浅析网络安全态势感知与防范技术

欧战祥，邓路华，陈金源
（郴州市公安局，湖南 郴州 423000）

摘 要：随着网络规模的扩大，网络安全面临着更加复杂的形势，快速感知、发现和处置网络安全遇到的风险显得愈发重要。本文介绍了网络安全态势感知平台的功能和架构，以及态势感知平台所使用的数据融合、数据挖掘、特征提取和态势分析等网络安全技术，并对态势感知攻击检测模型的创建和态势评估方法等网络安全领域的最新进展和关键技术进行综述，为网络安全领域的研究和实践提供有益的参考。

关键词：大数据；网络安全；态势感知平台；模型分析
中图分类号：TN915.08 **文献标志码：**A DOI:10.3969/j.issn.1672-8173.2023.05.006

当前，网络问题在全球范围内日益显著，包括恶意代码攻击、信息泄露等，已经给网络安全带来了巨大挑战。尽管已经存在多种网络安全检测技术，但随着网络环境的不断演变和日益增长的数据量，新的网络攻击方式具有更强的隐蔽性，传统的防御方法逐渐显得落后。因此，网络安全领域需要不断创新和改进方法，以更好地适应不断演进的网络威胁。已有研究表明，大数据技术的兴起为网络安全态势感知和技术分析提供了新的解决思路。然而，尽管大数据技术取得了一些进展，但仍然存在许多不足之处。首先，大数据在网络安全领域的应用尚未充分挖掘和利用，需要更多的研究来发掘其潜力[1]。其次，机器学习和数据挖掘等技术在网络安全中的具体应用和效果有待进一步验证和改进。此外，网络安全态势感知平台的构建和攻击事件的实时监测等方面仍存在挑战。因此，需要深入研究网络安全态势感知和技术，将大数据技术、机器学习等前沿技术与网络安全领域相结合，以更好地应对网络安全威胁[2]。

1 网络安全态势感知平台基本功能和架构

1.1 功能分析

网络安全态势感知平台的功能是及时应对潜在的网络攻击，确保网络和信息资产的安全。平台能够实时监控网络流量，不仅包括入站和出站的数据流，还包括内部网络流量，通过持续监控网络流量以检测异常或可疑的行为，如大规模数据传输、频繁的连接尝试或不寻常的数据包模式[3]。平台具备异常行为检测能力，能够自动识别与正常网络活动不符的运行模式[4]，这包括检测未经授权的访问、异常的登录尝试、异常的数据上传或下载、异常的系统资源使用等，通过实时监测和分析这些异常行为，平台能够快速定位潜在的风险因素[5]。平台可自主收集各种威胁情报，包括来自外部情报源、安全组织和内部网络的数据，这些情报涵盖已知的威胁签名、攻击者的行为模式、漏洞信息和恶意软件信息等，通过对上述安全情报进行实时分析，并与当前网络活动进行关联，以识别可能的威胁。

1.2 平台架构

网络安全态势感知平台架构由多个关键组件组成，主要包括数据采集组件、数据存储组件，还包括数据处理与分析组件、威胁情报组件、异常检测与行为分析组件、可视化与报告组件、告警与响应组件、用户权限与访问控制组件等。每个组件具有特定的任务和功能，以确保全面的网络安全监控和响应，是保障平台高效运作的关键。数据采集是网络安全态势感知平台运行的第一步，它涵盖了从各种网络设备和传感器中收集数据的具体过程，这些数据源包括防火墙、入侵检测系统、网络设备（如交换机和路由器）、终端设备、应用程序日志以及外部情报源，数据采集需要高度灵活性，以适应多样的数据格式和传输协议[6]。

收稿日期：2023-07-20
作者简介：欧战祥（1967—），男，湖南汝城人，高级工程师，研究方向为网络安全

数据存储是网络安全态势感知平台的核心，负责将大量的网络流量数据和日志数据存储在可扩展存储系统，上述数据均为海量存在，要求数据存储系统必须具备高度的可伸缩性和可靠性，常见的存储技术包括关系型数据库、分布式存储系统和云存储解决方案[7]。

1.3　平台运行机制

网络安全态势感知平台通过数据整合、数据挖掘、特征提取、态势分析、可视化技术、决策技术等多种网络安全技术来确保平台功能的实现。数据整合可将数据采集组件采集的多个源头的数据汇总到一个统一的平台中，建立全面的网络安全画像，让网络管理员和分析师能够更好地了解网络上发生的事情，包括潜在的威胁和异常活动。数据挖掘技术可在大量数据中发现模式、偏差和异常，识别潜在的攻击行为，如异常流量、登录尝试和恶意软件活动等，并自动检测和报警关键的网络安全事件，使网络管理员能够更快速地响应威胁。特征提取可从原始数据中提取有用信息，例如网络流量的统计数据、数据包的属性或事件的时间戳等，这些数据和特征多用于建立攻击检测模型，进而挖掘安全态势信息、深层次分析数据信息，从不同角度了解网络所面临的安全威胁情况，帮助识别攻击行为和异常活动。可视化技术可直观呈现数据，实现快速识别模式、趋势和异常行为。决策技术是由系统自动采取行动、自动处置威胁，能降低潜在风险。

2　态势感知平台的关键网络安全技术及作用

2.1　数据融合

网络安全数据多种多样，来自防火墙、入侵检测系统、网络流量分析、日志文件、终端设备和外部情报等多个源头，可能以不同的格式、协议和结构存在。数据融合将来自多个源头的数据整合到集中的平台，做到全面获取数据信息，统一处理这些异构数据，以提升网络安全洞察能力，例如，同时分析入侵检测系统的警报、防火墙的日志和外部威胁情报，以帮助确定潜在的威胁并识别已知攻击模式[8]。

2.2　数据挖掘

数据挖掘先要识别到异常行为。当网络中存在各种类型的异常数据，包括异常的流量模式、不寻常的登录尝试、频繁的访问和异常的系统资源使用时，采用数据挖掘技术，可自动检测上述异常行为，迅速发现潜在的网络攻击。数据挖掘针对不同来源的威胁情报，包括已知的攻击模式、漏洞信息和恶意软件样本，自动化分析这些情报，将其与实时网络活动相结合来识别潜在的威胁[9]，并基于数据做出决策。数据挖掘也有助于建立攻击检测模型，这些模型基于历史攻击数据和攻击者的行为模式，可以更好地理解潜在的攻击方式，通过分析攻击特征，可以提前预警可能的威胁，并采取相应的防御措施[10]。

2.3　特征提取

特征提取是以数学方法为基础，耦合海量网络数据，以整合生成数据结果。特征提取是为了将复杂的网络活动转化为可供分析的可量化特征，以便更好地理解网络状态和检测潜在的威胁[11]。此过程涉及从原始数据中提取有用信息，这些信息通常以特定的格式和结构表示。而网络数据是多维的，包括来自各种源头的信息，如网络流量、数据包的属性、时间戳、源和目标 IP 地址等，这些数据通常以原始格式存在而难以直接分析，特征提取的目标就是从这些数据中抽取出对网络安全分析有意义的信息，以便更好地理解网络的状态[12]。特征提取在数据分析和建立攻击检测模型方面发挥着重要的作用。

2.4　态势分析

态势感知是网络安全的第一道防线，它涉及对网络中的实时威胁和事件进行监测和感知的能力，这需要实时的数据收集和分析，以快速识别潜在的攻击。态势感知不仅包括监视网络流量、事件日志和威胁情报，还包括实时的警报和通知机制，能够迅速响应潜在的威胁并通过及时的态势感知以更早地发现攻击并采取适当的措施，以最小化潜在的损失[13]。若潜在的威胁被识别出来，接下来是态势评估的阶段，态势评估涉及对已识别的威胁进行深入分析，确定其严重性和影响，包括研究攻击的来源、攻击者的意图、受影响的系统和数据等方面，有利于优先处理最严重的威胁，确保资源的有效分配和响应计划的制定[14]。

2.5　可视化技术

可视化技术在网络安全领域的应用不可或缺，它为网络安全分析提供了一种直观、交互式的方式来呈现复杂的网络数据和威胁信息，可视化不仅美化了数据呈现，还能够极大地增强网络安全团队的洞察力，

促进更快速、更明智的决策制定。网络安全涉及到大量的数据，包括网络流量、事件日志、威胁情报等，这些数据通常是庞大而复杂的，难以在原始形式下理解[15]。可视化技术通过将数据转化为图形、表格和可交互的仪表盘，提供了一种直观的方式来呈现数据。这使得网络管理员和安全分析师能够在几秒钟内识别模式、趋势和异常行为，而不需要深入分析大量的原始数据[16]。

2.6 决策技术

决策技术涉及基于网络安全态势分析的结果来采取行动，以保护组织的网络和数据资产。网络安全威胁常常需要迅速响应，而人工干预可能耗费较长的时间，决策技术中的自动化响应机制允许系统自动采取行动，可减少潜在的损失，例如，当网络安全平台检测到特定的威胁模式时，它可以自动封锁恶意 IP 地址、隔离受感染的设备或禁用恶意软件，而无需等待人工干预。这种实时自动化响应有助于迅速遏制威胁，降低潜在风险[17]。

3 态势感知攻击检测模型创建与态势评估技术

3.1 数据预处理

数据预处理是将大量原始网络安全数据转化为可用于建模、分析和决策的格式，是网络安全工作的基石之一。数据预处理的质量直接关系到最终的威胁检测、攻击分析和安全决策的准确性和可靠性。数据预处理的首要任务是数据清洗，由于原始网络安全数据通常包含噪声、异常值和缺失值，所以需要对其进行清洗和修复，这包括去除重复数据、填充缺失值、纠正数据格式错误等。数据清洗确保数据的一致性和完整性，使其适合后续的工作。

3.2 创建攻击检测模型

网络安全攻击检测模型的作用在于实时监测网络中的威胁和事件，快速识别潜在的攻击。模型运行依赖于机器学习和深度学习技术，旨在训练检测模型来识别网络攻击，有效地捕捉潜在的威胁行为。随机森林和卷积神经网络（CNN）是两种机器学习方法，可用于构建网络安全的检测模型和实现攻击态势评估。

3.2.1 随机森林学习算法

随机森林是一种集成学习算法[18]，其核心思想是通过构建多个决策树来实现分类或回归任务。假设有一个包含 N 个样本的数据集，每个样本具有 M 个特征，目标是构建一个包含 T 棵决策树的随机森林。

步骤 1：随机抽样。对于每棵决策树的构建，首先从 N 个样本中随机抽取一个包含 n 个样本的训练集（$n<<N$）。这个抽样是有放回的，意味着同一个样本可以在训练集中多次出现。抽样的过程可以表示为

$$D_i = \{(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)\}, \quad (1)$$

其中：D_i 是第 i 棵决策树的训练集，x_n 表示第 n 个样本的特征，y_n 表示第 n 个样本的标签。

步骤 2：特征随机选择。对于每个决策树的构建，需要进行特征的随机选择。在每个节点分裂时，从 M 个特征中随机选择一个子集，通常称为“特征子空间”，假设每个决策树选择的特征子空间包含 m 个特征，其中 $m<<M$。

步骤 3：决策树构建。使用训练集 D_i 和选定的特征子空间，构建一棵决策树。决策树的构建过程通常采用递归分裂方法，基于信息熵或基尼不纯度等指标选择最佳分割点，直到达到停止条件（例如，树的深度达到预定值或节点包含的样本数小于某个阈值）。决策树的构建可以表示为

$$f_k(x) = \sum_{i=1}^{N_k} c_i I(x \in R_i), \quad (2)$$

其中：$f_k(x)$ 表示第 k 棵决策树的输出，N_k 表示叶子节点的数量，c_i 表示第 i 个叶子节点的区域，I 表示叶子节点的类别标签或回归值。

步骤 4：多棵决策树。重复步骤 1 和步骤 3，构建 T 棵不同的决策树，每棵树都使用不同的训练集 D_i 和特征子空间。

步骤 5：投票集成。对于分类任务，当需要对新样本进行分类时，每棵决策树都会投票给出其分类结果，最终的分类结果是多个决策树投票结果的总和，通常选择得票最多的类别作为最终的分类结果。

使用随机森林法构建检测模型要先做好数据准备，收集和整理网络安全数据，包括网络流量数据、日

志文件、攻击事件记录等，将数据进行清洗等预处理，以用于训练模型。根据数据中的正常行为和恶意行为，为每个数据点分配标签，例如正常或恶意，从准备好的数据中选择适当的特征，这些特征可以描述网络流量、行为模式、数据包属性等。将数据集分为训练集和测试集，使用训练集对随机森林模型进行训练，模型将学习如何区分正常行为和恶意行为；测试集评估模型的性能，包括准确率、召回率、精确度等指标。一旦模型被认为足够准确，可以将其部署到实际网络中进行实时威胁检测。

3.2.2　卷积神经网络（CNN）

CNN 是一种深度学习方法，卷积层是 CNN 的核心组成部分，用于提取图像或数据中的特征，在卷积层中使用卷积操作来对输入数据进行滤波，以检测图像中的不同特征，例如边缘、纹理等。

给定一个输入特征图 X 和一个卷积核（也称为滤波器）K，卷积操作可以表示为

$$S(i,j)=(X\times K)(i,j)=\sum_m\sum_n X(i-m,j-n)K(m,n)\,, \tag{3}$$

其中：$S(i,j)$ 是输出特征图中的元素，$X(i,j)$ 是输入特征图中的元素，$K(m,n)$ 是卷积核中的权重，m 和 n 是卷积核的索引。

池化层用于降低特征图的维度，减少计算量，并增加模型的平移不变性。常用的池化操作是最大池化（max pooling）和平均池化（average pooling）。

最大池化操作在一个窗口内选择最大值，并将其作为输出，表示为

$$Y(i,j)=\max_{m,n}X(i+m,j+n)\,, \tag{4}$$

其中：$Y(i,j)$ 是输出特征图中的元素，$X(i+m,j+n)$ 是输入特征图中的元素。

全连接层用于将前面的卷积和池化层的输出连接到输出层，以进行分类或回归任务，全连接层将特征图中的每个元素与一个权重连接，并通过激活函数进行非线性变换，全连接层的输出 Z 可以表示为

$$Z=\sigma(W\cdot X+b), \tag{5}$$

其中：σ 是激活函数（如 ReLU 或 Sigmoid），W 是权重矩阵，X 是前一层的输出，b 是偏置。

卷积神经网络在处理图像和序列数据方面非常有效。使用卷积神经网络构建检测模型同样需要准备网络安全数据，包括网络流量数据、日志文件、攻击事件记录等，并进行数据清洗和特征提取。为数据点分配标签，将数据标记为正常或恶意，将数据转换为适合卷积神经网络的格式（通常是图像数据或序列数据格式）。设计一个卷积神经网络架构，包括卷积层、池化层、全连接层等。使用训练集对卷积神经网络模型进行训练，模型将学习如何从数据中提取有用的特征并进行分类，测试集评估模型性能，通常采用交叉验证等技术来确保模型的泛化能力，一旦模型表现良好，可以将其部署到实际网络中进行威胁检测。

无论是使用随机森林还是卷积神经网络构建的检测模型，都可以用于实时监测网络流量和识别潜在的攻击行为，为攻击因子分析、态势评估提供支撑。

3.3　攻击因子分析

攻击因子分析是对攻击事件进行深入的解剖，以揭示攻击者的意图、方法和可能的动机的一种分析，有助于加强网络安全策略、改进防御措施，提前识别并应对潜在的攻击，攻击因子分析包括攻击来源分析、攻击目标分析等。攻击来源指的是攻击事件的起源，即攻击者的位置、身份和组织背景，攻击者可以是来自互联网的匿名黑客、内部员工、竞争对手，甚至是国家级的网络间谍机构，通过追踪攻击来源，组织可以更好地了解潜在的威胁来源，采取相应的对策。攻击目标是攻击者试图侵犯或破坏的系统、网络或数据，了解攻击目标可以帮助组织识别自身的薄弱点，并采取措施来保护关键资产[19]。

3.4　攻击态势评估

攻击态势评估包括感知、评估和预测等环节，是在网络安全事件发生后，对已识别的威胁进行深入分析，以确定其严重性和影响。攻击态势评估有助于全面理解攻击事件的性质，为网络管理员和安全分析师提供重要的信息，以便让他们能够迅速响应威胁、优先处理最严重的攻击事件，从而提高网络安全性。攻击态势评估要评估数据泄露的风险，这包括确定受影响的敏感数据类型、泄露的数据规模、数据的机密性等[20]，通过深入分析数据泄露事件，以量化潜在的数据损失，估算可能导致的法律责任和声誉损害，帮助组织制定适当的应对措施，包括通知相关当事人、改进数据安全措施和合规性措施等。态势评估要通过历史数据和机器学习模型，尝试预测未来可能的威胁和攻击趋势，以便组织可以提前做好准备和加强安全防护。

4 结束语

综上所述，有效的网络安全态势感知平台和先进技术能够帮助组织及时发现、应对和预测网络威胁，保护关键信息和系统的安全性。然而，网络威胁不断演进，变得更加隐蔽和复杂，因此网络安全领域需要不断进行研究与创新，以应对这些威胁的新形态。在未来的发展中，利用人工智能和自动化技术，网络安全系统将能够更快速地检测和应对威胁，减少对人工干预的需求。随着边缘计算的兴起，边缘安全将成为一个重要领域，以确保连接到边缘设备的数据和通信的安全性。 生物识别技术将广泛用于用户身份验证，提高了身份验证的安全性，如指纹、虹膜和声纹识别。随着云计算的持续增长，云安全将继续是焦点，包括数据加密、访问控制和云安全监测。 跨组织的威胁情报共享将成为一种标准实践，以加强整个生态系统的安全性。

参考文献

[1] 罗宏芳 , 王春枝 . 基于贝叶斯攻击图的光网络安全态势预测研究 [J]. 激光杂志 ,2023,44(8):134-138.
[2] 覃岩岩 , 郭舒扬 , 方雪琴 . 基于 RBF 神经网络的电力信息网络安全态势辨识研究 [J]. 电子设计工程 ,2023,31(16):143-146,152.
[3] 蒲伟华 . 基于大数据背景的网络安全态势感知技术分析 [J]. 网络安全和信息化 ,2023(8):127-129.
[4] 朱珅莹 . 网络安全态势感知及其应用实践研究 [J]. 软件 ,2023,44(7):157-159.
[5] 李泽慧 , 徐沛东 , 邬阳 , 等 . 基于大数据的网络安全态势感知平台应用研究 [J]. 计算机应用与软件 ,2023,40(7):337-341.
[6] 丁昊天 . 基于模糊 C 均值算法的多层次网络安全态势感知方法 [J]. 信息与电脑 (理论版),2023,35(12):76-78.
[7] 曾进 , 李皓杰 . 基于人工智能的网络安全态势感知技术研究 [J]. 信息与电脑 (理论版),2023,35(11):229-232.
[8] 刘秀娟 . 基于深度学习的网络安全态势感知算法研究 [D]. 太原：山西财经大学 ,2023.
[9] 张弛 . 基于机器学习的网络安全态势感知研究 [D]. 张家口：河北建筑工程学院 ,2023.
[10] 宿梦月 . 改进卷积神经网络在网络安全态势评估中的应用研究 [D]. 石家庄：河北师范大学 ,2023.
[11] 孙明伟 . 基于选择性卷积网络的网络安全态势感知研究 [D]. 石家庄：河北师范大学 ,2023.
[12] 吴亚星 . 基于融合模型的网络安全态势感知研究 [D]. 石家庄：河北师范大学 ,2023.
[13] 管磊 , 胡光俊 , 王专 . 基于大数据技术的网络安全态势感知平台研究 [J]. 保密科学技术 ,2016(5):13-19.
[14] 余晴 , 郑崇辉 , 杜晔 . 面向云平台虚拟层的安全态势评估关键技术研究 [J]. 信息网络安全 ,2020,20(7):53-59.
[15] 刘彬 . 网络安全态势评估与预测关键技术研究 [D]. 成都：电子科技大学 ,2021.
[16] 孙东良 . 基于 Spark 的网络攻击态势评估技术的设计与实现 [D]. 北京：北京邮电大学 ,2021.
[17] 韩晓露 . 大数据环境网络安全态势感知关键技术研究 [D]. 北京：北京交通大学 ,2023.
[18] 王奕森 , 夏树涛 . 集成学习之随机森林算法综述 [J]. 信息通信技术 ,2018,12(1):49-55.
[19] 解羽 . 基于神经网络的互联网安全态势预测技术研究 [D]. 沈阳：沈阳理工大学 ,2021.
[20] 钱诚 , 韩戴鸿 , 邬显豪 , 等 . 网络安全态势评估与预测关键技术分析 [J]. 网络安全技术与应用 ,2016(3):18,20.