基于大语言模型的网络威胁情报知识图谱构建技术研究

赖清楠，金建栋，周昌令
（北京大学计算中心，北京 100871）

摘　要：随着网络威胁的复杂性和精细度不断增加，将网络威胁情报整合到网络安全措施中变得至关重要。设计了一个基于大语言模型的网络威胁情报知识图谱构建框架AutoCTI2KG，通过指令提示和上下文学习，自动从网络威胁情报中生成网络安全知识图谱和攻击知识图谱，并提供可操作的防护建议。实验结果表明，所提出的框架在网络安全知识图谱和攻击知识图谱构建方面表现出色，F1值在0.90左右，展示了大语言模型在网络安全领域知识图谱构建的潜力。所提出的框架不仅推进了网络安全知识图谱构建的前沿技术，还为网络安全专业人员提供了一个实用工具，以更好地理解和降低网络风险。

关键词：知识图谱；大语言模型；威胁情报；网络安全；人工智能
中图分类号：TP181
文献标志码：A
DOI：10.11959/j.issn.1000-436x.2024225

Research on knowledge graph construction technology for cyber threat intelligence based on large language models

LAI Qingnan, JIN Jiandong, ZHOU Changling
Computing Center, Peking University, Beijing 100871, China

Abstract: As the complexity and sophistication of cyber threats continue to increase, integrating cyber threat intelligence into cybersecurity measures has become crucial. A framework called AutoCTI2KG was proposed, which was based on large language models for constructing cyber threat intelligence knowledge graphs. Through instruction prompts and context learning, AutoCTI2KG automatically generated cybersecurity and attack knowledge graphs from cyber threat intelligence and provided actionable defense recommendations. Experimental results show that the proposed framework performs excellently in constructing cybersecurity and attack knowledge graphs, with F1 scores around 0.90, demonstrating the potential of large language models in knowledge graph construction in the cybersecurity domain. This work not only advances the frontier of cybersecurity knowledge graph construction but also provides a practical tool for cybersecurity professionals to better understand and mitigate cyber risks.

Keywords: knowledge graph, large language model, threat intelligence, cybersecurity, artificial intelligence

0　引言

全维度、多视角地感知网络空间威胁，特别是智能化、系统性地认知高级可持续威胁（APT, advanced persistent threat）攻击关联的战术、技术、漏洞及产品等，有助于提升国家及企业对网络威胁的科学防御能力[1]。在此背景下，构建网络安全知识图谱成为提升网络安全防御能力的关键手段之一，网络安全知识图谱通过将海量的网络安全数据进行语义关联和深度挖掘，形成具有逻辑关系的知识网络，为安全分析、威胁检测和应急响应提供全面的数据支持，从而增强网络安全态势感知和威胁响应能力。

在网络安全领域，STIX（structured threat information expression）[2] 和 ATT&CK （adversarial

收稿日期：2024–10–22

tactics, techniques, and common knowledge）[3]是2个重要的框架，它们在威胁情报共享和攻击行为分析中发挥了关键作用。STIX是由OASIS组织提出的一个开源标准，用于描述、共享和分析网络威胁情报，STIX框架提供了一种结构化的方式来表示威胁情报，包括威胁行为、攻击模式、受害者信息等，使得不同机构之间可以高效地共享和利用威胁情报。ATT&CK是由MITRE公司开发的一个知识库，详细描述了网络攻击者在攻击过程中使用的战术、技术和程序TTP（tactics，techniques，and procedures）。ATT&CK框架通过对真实攻击案例的研究，系统化地分类和描述了攻击者的行为模式，为安全防护和检测提供了重要的参考依据。除此之外，MITRE公司维护的攻击模式库CAPEC[4]、弱点分类库CWE[5]和通用漏洞披露CVE[6]、美国国家标准与技术研究院（NIST）维护的国家漏洞数据库NVD[7]和中国信息安全测评中心维护的国家信息安全漏洞共享平台CNVD[8]等也在网络安全威胁情报中扮演着重要角色。网络安全知识图谱的构建，可以整合和利用现有的网络安全框架和平台，通过智能化的数据处理和分析，提升网络安全防护的整体水平。

近年来，尤其是BERT[9]和GPT-2[10]发布以来，大语言模型（LLM, large language model）迎来了爆发式进步，GPT-3[11]更是将其推向了新高度。大语言模型的特点包括庞大的参数规模，通常在数十亿乃至万亿级别。大语言模型的多任务学习和上下文理解能力，使其在多种语言任务中表现出色。如今，以对话式语言大模型ChatGPT[12]、增强推理能力的多模态大模型GPT-4[13]为代表的大语言模型在各个领域的应用广泛且深远，如自然语言处理、医疗、法律、教育和内容创作等，显著提升了这些任务的性能和效率。大语言模型表现出来的能力，让研究人员不断探索更多的可能性，促进各行各业的生产力进步[14]。

传统的知识图谱构建面临严峻的技术挑战，如知识抽取、知识图谱补全、知识融合和知识推理等。大规模知识图谱的构建往往需要投入大量的人力、物力和时间成本，且依旧无法保证知识图谱质量和可用性。而大模型能有效解决这些问题，大语言模型内部存在海量的知识信息，在处理复杂的文本数据信息时，能够迅速地进行实体识别与关系抽取，有效应对知识图谱构建的挑战[15]。

本文将深入探讨如何利用大语言模型构建网络安全知识图谱和攻击知识图谱，为了实现这一目标，设计了一个基于大语言模型的框架AutoCTI2KG。该框架从网络威胁情报（CTI, cyber threat intelligence）中提取信息，经过预处理和结构化后，构建成网络安全知识图谱，涵盖多种实体（如威胁行为者、战术、攻击模式、恶意软件等）和关系。在此基础上，进一步将网络威胁情报映射到企业ATT&CK框架上，构建攻击知识图谱，帮助理解攻击者的战术、技术和程序，展示攻击路径和攻击手段。网络安全知识图谱包含了静态的威胁信息，攻击知识图谱能动态地展示攻击流程，两者结合提高了对威胁的感知能力。此外，本文还根据网络威胁情报和企业ATT&CK框架中的缓解措施，生成针对具体威胁情报的缓解措施，这些措施为网络安全管理人员提供了具体的操作参考，更有效地应对潜在威胁。

1 相关研究

数字化时代，网络威胁情报共享和网络安全知识图谱的构建成为保障网络安全的重要手段。网络威胁情报旨在通过收集、分析和共享有关潜在威胁的信息，帮助组织预防和应对网络攻击。网络安全知识图谱通过结构化和语义化的数据表示，提供全面的网络威胁情报视图，支持自动化威胁检测和响应。两者结合不仅提升了网络威胁情报的有效性和及时性，还促进了网络安全防御措施的智能化和自动化。

在网络安全知识图谱构建方面，史慧洋等[16]提出了由情报搜集、信息抽取、本体构建和知识推理构建威胁情报知识图谱的框架，可实现情报中重要指标的搜索和关联，基于预训练语言模型Bert，循环神经网络BiLSTM和条件随机场（CRF）的失陷指标识别抽取方法，加以正则匹配机制进行输出限制，用于从文本信息中识别抽取失陷指标信息，并进行结构化威胁信息表达标准格式转换。黄智勇等[17]针对网络安全领域的图谱构建任务，基于循环神经网络BiLSTM和CRF模型引入了外部网络安全词典来加强网络安全文本的特征，并结合多头注意力机制提取多层特征，最终在网络安全数据集上取得了更优异的结果。但所提的模型在面对一些无

特定语法的跨语言网络安全实体的抽取上表现并不如意。唐思宇等[18]将权威知识库作为数据源，利用Scrapy爬虫框架采集网络安全数据并进行知识抽取，使用图数据库实现网络安全知识图谱的构建，为安全人员提供直观、可靠的安全知识查询，但由于研究中抽取出的数据量相对较少，对于更加复杂的应用场景的表现还有待考证，并且算法的可移植性有待进一步提高。丁兆云等[1]和王晓狄等[19]，对网络安全知识图谱的相关研究进行了综述，介绍了构建网络安全知识图谱的关键技术，国内外最新的研究现状，并总结了网络安全知识图谱构建面临主要挑战，如数据源质量参差不齐，构建过程未考虑本体模型的动态演变，实例节点的语义特性难以表示，传统知识图谱推理模型不能够较好地适用于网络空间威胁知识推理等。

攻击知识图谱是一种用于建模和分析网络攻击路径及其演变的工具，与网络安全知识图谱不同，攻击知识图谱能反映攻击行为的动态变化，提供更为准确和及时的安全态势感知，帮助组织更好地应对不断变化的网络威胁。

在攻击知识图谱构建方面，Gao等[20]设计了一个用于自动化开源网络威胁知识收集和管理的系统THREATKG，自动从各种来源收集大量开源的网络威胁情报，提取高保真度的威胁知识，构建威胁知识图谱，并通过持续摄取新知识来更新知识图谱。Li等[21]提出了AttacKG，自动从网络威胁情报中提取结构化的攻击行为图，识别所采用的攻击技术，在报告之间聚合网络威胁情报，将攻击行为图增强为技术知识图谱。Zhang等[22]提出了一个基于大语言模型的全自动框架AttacKG+来构建攻击知识图谱，框架由重写器、解析器、标识器和总结器4个模块组成，每个模块通过指令提示和由大语言模型支持的上下文学习来实现。

大语言模型构建知识图谱是一项前沿的研究方向，结合自然语言处理技术，以自动化和智能化的方式生成和维护知识图谱，大语言模型具备强大的语言理解和生成能力，能够从海量文本中自动提取信息，显著提升知识图谱构建的效率和准确性，同时可以处理多种语言和领域的文本，具有广泛的适应性和扩展性。

在大语言模型构建知识图谱方面，Monica等[23]介绍了如何使用大语言模型在临床文本中进行零样本和少样本信息提取，以及如何利用指导式提示和解析器来提高输出的结构性。在3个新的标注数据集上评估了大语言模型，发现它们在临床信息提取任务上表现出色，甚至超过了现有的零样本和少样本基线。Wei等[24]探讨了是否可以通过直接提示大语言模型来构建强大的信息抽取（IE, information extraction）模型，将零样本信息抽取任务转化为一个多轮问答问题，并利用ChatGPT的强大功能设计了框架ChatIE，在3个信息抽取任务上进行了评估：实体关系三元组抽取、命名实体识别和事件抽取。在跨2种语言的6个数据集的实证结果表明，ChatIE在多个数据集上取得了令人印象深刻的性能，甚至在某些数据集上（如NYT11-HRL）超越了一些全样本模型。Polak[25]提出了ChatExtract方法，该方法利用先进的对话式大语言模型，在最少的背景知识下，完全自动化地进行非常准确的数据提取。ChatExtract可以应用于任何对话式大语言模型，并能实现非常高质量的数据提取。Chen等[26]引入了AutoKG，一种轻量且高效的自动化知识图谱构建方法，对由文本块组成的知识库，AutoKG使用大语言模型提取关键词，通过图拉普拉斯学习评估每对关键词之间的关系权重。初步实验表明，AutoKG提供了一种比语义相似性搜索更全面、更互联的知识检索机制，增强了大语言模型在生成更有洞察力和相关输出方面的能力。

大语言模型如GPT-4，因其广泛的适用性，在自然语言处理和人工智能领域掀起了新的浪潮。大语言模型作为黑箱模型，往往难以捕捉和访问事实性知识，知识图谱是结构化的知识模型，能够显式地存储丰富的事实性知识。知识图谱可以通过提供外部知识来增强大语言模型的推理能力和可解释性，因此，将大语言模型和知识图谱统一起来，利用它们互补的优势，是一个很有前景的方向[27]。

2　任务描述

网络威胁情报是记录和分析网络威胁活动的重要文档，这些报告通常是非结构化的，可以从多种来源获取，例如安全公司的威胁研究报告、开源情报（OSINT）、政府机构发布的安全通告以及社区共享的威胁情报平台[1]。网络威胁情报的格式和质量参差不齐，有些报告详细且结构清晰，有些则较为简略且格式混乱。尽管如此，网络威胁情报在网

络安全领域具有重要的作用，包括帮助安全团队识别和应对新兴威胁、制定防御策略以及提升整体安全态势感知。基于此，本文基于大语言模型技术对网络威胁情报进行解析，主要完成以下3个任务。

1) 构建网络安全知识图谱。网络安全知识图谱能够直观地展示各种威胁实体及其关系，如威胁行为者、恶意软件、攻击目标等，为网络安全团队提供一个全局视角，提升对威胁的理解和应对能力。从网络威胁情报中提取关键信息并构建网络安全知识图谱，将分散的威胁情报进行结构化整合，不仅有助于全面了解当前的威胁态势，还能为后续的安全分析和决策提供坚实的数据基础。

2) 构建攻击知识图谱。攻击知识图谱可以直观地展示攻击者的攻击步骤和使用的战术，有助于网络安全团队更好地理解攻击路径和攻击手段，从而制定更有效的防御措施。攻击知识图谱能够揭示攻击者的行为模式，帮助预测未来可能的攻击行为，提高防御的前瞻性和针对性。

3) 提出针对性的缓解措施。结合网络威胁情报中使用的战术，提出针对性的缓解措施，可以为网络安全管理人员提供具体的操作指导，这些措施能够帮助安全团队迅速采取行动，减少威胁造成的损害，提高网络防御的有效性和响应速度。

完成这3个任务的主要难点和困难如下。

1) 数据获取和处理。网络威胁情报的格式和质量参差不齐，筛选出高质量的报告是一个主要难点。

2) 实体识别和关系抽取。从非结构化文本中准确识别实体和关系需要复杂的自然语言处理技术，网络安全领域实体和关系的抽取还需要相应的背景知识。

3) 战术和技术的识别。将自然语言描述的信息映射到企业ATT&CK框架结构化表示需要深入理解企业ATT&CK框架战术和技术的定义。

深度学习和机器学习技术在处理大量数据和复杂模式识别方面虽然具有显著优势，能够自动化地进行实体识别和关系抽取，但这些方法通常需要大量标注数据进行训练。而大语言模型作为预训练模型，在自然语言文本处理方面表现出色，能够从非结构化文本中提取丰富的语义信息，采用大语言模型的好处如下。

1)高效的信息提取。大语言模型经过预训练具备网络安全背景知识，能够快速处理大量网络威胁情报，自动进行实体识别和关系抽取，提高知识图谱构建的效率和准确性。

2)持续学习和更新。大语言模型可以通过增量训练、持续学习等方式不断学习新的知识，适应不断变化的威胁环境。

3）捕捉实体的真实语义。大语言模型具备语义理解能力，能深入理解企业ATT&CK框架中战术和技术的定义。

3 方法步骤

本文设计了一个基于大语言模型的全自动框架AutoCTI2KG，通过调用大语言模型的接口，利用指令提示和上下文学习来实现，框架由5个模块组成：重写模块、网络安全知识图谱生成模块、攻击知识图谱生成模块、防护建议生成模块和总结输出模块，该框架的整体设计如图1所示，随着大语言模型对长文本的理解和处理能力的逐步增强[28-29]，可以将网络威胁情报、输出模板和企业ATT&CK框架等背景知识作为输入，通过给定的提示词，让大语言模型处理后按照模板进行输出。

3.1 重写模块

网络威胁情报在语言风格、写作逻辑等方面具有不同的特征，从中准确提取威胁行为仍然具有挑战性。重写模块的主要作用是对网络威胁情报进行处理，使其内容更加清晰和结构化，重写的目的主要有2个。

1) 去除噪声信息：保留原报告主要内容，提高可读性和专业性。

2) 提取攻击步骤：识别并总结攻击步骤，将攻击步骤对应到企业ATT&CK框架的战术和技术上，为后续的知识图谱生成提供基础数据。

重写模块不仅提高了数据的一致性和准确性，还能帮助网络安全团队更好地理解和分析威胁情报，确保信息在传递过程中不丢失关键细节。

3.2 网络安全知识图谱生成模块

网络安全知识图谱的构建和通用知识图谱构建过程类似，可以采取自顶向下的模式，本体的设计参考了ATT&CK、STIX及相关研究[1,19]。知识图谱本体是一个详细的指南，描述了如何组织和记录相应的信息，定义了类别（如漏洞、攻击者、防御措施）、类别的属性（如编号、严重程度）以及类别之

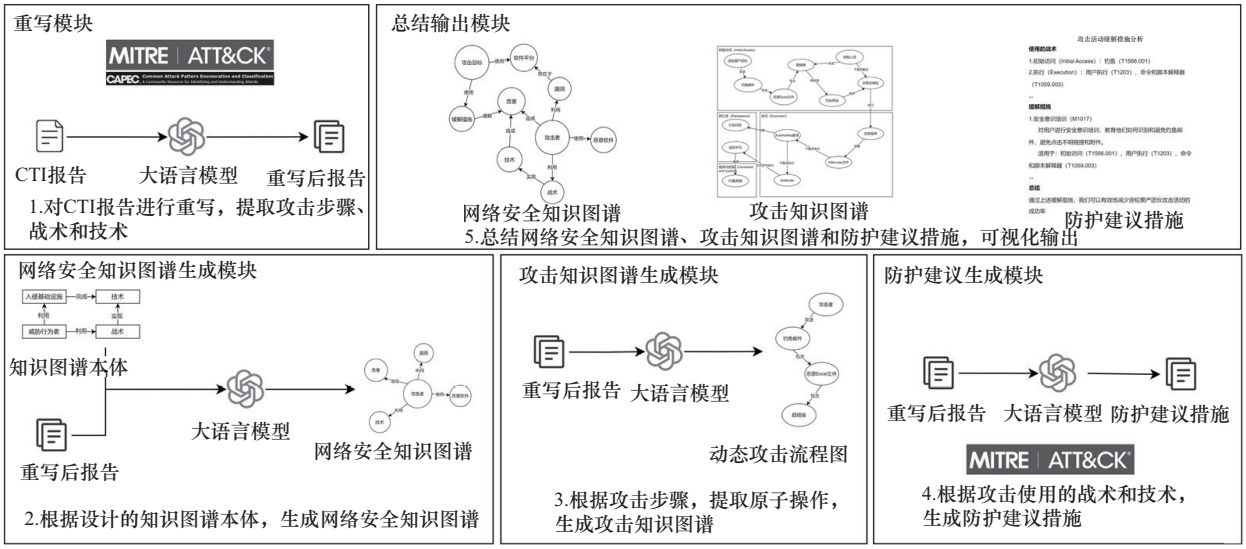

图1　AutoCTI2KG 框架的整体设计

间关系（如利用、修复）。基于本体，就能构建一个结构化、易于理解和使用的知识图谱。本文对网络安全知识图谱本体的设计如图2所示，每个方块表示一个类别，箭头及文字表示2个类别之间的关系。

网络安全知识图谱生成流程如图3所示，将企业 ATT&CK 框架 14 个战术及 235 项技术的定义、CAPEC 的 559 种攻击模式、知识图谱本体信息、重写后的报告输入大语言模型，让大语言模型进行实体识别和关系抽取，并给定大语言模型输出模板进行结果的输出。

图4为使用某网络威胁情报利用大语言模型自动完成知识图谱构建的结果，一共提取了 39 个实体，57 种关系，左上角和右下角分别代表了威胁行为者和攻击目标，两者之间包括威胁行为者使用的战术、技术和攻击模式，造成的危害和可采取的缓解措施等。网络安全知识图谱能够静态地展示攻击信息，但无法体现攻击流程，因此本文采用攻击知识图谱来进一步展示攻击的动态变化情况。

网络安全知识图谱生成模块的作用是将非结构化的网络威胁情报转化为结构化的知识图谱，展示各种威胁实体及其关系。通过构建网络安全知识图谱，安全团队可以全面了解当前的威胁态势，为安全分析和决策提供坚实的数据基础。网络安全知识图谱不仅有助于发现潜在的威胁模式，还能揭示不同威胁之间的关联，提升整体的威胁感知能力和响应效率。

3.3　攻击知识图谱生成模块

在网络安全领域，攻击事件是指网络攻击者或恶意行为者对信息系统、网络、应用程序或数据进行的任何未经授权的访问、使用、披露、破坏、修改或拒绝服务的行为。攻击事件通常由多个攻击步骤组成，每个攻击步骤包含一个或者多个原子操作。原子操作是指在攻击过程中执行的最小、不可再分的操作单元，原子操作表示为相互连接的三元组(s, a, o)，其中：s（subject）表示操作的发起者

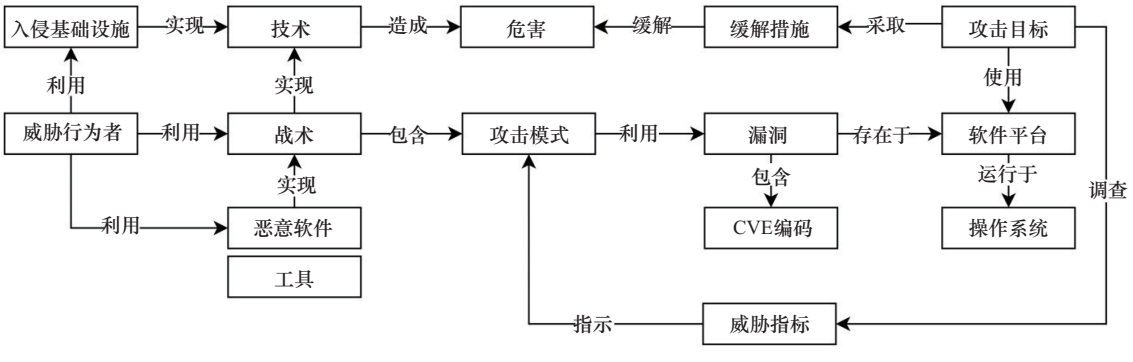

图2　网络安全知识图谱本体设计

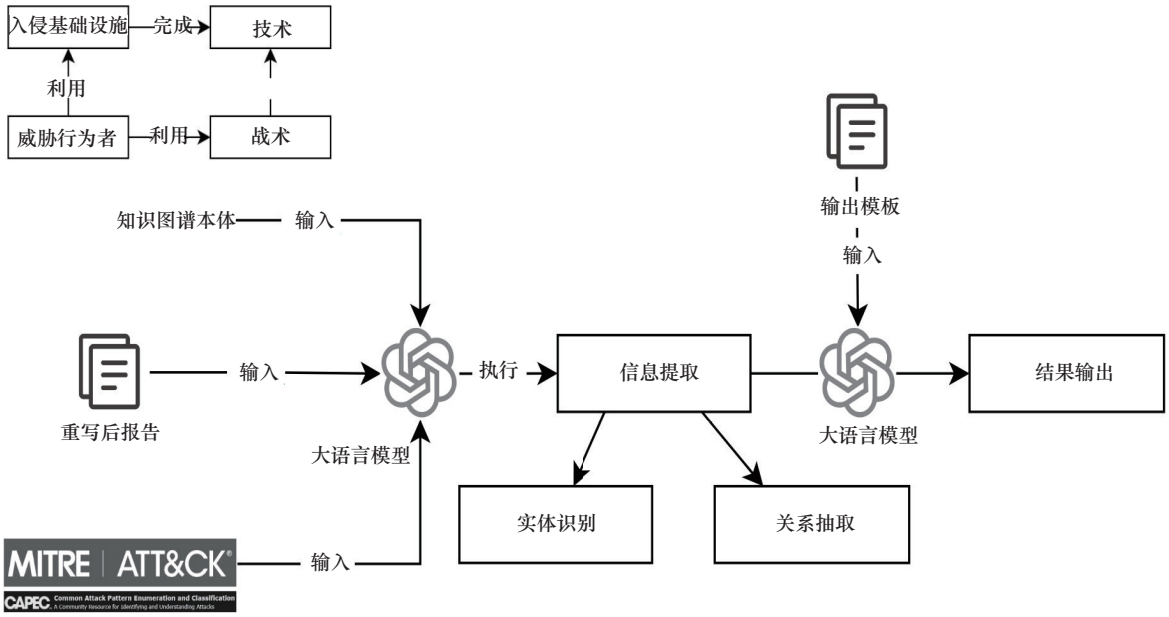

图3　网络安全知识图谱生成流程

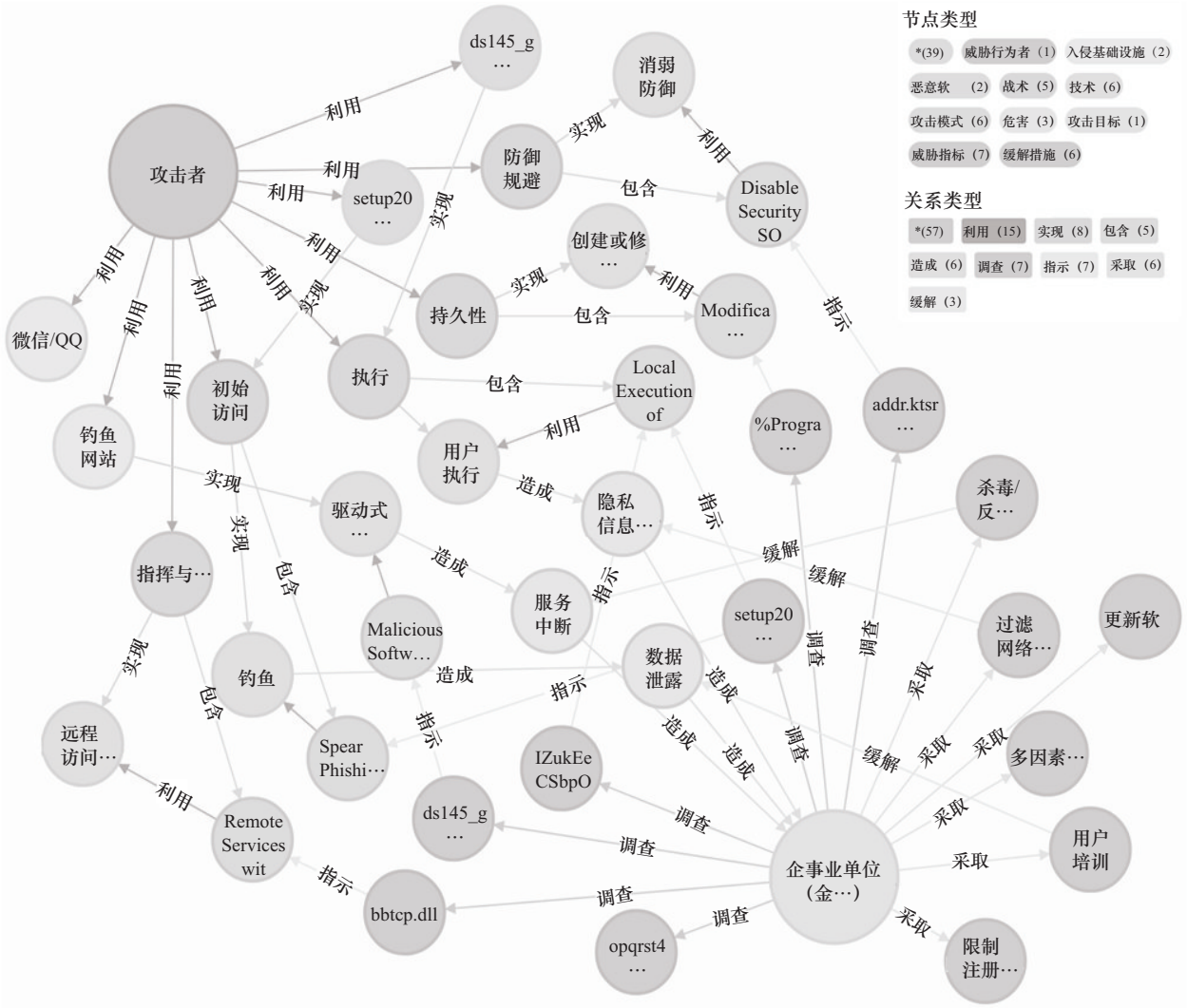

图4　网络威胁情报生成的知识图谱示例

或源实体，a（action）表示操作的具体行为或动作，o（object）表示操作的目标或目的实体。

例如，某个攻击步骤为“攻击者通过钓鱼邮件发送恶意 Excel 文档，诱导财税人员点击其中的超链接”，提取出其中的原子操作为：(s:攻击者, a:发送, o:钓鱼邮件)、(s:钓鱼邮件, a:包含, o:恶意 Excel 文档)、(s:恶意 Excel 文档, a:包含, o:超链接)、(s:财税人员, a:点击, o:超链接)。攻击步骤对应企业 ATT&CK 框架中的初始访问（Initial Access）阶段，因此，生成的攻击图如图 5 所示，直观展示了攻击步骤对应的攻击过程。

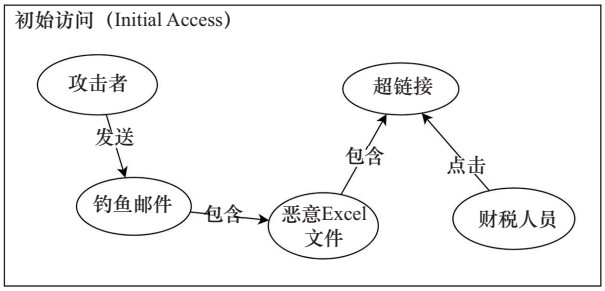

图 5　生成的攻击图

攻击知识图谱生成模块的流程如图 6 所示，根据网络威胁情报的攻击步骤进行原子操作的识别，识别后按照输出模板进行输出。

图 7 为根据网络威胁情报生成的完整攻击知识图谱示例，可以看出，这个攻击过程采用了 4 个战

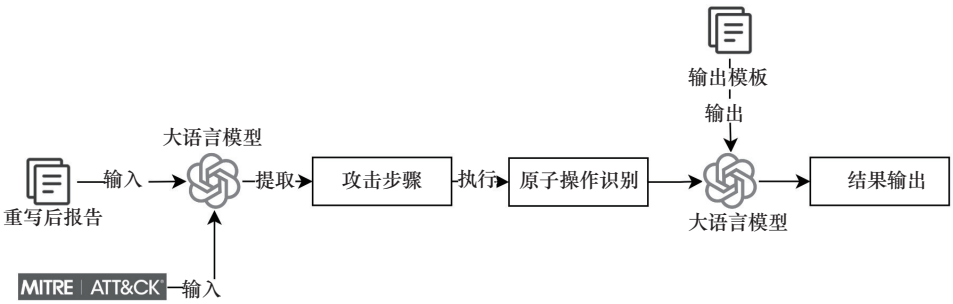

图 6　攻击知识图谱生成模块的流程

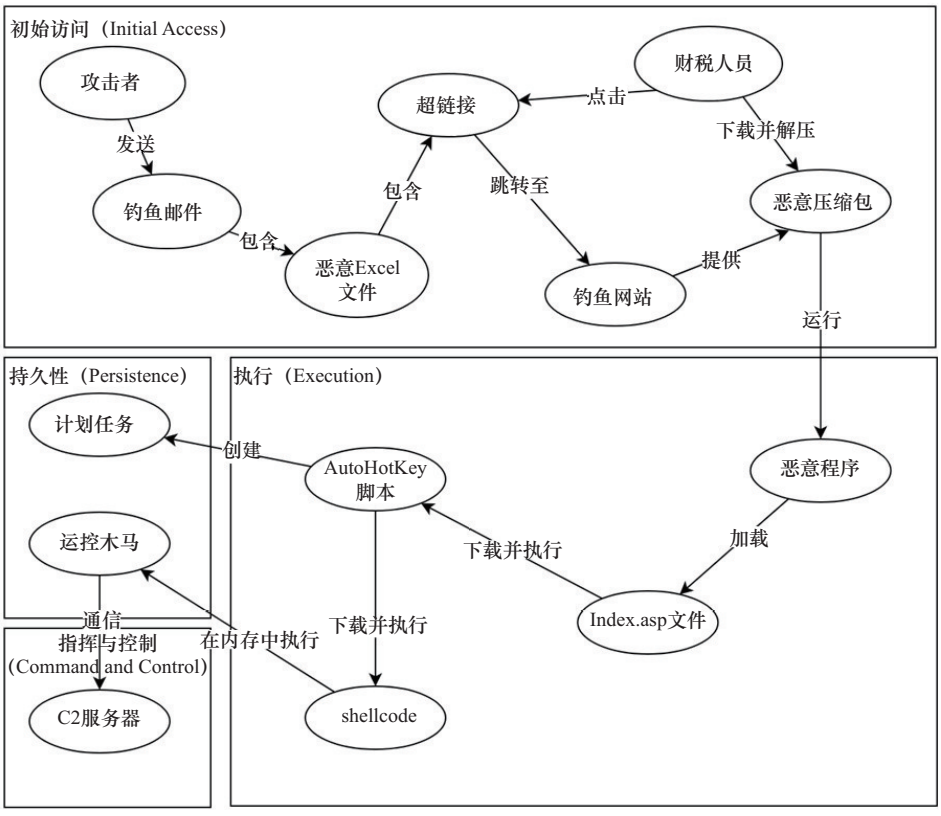

图 7　根据网络威胁情报生成的完整攻击知识图谱示例

术，包括初始访问、执行、持久性、指挥和控制，每个阶段包含若干原子操作，攻击知识图谱可以直观地看到整个网络威胁报告的攻击流程。

攻击知识图谱生成模块的作用是直观地展示攻击者的行为模式和攻击路径，帮助安全团队更好地理解攻击手段和制定防御措施。将攻击流程映射到企业ATT&CK框架，利用已有的知识库，提高攻击图的准确性和实用性。攻击知识图谱不仅可以揭示攻击者的策略，还能帮助预测未来可能的攻击行为，提升防御的前瞻性和针对性。

3.4 防护建议模块

企业ATT&CK框架定义了43种缓解措施用来针对技术或者子技术的实施，从而达到防御的效果。在重写模块，从网络威胁情报中提取出了攻击者采用的战术和技术，针对这些战术和技术，可以直接找到对应的缓解措施。利用大语言模型的语义理解和文本生成能力，可以根据具体的网络威胁情报对使用的战术和缓解措施进行总结，并生成报告。

防护建议模块的作用是为网络安全管理人员提供具体的操作指导，帮助他们迅速采取行动，减少威胁造成的损害，结合网络威胁情报，提出更加针对性和有效的防护措施。不仅有助于提高防御的有效性和响应速度，还能为网络安全策略的制定提供科学依据，确保防护措施的全面性和可行性。

3.5 总结输出模块

为了实现自动化输出及可视化，通过总结输出模块进行数据汇总，知识图谱可视化采用Neo4j[30]，动态攻击图采用Graphviz[31]，防护建议措施生成pdf文档，总结输出模块的作用是将所有信息整合在一起，形成一份全面的报告。这份报告可以为安全团队提供全局视角，帮助他们全面了解威胁情报、攻击手段和防护措施，从而提高整体网络安全防护水平。通过系统化地汇总和分析网络威胁情报，网络安全团队可以更准确地评估风险，制定更加有效的应对策略。

4 实验验证

在本节中，通过实验来验证AutoCTI2KG作为网络威胁情报解析器构建网络知识图谱和攻击知识图谱的准确性，以判断其性能和有效性。

4.1 实验设置

从开源情报网站Freebuf[32]收集了部分真实的网络威胁情报，这些报告描述了实际发生的高级可持续威胁APT活动，利用大语言模型GPT-4的API来实现框架。为了确保评估的准确性，选取了10篇高质量的网络威胁情报，由3位具有网络安全背景的技术人员对这些报告进行了详细分析，按照预先设计的本体构建了网络安全知识图谱和攻击知识图谱，形成标注数据。随后，将标注的结果与AutoCTI2KG自动化生成的结果进行对比分析。

4.2 实验结果

网络安全知识图谱实体识别、关系抽取和攻击知识图谱原子操作识别结果如表1所示，表中−2（+2）分别表示假负例和假正例，在网络安全知识图谱实体识别任务F1值为0.899，在关系抽取任务F1值为0.915，在攻击知识图谱中原子操作识别任务F1值为0.913，表现了较高的准确性。

在战术、技术和攻击模式识别结果如表2所示，战术识别的F1值为0.906，技术识别F1值为0.879，攻击模式识别F1值为0.886，同样表现出了优异的性能。

4.3 实验分析

1) 数据标注的挑战

数据标注对技术人员来说是一个很大的挑战，在标注过程中，不仅要求技术人员熟悉企业ATT&CK框架和CAPEC攻击模式，还需要对网络威胁情报进行深度分析。这是一项耗时耗力的工作，完成一篇网络威胁情报的标注通常需要4小时以上，且人工标注的数据也不能保证100%的准确。

2) 提示词设计的重要性

大语言模型的任务表现高度依赖于输入的指令提示，不同的提示词会导致结果存在显著差异。这表明，在使用大语言模型进行自动化任务时，提示词的设计至关重要，直接影响到模型的输出质量和准确性。因此，在实际应用中，需要对提示词进行精心设计和优化，以确保模型能够生成高质量的结果。

3) 自动化知识图谱构建的技术意义

自动化知识图谱构建技术在网络安全领域具有重大意义，通过自动化工具，可以大幅提升知识图谱构建的效率和准确性，并且能够在大量的网络威胁情报中快速识别和抽取有价值的信息。这不仅提高了网络安全专家的工作效率，还为快速响应和处理网络威胁提供了有力支持。

表1　　实体识别、关系抽取和攻击知识图谱原子操作识别结果

报告	实体识别		关系抽取		原子操作	
	标注	AutoCTI2KG	标注	AutoCTI2KG	标注	AutoCTI2KG
1	39	−2(+2)	55	−2(+3)	13	−1(+2)
2	31	−3(+3)	29	−1(+3)	16	−2(+1)
3	32	−3(+5)	46	−2(+6)	12	−0(+3)
4	35	−2(+3)	39	−3(+2)	17	−1(+0)
5	33	−5(+3)	38	−2(+4)	11	−0(+2)
6	23	−3(+4)	25	−5(+3)	14	−1(+1)
7	46	−5(+2)	40	−3(+5)	15	−2(+2)
8	25	−2(+3)	36	−2(+2)	21	−1(+0)
9	34	−3(+6)	37	−4(+3)	7	−0(+1)
10	30	−4(+2)	32	−2(+5)	12	−2(+2)
准确率	1.000	0.901	1.000	0.905	1.000	0.932
召回率	1.000	0.896	1.000	0.925	1.000	0.894
F1 值	1.000	0.899	1.000	0.915	1.000	0.913

表2　　战术、技术和攻击模式识别结果

报告	战术识别		技术识别		攻击模式	
	标注	AutoCTI2KG	标注	AutoCTI2KG	标注	AutoCTI2KG
1	6	−0(+0)	6	−0(+0)	5	−0(+1)
2	5	−1(+1)	6	−0(+1)	6	−1(+0)
3	5	−0(+1)	6	−1(+0)	4	−0(+0)
4	4	−0(+1)	5	−0(+2)	6	−0(+1)
5	6	−0(+1)	6	−1(+0)	4	−1(+2)
6	6	−1(+0)	7	−2(+1)	5	−0(+1)
7	3	−0(+2)	4	−0(+1)	7	−1(+0)
8	6	−1(+0)	6	−2(+1)	5	−0(+1)
9	4	−0(+1)	7	−0(+1)	3	−1(+0)
10	6	−0(+0)	4	−0(+2)	4	−0(+2)
准确率	1.00	0.947	1.00	0.905	1.000	0.911
召回率	1.00	0.869	1.00	0.855	1.000	0.862
F1 值	1.00	0.906	1.00	0.879	1.000	0.886

4.4　案例研究

尽管AutoCTI2KG在实体识别、关系抽取和攻击模式识别任务中表现出较高的准确性，但在某些情况下仍存在识别错误的情况。以下是一些失败案例。

1) 实体识别。“恶意软件通过JSE脚本执行并释放诱饵PDF”的描述，AutoCTI2KG错误地将“诱饵PDF”识别为一个独立的恶意软件实体。实际上，“诱饵PDF”是恶意软件执行的一个结果，并不是一个独立的恶意软件实体。

2) 技术识别。AutoCTI2KG将“木马与C2服务器通信并接收指令”识别为“网络嗅探（Network Sniffing）”技术，而实际上它应该是“应用层协议（Application Layer Protocol）”技术。

3) 攻击模式识别。“在使用白加黑技术和VMP加壳保护恶意程序的情况下”，AutoCTI2KG将其识别为“使用恶意文件（Using Malicious Files）”，而实际上它应该被识别为“混淆文件或信息（Obfuscated Files or Information）”。

这些失败案例表明，完成识别任务需要很强的专业性和领域知识，尽管如此，大语言模型在这些任务中仍表现出较高的准确性，为未来大语言模型在网络威胁情报分析中的应用提供了重要参考依据，并表明可以更有效地利用自动化工具来辅助网络安全专家进行威胁情报分析。

5 结束语

网络安全知识图谱和攻击知识图谱在网络安全威胁情报分析中具有重要的意义和应用价值，这些知识图谱能够系统化地表示和组织复杂的网络威胁情报信息，帮助安全专家更高效地理解、分析和应对各种网络威胁。本研究基于大语言模型设计并实现了一个自动化框架AutoCTI2KG，旨在自动化地完成网络安全知识图谱和攻击知识图谱的构建任务。实验结果表明，AutoCTI2KG在实体识别、关系抽取和攻击模式识别等任务中表现出较高的准确性，各项任务的F1值均在0.9左右，验证了其有效性和实用性。

尽管如此，本文仍存在一些不足之处。例如，网络威胁情报中的图表信息尚未得到充分处理，这可能会影响某些复杂情报的全面解析。未来计划进一步扩展研究工作，通过大批量分析网络威胁情报，研究并发现其中的一些规律，为网络安全提供新的思路和策略。同时，尝试使用本地化的大语言模型来完成这些任务，以提高模型的适应性和性能。此外，还将探索如何更好地处理图表信息，以提升整体情报分析的全面性和准确性。通过这些努力，希望能够推动网络威胁情报自动化分析技术的发展，为网络安全防护提供更强有力的支持。

参考文献：

[1] 丁兆云, 刘凯, 刘斌, 等. 网络安全知识图谱研究综述[J]. 华中科技大学学报(自然科学版), 2021, 49(7): 79-91.
DING Z Y, LIU K, LIU B, et al. Survey of cyber security knowledge graph[J]. Journal of Huazhong University of Science and Technology (Natural Science Edition), 2021, 49(7): 79-91.

[2] OASIS. STIX™ Version 2.1[EB/OL]. (2021)[2024-8-10].

[3] MITRE. MITRE ATT&CK®[EB/OL]. (2024)[2024-08-10].

[4] MITRE. Common Attack Pattern Enumeration and Classification (CAPEC)[EB/OL]. (2024)[2024-08-10].

[5] MITRE. Common Weakness Enumeration (CWE) [EB/OL]. (2024) [2024-08-10].

[6] MITRE. Common Vulnerabilities and Exposures (CVE)[EB/OL]. (2024) [2024-08-10].

[7] NIST. National Vulnerability Database (NVD) [EB/OL]. (2024) [2024-08-10].

[8] CNVD. China National Vulnerability Database (CNVD) [EB/OL]. (2024)[2024-08-10].

[9] DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding[C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Stroudsburg: ACL Press 2019: 4171-4186.

[10] RADFORD A, WU J, CHILD R, et al. Language models are unsupervised multitask learners[EB/OL]. (2019)[2024-08-10].

[11] BROWN T, MANN B, RYDER N, et al. Language models are few-shot learners[J]. Advances in Neural Information Processing Systems, 2020, 33: 1877-1901.

[12] OpenAI. ChatGPT: optimizing language models for dialogue[EB/OL]. (2022)[2024-08-10].

[13] OpenAI, ACHIAM J, ADLER S, et al. GPT-4 technical report[J]. arXiv Preprint, arXiv: 2303.08774v6, 2023.

[14] 车万翔, 窦志成, 冯岩松, 等. 大模型时代的自然语言处理: 挑战、机遇与发展[J]. 中国科学: 信息科学, 2023, 53(9): 1645-1687.
CHE W X, DOU Z C, FENG Y S, et al. Towards a comprehensive understanding of the impact of large language models on natural language processing: challenges, opportunities and future directions[J]. Scientia Sinica (Informationis), 2023, 53(9): 1645-1687.

[15] 黄勃, 吴申奥, 王文广, 等. 图模互补: 知识图谱与大模型融合综述[J]. 武汉大学学报(理学版), 2024, 70(4): 397-412.
HUANG B, WU S A, WANG W G, et al. KG-LLM-MCom: a survey on integration of knowledge graph and large language model[J]. Journal of Wuhan University (Natural Science Edition), 2024, 70(4): 397-412.

[16] 史慧洋, 魏靖烜, 蔡兴业, 等. 威胁情报提取与知识图谱构建技术研究[J]. 西安电子科技大学学报, 2023, 50(4): 65-75.
SHI H Y, WEI J X, CAI X Y, et al. Research on threat intelligence extraction and knowledge graph construction technology[J]. Journal of Xidian University, 2023, 50(4): 65-75.

[17] 黄智勇, 余雅宁, 林仁明, 等. 基于改进BiLSTM-CRF模型的网络安全知识图谱构建[J]. 现代电子技术, 2024, 47(6): 15-21.
HUANG Z Y, YU Y N, LIN R M, et al. Knowledge graph construction for network security base on modified BiLSTM-CRF[J]. Modern Electronics Technique, 2024, 47(6): 15-21.

[18] 唐思宇, 李赛飞, 张丽杰. 基于Neo4j的网络安全知识图谱构建分析[J]. 信息安全与通信保密, 2022, 20(8): 60-70.
TANG S Y, LI S F, ZHANG L J. Research on the construction of cyber security knowledge graph based on Neo4j[J]. Information Security and Communications Privacy, 2022, 20(8): 60-70.

[19] 王晓狄, 黄诚, 刘嘉勇. 面向网络安全开源情报的知识图谱研究综述[J]. 信息网络安全, 2023, 23(6): 11-21.

WANG X D, HUANG C, LIU J Y. A survey of cyber security open-source intelligence knowledge graph[J]. Netinfo Security, 2023, 23(6): 11-21.

[20] GAO P, LIU X Y, CHOI E, et al. ThreatKG: an AI-powered system for automated open-source cyber threat intelligence gathering and management[J]. arXiv Preprint, arXiv: 2212.10388v2, 2022.

[21] LI Z Y, ZENG J, CHEN Y, et al. AttacKG: constructing technique knowledge graph from Cyber threat intelligence reports[C]//Proceedings of Lecture Notes in Computer Science. Cham: Springer International Publishing, 2022: 589-609.

[22] ZHANG Y H, DU T W, MA Y S, et al. AttacKG+: boosting attack knowledge graph construction with large language models[J]. arXiv Preprint, arXiv: 2405.04753v1, 2024.

[23] AGRAWAL M, HEGSELMANN S, LANG H, et al. Large language models are few-shot clinical information extractors[J]. arXiv Preprint, arXiv: 2205.12689v2, 2022.

[24] WEI X, CUI X Y, CHENG N, et al. ChatIE: zero-shot information extraction *via* chatting with ChatGPT[J]. arXiv Preprint, arXiv: 2302.10205v2, 2023.

[25] POLAK M P, MORGAN D. Extracting accurate materials data from research papers with conversational language models and prompt engineering[J]. Nature Communications, 2024, 15(1): 1569.

[26] CHEN B H, BERTOZZI A L. AutoKG: efficient automated knowledge graph generation for language models[C]//Proceedings of the 2023 IEEE International Conference on Big Data (BigData). Piscataway: IEEE Press, 2023: 3117-3126.

[27] PAN S R, LUO L H, WANG Y F, et al. Unifying large language models and knowledge graphs: a roadmap[J]. IEEE Transactions on Knowledge and Data Engineering, 2024, 36(7): 3580-3599.

[28] LI J Q, WANG M M, ZHENG Z L, et al. LooGLE: can long-context language models understand long contexts? [J]. arXiv Preprint, arXiv: 2311.04939v2, 2023.

[29] DONG Z C, TANG T Y, LI J Y, et al. BAMBOO: a comprehensive benchmark for evaluating long text modeling capacities of large language models[J]. arXiv Preprint, arXiv: 2309.13345, 2023.

[30] Neo4j. Neo4j[EB/OL]. (2024)[2024-08-10].

[31] The Graphviz Authors. Graphviz[EB/OL]. (2024)[2024-08-10].

[32] FreeBuf. FreeBuf[EB/OL]. (2024)[2024-08-10].

[作者简介]

赖清楠（1990–），男，江西兴国人，北京大学工程师，主要研究方向为网络安全、人工智能、安全大数据分析等。

金建栋（1994–），男，蒙古族，内蒙古通辽人，北京大学工程师，主要研究方向为网络空间安全、开源情报、智能推理决策等。

周昌令（1977–），男，重庆人，博士，北京大学高级工程师，主要研究方向为网络安全、人工智能、安全大数据分析等。