Doi:10.3969/j.issn.2097-0307.2025.02.002

基于LSTM-GRBM的海洋云平台虚拟机异常检测算法

韩泽欣，吴永芳，张镭，司佳，陈刚
（国家海洋信息中心，天津 300171）

摘 要：在利用海洋云平台虚拟机部署各类海洋应用系统过程中会遇到软硬件故障、网络攻击等问题，准确进行虚拟机异常检测有利于保障海洋业务顺利开展。本文提出了一种基于长短期记忆网络-高斯受限玻尔兹曼机（LSTM-GRBM）的海洋云平台虚拟机异常检测算法，首先通过LSTM模型提取虚拟机性能指标数据的时序特征，然后利用Dropout技术避免出现数据过拟合，最后通过GRBM计算得出自由能并与训练得到参数基准模型进行对比来判断虚拟机是否出现异常。为验证算法有效性，使用真实海洋云平台虚拟机性能指标数据进行实验，结果表明提出的LSTM-GRBM模型具有更好的异常检测性能。

关键词：异常检测；海洋云平台虚拟机；长短期记忆网络；高斯受限玻尔兹曼机
中图分类号：TP393 **文献识别码：**A **文章编号：**2097-0307(2025)02-0074-08

Anomaly detection algorithm of marine cloud platform virtual machine based on LSTM-GRBM

HAN Zexin, WU Yongfang, ZHANG Lei, SI Jia, CHEN Gang
（National Marine Data and Information Service, Tianjin 300171, China）

Abstract: In deploying marine application systems via Virtual Machines (VMs) on marine cloud platforms, operational disruptions such as hardware/software failures or cyberattacks may impede maritime services, thus necessitating accurate VM anomaly detection. This paper proposes an Long Short-Term Memory-Gaussian Restricted Boltzmann Machine (LSTM-GRBM) hybrid algorithm for VM anomaly detection in marine cloud environments. The methodology involves:(1) Extracting temporal features from VM performance metrics using LSTM networks; (2) Mitigating overfitting via Dropout regularization; (3) Computing free energy values through GRBM and comparing them against a pre-trained parametric benchmark model to identify anomalies. Validation experiments utilizing real-world VM performance datasets from marine cloud platforms demonstrate the proposed model's superior detection accuracy compared to baseline methods.

Keywords: anomaly detection; marine cloud platform virtual machine; Long Short-Term Memory; Gaussian Restricted Boltzmann Machine

海洋云平台是基于服务器、存储阵列、交换机等硬件资源和应用软件、操作系统、集成开发环境等软件资源，运用虚拟化、分布式计算、软件定义网络和自动化管理等技术构建的云服务中枢平台。其能够支持负载均衡和内容分发等网络功能，提供可扩展的存储服务，实现弹性的计算资源分配和管理。目前海洋云平台已具备一定的基础和规模，实现了计算、存储和方法模型资源的统一管理和运维调度，有力保障了海洋大数据处理分析和海洋监管网络建设等海洋信息化产业的发展[1-2]。

海洋云平台虚拟机是在物理计算节点上创建的虚拟计算环境，通过虚拟化软件对物理计算节点的CPU、内存单元和存储介质等资源进行管理

收稿日期：2024-09-18；修订日期：2024-12-31
作者简介：韩泽欣（1991—），工程师，主要从事海洋云平台技术研究与应用，电子邮箱：hanzexin@nmdis.org.cn
通信作者：吴永芳，工程师，主要从事数据存储与备份技术研究与应用，电子邮箱：wuyongfang@nmdis.org.cn

和调度，实现资源的共享和高效利用，具有伸缩弹性、可用性高、配置灵活、资源优化和安全性强等优势[3]。每个虚拟机都有独立的操作系统、应用程序和数据存储，可以满足不同海洋应用场景的需求。但虚拟机在使用过程中，会遇到配置不当、应用程序代码错误、物理计算节点软硬件故障以及网络攻击等问题，将导致虚拟机性能降低甚至死机，影响海洋应用系统正常运行，可能造成数据丢失、泄露等严重事故。因此，需要采取有效手段进行海洋云平台虚拟机异常检测，进一步提升海洋云平台的稳定性。

目前，云平台虚拟机异常检测使用的方法主要有统计学分析、机器学习和深度学习等。由于虚拟机的运行情况是未知的，其运行状态信息难以服从概率统计分布模型，导致基于统计学的异常检测方法准确度较低[4]。基于机器学习或深度学习的异常检测方法可以避免运行情况未知的问题，通过分析云平台虚拟机的运行状态信息，能够较为准确地检测出虚拟机异常[5-6]。Khreich等[7]利用N-gram将信息序列的频率映射为特征向量，训练了基于SVM的检测模型。Mishra等[8]利用监控的虚拟机进程构建特征向量，并结合卷积神经网络（Convolutional Neural Networks，CNN）和长短期记忆（Long Short-Term Memory，LSTM）网络进行了异常检测。Zhang等[9]通过强化虚拟机动作行为之间的关联特征，使用深度神经网络实现了行为异常检测。贺寰烨等[10]将虚拟机的密度空间性质引入到LOF算法来检测负载不均异常。杨光[11]将虚拟机运行数据变换至时间维度和信息增益维度后再输入到神经网络进行计算。王开放等[12]根据虚拟机前后运行特征信息，使用多通道机制构建了Bi-LSTM模型来预测虚拟机故障。

本文综合运用LSTM与高斯受限玻尔兹曼机（Gaussian Restricted Boltzmann Machine，GRBM）两种网络模型，提出了基于LSTM-GRBM的海洋云平台虚拟机异常检测算法。首先将虚拟机性能指标数据集输入到LSTM网络进行时序特征提取，同时运用Dropout技术避免出现过拟合现象，然后在训练阶段利用GRBM最小化输入自由能来构建随时间变化的参数基准模型，最后在检测阶段通过GRBM计算得出自由能并与参数基准模型进行比较来判断虚拟机是否出现异常。

1　相关原理及技术

1.1　LSTM模型

Hochreiter等[13]在循环神经网络（Recurrent Neural Network，RNN）的基础上，提出了LSTM网络，以有效解决梯度消失和梯度爆炸等问题，其更适合处理长序列和长期依赖性任务。LSTM模型包含遗忘门、输入门和输出门3个门控单元，模型结构如图1所示。遗忘门f_t是网络中管理长期记忆的组件，用于去除当前单元的无用信息并保留关键信息。输入门i_t主要用于控制需要保存到当前单元的信息量，输出门o_t则决定当前单元的输出信息。

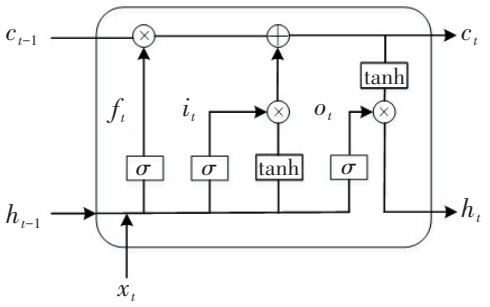

图1　LSTM模型结构图

注：f_t为遗忘门、i_t为输入门、o_t为输出门；h_t、h_{t-1}分别为当前时刻和上一时刻的隐藏状态信息；c_t、c_{t-1}分别为当前时刻和上一时刻的细胞状态信息；σ为Sigmoid激活函数；tanh为双曲正切激活函数；⊗为矩阵乘法；⊕为矩阵加法。

根据LSTM模型结构图，可以得到该神经网络的前向传播计算公式：

$$f_t = \sigma(W_f h_{t-1} + U_f x_t + b_f) \tag{1}$$

$$i_t = \sigma(W_i h_{t-1} + U_i x_t + b_i) \tag{2}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh(W_c h_{t-1} + U_c x_t + b_c) \tag{3}$$

$$o_t = \sigma(W_o h_{t-1} + U_o x_t + b_o) \tag{4}$$

$$h_t = o_t \odot \tanh c_t \tag{5}$$

式（1）—（5）中：h_t、h_{t-1}分别表示当前时刻和上一时刻的隐藏状态信息；x_t表示当前时刻的输入信息；c_t、c_{t-1}分别表示当前时刻和上一时刻的细胞状态信息；W_f、U_f表示f_t的权重参数，b_f是f_t的偏移量；W_i、U_i表示i_t的权重参数，b_i是i_t的偏移量；W_c、U_c表示c_t的权重参数，b_c是c_t的偏移量；W_o、U_o表示o_t的权重参数，b_o是o_t的偏移量；σ是Sigmoid激活函数；tanh是双曲正切激活函数。

1.2 Dropout技术

在深度学习模型训练过程中，Dropout技术会随机丢弃一部分神经元，更新模型参数得到不同的网络结构，这能够有效防止模型过于依赖某些特定神经元，从而使模型更加泛化[14]。神经网络前向传播序列，反向传递误差以更新参数，引入Dropout技术后，会随机删除隐藏层一半数量的神经元，如图2所示。

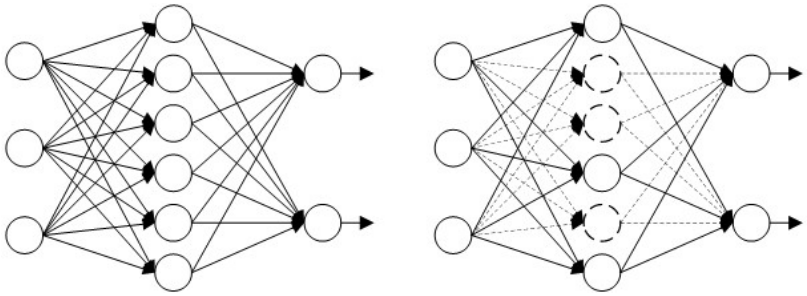

图2 标准神经网络（左）和引入Dropout技术后的神经网络（右）

设输入数据为$\boldsymbol{X}$，权重矩阵为$\boldsymbol{W}$，偏置矩阵为$\boldsymbol{b}$，激活函数为f，则神经网络的输出$\boldsymbol{Y}$为：

$$\boldsymbol{Y}=f(\boldsymbol{W}\times\boldsymbol{X}+\boldsymbol{b}) \tag{6}$$

在训练过程中使用Dropout技术，假设Dropout技术丢弃神经元的概率为p（$0\leqslant p\leqslant 1$），这时的神经网络输出为：

$$\boldsymbol{Y\_dropout}=f(\boldsymbol{W}\times\boldsymbol{X}+\boldsymbol{b}\times(1-\boldsymbol{dropout\_mask})) \tag{7}$$

式中：$\boldsymbol{dropout_mask}$是一个与输入数据$\boldsymbol{X}$大小相同的矩阵，其元素服从伯努利分布，成功概率为$1-p$，失败概率为$p$。当$\boldsymbol{dropout_mask}$的元素为1时，表示神经元被保留；元素为0时，表示神经元被丢弃。

1.3 GRBM模型

GRBM将受限玻尔兹曼机（Restricted Boltzmann Machine，RBM）中可见层的二值变量节点替换为带独立高斯噪声的线性变量节点[15]，模型结构如图3所示。GRBM模型克服了RBM模型输入节点二值约束的限制，能够提取连续过程数据的特征，准确进行异常检测。

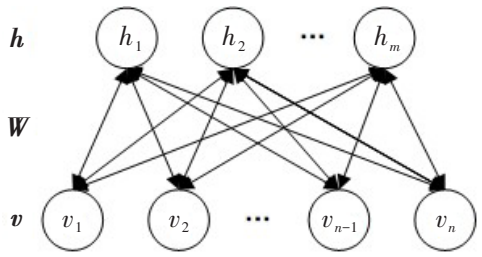

图3 GRBM模型结构图

注：向量$\boldsymbol{v}$和$\boldsymbol{h}$分别表示GRBM模型中的可见单元状态和隐藏单元状态，$\boldsymbol{W}$表示可见单元和隐藏单元之间的权重向量

GRBM模型的系统能量函数为：

$$E(v,h|\theta)=-\sum_{i=1}^{n}\frac{(v_i-a_i)^2}{2\sigma_i^2}-\sum_{j=1}^{m}b_jh_j-\sum_{i=1}^{n}\sum_{j=1}^{m}\frac{v_i}{\sigma_i}W_{ij}h_j \tag{8}$$

式中：v_i表示第i个可见单元；h_j表示第j个隐藏单元；$\theta=\{W_{ij},a_i,b_j\}$为GRBM模型的结构参数；$W_{ij}$表示$v_i$和$h_j$之间的连接权重；$a_i$和$b_j$分别表示$v_i$和$h_j$的偏置；$\sigma_i$表示高斯噪声标准差。

GRBM模型得出的自由能可以用于描述系统的稳定性和能量状态，通常表示为能量函数的相关函数，其值越低意味着系统性能越好。在海洋云平台虚拟机异常检测中，通过计算并最小化虚拟机性能指标数据的自由能，最终得到与实际输入序列最为接近的GRBM模型，使其能够准确拟合或预测虚拟机性能指标数据，从而实现对虚拟机状态的异常检测。

2 基于LSTM-GRBM的海洋云平台虚拟机异常检测

2.1 模型架构

海洋云平台中的虚拟机运行会产生大量的性能监测数据，机器学习方法具备高度自动化和智能化的特点，能够通过历史监测数据训练得到一个高效的分类模型，实现对海量数据的实时检测，及时发现异常虚拟机。因此，针对海洋云平台中虚拟机的性能指标，提出基于LSTM-GRBM的异常检测模型，其模型架构如图4所示。

基于LSTM-GRBM的海洋云平台虚拟机异常检测模型是由LSTM模型、Dropout技术和GRBM

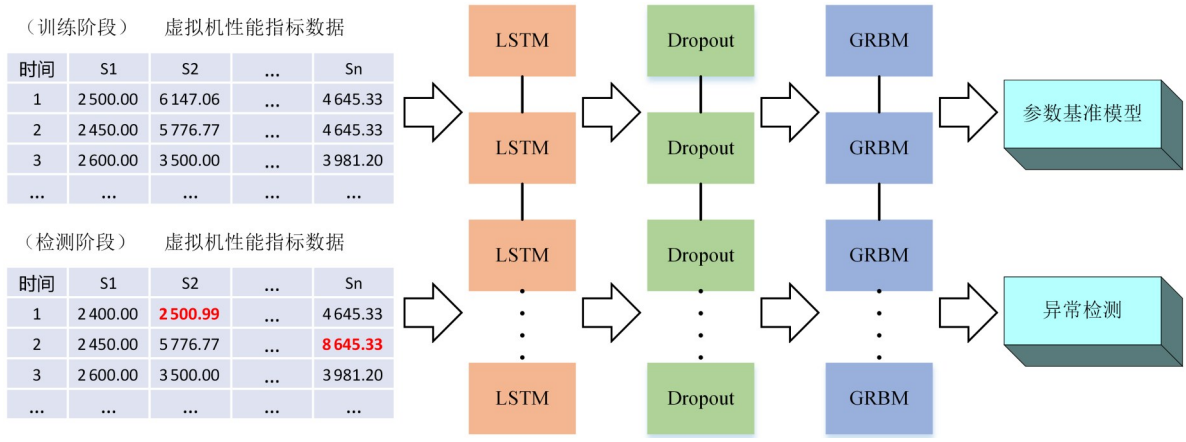

图4　基于LSTM-GRBM的模型架构

模型组合而成的端到端的深度学习网络。在训练阶段，将虚拟机正常运行状态下的性能指标数据集输入到网络，首先利用LSTM模型来对时序性数据进行处理以学习虚拟机性能指标特征，然后使用Dropout技术避免出现过拟合，最后运用GRBM模型最小化输入自由能以构建随时间变化的参数基准模型。在检测阶段，将虚拟机实时的性能指标数据集输入到LSTM-GRBM模型，通过对比得到自由能与参数基准模型来判断虚拟机是否出现异常。

2.2　工作原理

首先，从海洋云平台中获得虚拟机性能指标数据集，并将其输入到LSTM模型中得到时序数据特征向量。LSTM模型在5 min的时间窗口内对虚拟机性能指标数据进行多次采样，因此所提取的数据特征密度更高且时序性更为显著，大大提高了模型的异常检测精确度。

然后，为了规避模型训练过程中可能出现的数据过拟合问题，采用Dropout技术以概率$p(0 \leqslant p \leqslant 1)$来决定每个隐藏单元是否被丢弃。Dropout技术通过自适应的随机丢弃机制，在神经网络训练过程中选择性地忽略部分神经元来完成特征提取，而在异常检测时则保持输入虚拟机性能指标数据的完整性不受影响，能够充分利用LSTM模型所累积的全部知识，有效增强了模型的泛化能力。

最后，将虚拟机性能指标数据输入到GRBM模型中计算自由能。先根据公式（9）计算隐藏层节点的概率分布$P(h_j)$，并将得到的概率和预设阈值进行比较，如果超出阈值则激活该节点，并赋值为1，否则赋值为0。

$$P(h_j) = sigmoid\left(\sum_i w_{ij}\frac{v_i}{\sigma_i} + c_j\right) \tag{9}$$

式中：v_i表示第i个可见单元；h_j表示第j个隐藏单元；w_{ij}表示权重，c_j表示偏置；σ_i表示高斯噪声标准差。

接着计算可见层节点的概率分布$P(v_i)$，通过高斯分布生成可见节点的值，计算公式为：

$$P(v_i) = \frac{1}{\sigma_i\sqrt{2\pi}}\exp\left(-\frac{1}{2\sigma_i^2}\left(v_i - d_i - \sigma_i\sum_j u_{ij}h_j\right)\right) \tag{10}$$

式中：u_{ij}表示权重；d_i表示偏置

然后再利用公式（9）更新隐藏层节点的概率分布确定激活节点。配置隐藏层的单元数量与可见层保持一致，以便深入学习并捕捉识别异常的关键特征，通过隐藏层与可见层中各个节点值的相互作用，计算得到虚拟机性能指标数据的能量。对GRBM模型的能量函数进行变换后得到自由能$G(v)$，更为有效地逼近对数似然函数，在训练过程中最小化自由能来构建较好的模型，计算公式为：

$$G(v) = \log \mathrm{e}^{-E(v,h|\theta)} \tag{11}$$

式中：$E(v,h|\theta)$为GRBM模型的系统能量函数。

2.3　参数基准模型

参数基准模型由训练阶段得到的自由能的标准差$\sigma(G_{Tr})$和参数$\gamma \in N$决定。先是将历史虚拟机性能指标数据集输入到LSTM-GRBM模型进行训练，计算得到自由能G_{Tr}，然后使用趋势线函数$Trendline(G_{Tr})$构造随时间变化的自由能基准线，最后通过公式（12）和公式（13）分别确定参数基准模型的上下基准线$BaseLine_{\max}$和$BaseLine_{\min}$。处于正常状态的海洋云平台虚拟机，其自由能应处于该上下基准线之间。

$$BaseLine_{\max} = Trendline(G_{Tr}) + \gamma\sigma(G_{Tr}) \quad (12)$$

$$BaseLine_{\min} = Trendline(G_{Tr}) - \gamma\sigma(G_{Tr}) \quad (13)$$

通过计算可以得出，参数基准模型的上下基准线与基准线的偏离程度由γ决定。γ取值越大，模型自由能的变化范围越大，说明自由能处于该范围内的虚拟机运行状态正常，模型的容忍度越高；γ取值越小，模型自由能的变化范围越小，模型的容忍度越低。将一周的历史虚拟机性能指标数据集输入到LSTM-GRBM模型进行训练，获得$\gamma = 3$时的基准模型，如图5所示。

2.4 异常检测

在异常检测阶段，首先通过人工注入异常的方式生成虚拟机出现异常状态的性能指标数据集，然后将其输入到LSTM-GRBM模型计算自由能。若虚拟机在运行过程中出现异常，其产生的性能指标数据则会受到影响，得到的自由能将不会出现在上下基准线范围内，因此可检测出虚拟机异常。实验中，将3 d的异常虚拟机性能指标数据集输入到LSTM-GRBM模型检测，将得到的自由能与参数$\gamma = 3$时的基准模型进行对比，结果如图6所示。

3 实验与分析

3.1 性能指标

虚拟机在运行过程中出现异常状况，一般直接体现在CPU资源、内存资源、磁盘资源及网络性能的异常消耗等方面[16]。例如，CPU利用率较高、进程数较多，说明系统比较繁忙；内存利用率较高，内存写回量较大，是因为内存资源不足造成的；磁盘利用率大小和每秒磁盘读写次数等反映了磁盘的健康状况；网络负载率、每秒发送数据包丢包数和每秒接收数据包丢包数等指标用于显示当前网络性能。为保证虚拟机异常检测模型的准确性，最终选取了68项性能指标数据作为模型输入（表1）。

3.2 数据集来源及实验设置

本文使用的虚拟机性能指标数据集来自于海洋云平台，选取操作系统为Windows Server 2016、Windows Server 2019及Windows Server 2022的虚拟机共计150台。通过在每台虚拟机上安装性能监控软件来采集性能指标数据，采集频率1次/5 min。本文共监测10 d的数据，将前7 d的数据作为训练数据集，用于模型的训练与学习；将后3 d的数据作为测试数据集，并通过人工注入异常方式来实时获取异常数据，以便准确评估模型在异常检测时的性能表现。

通过人工手动生成CPU、内存、磁盘和网络等方面的异常。在部分虚拟机上执行脚本，计算字符串哈希值，抢占CPU资源，实现CPU资源异常注入；在部分虚拟机上执行脚本，不断申请占

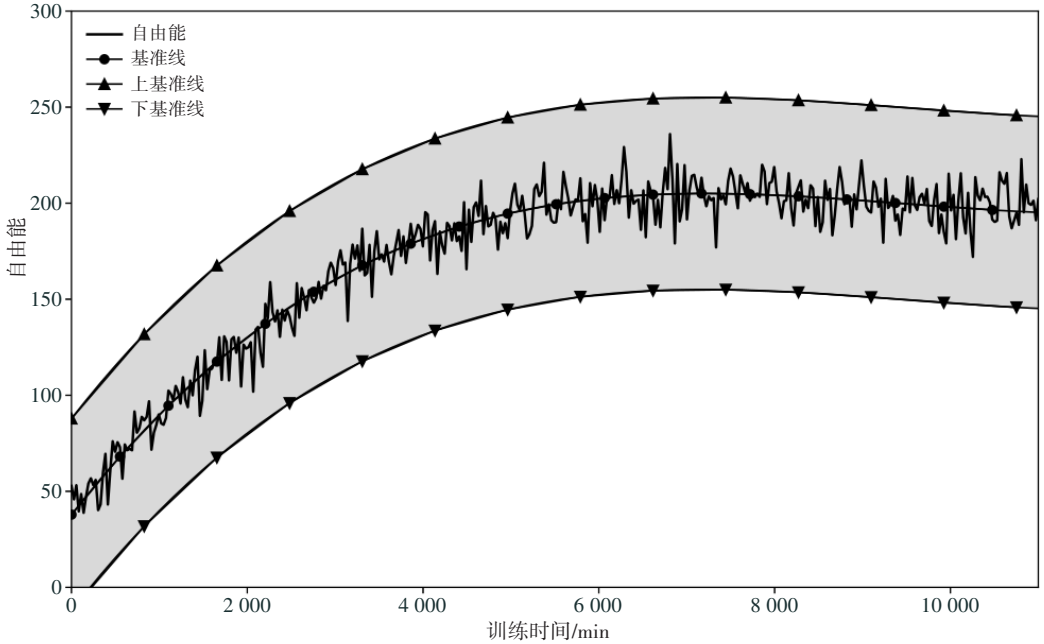

图5 参数基准模型

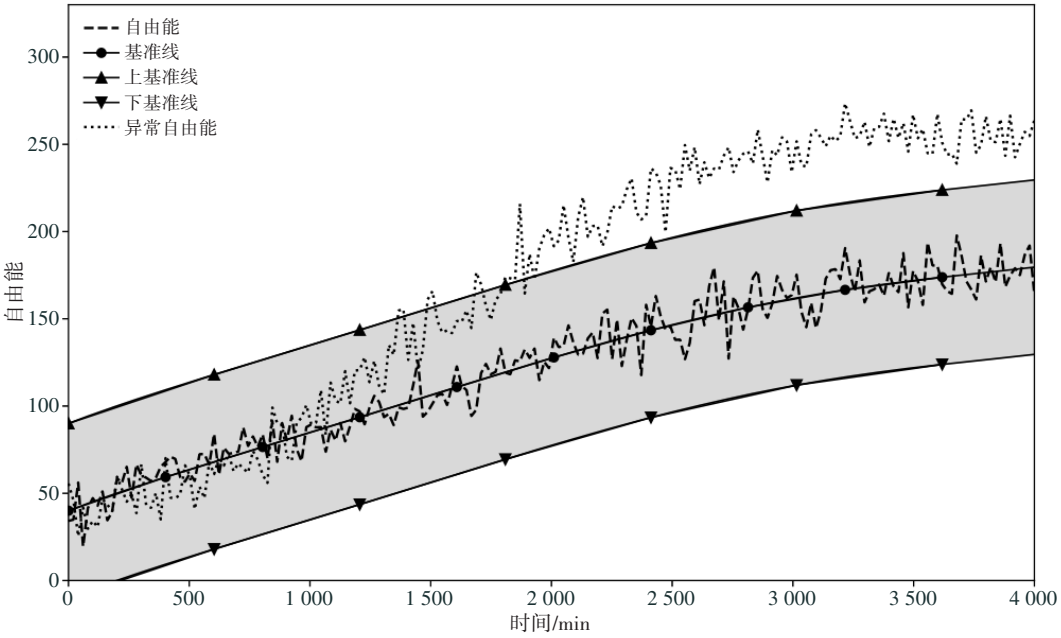

图6　异常检测

表1　虚拟机主要性能指标

CPU状态指标	内存状态指标	磁盘状态指标	网络状态指标
CPU利用率	内存利用率	磁盘利用率	网络负载率
响应I/O操作时间	每秒内存页换出次数	每秒磁盘读次数	每秒接收数据量
进程数	空闲内存大小	进程进行I/O操作等待时间	每秒接收数据包个数
进程的CPU百分比	物理内存大小	平均每次设备I/O服务时间	每秒接收数据包丢包数
响应硬件中断时间	虚拟机内存大小	每秒读I/O设备次数	每秒接收ICMP包个数
进程切换时间	虚拟机占用最大内存	平均I/O队列长度	每秒接收ICMP不可达数量
CPU运行时间	字符设备的缓存容量	虚拟块设备每秒读次数	虚拟网络发送数据量
进程等待CPU空闲时间	内存空间映射情况	每秒磁盘写次数	虚拟网络发送数据量比例
CPU核数	共享内存容量	字节导入	每秒发送数据量
过去20 min系统负载	占物理内存百分比	C盘剩余空间百分比	UDP数据包丢包数
CPU系统中断百分比	每秒内存页换入次数	平均每次设备I/O等待时间	每秒发送数据包个数
CPU空闲百分比	已使用交换空间数量	进行I/O操作的时间百分比	每秒发送数据包丢包数
内核操作CPU比例	块设备的缓存	每秒写I/O设备次数	每秒发送ICMP包个数
CPU等待I/O操作时间	内存写回量	平均每次设备I/O数据大小	每秒发送ICMP不可达数量
CPU响应软件中断时间	通过压缩所节省的内存率	所有磁盘的最大时延	虚拟网络接收数据量
可以访问CPU的时间	总压缩物理内存	字节导出	虚拟网络接收数据量比例
虚拟机所消耗的CPU	内存硬件损坏百分比	虚拟块设备每秒写次数	每秒接收数据量

用内存但不释放，导致内存利用率过高，实现内存资源异常注入；在部分虚拟机上运行磁盘文件读写程序，产生大量的磁盘读写请求，实现磁盘资源异常注入；利用httperf工具[17]向部分虚拟机发送大量的http请求，致使网络负载率过高，实现网络异常注入。

本文使用深度学习框架TensorFlow和Keras进行虚拟机异常检测模型训练，并在GPU服务器上部署运行以提升学习速率。在计算过程中，将LSTM模型的迭代次数配置为时间窗口长度，批处理输入样本数量为128，LSTM模型隐藏层维度为64，Dropout比率为0.2。GRBM模型的可见单元与隐藏单元数量均为64，学习率为0.001，目标函数如公式（11）所示，使用Adam方法进行优化。

3.3　评价指标

对海洋云平台虚拟机进行异常检测时，会出

现如表2所示的4种情况。

为了验证海洋云平台虚拟机异常检测模型的有效性，本文使用异常检测领域常用的评价指标召回率、精确率、F1分数、误报率和正确率来进行评价。召回率*REC*表示虚拟机异常并被正确检测出异常的概率；精确率*PRE*表示所有被检测异常的虚拟机中真实状态为异常的概率；F1分数*F1*为召回率和精确率的调和平均值，值越大说明召回率和精确率的兼顾性越好；误报率*FPR*表示虚拟机正常而被错误检测为异常的概率；正确率*ACC*表示预测结果和虚拟机真实状态一致的概率。各评价指标的计算公式如下所示：

$$REC = \frac{TP}{TP + FN} \tag{14}$$

$$PRE = \frac{TP}{TP + FP} \tag{15}$$

$$F1 = \frac{2REC \times PRE}{REC + PRE} \tag{16}$$

$$FPR = \frac{FP}{FP + TN} \tag{17}$$

$$ACC = \frac{TP + TN}{TP + FN + TN + FP} \tag{18}$$

3.4 结果及分析

为得到最优参数基准模型，通过召回率*REC*、误报率*FPR*和正确率*ACC*来确定参数γ的取值。在海洋云平台虚拟机异常检测中，γ值越大，说明虚拟机被判定为正常的概率越大，因此*REC*会越高，*FPR*会越低。将γ设置为1、2、3和4，然后分别计算基准模型的召回率*REC*、误报率*FPR*和正确率*ACC*，结果如图7所示。当γ取值为4时，将所有的虚拟机判定为正常，得出*REC*=1，*FPR*=0。在海洋云平台运行中，虚拟机出现异常是小概率事件，所以当γ取值为4时，*ACC*虽然最小但仍达到0.927，与γ取值为1时得出的*ACC*值0.934差距不大。当γ取值为3时，*ACC*最高，达到0.979，此时*REC*=0.974，*FPR*=0.017。通过上述分析，将参数γ设置为3。

为验证所提模型的有效性，对Bi-LSTM模型[12]、本文提出的LSTM-GRBM模型、LSTM模型以及GRBM模型进行实验比较。在海洋云平台虚拟机异常检测中，得出不同模型的召回率*REC*、精确率*PRE*、F1分数*F1*、误报率*FPR*和正确率*ACC*，结果如表3所示。

表2 异常检测出现的四种情况

真实状态	预测结果	
	异常	正常
异常	*TP*：虚拟机异常并被检测为异常	*FN*：虚拟机异常并被检测为正常
正常	*FP*：虚拟机正常并被检测为异常	*TN*：虚拟机正常并被检测为正常

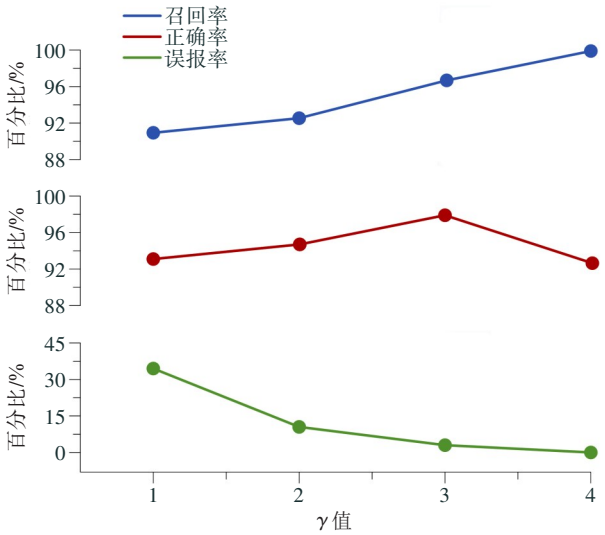

图7 不同参数基准模型的召回率、正确率和误报率

由表3可以看出，通过引入多通道机制，Bi-LSTM模型的整体检测效果明显优于LSTM模型；单独使用GRBM模型的检测效果优于单独使用LSTM模型；Bi-LSTM模型在*PRE*和*FPR*方面稍逊于GRBM模型，其他指标均优于GRBM模型；通过综合运用LSTM与GRBM两种网络模型，能够大幅提升单独使用LSTM模型或GRBM模型的异常检测能力，表3直观地表明本文提出的LSTM-GRBM模型的异常检测效果最好。

表3 不同模型的异常检测性能比较

模型	召回率	精确率	F1分数	误报率	正确率
Bi-LSTM	0.928	0.965	0.971	0.057	0.962
LSTM	0.899	0.936	0.946	0.231	0.939
GRBM	0.922	0.968	0.952	0.052	0.958
LSTM-GRBM	0.974	0.983	0.977	0.017	0.979

4 结语

本文提出了一种基于LSTM-GRBM模型的海

洋云平台虚拟机异常检测算法，通过LSTM模型处理虚拟机时序性能指标，并利用GRBM模型构建参数基准模型，能够准确检测出虚拟机异常，有效保障了海洋云平台稳定运行。本文将提出模型LSTM-GRBM与Bi-LSTM、LSTM及GRBM模型进行实验比较，通过*REC*、*PRE*、*F*1、*FPR*和*ACC*性能评价指标测算表明，LSTM-GRBM模型在海洋云平台虚拟机异常检测中表现最好。本文算法可以应用于不同场景及不同领域的云平台虚拟机异常检测，以便及时发现并处理异常行为。本文目前仅对海洋云平台中操作系统为Windows Server 2016、Windows Server 2019和Windows Server 2022的虚拟机进行异常检测，后续将进一步改进模型以增强其泛化能力，实现操作系统为其他版本及类型虚拟机的异常检测。

参 考 文 献

[1] 相文玺, 臧琦, 韩志聪, 等. 海洋监管信息化建设技术体系框架探究[J]. 海洋信息技术与应用, 2024, 39(2): 98-105.

[2] 杨锦坤, 韩春花, 韦广昊, 等. 海洋大数据的内涵、现状与发展趋势展望[J]. 海洋信息技术与应用, 2023, 38(1): 1-8.

[3] 刘芳, 曹进克. 云服务器虚拟机通信串口数据安全性监控仿真[J]. 计算机仿真, 2023, 40(8): 174-177, 190.

[4] 邢凌凯, 张健. 基于HPC的虚拟化平台异常检测技术研究与实现[J]. 信息网络安全, 2023(10): 64-69.

[5] GHAHRAMANI Z. Probabilistic machine learning and artificial intelligence[J]. Nature, 2015, 521(7553): 452-459.

[6] MOUSAVI A, PATEL A B, BARANIUK R G. A deep learning approach to structured signal recovery[C]//53rd Annual Allerton Conference on Communication, Control, and Computing, September 29-October 2 2015, Monticello, Illinois, USA: IEEE: 1336-1343.

[7] KHREICH W, KHOSRAVIFAR B, HAMOU L A, et al. An anomaly detection system based on variable N-gram features and one-class SVM[J]. Information and Software Technology, 2017, 91(11): 186-197.

[8] MISHRA P, KHURANA K, GUPTA S, et al. VMAnalyzer: Malware semantic analysis using integrated CNN and Bi-Directional LSTM for detecting VM-level attacks in cloud [C]//2019 Twelfth International Conference on Contemporary Computing (IC3), August 8-10, 2019, Noida, India: IEEE; 1-6.

[9] ZHANG H, ZHANG W, LV Z, et al. MALDC: a depth detection method for malware based on behavior chains[J]. World Wide Web, 2020, 23(2): 1-20.

[10] 贺寰烨, 林果园, 顾浩, 等. 云虚拟机异常检测场景下改进的LOF算法[J]. 计算机工程与应用, 2020, 56(23): 80-86.

[11] 杨光. 云架构下的虚拟机异常行为检测方法研究及系统实现[D]. 成都: 四川师范大学, 2022.

[12] 王开放, 姜瑛. 云环境下基于动态滑动窗口多通道Bi-LSTM的虚拟机故障预测模型[J]. 计算机应用研究, 2023, 40(3): 855-862.

[13] HOCHREITER S, SCHMIDHUBER J. Long Short-Term Memory[J]. Neural Computation, 1997, 9(8): 1735-1780.

[14] HINTON G E, SRIVASTAVA N, KRIZHEVSKY A, et al. Improving neural networks by preventing co-adaptation of feature detectors[J]. Computer Science, 2012, 3(4): 212-223.

[15] HINTON G E. A practical guide to training Restricted Boltzmann Machines[J]. Momentum, 2010, 9(1): 926-947.

[16] BENMAKRELOUF S, ST-ONGE C, KARA N, et al. Abnormal behavior detection using resource level to service level metrics mapping in virtualized systems[J]. Future Generation Computer Systems, 2020, 102(1): 680-700.

[17] 陈宣, 罗军, 谭郁松, 等. 集群系统中自适应负载反馈平衡策略的研究[J]. 计算机应用与软件, 2006, 23(8): 12-13, 29.

（本文编辑：李红军）