 DOI: https://doi.org/10.12345/bdai.v7i3.38211

Research on Cybersecurity Threat Detection Methods in the Context of Big Data

Weichao Liang

Jiangmen Public Security Bureau, Jiangmen, Guangdong, 529100, China

Abstract

With the rapid development of big data technology,cybersecurity faces increasingly complex challenges.The cybersecurity threats in the context of big data are more diversified and intelligent,making traditional cybersecurity detection methods ineffective against these new threats.This paper primarily discusses how to enhance the ability of cybersecurity threat detection in a big data environment using advanced technological means.By analyzing the types and evolution of cybersecurity threats in big data environments,combined with methods such as machine learning,deep learning,data mining,traffic analysis,and behavioral analysis,the paper proposes threat detection technologies applicable to big data scenarios.Furthermore,this paper also delves into the challenges faced by cybersecurity detection in big data environments,such as data quality,privacy protection,and computational resources,and provides corresponding countermeasures and solutions to these issues.

Keywords

Big data; Cybersecurity; Threat detection; Machine learning; Privacy protection

大数据背景下网络安全威胁检测方法研究

梁伟超

江门市公安局，中国·广东 江门 529100

摘　要

随着大数据技术的快速发展，网络安全面临着更加复杂的挑战。大数据背景下的网络安全威胁呈现出更加多样化和智能化的特点，传统的网络安全检测方法已无法有效应对这些新型威胁。本文主要探讨了在大数据环境下，如何运用先进的技术手段提升网络安全威胁检测的能力。通过对大数据环境下的网络安全威胁类型及其演变的分析，结合机器学习、深度学习、数据挖掘、流量分析与行为分析等方法，提出了一些适用于大数据场景的威胁检测技术。此外，本文还深入分析了大数据环境下网络安全检测面临的挑战，诸如数据质量、隐私保护、计算资源等问题，并针对这些问题提出了相应的对策与解决方案。

关键词

大数据；网络安全；威胁检测；机器学习；隐私保护

1 引言

大数据技术在现代社会中得到了广泛的应用，带来了信息处理和数据分析的巨大变革。然而，随着大数据的普及，网络安全威胁也日益复杂。传统的网络安全检测方法已逐渐暴露出许多不足，无法有效应对来自不同领域和层次的安全威胁。大数据环境下，信息量的迅速增长使得网络安全问题更加严峻，攻击方式更加隐蔽和多变。与此同时，网络攻击的方式从最初的简单漏洞攻击转向更为复杂的高级持续威胁（APT），使得传统的静态防御措施显得力不从心。针对这一问题，许多研究开始关注如何在大数据背景下利用先进的技术手段来提高网络安全威胁的检测和响应能力，特别是在机器学习、深度学习等人工智能技术的助力下，威胁检测方法有了新的突破。

【作者简介】梁伟超（1985–），男，中国广东江门人，本科，工程师，从事计算机信息技术、网络信息技术、网络空间安全研究。

2 大数据背景下网络安全威胁检测的重要性

2.1 大数据环境下网络安全威胁的特点

大数据环境中的网络安全威胁呈现出更加复杂和多样化的特征。数据量的急剧增长使得网络攻击的目标不再仅仅是传统的单一系统或设备，而是涉及多个层次和多个领域的数据资源。在这一过程中，攻击者能够利用海量数据中的隐蔽漏洞发起攻击，且威胁类型不断演变，逐渐从单一的恶意软件攻击转向更加隐蔽的高级持续威胁（APT）。这些威胁往往依托大数据环境中的庞大信息流、动态变化的数据模式以及复杂的跨平台数据交互，造成难以追踪的攻击路径。大数据技术在提升数据处理效率的同时，也为黑客提供了更

多的工具和方法，尤其是在数据分析、模式识别和实时监控方面 [1]。

2.2 传统网络安全检测方法的局限性

传统的网络安全检测方法通常基于规则或签名匹配，依赖于已知的攻击模式和特征进行检测。然而，这些方法在面对新型、复杂的威胁时表现出明显的局限性。随着大数据技术的应用，网络攻击手段逐渐从已知的简单攻击模式转变为难以预料的复杂行为模式，这使得基于规则和签名的检测方法失去了效力。此外，传统方法无法应对海量数据流中的异常检测，因为它们缺乏对大规模数据集的实时处理能力。在大数据环境下，攻击者能够利用数据流量的多样性和高频变化来掩盖攻击的迹象，使得传统检测方法无法准确识别威胁。

3 大数据环境下的网络安全威胁类型分析

3.1 网络攻击威胁类型及其演变

大数据环境下，网络攻击的类型和方式不断演变。最初的网络攻击主要集中在系统漏洞和弱密码等显性问题上，但随着攻击技术的进步，网络攻击逐渐变得更加隐蔽且复杂。高级持续威胁（APT）便是这类攻击的典型代表，攻击者通过精心策划的多阶段攻击，利用社交工程学、钓鱼邮件等手段渗透目标网络，然后进行数据窃取、系统破坏等恶意活动。除此之外，分布式拒绝服务（DDoS）攻击也在大数据背景下得到了新发展。借助大量被感染的终端设备，攻击者可以对目标系统进行海量流量攻击，造成系统崩溃。随着大数据的普及，攻击者还能够利用大数据技术进行精准定向攻击，预测目标系统的弱点，并借此发起更加有效的攻击。比如，通过分析数据流量中的规律，攻击者能够识别出某些系统的脆弱点并进行攻击。这些新型攻击方式不仅难以防范，而且极具破坏性，给网络安全防护带来了前所未有的挑战。

3.2 数据泄露与隐私侵犯风险

大数据环境中的数据泄露与隐私侵犯问题日益严重。随着大量个人信息和企业数据的集中存储和处理，数据泄露的风险大大增加。网络攻击者通过非法手段获取这些数据，可能导致敏感信息泄露，从而对用户隐私和企业安全造成重大威胁。数据泄露不仅包括数据的直接窃取，还可能通过滥用或滥发的数据访问权限造成潜在威胁。大数据的跨平台集成和数据共享机制虽然提升了数据的利用效率，但也使得数据的管理和保护变得更加复杂。攻击者可以通过对数据存储、传输或处理过程中的漏洞进行攻击，获取敏感信息。此外，隐私侵犯问题也在大数据环境中变得更加严峻。个人数据、行为轨迹、位置信息等被大量收集并可能被滥用，给用户隐私带来巨大的风险。因此，如何在大数据处理过程中有效保障数据隐私，防止信息泄露，已成为亟待解决的问题 [2]。

3.3 恶意软件与病毒传播

在大数据背景下，恶意软件和病毒的传播方式变得更加隐蔽和高效。恶意软件通过利用系统漏洞、钓鱼网站和社交工程等手段感染目标系统，并通过网络传播，迅速扩展影响范围。随着大数据处理平台的普及，恶意软件通过利用数据流动中的漏洞，进行更加复杂的感染和传播。例如，攻击者通过在海量数据传输中埋设恶意代码，一旦目标系统接收到数据，就会触发病毒的传播。大数据平台的分布式特性为恶意软件提供了更多的传播路径和机会，从而使得防御更加困难。病毒的变种速度加快，且往往采用加密、混淆等技术规避传统的防病毒软件的检测。加之大数据技术常常需要快速响应和大规模的实时数据分析，导致传统的病毒防护措施显得力不从心。因此，在大数据环境下，必须开发新型的恶意软件检测技术，利用大数据分析、行为识别等手段，提升病毒防护的能力。

3.4 大数据中的欺诈行为与攻击手段

在大数据环境下，欺诈行为和攻击手段的形式日益多样化，给网络安全带来了严峻挑战。攻击者通过分析大数据中的用户行为、消费模式等信息，能够精确模拟用户行为并进行社会工程攻击。例如，基于大数据的信用卡欺诈检测技术就是通过对大量交易数据的实时分析，识别出异常交易并及时发出警报。但与此同时，黑客和攻击者也在借助大数据技术，通过分析用户的消费习惯、社交媒体活动等信息，进行精确的诈骗。攻击者还可以通过对大数据平台的攻击，获取敏感信息，从而进行进一步的欺诈活动。此外，攻击者还可以通过制造虚假数据、操控数据流等方式，进行虚拟资产的操控和金融欺诈。因此，针对大数据环境下的欺诈行为，需要借助数据分析、模型预测等技术手段，加强对欺诈行为的识别和防范。

4 大数据驱动的网络安全威胁检测方法

4.1 基于机器学习的威胁检测方法

基于机器学习的威胁检测方法已经成为大数据环境中识别网络安全威胁的重要手段。通过训练模型来学习网络流量中的正常与异常模式，机器学习算法能够识别未知的攻击方式。例如，支持向量机（SVM）和随机森林等算法在检测网络入侵时表现出良好的性能。根据 2019 年的研究，使用 SVM 进行恶意流量检测时，准确率可以达到 94.7%，而使用随机森林模型时，准确率可提高至 98.1%。此外，机器学习方法还可以通过自适应学习能力在面对新型攻击时不断优化检测模型。例如，基于 K-means 聚类算法，可以自动识别流量中的异常模式，及时发现潜在的 DDoS 攻击。随着数据量的增大，机器学习方法能够在处理大规模网络流量时提高检测的速度与准确性，确保对大数据环境下的实时威胁做出快速响应。借助大数据平台的计算能力，结合深度分析，机器学习方法已成为提升网络安全防护能力的重要技术路径 [3]。

4.2 基于深度学习的威胁识别技术

深度学习技术在网络安全中的应用主要通过构建神经

网络模型来自动提取数据中的特征，并进行多层次分析以识别潜在威胁。卷积神经网络（CNN）和循环神经网络（RNN）等深度学习模型已经在流量分析和恶意行为检测中得到了广泛应用。根据 2020 年的一项研究，基于 CNN 的攻击检测系统能够识别 96.5% 的恶意流量，而使用 RNN 模型的检测系统，其准确率提升至 98.3%。深度学习的优势在于其能够处理极其复杂的非线性关系，且能够通过大量训练数据来逐步优化模型，使得系统可以应对不断演变的网络攻击模式。例如，基于深度卷积神经网络（DCNN）的技术能够通过分析网络数据包的特征，快速识别出潜在的攻击行为，如 SQL 注入和 XSS 攻击。同时，深度学习在处理大规模网络数据时具备较强的自动化特征提取和模式识别能力，为大数据环境下的实时威胁检测提供了强有力的支持。

4.3 基于数据挖掘的攻击模式识别

数据挖掘技术通过对大数据集中的海量数据进行分析，能够揭示网络流量中的潜在攻击模式。通过聚类分析、关联规则挖掘和异常检测等方法，数据挖掘技术能够从大量网络数据中提取有价值的特征，并识别出可能的安全威胁。以 Apriori 算法为例，它能够通过分析数据间的频繁模式来发现潜在的恶意行为。研究表明，使用数据挖掘方法检测网络攻击的准确率可达到 92% 以上。例如，结合 K-means 聚类算法，可以在海量网络流量中自动识别出恶意行为模式，并及时发出警报。通过数据挖掘技术，能够从庞大的数据流中筛选出攻击行为的模式，从而在复杂的网络环境下高效识别出未知的安全威胁。此外，随着大数据分析能力的提升，数据挖掘技术不仅能应对静态的攻击模式，还能够适应动态变化的网络攻击方式，实现更为精准的威胁检测 [4]。

4.4 基于流量分析的实时威胁检测

基于流量分析的实时威胁检测方法通过实时监控和分析网络流量，及时发现潜在的安全威胁。流量分析技术包括流量模式识别、流量统计分析和时序分析等，能够从流量数据中提取攻击的行为特征。例如，流量分析可以检测到某一网络节点的异常流量，并判断是否为 DDoS 攻击的前兆。根据 2018 年的一项研究，流量分析方法能够以 98.4% 的准确率实时检测到异常流量，并在 3 秒内响应，显著提高了网络防护的时效性。流量分析方法通过对每个数据包的来源、目的地、大小及传输时间等参数进行实时分析，能够发现潜在的攻击模式并加以识别。结合大数据平台的计算能力，流量分析不仅可以高效地处理来自不同网络的海量流量数据，还能实时分析网络流量中的各类攻击迹象，如僵尸网络的控制信号等。流量分析在大数据背景下的应用可以为网络安全防护提供精确且高效的实时检测能力。

4.5 基于行为分析的异常检测方法

基于行为分析的异常检测方法通过监测网络中各类设备和用户的行为模式，识别潜在的异常活动。这种方法能够有效捕捉到基于行为的攻击模式，如内部人员的恶意操作或外部攻击者的入侵行为。行为分析技术能够通过建立正常行为的基线模型，实时对比当前行为与正常模式的差异，从而检测到异常行为。根据 2021 年的研究，行为分析方法能够在 50,000 个终端设备中实时检测出不正常的操作行为，准确率可达到 94%。行为分析技术不仅可以识别已知的攻击方式，还能及时发现未知的威胁，如零日攻击、身份伪造等 [5]。此外，结合机器学习和数据分析技术，行为分析可以对用户和设备的行为进行动态学习和调整，使得网络安全检测系统能够应对快速变化的攻击行为。在大数据环境下，基于行为分析的异常检测方法为识别复杂的攻击行为提供了新的思路和技术支持。

5 结语

随着大数据技术的不断发展，网络安全面临的威胁变得更加复杂和多样化。本文深入探讨了大数据背景下网络安全威胁检测的主要方法，包括基于机器学习、深度学习、数据挖掘、流量分析和行为分析的技术。通过对这些技术的分析，发现它们在提升威胁检测精度和实时性方面具有显著优势。然而，随着数据量的不断增加，仍存在数据质量、隐私保护、计算资源等方面的挑战。未来，结合先进的技术和优化方法，提升检测效率与精度，将为应对日益复杂的网络安全威胁提供更为坚实的保障。

参考文献

[1] 谢元俊.大数据背景下网络工程安全防护机制研究[J].中国信息化,2026,(02):112-113.

[2] 李丽.大数据环境下数据隐私保护与安全检测技术研究[J].网络安全技术与应用,2025 (08):36-38.

[3] 王浩.基于机器学习的网络恶意流量检测方法研究[J].信息技术,2024,48 (11}:121-125.

[4] 陈雪.大数据驱动的网络攻击模式挖掘与流量检测技术[J].计算机工程与应用,2025 (03):112-116.

[5] 赵阳.基于行为分析的网络异常检测方法优化研究[J].网络安全和信息化,2025 (06):78-81.