计算机系统应用 ISSN 1003-3254, CODEN CSAOBN
Computer Systems & Applications,2022,31(2):246−252 [doi: 10.15888/j.cnki.csa.008380]

E-mail: csa@iscas.ac.cn
http://www.c-s-a.org.cn
Tel: +86-10-62661041

图嵌入和 LSTM 自动编码器结合的 BGP 异常检测①

张树晓, 唐　勇, 刘宇靖

(国防科技大学 计算机学院, 长沙 410073)
通信作者: 唐　勇, E-mail: ytang@nudt.edu.cn

摘　要: 针对 BGP 异常数据的检测问题, 依托互联网公开的真实 BGP 更新报文数据, 重点结合网络的拓扑特征及时序变化特点, 提出一种新的基于图嵌入特征和 LSTM 自动编码器的 BGP 异常检测方法. 首先利用 BGP 数据中 AS_PATH 属性信息, 构建基于时间序列的网络拓扑图的动态嵌入特征数据集, 然后使用 LSTM 自动编码器模型对数据进行检测, 发现异常数据. 在实际的异常事件数据中, 该方法成功检测到了异常数据, 并且相比传统的检测方法有较高的准确率.

关键词: 图嵌入特征; BGP; 异常检测; LSTM 自动编码器

引用格式: 张树晓,唐勇,刘宇靖.图嵌入和 LSTM 自动编码器结合的 BGP 异常检测.计算机系统应用,2022,31(2):246–252. http://www.c-s-a.org.cn/1003-3254/8380.html

Anomaly Detection of BGP Using Graph Embedding and LSTM AutoEncoder

ZHANG Shu-Xiao, TANG Yong, LIU Yu-Jing

(College of Computer Science and Technology, National University of Defence Technology, Changsha 410073, China)

Abstract: With the real border gateway protocol (BGP) update message data disclosed on the Internet, this study proposes a new BGP anomaly detection method based on graph embedding features and long short-term memory (LSTM) AutoEncoder, which focuses on the network topology and variation characteristics in time series. First, the AS_PATH attribute information of BGP data is used to construct a dynamic embedding feature dataset based on the network topology of time series, and then the LSTM AutoEncoder model is employed for data detection to find abnormal ones. For the actual data of abnormal events, the method successfully detects the abnormal data and has higher accuracy than traditional detection methods.

Key words: graph embedding; border gateway protocol (BGP); anomaly detection; long short-term memory (LSTM) AutoEncoder

1　引言

互联网是一个庞大复杂的系统, 根据数据统计, 其当前包含 7 万多个网络自治系统 (autonomous system, AS)[1], 且数量一直处于不断增加之中. AS 之间通过边界网关路由协议进行通信, 实现网络的互联互通[2]. BGP (border gateway protocol) 作为边界网关路由协议的事实标准, 有效实现了 AS 间乃至整个互联网通信的有效可靠运行. 然而, 由于原始设计本身的安全缺陷, BGP 很容易遭到恶意攻击或产生错误配置, 对互联网的正常运行造成重大影响[3].

根据相关研究[4], BGP 协议本身的安全缺陷主要包括: 对等体 (peer) 的真实性及其之间通信报文的完

① 基金项目: 国家自然科学基金 (61602503)
收稿时间: 2021-04-25; 修改时间: 2021-05-19; 采用时间: 2021-06-24; csa 在线出版时间: 2022-01-17

整性、及时性没有验证; 没有安全机制对 AS 宣告网络层可达信息 (NLRI) 的权限进行验证; 没有安全机制确保 AS 通告路径 (AS_PATH) 的真实性. 此外, 由于 BGP 协议使用 TCP 传输协议进行通信, TCP 协议的安全风险在 BGP 协议上同样存在.

由于上述的缺陷问题, 通常可以将对 BGP 协议的恶意攻击分为前缀劫持攻击和路由泄露攻击[5]. 其中, 前缀劫持攻击一般通过伪造 NLRI 信息或 NLRI 信息与 AS_PATH 信息来实现. 进一步细化原因分类, 可以分为源 AS 路由修改、伪造 AS_PATH 路径信息、错误配置 AS、错误 AS 路径前置等 4 种类型[6]. 前缀劫持攻击很容易导致互联网的大规模异常, 也是国内外众多研究人员关注的焦点. 然而, 现实表明, 每年依然会有很多机构发表关于 BGP 前缀劫持攻击事件的报告[7,8], 部分事件更是由于波及范围广, 影响时间久, 引起了社会大众的广泛关注. 数据表明[8], 有超过 40% 的网络运营商曾是劫持攻击的受害者. 此外, 通过 BGP 前缀劫持攻击, 恶意攻击者能够在用户毫无知觉的情况下窃听、操控、记录指定 Internet 流量, 进一步实现对网络的中间人攻击[2], 对信息安全造成严重危害. 路由泄露则是由于 AS 没有按照规定的路由通告策略执行而产生, 导致相关的网络流量重定向到非正常的 AS. 这在现实中很有可能导致非法路由的快速异常扩散, 引发大规模乃至全球范围内的“断网”事件.

因此, 对 BGP 异常事件的检测研究很有必要, 有助于运营商快速发现异常, 及时响应处理, 避免产生更大危害.

本文依托从互联网公开收集的 BGP 更新报文数据, 提出一种基于拓扑图特征的数据集构建方法, 重点提取分析报文中的 AS_PATH 属性信息, 获取网络拓扑图序列, 生成图嵌入特征的时间序列数据集, 进一步结合 LSTM 自动编码器模型架构, 实现一种新的 BGP 异常数据检测方法, 并在实际的异常事件数据中进行了验证, 取得了良好效果.

2 相关研究

对 BGP 攻击的研究主要从两个方面考虑: 一是攻击的预防, 二是攻击的检测.

对于攻击的预防, 主要从前文所述的 BGP 协议缺陷入手, 进行相应改进, 以求从根本上解决问题. 包括安全域间路由 (secure inter-domain routing, SIDR) 工作组的安全 BGP 研究[9]; 避免选择虚假路由的完美 BGP (pretty good BGP)[10]; 已经被加入 RFC 标准的 BGPsec[11]; 引入加密和路由冗余机制的 BGP[12] 等. 然而, 由于多种原因, 相关的改进一直未能在真正的互联网上大规模部署, BGP 协议的缺陷问题没有从根本上发生改变, 各类安全异常事件也时有发生.

通过对 BGP 数据进行异常检测来发现、分类以及处理攻击行为的研究同样得到了广泛开展. 综合现有研究成果, 按照异常检测所用的数据类型, 大致可以分成 3 类方法[13]: 控制平面方法、数据平面方法和混合方法. 控制平面方法是一种无源解决方案, 其依赖于第三方监控器和收集器提供的数据进行异常行为的检测. 现在常用的第三方工具框架有 BGPmon[14] 和 BGPStream[15], 二者均提供历史和实时 BGP 数据的解析运行情况, 记录相关的异常事件. 在此基础上, 文献 [16] 提出了前缀劫持报警系统 PHAS; 文献 [17] 提出了 ARTEMIS 系统, 其号称能够快速检测到前缀劫持异常, 并能通过多种手段迅速灵活的缓解劫持异常. 总之, 此类方法都属于对异常的响应处理, 以及时做出策略应对. 数据平面方法则从网络中数据转发的异常入手, 使用了常见的 ping、traceroute 等网络工具. 通过检测相应的前缀地址的可达性实现异常的检测[13]. 混合方法主要是考虑前面两种方法的优缺点, 将二者进行结合使用, 主要思路是发现控制平面数据的不一致性后, 进一步通知数据平面进行精确测量[13].

近年来, 随着人工智能的繁荣发展, 机器学习也广泛应用于异常检测领域, 对 BGP 异常数据的检测也同样包含其中. 大量的文献使用机器学习 (包含深度学习) 模型算法提高 BGP 异常检测的准确性以及分类识别 BGP 异常等. 文献 [18] 对 BGP 数据进行特征提取并建立数据集, 提出使用支持向量机 (SVM) 模型来训练和测试 BGP 数据集, 并使用隐藏马尔可夫模型 (HMM) 来评估数据特征的有效性. 文献 [19] 使用了决策树和模糊测试集方法进行特征选择, 然后使用决策树和极限学习机进行异常分类, 提升 BGP 异常检测的准确性. 文献 [20] 则是对常见机器学习算法进行性能评估, 其使用的数据来自路由信息库 (RIB), 分布使用了朴素贝叶斯 (NB), 支持向量机 (SVM) 和决策树 (J48) 分类器进行异常检测, 对比其检测效能. 文献 [21] 同样基于支持向量机模型进行检测, 其利用 Fisher 线性分析算法和马尔可夫随机理论对特征选取进一步优化, 提升检

测性能. 文献 [22] 提出了利用离散小波变换来体现 BGP 数据的时间序列信息, 并与多层 LSTM 模型结合, 在异常数据集的检测中取得了良好效果.

3 模型与方法

本文通过图嵌入特征学习算法提取 BGP 数据中的拓扑图特征, 生成多维时间序列数据, 并使用 LSTM 自动编码器模型实现异常数据的检测.

3.1 图嵌入特征

众所周知, 整个互联网作为互联互通的整体, 其必然存在完整的拓扑图, 但由于实际网络中的海量设备及复杂的连接关系, 该拓扑的绘制是无法实现的. 因此我们从构成互联网的 AS 角度考虑, 将每个 AS 看作互联网中的一个节点, 也就是拓扑图中的顶点, 将 AS 间互相连接关系看作拓扑图的边, 是可以构建出基本的连接关系的. BGP 更新报文中的 AS_PATH 属性恰好包含构成拓扑图的要素, 其中蕴含了网络的拓扑图特征, 同时, 每一条 AS_PATH 都可以看作整体拓扑图中的一个子图. 此外, 结合报文的定时更新特性, 按时间窗口累计的 AS_PATH 就可以构成相对完整的拓扑图, 整个网络的变化状态就演变成网络拓扑图的变化, 并通过图的分布式特征表现出来. 其示意图如图 1 所示.

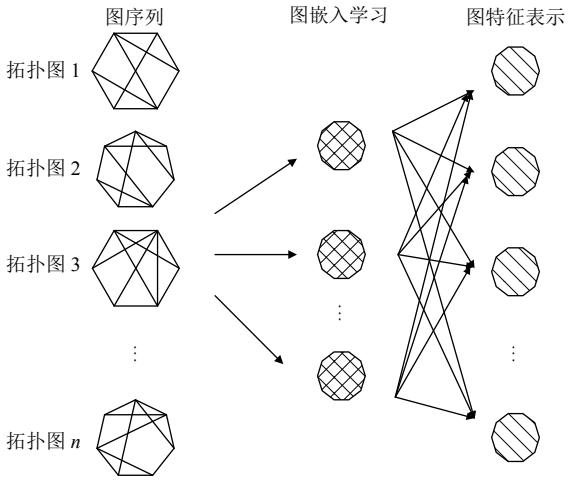


图 1 图嵌入特征结构

图 1 中, 拓扑图序列表示将 AS_PATH 属性按时间窗口划分聚合后的拓扑, 经过图嵌入学习后, 关联各时序的图特征, 得到嵌入特征, 每个图的特征可以用统一维度的向量表示, 最终获取多维的时间序列数据进行后续处理.

具体来说, 对于给定的图的集合 $\boldsymbol{G}=\{G_1, G_2, G_3, \cdots, G_n\}$, 每个图 G_i 都可以用 e 维向量表示 (e 为正整数), 那么所有图的集合可以使用矩阵 M 表示, 其中 $M\in\boldsymbol{R}^{|\boldsymbol{G}|\times e}$. 对具体的图 $G=\{N, E, \lambda\}$, N 表示节点集合, E 表示边的集合, 函数 λ 表示对图中每一个节点进行唯一标记. 同时, 定义子图 $sg_n(d)$, 其表示图中某个节点 n 的深度为 d 的子图. 其主要算法[23] 流程如算法 1.

算法 1. 图嵌入特征学习算法

输入: 图集合 $\boldsymbol{G}=\{G_1, G_2, G_3, \cdots, G_n\}$, 对任意图 $G_i=\{N_i, E_i, \lambda_i\}$; D: 有根子图的最大深度, 并产生包含所有有根子图的集合, $SG=\{sg_1, sg_2, \cdots, sg_n\}$; e: 向量维度; p: 周期轮数; r: 学习率.
输出: 图向量表示矩阵 M.

1) 随机初始化矩阵 M.
2) 对所有图随机排序.
3) 对每个图 G_i 的每个节点 N_i 进行处理.
4) 获取节点 N_i 周围所有深度为 d 的子图 $sg_i(d)$, $d\in(0, D)$.
5) 获取子图 $sg_i(d)$ 出现在图 G 中的概率 $R(M)=-\ln Pr(sg_i(d)|M(G))$.
6) 根据上述概率不断修正更新矩阵 $M=M-r\frac{\partial R}{\partial M}$.
7) 返回步骤 2), 进行新一轮计算直至 p 轮结束.
8) 返回最终矩阵 M.

子图的获取主要使用了经典的威斯费勒-莱曼算法. 当然, 对于图来说, 所有子图的规模是很大的, 因此采用了负采样优化的方法计算概率. 在学习率的选取上, 使用经验数值进行调整.

3.2 LSTM 自动编码器模型

3.2.1 自动编码器 (AutoEncoder)

自动编码器[24] 是一种对称结构的人工神经网络, 包括编码器和解码器两个部分, 其本质上是一种数据压缩的算法, 通常用于高维数据的处理. 对输入的数据, 其编码器部分会通过一个或多个隐藏层, 形成低维向量数据, 再通过解码器部分重建输入数据, 其框架如图 2 所示.

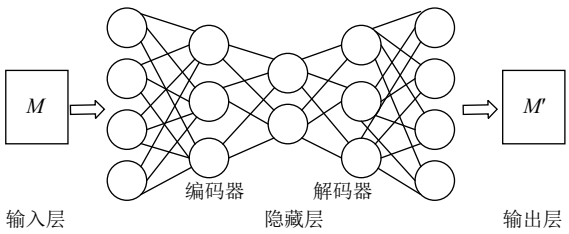


图 2 自动编码器结构

图 2 中, 输入层 M 表示多维向量的集合, 输出层 M' 表示重构向量的集合, 中间为隐藏层 H, 输入层与

隐藏层之间的组件为编码器, 隐藏层与输出层之间的组件为解码器. 通过自动编码器的处理, 我们可以重建数据的原始输入, 并学习隐藏层中的高维数据的编码表示. 对于正常数据与异常数据来说, 二者重构后的数据特征差异是巨大的, 这样我们就可以将异常数据检测出来.

通常, 编码层和解码层都可以用矩阵变换函数 Φ 和 Ψ 表示:

$$H = \Phi(M) \tag{1}$$

$$H = \varphi(W_1 M + b_1) \tag{2}$$

$$M' = \Psi(H) \tag{3}$$

$$M' = \psi(W_2 H + b_2) \tag{4}$$

其中, φ、ψ 表示激活函数, W_1 、W_2 为转移矩阵, b_1、b_2 为偏差向量. 具体来说, 编码层的变换将输入 M 压缩表示为 H, 我们选择使用常见的激活函数 ReLU, 其收敛速度和计算速度相对更快; 解码层的变换则是还原并输出 M'的过程, 同样使用 ReLU 作为激活函数.

对于输入和输出之间的重构误差, 则使用二阶距离度量, 记为:

$$D(M, M') = \|M - M'\|_2^2 \tag{5}$$

综上, 我们可以通过寻找合适的参数使得正常数据的重构误差最小, 便于将正常数据与非正常的数据区分出来. 总体而言, 自动编码器具备非线性的降维能力, 并能通过中间的隐藏层学习更多的特征, 与传统的主成分分析 (PCA) 算法相比, 更具灵活性, 也利于提高检测的正确率.

3.2.2 集成 LSTM 的自动编码器

LSTM (long short term memory network) 全称为长短期记忆网络[25], 是一种时间循环神经网络 (recurrent neural network, RNN), 相比普通的 RNN, LSTM 在更长的时间序列中有更好的表现.

LSTM 网络由 LSTM 单元构成, 每个 LSTM 单元会传递自身的状态和之前输入的状态, 选择忘记不重要的信息, 选择记忆重要的特征信息. 因此, 可以将 LSTM 单元集成到自动编码器中[26], 每个编码器和解码器部分都采用 LSTM 单元, 其结构示意如图 3 所示.

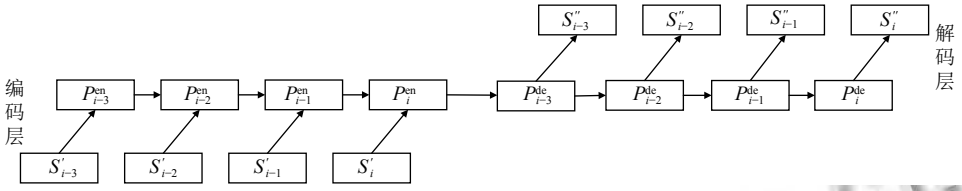


图 3 LSTM 自动编码器结构

图 3 中, $S=\{S_1, S_2, \cdots, S_n\}$表示输入的时间序列数据, 时刻 i 的序列状态记为 S_i, S'表示自动编码器训练过程中产生的序列. 图中给出了子序列长度为 4 的实例流程, 在编码器阶段, 时刻 i 的状态 P_i^{en} 受到 i–3 时刻到 i 时刻的状态影响, 而在传递到解码器时, 只需要传递 i 时刻的状态, 也就是在此时刻编码器状态与解码器状态是一致的, 解码器再进行反向的状态遍历, 还原时间序列.

4 实验

4.1 实验数据

在实验过程中, 使用的 BGP 数据均来自开源项目 RIPE RIS[27] 收集的真实互联网路由信息数据. RIPE RIS 包含多个部署在全球各地的远程采集器 (remote route collectors, RRC) 实时记录路由表及更新报文信息, 并提供 MRT 格式文件供下载利用. 根据已有的研究成果和最新的 BGP 安全事件, 我们选取了 3 个事件进行分析: 2003 年的 Slammer 蠕虫爆发、2017 年谷歌劫持 BGP 路由、2020 年 CenturyLink 网络中断事件. 其中, Slammer 蠕虫爆发时间最早, 且持续时间相对最长、波及范围最广, 是当时的年度重大安全事件, 相关的检测分析研究也较多, 便于与本文提出的方法进行对比; Google 劫持事件则是一次主要集中在日本范围内的异常, 且发现处理的速度较快, 影响相对较小; CenturyLink 网络中断则是 2020 年最新发生的全球性大规模异常事件, 反映了整个互联网依然存在的脆弱性问题. 各事件相关情况如表 1 所示.

表 1 中的数据来自于 RIPE RIS 项目位于日内瓦的采集器 RRC04 和东京大手町的 RRC06, 分别采集异常事件发生前后共计 5 天的 BGP 更新报文数据, 并按

每分钟作为间隔提取该时间范围内的 AS_PATH 属性信息, 按照异常的持续时间范围, 对相应的信息统一标记为 BGP 异常数据.

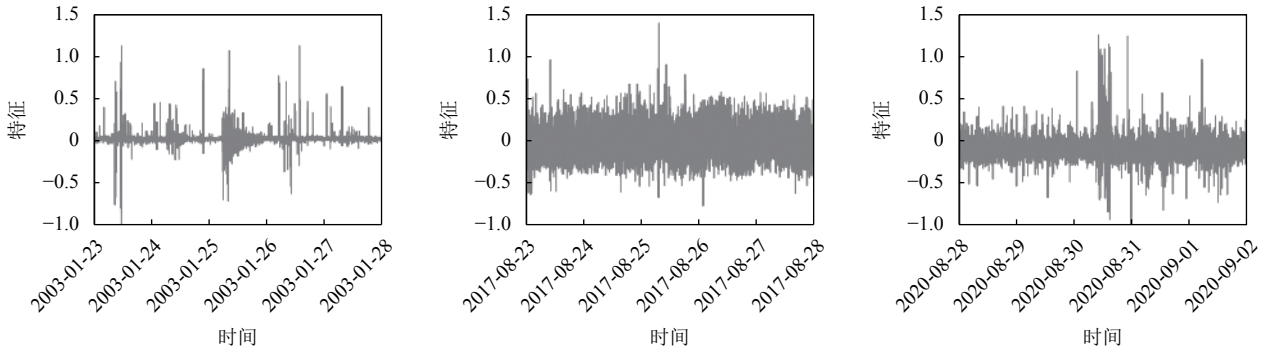
4.2 拓扑图特征

利用上文提到的图嵌入特征学习算法, 结合 AS_PATH 属性信息, 我们可以将相应时间段内的网络拓扑变化情况以向量特征表示出来. 具体结果如图 4.

从图 4 中可以看出, 在发生异常事件的时间段内, 网络的拓扑特征同样会发生明显变化.

表 1 数据集

异常事件	异常开始时间	持续时长 (min)	采集器
Slammer	2003-01-25 (05:30 UTC)	860	RRC04
Google	2017-08-25 (03:22 UTC)	39	RRC04
CenturyLink	2020-08-30 (10:04 UTC)	266	RRC04



(a) Slammer (b) Google (c) CenturyLink

图 4 异常事件拓扑图部分特征

4.3 评价指标

在模型数据处理中, 将数据集中的正常数据作为训练集和验证集进行模型的训练, 并使用标记的全部数据集作为训练集, 将模型的检测结果与实际数据集中的标记类型进行对比, 对检测效果进行评价.

对正常与异常数据的二分类问题, 通常采用正确率 (*Accuracy*)、准确率 (*Precision*)、召回率 (*Recall*)、*F* 值 (*F-measure*) 进行评价. 具体如下所示:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{6}$$

$$Precision = \frac{TP}{TP+FP} \tag{7}$$

$$Recall = \frac{TP}{TP+FP} \tag{8}$$

$$F\text{-}measure = \frac{2\times Precision\times Recall}{Precision+Recall} \tag{9}$$

其中, *TP* 表示检测正常为真, *TN* 表示检测异常为真, *FP* 表示检测异常为假, *FN* 表示检测正常为假. 正确率反映了整体检测的正确水平, 准确率和召回率只针对正常数据的检测情况, 反映精准度和错检情况, *F* 值则是综合衡量检测的性能.

此外, 由于异常检测中的数据是不平衡的, 同样需要关注对异常值的检测情况, 可以使用曲线下面积 AUC (area under curve) 进展评估, 此处的曲线指接收者操作特征曲线 ROC, 该曲线表示了两类预测分布间的差异性, 通常认为 AUC 值越接近 1, 分类的效果就越好.

4.4 实验结果及分析

首先, 我们使用上文提到的数据及模型, 对相关数据集进行检测分类, 以验证方法的有效性. 对于本文而言, 通过实验不仅可以验证模型的检测效果, 同时可以对文中提出的基于网络拓扑图嵌入特征的数据集构建方法进行有效验证. 实验的具体结果如表 2 所示.

表 2 LSTM 自动编码器检测结果 (%)

数据集	*F*值	正确率
Slammer	81.52	90.77
Google	61.71	99.37
CenturyLink	76.97	95.27

从表 2 中可知, 模型整体的正确率很高, 都在 90% 以上, 特别是在 Google 事件数据集中正确率达到了 99.37%, 但其 *F* 值相对较低, 主要是该异常事件持续时间短, 在按分钟级的数据提取情况下, 异常数据的量相对于正常数据所占比例是非常小的, 同时个别数据特别是在异常事件起始和终止时刻的数据难以精准的识别

判断, 容易出现错检情况, 从而对检测结果产生较大影响. 相对的, 对于持续时间较长的异常事件, 受到的干扰影响就比较小, 检测的结果就比较稳定. 具体到 F 值上看, 异常事件的持续时间越长, 异常检测的准确度也越高.

同时, 进一步使用 AUC 对 3 个异常事件的整体分类情况进行评估, 得到 AUC 曲线图如图 5 所示.

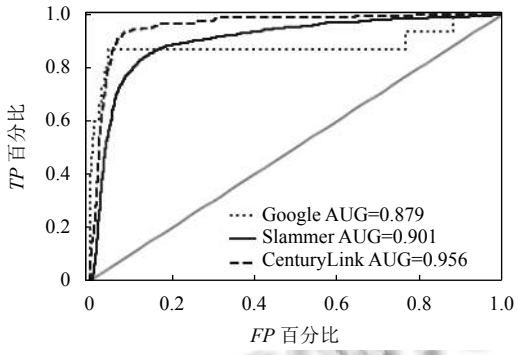


图 5 ROC 曲线图

从图 5 中可知, 在 CenturyLink 事件中的异常检测效果最好, 达到 95.6%. 这也与前表中该事件正确率与 F 值的水平都比较均衡有关, 整体上达到了很好的异常分类效果. 而在 Google 事件中, 可以明显观察到曲线的阶梯状上升, 这就是在极少量异常数据情况下单个异常点数据影响的结果.

综合上述检测结果, 我们认为构建的数据集和相应的检测模型都是有效的. 特别是对于 BGP 数据集的构建方法, 采用的拓扑图嵌入特征能够很好的反映异常事件中的数据变化情况, 相比于文献 [18] 中提出的基于 37 种 BGP 数据扩展特征构建的数据集或者从中选取部分扩展特征构建的数据集而言, 更加简单易于处理, 并且进一步结合检测模型能获得较为理想的检测结果.

为了进一步评估模型的效果, 使用经典的支持向量机 (support vector machines, SVM) 模型及前文提到的原始自动编码器模型对 Slammer 数据集进行了异常检测分类测试, 结果如表 3 所示.

从表 3 中可知, 文中提出的 LSTM 自动编码器模型的异常检测效果最好, 其 F 值和正确率都优于其他两种模型. 对于原始的自动编码器模型而言, 还是体现出了其作为深度学习算法在高维数据非线性处理上的优势, 结合数据的时间序列特征集成 LSTM 处理单元之后, 模型的检测能力得到了进一步的提高.

表 3 不同模型异常检测结果 (%)

模型	F值	正确率
SVM	63.21	83.36
AutoEncoder	75.92	90.57
LSTM-AutoEncoder	81.52	90.77

5 结论

本文针对 BGP 数据的异常检测问题, 提出了一种使用 LSTM 自动编码器模型进行异常事件检测的方法. 为了能够有效的检测出 BGP 异常事件, 从互联网中真实的历史数据入手, 利用图嵌入特征算法, 获取基于时间序列的网络拓扑变化情况, 构建基于网络拓扑图特征的数据集, 为模型处理提供高效简洁的数据. 我们通过选取 3 起有代表性的 BGP 安全事件及相关节点采集的数据进行了模型方法的验证.实验结果表明, 该数据集能够很好的反映异常事件中的网络特征变化情况, 使用的模型相比于传统的检测方法有更高的准确率, 能够有效检测出 BGP 异常事件, 便于后续及时响应处理. 下一步将结合具体的应用场景, 研究更大规模 BGP 时间序列数据情况下的检测与响应机制.

参考文献

1 CIDR REPORT for 21 Dec 20. https://www.cidr-report.org/as2.0/. [2020-11-20].
2 Shapira T, Shavitt Y. A deep learning approach for IP hijack detection based on ASN embedding. Proceedings of the Workshop on Network Meets AI & ML. Virtual: ACM, 2020. 35–41.
3 Vervier PA, Thonnard O, Dacier M. Mind your blocks: On the stealthiness of malicious BGP hijacks. Network and Distributed System Security Symposium. San Diego: Internet Society, 2015.
4 Murphy S. BGP security vulnerabilities analysis. RFC4272, 2006.
5 黎松, 诸葛建伟, 李星. BGP 安全研究. 软件学报, 2013, 24(1): 121–138. [doi: 10.3724/SP.J.1001.2013.04346]
6 Cho S, Fontugne R, Cho K, *et al*. BGP hijacking classification. 2019 Network Traffic Measurement and Analysis Conference (TMA). Paris: IEEE, 2019. 25–32.
7 Demchak CC, Shavitt Y. China's maxim–leave no access point unexploited: The hidden story of China Telecom's BGP hijacking. Military Cyber Affairs, 2018, 3(1): 7.
8 Sermpezis P, Kotronis V, Dainotti A, *et al*. A survey among network operators on BGP prefix hijacking. ACM

SIGCOMM Computer Communication Review, 2018, 48(1): 64–69.
9 Huston G, Bush R. Securing BGP and SIDR. IETF Journal. 2011, 7(1).
10 Karlin J, Forrest S, Rexford J. Pretty good BGP: Improving BGP by cautiously adopting routes. Proceedings of the 2006 IEEE International Conference on Network Protocols. Santa Barbara: IEEE, 2006. 290–299.
11 Lepinski M, Sriram K. BGPSEC protocol specification. RFC8205, 2017.
12 Subramanian L, Roth V, Stoica I, *et al*. Listen and whisper: Security mechanisms for BGP. Proceedings of the 1st Conference on Symposium on Networked Systems Design and Implementation. San Francisco: USENIX Association, 2004. 10.
13 Al-Musawi B, Branch P, Armitage G. BGP anomaly detection techniques: A survey. IEEE Communications Surveys & Tutorials, 2017, 19(1): 377–396.
14 BGPmon. https://www.bgpmon.net/. [2020-11-20].
15 BGPStream. https://bgpstream.caida.org/. [2020-11-20].
16 Lad M, Massey D, Pei D, *et al*. PHAS: A prefix hijack alert system. Proceedings of the 15th Conference on USENIX Security Symposium. Vancouver: USENIX Association, 2006. 11.
17 Sermpezis P, Kotronis V, Gigis P, *et al*. ARTEMIS: Neutralizing BGP hijacking within a minute. IEEE/ACM Transactions on Networking, 2018, 26(6): 2471–2486.
18 Al-Rousan NM, Trajković L. Machine learning models for classification of BGP anomalies. 2012 IEEE 13th International Conference on High Performance Switching and Routing. Belgrade: IEEE, 2012. 103–108.
19 Li Y, Xing HJ, Hua Q, *et al*. Classification of BGP anomalies using decision trees and fuzzy rough sets. 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC). San Diego: IEEE, 2014. 1312–1317. [doi: 10.1109/SMC.2014.6974096]
20 Ćosović M, Obradović S, Trajković L. Performance evaluation of BGP anomaly classifiers. 2015 3rd International Conference on Digital Information, Networking, and Wireless Communications (DINWC). Moscow: IEEE, 2015. 115–120.
21 Dai XB, Wang N, Wang WJ. Application of machine learning in BGP anomaly detection. Journal of Physics: Conference Series, 2019, 1176(3): 032015.
22 Cheng M, Li Q, Lv JM, *et al*. Multi-scale LSTM model for BGP anomaly classification. IEEE Transactions on Services Computing, 2021, 14(3): 765–778.
23 Narayanan A, Chandramohan M, Venkatesan R, *et al*. graph2vec: Learning distributed representations of graphs. arXiv: 1707.05005v1, 2017.
24 Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. Science, 2006, 313(5786): 504–507.
25 Hochreiter S, Schmidhuber J. Long short-term memory. Neural Computation, 1997, 9(8): 1735–1780.
26 Kieu T, Yang B, Jensen CS. Outlier detection for multidimensional time series using deep neural networks. 2018 19th IEEE International Conference on Mobile Data Management (MDM). Aalborg: IEEE, 2018. 125–134.
27 RIS Raw Data. https://www.ripe.net/analyse/internet-measurements/routing-information-service-ris/ris-raw-data. [2020-11-20].