异常检测在网络安全防护中的应用研究

Research on Application of Anomaly Detection Algorithm in Network Security Protection

刘　洋,翟　锐,巩　坤(中国联通山东分公司,山东 济南 250014)
Liu Yang,Zhai Rui,Gong Kun(China Unicom Shandong Branch,Jinan 250014,China)

摘　要:
图像异常数据作为网络安全检测的核心监控对象,面临着样本不均衡、数据缺乏标注以及异常形式多样化等挑战,针对这些问题,创新性地提出了自信息量挖掘模块,旨在学习已知类别样本的数据模式;同时提出了三元组信息量学习策略,优化类别信息学习和已知类别的数据模式学习,最终实现了在网络安全防护场景中对图像的未知类别样本的异常检测。实验结果表明,异常检测算法可以有效提升网络安全防护的准确性,在实际应用中表现出色。

关键词:
深度学习;异常检测;网络安全;数据模式
doi:10.12045/j.issn.1007-3043.2024.08.005
文章编号:1007-3043(2024)08-0024-05
中图分类号:TP181
文献标识码:A
开放科学(资源服务)标识码(OSID):

Abstract:
Abnormal image data, as the core monitoring target of network security detection, faces challenges such as sample imbalance, lack of data annotation, and diverse forms of abnormalities. To address these issues, it innovatively proposes a self-information mining module aimed to learn the data patterns of known-category samples. Simultaneously, a triplet information learning strategy is introduced to optimize category information learning and known-category data pattern learning, ultimately enabling the detection of abnormalities for unknown class samples of images in the context of network security protection. Experimental results show that the anomaly detection algorithm can effectively improve the accuracy of network security protection, demonstrating excellent performance in practical applications.

Keywords:
Deep learning;Anomaly detection;Network security;Data pattern

引用格式:刘洋,翟锐,巩坤. 异常检测在网络安全防护中的应用研究[J]. 邮电设计技术,2024(8):24-28.

1 概述

在数字化转型的时代浪潮中,网络安全已不仅仅是一个技术问题,更是涉及到国家安全、社会稳定、经济发展等多方面的战略性问题。随着计算机技术的不断发展和网络环境的日益复杂多变,传统的网络安全防护措施难以应对层出不穷的网络攻击手段。而随着深度学习技术的发展和引入,为网络安全领域带来了革命性的变革。

深度学习技术通过构建深度神经网络模型,能够自动学习和提取网络数据中的复杂特征,进而实现精准的识别、分类和预测。深度学习还能够与其他网络安全技术相结合,形成更加全面的安全防护体系。例如,可将深度学习模型与入侵检测系统相结合,实现对网络流量的实时分析和监控;也可将深度学习模型用于用户行为分析,通过识别异常行为来发现潜在的安全威胁。这种多层次的安全防护体系能够大大提升网络安全水平,有效防范各类网络攻击。当然,深

收稿日期:2024-06-14

度学习技术在网络安全领域的应用也面临着一些挑战，如何有效地收集、处理和分析网络安全领域的数据，是深度学习技术在网络安全领域应用的关键。

基于上述现状和问题，本文提出一种基于异常检测算法的网络安全技术。该方法使用聚类算法探索样本相关性，并提出了自信息量挖掘模块和三元组信息量学习策略，联合优化类别信息学习和已知类别的数据模式学习。实验结果表明，本文算法可以有效检测未知类别样本，对于防范异常信息入侵具有较好的效果，对于提高网络安全防御能力、网络安全响应效率和准确性具有重要意义。

2 异常检测方法

异常检测是指在所有的数据样本中，检测出对大部分数据样本具有显著偏离性的异常样本。几十年来，异常检测一直受到众多学者的关注，随着风险管理、安全、人工智能安全等领域的崛起，对于异常检测任务的需求也在不断增加。异常检测的方法主要分为以下几类（见图1）。

基于概率统计的异常检测方法可以在给定数据符合假设的分布时取得很好的效果并对结果给出非常好的解释，但是此类方法常常难以用于真实世界的高维数据[1]。基于重构的异常检测旨在构建可以重构正常图像的网络模型[2]。基于分类方法实现异常检测的基本假设是使用已有的特征向量或者变换过后的特征向量训练精准分类器从而检测出真正异常样本[3]。基于深度聚类的异常检测旨在学习表征，使异常在新学习的表征空间中明显偏离聚类集群[4-8]。

3 网络安全技术

目前网络安全防御技术是一个多层次、多方面的复杂体系，涵盖了多种技术手段和策略，旨在保护网络空间免受各种威胁和攻击。基础防御技术目前主要有如下几种：认证与授权、加密技术、防火墙。随着人工智能、大数据等技术的飞速发展，网络安全防御技术也得到了前所未有的进步和革新。以下是现有的主流网络安全防御技术。

a）智能威胁检测与防御。在网络安全领域，智能威胁检测与防御利用人工智能和机器学习技术，专注于对潜在的网络攻击进行高精度的识别和响应。

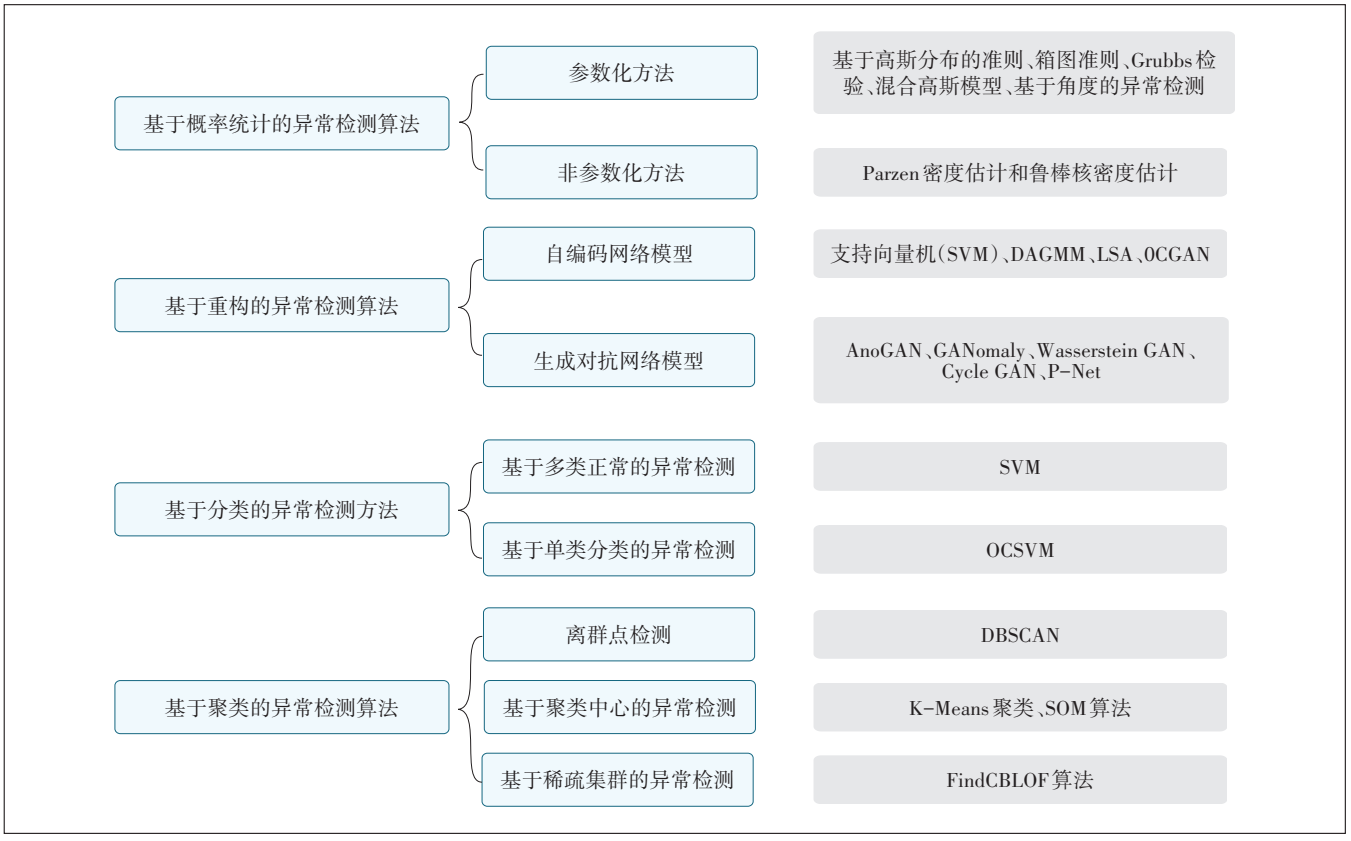

图1　异常检测方法分类

b)安全态势感知。利用大数据技术对网络环境进行全面的监控和分析,通过对海量网络数据的挖掘和分析以及安全态势感知能够发现网络安全的整体态势,如漏洞的分布、攻击的集中度、防御资源的分布等。

4 基于异常检测的网络安全技术

异常数据作为网络安全检测中的重点监控对象,存在样本不均衡、数据缺少标注以及异常多样化等问题。本文致力于研究网络安全领域的未知类别异常检测,用于检测不属于已知正常类别的异常类数据,为网络异常入侵提供安全预警信息,其主要包括3个部分:基于聚类方法的类别信息学习、基于自信息量的数据结构模式学习以及基于三元组的类别模式信息联合优化学习。

4.1 基于聚类方法的类别信息学习

基于聚类方法的类别信息学习在机器学习领域中占据重要位置。这是一种无监督学习方法,它突破了预定义类别标签的限制,通过数据对象间的相似性,智能地将海量数据划分为不同的类或簇。这种方法不仅有助于更深入地理解数据的内在结构和特征,还能为后续的数据分析和应用提供有价值的类别信息,在数据挖掘、图像处理、生物信息学等领域具有广泛的应用前景。

4.2 基于自信息量的数据结构模式学习

原始图像对经过几何变换的自身图像有很强的内在响应,即存在较大的自信息量。本文通过挖掘样本的自信息量,学习已知类别样本的数据结构模式信息,提高特征判别力。给定一个图像样本x和一个几何变换G,可以用$x' = G \cdot x$表示变换后的样本。一个训练良好的特征提取器f_θ应使这2个样本具有相同的标签并且$f_\theta(x) \approx f_\theta(x')$。因此,将$f_\theta(x)$和$f_\theta(x')$之间的距离作为特征不变损失:

$$\min_\theta \sum_{i=1}^{N} l_r\left[f_\theta(x_i),\ f_\theta(x_i')\right] \tag{1}$$

其中,l_r为l_2范数,用于测量原始样本和转换样本的预测结果之间的距离。x和由转换产生的$G \cdot x$可被视为“易正例对”,它可以很好地稳定训练,提高性能。

4.3 基于三元组的类别模式信息联合优化学习

本文最大限度地利用同一样本的深层特征和浅层特征之间的互信息,从而能够保证表征之间的一致性。将2个随机变量D和S之间的互信息转换为由J联合分布及其边缘M的乘积所产生的样本之间的JS散度(Jensen-Shannon divergence,JSD)。相应地,不同层的特征只有在属于同一样本特征时才服从联合分布,否则服从边缘分布。所以基于JS散度的互信息被定义为:

$$\mathrm{MI}^{\mathrm{JSD}}(D,S) = E_J\left\{-\mathrm{sp}\left[-T(d,s)\right]\right\} - E_M\left\{\mathrm{sp}\left[T(d,s)\right]\right\} \tag{2}$$

其中,d对应深层特征,s对应浅层特征,T是用来区分d和s是否是从联合分布中抽样的判别器,$\mathrm{sp}(z) = \log(1 + e^z)$是softflus函数。在鉴别器的输入中加入局部性知识可以提高表征的质量。

本文还引入了正样本对的互信息损失。通过选择具有相同锚点的正样本对和负样本对来构建三重相关性,从而将互信息监督提升为三级监督。

4.4 网络训练及异常检测

在网络训练的过程中,最终目标函数可以表示为:

$$\min_\theta L = \hat{L}_{\mathrm{PG}} + \alpha\hat{L}_{\mathrm{PL}} + \beta L_{\mathrm{MI}} \tag{3}$$

其中,α和β是平衡不同损失的常数,$\hat{L}_{\mathrm{PG}} = L_{\mathrm{PG}} + L'_{\mathrm{PG}}$为整体伪图损失,$\hat{L}_{\mathrm{PL}} = L_{\mathrm{PL}} + L'_{\mathrm{PL}}$为整体伪标签损失。通过计算高置信度伪图和伪标签来指导原始样本和转换样本之间的特征学习,研究不同样本之间的相关性和小扰动下的局部鲁棒性。同时,为了研究鉴别特征学习的特征对应关系,利用伪图选择高置信度的正负样本对进行三重互信息优化。

图2为整个网络的训练流程图。最终训练得到的网络模型“记忆”了已知类别的类别模式,对于已知类别样本判定为较大的自信息量,而未知类别样本则判定为较小的自信息量,由此实现了对未知类别数据样本的异常检测。

5 数据输出及结果验证

为验证本算法的有效性,训练及测试数据采用国际公开的CIFAR-10数据集[5]。本文将接收者操作特征曲线(Receiver Operating Characteristic,ROC)下的面积(Area Under Curve,AUC)作为最终的性能评估结果。为了验证所提算法的有效性和先进性,本文对比了近年来异常检测领域的多种优秀方法,分别是:OC-SVM[9]、KDE[10]、VAE[11]、PixCNN[12]、MemAE[13]、OCGAN[14]和GANomaly[15]。

表1展示了本文算法与对比算法在CIFAR-10数

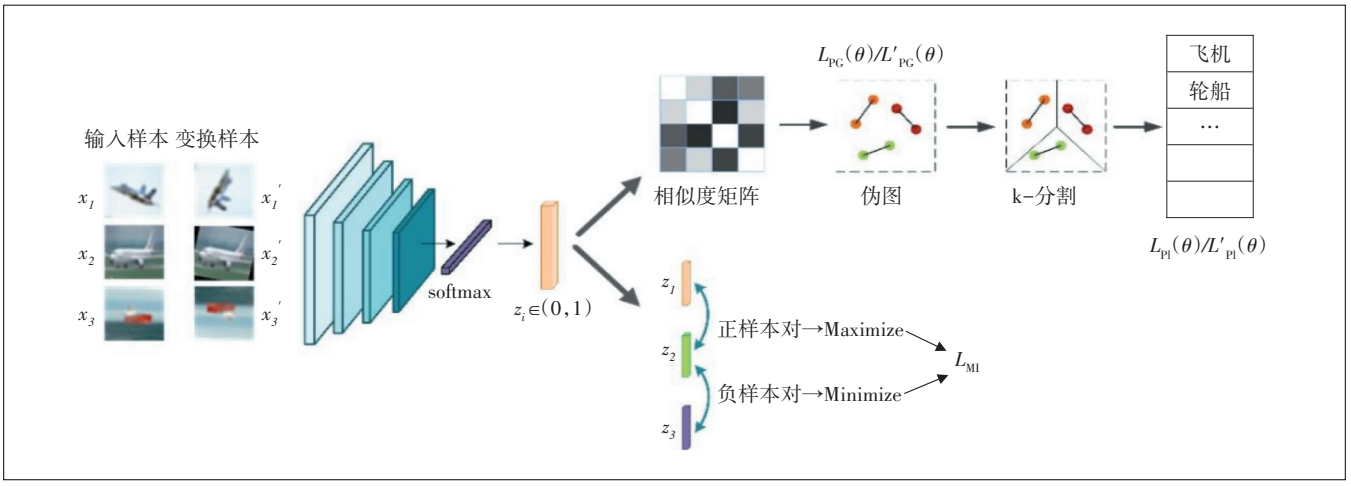

图2　网络训练流程

表1　本文算法与某些最新算法在CIFAR-10数据集上的不同类别的AUC值的比较结果（单位：%）

类别	OC-SVM[9]	KDE[10]	VAE[11]	PixCNN[12]	MemAE[13]	OCGAN[14]	GANomaly[15]	本文算法
飞机	63.00	65.80	70.00	78.80	–	75.70	–	**81.70**
汽车	44.00	52.00	38.60	42.80	–	53.10	–	**60.20**
鸟	64.90	65.70	67.90	61.70	–	64.00	–	**68.10**
猫	48.70	49.70	53.50	57.40	–	62.00	–	**66.34**
鹿	73.50	72.70	74.80	51.10	–	72.30	–	**75.30**
狗	50.00	49.60	52.30	57.10	–	62.00	–	**63.10**
青蛙	72.50	75.80	68.70	42.20	–	72.30	–	**73.80**
马	53.30	56.40	49.30	45.40	–	57.50	–	**65.10**
船	64.90	68.00	69.60	71.50	–	82.00	–	**83.60**
卡车	50.80	54.00	38.60	42.60	–	55.40	–	**66.31**
平均	58.56	60.97	58.33	55.06	60.88	65.66	69.51	**70.36**

据集上的评估结果，表1中数值为图像级的AUC指标，加粗显示代表效果优于其他算法。从表1可以看出，本文所提出的方法得到的图像级AUC平均为70.36%，高于二分类算法（OC-SVM）12个百分点，高于GANomaly方法1个百分点，高于OCGAN方法5个百分点。实验结果表明本文所提出的方法在CIFAR-10数据集上异常检测性能优于其他几个算法，证明本文方法能够有效判断类级别的异常情况。

6 异常检测技术在网络安全领域的落地应用

6.1 网络流量监控与分析

企业内部网络每天承载着大量的数据传输和访问请求。网络流量监控是网络安全的关键环节。通过对网络流量的实时监控和分析，可以检测异常流量模式，预防DDoS攻击、网络入侵等安全威胁。为了保障数据安全和防止未经授权的访问，企业可以部署基于图像异常检测的网络安全威胁感知系统（见图3）。

通过监控网络流量数据，将网络流量数据转化为图像形式，如流量热力图、流量时序图等，随后应用本文提出的图像异常检测算法，对这些图像进行分析，识别出与正常流量模式不同的异常图像。结合深度学习算法，对异常图像进行进一步分类和识别，确定异常流量的类型和来源。最后，根据检测结果，采取相应的安全措施，如阻断异常流量、发送告警信息等。

6.2 恶意代码检测

恶意代码是网络攻击的主要手段之一。通过检测网络中的恶意代码，可以有效预防病毒、木马等恶意软件的传播和破坏。利用图像识别技术，可将恶意代码的特征（如代码结构、函数调用关系等）转化为图像形式。随后应用图像异常检测算法，对这些图像进

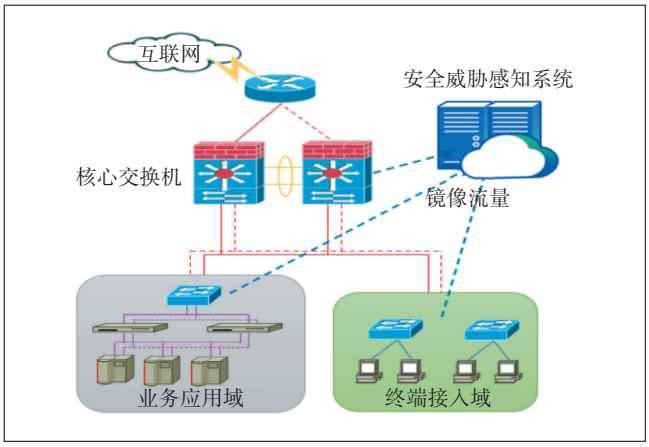

图3　网络安全威胁感知系统

行分析，识别出与正常代码不同的异常图像。结合机器学习算法，对异常图像进行进一步分类和识别，确定恶意代码的类型和危害程度。最终，根据检测结果，采取相应的处理措施，如隔离感染主机、清除恶意代码等。

6.3 用户行为分析

用户行为分析是网络安全管理的重要方面。通过对用户在网络中的行为进行分析，可以及时发现异常行为，预防内部攻击和数据泄露。利用图像识别技术，可将用户行为数据，如登录时间、操作记录等，转化为用户行为时间序列图、用户行为模式图等。应用图像异常检测算法，识别出与正常用户行为模式不同的异常图像。结合深度学习算法，对异常图像进行进一步分类和识别，确定异常行为的类型和潜在风险。最终，根据检测结果，采取相应的处理措施，如限制异常用户权限、进行安全审计等。

7 结论与展望

本研究基于异常检测技术，提出了一种更有效的网络安全检测实现方式。与传统的网络安全技术相比，基于异常检测的网络安全检测技术在无监督的条件下，采用聚类和信息量学习联合优化，挖掘数据模式信息，检测不属于已知正常类别的异常类数据。本研究为网络安全异常入侵检测提供了新的思路和方法，对于提高网络安全防御能力具有重要意义。

参考文献：

[1] NTALAMPIRAS S，POTAMITIS I，FAKOTAKIS N. Probabilistic novelty detection for acoustic surveillance under real-world conditions [J]. IEEE Transactions on Multimedia，2011，13(4)：713-719.

[2] AL-QATF M，LASHENG Y，AL-HABIB M，et al. Deep learning approach combining sparse autoencoder with SVM for network intrusion detection[J]. IEEE Access，2018(6)：52843-52856.

[3] SCHÖLKOPF B，PLATT J C，SHAWE-TAYLOR J，et al. Estimating the support of a high-dimensional distribution[J]. Neural Computation，2001，13(7)：1443-1471.

[4] JAING M F，TSENG S S，SU C M. Two-phase clustering process for outliers detection[J]. Pattern Recognition Letters，2001，22(6/7)：691-700.

[5] HE Z Y，XU X F，DENG S C. Discovering cluster-based local outliers [J]. Pattern Recognition Letters，2003，24(9/10)：1641-1650.

[6] JIANG S Y，SONG X Y，WANG H，et al. A clustering-based method for unsupervised intrusion detections[J]. Pattern Recognition Letters，2006，27(7)：802-810.

[7] SCHUBERT E，SANDER J，ESTER M，et al. DBSCAN revisited，revisited：why and how you should (still) use DBSCAN[J]. ACM Transactions on Database Systems(TODS)，2017，42(3)：1-21.

[8] KRIZHEVSKY A，NAIR V，HINTON G. The CIFAR-10 dataset[EB/OL]. [2024-01-10]. https://www.cs.toronto.edu/~kriz/cifar.html.

[9] SCHÖLKOPF B，WILLIAMSON R，SMOLA A，et al. Support vector method for novelty detection[C]//Proceedings of the 12th International Conference on Neural Information Processing Systems. Cambridge：MIT Press，1999：582-588.

[10] PARZEN E. On estimation of a probability density function and mode [J]. Annals of Mathematical Statistics，1962，33(3)：1065-1076.

[11] KINGMA D P，WELLING M. Auto-encoding variational bayes[EB/OL]. [2024-01-10]. https://arxiv.org/abs/1312.6114.

[12] VAN DEN OORD A，KALCHBRENNER N，VINYALS O，et al. Conditional image generation with PixeLCNN decoders[EB/OL]. [2024-01-10]. https://arxiv.org/abs/1606.05328.

[13] GONG D，LIU L Q，LE V，et al. Memorizing normality to detect anomaly：memory-augmented deep autoencoder for unsupervised anomaly detection[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway：IEEE，2019：1705-1714.

[14] PERERA P，NALLAPATI R，XIANG B. OCGAN：one-class novelty detection using GANs with constrained latent representations[C]// 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway：IEEE，2019：2893-2901.

[15] AKCAY S，ATAPOUR-ABARGHOUEI A，BRECKON T P. GANomaly：semi-supervised anomaly detection via adversarial training[C]//Computer Vision - ACCV 2018. Cham：Springer，2019：622-637.

作者简介：
刘洋，毕业于天津大学，助理工程师，硕士，主要从事AI信息安全管理与数字化赋能等工作；翟锐，毕业于中国科学院大学，工程师，博士，主要从事算力网络建设工作；巩坤，毕业于中国海洋大学，助理工程师，硕士，主要从事AI数字化建设相关工作。