DOI 10.15302/J-SSCAE-2023.06.009

多视角下的网络空间安全模型与体系化发展

孙帅[1]，张蕾[1*]，胡春卉[1]，傅首清[1]，卿昱[1]，崔勇[2]

（1. 中关村实验室，北京 100094；2. 清华大学计算机科学与技术系，北京 100084）

摘要：网络空间技术发展迅速，新应用、新技术衍生的网络空间安全风险趋于复杂化和隐蔽化；建立特有的网络空间安全模型是国内外应对复杂安全威胁的常规做法，但现有的网络空间安全模型存在未来发展方向不明确、新技术衍生风险分析能力不足、网络空间安全防御评估所需的安全能力缺失等问题。本文从技术、学科、产业等视角入手，开展了现有的网络空间安全模型评估，梳理了网络空间安全技术体系的自身特点及其发展脉络，阐明了网络安全应用存在的迫切问题；着重从网络空间安全技术视角出发，提出了一种基于技术要素的网络空间安全模型体系框架，并利用已有安全技术和新兴技术验证了体系框架的安全分析能力。从完善网络空间安全技术体系核心框架、促进网络空间安全领域“产学研”协同、推进网络空间安全核心技术标准制定、体系化解决人工智能安全威胁等方面提出了发展建议，以期高效应对网络空间安全威胁、提升我国网络空间安全保障能力。

关键词：网络空间；安全模型；安全能力；衍生风险

中图分类号：TP393　　**文献标识码：**A

Cyberspace Security Models and Systematic Development from Multiple Perspectives

Sun Shuai[1], Zhang Lei[1*], Hu Chunhui[1], Fu Shouqing[1], Qing Yu[1], Cui Yong[2]

(1. Zhongguancun Laboratory, Beijing 100094, China; 2. Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)

Abstract: As cyberspace technologies advance rapidly, cyberspace security risks derived from new applications and technologies are becoming more complex and hidden. Establishing a unique cyberspace security model is a common practice to deal with complex security threats in China and abroad. However, existing cyberspace security models have problems such as unclear development directions, insufficient ability to analyze risks derived from new technologies, and lack of security capabilities required for cyberspace security defense assessment. This study evaluates existing cyberspace security models from the perspectives of technology, discipline, and industry, sorts out the characteristics and development context of the cyberspace security technology system, and clarifies the urgent problems existing in cybersecurity applications. Focusing on the perspective of cyberspace security technology, this study proposes a cyberspace security model system framework based on technical elements, using existing security technologies and emerging technologies to verify the security analysis capabilities of the system framework. This study further proposes the following development suggestions: (1) improving the core framework of the cyberspace security technology system, (2) promoting the industry-university-research integration in the field of cyberspace security, (3) promoting the formulation of core technology standards regarding cyberspace security, and (4) addressing AI security threats, thus to effectively deal with cyberspace security threats and enhance the

收稿日期：2023-11-10；**修回日期：**2023-12-01

通讯作者：*张蕾，中关村实验室副研究员，研究方向为网络安全、网络优化等；E-mail: zhanglei@zgclab.edu.cn

资助项目：中国工程院咨询项目“网络空间安全技术体系与风险应对”(2022-JB-04)

本刊网址：www.engineering.org.cn/ch/journal/sscae

cyberspace security capabilities of China.
Keywords: cyberspace; security model; security capability; derivative risk

一、前言

网络空间是第五大主权领域空间，维护网络空间安全事关人民生命财产安全、关系到国家安全和社会稳定[1]。党的二十大报告提出，加快建设网络强国、数字中国，强化网络、数据等安全保障体系建设，坚决打赢关键核心技术攻坚战。保障网络空间安全，对维护国家网络空间主权、安全和发展具有重要意义[2]。目前，使用单一技术难以应对复杂多变的网络安全风险，需要全面理解和防范网络空间安全威胁，运用多种网络空间安全技术作为支撑，因此，系统梳理网络空间安全技术体系，明晰网络空间安全技术体系的要素，深入探究网络空间安全技术发展，对提升网络安全风险防范能力、促进网络空间安全技术的创新与发展尤为重要。

2010 年，国际电信联盟（ITU）将网络空间定义为：包括计算机、计算机系统、网络及其软件支持、计算机数据、内容数据、流量数据以及用户在内的全部或部分要素创建 / 组成的物理或非物理的领域[3]。早期网络空间重点关注单一的基本通信领域，研究主体和对象通常为电话、电报、传真等传统的通信基础设施。随着传输控制协议 / 网际协议（TCP/IP）通信技术和互联网技术的普及，网络空间的研究范围从传统的通信领域扩展到计算机与网络领域，研究主体也逐渐变为计算机、数据库、信息系统和互联网应用等。21 世纪以来，随着人工智能、大数据、物联网、工业互联网等技术和应用的不断成熟，网络空间的研究主体也日趋多样化，从互联网转变为关键基础设施、公共服务等。从网络空间的发展历程来看，随着技术的不断发展，网络空间的研究主体也发生变化[4]。

网络空间安全威胁会随着网络空间主体的不断变化而变化，在不同时期产生了应对不同威胁的网络空间安全防护技术[5]。例如，网络空间安全防护技术从语音窃听、信息保密等数据安全通信传输技术发展到网络空间关键基础设施安全防护技术[6]；网络空间安全技术的研究重点从通信保密、计算机安全、网络安全、信息安全保障发展到网络空间安全[7]，从单一化防御技术逐渐变为防御体系。网络空间安全模型是网络空间安全体系的具象化表现，其呈现形式和包含的技术内容也在不断更新。因此，把握安全模型的未来发展方向，建立一种可以适应网络空间安全研究主体变化的体系化模型，具有重要的研究价值。

目前，研究人员已开展了网络空间安全模型和框架构建的研究，如网络空间安全防御体系中的核心技术要点分析[8~10]、政策规章与安全模型发展之间的关系[11]等，但仍缺失从整体上立足国情开展的我国网络空间安全技术体系宏观研究。为此，本文系统梳理网络空间安全模型发展情况，阐明当前的能力缺失与现实问题，构建我国网络空间安全技术体系框架并展示实际应用场景，进一步提出我国网络安全模型的发展建议，以期更好地应对网络空间安全的新风险。

二、多视角下的网络空间安全模型评估

网络空间发展经历了多个阶段，相应的安全防护技术侧重点有所不同；描述各个安全模型的方法也不唯一，很难从单一视角全面阐述网络空间安全模型。本研究对典型的网络安全模型进行分析整理，梳理网络空间安全模型的发展脉络，分别从技术、学科和产业视角分析我国网络空间安全的发展现状，并将技术视角细化为信息安全保障、攻防对抗、关键信息基础设施安全、零信任、内生安全 5 个方面（见图 1）。

（一）技术视角

1. 信息安全保障模型

信息安全保障模型是从信息和信息系统防御角度出发建立的信息安全防御模型，以防护、检测和响应为核心，前期加入策略、预警识别等感知因素，后期加入恢复、反击等行动因素组成的防御架构。1996 年，美国国防部首次给出了信息安全保障的定义，并于同年提出了防护、检测、响应、恢复（PDRR）模型。PDRR 模型强调运用传统的单一安全防御思想，专注于信息安全保障，涵盖防护、检测、响应和恢复等环节[12]。之后，美国提出了动态

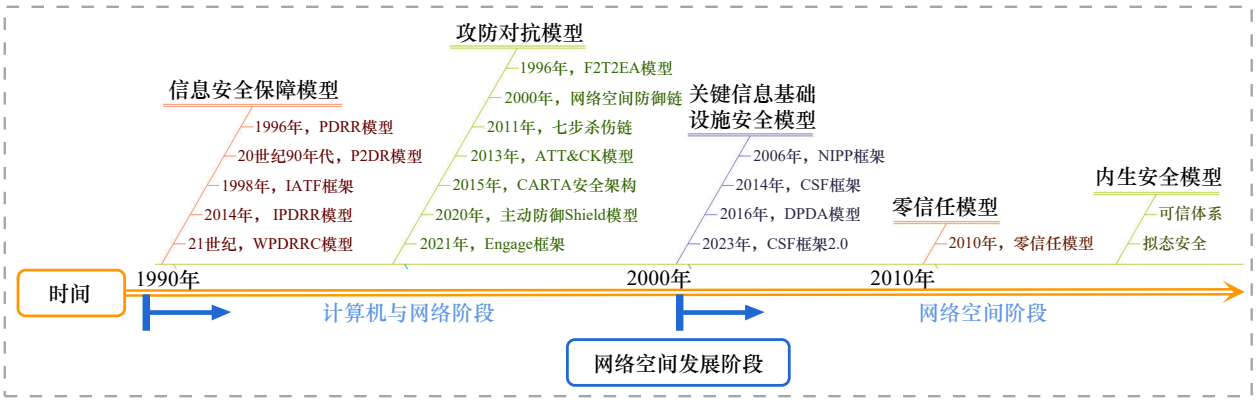

图1　技术视角下网络空间安全模型及其演进

注：P2DR模型表示安全策略、防护、检测、响应模型；IATF框架表示信息保障技术框架；IPDRR模型表示识别、防护、检测、响应、恢复模型；WPDRRC模型表示预警、保护、检测、响应、恢复、反击模型；F2T2EA模型表示经典杀伤链模型；ATT＆CK模型表示对抗战术、技术、常识模型；CARTA安全架构表示持续自适应的安全风险和信任评估架构；Engage框架表示交战框架；NIPP框架表示国家基础设施保护计划安全框架；CSF框架表示提升关键基础设施网络安全框架；DPDA模型表示美国联邦政府网络空间安全研发战略规划。

网络安全模型P2DR模型，即在防护、检测、响应（PDR）模型的防护环节前加入策略因素，并将其作为模型的核心，使防护、检测和响应环节都按照既定策略实施，体现了在防御之前利用风险评估来分析安全状态的动态过程[13]。为了保障用户信息及信息系统的安全，1998年美国国家安全局在P2DR模型的基础上，提出了IATF框架。该框架首次引入了“管理”的概念，将人为因素带入到网络安全防御模型中[14]，为美国政府和工业界的信息及信息系统安全提供指南。为了体现安全保障系统化的思想，2014年美国国家标准与技术研究院（NIST）提出了IPDRR模型。该模型在PDRR模型的基础上增加了风险评估环节，主要用于业务优先级确定、风险识别、资源优先级划分等。

在信息安全保障模型方面，我国在21世纪初结合国情，基于PDRR、P2DR等模型提出了WPDRRC模型。该模型包括预警、防护、检测、响应、恢复和反击等6个环节，能够全面反映信息系统安全保障的各方面能力，涵盖人员、策略和技术等方面的要素，以确保安全策略的贯彻执行[15]。从信息保障角度来看，网络空间安全模型以防护、检测、响应为核心，根据需求变化不断增加关键要素以改进模型。

2. 攻防对抗模型

攻防对抗模型是以武器化、漏洞利用和攻击对抗为核心组成的攻击架构，对应的安全防御模型通常分为防护、检测和响应3个阶段。最早的攻防对抗模型是1996年由美国空军提出的F2T2EA模型[16]。2000年、2011年美国又分别提出了基于F2T2EA模型的网络空间防御链模型、“七步杀伤链”模型[17]。其中，“七步杀伤链”模型明确提出，网络攻防过程中攻防双方虽各有优势，但重在识别和评估攻击链中的多个关键步骤，以支持安全专业人员的安全决策。2013年，美国MITRE公司首次提出ATT&CK模型[18]，该模型将已知攻击者行为转化为战术和技术的结构化列表，通过若干矩阵及结构化信息来表示攻防知识库，目前已在攻防能力覆盖评估、高级可持续威胁攻击（APT）情报分析、威胁狩猎系统等领域广泛应用。基于ATT&CK模型，后续也出现了CARTA安全架构（2015年）[19]、主动防御Shield模型（2020年）[20]以及Engage框架（2021年）等。

无论是以经典杀伤链为核心的网络攻防框架，还是以ATT&CK模型为核心提出的攻防知识库框架，均通过研究攻击过程以提升防御能力。以“七步杀伤链”模型为例，前期的侦查、武器化、投递阶段对应的防御手段有网络安全分析、防火墙访问控制等，可对应防御模型中的检测、保护阶段；后期的安装、指挥与控制、行动阶段对应的防御手段有拒绝访问、网络分割、入侵检测、权限降级等，可对应防御模型中的检测、防护、响应阶段。由此可见，攻防对抗模型中的每个防御步骤都与经典模型一致。

3. 关键基础设施安全模型

随着物联网、大数据等新应用的不断出现，物理空间逐步扩展到网络空间，关键基础设施安全成为网络空间重要的防护对象。传统单一模块化的网

络安全防御手段已经无法满足关键基础设施的安全防护需求[21]。美国发布了一系列针对关键基础设施保护的法案和技术框架。例如，2002 年，美国国土安全部颁布《国土安全法》，规定了针对关键基础设施保护的任务、设施、管理和技术要求[22]；2006 年，又发布《国家基础设施保护计划》（NIPP 框架），标志着美国全面开展针对关键基础设施的安全保护工作[23]。此后，美国 NIST 于 2014 年发布《提升关键基础设施网络安全框架》（CSF 框架）[24]。该框架以传统防御框架的识别、防护、检测、响应、恢复为技术核心，融合了可根据预期目标弹性化、定制化防御手段的配置文件层，并且融合了评估网络安全状态的实现层；通过搭建一个具有弹性又兼备评价反馈的复合型防御体系，满足网络空间阶段关键基础设施保护的需求。2016 年，美国提出美国联邦政府网络空间安全研发战略规划，其中的 DPDA 模型为联邦政府的网络空间安全技术发展指明了方向[10]。针对关键基础设施安全防护，美国先后提出了 CSF 的不同版本。2018 年，美国推出 CSF 框架 1.1 版本，添加了更全面的身份管理和供应链网络安全管理进程[25]，这对于组织在处理个人身份信息和符合隐私法规方面具有重要意义。2023 年 1 月，NIST 发布了 CSF 2.0 概念文件，并于 2023 年 8 月发布 CSF 2.0 的公开草案，再次明确了 CSF 模型的最新适用范围[26]。CSF 2.0 版本扩大了模型的适用范围，从保护医院和发电厂等关键基础设施扩展到为所有组织提供网络安全；同时，CSF 2.0 版本的核心层在传统的 5 个功能外还增加了第 6 个功能，即管理功能，使组织通过制定和执行内部决策以支持其网络安全战略，强调网络安全是企业风险的主要来源，与法律、财务和其他风险并列，是高层管理的考虑因素。由此可见，针对关键基础设施的安全防护模型是不断迭代发展的，以防御不同时期的新风险、满足不断增加的网络安全需求。

4. 零信任模型

云计算、虚拟化和分布式计算是网络空间技术不断演进的产物，致使网络边界和数据控制边界逐渐消失，网络安全的灵活性受到广泛关注。传统网络模型依赖大范围的访问权限，缺乏动态的上下文评估和持续验证，过度依赖被固化的网络边界，存在潜在风险和一定的安全漏洞。为了适应复杂的网络威胁环境，美国 Forrester Research 公司在 2010 年提出了零信任网络架构（ZTA），尝试摆脱传统网络边界的固有限制，在保证网络安全的情况下，实现更便捷、更灵活的接入和接出。零信任模型的核心策略是不信任任何人或设备，基于最小权限原则的访问控制，执行动态访问控制、持续身份验证等，核心技术为认证、授权、检测、分割和协作等[27]。

零信任安全与传统的“信任但验证”的安全观念相比，更加强调“动态防御”概念。零信任模型有助于提高网络安全水平，保障用户、设备或应用程序的数据安全，减少风险并适应日益复杂和威胁严重的网络环境。

5. 内生安全模型

为了解决传统网络安全模型的静态访问控制、依赖固定边界等局限性，内生安全概念于 2013 年首次提出[28]，并得到广泛关注。内生安全通过将安全性内置到系统、应用程序和数据中，不再依赖传统的外部边界和静态权限控制，从而提供了更加灵活、细粒度的自我保护安全策略，以适应现代网络和智能化的网络安全威胁的挑战。

为解决网络空间未知风险带来的安全黑洞问题，改变“易攻难守”的固有形势，我国学者提出了拟态安全[29]。这是一种不依赖漏洞后门检测和攻击特征分析等先验知识的内生安全理论与体系，核心策略为动态与弹性防御，利用虚假目标、诱饵和陷阱等手段扰乱攻击者的攻击行为，同时结合调度重构、同质异构和多模裁决等核心算法实现遭受攻击情况下仍能够“带菌生存”[30]。

此外，针对计算机系统的安全问题，我国学者提出了可信计算的概念，其思路是建立可信的计算环境，利用可信环境、身份验证、加密保护、安全监测和审计等手段，增强计算机和数据的安全性。通过应对未经授权的访问、数据泄露和篡改，确保计算平台的可信性，降低与计算环境不可信的相关安全风险，从而保障计算机系统安全以及数据的完整性和可靠性[31]。

（二）学科视角

2015 年，我国设立了网络空间安全一级学科，将网络空间安全划分为网络空间安全基础、系统安全、网络安全、应用安全、密码学及应用 5 个方向[32]。2018 年，国际上发布了网络空间安全学科知识体系（CSEC）。CSEC 主要面向本科教育，将

学科知识体系划分为数据安全、软件安全、组件安全、连接安全、系统安全、人员安全、组织安全和社会安全等8个部分。CSEC中不仅包含网络空间的科学和技术，还包含社会学、管理学等人文学因素。学科视角下的网络空间安全技术体系主要从学科知识体系出发，注重网络安全人才培养，将网络空间安全知识体系模块的逻辑性、系统性、完整性方面贯彻于课程体系建设及学生培养的各个环节。

（三）产业视角

网络空间安全产业与网络空间安全技术紧密结合，共同推动着网络空间安全的发展。从网络空间安全产业视角来看，美国的网络空间安全产业体系较为完备，上游产业涵盖技术研发和创新，提供安全防御的软硬件基础设备；中游产业主要完成技术集成，提供网络安全解决方案、网络安全产品和网络安全服务等；下游产业涵盖终端用户，包括企业、政府和用户，由传统互联网、云计算、移动互联网、物联网、大数据、工业等行业支撑，购买和使用中上游企业提供的网络安全技术和服务。美国网络空间安全产业生态系统中的各部分相互合作，确保网络及信息系统的安全性。《2022年中国网络安全产业研究报告》显示，我国网络空间安全产业链的上游可以提供基础理论、算法、芯片、高质量样本数据等产业基础能力，中游可提供各类网络安全专业设备、应用产品，下游可提供系统集成、工程建设、服务和产品分销[33]。

近年来，我国在网络空间安全战略性政策制定方面落实到位，市场规模占有率增长显著，但对比美国在网络空间安全领域和网络空间安全技术创新方面的先发优势来看，我国仍需加大网络空间安全产业投入，加快网络空间安全关键技术突破，加强网络空间安全产业链建设。

三、现有网络空间安全模型存在的问题

（一）现有模型缺乏对新技术、新应用的安全能力评估

现有的网络空间安全模型大多是为应对网络安全风险而制定的，如ATT&CK模型和基于PDR模型衍生的信息安全保障模型等。虽然基于既有攻防战术建立的防御模型或针对特定风险的网络安全防护模型在日常防御网络空间威胁时有其独特的优势，能够直观地发现防御链条的薄弱点、快速针对威胁行为作出应对，但上述模型缺乏针对新应用和新技术的必要安全能力分析。开展对新技术、新应用的安全能力评估能够提升网络空间技术的自身安全能力，从“头疼医头、脚疼医脚”的发展现状转变为增强自身安全能力，从根源上解决网络空间安全威胁。

（二）现有模型的体系化防护程度不足

现有的网络空间安全模型和框架是在特定历史阶段、瞄准相应时期的战略导向及需求制定的，体系化相对较弱。网络空间技术发展迅速，防护理念也从防御威胁本身发展到减轻网络风险的脆弱性，传统针对不同安全威胁建立的网络空间安全模型无法全面满足网络空间安全发展需要，固化的防御策略不具备风险形态的敏感性。在全面应对网络空间安全复杂多样的威胁时，现有单一的防御手段防护能力明显不足。网络空间安全模型需要进一步提升体系化程度，增加网络空间安全防御韧性，以适应网络空间风险的快速变化。

（三）网络空间安全技术体系的未来发展路径不清晰

当前，网络空间安全模型发展存在演进式发展和创新式发展两种发展路径。纵观美国网络空间安全模型的发展历史，现有的体系框架多数都是从原始核心框架不断衍生和迭代而来。以CSF框架为例，利用IPDRR模型构成CSF框架的核心层，而IPDRR模型是基于PDRR模型升级迭代而来；ATT&CK模型的初始版本只有9个战术，现已更新完善为14个战术，据此衍生的Shield知识库和Engage框架已被广泛应用。网络空间安全的内涵和范围一直在变化，网络空间安全模型也随着新观念、新理念的提出不断变化。多数网络空间安全模型根据网络空间安全内涵和外延的变化不断演进发展。与之相对，包括零信任模型和内生安全模型在内的创新性安全模型，则是以网络空间安全需求为驱动构建的一种网络空间安全模型。我国为了应对面临的网络空间安全威胁，有必要深入研究采用何种技术体系发展路径，以更好地适应网络空间的现实发展需要。

四、网络空间安全模型的体系化发展

本研究在分析已有网络空间安全模型的基础上，探讨网络空间安全技术体系框架可能的发展形态，构建可扩展、基于技术要素的网络空间安全技术体系框架，为从技术视角探究网络空间安全技术发展提供了新思路。同时，初步开展了典型应用的安全风险防范分析，为网络空间安全模型的体系化发展提供理论支撑。

（一）基于技术要素的网络空间安全技术体系框架

网络空间安全技术作为防御网络空间安全风险的重要手段，与网络空间技术呈现相伴相生、相辅相成的特性，网络空间安全技术的发展依托于网络空间技术的迭代更新。结合现有网络安全模型的特点，本研究从分析网络空间技术的构成要素出发，充分考虑安全技术的发展，构建了可扩展的网络空间安全技术体系框架。此框架模型从伴生角度出发，对网络空间技术本体进行剖析，有利于分析出关键技术存在的网络空间安全风险，也可以预测未来的应用安全风险，解决不断出现的网络空间新威胁，并提供技术分析体系支撑。网络空间技术体系具有可扩展的特性，因而从伴生角度构建的安全技术体系更易适应技术的发展而与时俱进。

结合经典互联网体系结构的分层思想，本研究将网络空间技术体系分为3层（见图2），自下而上依次为：由卫星、第五代移动通信（5G）、无线上网（WiFi）、光纤等物理通信硬件、拓扑结构构成的通信核心技术层；由各类计算机系统技术和互联网连接协议、技术构成的处理器（CPU）、操作系统（OS）和互联网核心技术层；基于云计算、工业互联网、物联网、人工智能等应用构成的上层应用核心技术层。基于网络空间技术与安全技术的关系，对应的网络空间安全技术分别为通信安全技术、系统安全技术、网络安全技术、应用安全技术。此外，网络空间安全基础研究对网络空间安全技术体系具有重要的理论支撑，因此在上述框架中也添加了基础理论核心要素。

在基于技术要素的网络空间安全技术体系框架中，将网络空间安全技术要素点按照对应的技术分层进行归类，如图3所示。其中，通信安全技术层包含通信线路安全、通信传输安全等技术；系统安全技术层包含芯片安全、操作系统安全、存储安全等技术；网络安全技术层包含路由安全、域名安全、入侵检测、访问控制等技术；应用安全技术层包含人工智能安全、云计算安全、区块链安全等通用技术以及金融、能源、工业互联网等领域的专有应用安全技术；基础理论层包含密码学、博弈论、信息论等网络空间安全基础理论。该技术体系框架遵循互联网体系结构的演进式发展思路，可以根据网络空间安全技术的不断变化进行扩展，包容不同领域应用的多样化特性。同时，从网络空间技术要

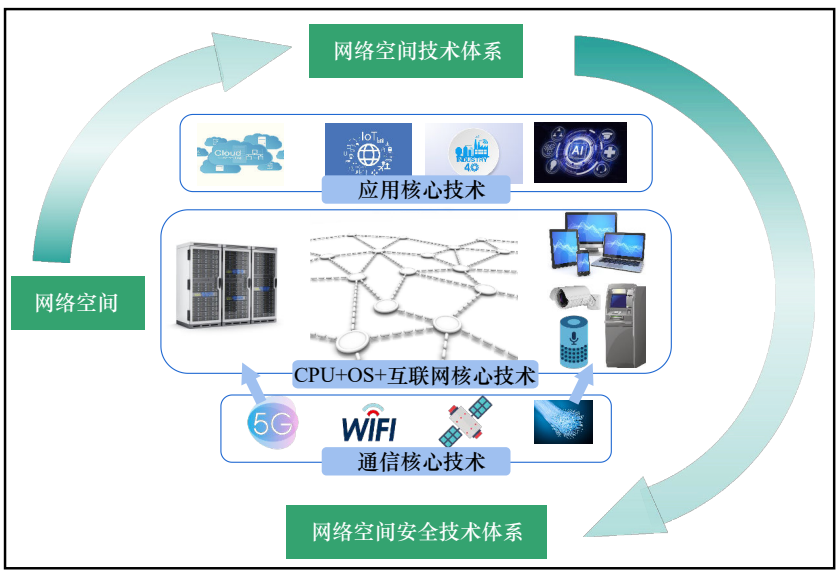

图2　基于技术要素的网络空间安全技术体系构建逻辑原理

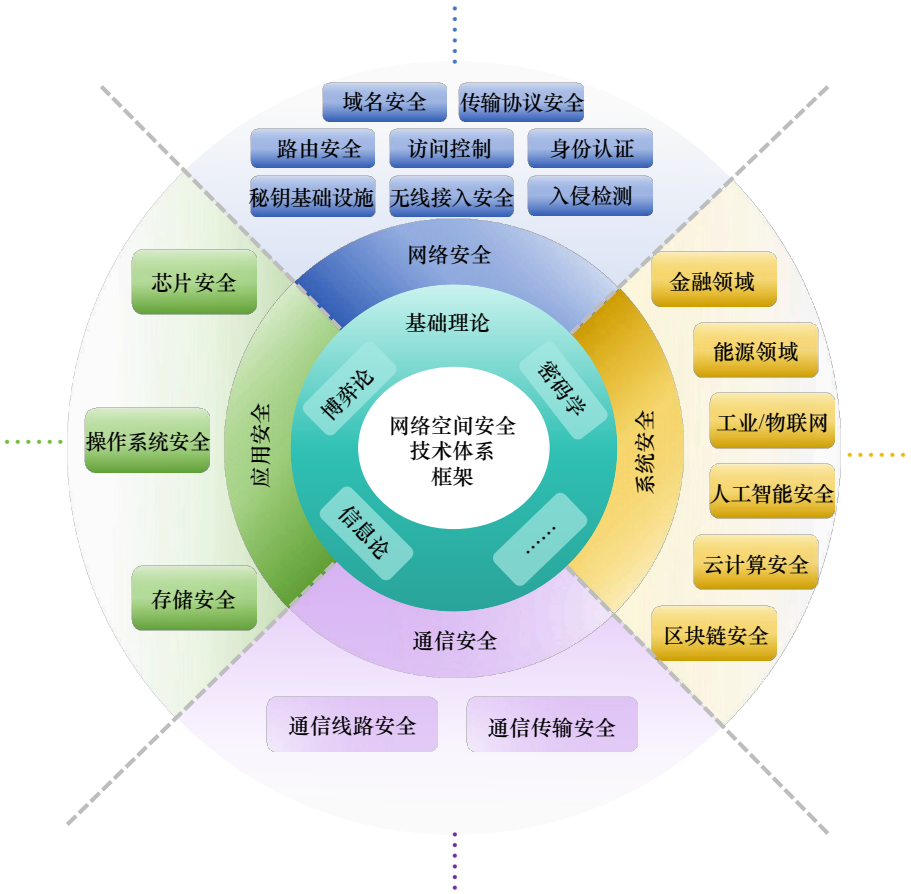

图3　基于技术要素的网络空间安全技术体系框架构成

素的角度来看，体现网络空间安全各个层面的技术发展趋势，能够较全面地分析新风险应对的防御技术，为网络空间安全技术提高风险和威胁应对能力提供有力支撑。下文将以防火墙应用和生成式人工智能应用为例，探究此技术体系框架用于分析安全风险及对应防御技术的过程。

（二）网络空间安全技术体系的应用分析

1. 已有安全技术的能力分析

以防火墙产品为例，按照实现技术的不同，分为应用程序代理防火墙、包过滤防火墙、检测防火墙[34]等。如图4所示，防火墙产品目前使用的网络空间安全技术包括应用代理技术、包过滤技术、状态检测技术、完全内容检测技术等[35]。

利用基于技术要素的网络空间安全技术体系框架分析防火墙产品，在应用安全技术层使用了防火墙的应用代理技术、完全内容检测技术中的内容过滤与应用识别功能，防火墙的应用安全技术层可以检查所有应用层的信息包，并将检查的内容信息放入决策过程从而提升安全性；在网络安全技术层使用包过滤技术、状态检测技术等，在网络中相互连接的设备上进行访问控制，对通过设备的数据包进行检查，同时检测数据包的状态，限制恶意数据包进出内部网络；在计算系统安全技术层使用内容检测技术中的防病毒检测、漏洞扫描修复等技术。在应对网络攻击事件时，可以利用上述框架中的要素点将需要升级加固的对象从一个技术层细化到一个技术点，进一步提升防火墙整体的防护效率，加快技术迭代效率，降低成本；可以评估每一个技术层的安全技术分布和技术更新，为保证防火墙的整体安全，对于安全技术较少、长时间没更新的技术要素需要更加关注，以防新风险产生。

2. 新兴技术的安全风险分析

利用基于技术要素的网络空间安全技术体系框架，以生成式人工智能应用为例（见图4），分析其在不同的网络空间安全技术层面可能产生的安全问题。生成式人工智能应用带来的安全问题主要有3个方面：① 大模型训练多采用分布式训练方法，

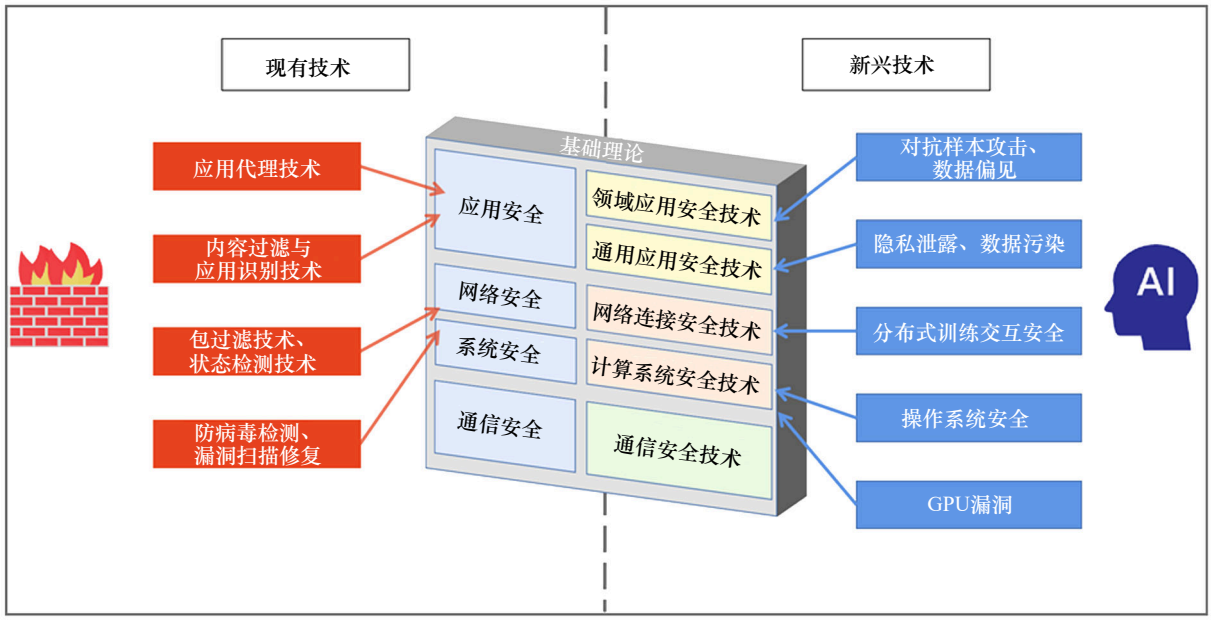

图4　网络空间安全技术体系框架的应用分析

训练结果通过网络传输至终端，而分布式存储和数据传输会带来网络流量侧的安全问题；② 大模型训练和数据运算需要在服务器上完成，致使服务器存在相应的安全风险；③ 在大模型训练后的部署阶段，应用端会存在投毒攻击、隐私泄露问题。对应到网络空间安全技术体系中，将参数服务器交互安全问题分类到网络安全技术层；将图形处理器及其操作系统的安全漏洞问题分类到系统安全技术层；将对抗样本攻击、数据偏见等投毒攻击产生的安全风险以及隐私泄露、数据污染等安全问题分类到应用安全技术层中。

除了生成式人工智能自身存在的安全风险，本研究探讨了人工智能对已有安全防御体系的威胁分析。人工智能赋能网络攻击的典型技术包括网络资产自动探测识别技术、智能社会工程学攻击技术、智能恶意代码攻击技术、自动化漏洞挖掘与利用技术等[36]。自动探测识别技术会威胁到数据安全，可划分到应用安全层内；社会工程学攻击技术一般应用于身份伪造中，可归类至网络安全层中；智能恶意代码攻击技术是对传统攻击的升级迭代，在应用安全层、网络安全层以及系统安全层均有体现；自动化漏洞挖掘与利用技术与传统漏洞利用技术相似，可归纳于系统安全层。

基于技术要素的网络空间安全技术体系框架可助力网络安全研究。通过对新兴技术可能面临的安全风险进行分析并梳理归类，从技术角度出发明确每一个技术层对应的安全风险问题，有利于定向开展相关问题的技术研究工作。

五、网络空间安全模型的发展建议

（一）完善网络空间安全技术体系核心框架

明确适合我国国情的网络空间安全技术体系发展生态，并不断改进和完善技术核心框架；结合多样化应用场景的变化和网络强国发展的实际需要，对技术核心框架进行迭代，形成有特色、有针对性的安全模型。鼓励学界、业界、政府机构等共同参与技术体系建设，组织网络空间安全领域核心技术攻关，取长补短，促进网络空间安全行业良性发展。此外，需要制定相应的评估与认证机制，用于验证技术体系的有效性和可信度。

（二）促进网络空间安全领域“产学研”协同

结合产业实际需要，优化网络空间安全技术体系，以确保技术发展与产业实际需求相契合；利用产业界的海量数据与真实案例，梳理技术体系的发展脉络、洞察未来发展方向。联合高校与科研院所，加强网络安全领域的教育与培训，培养具备创新思维和实践能力的专业人才，支撑产业良性发展，提升我国网络空间安全领域的软实力。鼓励

高校与产业界合作开展相关项目研究，深入了解实际需求，聚焦关键技术难题，开展前沿研究探索，共同解决实际问题，促进知识和技术的转化与应用。

（三）制定网络空间安全技术体系相关的技术标准

网络空间安全技术体系的建立离不开技术标准的支撑。技术标准定义了必要的规范、流程和要求，可以确保系统和设备在防御威胁、检测攻击和应对事件方面具备一致性和有效性，在网络空间安全领域发挥着关键的作用。为了提升我国网络空间安全的国际影响力，争夺科技话语权，需要在网络空间安全技术体系下补充对应的技术标准，制定符合国际通用标准的技术规范，助力我国企业和研究机构更好地融入国际市场，参与全球合作与竞争。此外，技术标准还可以提供通用的参考框架，促进不同组织及行业之间的合作与协同。通过制定一致的规范和接口标准，不同的企业和组织可以高效进行信息交流、技术对接和资源共享，提高整个网络空间安全体系的鲁棒性和防御能力。

（四）全方面体系化解决人工智能的安全威胁

为应对人工智能对已有网络空间安全防御体系的威胁，应从安全技术与政策法规两方面解决人工智能安全问题。在网络空间安全技术发展方面，应系统分析人工智能对网络空间安全的影响和威胁，梳理网络空间安全技术脆弱点，推动网络空间安全技术快速发展。在政策法规制定方面，推动人工智能相关政策和立法工作，加强重点行业和关键基础设施的体系化防护建设，全面提升网络空间安全防御能力，切实减少并阻断人工智能对网络空间安全的负面影响。

利益冲突声明

本文作者在此声明彼此之间不存在任何利益冲突或财务冲突。

Received date: November 10, 2023; Revised date: December 1, 2023
Corresponding author: Zhang Lei is an associate research fellow from Zhongguancun Laboratory. Her major research fields include network security, network optimization, etc. E-mail: zhanglei@zgclab.edu.cn
Funding project: Chinese Academy of Engineering project “Cyberspace Security Technology System and Risk Response” (2022-JB-04)

参考文献

[1] 冯登国. 准确把握网络空间安全技术发展的新特征 全力助推国家安全体系和能力现代化 [J]. 中国科学院院刊, 2022, 37(11): 1539–1542.
Feng D G. Accurately grasp the new features of cybersecurity technology development and fully promote the modernization of national security system and capabilities [J]. Bulletin of Chinese Academy of Sciences, 2022, 37(11): 1539–1542.

[2] 方滨兴, 杜阿宁, 张熙, 等. 国家网络空间安全国际战略研究 [J]. 中国工程科学, 2016, 18(6): 13–16.
Fang B X, Du A N, Zhang X, et al. Research on the international strategy for national cyberspace security [J]. Strategic Study of CAE, 2016, 18(6): 13–16.

[3] International Telecommunication Union. Toolkit for cybercrime legislation [EB/OL]. (2010-07-31)[2023-10-20]. https://www.itu.int/ITU-D/cyb/cybersecurity/docs/flyer-regulatory-resources.pdf.

[4] 崔保国. 世界网络空间的格局与变局 [J]. 新闻与写作, 2015 (9): 25–31.
Cui B G. The pattern and change of the world cyberspace [J]. News and Writing, 2015 (9): 25–31.

[5] 王文杰. 跨国网络空间安全治理的困境与中国对策研究 [D]. 济南: 山东大学 (硕士学位论文), 2019.
Wang W J. The research on the dilemmas of multinational cyberspace security governance and China’s countermeasures [D]. Jinan: Shandong University (Master’s thesis), 2019.

[6] 周文. 关键信息基础设施整体安全保障思路 [J]. 信息安全研究, 2016, 2(10): 946–951.
Zhou W. The totlal solution of cyber security in critical information infrastructure [J]. Journal of Information Security Research, 2016, 2(10): 946–951.

[7] 吕欣. 网络空间安全保障体系研究 [J]. 信息安全研究, 2015, 1(1): 37–43.
Lyu X. Studies on cybersecurity assurance system [J]. Journal of Information Security Research, 2015, 1(1): 37–43.

[8] 张军. 网络空间安全体系与关键技术分析 [J]. 中国新通信, 2021, 23(14): 129–130.
Zhang J. Analysis of cyberspace security system and key technologies [J]. China New Telecommunications, 2021, 23(14): 129–130.

[9] 谭可, 马清勇, 谢曦, 等. 网络空间安全体系与关键技术分析 [J]. 通讯世界, 2020, 27(6): 133, 135.
Tan K, Ma Q Y, Xie X, et al. Analysis of cyberspace security system and key technologies [J]. Telecom World, 2020, 27(6): 133, 135.

[10] 罗军舟, 杨明, 凌振, 等. 网络空间安全体系与关键技术 [J]. 中国科学: 信息科学, 2016, 46(8): 939–968.
Luo J Z, Yang M, Ling Z, et al. Architecture and key technologies of cyberspace security [J]. Scientia Sinica Informationis, 2016, 46(8): 939–968.

[11] 张弛, 左晓栋. 美国2019联邦网络安全研发战略计划解读 [J]. 网络空间安全, 2020, 11(3): 90–95.
Zhang C, Zuo X D. Analysis of U.S. 2019 federal cybersecurity research and development strategic plan [J]. Cyberspace Security, 2020, 11(3): 90–95.

[12] 熊小兵. 一种基于PDRR模型的静态数据完整性保护方案 [J]. 计算机与信息技术, 2006 (11): 51–52, 55.
Xiong X B. A static data integrity protection scheme based on PDRR model [J] Computer and Information Technology, 2006 (11): 51–52, 55.

[13] 刘峰, 林东岱, 等. 美国网络空间安全体系 [M]. 北京: 科学出版社, 2015.
Liu F, Lin D D, et al. Overview of the cybersecurity system in USA [M]. Beijing: Science Press, 2015.

[14] 王妍, 孙德刚, 卢丹. 美国网络安全体系架构 [J]. 信息安全研究, 2019, 5(7): 582–585.
Wang Y, Sun D G, Lu D. American network security architecture [J]. Journal of Information Security Research, 2019, 5(7): 582–585.

[15] 贾浩淼, 王石. NIST《改进关键基础设施网络安全框架》分析 [J]. 信息技术与标准化, 2014 (4): 47–50, 73.
Jia H M, Wang S. Analyzing on *NIST framework for critical infrastructure cyber security* [J]. Information Technology & Standardization, 2014 (4): 47–50, 73.

[16] TIRPAK J A. Find, fix, track, target, engage, assess [J]. Air Force Magazine, 2000, 83(7): 24–29.

[17] Hutchins E, Cloppert M, Amin R. Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains [C]. Washington DC: 6th International Conference on Information Warfare and Security (ICIW 2011), 2011.

[18] Strom B E, Applebaum A, Miller D P, et al. MITRE ATT&CK: design and philosophy [EB/OL]. (2018-07-20)[2023-11-20]. https://pdfslide.net/documents/mitre-attcka-design-and-philosophy-attck-was-created-out-of-a-need.html?page=1.

[19] 江欣. 基于EDR与CARTA模型的动态主机安全防护平台研究 [J]. 网络安全技术与应用, 2020 (9): 47–48.
Jiang X. Research on dynamic host security protection platform based on EDR and CARTA model [J]. Network Security Technology & Application, 2020 (9): 47–48.

[20] Fowler C, Goffin M, Hill B, et al. An introduction to MITRE shield [EB/OL]. (2020-08-26)[2023-11-20]. https://shield.mitre.org/resources/downloads/Introduction_to_MITRE_Shield.pdf.

[21] 张大伟, 沈昌祥, 刘吉强, 等. 基于主动防御的网络安全基础设施可信技术保障体系 [J]. 中国工程科学, 2016, 18(6): 58–61.
Zhang D W, Shen C X, Liu J Q, et al. TC assurance architecture for cybersecurity infrastructure based on active defense [J]. Strategic Study of CAE, 2016, 18(6): 58–61.

[22] Brook D A, King C L. Civil service reform as national security: The homeland security act of 2002 [J]. Public Administration Review, 2007, 67(3): 399–407.

[23] 孔勇, 范佳雪. 美国《国家基础设施保护计划》2013 更新版解读 [J]. 中国信息化, 2023 (1): 49–52.
Kong Y, Fan J X. Interpretation of the 2013 update of *the national infrastructure protection plan* of the United States [J]. Zhongguo Xinxihua, 2023 (1): 49–52.

[24] 李留英. 美国网络威胁情报共享实践研究 [J]. 信息安全研究, 2020, 6(10): 941–946.
Li L Y. Research on the practice of cyber threat intelligence sharing in the United States [J]. Journal of Information Security Research, 2020, 6(10): 941–946.

[25] Krumay B, Bernroider E W N, Walser R. Evaluation of cybersecurity management controls and metrics of critical infrastructures: A literature review considering the NIST cybersecurity framework [C]// Gruschka N. Nordic Conference on Secure IT Systems. Cham: Springer, 2018: 369–384.

[26] National Institute of Standards and Technology of US Department of Commerce. Public draft: The NIST cybersecurity framework 2.0 [EB/OL]. (2023-08-08)[2023-11-20]. https://www.nist.gov/system/files/documents/2023/11/17/11032023%20CSA%20Public%20Draft_%20The%20NIST%20Cybersecurity%20Framework%202%20CSA%20Comments.pdf.

[27] STAFFORD V. Zero trust architecture [R]. Gaithersburg: The National Institute of Standards and Technology of US Department of Commerce, 2020.

[28] Sabnis S, Verbruggen M, Hickey J, et al. Intrinsically secure next-generation networks [J]. Bell Labs Technical Journal, 2012, 17(3): 17–36.

[29] 邬江兴. 网络空间拟态安全防御 [J]. 保密科学技术, 2014 (10): 1, 4–9.
Wu J X. Cyberspace mimicry security defense [J]. Secrecy Science and Technology, 2014 (10): 1, 4–9.

[30] 康友春. 网络空间安全创新理论——拟态防御 [J]. 智能建筑, 2018 (6): 54–59.
Kang Y C. The innovated theory of network space security—Mimicry defense [J]. Intelligent Building, 2018 (6): 54–59.

[31] 沈昌祥, 张焕国, 王怀民, 等. 可信计算的研究与发展 [J]. 中国科学: 信息科学, 2010, 40(2): 139–166.
Shen C X, Zhang H G, Wang H M, et al. Research and development of trusted computing [J]. Scientia Sinica Informations, 2010, 40(2): 139–166.

[32] 李建华, 邱卫东, 孟魁, 等. 网络空间安全一级学科内涵建设和人才培养思考 [J]. 信息安全研究, 2015, 1(2): 149–154.
Li J H, Qiu W D, Meng K, et al. Discipline construction and talents training of cyberspace security [J]. Journal of Information Security Research, 2015, 1(2): 149–154.

[33] 中国网络安全产业创新发展联盟, 中国信息通信研究院.中国网络安全产业研究报告(2022 年) [R]. 北京: 中国信息通信研究院, 2023.
China CybersecurityIndustry Alliance of Innovation and Development (CIIDAlliance), China Academy of Information and Communications Technology. Research report on China's cybersecurity industry (2022) [R]. Beijing: China Academy of Information and Communication Technology, 2023.

[34] 王吉. 基于计算机防火墙安全屏障的网路防范技术 [J]. 信息与电脑(理论版), 2014 (1): 132–133.
Wang J. Network prevention technology based on computer firewall security barrier [J]. China Computer & Communication, 2014 (1): 132–133.

[35] 方自远, 王栋. 网络防火墙技术 [J]. 电脑知识与技术, 2018, 14(32): 30–32.
Fang Z Y, Wang D. Network firewall technology [J]. Computer Knowledge and Technology, 2018, 14(32): 30–32.

[36] 方滨兴, 时金桥, 王忠儒, 等. 人工智能赋能网络攻击的安全威胁及应对策略 [J]. 中国工程科学, 2021, 23(3): 60–66.
Fang B X, Shi J Q, Wang Z R, et al. AI-enabled cyberspace attacks: Security risks and countermeasures [J]. Strategic Study of CAE, 2021, 23(3): 60–66.