768，获取信息不如 CSEM 全面。

推理效率方面，上述机器学习模型推理一般分为生成特征和模型推理两步。其中，基于统计特征的模型如随机森林，在模型推理阶段的预测效率一般是最高的；而基于原始输入的深度学习的模型在模型推理阶段速度较慢，本文实现的 CSEM 在模型推理阶段使用是深度学习模型，推理阶段与 ET-Bert 类似，效率无明显差异。但是在生成特征阶段，由于 CSEM 需要对加密流量载荷部分进行统计计算，本文使用了 Python 进行实现，而没有使用 C++进行实现，因此在生成特征阶段十分缓慢，对 DataCon 数据进行预处理大概需要花费 23h，故本文不进行预测效率对比，后续应用在生产环境推理时，会使用 C++重写载荷统计特征计算部分。

4.4 消融实验

为了验证会话统计编码器是否真的能够提升模型对会话的表征能力，在 DataCon 数据集上进行消融实验。消融实验会从以下 3 个方面进行：1）仅使用会话 encoder 进行训练；2）仅使用会话统计进行训练；3）使用混合特征进行训练。

如图 6 所示，单独使用会话 encoder 训练或仅使用会话统计进行训练，模型的波动都比较大，从 F1 值可以看出，随着训练轮数的增加，会话 encoder 的效果更好，可见会话 encoder 的表征能力更强。而如果使用完整模型进行训练，会使得在验证集上的 F1 值较为稳定地提升，而不是像仅使用 encoder 模型时，波动较大。这部分在整个实验过程中的表现都比较明显，如果仅使用会话 encoder 进行训练，波动明显的原因可能是 encoder 模型参数相对较多，训练收敛较为困难，而由会话统计特征进行辅助较为容易收敛。

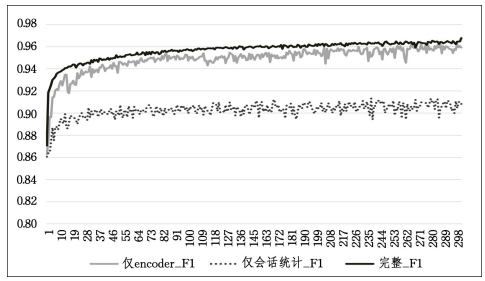


图 6 消融实验

Fig. 6 Ablation experiments

4.5 超参选择

本文评估 encoder 超参数的优劣主要通过模型训练的稳定性，而不是训练中某个时刻模型的预测效果。这主要是由于使用参数量较大的模型时，训练容易出现抖动或陷入某个区间最优解中，故 *encoder* 最后设置为 $head=8$，$hidden\_size=128$，堆叠层数 $layer=2$。

projection 最后选择使用单层 MLP，隐藏层大小为 2000。sum 层使用动态学习参数，两个 projection 输出权重和为 1，encoder 输入序列长度最大为 25，为 DataCon2020-加密恶意流量数据集中会话包个数的 90 分位数。学习率选择 1×10^{-4}，优化器选择为 Adam，$batch\_size=1\,024$，训练 *epoch* 为 300。

结束语 本文提出了基于会话统计编码器的恶意加密流量分类检测方法，使用统计特征描述包的 payload 特征，结合网络层与传输层基本特征，构建包特征向量，使得包特征向量长度固定，且不需要依靠 0 填充等手段，结合会话过程特征，训练后对恶意加密流量进行检测。实验结果显示，该模型的召回率达到了 0.9911，精确率达到了 0.9407，F1 值达到了 0.9652，相比随机森林模型 F1 值提升了 9%，达到了 SOTA 效果。

后续会从以下3个方向进行进一步的研究：1)完善统计特征部分，测试分析每个统计特征的作用和效果，并且新增合适的新的统计特征，并使用C++重写载荷统计特征提取；2)构建大规模数据集，训练出可以应用于真实业务场景的模型；3)尝试应用无监督学习，如对比学习，使得会话encoder模型可以应用在0-shot和few-shot任务上。

参 考 文 献

[1] CNCERT. Analysis Report on China's Internet Network Security Monitoring Data in the First Half of 2021 [EB/OL]. (2021-07-31) [2023-08-15]. https://www.cert.org.cn/publish/main/46/2021/20210731090556980286517/20210731090556980286517_.html.

[2] JON O. Network Traffic Analysis (NTA): A Cybersecurity 'Quick Win' [EB/OL]. [2023-08-15]. https://www.cisco.com/c/dam/en/us/products/collateral/security/stealthwatch/stealthwatch-esg-wp.pdf.

[3] LI Y, GUO H, HOU J, et al. A Survey of Encrypted Malicious Traffic Detection[C]// 2021 International Conference on Communications, Computing, Cybersecurity, and Informatics. IEEE-Computer Society, 2021: 1-7.

[4] FANG Y, XU Y, HUANG C, et al. Against malicious SSL/TLS encryption: identify malicious traffic based on random forest [C]// Fourth International Congress on Information and Communication Technology. Springer, 2020: 99-115.

[5] KHRAISAT A, GONDAL I, VAMPLEW P. An anomaly intrusion detection system using C5 decision tree classifier[C]// Pacific-Asia Conference on Knowledge Discovery and Data Mining. Springer, 2018: 149-155.

[6] LI Y, XIA J, ZHANG S, et al. An efficient intrusion detection system based on support vector machines and gradually feature removal method[J]. Expert Systems with Applications, 2012, 39(1): 424-430.

[7] LIN W, KE S, TSAI C. CANN: An intrusion detection system based on combining cluster centers and nearest neighbors[J]. Knowledge-based Systems, 2015, 78: 13-21.

[8] ASHKARI A H. CICFlowmeter-V4.0 (formerly known as ISCXFlowMeter) is a network traffic Bi-flow generator and analyser for anomaly detection [EB/OL]. [2021-07-05]. https://github.com/ahlashkari/CICFlowMeter.

[9] ASHKARI A H, DRAPER-GIL G, MAMUN M S I, et al. Characterization of tor traffic using time based features[C]// International Conference on Information Systems Security and Privacy. 2017: 253-262.

[10] WANG W, ZHU M, WANG J, et al. End-to-end encrypted traffic classification with one-dimensional convolution neural networks[C]// 2017 IEEE International Conference on Intelligence and Security Informatics. IEEE, 2017: 43-48.

[11] BAZUHAIR W, LEE W. Detecting malign encrypted network traffic using perlin noise and convolutional neural network [C]// 2020 10th Annual Computing and Communication Workshop and Conference. IEEE, 2020: 200-206.

[12] CHENG J, HE R, E Y P, et al. Real-time encrypted traffic classification via lightweight neural networks[C]// GLOBECOM 2020-2020 IEEE Global Communications Conference. IEEE, 2020: 1-6.

[13] ZOU Y, ZHANG J, JIANG B. Detection Of Malicious Encrypted Traffic Based on Lstm Recurrent Neural Network [J]. Computer Applications and Software, 2020, 37(2): 308-312.

[14] LIN X, XIONG G, GOU G, et al. ET-BERT: A Contextualized Datagram Representation with Pre-training Transformers for Encrypted Traffic Classification[C]// Proceedings of the ACM Web Conference 2022. 2022: 633-642.

[15] ZENG Y, GU H, WEI W, et al. a deep learning based network encrypted traffic classification and intrusion detection framework[J]. IEEE Access, 2019, 7: 45182-45190.

[16] BADER O, LICHY A, HAJAJ C, et al. MalDIST: From encrypted traffic classification to malware traffic detection and classification[C]// 2022 IEEE 19th Annual Consumer Communications & Networking Conference. IEEE, 2022: 527-533.

[17] GU Y H, XU H, ZHANG X Q. Encrypted malicious traffic detection based on multi-granularity characterization learning[J]. Journal of Computing, 2023, 46(9): 1888-1899.

[18] WEI J H, ZHENG R F, LIU J Y. Research on malicious TLS traffic identification based on hybrid neural network[J]. Computer Engineering and Applications, 2021, 57(7): 107-114.

[19] DEVLIN J, CHANG M, LEE K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding [J]. arXiv: 1810.04805, 2018.

[20] 国家密码管理局. 随机性检测规范[EB/OL]. (2021-10-19) [2023-08-15]. https://std.samr.gov.cn/hb/search/stdHBDetailed?id=E66CC4F6F8D78B7FE05397BE0A0A6C55.

[21] ANDREW R, JUAN S, JAMES N, et al. SP 800-22 Rev. 1a, A Statistical Test Suite for RNGs and PRNGs for Crypto Apps | CSRC [EB/OL]. (2010-04-01) [2023-07-16]. https://csrc.nist.gov/publications/detail/sp/800-22/rev-1a/final.

[22] CACHIN C. Smooth entropy and Rényi entropy [C]// International Conference on the Theory and Applications of Cryptographic Techniques. Springer, 1997: 193-208.

[23] DataCon 社区. DataCon 开放数据集-DataCon2020-加密恶意流量数据集方向开放数据集[EB/OL]. (2021-11-11) [2023-08-15]. https://datacon.qianxin.com/opendata/openpage?resourcesId=6.

GONG Siyue, born in 1996, postgraduate. His main research interests include natural language processing and malicious traffic detection.

WANG Baohui, born in 1973, Ph.D, professor. His research interests include big data, artificial intelligence and network information security.

(责任编辑：柯颖)