信息融合 97 (2023) 101804

可在ScienceDirect上获取目录列表

信息融合

期刊主页：www.elsevier.com/locate/inffus

人工智能在网络安全中的应用：文献综述与未来研究方向

Ramanpreet Kaur, Du ˇsan Gabrijel ˇci ˇc, Toma ˇz Klobu ˇcar

Jo ˇzef Stefan Institute 的开放系统和网络实验室，斯洛文尼亚卢布尔雅那

文章信息

关键词：
检测
保护
响应
恢复
识别
学习
网络攻击
分类法

摘要

人工智能（AI）是一种强大的技术，可以帮助网络安全团队自动化重复任务，加速威胁检测和响应，并提高其行动的准确性，以加强安全防御，应对各种安全问题和网络攻击。 本文对人工智能在网络安全供应中的应用进行了系统的文献综述和详细分析。 综述结果共有2395项研究，其中236项被确定为主要研究。 本文根据 NIST 网络安全框架，采用主题分析方法对确定的人工智能应用进行了分类。 这个分类框架将为读者提供人工智能在不同背景下提升网络安全的潜力的全面概述。 综述还确定了新兴网络安全应用领域、先进的人工智能方法、数据表示以及为成功采用基于人工智能的网络安全技术而开发的新基础设施等未来研究机会。这些研究机会发生在当今数字转型和多重危机的时代。

1. 引言

网络安全一词指的是一套技术、流程和实践，用于保护和防御网络、设备、软件和数据免受攻击、损坏或未经授权访问[1]。 由于互联设备、系统和网络的指数增长，网络安全变得复杂起来。 这还受到数字经济和基础设施的进步的影响，导致网络攻击的数量大幅增长，并带来严重后果。 此外，研究人员报告称，与国家相关的对手和犯罪对手的持续演变，以及网络攻击的日益复杂，使得即使是最精明的目标也能找到新的侵入方式[2]。

这种演变推动了网络攻击的数量、规模和影响的增加，
并且 需要 实施 基于智能的网络安全
以提供对不断演变的网络攻击的动态防御，并管理大数据。 咨询机构，如国家标准和技术研究所（NIST），还鼓励采用更主动和适应性的方法，通过实时评估、持续监控和数据驱动的分析来识别、保护、检测、响应和记录网络攻击，以防止未来的安全事件[3]。

人工智能是一种引人入胜的工具，可以通过快速分析数百万个事件和追踪各种网络威胁来提供分析和智能，以保护免受不断演变的网络攻击。为了提前预防和应对问题，人工智能可以预测和采取行动。 因此，人工智能越来越多地被整合到网络安全体系中，并在各种用例中用于自动化安全任务或支持人工安全团队的工作。

网络安全领域蓬勃发展，来自人工智能和网络安全领域的研究人员的热情高涨，已经进行了大量研究，以解决与识别、保护、检测、响应和恢复网络攻击相关的问题。

近年来，已经发表了几篇关于网络安全和人工智能应用的综述文章[4–7]。 然而，据我们所知，还没有一篇全面的综述文章涵盖了最新研究，以解释人工智能技术在网络安全活动中的应用细节。 因此，我们的目标是提供一篇系统综述，全面介绍人工智能在网络安全中的用例，并讨论与适应和使用人工智能进行网络安全相关的研究挑战，以供未来研究人员和从业者参考。 表1显示了该研究与近年来综述文章的比较。

我们对人工智能在网络安全提供方面进行了系统的文献综述（SLR），特别关注了NIST网络安全框架[3]定义的五个不同网络安全功能（识别、保护、检测、响应和恢复）内的实际应用。 本研究解决的具体研究问题是

*通讯作者。
电子邮件地址：raman@e5.ijs.si（R. Kaur）。

https://doi.org/10.1016/j.inflflus.2023.101804
收到日期：2023年1月19日；修订日期：2023年3月16日；接受日期：2023年4月6日
在线发布日期：2023年4月7日

SLR是：

- RQ1：人工智能在网络安全提供方面的分类表示是什么？
- RQ2：人工智能在网络安全中的具体用例是什么？
- RQ3：与网络安全相关的人工智能当前研究趋势是什么？
- RQ4：采用人工智能进行网络安全的研究热点和未来研究方向是什么？

为了回答这些研究问题并为研究界提供有价值的成果，在2022年2月之前，我们检查了236篇文章。 然后，进一步分析选定的研究，以确定人工智能在哪些网络安全应用中被使用，选择的人工智能领域以及其产生的影响。 系统性文献综述得出以下结论：

- 一个关于人工智能在网络安全中的分类法，根据网络安全功能、解决方案类别和具体应用案例对所审查的文章进行多层次分类。
- 人工智能在网络安全中的具体应用案例，揭示了利用人工智能能力的潜在领域。
- 文献综述的描述性分析，探索了人工智能在网络安全中的研究趋势。
- 对现有文献的批判性分析，识别出研究领域的研究空白，以激发未来的研究。

本文的其余部分结构如下。 第2节讨论相关背景，介绍和概念化网络安全和人工智能主题，并解释与人工智能在网络安全领域相关的分类范式。

第3节描述了采用的研究方法，以进行
SLR。第4节讨论了数据提取过程，以供
描述性分析和第5节中呈现的最新研究使用。
第6节提供了综合文献的描述性分析。 第7节确定了新研究可以针对的各种研究空白，而第8节指出了我们研究的局限性。 最后，第9节提出了主要结论和研究对网络安全的影响。

2. 背景

本节旨在分析与本综述的关键概念相关的背景信息，包括使用NIST网络安全框架[3]对网络安全的操作定义和AI Watch提出的AI分类法[8]以澄清网络安全的不同应用。

2.1. 网络安全

网络安全制定政策、程序和技术机制，以保护、检测、纠正和防御信息和通信系统及其所包含信息的损害、未经授权的使用或修改，或信息的利用。 技术变革和创新的快速步伐，以及网络威胁的快速演变，进一步复杂化了情况。 为了应对这一前所未有的挑战，基于人工智能的网络安全工具已经出现，帮助安全团队高效地减轻风险并提高安全性。 鉴于人工智能和网络安全的异质性，需要一个统一接受和整合的分类法来研究应用人工智能进行网络安全的文献。 这个结构化的分类法将帮助研究人员和从业人员对需要使用人工智能改进的技术程序和服务达成共识，以实施有效的网络安全。

为了达到这个目的，使用了由NIST提出的一个著名的网络安全框架来理解保护、检测、反应和防御网络攻击所需的解决方案类别[3]。 NIST网络安全框架的核心描述了改善任何组织的网络安全的实践。 该框架的核心包括四个要素：功能、类别、子类别和信息参考。

NIST框架的前两个层级包括5个网络安全功能和23个解决方案类别，用于对已识别的人工智能应用案例进行分类。 这些功能提供了一个全面的视角，以便随着时间的推移管理网络安全的生命周期。 每个功能下列出的解决方案类别为确定改善网络安全的人工智能应用案例提供了一个良好的起点。 选择这两个层级的主要目的是为了将现有的人工智能在网络安全中的文献清晰而直观地分类到相应的解决方案类别中。 所提出的分类法引入了一个与前两个层级一致的第三层，通过指定与网络安全框架的每个层级相对应的基于人工智能的应用案例，如图1所示。 本文提供了所提出的分类法的详细描述，并对人工智能在网络安全中的最新研究进行了综述，详见第5节。

这个分类法为我们的系统性文献综述提供了基础，通过提供相关子领域的描述，涵盖了网络安全解决方案类别的主要方面和基本关键词。 关键词选择的详细描述可以在第3节中找到。

2.2. 人工智能

可以找到与AI系统相关的几个定义，涉及到(a)的

表1
与现有研究的比较。

参考文献	分类法	用例识别	基于防御解决方案的分类			覆盖范围	研究空白	目的
			功能	AI领域				
Wiafle等[4]	否	否	否	是	是	IEEE和ACM数字图书馆	否	提供了关于AI在网络安全中现有研究的概述Zhang等[5]
	否	否	否	否	是	Google Scholar，SpringerLink，ScienceDirect，IEEE和ACM数字图书馆	否	提供了一个受限于用户认证、危险行为监测、网络情况感知和异常流量识别等领域中AI应用的综述Torres等[6]
	否	否	否	否	否	Nil	否	回顾了机器学习技术在垃圾邮件、恶意软件和钓鱼检测中的应用Truong等人[7]
	否	否	否	否	否	Nil	否	回顾了人工智能技术在入侵、恶意软件、高级持续性威胁和钓鱼检测中的应用提出研究
	是	是	是	是	是	Scopus数据库	是	从描述性角度和详细的现状分析探讨了2010年至2022年2月与网络安全中人工智能应用相关的研究

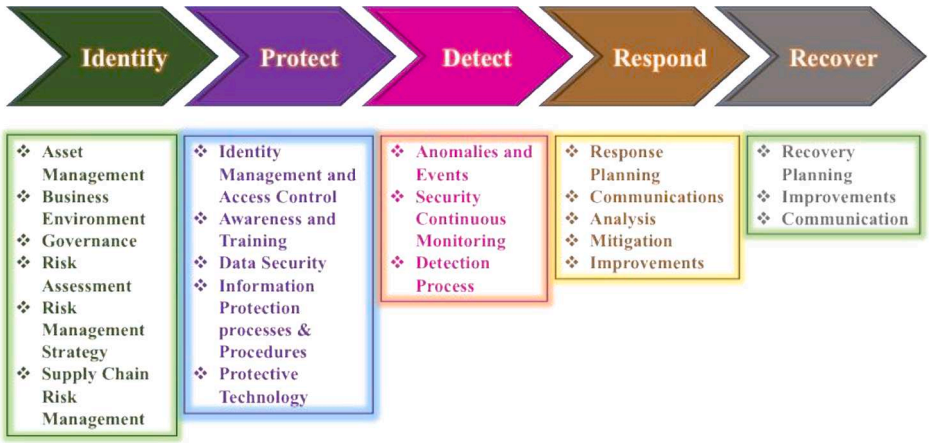


图1. NIST网络安全框架。

它们被使用的领域和（b）AI系统的生命周期阶段，如研究、设计、开发、部署和使用。由于本文的重点是网络安全中的人工智能应用，因此采用了一种普遍但简化的AI定义：“通过分析环境并具有一定程度的自主性来展现智能行为并采取行动以实现特定目标的系统” [9]。 在实际应用中，AI指的是一系列不同的技术和应用，以各种方式使用。 网络安全中的AI应用案例描述了什么样的环境情况是理想的和不理想的，并为序列分配行动。

对于这个SLR，使用Samoili等人提出的AI分类法[8]，该分类法定义了核心和横向AI领域和子领域。
核心AI领域，即推理、规划、学习、通信和感知，被发现非常有用，因为它们涵盖了AI的主要科学领域。推理涉及知识表示和不同的推理方式，而规划还涵盖了搜索和优化。 学习包括机器学习；通信与自然语言处理相关；感知涉及计算机视觉和音频处理[8]。 构成这些AI领域的方法和技术包括但不限于模糊逻辑、基于案例的推理、遗传算法、贝叶斯优化、进化算法、规划图、人工神经网络、深度学习、支持向量机、自然语言处理、文本挖掘、情感分析、图像处理、传感器网络、物体识别和语音处理。

人工智能是一个庞大的、多学科的研究领域，有大量的文献从各种角度讨论其应用和影响，例如技术、运营、实践和哲学。本研究聚焦于文献中讨论上述方法和人工智能在网络安全场景中应用的含义。 它详细分析了人工智能方法如何在网络安全领域中用于识别、保护、检测、响应和恢复。

3. 研究方法

系统性文献综述旨在识别、评估和解释领域内所有可用的研究，以确定潜在的研究空白并突出知识的前沿。 它提供了一个高质量、透明和可复制的综述，以总结大量的研究研究。 本研究采用系统性文献综述方法，原因如下：(i) 人工智能在网络安全领域是一个多样化的领域，有大量的文献；(ii) 本研究旨在回答特定的研究问题；(iii) 它提供的严谨性和可复制性使其成为一项无偏的科学研究。 系统性文献综述的程序将在下面详细描述。

3.1. 文献计量数据库的选择

Scopus和Web ofl Science（WoS）是两个最受欢迎的文献计量数据库。 本研究选择了Scopus数据库，因为其覆盖范围比WoS大近60%[10]。 此外，由于其更广泛的覆盖范围、高级搜索过滤器和数据分析网格，Scopus提供了更好的数据管理。

3.2. 搜索策略

在2021年11月至2022年2月期间，为了进行对人工智能对网络安全影响的全面文献综述，进行了与人工智能和网络安全相关的术语的综合搜索。 搜索使用了针对人工智能和网络安全领域的明确指定的搜索术语，如表2所示。 使用逻辑AND运算符组合了人工智能和网络安全领域的关键词。 在不同关键词内部使用逻辑OR运算符，以找到与每个领域中任何术语相关的研究。 具体而言，人工智能关键词对应于AI Watch提出的人工智能分类法[8]，而网络安全关键词则来自NIST网络安全框架[3]。

3.3. 包含和排除标准

在搜索阶段之后，筛选出了识别出的研究，以排除无关的工作。 为了找到回答研究问题的相关论文，早期阶段收集到的研究需要符合包含和排除标准。 在这一点上，必须确保选择了一些重要但可管理的研究。 进行的搜索不限于特定时期，还考虑了早期出版物，以避免忽视任何重要的研究。 包含标准如下：

- 文章是用英语写的。
- 文章是一篇完整的研究论文（即不是演示或海报的补充）。
- 本文应明确表明人工智能是其主要重点，或将人工智能作为方法论的重要组成部分。 例如，明确将机器学习作为方法论/研究的核心组成部分的出版物。
- 本研究提出的一个或多个研究问题在本文中得到了直接回答。
- 对于出现在多个期刊或会议中的研究，选择最近的版本。

以下出版物被排除在进一步审查之外：

表2
搜索字符串。

人工智能关键词	网络安全关键词
(("推理" OR "优化" OR "机器学习" OR "人工智能" OR "自然语言处理" OR "文本挖掘" OR "分类" OR "特征提取" OR "数据挖掘" OR "情感分析" OR "计算机视觉" OR "识别" OR "遗传" OR "过滤" OR "GAN" OR "深度学习" OR "强化学习" OR "数据驱动" OR "主题建模")	("网络安全") AND (("资产管理" OR "清单" OR"配置" OR "安全控制验证" OR "评估" OR "资产"OR "安全控制测试" OR "安全姿态" OR "业务影响" OR "治理" OR "风险管理" OR"团队" OR "风险指标" OR "风险评估" OR "自动化漏洞" OR "漏洞" OR"模糊测试" OR "渗透测试" OR"漏洞严重性" OR "漏洞管理" OR "威胁猎杀" OR"自动化渗透" OR "攻击图" OR ("风险" AND "投资") OR "风险量化" OR ("风险" AND"供应链") OR "角色挖掘" OR "角色维护" OR "多因素身份验证" OR "身份验证" OR“身份” OR ("上下文相关的" AND"身份验证") OR "访问控制" OR"未经授权的访问" OR “VPN” OR("基于属性的访问") OR "基于角色的访问" OR "分离" OR "隔离"OR "隔离" OR "网络分割"OR "数据丢失" OR "数据泄露" OR "SQL注入攻击" OR "高级持续性威胁" OR "电子邮件" OR"恶意域名" OR ("完整性" AND"文件") OR ("完整性" AND ("监控"OR "审计")) OR ("自动化" AND"配置") OR "虚假新闻" OR"备份" OR (("备份") AND (“数据” OR “代码")) OR "计划" OR (("业务连续性" OR "灾难恢复" OR"事件响应") AND ("自动化"))OR "风险评分" OR "风险优先级"OR "漏洞利用" OR(("风险") AND ("修复")) OR (("日志"OR "审计") AND ("分析")) OR "安全信息与事件管理"OR "虚拟专用网络" OR "防火墙" OR "入侵防御系统" OR"防病毒软件" OR "反恶意软件" OR"免疫系统" OR (((“异常") OR("事件") OR ("入侵") OR ("欺诈"))AND ("检测") OR (("事件关联") OR ("安全智能")OR ("事件分析") OR ("关联")OR ("安全信息与事件管理") OR ("安全运营中心") OR("监控") OR ("基于行为的") OR ("社交网络") OR ("威胁情报") OR ("暗网") OR ("聊天噪音") OR ("翻译") OR ("主题建模") OR ("情感分析") OR ("网络陷阱") OR ("威胁情报 ") OR ("暗网") OR ("深网") OR ("社交网络") OR ("情感") OR ("蜜罐")) OR(("事件") AND (("检测") OR("响应") OR ("手册") OR ("基于案例") OR ("案例") OR ("识别")OR ("评估") OR ("分类")OR ("分类") OR("分类"))) OR (("网络安全")AND ("分流")) OR (("取证") AND(("智能") OR ("事件")) OR"隔离" OR "修复" OR "风险量化" OR "推荐系统" OR (("事件") AND(("分析") OR ("报告") OR("文档") OR ("信息"))) OR(("恢复") AND (("规划") OR("动态")) OR ("安全")) OR "恢复")

- 非英文写作的研究；
- 提供对不同网络安全领域中人工智能的综述或调查的研究；
- 同一作者在不同会议或期刊中发表的代表相同工作的文章也被过滤掉以去除重复；
- 提供对不同人工智能模型或现有网络安全技术进行比较分析的文章；
- 提高人工智能技术安全性以使其具有抵抗攻击能力的文章；
- 仅提供网络安全建议、指南或原则的论文（非科学性的）；
- 社论、书籍、章节和研讨会摘要；
- 未提供足够信息的研究；
- 少于5页的研究；
- 无法找到完整文本的研究。

3.4. 主要研究的选择

图2详细展示了研究的选择过程。 在确定和应用搜索词的初始步骤之后，从Scopus数据库中检索到的2395项研究经过了包含和排除标准的应用，以进一步细化。 基于排除非英文论文、海报、评论、调查、非科学出版物、社论、书籍、章节、研讨会和摘要、重复、指南文件和比较研究，共删除了366篇文章，剩下2029篇。 这些2029项研究基于标题和摘要进行了分析。 标题和摘要清楚地表明了研究是否超出了综述的重点，因此可以被排除。 如果标题或摘要没有清楚地指示研究的应用领域或贡献，它将被纳入综述的后续步骤中，其中将检查文章的全文。 基于标题和摘要分析，将这2029项研究进一步缩小到638项。 在对全文进行彻底检查后，又排除了402项研究。 因此，共有236项主要研究作为本次系统文献综述的基础。下一节将介绍这236项主要研究的发现和分析。

4. 数据提取

在选择了主要研究后，开始进行数据提取以供最新技术和描述性分析阶段使用。 数据提取的主要目标是将每个研究分解为其组成部分，并描述其整体关系和连接。 数据提取参数（在表3中解释）从选择的SLR主要研究中收集定性和上下文数据。定性数据被收集用于撰写每个主要研究的简要摘要，以展示其贡献以及人口统计信息。 上下文数据包括有关网络安全功能、解决方案类别、使用案例和核心人工智能领域的详细信息，以便对现有文献有清晰的理解。 进一步检查这些定性和上下文数据，以确定不同研究之间的关系。

5. 最新技术

为了确定评估人工智能在网络安全中应用的研究，提出了一个分类法来对应答第一个和第二个研究问题（RQ1和RQ2）的研究进行分类。分类法的前两个层次采用了NIST网络安全框架。 第一层将网络安全文献分为五个核心功能：识别、保护、检测、响应和恢复。 这五个网络安全功能涵盖了从防止安全攻击到主动寻找新威胁和反击的更复杂机制的人工智能任务的使用。 这些功能应对网络安全攻击生命周期的不同方面，以实现有效的防御。 分类法的第二层使用了NIST框架的类别来

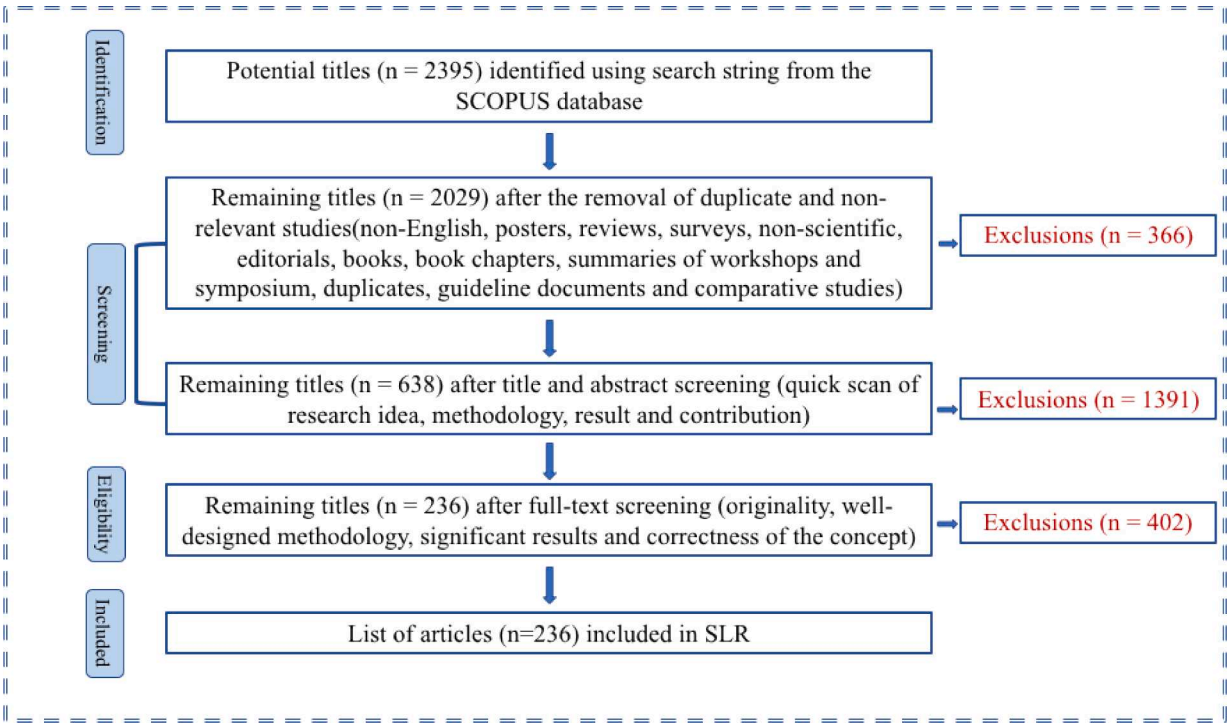


图2.SLR协议的每个阶段的选择过程和研究数量。

表3
从每个主要研究中提取的数据。

数据类型	数据项	描述
定性数据	标题	主要研究的标题
	作者	研究的作者
	发表年份	研究的发表年份
	文章类型	发表类型，例如会议、期刊
	来源	发表了研究的期刊/会议名称地理区域
		主要研究的作者所在的地理区域摘要
		论文的摘要，包括主要贡献。
上下文数据	网络安全功能	主要研究中的网络安全活动类型。
		NIST分类法将网络安全活动定义为5个功能：识别、保护、检测、响应、恢复。解决方案类别
		确定主要解决方案类别的标识。
		NIST分类法将每个网络安全功能细分为一组网络安全解决方案类别，例如，检测功能分为3个类别：异常和事件、安全持续监控和检测过程。
	具体应用案例	主要研究人工智能在网络安全中的特定应用案例，以匹配功能和解决方案类别。
	核心人工智能领域	主要研究所使用的人工智能技术的核心领域，根据AI Watch [7] 的定义。

将核心功能扩展为不同的网络安全解决方案，与紧密相关的编程需求和特定活动。 分类法的最后一级呈现了与上一级分类法相关的人工智能应用案例，并将系统化文献综述与每个确定的应用案例联系起来。图3总结了所提出的分类法，并详细描述了网络安全功能的逻辑进展。

使用人工智能技术实施的不同类别的网络安全解决方案。

5.1. 鉴定

鉴定功能通过确定与系统、人员、资产和数据相关的关键功能和风险，为其他网络安全功能提供基础。 这有助于了解当前网络安全的状况，确定差距，并制定适当的风险管理策略，以根据组织自身的需求、风险和预算实现所需的安全性。 表4总结了鉴定功能中每个主要研究的主要贡献。 以下详细介绍了该功能中的各种网络安全解决方案类别。

5.1.1. 资产管理

资产管理是识别和跟踪帮助组织实现其目标并与资产的相对重要性和风险策略成比例的信息、人员、设备、系统和建筑物的过程。 它包括发现、清点、管理和跟踪资产以保护它们。 随着组织拥有比以往更多的平台，网络安全资产管理变得越来越复杂：从运营技术系统和物联网到本地和基于云的服务。 新资产类型的大量增加和远程工作的能力导致了高度分布的资产，这些资产难以管理和清点。

基于人工智能的资产管理系统可以通过为人工团队提供新的智能水平来解决许多这些挑战。

5.1.1.1. 资产清单管理。资产发现和管理对于确保对扩展网络中的所有资产具有完全可见性和控制至关重要。 人工智能可以帮助持续和自动地发现所有设备、应用程序和用户，以及对它们进行分类和关键性评估以进行运营。 通过准确和及时的清单，可以对资产进行跟踪和分析以进行风险评估。

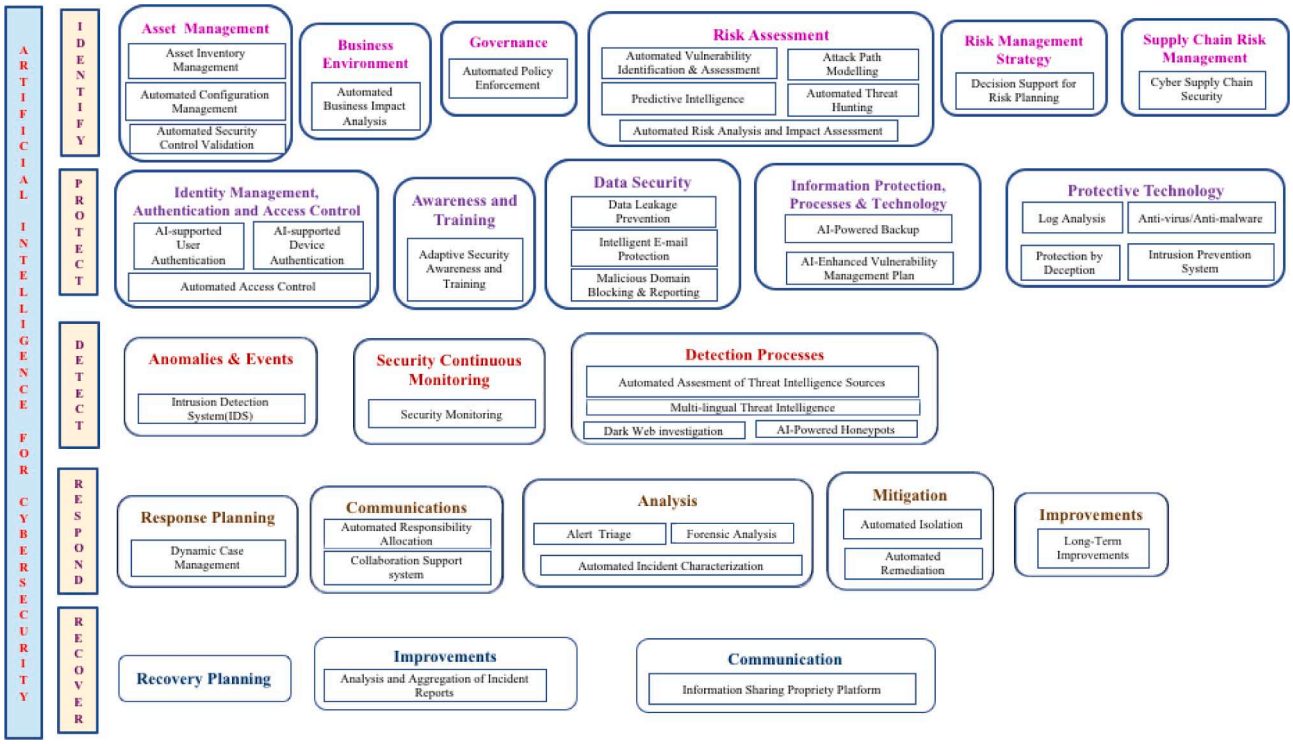


图3。 提出的网络安全领域AI技术分类法。

针对已知攻击向量的防护措施，合规性监控可以检测到恶意资产和未经授权的使用。

研究人员使用机器学习算法开发了不同的资产分类方法。Promyslov等人 [11] 使用k-means聚类将资产根据其在核电厂中的安全性、功能性和完整性进行分类。Millar等人 [12] 提出了一种基于随机森林的机器学习分类器，用于操作系统分类和网络上易受攻击设备的识别。 几项研究[13–15] 侧重于基于网络流量特征识别和分类物联网设备。Aksoy等人 [13] 和Sivanathan等人 [14] 分别使用多个和多阶段的机器学习方法进行单设备识别和分类，仅适用于小型物联网网络。Cviti´c等人 [15] 提出了一种解决在快速演变、异构和动态环境中的分类问题的方法，该方法使用监督式机器学习方法，能够根据流量流的值将物联网设备分配到预定义的类别中。

研究人员还在努力识别和阻止
受恶意软件感染的资产[16]，确定资产的重要性[17]，
以及评估个别资产的风险[18]，以管理和确保
它们的安全。

5.1.1.2. 自动配置管理。配置管理是一种治理过程，用于定义和维护系统的期望状态，并及时提供任何配置错误的警报。

自动配置管理系统将始终按照规定定义系统设置并进行系统维护，只允许在受控和授权的环境中进行更改。

为了减少由于手动或次优配置设置而导致的人为错误，定制系统的配置以确保所需的性能和安全水平非常重要。 研究人员[19,20]正在基于系统特性和操作环境，使用多目标强化学习和遗传算法，分别为在线文件共享系统和分布式云存储系统开发动态配置系统。Sharifli等人[21]而Bringhenti等人[22]提出了一个完全自动化的框架，通过观察用户的行为并通过以人类友好的语言表达的高级安全要求的细化来定制安全控制。

自动化配置评估使合规团队能够持续审查和测试配置，及时识别脆弱的配置，以减少或避免网络安全事件。 Varela-vaca [23,24]提出了一种基于软件产品线技术的方法，用于自动分析系统的易受攻击配置。另一方面，Liu等人[25]利用随机森林模型预测基于DNS和BGP协议的配置错误以及网络中可见的恶意活动引发的网络安全事件。

5.1.1.3. 自动化安全控制验证。 安全控制验证的自动化将在不断变化的环境和威胁背景下提供实时的安全监控。. 研究人员正在利用网络望远镜数据[26]、构建网络安全框架[27]或通过相关威胁、漏洞和安全措施[28]来实现对系统整体安全性的明确评估。

5.1.2. 商业环境

商业环境类别被定义为识别在不利情况下确保业务连续性的关键流程和应用程序。 这些信息对于业务的可持续性至关重要，并成为制定有效的响应和恢复策略的基础。 AI技术可以通过以下用例来自动化这个过程。

5.1.2.1. 自动化业务影响分析。业务影响分析是通过评估网络安全事件对业务的影响来确定业务环境中的关键功能和应用程序的最重要技术。AI技术可以通过评估经济风险来自动化业务影响分析。

表4
关注识别功能的主要研究总结。

解决方案类别	用例	贡献	AI领域	作者
资产管理	资产清单管理	资产分类	学习	Promyslov等人[11]，Millar等人[12]
		资产分类	规划 & 学习	Aksoy等人[13]
		资产分类	通信 & 学习	Sivanathan等人[14]
		资产识别 &监测	学习	Cviti ˙c等人[15]
		恶意软件感染资产的识别	学习	Cam [16]
		资产重要性 & 风险预测	推理 & 学习	Kure等人[17]
		威胁 & 风险评估	推理	Vega-Barbas等人[18]
	自动化配置管理	动态配置系统	学习	Tozer等人[19]
		动态配置系统	规划	García-Hernández等人[20]
		配置的定制化	推理	Sharifli等人[21]
		定制化的最优分配和网络安全配置	规划	Bringhenti等人[22]
		自动化配置评估	推理	Varela-Vaca等人[23,24]
		自动配置评估	学习	Liu等人[25]
	自动安全控制验证	网络安全态势的特征化	学习	Zhan等人[26]
		自我评估以确定网络安全态势	推理	Gourisetti等人[27]
		安全分析师工作的自动化	规划	Stepanov等人[28]
商业环境	业务影响分析	经济风险的建模和测量	学习	Narasimhan [29]
		安全事件发生的概率估计	规划	Nguyen & Nicol [30]
		对业务资产进行攻击和影响评估的可行性	推理	Ponsard等人[31]
治理	自动策略执行	对网络流量进行自动策略执行	规划	Odegbile等人[32]
风险评估	自动漏洞识别和评估	漏洞检测	通信 & 学习	Nembhard等人[33]
		漏洞检测	通信	Liu等人[34]，Jeon & Kim [35]
		漏洞跟踪	通信 & 学习	Huflfl等人[36]
		漏洞跟踪	通信	Iorga等人[37]
		漏洞识别	学习	Saha等人[38]
		基于人工智能的模糊测试	推理	Wang等人[39]
		基于人工智能的模糊测试	学习	Wang等人[40]，Godeflroid等人[41]，Cummins等人[42]，Xu等人[43]，She等人[45]，Liu等人[46]
		基于人工智能的模糊测试	规划 & 学习	Chen等人[44]
		自动化渗透测试	学习	Zhou等人[47]，Gangupantulu等人[48]，Neal等人[49]
		漏洞分类	通信	Russo等人[50]，Aota等人[51]，Vanamala等人[52]
		漏洞探索	推理	Bakirtzis等人[53]
		漏洞探索	通信	Kuppa等人[54]
		漏洞探索	学习	Chatterjee & Thekdi [55]
		漏洞评估 & 修复	通信	Jiang & Atifl [56]，Samtani等人[57]
		漏洞评估 & 修复	规划	Brown等人[58]
	自动化威胁猎杀	开源网络威胁情报(OSCTI)	通信	Gao等人[59]
	攻击路径建模	使用入侵警报进行路径建模	学习	Nadeem等人[60]
		使用漏洞描述进行路径建模	通信 & 学习	Binyamini等人[61]
		使用漏洞描述进行路径建模	规划	Falco等人[62]
		模拟攻击者/防御者活动和预防措施	学习	Cam [63]
		模拟攻击者/防御者活动和预防措施	规划	Wollaber等人[64]
	自动化风险分析和影响评估	自动计算风险评分	学习	Sancho等人[65]
		自动计算风险评分	推理	Tubis等人[66]
		推断安全事件的概率	学习	秦等人[67]
		关键漏洞风险指标的识别	学习	法尔科等人[68]
		关键漏洞风险指标的识别	推理	维加-巴尔巴斯等人[69]
		自动化风险评估和决策分析	学习	卡林宁等人[70]，阿尔-哈德拉米等人[72]

（续下页）

表4 (续)

解决方案类别	用例	贡献	AI领域	作者
		自动化风险评估和决策分析	通信	Biswas等人[71]
	预测性智能	入侵警报预测	学习	Ansari等人[73]，Wang &Jones [74]，Najada等人[75]，Mueller等人[76]恶意软件预测
			学习	Rhode等人[77]
		攻击预测	通信	Perera等人[78]
		攻击预测	推理	Marin等人[79]
		攻击预测	学习	Polatidis等人[80]
风险管理策略	风险规划的决策支持	风险规划的决策支持系统（DSS）	规划	Rees等人[81]，Paul &Wang [82]，Paul & Zhang [83]用于风险规划的攻击图建模
			规划	Zheng等人[84]
供应链风险管理	基于人工智能的供应链安全	威胁分析 & 预测	学习	Yeboah-Oflori等人[85]
		最佳网络安全投资	规划	Sawik [86,87]，Sawik &Sawik [88]
		网络韧性评估	推理	Rahman等人[89]

基于已知攻击向量或计算威胁可行性，以及对关键业务领域上高影响安全事件的概率进行评估。 研究人员使用不同已知攻击配置文件的建模[29]、罕见事件模拟[30]或将业务目标与攻击者的能力联系起来，以指导情景分析[31]，以确定其对业务资产的影响，从而测量不同企业中网络安全的经济风险。

5.1.3. 治理

治理涉及组织了解环境和运营要求、监控法规要求的政策、程序和流程。 这有助于确定组织的责任，并向管理层提供有关网络风险的信息。 人工智能可用于政策执行或自动检索关键风险指标。 虽然有关政策执行的一些研究，但在本研究中没有找到关于实时测量风险指标的研究文章。 因此，开发一个早期警示系统以指示由于政策违规、红旗或其他症状而随时间发展的风险，是一个诱人的未来研究方向。 自动检索关键风险指标，如故障间隔时间、未打补丁系统的存在、风险偏好或尝试入侵的次数，并将其转化为知识，将有助于通过快速修复风险来防止网络安全漏洞。

5.1.3.1. 自动策略执行。策略执行对于组织来说至关重要，以确保其符合法规和适当的风险管理。 *AI*正在通过使用控制器和策略代理在传统非SDN网络中进行自动策略执行[*32*]。 控制器是用于管理传统路由器的软件定义中间盒的集中式管理服务器，策略代理将识别受策略约束的流量并协助其执行策略。

5.1.4. 风险评估

风险评估是识别、估计和优先考虑与当前或近期操作、操作资产和个人相关的网络安全风险的过程。 它需要对威胁、漏洞和攻击信息进行仔细分析，以确定网络安全事件可能对组织产生不利影响的程度，以及此类事件发生的可能性。 由于风险因素众多，手动风险评估过程复杂、昂贵且耗时，因此需要在每个阶段积极参与人员。 基于AI的风险评估过程通过支持风险管理团队在以下用例中解决这些挑战。

5.1.4.1. 自动化漏洞识别和评估。 自动化漏洞评估是使用自动化工具系统地审查系统中的安全弱点的过程，以进行漏洞识别、分类、探索和优先级排序。这些自动化工具依赖于漏洞库、供应商漏洞公告、资产管理系统和威胁情报源来识别、分类和评估严重性，并提出修复建议。

- 自动化漏洞检测：这是识别组织的应用程序、服务器或其他系统和资产的漏洞的重要步骤。 研究人员通过检查源代码使用深度学习和迁移学习[33-35]来进行软件漏洞检测。 这些研究采用文本挖掘技术，将基于机器学习的漏洞检测模型与推荐系统结合起来，帮助程序员编写安全代码。

研究人员还通过漏洞库或社交网络等方式，检测软件[36]和网络基础设施[37]的新出现的漏洞。 Saha等人[38]提出了一种新的方案，通过对系统和网络层次下受到攻击的物联网/物理系统（CPS/IoT）的行为进行建模，然后使用机器学习来发现任何潜在的攻击空间。

研究人员使用基于人工智能的模糊测试来发现软件和硬件接口以及应用程序中的漏洞。 这是通过向程序或接口中注入错误、意外或随机生成的数据，然后监视崩溃、代码断言失败、未记录的跳转或调试例程以及可能的内存泄漏等事件来完成的。 如图4所示，利用人工智能技术开发了一个自动化系统，用于识别潜在的攻击选项、生成输入、生成可能的测试用例并分析崩溃情况。 研究人员[39,40]使用推理和自然语言处理技术进行种子生成，以增加代码覆盖率，从而为智能模糊测试系统的基本步骤之一提供更多独特的执行路径。 测试用例生成是基于人工智能的模糊测试在Web浏览器[41]、编译器[42,43]、网络物理系统（CPS）[44]、软件库[45]和简单计算机程序[46]中广泛研究的领域之一。

自动化渗透测试是一种高效智能的尝试，通过利用已知或零日漏洞来渗透攻击面，以确定攻击者可以从当前环境中获得什么。 研究人员正在研究使用强化学习进行大型网络[47,48]和微电网控制算法[49]的自主渗透测试。

- 自动化漏洞分类：漏洞分类是一个重要的步骤，可以加快对安全相关信息的深入理解，以加速漏洞的修复。

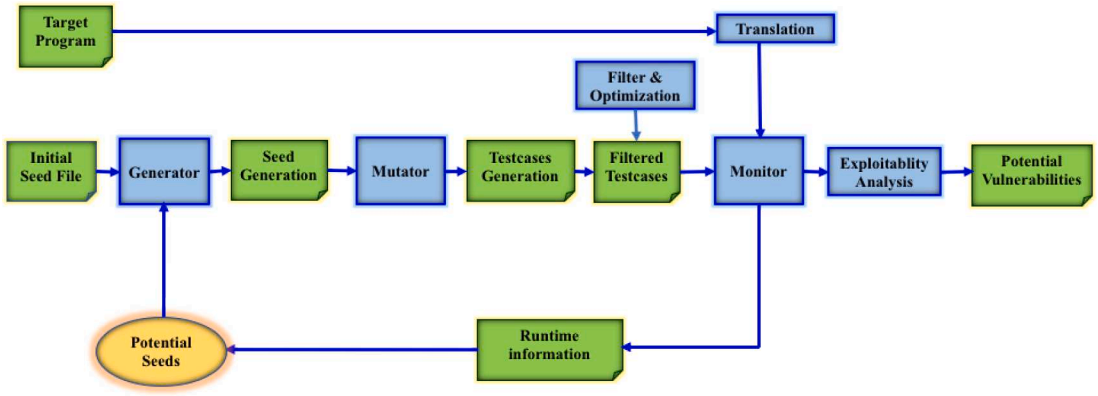


图4. 漏洞检测的智能模糊处理过程。

评估。 研究人员正在开发自动漏洞分类系统，并为漏洞报告中的漏洞描述进行标记。Russo等人[50]提出了一种方法，根据行业定义的分类模型，对每日发布的漏洞进行总结和分类。 Aota等人[51]使用文本挖掘来对常见漏洞和暴露（CVE）列表中提供的描述进行分类。Vanamala等人[52]还将CVE条目分类到Open Web应用安全项目（OWASP）的前10个风险中。

- 漏洞探索：漏洞评估中的一个重要步骤是识别可能利用漏洞的潜在攻击向量，以有效评估和管理它们。 为了实现这个目的，一些研究人员使用MITRE公司的数据集[53,54]，而其他人使用由事件响应和安全团队论坛提供的CVSS[55]。Bakirtzis等人[53]和Kuppa等人[54]使用基于模型的过程将对抗策略、技术和常识自动映射到给定的系统模型。 相比之下，Chatterjee和Thekdi使用概率模型来快速适应网络和攻击特征的变化，评估和管理系统漏洞[55]。

- 漏洞评估和优先级排序：这一步的主要目标是对漏洞进行优先级排序，并根据其严重性和系统的漏洞暴露情况提供评估报告。 AI技术被用于根据系统、数据和业务风险以及攻击的难易程度、严重性和潜在损害来为每个漏洞分配严重性评分。

Jiang和Atifl [56]通过基于漏洞严重性和威胁配置文件指标的机器学习流程，对冲突漏洞报告中的漏洞严重性进行了自动评估和协调。 Samtani等人从Shodan数据集中评估了SCADA设备的漏洞，并将其分为四个主要风险级别：关键、高、中和低[57]。Brown等人计算了攻击图中每个物联网设备的漏洞和利用风险分数，该攻击图是根据网络管理员指定的网络拓扑创建的[58]。

5.1.4.2. 自动化威胁搜索。 自动化威胁搜索是一种主动的安全搜索方法，用于在组织内的网络、终端和数据集中检测潜在的恶意、可疑或风险活动。. 它利用已经收集到的数据上的最新威胁情报，提前识别和分类潜在威胁。 威胁搜索是一个相对较新的应用领域，对于早期检测非常重要。 然而，现有方法仍然基于异常检测进行威胁检测，并忽视了开源网络威胁情报（OSCTI）提供的丰富外部知识。

5.1.4.3. 攻击路径建模。攻击路径建模是一种积极的风险减少方法，通过映射网络中的易受攻击路径来评估风险，识别漏洞，并采取对策来保护关键资产。 研究人员使用入侵警报[60]或漏洞描述[61,62]等AI技术进行路径建模。 一些研究人员使用所有的网络数据，包括警报、漏洞、日志和网络流量，以模拟攻击者/防御者的活动，并实时采取预防措施[63,64]。

5.1.4.4. 自动化风险分析和影响评估。自动化风险分析和影响评估通过智能地利用内部和外部可用的风险数据，实时评估风险和相关指标，加强风险管理团队。 人工智能是加速风险管理进展的催化剂，通过自动计算风险评分[65,66]、推断安全事件发生的概率[67]、识别关键漏洞风险指标[68,69]以及使用日志数据和威胁情报进行风险评估和决策分析[70–72]，在组织内外部发挥作用。

5.1.4.5. 预测智能。 预测智能是在特定环境中可操作和相关的智能，可以用来预测攻击。. 入侵预测工具有助于提前预测入侵的类型、强度和目标，从而对未来的攻击提供主动防御。 研究人员正在使用深度学习[73–76]方法来预测恶意来源[73]或给定目标[74–76]的警报，使用先前警报[73]、历史垃圾邮件[74]和网络流量[75,76]数据的序列。

恶意软件预测涉及预测和阻止恶意文件在完全执行其有效负载之前，以防止恶意软件攻击，而不是事后补救。 在这方面，Rhode等人开发了一个基于循环神经网络（RNNs）模型的恶意软件预测模型，使用机器活动数据[77]来预测恶意行为。

攻击预测被认为在积极推动网络安全性方面具有巨大潜力。 研究人员通过利用从新闻网站和网站[78]、暗网论坛[79]、国家漏洞数据库[79]、事件报告[79]和常见漏洞和曝光数据库[80]中检索到的不同类型的数据，提出了攻击预测方案。

5.1.5. 风险管理策略

风险管理策略通过确定优先级、风险容忍度和限制来辅助操作风险决策。 它需要确保建立和记录可接受的风险水平，以及合理的解决时间和投资。 人工智能有潜力通过自动化以下活动来改变这个领域。

5.1.5.1. 风险规划的决策支持。网络安全风险规划涉及在预定预算内实施所需的一系列对策。 形式化的决策支持系统[81–83]和攻击图模型[84]可以帮助安全规划者与对策成本和可用风险预算进行经济比较。

网络安全风险规划中的决策问题很重要，因为风险计划对决策者对风险的态度及其与可用预算的相互作用非常敏感。 因此，为了估计组织在网络攻击下面临的不确定风险，考虑到不确定的威胁率、对策成本和对资产的影响，实施决策支持系统非常重要。 在这方面，Rees等人[81]使用遗传算法找到了最佳的对策组合来阻止或减轻安全攻击，允许用户确定投资成本和产生风险之间的首选权衡。 Paul &Wang [82]和Paul & Zhang [83]使用鲁棒优化来研究预防、检测和遏制网络安全不确定性之间的最佳平衡。

Zheng等人[84]使用攻击图模型来识别一组安全控制措施以降低风险。 他们的模型用于选择最佳的控制措施集，以确保总成本不超过组织预算。

5.1.6. 供应链风险管理

供应链风险管理支持特定于识别、评估和管理供应链风险的风险决策。 在供应链中有效管理网络安全风险需要全面了解威胁和漏洞、成本效益的供应链风险规划策略以及对供应链的网络安全弹性进行评估。 研究人员正在积极使用人工智能技术自动化威胁分析和预测[85]、优化网络安全投资[86–88]以及评估供应链的网络安全弹性[89]。

5.1.6.1. 网络供应链安全。 网络供应链安全需要在进出链的子系统之间建立安全的集成网络。. 因此，了解和预测威胁对于使用内部和威胁情报资源来限制业务中断至关重要。*Yeboah-oflori*等人整合了网络威胁情报数据，并使用机器学习技术预测了网络供应链系统上的网络攻击模式[85]。

优化网络安全投资是工业4.0供应链网络安全中的一个重要领域，旨在快速检测、减轻和平衡安全漏洞对可用预算的影响。 在这个方向上，Sawik [86–88] 提出了不同的模型，以确定在有限预算和安全控制组合的情况下的最佳网络安全投资，以平衡供应链中的网络安全。

评估供应链的网络弹性是确保供应链免受网络入侵并获得竞争性业务优势的关键任务。2021年，Rahman等人 [89] 提出了一种基于Dempster-Shafler (D-S)理论的综合方法，用于构建评估增材制造供应链网络弹性的框架。

5.2. 保护

保护功能帮助规划和实施适当的控制措施，以限制或遏制潜在网络安全事件的影响。 这包括一系列技术和程序控制措施，以主动防范内部和外部网络威胁。 人工智能可以通过对用户、设备和其他资产进行身份验证、监控用户行为、自动访问控制、自适应训练、数据泄漏预防和完整性监控、自动信息保护以及流程和提供保护来提高系统的弹性。

主动保护系统的解决方案。 表5提供了每个主要研究关注保护功能的摘要。 下面介绍了解决方案类别以及每个类别中人工智能用例的详细概述。

5.2.1. 身份管理、认证和访问控制身份管理、认证和访问控制负责将对资产和相关设施的访问限制在授权用户、进程或设备以及授权活动范围内。 人工智能可以用于使用智能用户认证、智能设备认证、使用授权进行自动访问控制以及访问权限来管理和保护物理和远程访问，以防止未经授权的访问及其后果。

5.2.1.1. 支持人工智能的用户认证。人工智能可以通过使用生物特征识别[90]、行为生物特征识别[91–94]或多因素认证[95,96]来改进用户认证，而不是使用容易被破解的用户名、密码甚至一次性文本令牌。

物理生物特征是指用户固有的身体特征，如指纹、虹膜和生物信号，可用于身份识别。2020年，Siam等人提出了一种基于深度学习的PPG（光电脉搏图）生物特征人体认证系统[90]。

相比之下，行为生物特征与人类活动中独特可识别和可测量的模式有关，可以提供连续和用户友好的安全性。 行为生物特征包括使用行为[91–93]和步态[94]等多种形式。 连续认证系统的主要基础是与用户自己设备的交互相关的行为模式。 在这个方向上，Valero等人建议使用移动功能和使用数据，例如加速度计、陀螺仪、磁力计以及与不同应用程序的交互统计数据，来确定当前用户是否与先前认证的用户相同[91]。2019年，Sanchez等人设计并实现了一种基于用户与不同办公设备的交互行为模式的连续认证机制，该机制适用于智能办公室，采用了云计算模式和随机森林（RF）算法[92]。 此外，这些解决方案逐渐被应用于联合身份管理解决方案，进一步增加了对它们的兴趣[93]。

步态认证是一种非侵入式、透明和连续的移动设备认证方法，通过捕获用户行走时所需的信息来验证用户的真实性。 2022年，Alobaidi等人[94]通过使用每个步态周期的时域和频域特征提取，研究了步态认证在不受控制的真实世界中的可行性。

多因素认证是一种分层方法，用于保护需要两个或更多凭据来验证用户身份或登录的数据和应用程序。 研究人员[95,96]将击键动态作为网络用户认证和设备认证的第二层安全措施。

5.2.1.2. 基于人工智能的设备认证。 智能设备认证是基于设备在网络中的凭据或行为进行认证的过程，以确保机器间通信的安全性。. 研究人员正在积极研究传感器识别和认证领域，以确保网络物理系统或汽车行业的安全性。 通道[97]和传感器[98,99]的缺陷被用于找到瞬态和稳态参数，作为机器学习模型的输入，用于传感器识别。

5.2.1.3. 自动访问控制。自动访问控制根据组织内的情况或角色和规定，限制系统访问仅限于授权用户。 研究人员积极使用人工智能技术来维护访问控制状态[100]，角色挖掘[101]和情境感知决策[102,103]，以预防

表5
关于保护功能的主要研究总结

解决方案类别	用例	贡献	AI领域	作者
身份管理，认证和访问控制	AI支持的用户认证	基于物理生物特征的认证	学习	Siam等人[90]
		基于行为生物特征的认证	学习	Valero等人[91]，Sanchez等人[92]，Martin等人[93]基于行为生物特
		征的认证	感知	Alobaidi等人[94]
		多因素认证	推理	Rahman等人[95]，Shaot &Schmidt [96]A
	I支持的设备认证	传感器识别 & 认证	学习	Hafleez等人[97]，Baldini等人[98]
		分布式同步相量测量装置的源认证	学习	Cui等人[99]
	自动化访问控制	基于角色的访问控制	规划	Benedetti和Mori [100]，Abolflathi等人[101]基于属性的访问控制
			规划	Chukkapalli等人[102]
		基于属性的访问控制	推理	Leander等人[103]
意识和培训	自适应安全意识和培训	自适应网络安全培训	通信	Tan等人[104]
		安全编码推荐系统	通信	Nembhard等人[105]
		安全编码意识	学习	Gasiba等人[106]
数据安全	数据泄漏预防	监控数据访问、数据移动和用户活动	学习	Le和Zincir-Heywood [107]，Kim等人[108]，Al-Shehari等人[109]
		自动数据敏感性检测	通信	Alzhrani等人[110]，Guo等人[111]
		高级持续性威胁检测	学习	Li等人[112]，Alghamdi & Reger [113]
	智能电子邮件保护	恶意垃圾邮件检测	学习	Gallo等人[114]
		恶意垃圾邮件检测	通信	Wu等人[115]。
		钓鱼邮件检测	通信	Gualberto等人[116]，Nguyen等人[117]
	恶意域名阻止 & 报告	恶意网站检测的网站设计特征	学习	Cohen等人[118]
		基于域名的恶意网站检测特征	学习	Marques等人[119]，Yu等人[120]，Spaulding &Mohaisen [121]基
		于URL的恶意网站检测特征	学习	Indrasiri等人[122]，Vinayakumar等人[123]混合特征用于恶意网站检测
			学习	Li等人[124]，Alotaibi [125]
信息保护流程 & 程序	基于AI的备份	动态备份调度	推理	Qin等人[126]
		智能备份调度	学习	Van de Ven等人[127]
	增强型AI漏洞管理计划	基于上下文的漏洞风险评分	推理 & 学习	Zeng等人[128]
		漏洞利用趋势	学习	Yin等人[129]
		漏洞利用趋势	通信	Yin等人[130]
防护技术	日志分析	证据提取	学习	Bai等人[131]
		数据展示技术	通信	Aflzaliseresht等人[132]
		处理多样性和互操作性问题	通信	Torre-Abaitua等人[133]，Eljasik-Swoboda和Demuth [134]异构日志数据的自动
		化安全分析	学习	Sisiaridis和Markowitch [135]
	IPS	用于电子控制单元的IPS	学习	De Araujo-Filho等人[136]
		用于物联网网络的IPS	学习	Constantinides等人[137]
	反病毒/反恶意软件	恶意软件的作业方式分析	学习	De Lima等人[138]
		动态数据分析	学习	Marques等人[139]
	通过欺骗进行保护	诱饵文本生成	规划	Karuna等人[140]

未经授权的访问及其后果。

基于角色的访问控制（RBAC）根据用户在组织中的角色授予不同的访问权限。 Benedetti和Mori提出了使用人工智能技术来更新和维护访问控制状态的方法，当出现异常或违规时报告[100]。 他们主要致力于提供一个优化的行动计划，以重新配置RBAC状态，以便促进维护过程。Abolflathi等人提出了一种可扩展和最优的角色挖掘方法，从现有的访问控制列表中提取用户角色和角色权限关系[101]。

基于属性的访问控制考虑了与用户、环境和被访问资源相关的各种预配置属性特征。Chukkapalli等人[102]和Leander等人[103]分别在智能渔业和智能制造系统中测试了基于属性的访问控制的情境感知决策性能。

5.2.2. 意识和培训

这个解决方案类别涉及网络安全意识和通过自然语言处理算法[104,105]自动选择内容，或者通过提供机器学习启用的智能教练[106]来进行自适应和个性化的网络安全培训、意识或建议，可以为人员和合作伙伴提供培训，以便他们按照政策和程序履行他们的信息安全职责和责任。

5.2.2.1. 自适应安全意识 &培训。 自适应培训和意识对于克服过时的培训内容、培训材料的选择以及可接受的培训方法的选择是重要的。T*an*等人[*104*]创建了一个自适应基于DB*pedia*的网络学习系统，该系统从中获取最新的培训内容并根据学习者对信息安全的先前知识进行自动内容选择。. 另一方面，N*embh**ard*等人[*105*]和G*asiba*等人[*106*]通过主题建模或使用严肃游戏技术来帮助程序员推荐或提高对安全编码实践的意识。

5.2.3. 数据安全

数据安全根据风险策略来管理信息管理，以保护敏感信息。 这包括保护数据在静态和传输中的安全，以及管理持有信息的资产的生命周期，包括其废弃或处置。研究人员正在积极使用人工智能技术进行数据泄漏预防，智能电子邮件保护，恶意域名阻止和报告，以及基于代理的完整性监控，以确保数据的机密性，完整性和可用性。

5.2.3.1. 数据泄露预防。数据泄露预防涉及检测和保护数据泄露、数据外泄或数据意外破坏。 人工智能技术用于监控数据访问、数据移动和用户活动[*107*–*109*]，自动检测数据敏感性[*110*,*111*]，以及高级持续性威胁（APT）检测[*112*–*114*]以防止数据泄露。

通过观察授权个体的相关行为或活动，对其使用敏感信息的识别提供了数据泄露预防的准确洞察。 研究人员[107–109]正在积极使用人工智能技术监控用户活动，通过将来自多个来源的数据相关联，识别其异常行为，如活动激增或异常活动。 在这方面，研究人员使用由CERT提供的内部威胁测试数据集，通过不同的时间表示形式[107]或每日活动摘要、电子邮件内容和电子邮件网络[108]，为数据泄露预防提供洞察。相比之下，Al-shehari和Alsowail [109]提出了一个仅适用于在员工离职前的敏感期间识别数据泄露事件的模型。

自动化数据敏感性检测是一种通过分析、标记和组织数据，根据共享特征将数据识别和分类为相关类别（机密、私人和公共）的方法。 它可以赋予数据泄露预防技术监控用户对特定敏感数据部分的行为的能力，而不是始终跟踪所有数据。Alzhrani等人[110]提出了一种基于安全相似性的自动分类技术（ACESS），用于减轻内部人员泄露敏感数据的威胁。2021年，Guo等人[111]指出，敏感信息通常存在于非结构化数据中，使得意外数据泄露更容易发生。 因此，他们提出了一种基于内容和上下文的敏感信息识别方法，使用BiLSTM和注意机制。

高级持续性威胁（APTs）是一种持续时间长且目标网络未察觉的有针对性网络攻击。 这种攻击的主要目的是窃取数据而不是造成任何破坏。 研究人员正在努力从终端、网络和云端高效捕获遥测数据，将这些多样化的遥测数据整合和分析，以发现异常、威胁指标（IoCs）和其他感兴趣的行为[112,113]。

5.2.3.2. 智能电子邮件保护。智能电子邮件保护是一类软件解决方案，用于防止针对电子邮件的复杂网络攻击。 传统上，垃圾邮件被用作通过向大量列表发送未经请求的电子邮件来推销商品和服务的策略。

然而，如今，它被积极用于传播恶意软件、窃取身份验证凭据或进行金融欺诈。 人工智能技术正在用于自动化防护恶意垃圾邮件。

研究人员正在使用监督分类[114]和基于深度学习[115]的技术，通过分析动态传入的电子邮件数据，包括常规、查看、主题、附件和内容相关特征，实时识别垃圾邮件。 然而，Gualberto等人[116]和Nguyen等人[117]将他们的研究工作限制在通过分析电子邮件内容来检测钓鱼邮件。

5.2.3.3. 恶意域名阻止和报告。恶意域名阻止和报告为电子邮件保护提供了更高级别的安全性，通过捕捉由打开垃圾邮件或附件引起的任何恶意网络流量。 研究人员使用人工智能来识别每个DNS查询的可疑网站，并阻止访问与恶意软件、网络钓鱼、勒索软件和其他网络威胁相关的恶意网站。

恶意网站的检测是通过使用丰富的恶意和非恶意网站特征对机器学习算法进行训练来实现的。 这些特征可以分为四个主要类别：网站设计特征[118]、基于域名的特征[119–121]、基于URL的特征[122,123]和混合特征[124,125]。

2021年，Cohen等人[118]提出了一种新的网站分类方法，通过自动抓取和处理数千个视觉和非视觉设计特征来识别恶意软件或破解网站。基于域名的特征在研究领域中用于检测恶意网站非常流行。 研究人员正在使用监督式机器学习模型[119]和深度学习[120,121]技术，使用经典域名特征来检测恶意域名。 从URL字符串中提取的特征，包括语言、词汇、上下文和统计信息，用于确定恶意网站。Indrasiri等人[122]和Vinayakumar等人[123]分别使用基于URL的特征作为集成机器学习和深度学习模型的输入来进行恶意网站的检测和分析。

一些研究人员还使用混合特征来进行恶意网站识别，以了解僵尸网络检测[124]或钓鱼网站检测[125]。 这些技术使用与域名结构和DNS响应相关的特征组合，例如解析源、每日解析量等。

5.2.4. 信息保护、流程和程序该领域涉及与定义的安全策略、流程和程序一致的信息源和资产的保护。

它包括信息的保护，以及响应、恢复和漏洞管理计划的建立、管理和实施。 研究人员正在研究应用人工智能技术进行基于人工智能的备份，以及增强的漏洞管理计划，以维护信息保护的流程和程序。

5.2.4.1. 基于人工智能的备份。基于人工智能的备份解决方案正在兴起，根据优先级和要求备份关键数据和软件组件，以确保高效备份。 人工智能技术正在用于动态备份调度[*126*]和优化备份调度[*127*]。

秦等人[126]设计了一个具有智能调度算法的动态备份系统，以提高备份环境的稳定性和可预测性。 所提出的系统通过确定哪个备份先开始以及分配给该备份的存储来高效地调度备份。 相比之下，Van deVen等人[127]使用二维马尔可夫链来建模数据备份并研究备份调度的优化。在每个时间槽，所提出的技术都会检查概率备份策略来启动备份，而不考虑备份的大小。

5.2.4.2. 增强型人工智能漏洞管理计划。漏洞管理计划是一个旨在主动减少系统风险暴露的框架。

随着近年来报告的漏洞数量增加，将漏洞管理计划与系统的要求和关键成功因素相一致变得更加重要。 研究人员正在使用人工智能技术来确定基于上下文的漏洞风险评分和漏洞利用趋势，以实时保护资产和信息系统。

基于上下文的漏洞风险评分将帮助分析师在特定资产或信息系统的背景下优先考虑风险，并使他们能够采取保护措施。曾等人[128]提出了一种新的风险优先级评估方法，通过整合攻击者模型来捕捉攻击者对利用漏洞的偏好。 风险评分由利用的关键性和利用的可能性定义，使用逻辑推理引擎。

漏洞利用趋势将帮助分析师通过预测最有可能被利用的漏洞来优先修补和纠正。 研究人员正在使用新颖的人工智能技术来预测利用性[129]，并解决类别不平衡问题以提高机器学习算法的性能[130]。在这里，尹等人[129]专注于利用迁移学习来解决漏洞利用预测问题，以帮助专家优先应用补丁。尹等人[130]还提出了一种新颖的顺序批量学习技术，称为实时动态概念自适应学习，以解决利用性预测中的概念偏差和动态类别不平衡问题。

5.2.5. 保护技术

保护技术提供系统和资产的安全性和弹性。 这些技术使用特定的防篡改特征来识别和阻止对组织资产的入侵、更改、渗透和信息提取的企图。 人工智能可以用于提供以日志分析工具、入侵预防系统、防病毒/防恶意软件解决方案和欺骗保护为形式的保护解决方案。

5.2.5.1. 日志分析。日志分析是审查计算机生成的事件日志以主动识别错误、安全问题或其他风险的过程。 基于人工智能的日志分析工具可以自动化处理大量分布式日志数据的例行和重复任务。

Bai等人[131]测试了各种监督机器学习方法在检测恶意远程桌面协议（RDP）会话的证据方面的性能，使用了Windows RDP事件日志。

另一方面，Aflzaliseresht等人[132]提出了一种新的方法，使用叙事技术生成自然语言报告，根据用户的知识水平识别网络威胁信息。

研究人员还致力于解决日志管理中的多样性和互操作性问题。 De la Torre-Abaitua等人[133]和Eljasik-Swoboda和Demuth[134]通过智能方法从不同来源提取和处理文本数据，以辅助信息检索方法实现可接受的日志特征表示，解决了多样性问题。 同样，Sisiaridis和Markowitch[135]通过采用自动特征提取和特征选择技术，对来自不同网络传感器的异构日志数据进行安全分析。

5.2.5.2. 入侵预防系统。 入侵预防系统监控网络流量，然后采取适当的措施来阻止攻击，如报告、阻止、丢弃或重置连接。. 研究人员提出了无监督的孤立森林[136]和基于自组织增量神经网络和支持向量机的嵌入式系统入侵预防系统[137]，分别用于汽车电子和物联网网络。

5.2.5.3. 反病毒/反恶意软件解决方案。基于人工智能的反病毒/反恶意软件解决方案可以分析数千个文件并提取有用的特征来对其进行分类，判断其是否为良性文件或恶意软件。 研究人员已经创建了用于检测恶意软件的反病毒程序，这些程序使用从可执行文件[138]或动态数据分析[139]中提取的特征作为输入，分别使用人工神经网络（ANNs）或循环神经网络（RNN）模型进行分析。

人工神经网络（ANNs）或循环神经网络（RNN）模型。

5.2.5.4. 通过欺骗进行保护。欺骗保护是一种高级技术，用于在攻击者渗透网络后保护关键文档。 人工智能已被用于生成可信的虚假文档，以误导网络攻击。Karuna等人[140]提出了在对手已经进入系统时创建诱饵文件，将对手引开真正的目标。 他们的诱饵文本生成方法使用遗传算法来操作真实文档的可理解性，以生成难以理解但可信的虚假文档。

5.3. 检测

通过开发和实施适当的活动来及时发现网络安全事件的**detect**函数，能够识别它们的发生。 这个函数对于安全性至关重要，因为及时的检测将最大程度地减少干扰。 它包括及时检测入侵和异常的活动，以及影响评估、实施安全持续监控以验证保护措施的有效性，以及适当维护检测过程以确保对网络事件的意识。 人工智能可以通过监控内部和外部信息源，并迅速将这些信息相关联，以检测异常活动，从而最小化影响。 表6总结了每个研究对检测功能的主要贡献，以及解决方案类别、人工智能用例和使用的人工智能领域的详细信息。 下面提供了解决方案类别以及人工智能用例的详细审查。

5.3.1. 异常和事件

该领域的解决方案通过建立和管理来自多个来源的操作和数据流的基线，来解决异常活动的检测和分类问题。 然后使用这些基线来检测和分析事件，以了解攻击目标和方法。

5.3.1.1. 入侵检测系统（*IDS*）。*IDS*是一组工具和技术，用于监视系统和网络流量，分析异常和可疑活动，旨在检测可能针对系统的入侵。 在最先进的研究中，*IDS*从三个角度实现：二元分类、多类别分类或二者兼有。 二元分类假设有两个标签：正常和攻击。 另一方面，多类别分类处理将问题分类为三个或更多类别的问题。 在*IDS*的情况下，多类别分类区分不同类型的攻击，并为用户提供更多信息来应对攻击。 有效开发和评估二元和多类别入侵检测系统需要基准数据集。 因此，本综述仅包括使用基准数据集评估其方法的研究。

在这个方向上，研究人员使用了ADFA-WD和ADFA-WD: SAA数据集进行基于系统调用的入侵检测系统[141]，Aegean Wi-Fi (AWID)无线网络数据集[142,172,184]，BGP RIPE路由信息服务数据集[143]，加拿大网络安全研究所提供的数据集[151-153,178,180-184,192-194]，CIFAR-10图像数据集[147,148]，CTU-13各种僵尸网络场景数据集[149,150]，以太坊经典数据集描述了对基于开源区块链的分布式计算框架和智能合约的攻击[151]，ICS Cyberattack天然气管道数据集[152,153,178,188,189]，KDD99和NSL-KDD网络流量记录数据集[146,148,154-160,168,176,177,179,180,182,184,186,187,190-192,194]，Coburg入侵检测数据集(CIDDS)用于异常检测的标记流量数据集[156]，UNSW-NB15原始流量文件数据集用于不同类型的攻击[155,157,158,160-163,176,190,191]，

表格 6
主要研究集中在检测功能上的总结

解决方案类别	用例	贡献	人工智能领域	作者
异常 & 事件	入侵检测系统	二元分类	学习	Ajayi 和 Gangopadhyay [141]，Li 等 [143]，Almlani 等 [144]，Corsini 等 [145]，Kumar 等 [147]，Maímo 等 [149]，Le 等 [150]，Saveetha &Maragatham 等 [151]，AL-Hawawreh 等 [152]，Zhang 等 [154]，Blanco 等 [155]，Nguyen 等 [156]，Alhowaide 等 [157]，Dutta 等 [161]，Perez 等 [162]，Singh 等 [163]，Catillo 等 [164]，Zhao 等 [165]，Nedeljkovic & Jakovljevic 等 [166]，Liu 等 [168]，Vidal 等 [169]，Latifl 等 [170]
		二元分类	规划 & 学习	Granato等人[142]
		二进制分类器的超参数调整	学习	Choras & Pawlicki [146]
		二元分类	感知 & 学习	Wu等人[148]
		二元分类	规划 & 学习	Vavra等人[153]
		解决二进制分类数据集中的类别不平衡问题	学习	Binbusayyis和Vaiyapuri [158]
		从二进制分类数据集中提取特征	学习	Herrera-Semenets等人[159]，Rashid等人[160]，Elnour等人[167]，Leevy等人[171]
		多类别分类	学习	Iwendi等人[173]，Toupas等人[174]，Jagtap等人[178]，Asifl等人[179]，Liu等人[180]从多类别分类数据集中提
		取特征	学习	Abdulhammed等人[172]，D'hooge等人[175]，Shafliq等人[183]解决多类别分类数据集中的类别不平衡问题
			学习	Huang & Lei [176], Gupta et al. [177]
		从数据集中提取特征进行多类别分类	规划 & 学习	Blanco et al. [181]
		多类别分类器的超参数调整	学习	Pawlicki et al. [182]
		二分类和多类别分类	学习	Mikhail et al. [184], Basnet & Ali [185], Diallo and Patras [186], Ullah & Mahmoud [188], Li et al. [189], Zhang et al. [190]
		解决数据集中二分类和多类别分类的类别不平衡问题	学习	Gupta et al. [187]
		数据可视化和二分类、多类别分类	学习	Zong et al. [191]
		从数据集中提取特征进行二分类和多类别分类	学习	Ieracitano et al. [192], Liu et al. [193], Xuan et al. [194]
安全持续监控	安全监控	数据处理和相关性	学习	Grammatikis等人[195]，Fausto等人[196]
		情境感知	学习	Kodituwakku等人[197]，Nikoloudakis等人[198]，Zhang等人[199]，Marino [200]
检测过程	暗网调查	情感分析	通信	Al-Rowaily等人[201]，Deb等人[202]
		主动威胁情报	学习	Ishikawa等人[203]
		主动威胁情报	通信	Pantelis等人[204]，Schäfler等人[205]
		探索讨论	通信	Fang等人[206]，Huang &Ban [207]
	自动评估不同威胁情报来源	报告的自动分析	通信	Kim等人[208]，Sarhan & Spruit [209]
		用于威胁信息提取的推文处理流程	通信 & 推理	Alves等人[210]
		用于威胁信息提取的推文处理流程	通信	Dionísio等人[211]
		挖掘推文以提取威胁信息	通信	Saura等人[212]
		漏洞情报	通信	Georgescu等人[213]
		威胁演化的识别	通信 & 推理	Sleeman等人[214]
		生成结构化的网络威胁情报记录	通信	Sun等人[215]
		发布威胁警告	通信	Sapienza等人[216]
		对中国威胁情报的分析	通信	Tsai等人[217]
	多语言威胁情报	用于俄语的威胁情报工具	通信	Ranade等人[218]
	基于人工智能的诱饵系统	基于诱饵系统的僵尸网络检测	学习	Memos &Psannis [219]
		基于分布式诱饵网络的早期警报系统	学习	Chatziadam等人[220]

BOT-IoT数据集包含物联网网络中的正常流量和僵尸网络流量[157,183]，基于Hadoop日志的异常检测基准数据集[164]，Unix用户行为日志的SEA数据集[165]，从不同传感器和执行器中收集的水处理厂网络数据的Secure水处理数据集（SWaT）[153,166,167,178]，从第三层互联网服务提供商收集的标记良好的流量的UGR'16数据集[168]，DARPA'99数据集包含实验环境中的真实/合成样本的在线和离线收集[169]，UCM 2011数据集包含真实流量跟踪[169]，以及面向工业4.0/IoT、工业物联网（IIoT）的TON_IoT数据集[170]。

二元分类是基本的分类类型，在文献中对入侵检测问题进行了最多的研究。 大多数研究者致力于应用不同的机器学习分类器[148-152,154-164,168-173,175-177]进行误用检测，但也有少数研究者致力于解决分类器的超参数优化问题[146]、数据集中的类别不平衡问题[158]以及从数据集中提取特征[159,160,167,171]。

在多类别分类中，数据集包含多个不相交的类别，并且属于每个类别的数据被赋予相同的标签。 在多类别分类问题中，除了正常流量之外，还有其他几类流量，包括拒绝服务（DoS）、分布式拒绝服务（DDoS）、远程到本地或用户到根的攻击。 入侵检测的多类别分类还探索了应用不同的分类器[173,174,178,179,180]、特征提取问题[172,175,181,183]、分类器的超参数调优[182]以及处理数据集中类别不平衡问题[176,177]。

一些研究人员在他们的方法中使用了二进制和多类别分类[191–201]。 他们为入侵检测问题提出了不同的分类器[184-186,188-190]，并解决了类别不平衡[187]、三维数据可视化[191]和特征提取[192–194]问题。

5.3.2. 安全持续监控

安全持续监控是对信息系统和资产进行实时监控，以深入了解其环境并检测安全事件。 人工智能可以通过使用动态、异构的信息网络来自动化监控，从而提供安全智能，该网络将处理从物理环境、网络、服务提供商、用户和包含敏感信息的系统的监控生成的日志数据。

5.3.2.1. 安全监控。安全监控是一个涉及从各种来源收集、分析和呈现数据的过程，旨在开发一种通用解决方案，揭示攻击者的操作方式和意图。 在这个方向上，研究人员正在积极处理和关联来自异构来源的数据[195,196]和情境感知[197–200]，以了解安全信息。

这个领域中的一个重要问题是从不断变化的来源和算法中处理大规模、动态和异构的安全信息。 因此，人工智能技术被用于详细分析以细粒度和可靠的方式跟踪相关的安全事件。Grammatikis等人提出了一个专门针对智能电网的安全信息和事件管理系统，用于检测、规范化和关联智能电网应用层协议的网络攻击和异常[195]。 同样，Fausto等人[196]关注将属于物理和网络领域的日志集成起来，并将它们的数据进行关联以检测关键基础设施中的潜在异常。

情境感知提供了对大规模网络的整体和具体视图，使安全分析师和研究人员能够实时识别、处理和理解信息。 为此，Koditu-wakku等人[197]引入了一个平台来处理和可视化实时的大规模网络数据，不仅可以监视和研究网络流量数据，还可以开发新的分析方法。 类似地，Nikoloudakis等人[198]提出了一个自动化的情境感知平台，利用软件定义网络（SDN）范式提供的实时感知功能，对网络可用实体进行漏洞评估，将它们分配到适当的连接片段，并持续监视底层基础设施。 研究人员[199,200]正在积极研究在网络物理系统中维持整体情境感知（网络和物理系统）的方法，因为网络和物理系统在任务关键应用中紧密集成。

5.3.3. 检测过程

检测过程涉及确保检测程序的维护和准备工作，可可靠地提供网络安全事件的威胁情报和意识。 这涉及持续改进和测试检测过程以实现高效工作。 人工智能可以用于通过从各种网络和内部资源中提取自动化威胁情报来提供互联网的积极警戒。 这些资源包括暗网、威胁情报共享平台和诱饵系统。 以下用例详细说明了使用人工智能技术维护检测过程的应用。

5.3.3.1. 暗网调查。暗网调查是一种监控与网络犯罪相关的暗网资源的过程。
该过程持续监控犯罪论坛和黑市，以便检测非法活动并采取适当措施以降低风险。 研究人员通过对暗网论坛的文本数据进行情感分析[201,202]、使用暗网数据进行威胁情报[203–205]以及识别关键攻击者、他们的资产和专业领域[206,207]来进行暗网调查。

需要自动化分析工具来识别潜在威胁，通过分析黑暗网络上的帖子的语言和目标，而无需手动监控大量帖子。情感分析用于使用自然语言处理（NLP）自动挖掘文本中的观点、意见和情绪。 在这个方向上，Al-Rowaily等人报道了一个用于分析与网络威胁、激进主义和冲突相关的英文和阿拉伯文文本情感的双语词汇资源（BiSAL）[201]。Deb等人构建了一种通过对黑客论坛上的帖子进行情感分析来预测恶意网络事件的方法学[202]。 这些论坛在表面网络和暗网上都有一定的预测能力，可以作为网络外部的信号来预测攻击，使用时间序列模型。

积极的威胁情报是在攻击发生之前通过收集来自黑客论坛和市场的信息来识别和解决安全风险。石川等人通过暗网流量分析提出了一种机器学习方案来跟踪攻击活动和感染设备的演变[203]。 他们的分析基于对扫描数据包的目标端口号所指示的目标网络服务之间的相关性的探索。 相比之下，Pantelis等人和Schafler等人致力于为暗网开发专门的信息检索技术，以支持网络威胁文本挖掘和商业情报。

探索黑暗网络论坛的讨论是提取关键思想、揭示热门话题、新兴威胁和攻击者社区中的关键角色的重要问题，以造福网络安全专业人员。研究人员[206,207]使用主题建模来提取主题、追踪主题的演变，并识别具有特定主题专长的关键黑客，揭示他们在地下市场中的角色。

5.3.3.2. 自动评估不同的威胁情报来源。 自动评估不同的威胁情报来源将有助于从各种来源中提取有用信息，如漏洞数据库、Twitter、新闻网站、事件报告和研究文章，以及时采取行动，确保系统的整体安全。

这涉及从多个来源处理基于证据的关于威胁和行为者的知识，以改善安全性和决策过程。 研究人员正在努力解决网络威胁情报来源的数量和异构性问题，以及它们的格式，以提供可操作的情报。 研究人员正在努力解决网络威胁情报来源的数量和异构性问题，以及它们的格式，以提供可操作的情报。

安全专业人员在分析网络安全报告时面临一个基本挑战，因为每天产生着无法估量的大量网络信息，需要自动化信息提取技术来促进数据检索和查询。 在这方面，研究人员[208,209]提出了创新的方法，利用命名实体识别从网络威胁情报报告中提取信息，帮助安全分析师尽快获得准确的威胁信息。

及时获取来自用户、安全组织和研究人员每天发布的开放源情报（OSINT）中的相关信息对于保持高水平的安全至关重要。 Twitter是一个重要的OSINT平台，也是网络安全情报的中心和聚集地，因为它具有自然的聚合能力、及时性、公共和私人意见的中心以及最重要的网络安全信息源（例如NVD、ExploitDB、CVE、Security Focus）。在这方面，已经提出了推文处理流程[210,211]，以及通过挖掘推文来提取安全问题[212]和用于情报验证的来源相关性评分[213]。

漏洞情报是从公共漏洞数据集（如CVE和NVD）中提取有关软件和系统漏洞的信息。这将有助于确定漏洞和攻击向量，以优先考虑安全工作和修补计划。

Georgescu等人[213]提出了一种自动诊断和检测物联网系统中潜在漏洞的系统，该系统使用物联网特定本体论和基于内容的提取最合适的漏洞，考虑到CVE数据库的持续变化和物联网系统的当前情况。

指定平台、暗网和社交网站上网络安全信息的爆炸性增长需要开发自动化工具来识别威胁演变[214]、生成结构化的网络威胁情报记录[215]、威胁警告[216]以及从本地威胁情报源[217]进行网络威胁情报（CTI）分析。 为了识别威胁演变，Sleeman等人[214]使用动态主题建模来展示时间戳集合中网络安全文件的关键主题的演变。 开源威胁情报发布平台（OSTIPs）上共享信息的增长和非结构化特性使得自动收集CTI记录变得具有挑战性。 在这种情况下，Sun等人[215]提出了一种从OSTIPs自动生成结构化CTI数据的方法，该方法结合了机器学习和自然语言处理，以实现准确、结构化和详细的数据，可供安全工具和分析师用于威胁缓解。

早期威胁预警系统通过提供及时通知可能发生的事件和安全问题的信息源，有助于防御网络攻击。Sapienza等人[216]通过挖掘安全专家的推文和与网络安全相关的博客来发布网络威胁警报。 威胁情报工具针对特定语言，帮助威胁情报专业人员更好地了解本地语言的网络威胁，以获取特定国家的知识。 在这个方向上，Tsai等人[217]开发了一个自动化系统，分析中国的网络威胁情报，以提高威胁情报的可见性。 这包括开发自动分类系统、推荐系统和威胁标记技术。

5.3.3.3. 多语言威胁情报。互联网的多语言特性要求将威胁情报来源进行翻译，以得出可靠的结论。 在这里，第三方翻译引擎不适用因为它们缺乏网络安全术语，并且其隐私和保密政策不足。Ranade等人[218]强调了为非英语语言开发威胁情报工具的重要性。 然而，他们的威胁情报工具只适用于俄语。

5.3.3.4. 基于人工智能的蜜罐。 使用蜜罐的主要目标是研究网络攻击的技术和行为，以改进现有的安全系统，并为这些类型的攻击做好准备。. 基于人工智能的蜜罐使用机器学习算法，利用来自多个蜜罐的数据[219]或来自暗网站的威胁情报数据[220]来预测攻击的概率，以尽早防止大规模的安全事件。 提出了一种结合机器学习和基于蜜网的检测方法的方法，用于确定物联网设备是否可能成为僵尸网络的组成部分[219]。 相比之下，Chatziadam等人[220]提出了一种基于分布式蜜罐网络的早期警报入侵检测系统，该系统利用暗网进行数据收集。

5.4. 响应

响应功能创建了管理和限制潜在网络安全事件影响的路线图。 这个功能非常重要，因为它代表了事件处理的第一道防线，并为未来的风险缓解方法的制定提供了支持。 这个功能包括提前规划，开发有效的处理问题的流程，分析事件以确定其原因、范围和影响，事件遏制，以及在攻击期间和之后的协调沟通。 通过使用人工智能技术进行响应活动，可以更快地解决事件，并减少安全分析师的时间和精力投入。 表7提供了关于响应功能的主要研究的摘要。 下面详细介绍了各个类别中的各种网络安全解决方案和人工智能应用案例。

5.4.1. 响应计划

这个类别是关于规划良好的响应程序，在事故发生期间和之后遵循，以限制其范围和影响。
这包括制定一个包含各种攻击场景和适当响应措施的应急计划，并结合持续的事件响应活动中所学到的经验来更新计划。 人工智能可以用于自动化响应计划，通过建立一个动态案例管理工具来记录、执行和更新应急计划。

5.4.1.1. 动态案例管理。 动态案例管理工具基于历史安全漏洞，记录不同的攻击场景，并在事故发生之前推荐适当的响应措施。. 这有助于针对特定类型的漏洞规划响应活动，并在事故关闭后进行知识管理记录。在这个领域的研究主要集中在通过基于案例推理，将最相似的事件与知识管理者进行匹配，并在事件之后修订知识管理者的自动化响应建议。

研究人员使用基于案例的网络安全事件解决系统，领域专家描述先例模型以保存和检索知识库中的先例。 研究人员积极使用分层结构[221,222]、机器学习[223,224]和本体论方法[225]来形式化先例库的基础。Kim等人[221]描述了一个包含RFM（最近性、频率和货币）技术属性的分层结构，以便快速响应安全漏洞。 他们的方法考虑了安全事件的情况，使用其频率和不同的属性值。 相比之下，Jiang等人[222]使用分层结构来存储潜在攻击场景的属性，例如

表7
关注响应功能的主要研究总结

解决方案类别	用例	贡献	AI领域	作者
响应计划	动态案例管理	使用RFM的基于案例的事件解决系统	推理	Kim等人[221]
		使用攻击情况的基于案例的事件解决系统	推理	Jiang等人[222]
		使用事件对象描述交换格式（IODEF）来保留、重用和共享问题解决经验	推理	Nunes等人[223]
		事件解决推荐系统	学习	Kraeva & Yakhyaeva [224]安全事件
		的表示	推理	Ping等人[225]
通信	自动化责任分配	资源分配的自适应和动态决策模型	学习	Shah等人[226]
	协作支持系统	异步协作支持系统	学习	Lin等人[227]
		同步协作支持系统	推理	Thomas等人[228]
分析	自动事件特征化	使用多类别分类进行严重性分配	学习	Decastro-Garcia等人 [229]
	警报分类	自动知识推理	学习	Husak等人[230]
		警报分组	学习	Manganiello等人[231]
		警报优先级排序	学习	Dey等人[232]
	取证分析	智能归因	学习	Chen [233]
		取证时间线中的异常识别	通信	Studiawan & Sohel [234]
		来自不同取证设备的证据关联	通信 & 推理	Amato等人[235]
		优化取证调查的决策框架	推理	Nisioti等人[236]
缓解	自动隔离	物理网络上的攻击定位	学习	Sakhnini等人[237]
		感染设备的隔离和替换	学习	Maimo等人[238]
	自动修复	选择最佳的一组对策	学习	Nespoli等人[239]
		安全分析师的推荐系统	学习	Husak等人[240]
		预防恶意软件传播的推荐系统	学习	Husak [241]
改进	长期改进	自动知识提取	通信	Piplai等人[242]，Peng 等人[244]
		自动知识提取	学习	Woods等人[243]

目标组织、攻击者信息、受影响资源以及对目标的潜在影响。Nunes等人[223]提出了使用基于案例的推理和事件对象描述交换格式来保留、重用和共享网络安全事件解决方案的问题解决经验。 K最近邻算法被用来计算案例相似度。 Kraeva和Yakhyaeva [224]提出了一个推荐系统，该系统使用神经网络将安全事件映射到嵌入向量中，然后找到最近的事件嵌入向量来推荐类似的案例进行解决方案推荐。

Ping等人[225]采用本体论方法提供了一个标准化的安全事件表示，以形式化先例库。

5.4.2. 通信

这一活动有助于在安全事件期间和之后协调各方之间的沟通。 这包括支持安全分析师在攻击期间的协作沟通，以及跨部门的威胁情报共享，以提高应急保护团队的响应能力。 这一活动还将确保在需要响应时分配备用角色和责任。 人工智能可以在以下两个用例中支持这一活动。

5.4.2.1. 自动责任分配。自动责任分配可以作为一个智能和适应性的决策支持工具，帮助安全运营中心（SOC）经理根据事件的性质、员工的专业知识和可用性分配事件响应职责。 在系统性文献综述中没有针对这个问题的具体研究。然而，Shah等人[*226*]解决了在网络安全运营中心面临多种因素干扰（如更高的警报生成率、新的警报模式和分析师缺席）时，如何做出最佳决策来分配资源（时间或额外的工作人员）以保持最佳运营效果的问题。 他们开发了一种基于随机、动态规划的自适应和动态决策模型，通过强化学习来解决。

5.4.2.2. 协作支持系统。 协作支持系统是一种信息系统，可以促进涉及事件响应的各方之间高效共享数据、信息和知识。. 这些参与者可以是组织内外的团队和员工。 人工智能技术被用于支持跨部门威胁情报共享和社区信息共享的网络防御、协作分析支持系统[*227*,*228*]。

这些协作支持系统分为两大类：异步和同步。 异步协作支持系统不提供实时通信，团队成员可以在方便的时候查看信息，并根据需要加入和退出对话。 林等人提出了一种基于留言板的跨部门威胁情报共享协作支持系统[227]。 他们设计了一个基于多Agent系统的黑板共享模块的监控机制，以解决威胁情报共享的并发问题，并提高任务执行效率。 相比之下，同步协作系统提供实时通信，支持实时响应，允许安全分析师快速交流发现，并引入有效的任务分工。 在这个方向上，Thomas等人提出了一个系统，提供可视化、实时通信和大量数据的高效转换，以提供对威胁和相应行动的全面理解。

5.4.3. 分析

分析是审查安全事件和响应活动的过程，以确保正确的处理过程被遵循。 这涉及收集和分析有关事件的信息，以支持事件的特征化和警报调查，以确定事件的严重性和影响。 它还通过取证分析来收集和保留未来诉讼的证据，支持恢复活动。 人工智能可以支持以下用例的分析过程。

5.4.3.1. 自动事件特征化。 事件特征化涉及识别事件类别的过程，根据响应计划进行。. 这包括识别事件的重要性以及其与其他事件的关系，以自动化地优先考虑进一步调查的事件。D*ecastro-Garcia*等人[229]提出了使用机器学习技术进行多类别分类的自动化模型，用于为不同类型的网络安全事件分配严重程度。

5.4.3.2. 警报分类。 警报处理和分类是一种高效准确地调查入侵警报、确定是否升级为事件响应的方式。. 通过提供自动化的知识推理[*230*]、警报分组[*231*]和警报优先级工具[*232*]，可以利用*AI*技术实现有效的警报分类，以识别威胁警报并升级进行进一步调查。

知识推理是将逻辑规则应用于知识库以评估和解释新信息的过程。

基于此，Husak等人[230]提出了一种使用顺序规则挖掘从共享入侵检测警报中提取知识并用于创建预测性和定制化黑名单的方法。

警报分组和警报优先级可以通过聚合或分类重要的威胁警报日志来解决威胁警报疲劳问题。 提出了使用自组织映射和无监督聚类算法对可能属于同一攻击场景的安全警报进行分组[231]，同时定义了分析事件每个属性的异常分数来衡量安全事件的优先级并找到潜在的异常事件[232]。

5.4.3.3. 取证分析。取证分析是一种事后技术，用于确定攻击的时间线，并揭示入侵的程度和来源，以及完全消除威胁并防止其再次发生的工具和方法。 这项调查还通过连接攻击者留下的证据碎片来创建一个可以作为法庭证据或支持起诉的踪迹。 人工智能技术可以用于帮助事件响应团队进行智能归因[*233*]，在取证时间线中进行异常识别[*234*]，从不同的取证设备中进行证据关联[*235*]，以及用于优化取证调查的决策框架[*236*]。

智能归因有助于推断安全事件的原因，通过发现不同实体和事件之间的关系。
基于上下文的学习可以用于捕捉状态并为源归因提供提示[233]。

取证时间线提供了关于网络安全事件发生前、期间和之后活动的信息。 在取证时间线中记录的安全事件可以被视为需要识别的异常。 可以使用深度自编码器来构建日志文件中正常事件的基准模型，并基于构建的基准模型为重构值设置异常阈值，以识别异常[234]。

取证调查员的主要目标是检测和分析欺诈活动，以便为法庭案件准备报告。 调查员可以利用各种取证工具来检查设备，但这些工具生成的数据格式不同，进一步复杂化了分析过程。因此，Amato等人[235]提出了一种基于语义的新方法，通过关联使用不同取证工具发现的证据，协助调查员进行分析过程。

新的对抗技术和技术的出现挑战了取证分析团队在有限资源下及时有效地调查事件的能力。 为了应对这些挑战，提出了一种新颖的决策支持系统模型，利用从已知对抗策略、技术和程序库中收集的威胁情报信息来推荐检查行动，以优化取证调查。

该模型考虑了潜在攻击活动、当前调查发现和可用于调查的预算之间的概率关系和接近程度值。 该模型考虑了潜在攻击活动、当前调查发现和可用于调查的预算之间的概率关系和接近程度值。

5.4.4. 缓解措施

缓解措施涉及一系列活动，以防止安全事件的扩大，并消除安全漏洞的任何长期后果。 这是一个关键步骤，不仅包含了事件，还减轻了新的漏洞或将其作为已接受的风险进行记录。 以下用例解释了在安全事件缓解中使用人工智能技术的情况。

5.4.4.1. 自动隔离。 自动隔离基于自动隔离设备或一组设备的原则，以响应威胁指标（IoC）的检测。. 这包括在感染后断开设备或一组设备的连接，或者找到高风险感染用户以给予他们更多关注。 自动隔离通过将攻击局限在物理网络上，并使用网络功能虚拟化（NFV）和软件定义网络（SDN）来隔离或替换感染的设备或用户。

将攻击定位到系统中的特定特征或测量值对于协助网络安全专业人员减轻攻击对通信网络的影响至关重要，通过将攻击与物理系统中的特定位置联系起来。 Sakhnini等人[237]提出了一个针对智能电网物理层的攻击分类和定位模型。 该模型使用集成和表示学习进行攻击分类，并使用卡方算法将攻击场景与特定特征相关联，并将攻击定位到一组特定测量值或系统中的特定位置。

在集成临床环境中，还需要一个实时的缓解系统来快速高效地处理网络安全事件。Maimo等人[238]提出了使用机器学习技术来检测和分类勒索软件攻击的传播阶段，并使用NFV和SDN范式来通过隔离和替换受感染设备来控制勒索软件的传播。

5.4.4.2. 自动修复。自动修复是一个引导性的问题解决过程，它可以通过简单的脚本或强大的上下文感知推荐系统自动执行补救措施。 人工智能技术被用于选择最佳的威胁消除对策[*239*]，在推荐系统中协助安全分析员进行弹性配置和安全工具的编排[*240*]，以及追踪攻击者的横向移动[*241*]。

快速有效地应对破坏性网络攻击是防御性网络安全的基本支柱，选择最佳的对策组合以应对威胁是一个重要的研究问题，以全自动的方式确定最佳反应。Nespoli等人[239]提出了一种基于人工免疫系统的新型网络安全反应技术，用于选择并执行对受风险的受保护系统的资产最佳组合的原子对策。

另一个重要的研究问题是决策支持系统和智能推荐系统，可以帮助安全分析师快速有效地保护资源和服务免受网络攻击。最近，Husak等人[240]提出了一个推荐系统，用于为关键基础设施建议最强大的配置，然后协调网络安全工具以进行快速缓解措施。 推荐系统被用于迅速阻止恶意软件的传播或跟踪攻击者的横向移动[241]。

5.4.5. 改进

改进有助于从事件检测和响应活动中吸取教训。 这包括根据吸取的教训更新响应计划和策略。 研究人员已经使用人工智能来进行自动化知识提取[242– 244]，基于事件报告。

5.4.5.1. 长期改进。从事件和威胁情报报告中提取知识为安全分析师提供可靠的信息，可用于检测或发现指示网络攻击的模式。*Piplai*等人[*242*]提出了一个系统，用于从事后行动报告中提取知识，通过将可比较的项目分组来汇总它，并在网络安全知识图中显示检索到的知识。 这些图帮助安全分析师找到不同网络攻击之间的相似之处。 相比之下，*Woods*等人[*243*]描述了一种基于数据挖掘技术的知识提取过程，用于识别指标和事件领域的子集，这些子集的完整事件信息可能对安全分析师和决策者有用。*Peng*等人[*244*]致力于从威胁相关文章中提取关于威胁行动的知识，基于条件共现度。 他们的工作还应用了机器学习，根据文章的相似特征对不同的威胁相关文章进行分类。

5.5. 恢复

恢复功能的主要目标是通过网络安全事件来维护弹性规划和及时恢复受损的能力或服务。 这鼓励及时返回正常运营，以减轻网络安全事件的影响，并以教训的形式提炼重要信息。 这个功能可以作为返回正常运营的路线图，借助以下网络安全解决方案的帮助。 表8总结了主要研究关注的恢复功能的主要贡献。

5.5.1. 恢复规划

恢复规划涉及维护、测试和执行恢复受网络安全事件影响的系统或资产的过程和程序。 这包括及时恢复丢失的数据和受损的能力，以确保一切正常运行。 基于人工智能的恢复规划可以在网络安全事件中自动化数据和系统恢复，并删除恶意软件或受污染的数据。 然而，在这个领域没有找到重要的研究工作。

5.5.2. 改进

这个解决方案类别涉及对安全事件的审查，通过从安全漏洞中学习来改进恢复计划过程。 它涉及根据经验教训对恢复计划和过程进行修订，并审查现有策略以匹配安全目标和目标。 人工智能可以用于自动审查现有策略、事件报告和审计日志，以寻找未来响应计划中的改进机会。

表8
关于恢复功能的主要研究总结

解决方案类别	用例	贡献	AI领域	作者
改进	事件报告的分析和聚合	漏洞的事后分析	通信	Meyers和Meneely [245
		]网络安全事件报告的聚合	规划 & 学习	Carriegos等人 [246]

从最新的安全漏洞中获取改进机会，用于未来的响应计划。

5.5.2.1. 事件报告的分析和聚合。 安全事件数据和报告的分析和聚合可以提供有价值的见解，可用于提供推荐和指导，将网络安全推进到更高水平。. 然而，事件数据的管理和分析是一项艰巨且耗时的任务。 人工智能技术可以用于高效的数据收集、聚合、信息提取、可视化和异构事件数据的预测[*245*,*246*]。 *Meyers*和*Meneely* [*245*]开发了一种自动化方法，用于对漏洞进行事后分析，以找到它们之间的复杂关系，使用自然语言处理。*Carriegos*等人[246]提出了一种有效的方法，用于聚合网络安全事件报告，以确定和定义准确的网络安全事件度量，并相应地进行预测和部署安全策略。

5.5.3. 通信

通过协调内部和外部各方之间的通信活动，通信有助于恢复。 虽然在这个系统化文献综述中没有在这个领域找到重要的研究文章，但提供专有平台以共享最新的安全漏洞或威胁信息是一个具有很高潜力的研究领域，有助于确保关键基础设施的网络安全。

6. 描述性分析

在现有技术分析之后，通过分类法、所使用的人工智能技术、出版类型、出版年份和研究进展的地理分布来回答RQ3，展示了主要研究的统计分布情况。

6.1. 文章类型分布

在为系统化文献综述选择的236篇文章中，101篇文章（43%）来自会议论文集，135篇（57%）来自同行评审的期刊，如图5所示。

6.2. 出版年份分布

本综述的时间跨度为2010年至2022年2月。如图6所示，直到2016年，人工智能在网络安全方面仍然是一个相对研究较少的主题，只有少数几篇研究发表在同行评审的期刊和会议上。 只有在过去的四年（2018年至2021年）中，对人工智能作为网络安全研究的兴趣才有所增加。

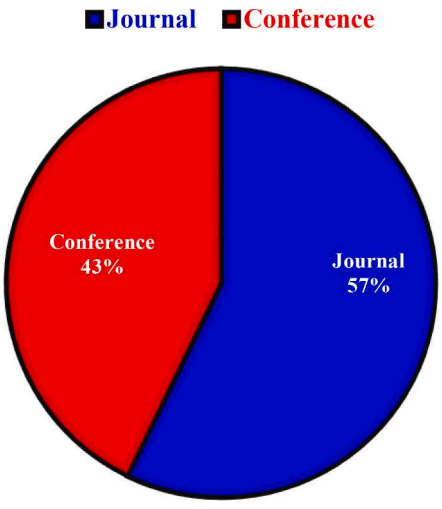


图5. 按出版物类型分布的文章

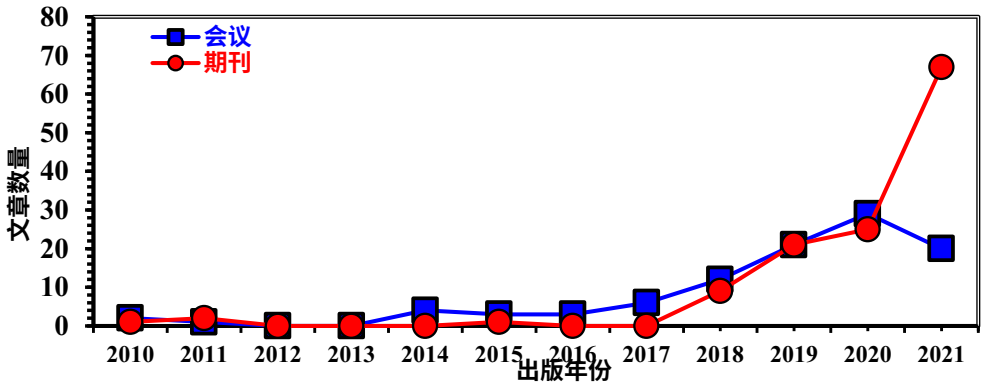


图6。 文章的年度分布。

主题。图6还显示了Covid-19对2021年会议出版物数量的影响，相比之前的年份，数量较少。
相反，2021年期刊出版物的数量比去年增加了近2.6倍。

6.3. 按地理区域分布

引用文章的作者的地理分布按照五大洲进行展示：亚洲、美洲、欧洲、非洲和大洋洲。来自不同洲的作者合作的文章被作为合作地区展示。对于选定的期刊出版物（参见图7(a)），30%的研究人员位于欧洲，其次是22%的研究人员位于亚洲，22%位于美洲。大洋洲在该主题上的研究论文相对较少，仅为4%。 非洲在选定的文章池中没有期刊出版物。其余22%的文章是来自不同洲的研究人员的合作努力。

关于会议出版物（见图7(b)），45%的作者位于美国，30%位于欧洲，14%位于亚洲。大洋洲在这个主题上的会议出版物非常少，只有3%。 与期刊出版物类似，非洲在所选研究中没有任何出版物。 其余8%的文章是来自不同洲的研究人员的合作工作。

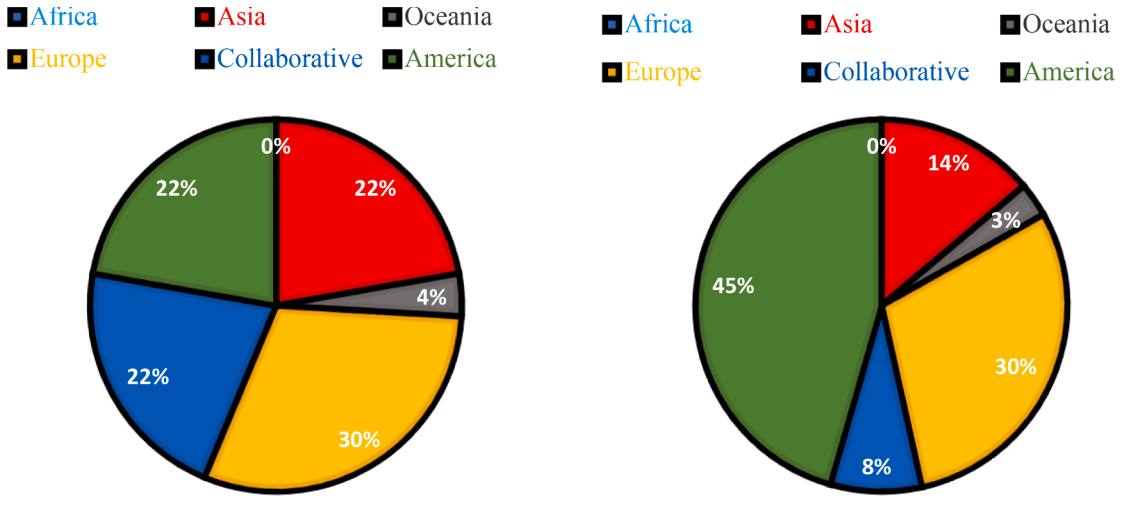
6.4. NIST网络安全功能的分布

完整阅读了主要研究，并提取和总结了相关数据，详见第5节。 确定每个主要研究都专注于特定的网络安全功能。 根据主要关注的网络安全功能，将识别、保护、检测、响应和恢复等五个主要类别的研究进行分类。图8显示了这五个类别的研究分布情况。 在期刊出版物中（见图8(a)），36%的研究集中在异常检测和网络安全事件上。 由机器学习算法控制的检测可以及时实现自动攻击检测和防御。 第二受欢迎的类别是识别，占28%，其次是保护和响应，分别占25%和10%。

2021年和2022年只有很少的研究关注使用人工智能进行恢复。

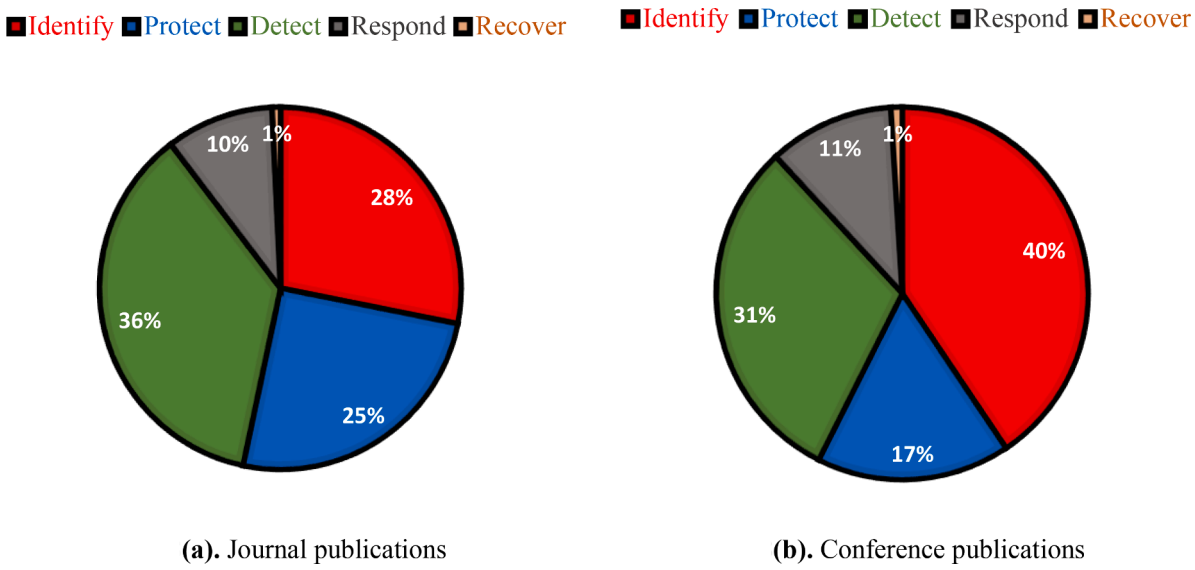
会议出版物显示了相同的趋势，但具有不同的百分比，如图8(b)所示。 关于人工智能在网络安全中的应用，大多数会议文章主要集中在三个主要类别：识别（40%），检测（31%）和保护（17%）。 响应和恢复分别占剩余的12%中的11%和1%。

图9(a)和9(b)展示了过去6年的百分比分布情况。 数据显示，随着时间的推移，识别、保护和检测类别的出版物数量逐渐增加，而响应和恢复功能则最近才开始受到关注。



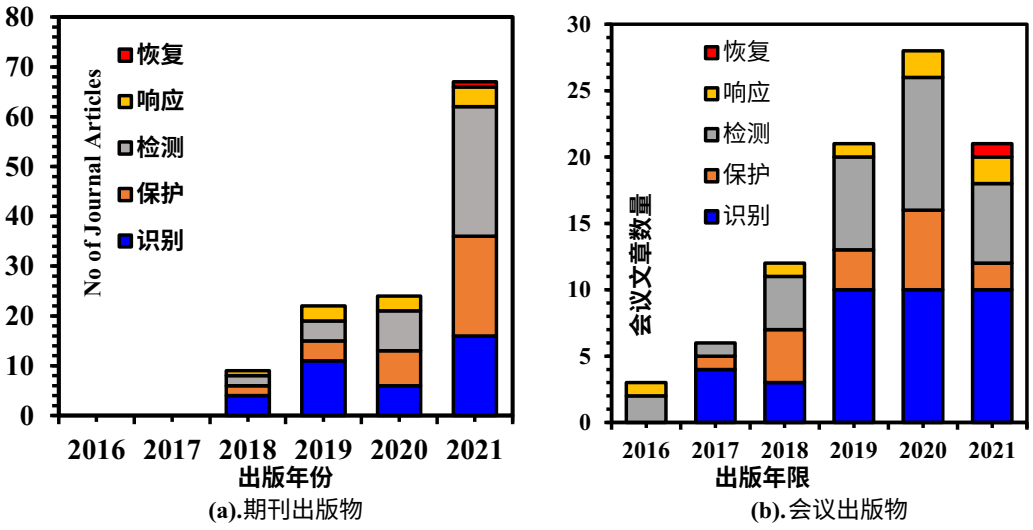
(a). Journal publications **(b).** Conference publications

图7。 与人工智能在网络安全领域相关的主要研究的地理分布。



(a). Journal publications **(b).** Conference publications

图8. 与NIST功能相关的主要研究分布。



(a).期刊出版物 (b).会议出版物

图9.过去6年期刊和会议出版物中与NIST功能相关的主要研究分布。

6.5. 使用的人工智能技术的分布

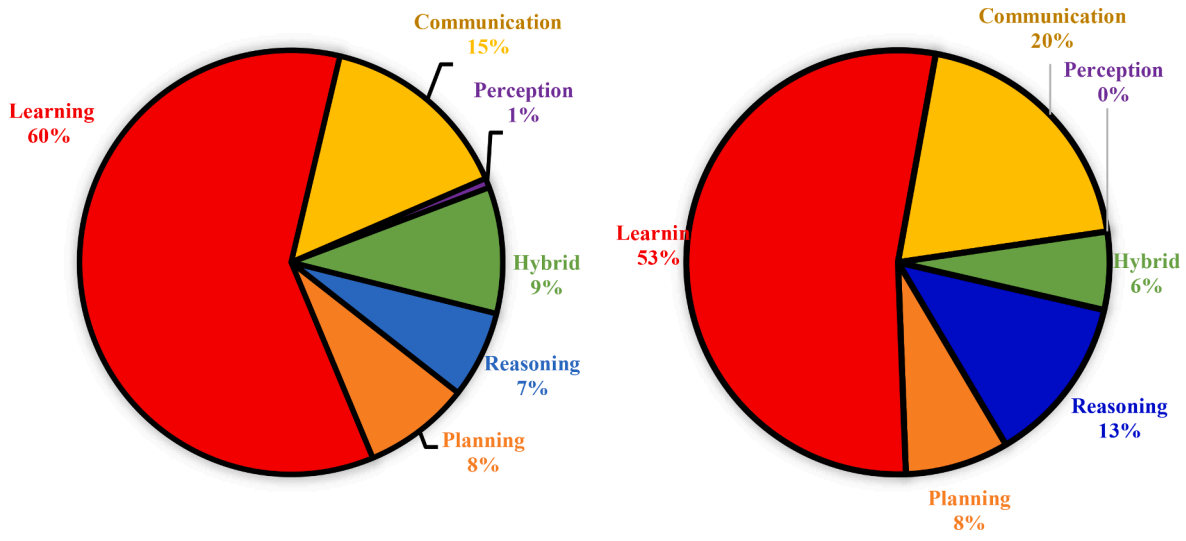
定性分析中研究的另一个特征是所使用的人工智能技术。 人工智能技术指的是可以指定的方法论方法，包括算法、架构、数据或知识形式和算法。 AI Watch提出的核心人工智能领域，即推理、规划、学习、通信和感知，被用于分析。 推理领域的主要研究解决了机器将数据转化为知识或从数据中推断事实的方式。 规划研究侧重于自动规划设计和执行具有精心优化解决方案的策略。 学习领域的研究解决了自动学习、预测、适应和对变化做出反应的解决方案。

通信领域的研究强调机器在口头或书面人类对话中识别、解释、理解或产生信息的能力。 感知领域的主要研究解决了通过视觉和听觉感知周围环境的问题。 确定了每个主要研究的人工智能领域重点。图10(a)和10(b)说明了期刊和会议中关于五个确定的人工智能领域的研究分布。

图11(a)和11(b)显示了过去6年在期刊和会议论文中使用的人工智能技术的份额。 从图表中可以清楚地看出，学习是网络安全应用中最广泛使用的技术，其次是通信。

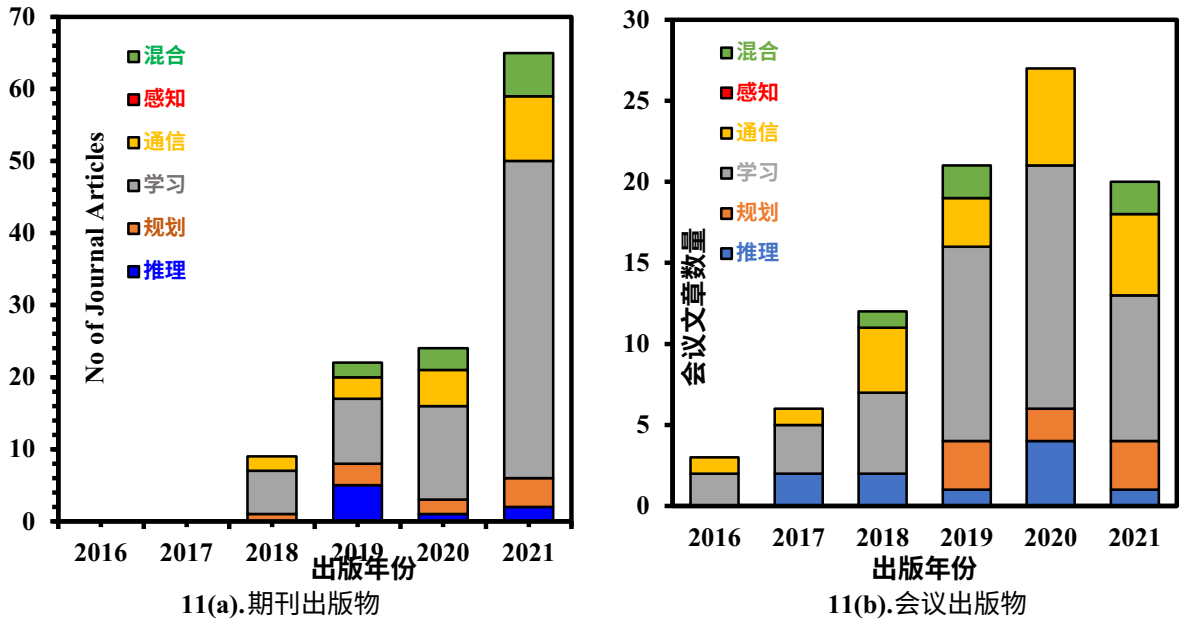
7. 研究空白

为了回答本文的第四个也是最后一个研究问题（RQ4），我们审查了与我们的研究问题相关的文献，以突出潜在的研究空白，并确定未来网络安全人工智能研究的机会。 进行网络安全人工智能研究的关键要素是确定新兴的应用领域、适当的资源（例如数据来源和管理、计算基础设施等）以及先进的人工智能技术，以成功采用人工智能来进行网络安全。 本节提供了未来研究的有用方向，主要包括：（i）新兴的网络安全应用领域，（ii）数据表示，（iii）网络安全的先进人工智能方法，以及（iv）新基础设施的研究和开发。



(a). Journal Publications **(b).** Conference Publications

图10.主要期刊和会议研究的人工智能技术分布。



11(a).期刊出版物 11(b).会议出版物

图11.过去6年中AI领域的使用分布。

7.1. 网络安全应用的新兴领域

推进网络安全的人工智能需要一个坚实的应用领域基础。 此外，未来的展望需要对新的网络安全活动进行自动化，并持续改进现有活动。研究界可以解决的新兴应用领域如下：

- 关键风险指标的自动检索：在SLR文献中没有涉及实时自动检索风险指标的研究文章。 因此，开发一个早期预警系统来指示由于政策违规、红旗或其他症状而导致的风险发展是一个诱人的未来研究方向。 自动检索关键风险指标，如未打补丁的系统存在、尝试入侵的次数、故障间平均时间等，并将其转化为知识，将是一个重要的研究方向。

及时修复风险有助于防止网络安全漏洞。

- 检测新攻击：防御零日攻击是现代网络安全中最具挑战性的方面之一。 零日攻击是指针对新的、尚未广为人知的软件漏洞的网络攻击。 当然，防御你不知道存在的东西是一个重大障碍。 为此，需要对整个信息技术环境进行全面的可见性，包括终端、网络和云。

- 预测性智能：预测分析可以促进例行网络安全任务的自动决策，包括攻击路径预测、恶意软件预测、数据分类、垃圾邮件过滤、漏洞分类、安全预测和任务映射。 尽管这些任务被广泛使用且定义明确，但许多常用技术要么不自动化，要么误报率很高。

基于人工智能的预测分析可以用于帮助解决其中一些特定问题。 深度贝叶斯预测、突发检测、带有时间约束的深度生成建模、基于时间的神经图网络和其他技术是有前景的预测分析技术。 每种策略都可以通过纳入行业特定的政策、信息、任务和需求来改进。 安全运营中心（SOC）分析师和CTI专业人员，特别是那些在企业内部的战术和操作层面的人员，是可以从改进的预测中受益的人群之一。

- 多语言威胁情报：在SLR语料库中，只有一项研究分析了对陌生语言的威胁情报理解。 互联网的多语言特性使得网络安全社区在从社交媒体、博客和暗网市场中战略性地挖掘威胁情报方面面临更大的挑战。 根据赛门铁克的研究[247]，非英语国家既是最主要的攻击源，也是最主要的攻击目标。 因此，应该使用包含非英语内容的数据集来评估在其他语言中进行文本挖掘的有效性。 此外，大多数预处理工具和库只支持英语语言。

新的研究可以提供针对其他语言的工具，或者为私人安全设置创建语言翻译器。

- 基于人工智能的网络防御和弹性：组织可以利用其分析数据自动应用适当的安全控制。 例如自动化威胁建模、自动化补丁、自动化修复和缓解以及自动化网络分割和重组。 SOC分析师和操作员可以从智能工作自动化中获得巨大的好处。 增强的人工智能代理、强化学习、演员关键网络、选定的防御、对抗学习技术和贝叶斯网络是一些现代人工智能技术，使得这些任务成为可能。 未来的研究可以探讨每种技术如何为各个行业（例如科学网络基础设施、企业IT、基于传感器的环境等）提供适当的网络安全保护。

- 数据泄露的预防和发现：近年来，数据泄露的普遍发生导致了企业和消费者的损失。 这个主题需要得到网络安全研究人员的重视，因为现有研究大多集中在观察内部人员的行为和活动，或者通过分析终端的遥测数据来检测内部威胁和高级持续性威胁（APTs）的存在。 然而，对于防止意外数据泄露的敏感数据识别的研究仍然不足。 基于机器学习和自然语言处理的人工智能技术可以用于敏感数据的发现，以及监控和控制大数据场景中的数据流动，以防止或分析数据泄露。 研究还应考虑分析暗网以发现任何意外的数据泄露，以恢复控制并保护声誉。

- 虚假文件生成：在这个网络战时代，保护重要的数字资产，如知识产权和国家安全数据，至关重要。 在SLR语料库中，只有一项研究报告了在网络渗透后使用自动生成可信和交互式虚假文件来保护重要文件的情况。 这种使用人工智能进行虚假文件生成的概念是为了通过伪造信息创建大量虚假版本的任何文件来保护敏感材料的相对较新的概念。

因此，这个话题需要得到网络安全社区的更多关注。

- 基于上下文的警报处理 &分类：每天都有成千上万个警报和事件被精心收集用于威胁分析，这使得安全团队不堪重负，需要有经验的威胁情报专家的关注。 当前的研究文章已经尝试通过高级管理来关联安全警报，考虑它们的逻辑关系，以克服这个问题。

在将它们转发给用户之前，需要对事件的关系进行优先级排序。 然而，它们仍然缺乏对特定网络环境中不同事件上下文的考虑。 研究人员可以使用不同的语言模型来实现事件上下文的表示学习，并致力于开发适应动态网络环境的自适应方法。 此外，在数据可视化和在线更新能力方面还需要进一步努力，以可视化警报之间的关系，并开发高效的警报分类系统。

- 基于人工智能的事件响应：由于攻击者从最初的入侵到完全控制企业基础设施所需的时间大大缩短，因此自动化响应工作对于有效地减轻、遏制或智能化攻击也至关重要。 事件响应过程的自动化需要记录过去安全事件所获得的知识，以及应用特定解决方案触发的事件，并随时间记录新的威胁模式及其特征。 然后，可以利用这些知识创建自动化的事件响应手册，根据其专业知识、可用性或案例历史进行建议、资源分配或责任分配。 这些自动化安全手册将推动积极的防御，并帮助在不断复杂和不断演变的威胁面前保持领先。 此外，将来，在威胁情报平台上共享这些标准化的安全手册将允许不同组织在机器时间内对事件做出响应并消化这些信息。

7.2. 数据表示

为了使人工智能发挥作用，拥有良好的数据至关重要。 最大的挑战在于选择适当的训练数据集以及处理数据的多样性和速度。 因此，下面描述了数据表示和质量、最新挖掘以及上下文感知对于网络安全应用中的人工智能模型的训练和建模的重要性：

- 精细化的数据表示。数据的表示对于人工智能算法的性能至关重要。目前，将数据展示为扁平化的特征向量是最常用的描述网络安全的方法。尽管被广泛使用，但这种方法忽略了数据中明显存在的重要联系（例如序列）。 因此，在实际应用中使用这种表示方法可能会产生明显更差的结果。 未来从事网络安全人工智能研究的学者可以仔细评估他们感兴趣的环境中网络安全数据的存在方式，并选择最适合的替代方法来更好地描述所关注的现象，以减轻这种困难。 例如，可以将虚拟机中的文件系统和应用程序可视化为树或图，以捕捉它们的依赖关系（由于它们的层次结构性质）。 网格、序列和非欧几里德表示是其他潜在的表示方法（例如张量、立方体）。 此外，在选择合适的表示方法时，还可以考虑关键数据属性、组织需求和相关的社会行为经济学理论。

- 网络安全中的上下文感知：目前关于网络安全的研究通常从相关的网络数据开始，其中包含一些低级特征。 这些数据集可以用于应用数据挖掘和机器学习方法，以找到一个连贯的模式来准确解释它们。 然而，要决定是否存在可疑活动，还可以使用更广泛的上下文信息，如事件之间的时间和空间关系，或者连接和依赖关系。 例如，虽然安全专业人员可能不认为单个连接是恶意的，但其他方法可能将其分类为DoS攻击。 因此，以往网络安全工作中无法利用上下文知识预见危险或攻击是一个重大弱点。上下文-

因此，意识到适应性网络安全解决方案可能是人工智能在网络安全研究领域的另一个研究方向。

- 增量学习 &最新性挖掘：为了提供数据驱动的决策，基于机器学习的安全模型通常使用大量的静态数据。 然而，用户和恶意对手的行为模式可能不是静态的，而是随时间变化而大不相同。

  因此，在解决常见网络安全任务（如数据分类、垃圾邮件过滤、漏洞分类和任务映射）的预测分析中，最近的行为模式和相应的机器学习规则比较老旧的规则更有趣且更重要。 因此，在网络安全研究中，另一个挑战可能是有效地将最新性分析的思想应用于网络安全解决方案。

7.3. 网络安全的高级人工智能方法

为了充分发挥前述数据源、应用领域和数据表示的潜力，需要更复杂的人工智能技术。 在众多选择中，三种关键的新技术——多数据源分析、可解释人工智能（XAI）和增强智能（人工智能界面）——对于实际可用的网络安全人工智能的发展具有重大影响。

- 多数据源分析：当前网络安全领域的人工智能研究和实践存在一个重大缺陷，即孤立使用单个数据集。 这往往是由于无法访问多个数据集（在学术界很常见）或无法理解不同数据集之间的关系。 不同时处理多个数据集可能导致对环境的不完整评估。 未来的网络安全人工智能研究可以尝试更全面地利用不同数据源的特点来解决这个问题。

  基于深度学习、短文本匹配算法（如深度结构语义模型）、多视图方法（如多源）和多任务学习技术的实体匹配是多数据源分析的有希望的方法。 成功融合多个数据集可以产生新的衍生属性，增强风险管理（如漏洞评估）并全面了解组织的网络安全状况。

- 可解释人工智能（**XAI**）的应用：了解算法如何以及为什么得出初始结论在网络安全领域至关重要。 不幸的是，当前基于人工智能的算法在决策过程中缺乏透明度。 尽管在暗网调查、漏洞评估等高影响力网络安全应用中表现出色，但它们以其“黑盒子”性质而臭名昭著。 未来网络安全的人工智能研究可以探索可解释和可解释的人工智能如何提高算法的性能，并打开它们的黑盒子特性，以减少这些限制，增加其在重要网络安全利益相关者中的接受度和可信度。

- 增强智能（人工智能界面）：根据许多网络安全专家的观点，基于人工智能的算法和系统不应仅用于网络安全决策。 相反，为了实现更好的决策过程，基于人工智能的方法应与人类行动紧密结合（例如，在分析过程中有安全分析师作为积极成员）。 这些方法，也称为增强智能或人工智能界面，有潜力显著超越算法或单个人类的应用。 人工智能与人类在关键和基本的网络安全任务中的互动的广度、范围和深度尚未得到充分研究，但这是迫切需要的。

  这样的研究必然需要采用多学科的方法，特别是结合心理学、认知科学、人机交互等领域的观点。

7.4. 新基础设施的研究与开发

人工智能逐渐成为网络安全的关键组成部分，以提高各种规模和行业的组织的网络安全效率。 因此，有必要研究和开发新的基础设施，以支持人工智能技术，处理大量的内部系统数据以及外部安全研究数据源，以提供全球和内部安全事件的实时网络安全。 成功实施人工智能在组织和网络安全研究人员中的关键研究差距如下：

- 缺乏威胁情报平台（在国家和国际层面上）：网络现实非常复杂和动态；总是有新的威胁，攻击专门设计来规避已知的潜在情景。 因此，需要专有平台来促进指定同行之间的合作，讨论和共享最新的威胁数据。 目前缺乏设计灵活、适应性强且网络化的威胁情报平台，这些平台主要依赖于国家和国际层面上的指定信息共享中心，但不仅限于此。 政府、关键基础设施的所有者和运营商以及其他实体将从威胁情报平台共享的准确、可用、及时和相关的威胁信息中受益。 通过提高情境意识和促进有效的风险知情决策，这种共享改善了关键基础设施的安全性和可靠性。

- 缺乏新的、实时或更广泛的数据集：数据集是网络安全中人工智能的最重要组成部分。 大多数可用的数据集已经过时，可能不足以理解不同网络攻击的最新行为模式。 本调查发现许多研究都将人工智能技术应用于同一数据集。 例如，大多数研究使用DARPA98、KDD99、NSLKDD和CICIDS2017进行检测。 这些研究明显缺乏使用最新和不同数据集对其技术进行评估。 此外，通过在多个数据集上验证特定上下文研究，可以在不同场景下进行分析。

8. 限制

所提供的系统性文献综述为网络安全和人工智能技术的交叉点提供了有价值的信息，并确定了未来研究的研究空白。 然而，我们的研究遗漏了那些发表在Scopus以外的科学数据库中或使用不同关键词的文章。 此外，由于需要花费时间分析所选的主要研究以获得可靠的结果，因此本研究不包括最近的出版物（2022年2月之后的出版物）。

9. 结论

本文系统地研究了人工智能在网络安全中的应用，并探讨了未来的研究方向。 通过从Scopus数据库中选择了2395篇相关文章中的236篇主要研究，覆盖了2010年至2022年2月的13年时间段，实现了这一目标。 本研究讨论了在网络安全领域应用的不同人工智能技术以及哪些网络安全活动已经利用了人工智能技术。 通过对选定的文献进行分析，主要从以下几个方面进行：（一）人工智能在网络安全中的分类；（二）按年份划分的发表频率；（三）按地理区域划分的发表频率；（四）网络安全的贡献类型；（五）使用的人工智能技术类型。

本文通过深入探讨特定应用案例和研究的理论基础，全面研究了人工智能在网络安全中的现有研究的“如何”和“什么”。 本研究对于人工智能在网络安全中的应用做出了贡献。

通过分析人工智能在网络安全领域的应用演变和识别研究空白，总结了相关知识体系。 研究了人工智能在网络安全中的演变，涉及不同功能、解决方案类别、具体用例和人工智能技术类型。

分析结果显示，相关出版物数量正在增加，但在实施基于人工智能的网络安全解决方案时，需要更加关注获取和表示与不同网络安全功能相关的历史数据。 本研究的主要贡献是对主要研究进行分类，以整合该领域的文献状况，并理解人工智能在网络安全中的重要性。 此外，本文提出了未来的研究方向，以解决人工智能在网络安全成功应用中出现的新问题。

CRediT作者贡献声明

Ramanpreet Kaur：概念化、方法论、验证、调查、数据整理、撰写原稿、审阅和编辑、可视化。杜圣加布里耶尔奇：方法论、项目管理。托马斯·克洛布卡尔：方法论、审阅和编辑、监督、项目管理、资金获取。

竞争利益声明

作者声明以下可能被视为潜在竞争利益的财务利益/个人关系：Tomaˇz Klobuˇcar 报告斯洛文尼亚研究机构提供了财务支持。Tomaˇz Klobuˇcar 报告斯洛文尼亚政府信息安全办公室提供了财务支持。

数据可用性

本文研究未使用任何数据。

致谢

本研究得到了斯洛文尼亚研究机构ARRS（V2–2147和P2–0037）和斯洛文尼亚共和国政府信息安全办公室（V2–2147）的支持。

参考文献

[1] Bhardwaj, M.D. Alshehri, K. Kaushik, H.J. Alyamani, M. Kumar, Secure flramework against cyber-attacks on cyber-physical robotic systems, J. Electron. Imaging 31 (6) (2022), 061802-061802.
[2] P. Chithaluru, A.T. Fadi, M. Kumar, T. Stephan, 计算智能启发的自适应机会式聚类方法用于工业物联网网络，IEEE物联网杂志（2023），https://doi.org/10.1109/JIOT.2022.3231605。
[3] M. Barrett, 技术报告，国家标准与技术研究所，美国，马里兰州盖瑟斯堡，2018年。
[4] I. Wiafle, F.N. Koranteng, E.N. Obeng, N. Assyne, A. Wiafle, S.R. Gulliver, 人工智能在网络安全中的应用：文献系统映射，IEEE Access 8（2020）146598–146612。
[5] Z. Zhang, H. Ning, F. Shi, F. Farha, Y. Xu, J. Xu, F. Zhang, K.K.R. Choo, 人工智能在网络安全中的应用：研究进展，挑战和机遇，人工智能评论55（2022）1029–1053。
[6] J. Martínez Torres, C. Iglesias Ćomesana, P.J. García-Nieto, 机器学习技术在网络安全中的应用，《机器学习与网络安全国际期刊》10 (10) (2019) 2823–2836.
[7] T.C. Truong, I. Želinka, J. Plucar, M. Candík, V. Sulc, 人工智能和网络安全：过去、现在和未来，《人工智能和进化计算在工程系统中的应用》, 2020, pp. 351–363.
[8] S. Samoili, M.L. Cobo, E. Gomez, G. De Prato, F. Martinez-Plumed, B. Delipetrev, A.I. Watch, 技术报告，联合研究中心（塞维利亚站）, 2020.
[9] 人工智能高级专家组（HLEG AI），人工智能的定义：主要能力和学科，（2019）。从布鲁塞尔检索 https://ec.europa.eu/newsroom/dae/document.cflm?doc_id=56341。
[10] D. Zhao, A. Strotmann，引文网络的分析与可视化，信息概念、检索和服务综合讲座，7 1 (2015) 1–207。
[11] V.G. Promyslov, K.V. Semenkov, A.S. Shumov，一种资产网络安全分类的聚类方法，IFAC-PapersOnLine 52 (13) (2019) 928–933。
[12] K. Millar, A. Cheng, H.G. Chew, C.C. Lim，操作系统分类：一种简约方法，2020年机器学习和控制论国际会议（ICMLC），2020年，pp. 143–150。
[13] A. Aksoy, M.H. Gunes, 使用网络流量自动识别物联网设备，在：IEEE国际通信大会（ICC），2019年，第1–7页。
[14] A. Sivanathan, H.H. Gharakheili, F. Loi, A. Radflord, C. Wijenayake,A. Vishwanath, V. Sivaraman, 使用网络流量特征对智能环境中的物联网设备进行分类，IEEE移动计算期刊18（8）（2018）1745–1759。
[15] I. Cviti´c, D. Perakovi´c, M. Peri˙sa, B. Gupta, 集成机器学习方法用于智能家居中物联网设备的分类，国际机器学习与控制期刊12（11）（2021）3179–3202。
[16] H. Cam, 在IEEE军事通信会议（MILCOM）中对恶意软件感染资产进行在线检测和控制，2017年，第701–706页。
[17] H.I. Kure, S. Islam, M. Ghazanflar, A. Raza, M. Pasha, 资产重要性和风险预测，以实现对网络物理系统的有效网络安全风险管理，神经计算应用。34（1）（2022）493–514。
[18] M. Vega-Barbas, V.A. Villagr´a, F. Monje, R. Riesco, X. Larriva-Novo, J. Berrocal,基于本体的动态风险管理系统在行政领域的应用，应用科学。9（21）（2019）4547。
[19] B. Tozer, T. Mazzuchi, S. Sarkani, 使用多目标强化学习优化攻击面和配置多样性，在IEEE第14届国际机器学习和应用会议上，2015年，第144–149页。
[20] L.E. García-Hernandez, A. Tchern´ykh, V. Miranda-L opez, M. Babenko,A. Avetisyan, R. Rivera-Rodriguez, G. Radchenko, C.J. Barrios-Hernandez,H. Castro, A.Y. Drozdov, 多目标配置安全分布式云数据存储, in: 拉丁美洲高性能计算会议,2019, pp. 78–93. 九月.
[21] M. Sharifli, F. Eugene, J.G. Carbonell, 个性化安全设置的学习, in: IEEE国际系统、人和控制会议, 2010, pp. 3428–3432.
[22] D. Bringhenti, G. Marchetto, R. Sisto, F. Valenza, F.J. Yusupov, 迈向完全自动化和优化的网络安全功能编排, in: 第四届计算、通信和安全国际会议 (ICCCS), 2019, pp. 1–7.
[23] A.J. Varela-Vaca, R.M. Gasca, J.A. Carmona-Fombella, M.T. G ómez-L opez, AMADEUS：面向自动化安全测试的研究，第24届ACM系统和软件产品线会议论文集，2020年，第1–12页。
[24] A.J. Varela-Vaca, R.M. Gasca, R. Ceballos, M.T. G ómez-L opez, P.B. Torres, CyberSPL：使用软件产品线验证系统配置的网络安全策略合规性框架，Appl. Sci. 9 (24) (2019) 5364.
[25] Y. Liu, A. Sarabi, J. Zhang, P. Naghizadeh, M. Karir, M. Bailey, M. Liu, Cloudy with a chance ofl breach：预测网络安全事件，第24届USENIX安全研讨会（USENIX Security 15），2015年，第1009–1024页。
[26] Z. Zhan, M. Xu, S. Xu, 从网络望远镜数据中表征网络安全姿态, 2014年可信系统国际会议, 第105–126页, 2014年.
[27] S.N.G. Gourisetti, M. Mylrea, E. Gervais, S. Bhadra, 基于多场景用例的建筑物网络安全框架Web工具演示, 2017年IEEE计算智能研讨会系列(SSCI), 第1–8页.
[28] L.V. Stepanov, A.S. Koltsov, A.V. Parinov, 基于遗传算法评估企业网络安全性, 2020年国际俄罗斯自动化会议, 第580–590页.
[29] V.L. Narasimhan, 使用深度学习评估虚拟电厂的网络安全经济风险, 2021年第7届国际电力能源系统会议(ICEES), 第530–537页.
[30] H.H. Nguyen, D.M. Nicol, 在不确定性存在的情况下估计网络攻击造成的损失，在2020年IEEE第19届计算与通信中的信任、安全和隐私国际会议（TrustCom）上，2020年，第361–369页。
[31] C. Ponsard, V. Ramon, M. Touzani, 通过结合使用i*和基础设施模型来改进网络安全风险评估，在第14届国际iStar研讨会上，2021年，第63–69页。
[32] O. Odegbile, S. Chen, Y. Wang, 在传统非SDN网络中实施可靠的策略执行，在2019年IEEE第39届分布式计算系统国际会议（ICDCS）上，2019年，第545–554页。
[33] F.D. Nembhard, M.M. Carvalho, T.C. Eskridge, 将推荐系统应用于安全编码，EURASIP J. Infl. Security 1 (2019) 1–24页。
[34] 刘胜, 林刚, 韩庆龙, 温松, 张军, 向阳, DeepBalance: 深度学习和模糊过采样在漏洞检测中的应用, IEEE模糊系统学报, 2019年, 28(7): 1329–1343.
[35] 全胜, 金海克, AutoVAS: 一种基于深度学习的自动漏洞分析系统, 计算机安全, 2021年, 106: 102308.
[36] 哈夫, 麦克兰纳汉, 莱, 李琪, 一种用于跟踪漏洞的推荐系统, 第16届可用性、可靠性和安全性国际会议, 2021年, pp. 1–7.
[37] 伊奥尔加, 科尔拉特斯库, 格里戈雷斯库, 桑德斯库, 达斯卡卢, 鲁吉尼斯, Yggdrasil—从Twitter中早期发现网络漏洞, 第23届控制系统与计算机科学国际会议, 2021年, pp. 463–468.
[38] T. Saha, N. Aaraj, N. Ajjarapu, N.K. Jha, SHARKS: 基于机器学习的物联网和网络物理系统的风险扫描的智能黑客方法，IEEE Trans. Emerg. 10 (2) (2021) 870–885.

[39] Y. Wang, Z. Wu, Q. Wei, Q. Wang, NeuFuzz: 基于深度神经网络的高效模糊测试，IEEE Access, 7 36340–36352.
[40] J. Wang, B. Chen, L. Wei, Y. Liu, Skyflire: 基于数据驱动的模糊测试种子生成，in: 2017 IEEE 安全与隐私研讨会 (SP), 2017, pp. 579–594.
[41] P. Godefiroid, H. Peleg, R. Singh, 学习&Fuzz: 输入模糊测试的机器学习方法，in: 第32届IEEE/ACM国际自动化软件工程大会 (ASE), 2017, pp. 50–59.

[42] C. Cummins, P. Petoumenos, A. Murray, H. Leather, 通过深度学习进行编译器模糊测试, 在：第27届ACM SIGSOFT国际软件测试与分析研讨会, 2018, pp. 95–105.
[43] H. Xu, Y. Wang, S. Fan, P. Xie, A. Liu, DSmith: 基于生成式深度学习模型和注意力机制的编译器模糊测试, 在：2020年国际神经网络联合会议(IJCNN), 2020, pp. 1–9.
[44] Y. Chen, C.M. Poskitt, J. Sun, S. Adepu, F. Zhang, 基于学习引导的网络模糊测试用于测试网络物理系统防御, 在：第34届IEEE/ACM自动化软件工程国际会议(ASE), 2019, pp. 962–973. 11月.

[45] D. She, K. Pei, D. Epstein, J. Yang, B. Ray, S. Jana, NEUZZ: 基于神经程序平滑的高效模糊测试, 在: IEEE安全与隐私研讨会 (SP), 2019, pp. 803–817.

[46] X. Liu, X. Li, R. Prajapati, D. Wu, DeepFuzz: 自动生成语法有效的C程序进行模糊测试, 在: AAAI人工智能会议论文集 33, 2019, pp. 1044–1051.

[47] S. Zhou, J. Liu, D. Hou, X. Zhong, Y. Zhang, 基于改进的深度Q网络的自主渗透测试, 应用科学. 11 (2021) 8823.
[48] R. Gangupantulu, T. Cody, A. Rahm, C. Redino, R. Clark, P. Park, 使用攻击图与强化学习进行重要资产分析, 在: IEEE计算智能研讨会系列 (SSCI), 2021, pp. 1–6.

[49] C. Neal, H. Dagdougu, A. Lodi, J.M. Fernandez, 基于强化学习的微电网控制算法渗透测试, in: IEEE第11届年度计算与通信研讨会和会议（CCWC），2021年，pp. 0038–0044.

[50] E.R. Russo, A. Di Sorbo, C.A. Visaggio, G. Canflora, 支持专家在漏洞评估活动中的漏洞描述总结，J. Syst.Softw. 156 (2019) 84–99.

[51] M. Aota, H. Kanehara, M. Kubo, N. Murata, B. Sun, T. Takahashi, 使用机器学习从漏洞描述中自动分类漏洞, in: IEEE计算机与通信研讨会（ISCC），2020年，pp. 1–7.
[52] M. Vanamala, X. Yuan, K. Roy, 主题建模和常见漏洞和曝光数据库的分类, in: 国际人工智能、大数据、计算和数据通信系统会议 (icABCD),2020, pp. 1–5.

[53] G. Bakirtzis, B.J. Simon, A.G. Collins, C.H. Fleming, C.R. Elks, 数据驱动的设计阶段系统分析漏洞探索, IEEE Syst. J. 14 (2019) 4864–4873.
[54] A. Kuppa, L. Aouad, N.A. Le-Khac, 将CVE '与MITRE ATT &CK技术相关联, in: 第16届可用性、可靠性和安全性国际会议, 2021, pp. 1–12.
[55] S. Chatterjee, S. Thekdi, 一种迭代学习和推理方法来管理复杂系统的动态网络安全漏洞, Reliab. Eng. Syst. 193 (2020), 106664.
[56] Y. Jiang, Y. Atifl, 一种用于认知网络安全分析的选择性集成模型, J. Netw. Comput. Appl. 193 (2021), 103210.
[57] S. Samtani, S. Yu, H. Zhu, M. Patton, J. Matherly, H. Chen, 通过文本挖掘方法识别物联网中的SCADA系统及其漏洞, IEEE Intell. Syst. 33 (2) (2018) 63–73.

[58] J. Brown, T. Saha, N.K. Jha, GRAVITAS: 用于物联网聚合安全的图形化攻击向量, IEEE Trans. Emerg. 10 (3) (2022)1331–1348.

[59] P. Gao, F. Shao, X. Liu, X. Xiao, Z. Qin, F. Xu, P. Mittal, S.R. Kulkarn, D. Song, 利用网络威胁情报实现高效的网络威胁猎杀, in: IEEE 第37届国际数据工程大会 (ICDE), 2021, pp. 193–204.
[60] A. Nadeem, S. Verwer, S. Moskal, S.J. Yang, 使用S-PDFA进行基于警报的攻击图生成，IEEE Trans. Dependable and Secure Comput. 19 (2022) 731–746.
[61] H. Binyamini, R. Bitton, M. Inokuchi, T. Yagyu, Y. Elovici, A. Shabtai, 从安全漏洞描述中建模网络攻击技术的框架，第27届ACM SIGKDD知识发现与数据挖掘会议论文集，2021年，第2574–2583页.

[62] G. Falco, A. Viswanatha, C. Caldera, H. Shrobe, 用于智能城市的基于AI的自动化攻击规划器的主攻击方法论，IEEE Access 6 (2018) 48360–48373.

[63] H. Cam，模型引导的感染预测和使用上下文特定的网络安全观察进行主动防御，在：MILCOM 2019-2019 IEEE军事通信会议（MILCOM），2019年，第1–6页。
[64] A. Wollaber，J. Penna，B. Bleas，L. Shing，K. Alperin，S. Vilvovsky，P. Trepagnier，N. Wagner，L. Leonard，通过高性能计算实现主动网络态势感知，在：IEEE高性能极限计算会议（HPEC），2019年，第1–7页。

[65] J.C. Sancho，A. Caro，M. Avila，A. Bravo，基于安全事件管理的威胁分类和安全风险估计的新方法，Future Gener.Comput. Syst. 113 (2020) 488–505。

[66] A.A. Tubis, S. Werbińska-Wojciechowsk, M. G oralczyk, A. Wr oblewski, B. Zi ętek ,基于模糊理论的矿山不同自动化水平下的网络攻击风险分析方法，传感器2020年第20卷第24期7210页。

[67] Y. Qin, Y. Peng, K. Huang, C. Zhou, Y.C. Tian, 基于关联分析的工控系统网络安全风险评估，IEEE系统杂志2020年第15卷第1期1423–1432页。

[68] G. Falco, C. Caldera, H. Shrobe, SCADA系统的工控物联网网络安全风险建模，IEEE物联网杂志2018年第5卷第6期4486–4495页。
[69] M. Vega-Barbas, V.A. Villagr´a, F. Monje, R. Riesco, X. Larriva-Novo, J. Berrocal,基于本体的行政领域动态风险管理系统，Appl. Sci. 9 (21) (2019) 4547.

[70] M. Kalinin, V. Krundyshev, P. Zegzhda, 智能城市基础设施的网络安全风险评估，Machines 9 (4) (2021) 78.
[71] B. Biswas, A. Mukhopadhyay, S. Bhattacharjee, A. Kumar, D. Delen, 基于文本挖掘的网络风险评估和缓解框架，用于在线黑客论坛的关键分析，Decis. Support Syst. 152 (2022), 113651.
[72] N. Al-Hadhrami, M. Collinson, N. Oren, 一种主观网络方法用于网络安全风险评估，在：第13届国际信息与网络安全会议，2020年，pp. 1–8.

[73] M.S. Ansari, V. Barto˘s, B. Lee, 基于GRU的深度学习方法用于网络入侵警报预测, 未来计算系统. 128 (2022) 35–47.
[74] L. Wang, R. Jones, 大数据分析在网络安全中的应用：网络数据和入侵预测, 在: IEEE第10届普适计算、电子与移动通信会议(UEMCON), 2019, pp. 0105–0111.

[75] H. Al Najada, I. Mahgoub, I. Mohammed, 使用深度学习和分布式大数据处理的网络入侵预测和分类系统, 在: IEEE计算智能研讨会系列(SSCI), 2018, pp. 631–638.
[76] W.G. Mueller, A. Memory, K. Bartrem, 使用LSTMs从安全日志中预测网络入侵, 在: 国际部署机器学习用于安全防御研讨会, 2020, pp. 122–137.

[77] M. Rhode, P. Burnap, K. Jones, 使用循环神经网络进行早期恶意软件预测，计算机安全，2018年，578–594。
[78] I. Perera, J. Hwang, K. Bayas, B. Dorr, Y. Wilks, 通过公共文本分析和小型理论进行网络攻击预测，IEEE大数据国际会议（Big Data），2018年，3001–3010。

[79] E. Marin, M. Almukaynizi, P. Shakarian, 归纳和演绎推理在网络攻击预测中的应用，第10届计算与通信研讨会和会议（CCWC），2020年，0262–0268。

[80] N. Polatidis, E. Pimenidis, M. Pavlidis, S. Papastergiou, H. Mouratidis, 从产品推荐到网络攻击预测：生成攻击图并预测未来攻击，进化系统，第11卷第3期，2020年，479–490。
[81] L.P. Rees, J.K. Deane, T.R. Rakes, W.H. Baker, 决策支持网络安全风险规划, 决策支持系统. 51 (3) (2011) 493–505.
[82] J.A. Paul, X.J. Wang, 社会最优的网络安全IT投资, 决策支持系统. 122 (2019), 113069.
[83] J.A. Paul, M. Zhang, 网络安全风险规划的决策支持模型：一个特色是公司、政府和攻击者的两阶段随机规划框架, 欧洲运筹学杂志. 291 (1) (2021) 349–364.

[84] K. Zheng, L.A. Albert, J.R. Luedtke, E. Towle, 一个预算最大多重覆盖模型用于网络安全规划和管理, IISE Trans. 51 (12)(2019) 1303–1317.

[85] A. Yeboah-Oflori, S. Islam, S.W. Lee, Z.U. Shamszaman, K. Muhammad, M. Altafl, M.S. Al-Rakhami, 用于改善网络供应链安全的网络威胁预测分析, IEEE Access 9 (2021) 94318–94337.
[86] T. Sawik, 用于工业4.0供应链中最佳网络安全投资的线性模型, Int. J. Prod. Res. 60 (4) (2022) 1368–1385.
[87] T. Sawik, B. Sawik, 使用具有最大网络安全价值的安全控制组合的网络安全投资初步研究, Int. J. Prod. Res. 60 (21)(2022) 6556–6572.

[88] T. Sawik, 在直接和间接网络风险下平衡供应链中的网络安全, Int. J. Prod. Res. 60 (2) (2022) 766–782.
[89] S. Rahman, N.U. Hossain, K. Govindan, F. Nur, M. Bappy, 评估增材制造供应链的网络弹性，利用数据融合技术：一个生成供应链网络弹性指数的模型，CIRP J. Manufl. Sci. Technol. 35 (2021) 911–928.

[90] A.I. Siam, A. Sedik, W. El-Shaflai, A.A. Elazm, N.A. El-Bahnasawy, G.M. El Banby, 基于卷积神经网络的生物信号分类用于人体识别，Int. J. Commun. Syst. 34 (7) (2021) 1–22.

[91] J.M. Jorquera Valero, P.M. Sanchez Sánchez, L. Fernández Maimo, A. Huertas Celdrán, M. Arjona Fernández, S. De Los Santos Vílchez, G. Martínez Pérez, 通过智能和自适应的连续认证系统提高移动设备的安全性和用户体验，Sensors 18 (11) (2018) 3769.
[92] P.M. Sánchez, A. Huertas Celdran, L. Fernández Maimo, G. Martínez Pérez, G. Wang, 通过智能和多设备的持续认证系统保护智能办公室, 在: 国际智能城市和信息化会议, 2019, pp. 73-85.

[93] A.G. Martín, M. Beltran, A. Fernández-Isabel, I.M. de Diego, 在身份联合中检测用户行为异常的方法, 计算机安全 108 (2021), 102356.

[94] H. Alobaidi, N. Clarke, F. Li, A. Alruban, 现实世界中基于智能手机的步态识别, 计算机安全 113 (2022), 102557.
[95] K.A. Rahman, D. Neupane, A. Zaiter, M.S. Hossain, 使用选择的单词击键动态进行网络用户身份验证, 在: 第18届IEEE国际机器学习与应用会议(ICMLA), 2019, pp. 1130–1135.
[96] A. Shaout, N. Schmidt, 使用模糊逻辑的击键识别器以提高密码安全性, 在: 第21届国际阿拉伯信息技术会议(ACIT), 2020, pp. 1–8.

[97] A. Hafleez, K. Topolovec, S. Awad, 通过参数信号建模和人工神经网络对ECU指纹进行防止欺骗攻击的车辆安全性研究, 在: 第15届国际计算机工程会议(ICENCO), 2019, pp. 29–38.

[98] G. Baldini, R. Giuliani, M. Gemo, F. Dimc, 将传感器身份验证与固有物理特征应用于车辆安全性, 计算机与电子工程, 91 (2021), 107053.

[99] Y. Cui, F. Bai, R. Yan, T. Saha, R.K. Ko, Y. Liu, 微电网网络安全中分布式同步相量测量装置的源身份验证, IEEE Trans. SmartGrid 12 (5) (2021) 4577–4580.

[100] M. Benedetti, M. Mori, 关于在RBAC维护中使用Max-SAT和PDDL的研究, Cybersecurity 2 (1) (2019) 1–25.

[101] M. Abolflathi, Z. Raghebi, H. Jaflarian, F. Banaei-Kashani, 一种适用于大型组织的可扩展角色挖掘方法, in: Proceedings ofl the 2021 ACM Workshop onSecurity and Privacy Analytics, 2021, pp. 45–54.

[102] S.S. Chukkapalli, S.B. Aziz, N. Alotaibi, S. Mittal, M. Gupta, M. Abdelsalam, 基于本体论的人工智能和访问控制系统用于智能渔业, in: Proceedings ofl the 2021 ACM Workshop on Secure and Trustworthy Cyber-Physical Systems, 2021, pp. 59–68.

[103] B. Leander, A. Čau ševi´c, H. Hansson, T. Lindstrom, 智能制造系统的访问控制, in: 欧洲软件架构会议, 2020,pp. 463–476.

[104] Z. Tan, R. Beuran, S. Hasegawa, W. Jiang, M. Zhao, Y. Tan, 使用链接开放数据数据集的自适应安全意识培训, Educ. Infl. Technol. 25 (6)(2020) 5235–5259.

[105] F. Nembhard, M. Carvalho, T. Eskridge, 提高程序安全性的混合方法, in: IEEE计算智能研讨会 (SSCI), 2017,pp. 1–8.

[106] T.Espinha Gasiba, U. Lechner, M. Pinto-Albuquerque, Siflu-一种具有挑战评估和智能教练的网络安全意识平台,Cybersecurity 3 (1) (2020) 1–23.

[107] D.C. Le, N. Zincir-Heywood, 使用无监督集成进行内部威胁的异常检测, IEEE Trans. Netw. Service Manag. 18 (2) (2021)1152–1164.

[108] J. Kim, M. Park, H. Kim, S. Cho, P. Kang, 基于用户行为建模和异常检测算法的内部威胁检测, Appl. Sci. 9 (19) (2019)4018.

[109] T. Al-Shehari, R.A. Alsowail, 使用独热编码、合成少数类过采样和机器学习技术进行内部数据泄露检测, Entropy 23 (10) (2021) 1258.

[110] K. Alzhrani, E.M. Rudd, T.E. Boult, C.E. Chow, 自动化的大文本安全分类, in: IEEE Conflerence on Intelligence and Security Inflormatics (ISI),2016, pp. 103–108.

[111] Y. Guo, J. Liu, W. Tang, C. Huang, Exsense: 从非结构化数据中提取敏感信息, 计算机安全. 102 (2021), 102156.

[112] H. Li, J. Wu, H. Xu, G. Li, M. Guizani, 可解释的智能驱动的高级持续性威胁防御机制: 一种联合边缘博弈和人工智能方法, IEEE可靠安全计算. 19 (2) (2022) 757–775.

[113] A.A. Alghamdi, G. Reger, 通过无监督学习提取多阶段威胁行为的模式, 在: 国际网络态势感知、数据分析和评估会议(CyberSA), 2020, pp. 1–8.

[114] L. Gallo, A. Maiello, A. Botta, G. Ventre, 在一家大公司的反钓鱼小组工作两年, 计算机安全. 105 (2021), 102259.

[115] D. Wu, W. Shi, X. Ma, 一种新型实时反垃圾邮件框架, ACM Trans. Internet Technol. (TOIT) 21 (4) (2021) 1–27.

[116] E.S. Gualberto, R.T. De Sousa, T.P. Vieira, J.P. Da Costa, C.G. Duque, 答案在文本中：基于特征工程的多阶段网络钓鱼检测方法, IEEE Access 8 (2020) 223529–223547.

[117] M. Nguyen, T. Nguyen, T.H. Nguyen, 一种具有分层lstms和监督注意力的深度学习模型用于反网络钓鱼, 在：第一届反网络钓鱼共享任务在第四届ACM IWSPA上的试点，2018年，第29–38页.

[118] D. Cohen, O. Naim, E. Toch, I. Ben-Gal, 通过设计属性学习进行网站分类计算机安全 107 (2021), 102312.

[119] C. Marques, S. Malta, J.P. Magalh˜aes, 恶意域名检测的DNS数据集, Data Br 38 (2021), 107342.

[120] B. Yu, J. Pan, D. Gray, J. Hu, C. Choudhary, A.C. Nascimento, M. De Cock, 弱监督深度学习用于域名生成算法检测, IEEEAccess 7 (2019) 51542–51556.

[121] J. Spaulding, A. Mohaisen, 使用D-FENS防御物联网中的恶意域名名称, in: IEEE/ACM边缘计算研讨会 (SEC), 2018, pp. 387–392.

[122] P.L. Indrasiri, M.N. Halgamuge, A. Mohammad, 强大的集成机器学习模型用于过滤钓鱼URL：可扩展的随机梯度堆叠投票分类器 (ERG-SVC), IEEE Access 9 (2021) 150142 –150161.

[123] R. Vinayakumar, K.P. Soman, P. Poornachandran, 评估深度学习方法以表征和分类恶意URL's, J. Intell. Fuzzy Syst. 34(3) (2018) 1333–1343.

[124] W. Li, J. Jin, J.H. Lee, 物联网网络安全中的僵尸网络域名分析, IEEE Access 7 (2019) 94658–94665.

[125] B. Alotaibi, M. Alotaibi, 网络钓鱼的共识和多数投票特征选择方法和检测技术, J. Ambient. Intell. Humaniz. Comput. 12 (1) (2021) 717–727.

[126] Y. Qin, B. Hoflflmann, D.J. Lilja, Hyperprotect: 使用智能调度提高动态备份系统的性能, in: IEEE第37届国际性能计算与通信会议（IPCCC），2018年,pp. 1–8.

[127] P.M. Van de Ven, B. Zhang, A. Schorg endorfler, 分布式备份调度：建模与优化，在：IEEE INFOCOM 2014-IEEE计算机通信会议，2014年，第1644–1652页。

[128] Z. Zeng, Z. Yang, D. Huang, C.J. Chung, LICALITY–Likelihood and criticality：逻辑推理和深度学习的漏洞风险优先级排序，IEEETrans. Netw. Service Manag. 19 (2) (2021) 1746–1760页。

[129] J. Yin, M. Tang, J. Cao, H. Wang, 将迁移学习应用于网络安全：通过描述预测漏洞的可利用性，知识基础系统。210(2020), 106529页。

[130] J. Yin, M. Tang, J. Cao, H. Wang, M. You, 一种实时动态概念自适应学习算法用于可利用性预测，神经计算。472 (2022)252–265页。

[131] T. Bai, H. Bian, M.A. Salahuddin, A. Abou Daya, N. Limam, R. Boutaba, 使用机器学习进行基于RDP的横向移动检测，计算机通信，165 (2021) 9–19.

[132] N. Aflzaliseresht, Y. Miao, S. Michalska, Q. Liu, H. Wang, 从日志到故事：面向人的数据挖掘用于网络威胁情报，IEEE Access 8 (2020) 19089–19099.

[133] G. De la Torre-Abaitua, L.F. Lago-Fern´andez, D. Arroyo, 一种基于压缩的文本数据异常检测方法，熵，23 (5) (2021) 618.

[134] T. Eljasik-Swoboda, W. Demuth, 利用聚类和自然语言处理克服日志管理中的多样性问题，ICAART，2020，pp. 281–288.

[135] D. Sisiaridis, O. Markowitch, 减少大数据安全分析中特征提取和特征选择的数据复杂性, 在：2018年第一届数据智能与安全国际会议(ICDIS), pp. 43–48.

[136] P.F. De Araujo-Filho, A.J. Pinheiro, G. Kaddoum, D.R. Campelo, F.L. Soares, 一种用于CAN的高效入侵预防系统：通过低成本平台阻碍网络攻击, IEEE Access 9 (2021) 166855–166869.

[137] C. Constantinides, S. Shiaeles, B. Ghita, N. Kolokotronis, 一种新颖的在线增量学习入侵预防系统, 在：2019年第10届IFIP新技术、移动性和安全国际会议(NTMS), pp. 1–6.

[138] S.M. de Lima, H.K. Silva, J.H. Luz, H.J. Lima, S.L. Silva, A. de Andrade, A.M. daSilva, 基于人工智能的防病毒软件以预防性地检测恶意软件, Prog. Artifl. Intell. 10 (1) (2021) 1–22.

[139] P. Marques, M. Rhode, I. Gashi, 不浪费：利用超参数搜索得到的多样化神经网络来提高恶意软件检测能力, Comput. Secur. 108(2021), 102339.

[140] P. Karuna, H. Purohit, S. Jajodia, R. Ganesan, O. Uzuner, 通过操纵文本可理解性进行网络欺骗的虚假文档生成, IEEE Syst.J. 15 (1) (2020) 835–845.

[141] O. Ajayi, A. Gangopadhyay, DAHID: 领域自适应主机入侵检测, in: IEEE International Conflerence on Cyber Security and Resilience(CSR), 2021, pp. 467–472.

[142] G. Granato, A. Martino, L. Baldini, A. Rizzi, 通过模块化和优化的分类器集成在Wi-Fi网络中进行入侵检测，在：IJCCI，2020年，第412–422页。

[143] Z. Li, A.L. Rios, L. Trajkovi´c, 用于检测通信网络中异常和入侵的机器学习方法，IEEE J. Sel. Areas Commun. 39 (7) (2021)2254–2264页。

[144] M. Almiani, A. AbuGhazleh, Y. Jararweh, A. Razaque, 使用深度卡尔曼反向传播神经网络在5G-启用的物联网网络中进行DDoS检测，Int. J. Mach. Learn. Cybern. 12 (11) (2021) 3337–3349页。

[145] A. Corsini, S.J. Yang, G. Apruzzese, 对于网络入侵检测的顺序机器学习评估, 在：第16届可用性、可靠性和安全性国际会议, 2021, 第1–10页.

[146] M. Chora´s, M. Pawlicki, 基于优化人工神经网络的入侵检测方法, Neurocomputing 452 (2021) 705–715.

[147] K.S. Kumar, S.A. Nair, D.G. Roy, B. Rajalingam, R.S. Kumar, 使用联邦机器学习的安全和隐私感知的人工入侵检测系统, Comput. Electr. Eng. (96)107440.

[148] J.C. Wu, S. Lu, C.S. Fuh, T.L. Liu, 通过新颖性归一化进行单类异常检测, Comput. Vis. Image. Underst. 210 (2021), 103226.

[149] L. Fernández Maimo, A. Huertas Celdrán, M. Gil Pérez, F.J. García Clemente, G. Martínez Pérez, 5G网络基于深度学习的异常检测系统的动态管理，J. Ambient. Intell. Humaniz. Comput. 10 (8) (2019) 3083–3097.

[150] D.C. Le, A.N. Zincir-Heywood, M.I. Heywood, 基于网络流量的僵尸网络行为检测的数据分析，IEEE计算智能研讨会系列(SSCI)，2016，pp. 1–7.

[151] D. Saveetha, G. Maragatham, 基于区块链的深度学习入侵检测模型设计，Pattern Recognit. Lett. 153 (2022) 24–28.

[152] M. Al-Hawawreh, E. Sitnikova, F. den Hartog, 一种用于棕地工业物联网边缘系统的高效入侵检测模型, 第三届大数据与物联网国际会议论文集, 2019, 页码 83-87.

[153] J. Vavra, M. Hromada, L. Lukats, J. Dworzecki, 基于机器学习算法的工业控制环境自适应异常检测系统, 重要基础设施国际期刊, 34卷 (2021), 100446.

[154] Y. Zhang, L. Wang, W. Sun, R.C. Green II, M. Alam, 智能电网多层网络架构中的分布式入侵检测系统, IEEE智能电网期刊, 2卷4期 (2011) 796-808.

[155] R. Blanco, P. Malagon, S. Briongos, J.M. Moya, 使用高斯混合概率模型进行异常检测以实现入侵检测系统, 在：混合人工智能系统国际会议, 2019, pp. 648-659.

[156] G.N. Nguyen, N.H. Le, M. Viet, K. Elhoseny, B.B. Shankar, A.A. Gupta, Abd El-Latifl, 使用深度置信网络和ResNet模型的安全区块链启用的医疗保健网络物理系统, J. Parallel. Distrib. Comput. 153 (2021) 150-160.

[157] A. Alhowaide, I. Alsmadi, J. Tang, 用于物联网入侵检测的集成检测模型, 物联网 16 (2021), 100435.

[158] A. Binbusayyis, T. Vaiyapuri, 结合卷积自编码器和单类支持向量机的无监督深度学习方法用于网络入侵检测, Appl. Intell. 51 (10) (2021) 7094-7108.

[159] V. Herrera-Semenets, L. Bustio-Martínez, R. Hernández-Leon, J. van den Berg ,一种用于高效入侵检测的多度量特征选择算法，知识基础系统。227 (2021), 107264。

[160] M. Rashid, J. Kamruzzaman, T. Imam, S. Wibowo, S. Gordon, 一种基于树的堆叠集成技术与特征选择的网络入侵检测方法，应用智能。52 (2022) 9768–9781。

[161] V. Dutta, M. Chora´s, R. Kozik, M. Pawlicki, 一种改进网络入侵检测分类效果的混合模型，计算信息系统安全的智能会议，2019年，pp. 405–414。

[162] S.I. P´erez, S. Moral-Rubio, R. Criado, 一种结合多重网络和时间序列属性的新方法：在网络安全中构建入侵检测系统（IDS），混沌，孤立子和分形 150 (2021), 111143.

[163] P. Singh, A. Pankaj, R. Mitra, Edge-detect: 使用深度神经网络的以边为中心的网络入侵检测, 在IEEE第18届年度消费者通信与网络会议上, 2021, pp. 1–6.

[164] M. Catillo, A. Pecchia, U. Villano, AutoLog: 通过系统日志的深度自动编码进行异常检测, 专家系统与应用 191 (2022), 116263.

[165] R. Zhao, Y. Yin, Y. Shi, Z. Xue, 基于联邦学习辅助的长短期记忆的智能入侵检测, 物理通信 42 (2020), 101157.

[166] D. Nedeljkovic, Z. Jakovljevic, 基于CNN的工业控制系统网络攻击检测算法开发方法, 计算机安全, 2022, 102585.

[167] M. Elnour, N. Meskin, K.M. Khan, 基于1D-CNN和孤立森林的工业控制系统混合攻击检测框架, IEEE控制技术与应用会议(CCTA), 2020, pp. 877–884.

[168] H. Liu, C. Zhong, A. Alnusair, S.R. Islam, FAIXID: 一种通过数据清洗技术提高入侵检测结果的AI可解释性框架, 网络系统管理杂志, 2021, 29 (4), 1–30.

[169] J.M. Vidal, M.A. Monge, S.M. Monterrubio, EsPADA: 针对恶意软件检测的增强负载分析器, 能抵御对抗性威胁, 未来计算机系统, 2020, 104, 159–173.

[170] S. Latifl, Z.E. Huma, S.S. Jamal, F. Ahmed, J. Ahmad, A. Zahid, K. Dashtipour, M.U. Afltab, M. Ahmad, Q.H. Abbasi, 使用密集随机神经网络的物联网入侵检测框架, IEEE Trans. Industr. Infform. 18(9) (2021) 6435–6444.

[171] J.L. Leevy, J. Hancock, R. Zuech, T.M. Khoshgofltaar, 使用LightGBM和XGBoost学习器检测网络安全攻击, 在:IEEE第二届认知机器智能国际会议(CogMI), 2020, pp. 190–197.

[172] R. Abdulhammed, M. Faezipour, A. Abuzneid, A. Alessa, 使用减少属性集的机器学习分类增强无线入侵检测, 在: 第14届国际无线通信 &移动计算会议(IWCMC), 2018, pp. 524–529.

[173] C. Iwendi, S.U. Rehman, A.R. Javed, S. Khan, G. Srivastava, 使用人工智能架构实现物联网的可持续安全性，ACM Trans.Internet Technol. 21 (3) (2021) 1–22.

[174] P. Toupas, D. Chamou, K.M. Giannoutakis, A. Drosou, D. Tzovaras, 基于深度神经网络的多类分类入侵检测系统，第18届IEEE国际机器学习与应用会议(ICMLA)，2019年，第1253–1258页。

[175] L. D'hooge, M. Verkerken, T. Wauters, B. Volckaert, F. De Turck, 数据高效入侵检测建模的分层特征块排序，Comput.Netw. 201 (2021)，108613。

[176] S. Huang, K. Lei, IGAN-IDS:一种面向自组织网络入侵检测系统的不平衡生成对抗网络，Ad Hoc Netw 105 (2020)，102177。

[177] N. Gupta, V. Jindal, P. Bedi, CSE-IDS:使用成本敏感的深度学习和集成算法处理网络入侵检测系统中的类别不平衡，Comput. Secur. 112 (2022)，102499。

[178] S.S. Jagtap, S.S. VS, V. Subramaniyaswamy,一种基于超图的Kohonen映射用于检测网络–物理系统流量中的入侵，Future Gener.Comput. Syst. 119 (2021) 84–109。

[179] M. Asifl, S. Abbas, M.A. Khan, A. Ftima, M.A. Khan, S.W. Lee,基于MapReduce的智能入侵检测模型，使用机器学习技术，J. King Saud Univ. - Comput. Infl. Sci. 34 (2022) 9723–9731。

[180] J. Liu, B. Kantarci, C. Adams, 基于机器学习的入侵检测，针对Contiki-NG-based IoT网络暴露在NSL-KDD数据集中，在第二届ACM无线安全和机器学习研讨会上，2020年，第25–30页。

[181] R. Blanco, P. Malagon, J.J. Cilla, J.M. Moya, 使用遗传算法调整的CNN多类网络攻击分类器，第28届国际功率和时序建模、优化和仿真研讨会（PATMOS），2018年，第177–182页。

[182] M. Pawlicki, R. Kozik, M. Chora´s, 用于网络入侵检测的人工神经网络超参数优化，在智能计算国际会议上，2019年，第749–760页。

[183] M. Shafliq, Z. Tian, A.K. Bashir, X. Du, M. Guizani, 使用基于包装器的特征选择机制进行物联网恶意流量识别, 计算机安全, 94 (2020), 101863.

[184] J.W. Mikhail, J.M. Fossaceca, R. Iammartino, 一种基于敏感性加权二值化的多域网络入侵检测半提升嵌套模型, ACM智能系统与技术交易, 10 (3) (2019) 1–27.

[185] M. Basnet, M.H. Ali, 基于深度学习的电动汽车充电站入侵检测系统, 在: 第二届智能电力 &互联网能源系统国际会议(SPIES), 2020, pp. 408–413.

[186] A.F. Diallo, P. Patras, 自适应聚类的网络边缘恶意流量分类, 在: IEEE INFOCOM 2021-IEEE计算机通信会议, 2021, pp. 1–10.

[187] N. Gupta, V. Jindal, P. Bedi, LIO-IDS: 使用LSTM和改进的一对一技术处理类别不平衡的入侵检测系统, 计算机网络. 192(2021), 108076.

[188] I. Ullah, Q.H. Mahmoud, 一种用于SCADA网络异常基于入侵检测的混合模型, 在IEEE国际大数据会议上 (Big Data), 2017,pp. 2160–2167.

[189] G. Li, Y. Shen, P. Zhao, X. Lu, J. Liu, Y. Liu, S.C. Hoi, 使用在线学习算法在工业控制系统中检测网络攻击, 神经计算. 364(2019) 338–348.

[190] J. Zhang, F. Li, F. Ye, 一种基于集成的网络入侵检测方案, 使用贝叶斯深度学习, 在IEEE国际通信会议上 (ICC), 2020, pp. 1–6.

[191] W. Zong, Y.W. Chow, W. Susilo, 交互式三维可视化的网络入侵检测数据用于机器学习, 未来计算机系统 102 (2020) 292–306.

[192] C. Ieracitano, A. Adeel, F.C. Morabito, A. Hussain, 一种新的统计分析和自编码器驱动的智能入侵检测方法, 神经计算 387 (2020) 51–62.

[193] Q. Liu, D. Wang, Y. Jia, S. Luo, C. Wang, 一种基于多任务的深度学习方法用于入侵检测, 基于知识的系统 238 (2022), 107852.

[194] D. Xuan, H. Hu, B. Wang, B. Liu, 基于RF-SVM模型和特征选择优化的入侵检测系统, 在: 国际通信、计算、网络安全和信息学会议 (CCCI), 2021,pp. 1–5.

[195] P. Radoglou-Grammatikis, P. Sarigiannidis, E. Iturbe, E. Rios, S. Martinez, A. Sarigiannidis, G. Efltathopoulos, Y. Spyridis, A. Sesis, N. Vakakis, D. Tzovaras, Spear siem: a security inflormation and event management system flor the smart grid, Comput. Netw. 193 (2021), 108008.

[196] A. Fausto, G.B. Gaggero, F. Patrone, P. Girdinio, M. Marchese, Toward the integration ofl cyber and physical security monitoring systems flor critical inflrastructures, Sensors 21 (21) (2021) 6970.

[197] H.A. Kodituwakku, A. Keller, J. Gregor, InSight2: a modular visual analysis platflorm flor network situational awareness in large-scale networks, Electronics (Basel) 9 (10) (2020) 1747.

[198] Y. Nikoloudakis, I. Keflaloukos, S. Klados, S. Panagiotakis, E. Pallis, C. Skianis, E.K. Markakis, 为网络安全而基于机器学习的情境感知框架的研究：一个SDN实现, Sensors 21 (14) (2021) 4939.

[199] F. Zhang, H.A. Kodituwakku, J.W. Hines, J. Coble, 基于网络、系统和过程数据的多层数据驱动的工业控制系统的网络攻击检测系统, IEEE Trans. Industr. Infform. 15 (7) (2019) 4362–4369.

[200] D.L. Marino, C.S. Wickramasinghe, B. Tsouvalas, C. Rieger, M. Manic, 基于数据驱动的综合系统健康监测中的网络和物理异常相关性, IEEE Access 9 (2021) 163138–163150.

[201] K. Al-Rowaily, M. Abulaish, N.A. Haldar, M. Al-Rubaian, BiSAL–一个用于分析暗网论坛的双语情感分析词典，用于网络安全，数字调查14 (2015) 53–62.

[202] A. Deb, K. Lerman, E. Ferrara, 通过利用黑客情绪来预测网络事件，信息9 (11) (2018) 280.

[203] S. Ishikawa, S. Ozawa, T. Ban, 用于暗网流量特征和扫描攻击聚类的端口嵌入，国际神经信息处理会议，2020年，pp. 593–603.

[204] G. Pantelis, P. Petrou, S. Karagiorgou, D. Alexandrou, 通过识别暗网上的数据泄露、盗窃凭证和非法活动来加强中小企业的威胁情报和意识，第16届可用性、可靠性和安全性国际会议，2021年，pp. 1–7.

[205] M. Sch¨afler, M. Fuchs, M. Strohmeier, M. Engel, M. Liechti, V. Lenders,Black Widow: 监控暗网以获取网络安全信息, 第11届国际网络冲突会议(CyCon), 2019, pp. 1–21.

[206] Z. Fang, X. Zhao, Q. Wei, G. Chen, Y. Zhang, C. Xing, W. Li, H. Chen, 在中国黑客社区中探索关键黑客和网络安全威胁, IEEE智能与安全信息学会议(ISI), 2016, pp. 13–18.

[207] S.Y. Huang, T. Ban, 一种基于主题的无监督学习方法用于在线地下市场探索, 第18届IEEE信任、安全与隐私计算与通信国际会议/第13届IEEE大数据科学与工程国际会议(TrustCom/BigDataSE), 2019, pp. 208–215.

[208] G. Kim, C. Lee, J. Jo, H. Lim, 使用深度Bi-LSTM-CRF网络自动提取网络威胁的命名实体, Int. J. Mach. Learn. Cybern. 11 (10)(2020) 2341–2355.

[209] I. Sarhan, M. Spruit, Open-cykg: 一个开放的网络威胁情报知识图谱, Knowl. Based Syst. 233 (2021), 107524.

[210] F. Alves, A. Bettini, P.M. Ferreira, A. Bessani, 用于网络安全威胁意识的推文处理, Infl. Syst. 95 (2021), 101586.

[211] N. Dionísio, F. Alves, P.M. Ferreira, A. Bessani, 使用深度神经网络从Twitter中检测网络威胁, in: International Joint Conflerence on NeuralNetworks (IJCNN), 2019, pp. 1–8.
[212] J.R. Saura, D. Palacios-Marqu´es, D. Ribeiro-Soriano, 使用数据挖掘技术在Twitter中探索智能生活环境中的安全问题，计算机通信。179（2021）285–295。

[213] T.M. Georgescu, B. Iancu, M. Zurini, 基于命名实体识别的物联网网络安全情况诊断自动化系统，传感器。19（15）（2019）3380。

[214] J. Sleeman, T. Finin, M. Halem, 通过动态主题建模了解网络安全威胁趋势，Big Data 4（2021）。
[215] T. Sun, P. Yang, M. Li, S. Liao, 基于多源信息融合的网络威胁情报记录自动生成方法，未来互联网13（2）（2021）40。

[216] A. Sapienza, S.K. Ernala, A. Bessi, K. Lerman, E. Ferrara, 发现：挖掘在线聊天以寻找新兴的网络威胁, 在：同行会议论文集, 2018, pp. 983–990.

[217] C.E. Tsai, C.L. Yang, C.K. Chen, A.N.T. CTI, 寻找中国威胁情报, 在：IEEE国际大数据会议 (Big Data), 2020, pp. 1847–1852.

[218] P. Ranade, S. Mittal, A. Joshi, K. Joshi, 使用深度神经网络翻译多语言威胁情报, 在：IEEE智能与安全信息学国际会议 (ISI), 2018, pp. 238–243.

[219] V.A. Memos, K.E. Psannis, 基于人工智能的蜜罐用于增强物联网僵尸网络检测, 在：第三届通信工程世界研讨会 (WSCE), 2020, pp. 64–68.

[220] P. Chatziadam, I.G. Askoxylakis, A. Fragkiadakis, 一个用于早期警告入侵检测的网络望远镜, 在：国际人类信息安全、隐私和信任会议, 2014, pp. 11–22.

[221] H.K. Kim, K.H. Im, S.C. Park, 应用CBR和协作响应的计算机安全事件响应决策支持系统, 专家系统应用, 37 (1) (2010)852–870.

[222] F. Jiang, T. Gu, L. Chang, Z. Xu, 基于描述逻辑的网络安全应急响应案例检索, 在：国际智能信息处理会议, 2014, pp. 284–293.

[223] R.C. Nunes, M. Colom´e, F.A. Barcelos, M. Garbin, G.B. Paulus, L.A. Silva, 基于案例推理的网络安全事件记录和解决方法, 软件工程与知识工程国际期刊, 11 (12) (2019) 1607–1627.
[224] I. Kraeva, G. Yakhyaeva, 应用度量学习进行安全事件播放书推荐, 在：IEEE第22届青年专业人员电子器件和材料国际会议（EDM），2021年，第475–479页。

[225] L. Ping, Y. Haifleng, M. Guoqing, 基于CBR和本体论的事件响应决策支持系统, 在：国际计算机应用和系统建模会议（ICCASM 2010），2010年，第311–337页。

[226] A. Shah, R. Ganesan, S. Jajodia, H. Cam, 在逆境条件下动态优化CSOC的运营效能水平, ACM智能系统技术交易. 9 (5) (2018) 1–20.

[227] Y. Lin, H. Wang, B. Yang, M. Liu, Y. Li, Y. Zhang, 基于多Agent系统的社区网络威胁情报黑板共享机制, 在：机器学习与网络安全国际会议，2019年，第253–270页。

[228] L. Thomas, A. Vaughan, Z. Courtney, C. Zhong, A. Alnusair, 通过可视化分析推理过程支持网络安全分析师之间的协作，在：IEEE国际多媒体与博览会研讨会（ICMEW），2018年，第1-6页。

[229] N. DeCastro-García, A.L. Muñoz Castañeda, M. Fernández-Rodríguez, 机器学习用于自动分配网络安全事件严重程度，计算数学方法医学2（1）（2020）e1072。
[230] M. Husak, T. Bajtoš, J. Kašpar, E. Bou-Harb, P. Celeda, 预测性网络态势感知和个性化黑名单：一种顺序规则挖掘方法，ACM管理信息系统交易11（4）（2020）1-6。
[231] F. Manganiello, M. Marchetti, M. Colajanni, 多步攻击检测和入侵检测系统中的警报相关性, 在：国际信息安全与保障会议, 2011年, 第101–110页.

[232] A. Dey, E. Totel, S. Navers, 利用自动编码器进行异构安全事件优先级排序, 在：互联网和系统风险与安全国际会议, 2020年, 第164–180页.

[233] J.Q. Chen, 上下文绑定的智能定位, 在：未来技术会议 (FTC), 2016年, 第1040–1046页.
[234] H. Studiawan, F. Sohel, 基于深度自动编码器的法医时间线异常检测, 信息安全应用杂志 63 (2021), 103002.
[235] F. Amato, A. Castiglione, G. Cozzolino, F. Narducci, 基于语义的数字取证分析方法, J. Parallel Distrib. Comput. 138 (2020)172–177.

[236] A. Nisioti, G. Loukas, A. Laszka, E. Panaousis, 数据驱动的决策支持优化网络取证调查, IEEE Trans. Infl. Forensics Secur. 16(2021) 2397–2412.

[237] J. Sakhnini, H. Karimipour, A. Dehghantanha, R.M. Parizi, 基于集成深度学习的网络–物理网格中物理层攻击识别和定位方法, Phys. Commun. 47 (2021), 101394.

[238] L. Fernandez Maimo, A. Huertas Celdran, A.L. Perales Gomez, F.J. GarciaClemente, J. Weimer, I. Lee, 综合临床环境中智能动态勒索软件传播检测与缓解, Sensors 19 (5)(2019) 1114.

[239] P. Nespoli, F.G. Marmol, J.M. Vidal, 一种仿生反击网络攻击的方法：基于人工智能的最佳对策选择, IEEE Access 9 (2021) 60971-60996.

[240] M. Husak, L. Sadlek, S. Spaček, M. Laštovička, M. Javorník, J. Komarková, CRUSOE：一种用于网络态势感知和决策支持的工具集，在事件处理中的应用, Comput. Secur. 115 (2022), 102609.
[241] M. Husák, 面向处理勒索软件和类似事件的数据驱动推荐系统，IEEE国际智能与安全信息学会议(ISI)，2021，第1-6页.

[242] A. Piplai, S. Mittal, A. Joshi, T. Finin, J. Holt, R. Zak, 从恶意软件行动报告中创建网络安全知识图谱, IEEE Access 8 (2020)211691–211703.

[243] B. Woods, S.J. Perl, B. Lindauer, 高效协作信息发现的数据挖掘, in: Proceedings ofl the 2nd ACM Workshop on InflormationSharing and Collaborative Security, 2015, pp. 3–12.
[244] S. Peng, A. Zhou, S. Liao, L. Liu, 基于条件共现度的威胁行为提取方法, in: 7th International Conflerence onInflormation Science and Control Engineering (ICISCE), 2020, pp. 1633–1637.
[245] B.S. Meyers, A. Meneely, 利用自然语言词嵌入进行漏洞关系的自动化事后分析, Procedia. Comput. Sci.(2021) 953–958.

[246] M.V. Carriegos, A.L. Castañeda, M.T. Trobajo, D.A. De Zaballa, 对网络安全事件报告的聚合和预测, IEEE Access 9 (2021) 102636–102648.

[247] Symantec, 互联网安全威胁报告, 检索自https://www.symantec.com/content/dam/symantec/docs/reports/istr-24-2019-en.pdfl (2019).