声学所提出基于字嵌入的可识别长流的分层注意力模型

作者：writer 来源：中国科学院

本文原地址：https://www.iikx.com/news/progress/6247.html

声学所提出基于字嵌入的可识别长流的分层注意力模型。在网络安全研究中，基于深度学习的入侵检测方法因具有较强的检测能力而受到越来越多的关注。然而，大部分基于深度学习的入侵检测方法处理长度过长的网络流量数据时能力不足，它们选择只处理流量的包头部分，忽略流量载荷中有价值的信息，因此当黑客把攻击行为隐藏在流量的载荷中时，这些入侵检测方法就无法有效检测到恶意流量。

中国科学院声学研究所国家网络新媒体工程技术研究中心博士生韩陆超等人提出了一种能够检测不同长度流量的注意力模型，以检测基于流量载荷的恶意流量;同时设计了一种基于生成式对抗网络(Generative Adversarial Networks, GAN)的流量生成模型，可以从原始数据集生成新的网络流量数据，以增强数据的安全性并保护用户隐私。相关研究成果6月24日在线发表于国际学术期刊IEEE Access。

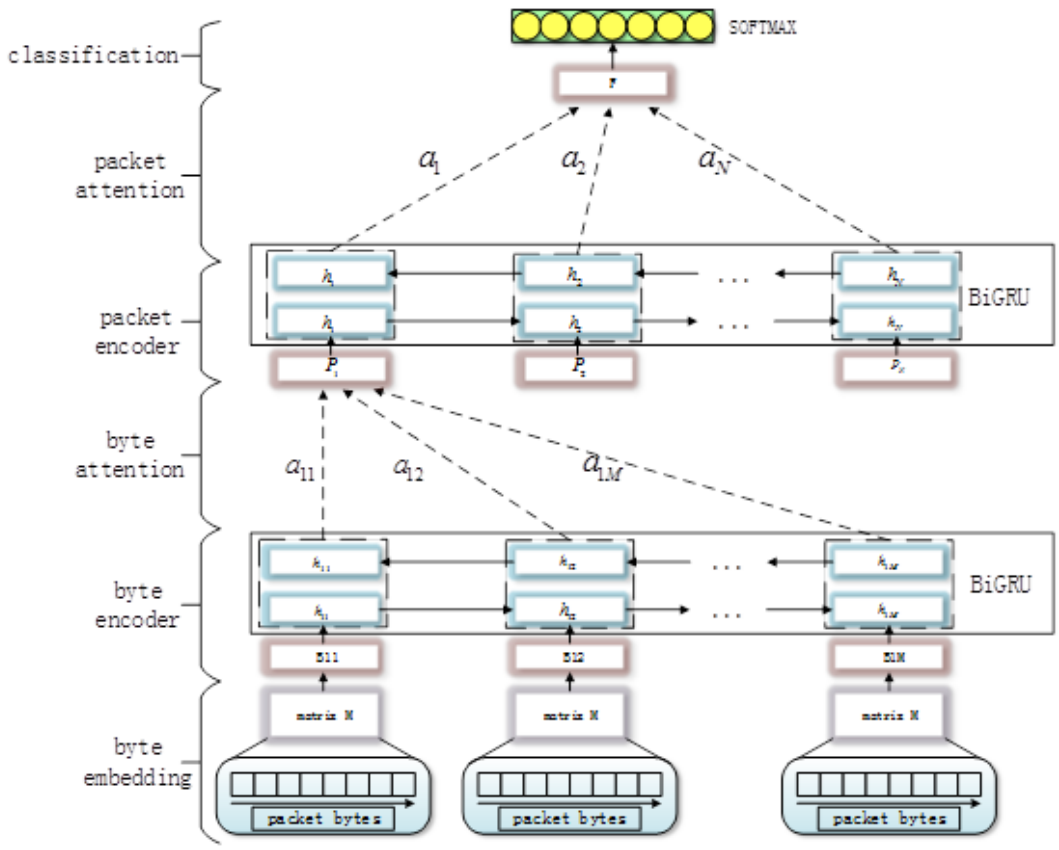
研究人员提出的分层注意力模型，可以从字节和数据包这两个层面学习流量信息。该模型使用双向GRU(Gated Recurrent Unit)构建字节表示，并通过注意机制给不同的字节分配不同的权重，一些与分类目标直接相关的关键字节在编码过程中被赋予更多权重。数据包表示的构建与此类似，最后使用注意力机制汇总构建整个TCP(Transmission Control Protocol)流的表示向量。

在入侵检测研究中经常遇到缺乏流量数据的问题，特别是在深度学习方法中，训练数据的局限性严重限制了模型的训练效果。此外，直接检测现实用户的网络流量可能会侵犯用户隐私。

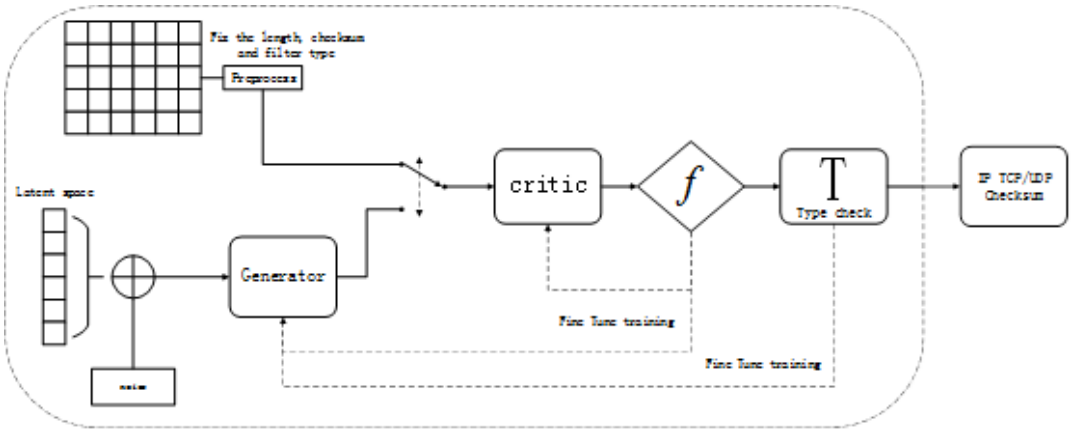
研究人员提出了Flow-WGAN(Wasserstein GAN)流量生成模型，从原始数据集中生成新数据。这种模型的结构和提取信息的方法与分类器不同，因此可以从同一原始训练集中学习新的特征并获得具有全新数据的网络流数据包。研究人员用此数据包来模拟新的网络应用流量类型，以评估分类器的性能或改进分类器。

基于ISCX-2012和ISCX-2017数据集的实验结果表明，与其他四种先进的深度学习方法相比，该分层注意力模型在准确性和真阳性率(true positive rate，TPR)方面具有更高的性能，且该模型在检测生成的数据包时所需训练时间比当前最先进的HSAT-IDS恶意流量检测模型减少30%。

论文链接



分层注意力模型的结构图(图/中科院声学所)



流量生成模型的原理图(图/中科院声学所)

更多 科学进展 请访问 https://www.iikx.com/news/progress/