文章编号:1006-9348(2024)07-0546-05

基于GAT-BiLSTM模型的日志异常检测方法

梁华雄[1],赵 刚[1],王兴芬[1,2]
(1. 北京信息科技大学信息管理学院,北京 100192;
2. 北京信息科技大学大数据研究院,北京 100192)

摘要:针对当前日志异常检测数据量大和人工构建特征困难的问题,提出一种基于特征融合的GAT-BiLSTM深度学习模型。模型利用图注意网络(GAT)充分提取全局日志间的信息,得到更为全面的日志特征表示;同时利用双向长短期记忆网络(BiLSTM)挖掘日志内容的序列特征信息,最后通过特征的自适应融合实现对日志特征的提取。实验结果表明,GAT-BiLSTM模型有效解决了日志文本数据在异常检测中准确率不足的问题,准确率达到82.10%,在日志异常检测领域具有一定的研究意义。

关键词:图结构;图注意力网络;特征融合;日志异常检测
中图分类号:TP391 **文献标识码**:B

A Log Anomaly Detection Method Based on GAT-BiLSTM Model

LIANG Hua-xiong[1],ZHAO Gang[1],WANG Xing-fen[1,2]
(1. School of Information Management,Beijing University of Information Science and Technology,
Beijing 100192,China;
2. Institute of Big Data,Beijing University of Information Science and Technology,Beijing 100192,China)

ABSTRACT:Aiming at the current problems of large data volume of log anomaly detection and difficulties in constructing features manually,a deep learning model of GAT-BiLSTM based on feature fusion is proposed. The model makes use of graph attention network (GAT) to fully extract the information among global logs and obtain a more comprehensive log feature representation; meanwhile,it makes use of bi-directional long and short-term memory network (BiLSTM) to mine the sequence feature information of log contents,and finally achieves the extraction of log features by adaptive fusion of features. The experimental results show that the GAT-BiLSTM model effectively solves the problem of insufficient accuracy of log text data in anomaly detection,with an accuracy rate of 82.10%,which has certain research significance in the field of anomaly log detection.

KEYWORDS:Graph structure;Graph attention networks;Feature fusion;Log anomaly detection

1 引言

随着互联网技术的发展,网络攻击的方式也呈现出自动化、多样化的发展趋势,网络设备的稳定运行面临着前所未有的挑战。通过对网络设备的日志检测来评价当前系统是否存在异常情况并及时采取措施规避风险,是网络安全运营人员常用的技术手段。

在传统的日志分析过程中,安全人员需根据专业领域知识和经验,人工构建异常日志的特征从而对日志进行检测。但随着黑客技术的发展,网络攻击由简单、易察觉、易防范逐渐演化为有组织、有规模、可持续威胁的攻击,导致传统检测方法的可适性逐渐降低。另一方面,网络攻击往往是交替发生且具有一定的步骤,导致产生的日志相互间存在一定的隐性关系,这进一步增大了人工构建日志特征的难度。而近些年来,深度学习模型从数据中自动学习特征的特性为解决上述问题提供了新的思路。

2 相关工作

利用机器学习和深度学习等技术对日志进行异常检测一直是当前研究的热点。程世文[1]提出正则表达式的方法检索日志以达到检测的目的,但该方法会随着正则表达式的复杂和日志的增多而让搜索时间呈指数级别增多。Du[2]等

基金项目:国家重点研发计划课题(2019YFB1405003)
收稿日期:2022-10-16 修回日期:2022-11-25

提出了使用基于长短期记忆网络的深度学习模型 DeepLog 对具有一定时序性的系统日志建模并进行异常检测,取得了不错的效果。梅御东等[3]使用深度学习中的 CNN-text 方法来对系统日志进行分类来达到日志异常检测的目的。房笑宇等[4]将日志序列输入基于注意力机制的生成对抗网络,通过对比生成器生成的序列和真实发生的序列是否一致判断该日志是否异常。仇媛等[5]提出了一种基于 LSTM 网络和滑动窗口的流数据异常检测方法,根据滑动窗口内 LSTM 的预测值和真实值的差值分布进行建模来计算数据流量的异常。Lin 等[6]则提出了一种基于知识库的聚类算法 LogCluster,利用历史数据中的序列化日志信息用作判断异常事件的基准。李海林等[7]提出了一种基于频繁模式发现的时间序列异常检测算法,利用历史输入的时间序列和当前新增的时间序列的相似度进行异常检测。Xia 等[7]提出基于生成对抗网络的日志异常检测模型 LogGAN。Tuor 等[8]提出一种基于循环神经网络的方法,将一个日志序列视为一个句子输入网络进行分类判断该序列是否异常。Huang[10]等通过日志序列编码器和参数编码器获得日志的表示,后利用注意力机制网络作为分类器对日志进行异常检测。Guo[11]等通过分析计算机系统日志和用户行为日志,建立不同日志间的多头注意力序列模型对日志进行检测。

上述对异常日志检测的研究虽然有效解决了人工提取特征困难的问题,但仍存在不足之处。模型侧重于日志的时间序列特征却忽略了日志间的空间位置特征,并没有将两者有效结合;将日志整体作为研究对象,没有考虑到日志内容细粒度的相互联系。因此,本文将结合图注意力网络[12]和双向长短期记忆网络[13]提取特征的优势,构建适用于日志异常检测的 GAT-BiLSTM 模型,以提高日志异常检测的准确率。本文贡献如下:

1)构建日志的图结构,并利用图注意力网络挖掘日志间非连续,长距离的信息;利用 BiLSTM 挖掘日志内容在序列上的特征信息。

2)并行结合 GAT 和 BiLSTM 结构,构建 GAT-BiLSTM 模型提取特征信息并用于日志异常检测。

3 GAT-BiLSTM 算法模型

本文提出的基于深度学习的 GAT-BiLSTM 结构如图 1 所示。模型主要由四个部分构成,分别为词嵌入、特征提取、自适应融合和模型训练。

3.1 词嵌入

深度学习模型的输入一般是数值化的向量,Word2Vec[14],Glove[15]是当前自然语言经常使用的词向量编码方式。考虑到 Glove 是基于全局词汇编码的特点,本文采用 Glove 的编码方式对日志文本向量化。

3.2 特征提取

3.2.1 图注意力网络

为了挖掘日志间的隐性关系,获得更全面的日志特征表达。本文构建图 G(V,E)来挖掘日志与日志间的隐藏信息。在构建的日志图结构[16]中,节点包含词节点和日志节点,边包含日志与词相连的边、词与词相连的边;然后通过图注意力网络汇聚邻接节点信息,获得日志的嵌入表达。其过程如图 2 所示。

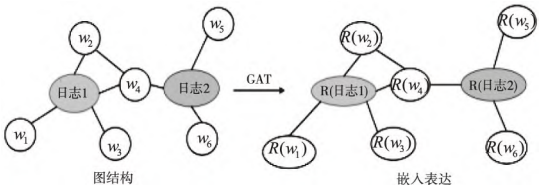

图 2 日志节点嵌入表达

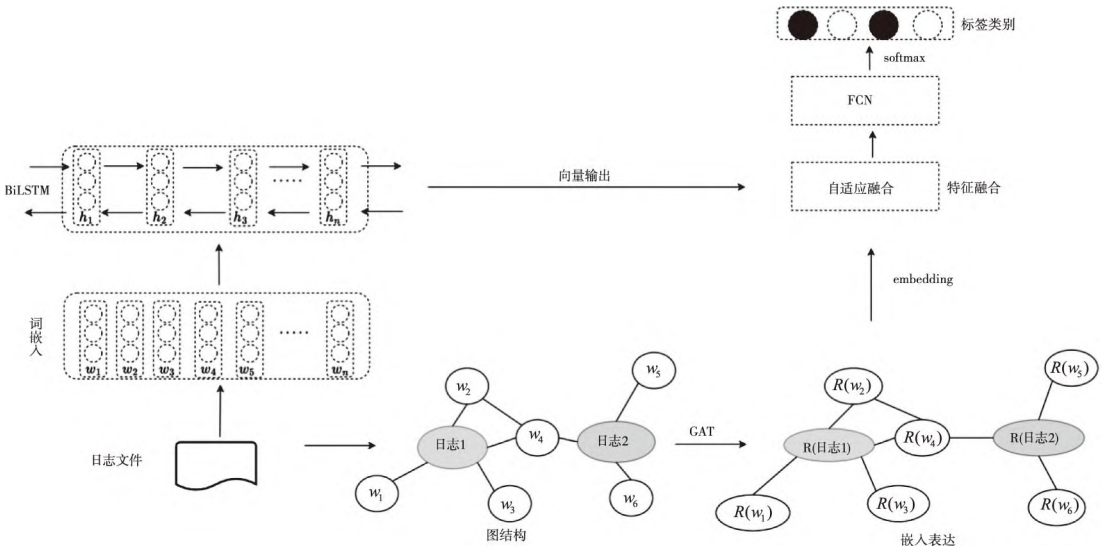

图 1 GAT-BILSTM 模型结构

在日志图结构中，邻接矩阵 A 的定义如下，D 为图的度矩阵。

$$A_{ij} = \begin{cases} PMI(i,j), & i,j \text{ 分别是词} \\ TF-IDF_{ij}, & i \text{ 是当前日志},j \text{ 是日志里面的词} \\ 1, & i=j \\ 0, & \text{其他} \end{cases} \tag{1}$$

$$A_{ij} = D^{-\frac{1}{2}} A_{ij} D^{-\frac{1}{2}} \tag{2}$$

在图结构中，使用词节点和日志节点的 one-hot 编码作为其输入 h，$h=\{h_1,h_2,\cdots,h_N\}$，N 是节点个数。图注意网络能为邻居节点分配不同的注意力分数，从而获得更好的日志节点嵌入，计算过程如下：

$$e_{ij} = Leaky\mathrm{Re}\,LU(a^T[A_{ij}Wh_i \parallel A_{ij}Wh_j]), j \in N_i \tag{3}$$

$$\alpha_{ij} = \frac{\exp(Leaky\mathrm{Re}\,LU(a^T[A_{ij}Wh_i \parallel A_{ij}Wh_j]))}{\sum_{k\in N_i}\exp(Leaky\mathrm{Re}\,LU(a^T[A_{ij}Wh_i \parallel A_{ij}Wh_k]))} \tag{4}$$

$$h'_i = \sigma\left(\sum_{j\in N_i}\alpha_{ij}Wh_j\right) \tag{5}$$

其中，$\parallel$ 表示拼接操作，a 为注意力机制，一般采用单层前馈神经网络，$j\in N_i$ 表示 j 是节点 i 邻居节点，e_{ij} 为节点 j 对节点 i 重要程度，α_{ij} 为注意力系数，h_i' 为 h_i 经过注意力系数更新的特征向量。

3.2.2　双向长短期记忆循环神经网络

BiLSTM 采用双向输入结构，以此捕捉上下文双向语义依赖关系，其结构如图 3 所示。

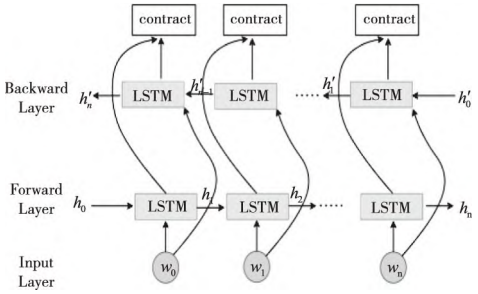

图 3　BiLSTM 提取特征过程

在图 3 的 LSTM 模块中，t 时刻输入词向量 w_t 的输出 h_t 计算过程如下

$$f_t = \sigma(W_f\cdot[h_{t-1},w_t] + b_f) \tag{6}$$

$$i_t = \sigma(W_i\cdot[h_{t-1},w_t] + b_i) \tag{7}$$

$$\tilde{C}_t = \sigma(W_c\cdot[h_{t-1},w_t] + b_c) \tag{8}$$

$$C_t = f_c * C_{t-1} + i_t * \tilde{C}_t \tag{9}$$

$$o_t = \sigma(W_o\cdot[h_{t-1},w_t] + b_o) \tag{10}$$

$$h_t = o_t * \tanh(C_t) \tag{11}$$

式中 $f_t,i_t,\tilde{C}_t,C_t,o_t$，分别为 t 时刻的遗忘门、输入门、临时细胞态、细胞态与输出门的值，W 和 b 分别是对应的权重矩阵和偏置向量，σ 为 sigmoid 函数，tanh 为双曲正切激活函数。同理，反向 LSTM 层网络中 t 时刻的输出信息为 h''。为了获得日志内容的序列信息，在日志最后一个词向量 w_n 经过 BiLSTM 后，将其输出结果 h_n 和 $h^{n\prime}$ 进行拼接，得到当前日志的特征向量 H，计算如下。

$$H = [h_n, h'_n] \tag{12}$$

3.3　自适应特征融合

在给定日志集合 $\{(D_1,l_1),(D_2,l_2),\cdots,(D_n,l_n)\}$ 中，其中 D_n 表示第 n 个日志，l_n 表示第 n 个日志的类别标签，$l_n\in\{0,1,\cdots,k\}$，k 是样本类别数。对于日志 (D_k,l_k)，其由一系列词向量组成 $\{w_{K_1},w_{K_2},\cdots,w_{K_n}\}$。在 *BiLSTM* 网络中，依次输入 $\{w_{K_1},w_{K_2},\cdots,w_{K_n}\}$ 得到日志序列特征 H_k；在图结构中，通过 *GAT* 网络汇聚邻居节点的信息，得到日志的全局嵌入表达 h'_k。为了提高模型的泛化性，对这两种特征向量采用自适应特征融合[17]的策略。日志 D_k 的自适应特征融合计算如下

$$\begin{cases} \alpha = \mathrm{sigmoid}(H_k\cdot W_1) \\ \beta = \mathrm{sigmoid}(h'_k\cdot W_2) \end{cases} \tag{13}$$

式中 α,β 分别为日志序列信息和日志间信息的重要度系数，W_1,W_2 为待训练的参数矩阵，融合后的特征最终表示为

$$\begin{cases} E_k = \alpha H_k + \beta h'_k \\ 1 = \alpha + \beta \end{cases} \tag{14}$$

3.4　模型训练

在自适应融合特征后，得到日志的特征向量矩阵 E。将 E 送入全连接层进行分类，计算过程如下。

$$y = soft\max(W_4\cdot\sigma(W_3\cdot E + b_1) + b_2) \tag{15}$$

σ 为 ReLU 函数，W_3 和 W_4 为待训练矩阵，b_1 和 b_2 为偏置向量。通过交叉熵损失函数计算其损失值 L：

$$L = -\sum_{i=1}^{k} y_i \log y'_i \tag{16}$$

其中 y_i 表示第 i 个日志类别的真实标签，y'_i 为模型的预测标签。通过损失函数最小化训练 GAT-BiLSTM 模型，并且采用 Adam[18] 作为优化器进行梯度下降。

3.5　模型算法步骤

GAT-BiLSTM 算法的输入是日志数据 D_{ata} 和日志标签 y。具体流程如下：

Step1：数据集文本图的构建。构建文本图结构并得到其度矩阵 D，利用 GAT 挖掘日志间的信息并获得特征向量 h'。

Step2：对数据集 D_{ata} 进行分词、去重、编码，得到各个词的编码向量 X，利用 BiLSTM 挖掘日志内容序列间的信息并获得特征向量 H。

Step3：融合特征向量得到向量 E。其中 E = Adaptive Fuion(H,h')。

Step4：模型的训练和参数更新。将 E 送进多层感知机中，得到预测结果 y_{pred}。通过损失函数最小化更新参数。

4　实验与分析

4.1　实验数据集和评价指标

实验数据集来源于公开数据集 Honeynet[19]，人工整理

并标记 7 种不同类别的 snort 报警日志,一共 7969 条记录。日志经过数据清洗后,使用 60%作为训练集,40%作为测试集。实验采取正确率(Acc),宏查准率(Macro_P),宏召回率(Macro_R)和宏 F 值(Macro_F)作为评价模型的性能指标,指标公式如下

$$Acc = (\sum_{i=1}^{k} TP_i)/N \tag{17}$$

$$P = \frac{TP}{TP + FP} \tag{18}$$

$$R = \frac{TP}{TP + FN} \tag{19}$$

$$Macro\_P = \frac{1}{k}\sum_{i=1}^{k} P_i \tag{20}$$

$$Macro\_R = \frac{1}{k}\sum_{i=1}^{k} R_i \tag{21}$$

$$Macro\_F = \frac{1}{k}\sum_{i=1}^{k} \frac{2 \cdot P_i \cdot R_i}{P_i + R_i} \tag{22}$$

其中 k 是样本类别数目,TP_i 是每个类别中样本为正且预测为正的样本数,N 是总样本数目,FP_i 是每个类别预测为正但实际样本为负的数目,FN_i 是每个类别预测为负但实际样本为正的数目,P_i 代表当前类别 i 的查准率,R_i 当前类别 i 的召回率。

4.2 模型分析实验

4.2.1 GAT-BiLSTM 模型检测效果对比实验

对比 TextCNN[20]、GAT、BiLSTM 和 GAT-BiLSTM 模型在数据集的实验指标,来衡量 GAT-BiLSTM 模型检测的有效性。为了更加客观衡量 GAT-BiLSTM 模型的检测能力,模型 GAT 和 BiLSTM 超参数都与 GAT-BiLSTM 模型中的 GAT 和 BiLSTM 的参数一致。GAT、BiLSTM 和 GAT-BiLSTM 分类层都使用两层全连接层网络。

表 1 模型参数

模型	参数项	参数值
GAT	输入	one_hot 编码
	head	4
	GAT 层数	2
	dropout	0.3
	激活函数	leaky_relu
	优化函数	Adam( )
	全连接层大小	128,64
	全连接层激活函数	tanh
BiLSTM	输入维度	50
	循环层大小	16+16
	激活函数	Sigmoid
	全连接层大小	128,64
	全连接层激活函数	tanh

各个模型的准确率、宏查准率、宏召回率和宏 F 值的实验结果如表 2 所示。

表 2 实验结果对比

模型	Acc	Macro_P	Macro_R	Macro_F
TextCNN	74.63%	72.49%	70.16%	70.96%
GAT	78.88%	77.28%	75.79%	76.38%
BiLSTM	79.11%	77.37%	78.24%	77.52%
GAT-BiLSTM	82.10%	79.14%	82.93%	80.74%

四种模型的结果对比如图 4 所示。可以看出,相较与 TextCNN、GAT 和 BiLSTM 模型,GAT-BiLSTM 深度学习模型在准确率、宏查准率、宏召回率和宏 F 值都有不同程度的提升,分别达到了 82.10%、79.14%、82.93%和 80.74%。相比于单个模型,经过特征组合后 GAT-BiLSTM 模型具有明显的优势。其原因是日志经过 BiLSTM 网络和 GAT 网络后,分别提取了日志内容的序列信息和日志间的关系信息。后经过特征自适应融合,一定程度上比单个模型输出的特征具有更好的表达效果。

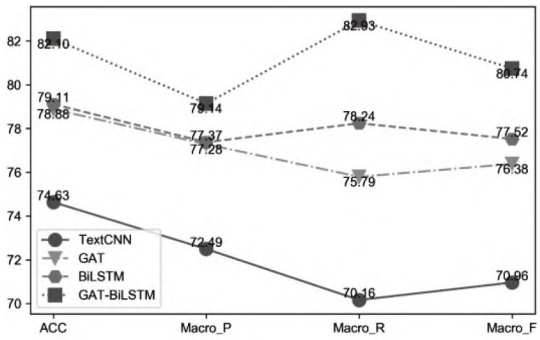

图 4 不同模型评价指标结果

4.2.2 GAT-BiLSTM 模型的消融实验

在不改变 GAT-BiLSTM 的结构和参数情况下,通过消融 GAT-BiLSTM 模型的部分结构,研究其对于模型的贡献。

1)消融 GAT-BiLSTM 模型词嵌入层。将数据集分别通过 Word2Vec 和 Glove 编码作为模型结构 BiLSTM 的输入,实验结果如表 3 所示。经过 Glove 编码的词向量在 GAT-BiLSTM 模型的宏准确率、宏查全率和宏 F 值分别比 Word2Vec 编码的 GAT-BiLSTM 模型提高了 1.2%、2.19%和 1.29%,由此可见不同的词向量输入对模型的实验指标具有一定的影响,且 Glove 编码的词嵌入表现更好。其原因可能是 Glove 是基于全局语料共现矩阵的编码;Word2Vec 由于滑窗的机制导致其是局部语料的编码,所以经过 Glove 编码的 GAT-BILSTM 模型在实验指标上表现更好。

表 3 词嵌入对实验指标的影响

词嵌入	Acc	Macro_P	Macro_R	Macro_F
Word2Vec	81.00	76.95	83.54	79.35
Glove	82.10	79.14	82.93	80.74

2)消融 GAT-BiLSTM 模型的 GAT 堆叠层数。为了探究 GAT 层数对模型的影响,分别设计一层 GAT、两层 GAT、三层 GAT 和四层 GAT 的 4 种 GAT-BiLSTM 模型。对比四个模型的实验指标,结果如图 5 所示。随着 GAT 层数的增长,准确率,宏查准率,宏召回率和宏 F 值都呈现出了先上升后下降的趋势。当 GAT 网络层堆叠到第二层的时候,各项实验验指标都达到了最好的效果。但是随着堆叠网络层的增多,各项实验指标反而在下降。原因是网络层数的增大,节点汇聚的邻接节点的信息更多,反而更容易使得每个节点同质化,导致提取的特征信息区分度不高,从而影响模型的判断。

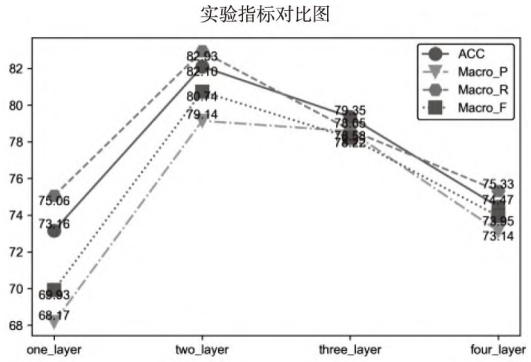

图 5　不同 GAT 层数的实验指标

为了更近一步观察 GAT 堆叠层数对实验模型的影响,将不同 GAT 层数的日志节点嵌入表达采用 TSNE 降维并可视化。T-SNE[21] 是一种非线性降维算法,能将高维数据依靠一定的分布降维到 2 维或者 3 维。日志节点嵌入表达的高维特征向量经过 TSNE 降维到 2 维空间的可视化结果如图 6 所示。可以看出,GAT 堆叠到两层的时候能学习到更多有辨识度的节点嵌入,不同类型间的嵌入表达区分明显。

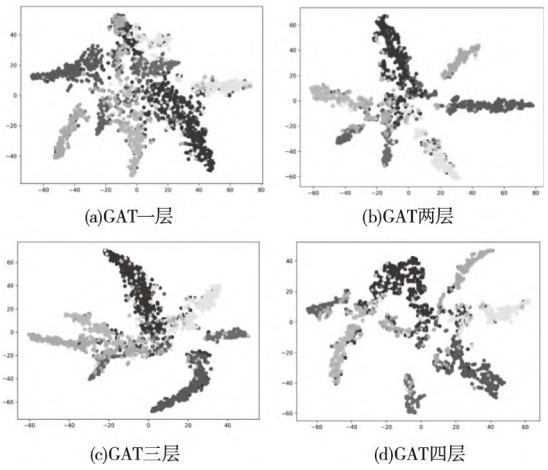

图 6　日志节点降维可视化

5　结束语

本文通过结合 GAT 和 BiLSTM 模型,构建 GAT-BiLSTM 模型用于异常日志的检测任务。首先构建日志的图结构,利用图卷积网络汇聚全图拓扑的信息,挖掘日志间的深层次信息;同时借助 BiLSTM 挖掘日志的序列关系,最后经过特征自适应融合的策略来提高日志检测的性能指标。通过在数据集上的对比实验,验证了 GAT-BiLSTM 模型在日志异常检测的有效性,对处理一些非结构数据具有一定的研究意义。

在以后的研究中,将考虑引入外部知识库来丰富嵌入学习的特征信息;另一方面,在 GAT-BiLSTM 模型的基础上,结合具体的工作场景,研究该模型在日志多标签上的适用性,提高模型解决系统异常问题的能力。

参考文献:

[1] 程世文,裴丹,王长进. 互联网软件错误日志聚类[J]. 小型微型计算机系统,2018,39(5): 865-870.

[2] Du M,Li F,Zheng G,et al. Deeplog: Anomaly detection and diagnosis from system logs through deep learning[C]. Proceedings of the 2017 ACM SIGSAC conference on computer and communications security. 2017: 1285-1298.

[3] 梅御东,陈旭,孙毓忠,等. 一种基于日志信息和 CNN-text 的软件系统异常检测方法[J]. 计算机学报,2020,43(2):366-380.

[4] 房笑宇,曹陈涵,夏彬. 基于注意力机制的大规模系统日志异常检测方法[J]. 南京大学学报(自然科学版),2021,57(5):785-792.

[5] 仇媛,常相茂,仇倩,彭程,苏善婷. 基于长短期记忆网络和滑动窗口的流数据异常检测方法[J]. 计算机应用,2020,40(5):1335-1339.

[6] Lin Q,Zhang H,Lou J G,et al. Logclustering based problem identification for online service systems[C]. 2016 IEEE/ACM 38th International Conference on Software Engineering Companion (ICSE-C). IEEE,2016:102-111.

[7] 李海林,邬先利. 基于频繁模式发现的时间序列异常检测方法[J]. 计算机应用,2018,38(11):3204-3210.

[8] Xia B,Yin J,Xu J,et al. Loggan: A sequence-based generative adversarial network for anomaly detection based on system logs[C]. International Conference on Science of Cyber Security. Springer,Cham,2019:61-76.

[9] Tuor A R,Baerwolf R,Knowles N,et al. Recurrent neural network language models for open vocabulary event-level cyber anomaly detection[C]. Workshops at the thirty-second AAAI conference on artificial intelligence. 2018.

[10] Huang S,Liu Y,Fung C,et al. Hitanomaly: Hierarchical transformers for anomaly detection in system log[J]. IEEE transactions on network and service management,2020,17(4):2064-2076.

(下转第 558 页)

[13] ZHOU J,KISHORE R,ZUO M,et al. Older adults in virtual communities: understanding the antecedents of knowledge contribution and knowledge seeking through the lens of socioemotional selectivity and social cognitive theories[J]. Journal of Knowledge Management,2021.
[14] LIN M J J,HUNG S W,CHEN C J. Fostering the determinants of knowledge sharing in professional virtual communities[J]. Computers in Human Behavior,2009,25(4): 929-39.
[15] GUAN T,WANG L,JIN J,et al. Knowledge contribution behavior in online Q&A communities: An empirical investigation[J]. Computers in Human Behavior,2018,81: 137-47.
[16] 张薇薇,朱杰,蒋雪. 社会学习对专业虚拟社区 不同类型用户知识贡献行为的影响研究[J]. 情报资料工作,2021,42(5): 94-103.
[17] ZHENG Y,ZHAO K,STYLIANOU A. The impacts of information quality and system quality on users' continuance intention in information-exchange virtual communities: An empirical investigation[J]. Decision support systems,2013,56:513-24.
[18] YAN J,LEIDNER D E,BENBYA H,et al. Social capital and knowledge contribution in online user communities: One-way or two-way relationship? [J]. Decision Support Systems,2019,127:113131.
[19] 姚慧丽,毛翔宇,金辉. 考虑平台影响因素的虚拟社区知识共享演化博弈研究[J]. 运筹与管理,2020,29(12): 82-8.
[20] 李从东,黄浩,张帆顺. 基于网络演化博弈的互动创新社区用户知识共享行为影响因素研究[J]. 现代情报,2021,41(4): 36-45.
[21] 王鹏民,侯贵生,杨磊. 基于知识质量的社会化问答社区用户知识共享的演化博弈分析[J]. 现代情报,2018,38(4):42-9.
[22] 杜智涛. 网络知识社区中用户"知识化"行为影响因素——基于知识贡献与知识获取两个视角[J]. 图书情报知识,2017,(2): 105-19.
[23] 高鹏,聂佳佳,杜建国,等. 消费者后悔预期对 IR 市场进入策略的影响[J]. 管理工程学报,2018,32(4):178-85.
[24] 高鹏,杜建国,聂佳佳,等. 消费者预期后悔对线下零售商服务提供策略的影响[J]. 预测,2018,37(3):62-8.
[25] 陈昕,钟英. 参与者情绪对虚拟社区知识共享的演化分析[J]. 山东科技大学学报(社会科学版),2020,22(05):74-86.
[26] ZHOU X,ZHAO R,CHENG L,et al. Impact of policy incentives on electric vehicles development: a system dynamics-based evolutionary game theoretical analysis[J]. Clean Technologies and Environmental Policy,2019,21(5):1039-53.

[作者简介]

梁　敬(1994-),女(汉族),山东省德州市人,博士研究生,主要研究领域为知识管理、系统建模与仿真、用户知识行为。

李　明(1981-),男(汉族),北京市人,教授,博士,博士生导师,主要研究领域为知识管理、数据挖掘、系统建模与仿真。

(上接第 550 页)

[11] Guo Y,Wen Y,Jiang C,et al. Detecting Log Anomalies with Multi-Head Attention (LAMA)[J]. arXiv preprint arXiv: 2101.02392,2021.
[12] Velickovic P,Cucurull G,Casanova A,et al. Graph attention networks[J]. arXiv preprint arXiv:1710.10903,2017.
[13] Xu G,Meng Y,Qiu X,et al. Sentiment analysis of comment texts based on BiLSTM[J]. Ieee Access,2019,7: 51522-51532.
[14] Church K W. Word2Vec[J]. Natural Language Engineering,2017,23(1):155-162.
[15] Pennington J,Socher R,Manning C D. Glove: Global vectors for word representation[C]. Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP). 2014:1532-1543.
[16] Yao L,Mao C,Luo Y. Graph convolutional networks for text classification[C]. Proceedings of the AAAI conference on artificial intelligence. 2019,33(01): 7370-7377.
[17] Wang N,Gong X. Adaptive fusion for RGB-D salient object detection[J]. IEEE Access,2019,7:55277-55284.
[18] Jais I K M,Ismail A R,Nisa S Q. Adam optimization algorithm for wide and deep neural network[J]. Knowledge Engineering and Data Science,2019,2(1):41-46.
[19] Thonnard O,Dacier M. A framework for attack patterns' discovery in honeynet data[J]. digital investigation,2008,5:S128-S139.
[20] Guo B,Zhang C,Liu J,et al. Improving text classification with weighted word embeddings via a multi-channel TextCNN model[J]. Neurocomputing,2019,363:366-374.
[21] Linderman G C,Rachh M,Hoskins J G,et al. Fast interpolation-based t-SNE for improved visualization of single-cell RNA-seq data[J]. Nature methods,2019,16(3): 243-245.

[作者简介]

梁华雄(1994-),男(汉族),广东阳春人,硕士研究生,主要领域为网络内容安全,网络安全,自然语言处理技术。

赵　刚(1965-),男(汉族),北京人,博士,教授,研究生导师,主要研究领域为管理科学与工程、网络空间安全学科,强化学习。

王兴芬(1968-),女(汉族),北京人,博士,教授,研究生导师,主要研究领域为管理科学与工程,计算机科学与技术。