

网络环境的人工智能防护：挑战与机遇

于洪鹏

(哈尔滨金融学院, 黑龙江 哈尔滨 150030)

【摘要】在人工智能技术不断进步的今天,网络安全已成为互联网发展中不可忽视的焦点。本文概述了网络安全的基本概念,并探讨了人工智能在网络攻击预防、APT攻击识别、漏洞自动挖掘、智能防火墙构建、反垃圾邮件技术,以及加密流量威胁检测等方面的应用。文章指出,通过实施严格的安全管理措施,可以有效地保护网络环境,促进其健康发展。

【关键词】人工智能;网络安全;网络管理

Artificial Intelligence Defense in the Network Environment: Challenges and Opportunities

Hongpeng Yu

(Harbin Finance College, Harbin, Heilongjiang 150030)

Abstract: With the continuous advancement of artificial intelligence technology, cybersecurity has become an indispensable focus in the development of the internet. This paper outlines the basic concepts of cybersecurity and explores the application of artificial intelligence in the prevention of network attacks, identification of APT attacks, automatic vulnerability mining, construction of intelligent firewalls, anti-spam technology, and threat detection in encrypted traffic. The article points out that by implementing strict security management measures, the network environment can be effectively protected and its healthy development promoted.

Keywords: Artificial Intelligence; Cybersecurity; Network Management

1 选题依据及选题意义

人工智能技术发展迅速,同时随着计算机计算能力的不断提高以及大数据分析的广泛应用,人工智能技术已经在多个领域取得极大的发展。从人类语言的语音识别处理到高级的图像绘画,再到具有深层理论架构分析能力的智能推荐系统,可见人工智能技术已经深入到我们生活的方方面面。

同时,人工智能应用的发展给许多行业带来了巨大的变革和机遇。在工业制造领域,人工智能通过优化工业生产流程能够极大促进生产效率,提高产品质量;在金融服务方面,人工智能通过对大数据模型的分析优化投资结构,提高了投资的风险管理能力;在医疗领域,人工智能通过对过往病例学习分析,能够更好的提供对疾病的预测诊断,极大的提高了居民医疗保障水平,提高诊断的准确性和效率。

然而,人工智能技术在信息安全方面也面临着极大的风险和挑战。例如,涉及网络数据的保护问题正随着人工智能技术的应用范围扩大而愈加明显,网络涉及的数据量越大,这也意味着网络安全问题变得更

加严重。其中涉及的大量个人隐私、商业机密和国家安全信息一旦遭到黑客恶意攻击而泄漏,后果将会十分严重。特别是近几年,我国已步入互联网高速发展阶段,“互联网+”的发展模式已融入各行各业,为人工智能技术的发展提供了有力支撑。同时,我们也应清醒地看到,网络安全问题主要源于人为因素导致的数据信息泄露、非法入侵等行为,严重破坏了网络环境的稳定性和保密性,对用户的信息安全和财产安全构成威胁。因此,运用人工智能技术构建完整、安全的网络体系显得尤为重要。通过人工智能技术的应用,可以显著提高安全工作的效率,弥补专业人员不足的问题,使网络安全的防御更加智能。

为应对这些潜在风险,人工智能技术开发者和应用者在使用技术的同时必须加强数据安全的保护。首先是对数据的加密,以及对数据的备份与恢复。其次,需要加强网络安全防护,包括实施安全检测、入侵检测以及打开防火墙等有效措施。除此之外,还应当加强安全管理,加强对信息技术方面安全培训,做出有针对性的安全风险评估。

随着人工智能技术应用的不断扩大，安全问题已成为互联网发展道路上急需解决和关注的重要议题，只有通过采取有效的安全措施，才能够确保人工智能技术的安全和稳定发展。

2 人工智能在网络安全中的应用

(一) 规则产生式专家系统

规则产生式专家系统是人工智能领域中常见的检测技术，这种系统基于推理和逻辑判断来模拟人类专家在特定领域的决策过程，讲这一技术应用到网络安全中，可以识别该系统的网络安全系数^[1]，检测网络安全系统功能，进而查清网络入侵可能的入口，从而达到防范于未然的效果。

同时该系统还拥有交互系统，能够对特定知识进行学习，通过知识库接口能够将知识体系灌输到系统中，存储的知识同时还会进行智能的推理学习，再通过用户接口实现交互，达到模拟决策的效果。

网络环境的整治离不开先进的网络技术支持，人工智能技术在这一过程中需充分发挥其功能。在传统观念中，网络安全保障的核心在于预防和控制的相互协调，而较少关注相关措施的学习与推理能力。尽管这种方式能基本满足安全防御需求，但未能从根本上确保网络安全。基于传统防治方法的局限性，网络信息处理存在较大不确定性。然而，引入人工智能技术后，这一问题得以有效解决，实现了网络防御与理论知识的有机融合，赋予了网络防御手段具备基本的学习与推理能力。

(二) 人工神经网络

人工神经网络^[2]通过学习和训练调整连接权重，进而达到对数据模型的分类、预测和识别。同时是一种类似于人脑神经网络结构的发展的人工智能方向，由大量的人工神经元组成，神经元之间通过连接进行信息传递和处理。通过该方向处理网路安全问题或进行网络安全管理拥有高适应的优点，能够快速识别和处理入侵项目，由于适应性能强大，所以在处理异常检测、资源优化、预测性维护以及自适应网络控制中具有强大优势。

随着网络环境的放大和数据流的急剧增长，传统的安全防护措施面临着前所未有的挑战。不确定性和未知风险成了网络运营的常态，而人工智能凭借其先进的模糊信息处理能力，能够在这种开放的网络环境中有效地识别与解析模糊数据，以便于快速精确地预测和防范潜在威胁。人工智能系统能够通过算法学习

和模式识别，从众多模糊、不完整的数据中挖掘出隐藏的信息，从而提高了网络安全的响应速度和准确度。这种能力使得人工智能不仅仅是防御工具，更是在维护网络安全稳定中发挥着策略形成者的作用。通过大规模的数据分析，人工智能助力网络安全管理变得更加精密和智能，从而大幅度提升防护效率，确保网络环境的安全与可靠。

人工智能技术借助先进算法，实现精准数据开发与计算，从而显著提升资源利用效率，并优化网络数据配置。计算成本的有效降低，很大程度上减少了软硬件系统开发成本，为人工智能技术的广泛应用奠定坚实基础。借助人工智能的智能计算能力，我们得以更高效利用资源，减少不必要成本支出，同时提升网络安全防御质量与效果。这一优势不仅在经济层面具有重要意义，也为实现社会效益提供了可行途径。

(三) 人工免疫技术

人工免疫技术灵感来源于人体免疫系统原理，通过防护检测手段进行网络防护。其中，异常检测通过模拟生物免疫系统的自我非自我识别机制，识别出网络中的异常流量，以此来抵御类似于DDoS攻击等基本网络攻击。

免疫技术中恶意代码检测能够识别恶意软件及病毒，通过即使发现和阻止传播的方式模拟免疫系统的自适应性和演化机制。还有入侵检测、安全策略优化、威胁情报分析等，都通过类模拟免疫系统的系统学习方式达到网络安全管理功能。

(四) 恶意软件生成

鉴于网络安全环境的日益恶化，人工智能在恶意软件的生成及变种研发过程中发挥了关键作用，为攻击者提供了提升威胁能力的新型工具。这主要得益于先进的人工智能算法，如深度学习和自然语言处理等，使得恶意软件能够模拟正常的网络流量模式，规避传统检测手段。通过机器学习，攻击者可以剖析安全软件的检测策略，并根据这些策略生成恶意代码，以规避检测。攻击代码生成器采用随机化手段，持续产生新颖且难以检测的恶意软件变体，使得基于签名的传统防御手段难以应对这些变化。在人工智能技术中，生成对抗网络(GANs)在恶意软件生成领域得到了广泛应用。恶意软件开发者通过训练生成器和判别器的GANs，使其互为竞争、互为学习，以生成具有特定目标的恶意代码。生成器根据所学模式和结构生成恶意代码变体，而判别器则力求识别这些变体与正常代码

的差异。这种对抗性学习过程使得恶意软件生成更具适应性和隐蔽性，传统安全防御机制难以侦测。

（五）APT 攻击检测

在网络安全领域，APT (Advanced Persistent Threat) 攻击因其定制化、多阶段的坚决攻击方式而形成独特的威胁模型，其中包含几个关键特征：强化的针对性、精密的组织协作、悠久的攻击持续期、精湛的隐蔽技巧和巧妙的间接入侵手段。攻击者为规避防御系统的检测而持续进化其策略，导致传统基于签名的检测机制难以适用。于此情景下，人工智能技术显露了其在识别和对抗 APT 攻击中的极大潜力。AI 系统利用机器学习算法，综合行为分析和特征识别，以非依赖特征匹配的方式识别出攻击者的模式，还结合了沙箱技术对可疑活动进行隔离和分析，及大数据技术以洞察全局威胁情报。特别是在检测 C&C (命令和控制) 异常流量、异常的 Web 请求、隐蔽的数据渠道、邮件钓鱼活动以及网络流量异常等方面，AI 已能有效挖掘出攻击行为的微妙迹象，以识别和阻断包括勒索软件、APT₂₉ (也称为“海莲花”)、Stuxnet (“震网”病毒)、BlackEnergy 和 Google Aurora 在内的复杂 APT 攻击。随着人工智能技术的不断深化和完善，这种适应性和前瞻性的检测方法被期待能为网络安全提供一个更具适应性、有效性以及及时性的防护系统。

（六）自动化漏洞挖掘

在当前的网络安全领域，人工智能技术正在引领一场自动化漏洞挖掘的革新。传统的漏洞挖掘方法受限于人工分析的深度和速度，难以应对日益增长且日趋复杂的安全威胁。然而，AI 技术凭借其卓越的数据处理和模式识别能力，使得自动化漏洞挖掘的效率和精准度大幅提升。通过运用先进的机器学习模型，特别是深度学习网络，现代自动化工具能够从海量的代码实例中快速识别和提取潜在的漏洞特征。这些工具不仅能够有效发现已知的漏洞类型，更能通过异常检测揭示出未知的安全弱点。例如，模糊测试 (Fuzzing) 技术与 AI 的结合，能够生成和测试海量随机的输入数据，从而挖掘出程序在处理异常输入时可能产生的漏洞。此外，AI 辅助的静态代码分析借助大数据挖掘和学习算法，精确地识别出代码中违反常规编程模式的部分。同时，动态行为分析技术通过模拟用户操作和网络环境，有效识别运行时漏洞和可能的利用途径。AI 在漏洞挖掘中的应用已经扩展到系统配置错误、恶意软件变种识别等多个维度，为网络安全防护提供了

新的视角和解决方案。为了应对不断演化的攻击手段，研究人员需要进一步加强算法和模型的深入研究，并注重与其他先进技术 (如区块链、云计算等) 的融合，以构建更为先进、全面且响应迅速的网络安全防护体系。

3 人工智能与网络安全管理

（一）威胁检测与预测

人工智能技术通过机器学习算法能够对网络流量、用户行为、系统日志等数据进行分析^[3]，从而识别异常活动，发现潜在威胁。对发现威胁的情报源进行分析和处理，指定与之匹配的安全防御策略以应对未来可能的攻击，同时对恶意代码进行分析，帮助识别代码变种，加强同一类型再次攻击时的系统应对能力。

（二）自动化安全响应

检测到异常之后自动响应处理，能够实现高度自动化，减少人工干预的依赖，通过智能算法分析自动评估和处理安全问题。

（三）异常行为检测

通过深度学习算法分析，识别与正常活动模式不符的网络活动，帮助发现潜在的安全威胁和内部代码恶意行为，从而提高系统网络安全可靠性。

（四）安全日志分析

通过数据挖掘技术对安全日志进行智能化处理，同时智能化分类和标记不同安全类型的事件，对相关特征和模式进行深入学习，能够更好的发现潜在系统网络安全问题，进而做出符合安全规范的处理。

（五）漏洞管理

通过漏洞扫描和分析，自动发现系统和应用程序中的漏洞，并提供自动化的修复建议和操作

（六）智能防火墙

防火墙技术是网络安全防御中的重要手段，近年来融合了人工智能技术取得了新的发展。以 MazeBolt 公司为例，他们利用人工智能技术开发了一款智能防火墙系统，能够实时检测和防御各类网络攻击。该系统通过分析网络流量、用户行为以及特定异常模式，自动学习和适应网络的正常工作方式，并对异常行为做出及时响应，有效应对各类安全威胁。防火墙技术涉及多个子技术，其中包过滤技术和状态检测技术是不可或缺的组成部分。包过滤技术在网络层对数据包进行筛选和分析，通过预设的安全策略检查外部数据包，保障内部通信的安全。同时，状态监测技术能够

全面分析数据流，及时识别并遏制异常连接，有效维护网络安全。相较于传统防御手段，融入人工智能的防火墙技术具有更高的灵活性和安全性。它不仅能够基于过往经验学习和适应新型攻击，还能够运用大数据分析和机器学习算法实现实时威胁识别与应对。因此，融入人工智能技术的智能防火墙在网络安全防御中发挥着关键作用，对保障网络系统的稳定和安全具有重要意义。

(七) 智能反垃圾邮件系统

随着电子邮件在日常生活和工作中的普及，垃圾邮件成了一个常见而令人困扰的问题。不法分子利用邮件中的漏洞传递不正当信息，给个人和企业带来经济损失和困扰。人工智能技术在网络安全防御中的应用为解决这一问题提供了新的途径。以 Tessian 公司为例，他们利用机器学习和深度学习算法开发了智能反垃圾邮件系统，能够自动检测并拦截垃圾邮件。该系统采用智能启发式扫描引擎，对邮件信息进行分析 and 评分，实现自动化的垃圾邮件拦截和删除，大大减轻了人工操作的负担，为信息安全提供了更可靠的保障。这样的智能反垃圾邮件系统不仅提高了工作效率，也有效地减少了垃圾邮件对个人和企业造成的潜在威胁，展现了人工智能技术在网络安全防御中的重要作用。

(八) 加密流量威胁检测

当前网络流量中的加密流量已超过 50%，加密作为保护隐私的重要手段，可以有效保护数据免受窥探。然而，加密也给不法分子提供了隐藏恶意软件的机会，从而带来一系列安全问题。传统的深度报文检测 (DPI) 需要解密原始报文才能进行分析，而这在实际中是不可行的。然而，依托人工智能的特征提取和行为分析方法，可以在不解密报文的情况下提取出恶意软件的特征，从而实现对加密流量中有害威胁的检测。举例来说，对于 HTTPS 流量，通过提取 TLS 连接的初始数据包信息、数据包长度和时间顺序、有效载荷上的字节分布等特征数据，并应用人工智能模型进行推理，可以准确地检测出其中的恶意流量，这种基于人工智能的加密流量威胁检测技术可以有效填补传统方法的

空白，保障网络安全。它不仅能够应对加密流量带来的挑战，还能够提升网络安全防御的准确性和效率，为保护敏感数据和应对网络威胁提供强有力的支持。

(九) AI 深度伪造技术带来的安全威胁

人工智能的迅猛发展使得深度伪造技术日臻完善，该技术通过深度学习算法伪造文本、音频、视频等多媒体内容，从而使欺诈行为的识别与防范变得愈发困难。近期，人工智能在网络诈骗领域的应用，特别是 AI 换脸技术，引起了广泛关注。有一个案例是利用深度伪造技术破解银行人脸识别系统，导致一名未授权的受害者遭受资金损失。这些事件凸显了人工智能在网络安全领域具有双重性质，尤其是 AI 深度伪造技术所带来的潜在威胁，亟待研究并提出有效的技术应对策略，以应对这一发展趋势。

(十) 针对深度伪造与生成式 AI 的应对措施

针对 AI 生成内容的安全威胁，美亚柏科早在 2017 年就成立 AI 研发中心，并于 2019 年专门组建了专项研究团队，致力于开发深度伪造检测技术。

该公司推出的 AI-3300 慧眼视频图像鉴真工作站和 AIGC 内容检测平台皆为业界典范。AI-3300 慧眼工作站整合了超过 40 种视频图像鉴定算法，能高效检测深伪视频图像篡改行为；AIGC 平台则专注于识别如 ChatGPT 所生成的中文文字内容。这些工具不仅提升了鉴定深度伪造内容的准确率，还极大地增强了司法鉴定人员和反诈专家在对抗 AI 诈骗时的技术支持。

4 结语

人工智能在网络安全领域具有广阔的前景和潜力，可以更高效的帮助提高网络安全的自动化管理能力。同时，网络安全管理也能确保人工智能项目在网络安全中的安全，可以相互进步和发展。

参考文献：

- [1] 耿聪聪. 基于人工智能的网络安全管理的研究 [D]. 中国石油大学 (华东), 2018.
- [2] 程银龙. 基于分布式人工智能的网络安全监测关键技术研究 [D]. 哈尔滨工程大学, 2023.
- [3] 廖军, 谈鹏驹, 张冬月等. 人工智能在电信网络中的应用研究 [J]. 信息通信技术与政策, 2020(09):31-34.



Copyright: © 2024 by the authors.

This is an open access article under the terms and conditions of the Creative Commons Attribution (CC BY) license <https://creativecommons.org/licenses/by/4.0/>.