A method for detecting abnormal traffic in working condition networks based on the fusion of information flow and state flow

Zhongjie Xiao

Science and Technology Information Department, Shanghai International Port (Group) Co., Ltd., Shanghai, 200080, China

Abstract

In the abnormal detection of working condition networks, based on the principle of information shortest path, the labels generated after the input of abnormal flow nodes lack complete node information, resulting in deviation in the calculation of abnormal scores, significant fluctuations in Acc values, and reduced detection accuracy. Therefore, a method for detecting abnormal traffic in working condition networks based on the fusion of information flow and state flow was designed. Based on the principles of information flow random propagation and shortest path propagation, establish an abnormal condition detection network that integrates information flow and state flow, and obtain the average probability of nodes traversing the detection network. Define the attribute set of the abnormal working condition detection network, and under the condition of information flow and state flow fusion, perform neighbor weighting on the attribute set to assign network abnormal detection labels. Based on the information status of upstream traffic, downstream traffic, data packets, etc., calculate the abnormal score of the abnormal traffic detection label in the working condition network, in order to achieve accurate detection of abnormal traffic in the working condition network. The final detection results showed that when the number of iterations reached 60, the Acc values of DoS Golden Eye, FTP Datator, DDoS and other datasets tended to stabilize, and the Acc value was greater than 90%. The accuracy of abnormal traffic detection was high, which played an important role in improving network security.

Keywords

Integration of information flow and state flow; Working condition network; Network abnormality; Abnormal traffic; test method

基于信息流和状态流融合的工况网络异常流量检测方法

肖中杰

上海国际港务（集团）股份有限公司科技信息部，中国·上海 200080

摘　要

工况网络异常检测中，依据信息最短路径原理，异常流量节点输入后生成的标签缺失节点信息，导致异常分数计算偏差，Acc值大幅波动，降低检测准确性。因此，设计了基于信息流和状态流融合的工况网络异常流量检测方法。根据信息流随机传播与最短路径传播原理，建立信息流与状态流融合的异常工况检测网络，获取节点遍历检测网络的平均概率。定义异常工况检测网络的属性集合，在信息流和状态流融合条件下，对属性集合进行邻居加权，分配网络异常检测标签。根据上行流量、下行流量、数据包等信息状态，计算工况网络异常流量检测标签异常分数，从而实现工况网络异常流量的精准检测。最终的检测结果显示，在迭代次数达到60次时，DoS Golden Eye 、FTP-Patator 、DDoS 等数据集的Acc值趋于稳定，且Acc值大于90%，异常流量检测的准确性较高，对于提高网络安全具有重要作用。

关键词

信息流和状态流融合；工况网络；网络异常；异常流量；检测方法

1 引言

工况网络是描述与监控特定设备运行状态的网络，通过收集与传输设备状态数据，能够实现对设备的运行管理[1]。在工况网络数据传输的过程中，受到设备故障、网络攻击、数据错误等影响，出现突发性、持续性、方向性等异常流量，可能会出现占用网络资源，引起设备故障，导致信息泄露等问题，影响网络数据的安全传输。针对此类问题，研发了多种异常流量检测方法。

李冲等提出了基于异常检测和 AUV 辅助的工况网络异常流量检测方法[2]。利用 RERA 协议，建立异常网络节点筛选机制，异常节点只转发时延敏感数据，时延不敏感的数据使用 AUV 辅助，通过节点转发数据类型，检测出异常节

【作者简介】肖中杰（1988–），男，中国河南信阳人，本科，工程师，从事网络安全、数据安全、工控安全研究。

点位置，从而实现异常流量检测。但是，该方法收集的数据需要进行有效的融合与分析，可能出现数据格式不兼容，降低数据质量的问题，导致数据集的 Acc 值大幅度波动，降低了异常流量检测的准确性。麻胜兰等提出了基于二维卷积神经网络的工况网络异常流量检测方法 [3]。利用二维卷积神经网络，检测结构异常的数据，并通过二维桁架数值模型，区分数据异常类型。并采用 5 层卷积层的二维网络，检测出流量异常情况。但是，该方法在网络流量数据分布发生较大变化时，泛化能力可能受到限制，从而出现数据集的 Acc 值大幅度波动的问题，影响异常流量检测的准确性。李贺等提出了基于图卷积自编码器的工况网络异常流量检测方法 [4]。利用 AMEAN 拆分出工况网络流量检测的多视图，并利用属性网络中的语义信息，对图卷积网络视图进行处理。引入自编码器模块，重构网络节点，异常程度越大，节点属于异常流量的可能性越大。但是，该方法的检测输出结果缺乏直观的解释性，检测性能受到限制，容易出现数据集的 Acc 值大幅度波动的情况，无法保证异常流量检测的准确性需求。王坤等提出了基于深度学习的工况网络异常流量检测方法 [5]。部署云端服务器的判别器与控制器，生成异常流量检测的对抗网络，利用异常流量训练样本，并在控制器上生成具有独立检测功能的检测代理，从而实现网络异常流量检测。但是，该方法对输入数据质量的依赖较大，一旦数据存在原始噪声或不平衡的情况，将会引入额外的检测误差，出现数据集的 Acc 值大幅度波动的问题，导致异常流量检测的准确性不高。

信息流和状态流融合，是一种数据处理技术，能够将网络中的运行信息、状态信息提取出来，找出网络潜在问题 [6]。本文将信息流和状态流融合，应用在工况网络异常流量检测方法中，获得更加全面、准确的工况网络状态信息，更加准确地检测出异常流量，实现工况网络的安全管理。

2 工况网络信息流和状态流融合异常流量检测方法设计

2.1 建立信息流与状态流融合的异常工况检测网络

工况网络较为复杂，网络中节点相互影响，实现信息交流 [7]。一般情况下，网络中的信息流，通过最短节点路径传播，节点的介数衡量了网络通过节点的流量大小。也就是说，信息流越大的节点，在工况网络中越重要，一旦该节点出现异常，将会影响网络性能。在实际情况下，网络中的信息普遍为随机传播的状态，由此建立异常工况检测网络，如图 1 所示。

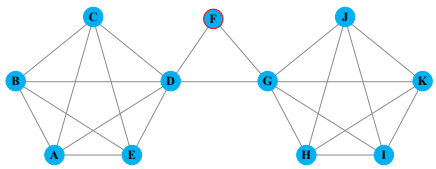

图 1 异常工况检测网络

如图 1 所示，A~K 为检测网络中的核心节点，连接左侧五边形的节点与右侧五边形的节点，最短路径均通过节点 D、G，二者之间存在一个很高的介数值 [8]。F 节点几乎未被经过，F 的介数存在一个很低的值。在异常流量检测网络中，经常会出现忽视 F，重视 D、G 的情况，容易出现异常流量检测偏差 [9]。本文综合考虑了传输最短的信息流与随机传输的信息流，将 A 作为源节点，G 作为目标节点，传输工况网络的状态数据。在状态流条件下，得到 A~G 传输的状态信息，公式如下：

$$m_{ij} = a_{ij} / R_{ij} \tag{1}$$

$$M = D_t\left[\left(A - M_t\right)^{-1}\right] \cdot Sm_{ij} \tag{2}$$

式（1-2）中，m_{ij} 为 A~G 在信息流中，节点 i 传输的第 j 条状态信息，a_{ij} 为网络邻接矩阵元素；R_{ij} 为 i 、j 的度；M 为信息传输的随机矩阵；D_t 为网络节点在 t 时刻的工作状态、连接状态、性能参数等状态数据；M_t 为网络节点在 t 时刻传输的数据包大小、传输速率、协议类型等信息数据；S 为信息流随机行走的概率 [10]。将信息流与状态流融合，得到 A~G 经过 F 的平均概率，公式如下：

$$S_i = \begin{cases} +1 & i = A \\ -1 & i = F \\ 0 & otherwise \end{cases} \tag{3}$$

$$P_i^{(AF)} = \frac{1}{2}\sum_j S_i \left| M_A - M_F \right| \tag{4}$$

式（3-4）中，S_i 为 A~G 为随机行走的概率；$P_i^{(AF)}$ 为 A~G 经过 F 节点，且遍历检测网络的平均概率；M_A 为最短路径传播信息流的介数；M_F 为随机路径传播的信息流介数 [11]。将异常工况检测网络设定为信息流、状态流随机传播的状态，设定 $P_i^{(AF)} > 80\%$，确保在不同工况下，网络异常流量检测均能够遍历不同节点，确保网络流量异常检测的全面性。

2.2 基于信息流和状态流融合分配网络异常检测标签

在工况异常检测网络中，各个网络节点存在 5 种属性，将 $P_i^{(AF)}$ 作为信息传输路径概率向量属性，其余属性向量分别为信息流相关的异常状态向量、状态流相关的异常状态集合、权重向量、特征变换集合等方面 [12]。根据信息流、状态流的融合情况，定义异常工况检测网络的属性集合，表达式如下：

$$f = \left[k', n', \omega, sf, P_i^{(AF)}\right] \tag{5}$$

式（5）中，f 为异常工况检测网络的属性集合；k' 为信息流相关的异常状态向量；n' 为状态流相关的异常状态集合向量；sf 为特征变换集合；ω 为权重向量。将 f 在检测网络中进行邻居加权，确保异常流量在网络层检测输出的分布被规范化，加速网络异常流量检测效率 [13]。在检测网络中，A、G 节点位置已知，将其作为网络异常流量节点，定义其邻居集合，公式如下：

$$\begin{cases} N=\{f(p,q)\} \\ p\in[A-k,A+k] \\ q\in[G-k,G+k] \end{cases} \quad (6)$$

式（6）中，N 为 A、G 节点的邻居集合；p 为 A 的邻居节点行索引；q 为 G 的邻居节点列索引；k 为给定的整数值。对于每个 p 、q 而言，计算其相对权重，公式如下：

$$\alpha_{p,q}=N\left(-\frac{d\left((A,G),(p,q)\right)^2}{2\delta^2}\right) \quad (7)$$

式（7）中，$\alpha_{p,q}$ 为邻居节点相对权重；$d\left((A,G),(p,q)\right)$ 为 (A,G) 与 (p,q) 之间的欧几里得距离；δ 为控制权重分布的超参数 [14]。计算 f 中各个属性集合的相对权重，并进行归一化处理，得到最终的权重，公式如下：

$$\beta=\frac{\alpha_{p,q}}{\sum_{(p',q')\in N}\alpha_{p',q'}} \quad (8)$$

式（8）中，β 为各个属性集合的综合权重；$\alpha_{p',q'}$ 为 f 中所有属性的相对权重；(p',q') 为检测网络中节点的所有邻居。根据 β ，分配网络异常检测标签，公式如下：

$$V_i=\beta\times\frac{1}{sf}\sum_{i=1}^{n}\left(k'-m[i]\right)^2+P_i^{(AF)}\times\frac{\beta}{\omega}\sum_{i=1}^{n}\left(n'-v[i]\right)^2 \quad (9)$$

式（9）中，V_i 为网络异常检测标签；n 为异常标签数量；$m[i]$ 为节点 i 被分配到异常标签的概率；$v[i]$ 为节点 i 被分配到正常标签的概率。将整个检测网络划分成多个节点子集批次，通过 V_i 标注节点异常情况，从而提高工况网络大规模异常流量检测的效率。

2.3 计算工况网络异常流量检测标签异常分数

在工况网络中，网络信息较为复杂，包括不同工况的状态信息。通过异常流量标签标记的数据集异常程度中，无法确定流量样本的显著差异。异常分数能够量化数据集中，各个流量样本异常标签的异常程度。因此，在异常流量检测标签分配完成之后，计算各个标签的异常分数，判断异常流量检测的准确性 [15]。本文根据上行流量、下行流量、数据包等信息状态，计算异常分数，公式如下：

$$Score_c(V_i)=-\sum_{r=1}^{R}\frac{s_i^{(r)}}{R} \quad (10)$$

$$Score_s(V_i)=\sum_{r=1}^{R}\frac{l_i^{(r)}}{R} \quad (11)$$

$$Score(V_i)=\alpha Score_c(V_i)+(1-\alpha)Score_s(V_i) \quad (12)$$

式（10-12）中，$Score_c(V_i)$ 为网络标签 V_i 的上行流量与下行流量的异常分数；$s_i^{(r)}$ 为网络流量标签与其对应数据包的匹配分数；R 为网络检测标签对应的采样子图；$Score_s(V_i)$ 为标签 V_i 的网络结构异常分数；$l_i^{(r)}$ 为网络流量标签与其对应数据包的预测误差；$Score(V_i)$ 为网络流量检测标签的异常分数；α 为信息流与状态流融合的超参数。在网络流量检测的过程中，$-1<Score(V_i)<1$，$Score(V_i)$ 趋近于 0，网络流量检测标签为正常，与其对应的数据包为正常状态；$Score(V_i)$ 趋近于 ±1，网络流量检测标签为异常，与其对应的数据包为异常状态。根据 $Score(V_i)$ 的变化情况，检测网络异常情况，从而实现工况网络异常流量的精准检测。

3 实验分析

本次实验用计算机为 DELL 工作站，CUDA 核心数为 3584 个，编程语言为 Python3.7。选用 CICIDS 网络安全研究数据集，其内收集了 DoS、DDoS 等多种攻击类型的异常流量数据，与 HTTP 请求、FTP 传输等正常流量数据，能够验证基于信息流和状态流融合的工况网络异常流量检测方法的有效性。

3.1 选取实验数据集

CICIDS 数据集模仿了现实世界的网络环境，数据集中的流量特征包括持续时间、流量包数量、字节大小等流量属性，能够为网络异常流量检测提供数据支持。网络流量数据集，如表 1 所示。

表 1 网络流量数据集表

类别名称	描述	数量
BENIGN	良性流量数据，能够与异常流量进行对比分析	2273097
DoS Hulk	分布式拒绝服务攻击	231073
Port Scan	端口扫描攻击	158930
DDoS	分布式拒绝服务攻击	128027
DoS Golden Eye	DoS 攻击变种，利用特定漏洞攻击	10293
FTP-Patator	针对 FTP 服务的暴力破解攻击工具	7938
SSH-Patator	针对 SSH 服务的暴力破解攻击工具	5897
DoS slowloris	低速率 DoS 攻击	5796
DoS Slow http test	发送慢速 HTTP 请求消耗网络资源的攻击	5499
Bot	机器人网络中的个体机器	1966
Brute Force	暴力破解攻击	1507
XSS	跨站脚本攻击	652
Infiltration	渗透测试或入侵行为	58
Sql Injection	SQL 注入攻击	34
Heartbleed	严重的安全漏洞	18

如表 1 所示，实验从原始的 pcap 数据中，提取了 DoS Hulk、Port Scan、DDoS、DoS Golden Eye、FTP-Patator、XSS、Infiltration、Sql Injection、Heartbleed 等数据类型作为异常网络流量数据，BENIGN 数据类型，作为正常网络流量数据。

3.2 网络流量标签异常分数检测结果

对于工况网络流量检测而言，将异常流量以标签的形式标注出来，并计算其异常分数，能够提高检测的准确性。本文以 DoS Hulk 为例，计算该数据集的 *Score*(*Vi*)。

本文采用红色框，表示 *Score*(*Vi*) 的正常区间；红色圆点表示 *Score*(*Vi*) 的异常情况。在 0~800s 的检测时间中，*Score*(*Vi*) 超出正常区间的异常流量数据均被标记，检测结果较为准确。

3.3 网络异常流量 ROC 曲线检测结果

为进一步分析工况网络异常流量检测情况，采用 ROC 曲线，描绘不同工况下，网络异常流量的真阳性率 TPR，假阳性率 FPR 之间的关系，直观地展示检测方法的检测性能。ROC 曲线作为最佳分类阈值，能够在网络中找出最佳的平衡点，将正常流量数据与异常流量数据有效分开，从而判断检测的准确性。在高负载工况、高实时性工况、强加密工况、分布式工况下，描述不同网络异常流量检测性能，检测结果如图 2 所示。

由图 2 可知，（a）为高负载工况网络，需要承受大量数据传输与并发访问；（b）为高实时性工况网络，需要实时传输与处理数据；（c）为强加密工况网络，需要保护敏感数据，避免网络数据被攻击；（d）为分布式工况网络，需要连接多个分散节点，确保网络的可扩展性。本文将 ROC 曲线对角线作为阈值，低于对角线时，证明检测性能不佳，ROC 曲线超出对角线时，且围成的面积越大时，证明检测性能越佳。从（a）、（b）、（c）、（d）中可知，在 BENIGN、DoS Hulk、Port Scan、DDoS、DoS Golden Eye、FTP-Patator 等数据集中，网络异常流量检测任务具有卓越表现，检测性能较佳。

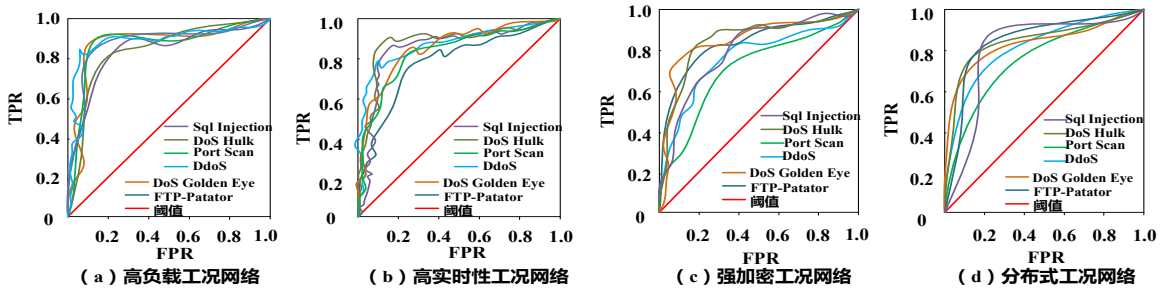

图 2 网络异常流量检测 ROC 曲线图

3.4 网络异常流量检测结果

在上述实验条件下，本文将 DoS Golden Eye、FTP-Patator、DDoS 等数据集进行平衡，按照检测准确个数与全部异常流量数据的比值，计算 Acc 值，得到工况网络异常流量检测结果。本文利用信息流和状态流融合的方式，将网络活动数据、网络流量变化进行深入分析，反映网络设备、网络流量的异常状态，从而帮助网络管理人员进行有效地修复措施。最终的检测结果显示，在迭代次数达到 60 次时，DoS Golden Eye 、FTP-Patator 、DDoS 等数据集的 Acc 值趋于稳定，且＞ 90%，异常流量检测的准确性较高。由此可见，使用本文设计的方法之后，能够全面地了解网络运行状态，发现潜在的异常情况，对于提高网络安全具有重要作用。

4 结语

本文设计的基于信息流和状态流融合的工况网络异常流量检测方法，将信息流与状态流融合，全面捕捉网络流量的状态变化，提供丰富的异常检测信息。通过部署检测网络，模拟网络流量的正常行为与异常行为。在检测网络中，定义异常行为检测标签，并根据网络流量数据的状态变化，分配对应的检测标签。通过计算标签数据点的异常分数，反映该数据点异常检测结果的偏离程度，为异常分数设定一个阈值，分数超过或低于阈值时，判断该标签对应的数据点为异常网络流量数据，从而实现了更加全面、准确的异常流量检测，为网络运行安全提供了保障。

参考文献

[1] 张乐,成玮,张硕,等.深度图网络驱动的核电系统多级异常检测方法[J].振动.测试与诊断,2025,45(01):88-94+202.

[2] 李冲,杜秀娟,王丽娟,等.基于异常检测和AUV辅助的水下传感器网络可靠节能路由协议[J].通信学报,2025,46(01):222-238.

[3] 麻胜兰,钟建坤,刘昱昊,等.基于二维卷积神经网络的结构加速度数据异常检测研究[J].建筑科学与工程学报,2025,42(01):112-120.

[4] 李贺,彭以冲,张万园,等.基于图卷积自编码器的多视图属性网络异常检测算法[J].中国科学:信息科学,2025,55(02):269-283.

[5] 王坤,付钰,段雪源,等.基于深度学习的SDN异常流量分布式检测方法[J].通信学报,2024,45(11):114-130.

[6] 王伟胜,王来花,贾晴,等.基于SE-U-Net预测网络的视频异常事件检测方法[J].计算机应用与软件,2024,41(12):154-160.

[7] 李为,袁泽坤,吴克河,等.基于注意力机制和多尺度卷积神经网络的容器异常检测[J].信息安全研究,2025,11(01):35-42.

[8] 胡文涛,徐靖凯,丁伟杰.基于溯因学习的无监督网络流量异常检测[J].信息网络安全,2024,24(11):1675-1684.

[9] 张宏伟,曼茂立,王宇,等.基于机器视觉与图卷积网络的矿区无人驾驶车辆异常行为检测[J].金属矿山,2024,(10):182-187.

[10] 杨浩然,谢辉,宋康,等.基于改进GAN的智能网联车CAN总线异常检测研究[J].汽车安全与节能学报,2024,15(05):660-669.

[11] 王泽鹏,马超,张壮壮,等.动态决策驱动的工控网络数据要素威胁检测方法[J].计算机研究与发展,2024,61(10):2404-2416.

[12] 岑俊杰,李永波. 链路失衡干扰下网络流量异常点挖掘仿真[J].计算机仿真,2024,41(2):397-400,405.

[13] 王宇飞,刘强,张唯贞,等.rtTorTIM:基于多模态特征融合和Stacking集成学习的实时Tor流量识别方法[J].计算机工程与科学,2025,47(02):238-246.

[14] 李聪聪,袁子龙,滕桂法.基于深度学习的时空特征融合网络入侵检测模型研究[J].信息安全研究,2025,11(02):122-129.

[15] 王梦雨,朱树永,张玉军.一种基于行为特征的网络靶场大规模攻击流量生成方法[J].高技术通讯,2024,34(11):1153-1163.