特色专题

2024年网络空间安全科技热点回眸

苏璞睿,冯登国*

摘要 2024年世界各国继续加大网络安全战略部署,针对软件供应链、大语言模型等新兴场景的安全问题发布了相关技术指南和管理政策。面对新形势、新问题,中国进一步完善网络空间安全方面的政策法规,以规划指导网络空间安全体系建设、规范网络空间安全产业发展。然而,当前网络攻击事件依然频发,APT攻击、勒索攻击等网络攻击对全球网络空间安全造成严重威胁。2024年,围绕数据安全、人工智能安全、量子计算和软件漏洞等热点领域有了系列突破。未来,亟需发展新的技术能力和技术体系,为构建安全、健康的网络空间环境提供技术保障。

关键词 网络空间安全;数据安全;人工智能安全;量子计算;软件漏洞

2024年,全球网络空间安全形势依然严峻,各种新兴威胁和技术挑战层出不穷。人工智能、量子计算等前沿技术给网络空间安全带来了新的挑战和新的机遇,数据安全、软件漏洞等经典问题依然话题不断。在这一背景下,各国不断加强网络空间安全技术的研究与应用,积极应对日益严峻的网络空间攻击威胁。

1 国际网络空间安全形势复杂

各国加大网络空间安全战略部署,纷纷发布人工智能、软件供应链等新场景的技术指南和技术标准。然而,全球范围内重大网络攻击事件依然频发,俄乌冲突等热点事件中网络空间的对抗与博弈继续加剧。

1.1 各国加大网络空间安全战略部署

世界各国高度重视网络空间安全问题,多国将其上升到国家战略高度,针对最新形势,出台一系列政策、制度,规划指导国家网络空间安全体系建设,规范网络安全产业发展,部署网络空间安全重点技术攻关方向。2024年,美国先后发布多项重要战略,旨在提升国家网络防御能力,例如,3月28日美国国防部发布《国防工业基地网络安全战略》[1],5月6日美国国务院发布《美国国际网络空间和数字政策战略》[2];针对软件供应链问题,6月18日美国能源部发布"供应链网络安全原则"[3],美国网络安全与基础设施安全局(CISA)10月15日发布《软件组件透明度框架》[4],旨在提高软件组件透明度,加强国家各类基础设施信息系统安全。欧盟也进一步加强网络安全措施,2024年1月7日欧盟新版《网络安全条例》生效[5],要求建立风险管理框架,并设立监督委员会;1月31日推出网络安全认证计划(EUCC)[6],旨在确保ICT产品符合安全标准;10月10日通过的《网络弹性法案》[7],用于确保数字产品在全生命周期内符合安全要求,作为欧盟首部对包含数字元素产品提出强制性网络安全要求的立法文件,已于12月10日正式生效。

中国科学院软件研究所, 北京 100190

收稿日期:2024-12-25;修回日期:2025-01-05

作者简介:苏璞睿,研究员,研究方向为网络空间安全,电子信箱: purui@iscas.ac.cn;冯登国(通信作者),研究员,中国科学院院士,研究方向为网络与信息安全,电子信箱:fengdg@263.net

引用格式:苏璞睿,冯登国. 2024年网络空间安全科技热点回眸[J]. 科技导报, 2025, 43(1): 102-117; doi: 10.3981/j.issn.1000-7857.2024.12.01867

针对以大语言模型(large language model, LLM)为代表的人工智能技术的快速发展与应用带来的新的安全挑战,各国出台了一系列针对性的技术指南和管理政策,以规范行业发展,构建健康生态,规避安全风险。2024年4月,五眼联盟国家发布《AI系统安全部署指南》[8],要求提高AI系统安全性、消除漏洞并实施多层防护。2024年10月24日,美国政府发布了首份关于人工智能的国家安全备忘录[9],旨在加强美国在AI领域的领导地位,利用AI实现国家安全目标,并提升AI的安全性、保障性和可信度。同日,美国白宫发布了《国家安全领域推进人工智能治理和风险管理框架》[10],进一步明确了AI风险管理和透明度要求。其他国家也发布了一系列相关指南。2024年1月,澳大利亚网络安全中心联合全球网络安全机构发布《参与人工智能指南》[11],指出AI常见的风险(如数据中毒和隐私问题),并提出缓解措施建议。2024年10月新加坡网络安全局发布的《人工智能系统安全指南》[12]强调了供应链攻击和对抗性机器学习等新兴风险,并提出了相关安全防控措施建议。

中国为进一步落实《网络安全法》《数据安全法》等法律法规,构建健康安全的网络空间环境,2024年出台了一系列技术指南和相关管理制度。例如,1月19日,工业和信息化部发布《工业控制系统网络安全防护指南》[13]提出33项安全防护要求,旨在提升工业控制系统的网络安全保障水平。5月15日,中央网信办等4部委发布《互联网政务应用安全管理规定》[14]要求政务应用应遵循法律法规要求,采取技术措施确保安全稳定运行。12月11日,国家发改委发布《电力监控系统安全防护规定》[15]。

数字经济快速发展,数据流动愈发广泛,确保数据安全可控是保障数字经济健康发展的重要基础。2024年1月4日,17个部门发布《"数据要素×"三年行动计划(2024—2026年)》[16],强调数据安全贯穿全过程,将完善法规和保护个人信息,提升整体数据安全水平。9月30日,国务院公布《网络数据安全管理条例》为数据处理者提供了具体的合规指引[17]。10月29日,中国工业和信息化部印发《工业和信息化领域数据安全事件应急预案(试行)》[18],加快推动工业和信息化领域数据安全应急处置工作制度化、规范化开展。11月16日,习近平主席在APEC会议的讲话[19]中提到,中国正在因地制宜发展新质生产力,深化同各方绿色创新合作,将发布《全球数据跨境流动合作倡议》,愿同各方共同促进高效、便利、安全的数据跨境流动,为亚太高质量发展贡献力量。

开源软件的使用不断增长,软件供应链攻击的防范已成为网络空间安全防御的关键任务。党的二十届三中全会审议通过的《中共中央关于进一步全面深化改革、推进中国式现代化的决定》[20]强调"健全提升产业链供应链韧性和安全水平制度",指出抓紧打造自主可控的产业供应链,健全强化集成电路、工业母机、医疗装备、仪器仪表、基础软件、工业软件、先进材料等重点产业链发展体制机制,全链条推进技术攻关、成果应用。2024年11月1日,中国出台的《网络安全技术 软件供应链安全要求》(GB/T 43698—2024)和《网络安全技术 软件产品开源代码安全评价方法》(GB/T 43848—2024)2项推荐性国家标准正式实施[21]。

人工智能技术的快速发展和广泛应用,也带来了新的问题和挑战。为规范行业发展,构建良好生态,中国出台了系列标准和规范。2024年2月29日,全国网络安全标准化技术委员会出台《生成式人工智能服务安全基本要求》[22],为生成式人工智能(artificial intelligence generated content, AIGC)服务提供安全评估指南,并为相关部门评判服务安全水平提供参考;9月9日,委员会发布《人工智能安全治理框架》1.0版[23],旨在推动人工智能创新与安全发展,框架强调通过风险导向、协同应对等原则,防范内生安全风险和应用层面风险,促进人工智能的健康发展和国际合作,推动全球人工智能安全治理体系的建设。

1.2 重大网络安全事件频发

2024年,微软蓝屏事件、黎巴嫩传呼机爆炸事件等供应链安全事件进一步展示了供应链安全的重要性,高级持续性威胁(advanced persistent threat, APT)攻击呈现前所未有的严峻态势,勒索攻击依旧猖獗,俄乌冲突等热点事件中网络战和信息战成为重要斗争手段。

1.2.1 供应链安全事件

2024年发生了多起供应链相关的重大安全事件。2024年3月爆出的XZ-Utils后门事件是一场持续多年的供应链代码投毒,一名使用假身份的恶意开发者通过多年参与项目开发赢得了XZ-Utils库开

发者的信任,被添加为该开源项目的维护者,最终在其试图向代码中添加后门时被发现[24]。该事件显示了开源生态系统面临的供应链攻击风险(图1[24])。7月19日的微软蓝屏事件,是由美国网络安全公司CrowdStrike分发异常更新导致的Windows操作系统蓝屏,最终致使银行、机场、电视台、医疗机构、酒店和企业的一系列的信息系统无法使用,全球多地航班停飞、医疗设备瘫痪、金融系统中断[25]。10月出现的Linux内核"清洗"俄罗斯开发者事件,俄罗斯程序员被从Linux内核开发者名单中除名,他们均为对Linux内核项目特定领域进行维护的关键成员,该事件进一步展示了当前在软件供应方面存在的巨大风险[26]。2024年9月17日和18日,黎巴嫩发生了连续的通信功能设备爆炸事件,涉及传呼机、对讲机、智能手机、太阳能板、收音机、汽车电池、指纹识别器、专线联系系统、无线电系统等[27]。该事件给全球敲响了警钟,安全可靠的软、硬件供应链才能为网络空间和物理空间提供安全保障。

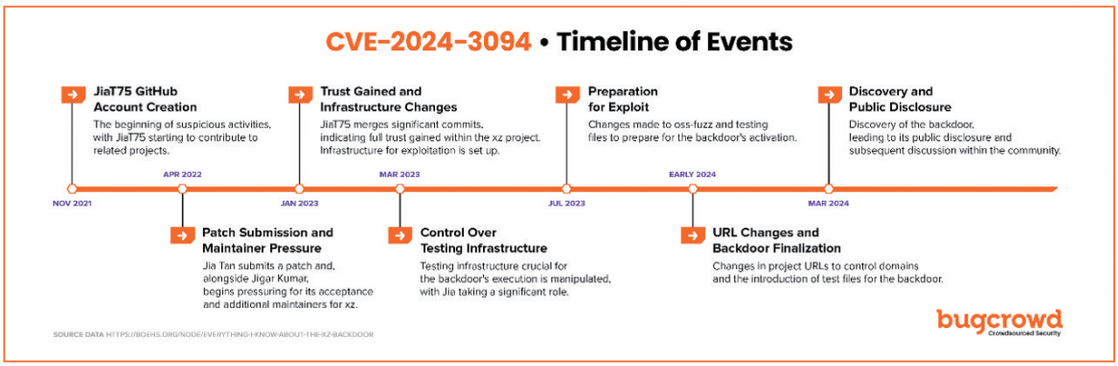

图1 XZ-Utils后门事件时间线

1.2.2 APT攻击

2024年,APT攻击显现出国家支持等特点,在资金、技术和人才优势的支撑下,能实施跨越地理边界的精准打击。《网络安全威胁2024年中报告》[28]显示,2024年上半年,政府、科研教育和国防军事是APT攻击的主要目标,分别占比31.3%、15.0%、13.8%。信息技术、制造业和新闻媒体等领域也频繁遭受攻击,至少48个国家受到波及,韩国、乌克兰、以色列和印度等地区受创严重(图2[29])。

其中,美国遭受的APT攻击主要来自国家支持的组织(如APT35)和跨境犯罪团体(如BogusBazaar和Smishing Triad),主要目标为政府、金融、科技、学术和基础设施。攻击手段包括AI诱饵、信息操控和零日漏洞(0day vulnerability)。中国也面临高频的APT攻击,主要集中在信息技术、政府、科研教育领域。除了传统APT组织外,多个未知威胁组织开始活跃,针对中国的关键领域进行多样化攻击,如银狐组织通过即时通信工具攻击财务人员,"amdc6766"通过供应链投毒攻击运维人员。攻击活动也向新能源、低轨卫星、人工智能等敏感领域扩展[30]。

图2 2024年上半年公开披露的高级威胁活动针对的国家和地区

1.2.3 勒索攻击

勒索软件已成为攻击者获取非法利益的主要工具,威胁着企业、政府和个人的数字资产安全。根据美国云安全公司Zscaler 2024年的报告[31]显示,全球勒索攻击数量同比增长18%,受影响企业增加58%,其中能源行业、食品服务行业、技术企业、医疗行业

等是攻击重点。

勒索软件组织数量和活跃度不断增加,英国威胁情报公司Searchlight Cyber报告[30]称,追踪到73个活跃组织,较去年增长56%。其中,LockBit组织仍占据受害者最多记录,2024年2月成立的RansomHub组织因为实施171次攻击而引起格外关注。当前暗网中已经出现了勒索软件即服务(Ransomware-as-a-Service, RaaS)的黑客工具,降低了勒索攻击的技术门槛,使网络犯罪分子能够快速构建并发起复杂的勒索攻击,2024年BlackBasta组织即利用RaaS针对北美、欧洲和澳大利亚的关键基础设施发起多次攻击[31]。

重大勒索事件也凸显了这一威胁的严重性。2024年2月,BlackCat组织攻击了美国联合健康集团子公司Change Healthcare[32],窃取了6 TB敏感数据,影响超过1/3的美国人。该公司支付2200万美元赎金以防数据泄露,部分服务中断超过9个月。Dark Angels组织在2024年3月成功勒索药品分销巨头公司Cencora 7500万美元[33],创下全球赎金新纪录。7月,美国电信巨头AT&T披露,其托管在Snowflake云服务公司的数据发生泄漏,涉及几乎所有客户,数量高达1.1亿人[34]。2024年8月针对俄罗斯Stoli集团的勒索软件攻击致使公司的ERP系统瘫痪[35],使得包括财务、供应链在内的核心业务全面转入手动操作模式,并导致其美国子公司运营中断,无法向贷方提供财务报告,因而被指控债务违约,直接申请破产。

1.2.4 俄乌冲突中的网络对抗

2024年,俄乌冲突中的网络对抗不断。根据乌克兰计算机应急响应小组(CERT-UA)报告[36],2024年上半年发生了1739起网络事件,比2023年下半年增加19%。然而,尽管网络事件增加,重大事件数量却大幅下降,2024年上半年的1739起网络事件中重大事件下降了90%。这表明攻击策略从广泛破坏转向持续的情报收集。公开新闻报道称,与以往的大规模基础设施攻击不同,俄罗斯的黑客组织开始采用更加隐秘和长期的渗透策略。

2024年1月,乌克兰国防情报局表示,黑客组织"BO Team"攻击了俄罗斯国家空间水文气象研究中心[37],导致该气象研究中心280台服务器被摧毁,并且丢失了2 PB、至少价值1000万美元的数据。3月,乌克兰宣布,其网络专家成功侵入俄罗斯国防部的服务器[38],并获得了大量机密文件。10月,乌克兰再次攻击俄罗斯金融和电信公司,造成支付系统故障,并摧毁了800多台服务器[39]。同时,为应对日益严重的网络威胁,乌克兰国防部在10月7日宣布建立了新的网络事件响应中心[40],加强网络防御能力。

与此同时,俄乌冲突中也爆发了信息战、舆论战等形式的网络空间安全对抗。2024年9月3日,美国兰德公司发布《乌克兰抵制俄罗斯虚假信息:未来冲突的教训》报告[41]指出,自俄乌冲突以来,俄罗斯发动了广泛、大量和多渠道的虚假信息运动。

2 数据安全已成为数字时代的关键议题

近年来,数据作为生产要素,其跨域、跨行业、跨境流通已成为全球发展趋势。为了应对数据外循环这一庞大需求,以隐私计算和机密计算为代表的数据安全关键技术在2024年迎来了迅猛发展。

2.1 隐私计算

隐私计算可为数据外部流通提供全流程的可信保障,已成为推动数据要素跨域流通与应用的关键技术方向。

针对云计算等典型数据流通和应用场景下的用户数据隐私保护问题,国内外互联网企业纷纷发布自研隐私计算服务,为产业界提供了高安全的数据保护方案。2024年6月,苹果公司在全球开发者大会上宣布推出私密云计算(private cloud compute, PCC)技术[42],旨在保证强大云端计算能力的同时,保护用户隐私,防止数据泄露和滥用。与传统云计算服务不同,私密云计算提供专用服务器和操作系统,确保数据"算后即焚",即服务器重启后所有数据自动清除。蚂蚁集团在2024年7月世界人工智能大会上发布了"隐语Cloud"大语言模型密算平台[43],通过软硬件结合的可信隐私计算技术,在大语言模型托管和大语言模型推理等环节实现数据密态流转,保护模型资产、数据安全和用户隐私(图3[43])。2024年10月中关村实验室、蚂蚁集团等联合发布了"星绽"操作系统和"星绽"机密计算两大项目[44],分别面向通用执行环境和可信执行环境提供安全原生的系统软件。

在隐私计算安全标准化方面,中国通信标准化协会已先后发布了3项针对隐私计算细分技术安全

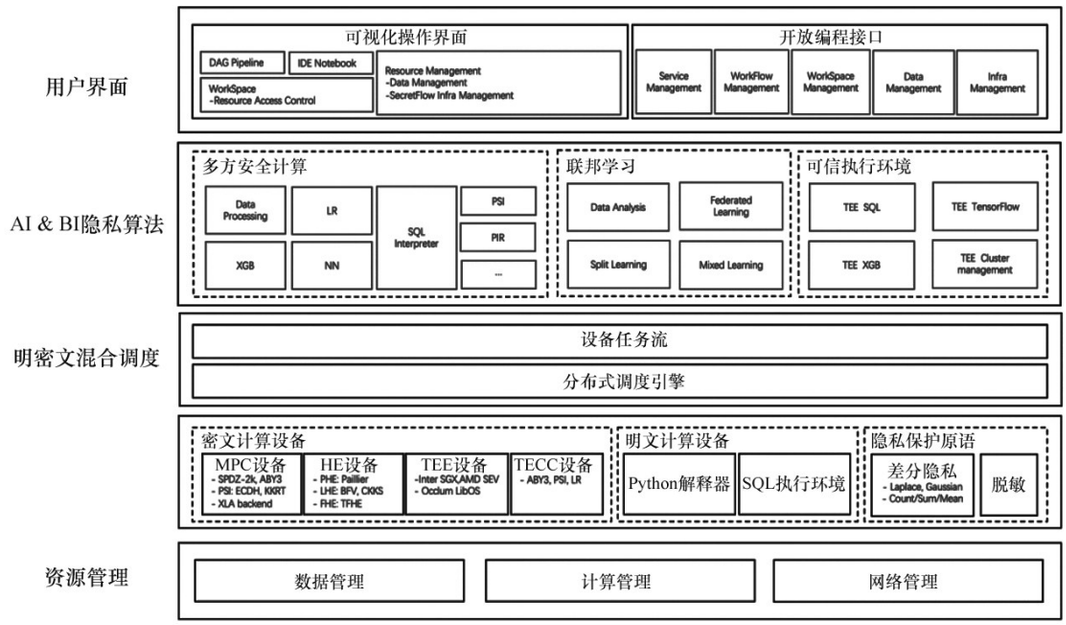

图3 蚂蚁集团隐私计算框架“隐语”的总体架构

的行业标准,包括YD/T 4690—2024《隐私计算 多方安全计算产品安全要求和测试方法》、YD/T 4691—2024《隐私计算 联邦学习产品安全要求和测试方法》、YD/T 4947—2024《隐私计算 可信执行环境产品安全要求》。为了进一步明确隐私计算产品的安全等级划分,2024年7月,中国通信标准化协会大数据技术标准推进委员会等多家单位共同发布《隐私计算产品通用安全分级白皮书(2024 年)》[45],指出了当前隐私计算产品在安全分级过程中面临的诸多挑战,并提出了通用安全分级的设计思路。2024年12月18日,中国网络空间安全协会发布了《隐私计算总体框架》(T/CSAC 005—2024)等6项团体标准[46]。

学术研究方面,研究人员提出了若干隐私计算的高效方案。在隐私保护密文检索技术方面,基于全同态加密的零隐私密文检索被越来越多的研究者关注。卡耐基梅隆大学研究者发现,现有所有基于同态加密的隐匿信息检索方案计算复杂度为线性,难以应对庞大数据库的快速计算需求,因此参考双服务器方案,提出了一种基于伪随机函数的单服务器隐匿信息检索方案PIANO[47]。里斯本新大学与Brave公司的研究者结合高效数据过滤器和FrodoPIR方案,提出了一种存储成本更低的关键词检索隐匿信息检索方案[48]。

2.2 机密计算

2024年10月,冯登国在2024年中国计算机大会(CNCC 2024)上阐述了机密计算发展历程与现状,并提出弹性机密计算概念与防护技术体系[49]。弹性机密计算旨在保障机密计算环境在面对各种攻击和漏洞时仍能保持安全性和可靠性,其核心思想是在可信执行环境(trusted execution environment, TEE)信任根攻击免疫的前提下,即使机密计算平台系统固件、虚拟机监控器、主机操作系统或部分TEE应用系统受到攻击或控制,整个机密计算系统依然能够保持其安全性和运行效率。弹性机密计算在现有的机密计算技术体系上,采用网络信任(CyberTrust)强化TEE内生安全能力,增强系统容错性、可恢复性等安全弹性,构建一个高安全性、高可信的计算环境。

弹性机密计算是围绕TEE的根信任,以及TEE信任的验证、扩展和保障建立的机密计算技术体系,其基本防护体系如图4[49]所示,包括4个安全防护层次,即硬件信任根(为系统提供最底层安全保障)、TEE基础防护(保护TEE初始环境)、软硬件协同防护(保护TEE系统的代码与关键数据)和可信状态监控(对TEE内部工作负载进行安全监控)。通过这4个层级的互相协作,弹性机密计算最终实现了启动时初始可信、计算时数据安全、互联时信道保护、运

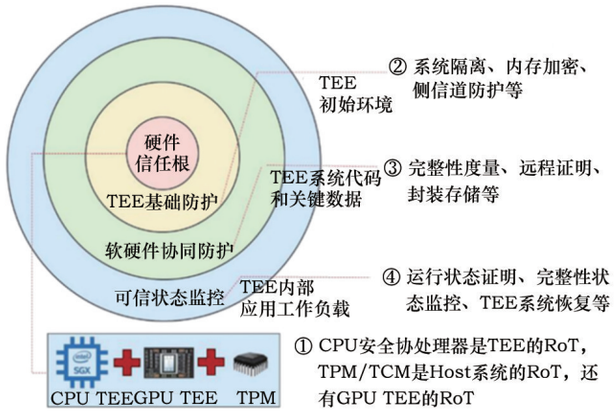

图4 弹性机密计算安全防护体系

行时动态监控的机密计算安全目标。

2024年,机密计算的形式化安全验证、机密AI、TEE运行时监控、异构计算单元的TEE等方面都取得了一系列进展和突破。

机密计算形式化安全验证方面,机密虚拟机是目前使用最广泛也最实用的机密计算环境之一,硬件厂商相继推出了各自的机密虚拟机架构,如AMD SEV-SNP、Intel TDX和ARM CCA,其虽然将虚拟机管理程序从TCB中移除了,但机密虚拟机内Guest镜像(特别是Guest内核)的代码量依然很大,为此AMD提出了SVSM技术在机密虚拟机内部进行进一步隔离。SVSM内隔离的安全模块为机密虚拟机环境提供安全功能,其安全性对于机密虚拟机架构而言至关重要。微软研究者在2024年提出了使用Rust实现的VeriSMo安全模块[50],并提供了对该安全模块功能的首次形式化验证,其验证涵盖了功能正确性、信息流安全性以及机密虚拟机本身的机密性和完整性,是机密虚拟机领域AMD SEV-SNP上第一个经过验证的机密虚拟机安全模块,该方案获得了2024年计算机系统领域会议OSDI的最佳论文奖,标志着机密计算在形式化验证上的一个突破。

机密AI方面,由于涉及大量不可信实体和复杂的软件/硬件基础设施,AI(特别是大语言模型)的广泛应用带来了新的攻击面,如模型窃取、成员推理等攻击。使用机密计算技术来保护AI的训练与推理已经成为一个新兴研究方向,主要包含3方面研究内容:1)分区执行。由于TEE环境受限,需要将复杂的AI模型进行切分,将部分过程放到TEE内执行;2)异构TEE。实现支持GPU的TEE,并设计CPU TEE与GPU TEE的协同计算,保障AI模型在CPU和GPU的计算安全;3)TEE辅助的AI计算。不完全信任TEE,在密码学-TEE混合计算安全模型下实现推理效率的提升和安全的加固。2024年9月阿里云宣布AI Infra全栈集成Confidential AI技术[51],是确保模型数据全生命周期安全的重要探索。

TEE运行时监控方面,机密计算技术基于硬件可信执行环境,通过隔离、完整性度量和远程证明等技术保障使用中数据的机密性和完整性,并免受特权敌手的攻击。然而,现有机密计算平台的完整性度量和远程证明机制主要针对启动时,而缺少运行时完整性保护,当用户工作负载潜在的内存漏洞被敌手利用时容易遭受控制流劫持等攻击。中国科学院软件研究所的研究人员针对该问题提出基于动态完整性度量的机密计算运行时监控方案,实现了机密计算平台内用户工作负载的运行时完整性保护[52]。

异构计算单元的TEE方面,当前各种异构计算单元在计算机中应用加速计算过程,例如GPU、DPU、NPU等,针对异构计算单元的TEE设计是近期的研究热点。2024年3月,英伟达发布Blackwell架构,进一步强化了机密计算能力,主要目的是为大语言模型提供高性能的安全性。为了支持边缘端设备的GPU TEE需求,南方科技大学研究者将统一内存GPU的工作流程与Arm CCA的系统设计结合,为Arm CCA的领域式架构(realm-style architecture)提供GPU机密计算支持,扩展了Arm CCA在异构计算支持方面的能力[53]。

3 人工智能已成为网络空间安全发展的关键变量

2024年人工智能技术继续高速发展,其中大语言模型技术的发展尤为瞩目。从聊天机器人、智能助理到包含多智能体和RAG技术的生成式AI系统,大语言模型已经应用于教育、医疗、金融等多个领域。随之而来的大语言模型安全和大语言模型赋能安全已成为工业界和学术界讨论的热点。

3.1 大语言模型安全

随着大语言模型的普及,安全问题和挑战日益

凸显。中关村实验室等研究者总结了大语言模型系统面临的12项风险及44项子类,如图5[54]所示。首先,LLM存在生成幻觉(Hallucinations)的风险,这种幻觉不仅表现在信息失真上,还可能导致模型输出不符合实际场景的误导性内容,从而在医疗诊断、金融咨询等关键领域产生严重后果,例如美联社2024年10月26日的报告[55]指出,OpenAI的语音转写工具Whisper出现严重的内容伪造问题,而Whisper已用于约700万次医疗就诊。其次,越狱(Jailbreak)攻击也是一个显著威胁,即通过特殊输入,诱导模型输出敏感、违法或危险信息。与此同时,隐私泄露问题日益严峻,大语言模型在处理用户数据时,可能无意中泄露个人敏感信息,如医疗记录、财务数据等,给用户隐私保护带来巨大压力。最后,生成模型引发的版权争议也愈发突出,特别是在艺术创作、文学作品等领域,大量训练数据和生成内容使版权归属变得模糊且难以界定,2024年国内外都出现了针对大模型训练侵权的诉讼[56]。

大语言模型同样面临软件供应链安全风险。将LLMs集成到现实世界应用中需要一系列开发和部署工具链,如数据处理(如用于数据质量保证的Cleanlab和用于数据管理的Hugging Face Datasets)、模型训练(如用于分布式训练的PyTorch Distributed)、优化(如用于模型量化的OmniQuant和用于模型合并的MergeKit)和部署(如用于Agent工作流编排的AutoGPT和用于检索增强生成的RAGFlow)。这些工具链的引入导致LLM应用开发、部署和维护的各个阶段都面临供应链风险。全球性安全组织OWASP已将供应链漏洞列入LLM应用10大安全威胁之一[57]。美国人工智能安全公司Lasso Security[58]研究发现,AI模型平台Hugging Face上存在API令牌漏洞,使攻击者能够访问谷歌、微软等公司的仓库,甚至窃取Meta的大语言模型,带来数据投毒和模型窃取等风险。

针对这些安全隐患,研究者已经开发并应用了一系列防御手段与检测方法。内容过滤技术是最常见的第一道防线,能够自动识别并屏蔽含有敏感、非法或虚假信息的内容,从而减少不当信息的传播,在

图5 大语言模型系统面临的12项风险及44项子类

一定程度上预防模型越狱攻击。对抗训练通过在模型训练过程中引入恶意输入样本,提升模型对异常情况的识别与应对能力,从而增强其鲁棒性。隐私保护技术方面,如差分隐私和联邦学习等方法,能够在不泄露用户数据的前提下提高模型性能,防止恶意攻击者对模型训练数据的窃取。此外,模型水印技术通过嵌入独特标识,使得模型来源验证变得可能,从而在版权保护方面发挥重要作用。实时监控和安全审计系统的引入,让企业能够更快速地发现和应对潜在风险,形成多层次的安全防护体系。2024年6月,OpenAI首次系统性地公布大语言模型开发安全方面的高级细节[59],包括基础架构、保护措施、敏感数据存储、开发人员访问管理等。2024年9月6日,世界数字技术院(WDTA)在外滩大会上正式发布国际标准《大语言模型供应链安全要求》[60],该标准由云安全联盟(CSA)大中华区联合蚂蚁集团、微软、谷歌、百度、Meta等数十家单位的专家共同编制。

在应对技术风险的同时,各国和国际组织也在加快对大语言模型的伦理研究与政策制定。中国颁布《生成式人工智能服务安全基本要求》,于2024年3月1日正式发布[61]。北京、上海、深圳等也出台大语言模型的区域发展政策,推动大语言模型在多个领域的有监管的应用落地。国际上,欧盟和美国也针对大语言模型的可信赖性、透明性颁布了多条法规与伦理准则。未来,随着大语言模型的普及,技术与伦理、政策的深度融合将成为推动安全发展的关键。

3.2 大语言模型赋能安全

随着信息系统复杂度的提升以及攻击手段不断演化,传统安全防护手段面临效率和适应性不足的问题。一方面,它们过度依赖专家知识,缺乏灵活性,这使得安全系统的反应速度较慢,难以快速应对复杂和快速变化的场景。另一方面,传统安全防护手段性能不足,容易产生误报和漏报,导致资源浪费并影响系统的有效性。大语言模型作为人工智能的重要技术进展,正在为网络安全带来革命性的变化。

凭借自然语言处理、代码理解和知识推理等优势能力,大语言模型在多种安全分析任务中取得了显著成果。开发了基于大语言模型的命令解释工具,浙江大学研究者利用其自然语言处理能力提升了对命令行为和意图的分析,增强了网络安全分析的自动化和标准化,有助于应对复杂的网络安全威胁。中国科学院信息工程研究所研究者利用大语言模型优化反编译器输出,减少了冗余信息和难以理解的变量,显著降低了逆向工程的认知负担,并提供了有意义的注释,提升了逆向分析效率。在漏洞挖掘方面,Google在2024年报告中指出,借助AI技术,Google实现了基于LLM的模糊测试对象生成与评估框架oss-fuzz-gen[62],能够为开源软件测试对象生成测试驱动,显著提高了OSS-Fuzz C/C++项目的测试覆盖率。2024年11月Google宣布,其开发的Big Sleep大模型辅助框架在SQLite开源数据库引擎中发现了一个零日漏洞,“是人工智能代理在广泛使用的现实软件中发现先前未知的可利用内存安全问题的第一个公开示例”[63]。在安全攻击实施方面,南洋理工大学研究者提出基于大语言模型的自动化渗透测试框架PentestGPT[64],通过3个自我交互模块分别处理渗透测试的不同任务,克服了上下文丢失等挑战,避免了传统渗透测试中过度依赖专业知识的问题。该框架在开源后12个月内获得超过6500个GitHub星标,在学术界和工业界受到广泛关注。

除了支持传统安全分析,大语言模型还为安全领域带来了新的发展范式。未来,大语言模型有望推动网络安全向更加智能化和自适应能力更强的方向发展。具体而言,大语言模型可以通过分析历史数据预测安全事件,识别潜在的攻击模式和漏洞风险,从而提供前瞻性防护。智能化安全运维助手将帮助运维人员快速获取安全报告和建议,提高决策效率。2024年4月,微软发布了Security Copilot[65],作为生成式AI安全解决方案,旨在支持安全专业人员进行事件响应、威胁搜寻、情报收集和态势管理。2024年9月,华清未央科技有限公司正式发布了面向机器语言模态的大语言模型(machine language model,MLM)[66],其基于大模型研发的产品矩阵涵盖逆向分析、代码转写与生态迁移、软件供应链分析、代码一致性检测、漏洞分析检测、恶意代码分析检测等诸多领域。

中国积极推动人工智能在网络安全方面的应用。2024年11月21日,世界互联网大会乌镇峰会网络安全技术发展与国际合作论坛以“智能向善,人工智能安全风险与治理”为主题,分享了国内外的经验

和最佳实践,探讨了各国各地区人工智能安全与治理的最新进展;发布了《人工智能赋能网络安全应急响应合作倡议》,以促进人工智能赋能网络安全应急能力建设,推动人工智能和网络安全融合创新[67]。

尽管大语言模型在安全领域的应用有很大前景,但目前仍存在限制和不足。林惠民在CNCC 2024上的报告[68]中指出,机器学习的概率特性会产生不可控问题,大语言模型可能带来虚假或错误的信息流传等问题。研究者也致力于研究模型的可解释性,以提高模型的透明性,相关进展有利于增强大语言模型在安全领域应用的实用性和安全性,最大化其效能并减少潜在负面影响。

4 量子化特征成为网络空间安全技术的新动力

量子信息技术的快速发展对网络空间安全技术带来了巨大的冲击,量子计算机计算能力的巨大潜力可能直接对现有安全技术(如算法、协议、方案)造成严重威胁,动摇其安全基础(如本原、困难问题)。为了应对这一挑战,政府、学术界和工业界积极推动后量子密码研究与标准制定,2024年各国纷纷部署了后量子密码迁移的方案。

4.1 量子计算机

2024年9月,IBM宣布,其位于美国纽约州波基普西的IBM量子数据中心的最新扩建工程已完成,该数据中心部署了IBM2023年底推出的IBM Quantum Heron处理器,成为全球单一地点运营数量最多的公用事业规模量子计算机[69]。2024年11月,IBM在量子开发者大会上宣布达成2年前所设定的100×100挑战(建立一个可以处理100个量子比特,并且进行100层量子操作的量子电路),并发布了具备5000个双量子比特闸操作能力的量子计算机[70]。

2024年12月10日,Google宣布了量子计算新芯片Willow[71],该芯片拥有105个量子比特。其主要实现了2项重大成就:随着量子比特的增加,Willow可以实现指数级的错误率降低,解决了量子纠错领域近30年来一直试图攻克的关键难题;在标准基准计算测试中,Willow也展示了非常高的性能。相关研究成果已发表于《Nature》。

中国在量子计算机方面也取得了系列进展。2024年1月6日,中国第三代自主超导量子计算机"本源悟空"上线运行。该量子计算机搭载72比特自主超导量子芯片"悟空芯"。2024年10月25日报道,中国科学家在量子计算机"本源悟空"上成功完成了量子计算流体动力学仿真[72]。

2024年4月中关村论坛重大成果专场中,"超大规模集成的光量子芯片"作为重大成果之一发布[73]。该成果由北京大学、中国科学院微电子研究所和浙江大学联合研发,被认为是量子计算机内核的关键技术突破。该芯片克服了大规模光量子芯片在设计、加工、调控和测量等方面的诸多难题。

4.2 后量子密码标准化与迁移

2024年10月24日,美国国家标准技术研究院(NIST)公布了抗量子密码标准化进程第二轮附加数字签名方案的候选算法[74]。NIST要求提交算法不能基于有代数结构的格密码假设,因此,公布的候选算法涵盖了6种不同技术路线下的14种算法,旨在为抗量子数字签名方案提供多样化的技术支持。入选算法包括基于编码的CROSS、LESS,基于同源的SQIsign,基于无代数结构格的HAWK,基于MPC-in-the-Head签名的Mirath、MQOM、PERK、RYDE、SDitH,基于多变量的MAYO、QR-UOV、SNOVA、UOV和基于对称密码的FAEST。通过引入多样化的技术路线,NIST希望进一步增强抗量子密码技术的灵活性和适应性,为未来全球向抗量子安全体系的迁移奠定坚实基础。

2024年11月12日,NIST发布的《Transition to post-quantum cryptography standards》报告,详细阐述了传统密码算法向后量子密码算法迁移的必要性、挑战和实施策略,以应对量子计算技术对现有加密系统的威胁(表1、表2)[75]。量子计算机能够通过"收集现在,解密未来"的方式破解目前广泛使用的公钥密码算法(如RSA和ECC),威胁长期敏感数据的安全性。因此,NIST制定的迁移计划目标是在2035年前完成美国联邦系统的全面迁移,并建议其他组织和行业尽快规划转型。报告明确了迁移的范围,重点包括数字签名算法(如RSA、ECDSA)和密钥建立方案(如Diffie-Hellman、RSA密钥传输),同时指出对称密码(如AES和SHA-2)由于量子攻击影响较小,

表1 易受量子攻击的数字签名算法

数字签名算法	安全强度	弃用时间	禁用时间
ECDSA	112比特	2030年	2035年
[FIPS186]	≥128比特	—	2035年
EdDSA	≥128比特	—	2035年
[FIPS186]			2035年
RSA	112比特	2030年	2035年
[FIPS186]	≥128比特	—	2035年

表2 易受量子攻击的密钥建立方案

密钥建立方案	安全强度	弃用时间	禁用时间
有限域DH和MQV	112比特	2030年	2035年
[SP80056A]	≥128比特	—	2035年
椭圆曲线DH和MQV	112比特	2030年	2035年
[SP80056A]	≥128比特	—	2035年
RSA	112比特	2030年	2035年
[SP80056A]	≥128比特	—	2035年

暂不需要立即替换。为确保平稳过渡,NIST建议在迁移过程中采用混合模式,即同时运行传统算法和后量子算法(如混合签名和密钥封装方案),以平衡安全性和兼容性。报告还提出了迁移的优先事项和时间表,明确声明在2035年全面淘汰RSA和椭圆曲线的经典密码算法,并强调应优先保护交互式协议(如TLS、IKE)免受量子攻击,同时升级PKI、网络安全协议以及文档、邮件加密等关键应用场景。NIST已公布的后量子密码算法标准有CRYSTALS-Dilithium(数字签名)、CRYSTALS-KYBER(密钥封装机制)和SPHINCS+(哈希签名),并根据对称密码的安全强度定义了5个量子安全类别(Category 1-5),以评估其安全性。然而,迁移过程面临诸多挑战,包括算法性能下降、系统兼容性问题和实施成本增加。报告建议组织提前规划迁移,优先保护敏感数据,采用混合模式作为过渡方案,并加强与行业标准组织和技术提供商的协作,共同推动后量子密码技术的普及应用。通过科学规划和逐步实施,组织可以在量子计算威胁到来前确保其信息系统的长期安全性。

2024年4月11日,欧盟委员会发布《向后量子密码迁移的协同实施路线图建议》[76],鼓励成员国制定统一战略,确保不同成员国及其公共部门之间向后量子密码的协调和同步迁移,包括明确的目标、关键里程碑和时间表,以实现保护欧盟公共管理部门数字基础设施及其他关键基础设施服务的目的。2024年10月,在瑞典斯德哥尔摩举行的ISO/IEC JTC1/SC6(系统间远程通信和信息交换)会议上,中国专家就如何设计抗量子攻击的通信网络安全协议提交提案并获会议一致通过,会议决议成立预备工作项目,由中国专家牵头推进制定协议设计指南[77]。

目前,密码敏捷性(cryptographic agility)被认为是后量子密码迁移的核心,并被认为是一门新的科学,值得进一步系统和深入研究。

5 软件漏洞依然是网络空间安全的重大威胁

软件作为信息技术关键载体,其安全性一直是现实网络攻击关注的重点。一方面软件系统的安全性经过多年的发展取得了显著提高,另一方面随着程序分析、人工智能等技术的发展,漏洞挖掘与利用的能力也得到提高。攻与防的博弈是软件安全方向经久不衰的主题。

5.1 漏洞挖掘

零日漏洞是指开发者不知道细节的漏洞,通常没有对应的补丁,因此往往具有很大的破坏性。模糊测试(fuzzing)是目前挖掘零日漏洞的主要方法。Google于2016年发起了OSS-Fuzz项目以应对著名的HeartBleed漏洞,旨在针对开源项目进行模糊测试,并向开发人员发出检测到的错误警报。截至2024年9月,OSS-Fuzz集群在超过1000个项目中发现了超过12000个错误[78],当前已成为开源社区的一项关键服务,支持C/C++、Go、Rust、Python等不同编程语言。除了针对开源软件,模糊测试已成为各大公司重要的安全测试手段,深度整合到持续集成/持续部署(CI/CD)管道中,使得每一次代码变更都能自动触发相应的安全检查,能够在早期阶段及时发现并修复潜在的安全风险。Google内部利用ClusterFuzz作为OSS-Fuzz的后端,对其产品(如Chrome、Android)进行模糊测试。国内众多行业场景中都进行了模糊测试工具的应用,例如智能网联汽车领域中通过协议模糊测试对WI-FI、蓝牙、NFC、CAN、V2X、充电桩协议等进行自动化测试;工业互联网领域中对工业控制系统的通信协议、外部接口、设备固

件和管理软件等测试;智能合约领域中模拟攻击者可能使用的输入,测试智能合约在面对这些输入时的反应。

1day/nday漏洞是指已经被公开的漏洞,然而公开并不意味着这些漏洞是没有威胁的,因为存在有相应的补丁但是未被及时部署的情况,还存在软件供应链场景中的组件漏洞未在软件系统中修复的情况。因此,这方面的研究涉及到代码相似性分析、软件成分分析、补丁部署检测等,国内外开展了大量研究工作,提出了BinaryAI[79]、CI-Detector[80]、δCFG[81]、CEBin[82]、HermesSim[83]等新的方法。

AI/LLM赋能的漏洞挖掘已成为研究热点。2024年出现了众多学术成果和开源工具,主要包括3个思路。一是利用大语言模型对程序功能的分析能力,辅助模糊测试中目标选择、测试种子生成、测试驱动生成等过程;二是利用AI/LLM在代码特征检测方面的能力,找到具有相似性或某种特征的漏洞,例如,根据函数库中报出的漏洞代码在其他软件中寻找是否存在相似的漏洞代码;三是利用大语言模型对人类行为的理解能力和与人的交互能力,解决软件需要交互过程才能触发的问题,例如在GUI测试中,利用大语言模型生成具有意义的复杂点击行为,并按照设计要求填写表格数据等。

传统程序分析技术在2024年也取得了进步,并进一步推动了漏洞挖掘方法的发展。污点传播、符号执行等程序分析技术很早就应用于漏洞挖掘,然而相关技术受限于分析能力和分析性能等问题,诸多方案停留在原型系统阶段。2024年,国内外研究者在相关基础程序分析方法方面实现了突破,例如AirTaint[84]、HardTaint[85]等污点传播改进工作,TACE[86]、SymFit[87]等符号执行改进工作。依靠传统程序分析技术的路线和基于AI/LLM的创新路线并不矛盾,相反,两者是互补的,当前AI/LLM增强的程序分析和程序分析增强的AI/LLM方案都表现出了更好的效果。

5.2 漏洞攻击及其防御

随着黑客技术的更新和黑客工具的传播,漏洞利用攻击威胁日益增大,给漏洞防御带来了巨大挑战。

在网络攻击中,0click漏洞、硬件漏洞等攻击技术不断发展和应用。0click漏洞是指攻击者无需受害者进行任何点击操作或其他交互动作,就能利用系统或软件中的漏洞来实现攻击目的。2024年初,安全专家演示了如何利用3个漏洞在未经用户确认的情况下诱骗蓝牙主机状态,并成功地配对一个假键盘并注入攻击代码以获取受害者的身份执行代码,该0click蓝牙漏洞影响苹果iOS和macOS、谷歌安卓以及微软Windows系统[88]。2024年8月,国内赛博昆仑公司的昆仑实验室安全研究员发现了一个由Windows TCP/IP协议栈IPv6相关数据处理代码整数下溢引起的堆栈溢出,导致可能无需认证就可以远程执行恶意代码[89]。2024年11月报道显示,俄罗斯黑客正在利用Mozilla Firefox浏览器和Windows系统的零日漏洞,实施针对Windows用户的0click漏洞利用攻击[90]。此外,2024年研究人员也多次展示了如何利用CPU、GPU等硬件单元漏洞进行密钥泄露等攻击,这些硬件漏洞仍存在隐藏深、修复难等技术难题。

为了防御高级漏洞利用技术,各大公司、政府部门积极推动漏洞悬赏、加强漏洞管理。微软在2024年的漏洞赏金计划金额增加到了1660万美元[91],颁给了来自55个国家的343名安全研究者,涵盖包括Azure、Microsoft Identity、Edge、Windows和Office 365等众多产品。微软还专门针对人工智能推出了一项新的漏洞赏金计划,为涉及其Copilot人工智能技术的漏洞报告提供高达1.5万美元的奖金。中国华为技术有限公司也开展了针对其终端、IoT产品和鸿蒙产品的漏洞奖励计划[92]。国家层面,美国漏洞数据库NVD、中国的3大漏洞库CNVD、CNNVD、NVDB等,不断提高数据库的覆盖面和数据质量,并部署了大规模软件缺陷库构建与应用等国家项目,推动漏洞管理与安全治理的进步。

为了防御供应链相关漏洞攻击,2024年许多大型企业和云服务提供商在软件供应链安全中采用了Zero Trust(零信任)原则与架构。这种架构通过严格的身份验证、最小权限访问和动态访问控制,限制了恶意软件和攻击者在供应链中的横向传播,增强了企业对软件开发流程的安全控制。2024年11月,美国政府发布了《联邦零信任数据安全指南》[93]。另一方面,随着AI和机器学习技术的成熟,2024年也出现了多个由AI驱动的软件成分分析工具和供应链风险识别工具,这些工具可以自动化地分析依赖关系,

发现潜在的安全漏洞和威胁。中国科学院软件研究所持续建设“源图”开源软件基础设施[94],为建设安全可靠的软件供应链体系提供有效支撑,保证软件行业健康发展。

为了研究如何利用大语言模型等AI技术实现漏洞攻防,美国DARPA部署了人工智能网络挑战(Artificial Intelligence Cyber Challenge,AIxCC)[95],探索利用人工智能发现漏洞、分析漏洞和利用漏洞,进而如何更好地修复漏洞(图6[95])。AIxCC半决赛于2024年8月在拉斯维加斯的DEFCON 2024会议上举行,近40个团队研发,基于真实世界开源项目进行了测试,例如Jenkins、Linux 内核等。AIxCC决赛将于2025年DEF CON大赛期间举行。

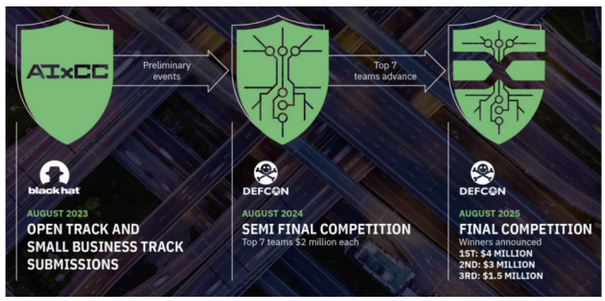

图6 DARPA AIxCC项目计划

6 结论

当前网络空间中,传统的安全问题仍形势严峻,多项重要难题有待解决。而新的应用场景和攻防场景又催生了新的安全需求,技术革新带来了新的安全挑战。2024年网络空间安全在数据安全、人工智能、量子密码、软件漏洞等方面均产生了众多热点话题和技术突破。未来,随着应用的发展,新技术、新场景的出现,网络空间安全形势仍然严峻,亟需发展新的技术能力和技术体系,以应对层出不穷的各类新型威胁,为构建安全、健康的网络空间环境提供技术保障。因此,针对国家战略需求和国际学科发展前沿,明确重点研究方向,着眼于关键交叉节点组织创新,才能抢占世界科技发展的制高点,维护国家网络空间安全。

致谢:张振峰研究员、陈恺研究员、秦宇研究员、张敏研究员、陈隆副研究员、李昊副研究员、冯伟副研究员、贾相堃副研究员、赵月副研究员、胡洁工程师等为文章撰写提供相关素材。

参考文献(References)

[1] Joseph Clark, DOD News. DOD Releases strategy to bolster cybersecurity across industrial base[EB/OL]. [2024-12-24]. https://www. defense. gov/News/News-Stories/Article/Article/3724118/dod-releases-strategy-to-bolster-cybersecurity-across-industrial-base/.

[2] Building digital solidarity: The United States International Cyberspace & Digital Policy Strategy-United States Department of State[EB/OL]. (2024-07-18)[2024-12-24]. https://www.state.gov/building-digital-solidarity-the-united-states-international-cyberspace-and-digital-policy-strategy/.

[3] DOE Leads Effort to Improve the cybersecurity of energy supply Chains[EB/OL]. [2024-12-24]. https://www. energy. gov/articles/doe-leads-effort-improve-cybersecurity-energy-supply-chains.

[4] Guidance: Framing software component transparency: Establishing a common software bill of materials[EB/OL]. [2024-12-28]. https://www.cisa.gov/news-events/alerts/2024/10/15/guidance-framing-software-component-transparency-establishing-common-software-bill-materials-sbom.

[5] New rules to boost cybersecurity of the EU institutions enter into force[EB/OL]. [2024-12-24]. https://ec.europa.eu/commission/presscorner/detail/en/ip_23_6782.

[6] First EU-wide cybersecurity certification scheme to make European digital space safer[EB/OL]. (2024-01-31)[2024-12-24]. https://digital-strategy. ec. europa. eu/en/news/first-eu-wide-cybersecurity-certification-scheme-make-europe-an-digital-space-safer.

[7] European Cyber Resilience Act(CRA)-Regulation EU 2024/2847[EB/OL]. [2024-12-24]. https://www. european-cyber-resilience-act.com/.

[8] Guidelines for secure AI system development[EB/OL]. [2024-12-24]. https://www.ncsc.gov.uk/collection/guidelines-secure-ai-system-development.

[9] 简报:拜登-哈里斯政府阐明协调性方针以利用人工智能之力促进美国国家安全[EB/OL]. [2024-12-24]. https://www. state. gov/translations/chinese/20241024-fact-sheet_ -biden- _harris-administration-outlines-coordinated-approach-to-harness-power-of-ai-for-u-s-national-security-chinese/.

[10] Framework to advance AI governance and risk management in national security[R/OL]. [2024-12-24]. https://ai.gov/wp-content/uploads/2024/10/NSM-Framework-to-Advance-AI-Governance-and-Risk-Management-in-National-Security.pdf.

[11] Engaging with Artificial Intelligence[EB/OL]. [2024-12-

24]. https://www.cyber.gov.au/resources-business-and-government/governance-and-user-education/artificial-intelligence/engaging-with-artificial-intelligence.

[12] Guidelines and companion guide on securing AI systems [EB/OL]. [2024-12-24]. https://www.csa.gov.sg/Tips-Resource/publications/2024/guidelines-on-securing-ai.

[13] 工业和信息化部关于印发工业控制系统网络安全防护指南的通知[EB/OL]. [2024-12-24]. https://www.gov.cn/zhengce/zhengceku/202402/content_6929643.htm.

[14] 四部门制定《互联网政务应用安全管理规定》[EB/OL]. [2024-12-24]. https://www.gov.cn/lianbo/bumen/202405/content_6952940.htm.

[15] 《电力监控系统安全防护规定》2024年第27号令[EB/OL]. [2024-12-24]. https://zfxxgk.nea.gov.cn/2024-12/12/c_1310787545.htm.

[16] 十七部门关于印发《"数据要素×"三年行动计划(2024—2026年)》的通知[EB/OL]. [2024-12-24]. https://www.cac.gov.cn/2024-01/05/c_1706119078060945.htm.

[17] 公布《网络数据安全管理条例》[EB/OL]. [2024-12-24]. https://www.moj.gov.cn/pub/sfbgw/gwxw/xwyw/202409/t20240930_507076.html.

[18] 工业和信息化部关于印发《工业和信息化领域数据安全事件应急预案(试行)》的通知[EB/OL]. [2024-12-24]. https://www.gov.cn/zhengce/zhengceku/202411/content_6984322.htm.

[19] 习近平出席亚太经合组织第三十一次领导人非正式会议并发表重要讲话[EB/OL]. [2024-12-24]. https://www.mfa.gov.cn/zyxw/202411/t20241117_11527668.shtml.

[20] 中共中央关于进一步全面深化改革 推进中国式现代化的决定[EB/OL]. [2024-12-24]. https://www.gov.cn/zhengce/202407/content_6963770.htm.

[21] 11月1日起,13项网络安全国家标准开始实施[EB/OL]. [2024-12-24]. https://www.chinanews.com.cn/cj/2024/10-31/10311099.shtml.

[22] TC260-003《生成式人工智能服务安全基本要求》发布[EB/OL]. [2024-12-24]. https://www.tc260.org.cn/front/postDetail.html?id=20240301164054.

[23] 《人工智能安全治理框架》1.0版发布[EB/OL]. [2024-12-24]. https://www.cac.gov.cn/2024-09/09/c_17275678 8619-9789.htm.

[24] Michael Skelton. Supply Chain Backdoors, xz/liblzma, CVE-2024-3094, and what we currently know[EB/OL]. [2024-12-24]. https://www.bugcrowd.com/blog/supply-chain-backdoors-xz-liblzma-cve-2024-3094-and-what-we-currently-know/.

[25] 2024 CrowdStrike-related IT outages: Wikipedia[EB/OL]. [2024-12-30]. https://en.wikipedia.org/wiki/2024_CrowdStrike-related_IT_outages.

[26] 开发者被Linux大清洗!俄罗斯宣告建立独立Linux开发社区_腾讯新闻[EB/OL]. (2024-10-30)[2024-12-24]. https://news.qq.com/rain/a/20241030A07WB100.

[27] 2024 Lebanon electronic device attacks-Wikipedia[EB/OL]. [2024-12-30]. https://en.wikipedia.org/wiki/2024_Lebanon_electronic_device_attacks.

[28] 奇安信威胁情报中心. 网络安全威胁2024年中报告[EB/OL]. [2024-12-30]. https://www.qianxin.com/threat/reportdetail?report_id=317.

[29] StopRansomware: Black Basta | CISA[EB/OL]. [2024-12-24]. https://www.cisa.gov/news- events/cybersecurity-advisories/aa24-131a.

[30] Searchlight Cyber. Ransomware in H1 2024 trends from the dark web[EB/OL]. [2024-12-30]. https://slcyber.io/whitepapers-reports/ransomware-in-h1-2024-trends-from-the-dark-web/.

[31] Zscaler. ThreatLabz 2024_Ransomware report[EB/OL]. [2024-12-30]. https://www.zscaler.com/resources/industry-reports/threatlabz-ransomware-report.pdf.

[32] Alder S. Nebraska sues change healthcare over february ransomware attack[EB/OL]. [2024-12-24]. https://www.hipaajournal.com/change-healthcare-responding-to-cyberattack/.

[33] Matos G. Cencora pays $75 million in Bitcoin in the largest known case of ransomware attack[EB/OL]. (2024-09-18) [2024-12-24]. https://cryptobriefing.com/cencora-bitcoin-ransom-payment/.

[34] 几乎所有客户被波及!美国电信巨头AT&T再曝重大数据泄漏事故[EB/OL]. [2024-12-30]. https://www.qianxin.com/news/detail?news_id=12226.

[35] Muncaster P. Vodka giant stoli files for bankruptcy after ransomware attack[EB/OL]. [2024-12-24]. https://www.infosecurity-magazine.com/news/vodka-stoli-bankruptcy-ransomware/.

[36] Mihir Bagwe. Russia's 2024 cyber offensive strategy favors espionage[EB/OL]. (2024-09-23)[2024-12-24]. https://thecyberexpress.com/russia-h1-2024-cyber-offensive-strategy/.

[37] Paganini P. Pro-Ukraine hackers wiped 2 petabytes of data from Russian research center[EB/OL]. (2024-01-27) [2024-12-24]. https://securityaffairs.com/158214/hacktivism/ukraines-ministry-of-defense-hit-russian-recent-center.html.

[38] Toulas B. Ukraine claims it hacked Russian Ministry of Defense servers[EB/OL]. (2024-03-04)[2024-12-24]. https://www.bleepingcomputer.com/news/security/ukraine-claims-

it-hacked-russian-ministry-of-defense-servers/.
[39] 消息人士称乌克兰情报部门攻击俄罗斯800多台服务器[EB/OL]. [2024-12-24]. https://news.cctv.com/2024/09/27/ARTIwgfCbzO4yXs8zTT4gRCS240927.shtml.
[40] Ukraine's defense ministry launches military CERT to counter Russian cyberattacks[EB/OL]. [2024-12-24]. https://therecord.media/ukraine-creates-military-cert.
[41] Helmus T C, Khrystyna H. Ukrainian resistance to Russian disinformation: Lessons for future conflict[EB/OL]. [2024-12-24]. https://www.rand.org/pubs/research_reports/RRA2771-1.html.
[42] Apple launches private cloud compute for privacy-centric AI Processing[EB/OL]. (2024-06-11)[2024-12-24]. https://thehackernews.com/2024/06/apple-integrates-openais-chatgpt-into.html.
[43] 机器之心. 蚂蚁集团开源可信隐私计算框架"隐语":开放、通用[EB/OL]. [2024-12-24]. https://news.qq.com/rain/a/20220705A048I600.
[44] 中关村实验室、蚂蚁等联合发布,"星绽"操作系统内核开源_腾讯新闻[EB/OL]. (2024-10-23)[2024-12-24]. https://news.qq.com/rain/a/20241023A03MAA00.
[45] 隐私计算产品 通用安全分级白皮书(2024年)[R/OL]. [2024-12-24]. https://www.shujiaowang.cn/uploads/20240923/b28ba9aa46dae72fb999fbb0fb01919a.pdf.
[46] 中国网络空间安全协会发布六项隐私计算系列团体标准[EB/OL]. [2024-12-24]. https://www.secrss.com/articles/73741.
[47] Zhou M, Park A, Zheng W, et al. Piano: Extremely simple, single-server PIR with sublinear server computation[C]// 2024 IEEE Symposium on Security and Privacy (SP). Oakland: IEEE, 2024: 4296-4314.
[48] Celi S, Davidson A. Call me by my name: Simple, practical private information retrieval for keyword queries[C]// Proceedings of the 2024 on ACM SIGSAC Conference on Computer and Communications Security. Salt lake: ACM, 2024: 4107-4121.
[49] 冯登国. 机密计算:进展与展望——CNCC2024 特邀报告[EB/OL]. [2024-12-24]. https://dl.ccf.org.cn/video/videoDetail.html?_ack=1&id=7227251887065088.
[50] Zhou Z Q, An J L, Chen W T, et al. VeriSMo: A verified security module for confidential vms. in 18th usenix symposium on operating systems design and implementation [R]. OSDI, Santa Clara, 2024.
[51] AI场景下确保模型数据安全,Confidential AI技术最佳实践解读-阿里云开发者社区[EB/OL]. (2024-11-06)[2024-12-24]. https://developer.aliyun.com/article/1634420.
[52] 李为, 冯伟, 秦宇, 等. 基于动态完整性度量的机密计算运行时监控方案[J]. 计算机研究与发展, 2024, 61(10): 2482-2500.
[53] Wang C, Zhang F, Deng Y, et al. CAGE: Complementing arm CCA with GPU Extensions[C]//Network and Distributed System Security (NDSS) Symposium. San Diego: NDSS, 2024.
[54] Cui T, Wang Y, Fu C, et al. Risk taxonomy, mitigation, and assessment benchmarks of large language model systems[J]. 2024, arXiv preprint arXiv: 2401.05778.
[55] OpenAI's Whisper experiencing"AI Hallucinations"despite high-risk applications[EB/OL]. [2024-12-24]. https://www.pcmag.com/news/openais-whisper-experiencing-ai-hallucinations-despite-high-risk-applications.
[56] AI生成版权纠纷涌现 著作权保护难题待解[EB/OL]. (2024-12-06)[2024-12-24]. https://finance.sina.com.cn/jjxw/2024-12-07/doc-incyqiwa8351891.shtml.
[57] Wang S, Zhao Y, Hou X, et al. Large language model supply chain: A research agenda[J]. ACM Transactions on Software Engineering and Methodology, 2024.
[58] +1500 HuggingFace API Tokens were exposed, leaving millions of Meta-Llama, Bloom, and Pythia users vulnerable [EB/OL]. [2024-12-24]. https://www. lasso. security/blog/1500-huggingface-api-tokens-were-exposed-leaving-mil - lions-of-meta-llama-bloom-and-pythia-users-for-supply-chain-attacks.
[59] Securing research infrastructure for advanced AI[EB/OL]. [2024-12-24]. https://openai.com/index/securing-research-infrastructure-for-advanced-ai/.
[60] WDTA发布《大模型供应链安全要求》推进AI安全可信负责任发展-新华丝路[EB/OL]. [2024-12-24]. https://www.imsilkroad.com/news/p/526784.html.
[61]《生成式人工智能服务安全基本要求》发布[EB/OL]. [2024-12-24]. https://www.secrss.com/articles/64121.
[62] google/oss-fuzz-gen: LLM powered fuzzing via OSS-Fuzz [EB/OL]. [2024-12-24]. https://github.com/google/oss-fuzz-gen.
[63] Project Zero. From naptime to big sleep: Using large language models to catch vulnerabilities in real-world code [EB/OL]. [2024-12-24]. https://googleprojectzero. blogspot.com/2024/10/from-naptime-to-big-sleep.html.
[64] GreyDGL/PentestGPT: A GPT-empowered penetration testing tool[EB/OL]. (2024-05-15)[2024-12-24]. https://github.com/GreyDGL/PentestGPT.
[65] Microsoft security copilot-microsoft adoption[EB/OL]. (2024-11-18) [2024-12-24]. https://adoption. microsoft. com/zh-cn/security-copilot/.
[66] 华清未央 MLM 机器语言大模型全球首发,开启软件智

能化时代[EB/OL]. (2024-09-09)[2024-12-24]. https://finance.sina.com.cn/tech/roll/2024-09-09/doc-incnpptn2855893.shtml.

[67] 2024年世界互联网大会乌镇峰会网络安全技术发展与国际合作论坛举行[EB/OL]. [2024-12-24]. https://www.cert.org.cn/publish/main/12/2024/2024112516152019 2572249/20241125161520192572249_.html.

[68] 林惠民. 计算 智能 安全: CNCC2024 特邀报告[EB/OL]. [2024-12-24]. https://dl.ccf.org.cn/video/videoDetail.html?_ack=1&id=7225805846317056.

[69] IBM Expands quantum data center in poughkeepsie, new york to advance algorithm discovery globally[EB/OL]. [2024-12-28]. https://newsroom.ibm.com/2024-09-26-ibm-expands-quantum-data-center-in-poughkeepsie,-new-york-to-advance-algorithm-discovery-globally.

[70] IBM Quantum delivers on 2022 100×100 performance challenge | IBM quantum computing blog[EB/OL]. [2024-12-28]. https://www.ibm.com/quantum/blog/qdc-2024.

[71] 谷歌推出突破性量子芯片[EB/OL]. [2024-12-20]. https://xinhuanet.com/20241211/6fee8a80e09c42d09120cb36c6d805c2/c.html.

[72] 科技日报. 全球最大规模量子计算流体动力学仿真完成[EB/OL]. http://kpzg.people.com.cn/n1/2024/1029/c404214-40349440.html.

[73] 就这一个字!"芯"——你见过量子计算机的内核吗[EB/OL]. [2024-12-24]. https://www.news.cn/tech/20240430/5ecf34fd9a3e4a3fa939c83c86a3e155/c.html.

[74] Post-quantum cryptography: Additional digital signature schemes | CSRC[EB/OL]. [2024-12-24]. https://csrc.nist.gov/Projects/pqc-dig-sig/round-2-additional-signatures.

[75] Moody D, Perlner R, Regenscheid A, et al. Transition to post-quantum cryptography standards[R]. National Institute of Standards and Technology, 2024.

[76] Recommendation on a coordinated implementation roadmap for the transition to post-quantum cryptography[EB/OL]. (2024-04-11) [2024-12-24]. https://digital-strategy.ec.europa.eu/en/library/recommendation-coordinated-implementation-roadmap-transition-post-quantum-cryptography.

[77] 中国将牵头制定抗量子攻击的通信网络安全协议设计指南[EB/OL]. [2024-12-24]. http://finance.people.com.cn/n1/2024/1028/c1004-40348899.html.

[78] google/oss-fuzz: OSS-Fuzz-continuous fuzzing for open source software[EB/OL]. [2024-12-24]. https://github.com/google/oss-fuzz.

[79] Jiang L, An J, Huang H, et al. BinaryAI: Binary software composition analysis via intelligent binary source code matching[C]//Proceedings of the IEEE/ACM 46th International Conference on Software Engineering, 2024: 1-13.

[80] Jia A, Fan M, Xu X, et al. Cross-inlining binary function similarity detection[C]//Proceedings of the IEEE/ACM 46th International Conference on Software Engineering. New York: ACM, 2024: 1-13.

[81] Wang J, Zhang C, Chen L, et al. Improving ML-based binary function similarity detection by assessing and deprioritizing control flow graph features[C]//33rd USENIX Security Symposium (USENIX Security 24). Philadelphia: USENIX, 2024: 4265-4282.

[82] Wang H, Gao Z, Zhang C, et al. CEBin: A cost-effective framework for large-scale binary code similarity detection [C]//Proceedings of the 33rd ACM SIGSOFT International Symposium on Software Testing and Analysis. Xian: ACM, 2024: 149-161.

[83] He H, Lin X, Weng Z, et al. Code is not natural language: Unlock the power of semantics-oriented graph representation for binary code similarity detection[C]//33rd USENIX Security Symposium (USENIX Security 24). Philadelphia: USENIX, 2024.

[84] Sang Q, Wang Y, Liu Y, et al. Airtaint: Making dynamic taint analysis faster and easier[C]//2024 IEEE Symposium on Security and Privacy (SP). San Francisco: IEEE, 2024: 3998-4014.

[85] Zhang Y, Liu T, Wang Y, et al. HardTaint: Production-run dynamic taint analysis via selective hardware tracing[J]. Proceedings of the ACM on Programming Languages, 2024, 8(OOPSLA2): 1615-1640.

[86] Jain R, Tihanyi N, Ndhlovu M, et al. Rapid taint assisted concolic execution (TACE)[C]//Companion Proceedings of the 32nd ACM International Conference on the Foundations of Software Engineering. New York: ACM, 2024: 627-631.

[87] Qi Z, Hu J, Xiao Z, et al. SymFit: Making the common (concrete) case fast for binary-code concolic execution[C]//33rd USENIX Security Symposium (USENIX Security 24). Philadelphia: USENIX, 2024: 415-432.

[88] 安卓手机"最受伤",专家演示零点击蓝牙漏洞攻击力[EB/OL]. (2024-01-25) [2024-12-24]. https://www.163.com/dy/article/IP9UNLR40511B8LM.html.

[89] 我国昆仑实验室发现! 微软Win10/Win11被曝9.8分漏洞:影响所有IPv6系统[EB/OL]. (2024-08-16)[2024-12-24]. https://finance.sina.com.cn/tech/discovery/2024-08-16/doc-inciuumq6022053.shtml.

[90] ESET Research discovers Mozilla and Windows zero day & zero click vulnerabilities exploited by Russia-aligned

RomCom APT group[EB/OL]. [2024-12-28]. https://www.eset.com/int/about/newsroom/press-releases/research/eset-research-discovers-mozilla-and-windows-zero-day-zero-click-vulnerabilities-exploited-by-russia-aligned-rom-com-apt-group/.
[91] 微软2024财年发放了约1.2亿元漏洞赏金:平均每个漏洞8.6万元[EB/OL]. [2024-12-24]. https://www.secrss.com/articles/69003.
[92] 华为安全奖励计划[EB/OL]. [2024-12-24]. https://bugbounty.huawei.com/#/home.
[93] Federal data, security leaders release zero-trust guide ahead of White House deadline[EB/OL]. [2024-12-30]. https://fedscoop.com/zero-trust-guide-federal-ciso-cdo-councils/.
[94] 2024中国数交会"软件促进数实融合支撑新型工业化论坛"举办-新华网[EB/OL]. [2024-12-28]. https://www.xinhuanet.com/tech/20241203/6f7224025fed4e2e9043d787f45e041f/c.html.
[95] AI Cyber Challenge[EB/OL]. [2024-12-24]. https://aicyberchallenge.com/.

Review of cybersecurity technology hotspots in 2024

SU Purui, FENG Dengguo*

Institute of Software, Chinese Academy of Sciences, Beijing 100190, China

Abstract In 2024, countries around the world continued to increase their deployment of cybersecurity strategies, issuing relevant technical guidelines or management policies for security issues in emerging scenarios such as software supply chains and big language models. China also further improved its policies and regulations on cybersecurity in response to new situations, new problems, and planned and guided constructions of cybersecurity systems, and regulated the development of the cybersecurity industry. In 2024, researchers achieved a series of breakthroughs around hot topics such as data security, artificial intelligence security, quantum computing, and software vulnerabilities, which were expected to further enhance the ability and level of cybersecurity governance. However, current cyber attacks continue to occur frequently, with APT attacks, ransomware attacks, and other cyber attacks posing a serious threat to global cyberspace security. In the future, it will be urgently necessary to develop new technical capabilities and technical systems to provide support for building a safe and healthy cyberspace environment.

Keywords cyberspace security; data security; artificial intelligence security; quantum computing; software vulnerabilities

（责任编辑　卫夏雯）