当前,智能化军事革命风起云涌。从无人蜂群作战到战场态势智能 分析,从自动化指挥决策到网络空间防御,人工智能系统日益成为军事 决策、目标识别和作战规划的重要支撑。"制智权"成为大国军事博弈的 战略焦点。

密码学,这门古老而又前沿的学科,不仅提供了丈量军事智能技术

安全尺度的数学"尺子",也给出了解决军事智能技术安全问题的数学可 解释性。有专家认为,密码学或将成为推动人工智能安全从经验性防御 向数学可验证安全范式跃迁的"金钥匙"

当古老的密码学遇到前沿学科人工智能技术,会发生怎样神奇的化 学反应?

打造人工智能的"密码盾牌"——

# 当密码学遇到人工智能

■李海亮 刘友良



## 人工智能安全面临 多重危机

今年夏季,北约一年一度的"锁定 盾牌"网络防御演习在爱沙尼亚塔林举 行。有专家认为,此次演习中一些"桥 段",展现了人工智能的"双刃剑"特征:

一方面,与人工智能相关的尖端技 术在应对网络攻击时,正在发挥越来越 重要的作用;另一方面,演习也暴露出 军事智能技术的脆弱性,来自对手的数 据投毒、模型窃取或对抗性攻击,都有 可能导致灾难性后果。

今天的人工智能,特别是深度学习 模型,就像一个在特定领域拥有超凡智 力、但心智尚未成熟的"天才少年"。它 强大却也异常敏感和脆弱,其固有的 "黑箱"特性和对数据的极度依赖,使其 在多个层面都可能被恶意利用或攻击。

——数据层。数据和信息是人工 智能系统的"血液"和"神经",是系统安 全的"生命线"。"数据投毒"攻击能从源 头上污染人工智能系统的"思想"。恶 意攻击者可以在庞大的数据集中悄悄 混入一些"毒数据",模型在不知不觉中 "吃"下这些毒数据,其"世界观"就会被 扭曲,导致在关键时刻做出灾难性错误 判断。军事智能系统的数据,如卫星图 像、雷达信号、通信截获、传感器网络数 据等,在采集、传输、存储和处理环节都 面临严重威胁。俄乌冲突中,乌克兰军 方使用的商用人工智能图像识别系统 就曾受到攻击,导致无人机误判战场目 标。美国国防高级研究计划局实验显 示,只需修改5%的坦克图像标签,就能 使智能系统目标识别准确率下降40%。

-模型层。智能模型是人工智 能系统的"大脑",往往会成为攻击者的 首选目标。当前,模型安全技术相对滞 后于模型能力跃升,给了攻击者可乘之 机。"模型逆向攻击"可通过模型的输出 结果推断其训练数据和参数,重构系统 内部逻辑和算法,使模型变成潜伏在系 统内部的"内鬼"。"对抗样本攻击"通过 精心设计的干扰,使人工智能产生错误 判断,甚至引发"自己人打自己人"的误 击悲剧。有实验表明,在停车标志上粘 贴特定图案,就能欺骗自动驾驶系统将 其识别为限速标志。在坦克上贴几张 特殊设计的贴纸,就可能诱导敌方的智 能目标识别系统把它识别成一辆无害

——系统层。基础软硬件及其所 依托的供应链如同"数字地基","地基" 一旦出了问题,就有可能引发系统安全 的全局性崩塌。智能系统依赖复杂的 软件栈、庞大的计算集群和高速网络通 信,为攻击者提供庞大的攻击面和众多 的攻击通道。"供应链攻击"通过在智能 芯片中植入硬件木马、预埋后门,或将 恶意代码隐藏在预训练模型和开源库 中,使智能系统在运行中悄然泄露数据



密码学让人工智能更安全。

AI图片

或执行恶意指令。从第三方IP核、开源 框架到云服务平台,任一环节遭污染, 都可能使上层的加密隔离形同虚设。 分布式智能系统还面临严重的"共谋攻 击"威胁,多个恶意节点的协作可严重 误导全局模型的判断。

面对这些威胁,传统"防火墙+杀毒 软件"式的安全体系很难发挥作用,需 要从人工智能的"基因"层面入手,在其 基础理论与架构设计中融入强大的安 全能力,构建全新内生安全体系,为人 工智能筑起从硅基石到云服务的"信任 长城"。

## 密码学为人工智能 注入数学"安全基因"

面对人工智能的内生安全困境,古 老的密码学正在焕发全新的生机。在 大众的认知里,密码学主要是加密解 密,是保护通信秘密的工具,这其实是 对密码学的一种片面窄化理解。现代 密码学的核心,远不止于"隐藏",更在 于"信任"的构建。它运用数学的严谨 性,提供了一整套关于数据隐私性、完 整性、真实性的解决方案。将这套"信 任科学"与人工智能相结合,就能为人 工智能注入全新的数学"安全基因",从 根本上改变其脆弱的本质。

同态加密,为数据戴上"隐形斗 篷"。同态加密是基于数学难题计算 复杂性理论的密码学技术。该技术允 许在加密数据上直接进行计算而无需

解密,为人工智能提供了革命性的安 全训练范式。想象一下,你有一个锁 着的保险箱,里面装着机密文件。你 想让一位专家处理这些文件,但又不 想让他看到文件的具体内容。同态加 密技术就像一个神奇的手套箱,专家 可以戴着特制的手套,伸进箱子里处 理文件,但他看不到里面的内容。操 作完成后,你就能用自己的钥匙打开 箱子,取出已经处理好的文件。北约 联合训练项目采用了同态加密技术, 各成员国能在不共享原始军事数据的 情况下,将加密后的训练数据上传至 中央服务器,共同训练高精度的威胁

零知识证明,为智能决策提供"可 信虎符"。零知识证明是一种在不泄露 信息本身的情况下,为互不信任的双方 提供可信验证的密码学工具。面对前 面提到的敌方坦克贴上特殊贴纸伪装 成校车,零知识证明就能派上用场。当 智能目标识别系统做出"这是一辆校 车"的判断时,它可以同时生成一个零 知识证明,向指挥中心证明:做出的判 断,并非基于某些微小的、可疑的、可能 是对抗性扰动的特征,而是基于该物体 宏观的、稳定的、符合校车定义的特 征。如果系统无法生成这样一个有效 的证明,那么就可以判定可能遭受了对 抗样本攻击,从而提醒操作员进行人工 复核。这相当于为人工智能的每一次 判断都增加了一个"可信度验算"步骤, 让敌人精心设计的"伪装"难以遁形。 零知识证明技术为可信人工智能提供 了有效的解决方案,使人工智能系统证

明其输出正确性而不泄露数据隐私和 推理过程。英国国防部已经将零知识 证明视为符合《致命性自主武器系统伦 理框架》的关键技术。

安全多方计算,为联合指挥提供 "保密圆桌"。安全多方计算是一种在 无可信第三方条件下,通过密码学协议 实现多方数据协同计算且不泄露数据 隐私信息的技术,允许多方在不泄露各 自输入内容的情况下共同计算结果。 在多国联合军事行动中,各国参谋人员 使用安全多方计算技术进行联合战役 规划,各国输入加密后的军力部署和作 战能力数据,系统输出优化的联合行动 方案,而任何参与方都无法获知他国的 具体军事机密。在2024年多个国家参 加的反恐行动中,借助该技术,在不暴 覆盖全球威胁模式的识别网络,显著提 高了对跨境恐怖分子的追踪效率。

后量子密码技术,为未来智能系 统穿上"密码铠甲"。当前的加密体 系,在未来的量子计算机面前将如同 "纸糊的窗户"。这意味着,未来战场 上所有基于现有加密通信、身份认证 和数据保护技术的人工智能系统都面 临被颠覆的风险。后量子密码技术通 过数学理论重构密码体系,能有效抵 御量子计算威胁。世界主要军事强国 都在大力发展和部署后量子密码。这 场围绕夺取未来制智权的"密码战争" 已经打响。

密码学在人工智能安全领域的应 用还有很多。如针对深度伪造的问题, 结合密码学原理与数字水印技术,可给

数据文件盖上一个无法伪造的"数字钢 印",也可通过嵌入密码学水印等技术, 为大模型生成的内容打上可验证的"数 字指纹"等等。密码学正从数据、模型 到输出的全链条,系统性地为人工智能 系统植入安全属性,使其从一个不稳定 的"黑箱",转变为一个安全、可靠、可控

## 构筑军事智能安全 的密码学"免疫体系"

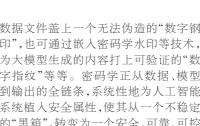
智能化浪潮风起云涌。未来战争 的胜负手不仅在于谁的人工智能更 聪明,更在于谁的人工智能更安全、 更可靠。从密码学的视角审视人工 智能安全,我们看到了一条从被动防 御走向主动免疫、从信任模型走向验 证模型、从当下安全走向未来安全的 清晰路径。

分层防御,构建军事智能安全纵深 矩阵。军事智能防护需贯彻"纵深防 御"理念,构建覆盖数据层、模型层、系 统层的全链路防御矩阵。在数据层,采 取同态加密、属性基加密等技术,确保 智能军事系统使用和生成的数据安全 可靠。在模型层,采用零知识证明、模 型水印等技术,增强模型的抗干扰、抗 攻击能力。在系统层,通过量子密钥分 发、安全多方计算等技术构建可信执行 环境。

动态防御,打造军事智能安全的移 动阵地。军事智能安全应借鉴军事上 的"机动防御"策略,采用动态密码学方 案,通过周期快速更新加密参数和算 法,快速切换防御阵地,缩小防御面,减 少攻击通道,增强抵御攻击的能力。还 可以采取可撤销生物特征、自适应差分 隐私预算等密码学技术,降低信息泄露 的风险和危害。据报道,美军的"联合 全域指挥控制系统"引入基于区块链的 密钥管理系统,实现加密策略的实时动 态调整,这种"动态安全"理念显著提升

主动防御,为军事智能系统接种 "疫苗"。智能系统安全应借鉴"积极防 御"的军事思想,在模型训练和部署之 前,主动在原始数据上添加精心设计的 微小扰动,这种扰动对于模型的正常工 作毫无影响,却能极大地破坏攻击者生 成有效对抗样本的路径,使军事智能模 型从建立之初就对"病毒"产生抵抗 力。攻击者即便拿到了模型,也很难制 作出能够欺骗它的"毒药"。当经过伪 装的对抗样本试图侵入时,模型免疫系 统就会被激活,在模型输入层自动识别 出异常的扰动。

人工智能的发展不应是一场"裸 奔"的冒险,为它穿上密码学"安全铠 甲",构建起"智能免疫系统",是我们迈 向一个安全、可信、可控的人工智能未 来的必由之路。当前,相关的密码技术 与人工智能的深度融合研究才刚刚起 步,仍有许多理论和工程挑战需要攻 克。但我们有理由相信,在不远的将 来,一个真正安全的智能时代将因此而 加速到来。





紧凑型聚变能实验装置(BEST)的杜瓦底座成功落位安装。

2025年10月1日,我国紧凑型聚变能 实验装置(BEST)迎来关键时刻——杜瓦 底座在安徽合肥精准落位。这个酷似"行 星发动机"的装置,引起了人们的关注。

杜瓦底座是BEST的核心部件,能 隔绝1亿摄氏度的等离子体,为超低温 工作的超导磁体提供隔热保护。作为 国内聚变领域最大的真空部件,它将 成为承载整个BEST 主机约 6700 吨重

量的"超级地基"。杜瓦底座的安装就 位,标志着我国新一代"人造太 阳"——BEST项目正式进入主机全面 组装阶段,为实现我国的聚变能源梦 想奠定坚实基础。

新华社发

在BEST之前,我国首台全超导托 卡马克装置EAST已屡创佳绩。就在今 年1月,EAST实现亿度千秒高约束模等 离子体运行,再次刷新世界纪录。

## BEST项目迎来关键时刻-

## 用核聚变点亮一盏灯

■辛宇恒 宣传杨

核聚变发电的关键在于实现能量 增益 Q>1——即输出能量超过输入能 量,并且稳定地转化为电能,才能具备 实际发电能力。然而,科学验证与工 程实践之间仍存在巨大鸿沟。当前阶 段,EAST还无法实现持续的高增益能 量输出,难以将理论成果有效转化为 实际应用。

如果把核聚变能源的探索比作人 类追寻飞行的历程,那么让"飞机离地 飞行"还远远不够,BEST要完成一次举 世瞩目的"载重航行"。今年5月,全球 首个采用全超导托卡马克技术路线的 紧凑型聚变能实验装置BEST正式进入 总装阶段,接下了EAST的"接力棒"。 EAST证明了我们能够实现高温、长脉 冲的稳态等离子体约束,BEST则在此 基础上进一步突破,力争通过燃烧等离 子体实现 Q>1,从而从聚变反应中获取 净输出能量。

如何填补"实验堆"到"示范堆"的 工程化空白,让核聚变从实验室走向实 际工程? BEST 提供了一种解决方案。 作为EAST的"升级版",BEST选择了一 条"小而精"的技术路线。它首次采用 紧凑高场超导托卡马克技术和第二代 高温超导带材,优化了超导磁体系统布 局和真空室结构。这使得BEST在相对 较小的体积内拥有了更强大的等离子 体约束效率,可以实现更高功率的聚变 反应。与国际热核聚变实验堆(ITER) 相比,BEST体积减小了40%,等离子体 约束时间却延长3倍。这种设计不仅降 低了超导磁体和低温系统的能耗,还显

著缩短了工程周期,减少了建设与运行

"点亮第一盏灯"是BEST项目最形 象的使命,也是我国聚变能源研究的中 期目标。BEST 计划于 2027 年完成建 设,并将在2030年实现世界首次聚变能 发电演示,其示范成果将为建设可控核 聚变电站提供核心技术支撑,加速全球 聚变能源商业化进程。

从 EAST 到 BEST, 我国正在实现从 科学验证到工程实践的关键转型。当 聚变工程迈向商业化应用,人类或将迎 来跨时代的能源革命。





科技连着你我他 ■本期观察:王 涛 汤继泽

2025年,人形机器人迎来了新一 波开发浪潮。全球范围内,多款人形机 器人相继发布。得益于人工智能、环境 感知与精密控制技术的发展,新一代人 形机器人正在以一种前所未有的敏捷 与智能,加速进入现实世界。本期科技 云,为您介绍3款新型人形机器人。

## 轮式人形机器人



我国农业机械企业中联重科成功 研发出一款轮式人形机器人,目前样 机已具备基础生活场景动作执行能力 及工业物流作业功能,可完成物料搬 运、智能分拣等典型仓储场景任务。

技术团队称,该型机器人研发过 程中实现多项核心技术突破:首创全 场景多模态跨尺度环境感知系统,结 合深度学习算法实现行为意图精准解 析;创新性构建"视觉+力觉+触觉"三 维感知融合体系,开发出通用型智能 抓取解决方案;攻克基于本体安全感 知的双臂协同运动规划 ……

该型轮式物流机器人主要面向仓 储物流场景,通过引入AI原生云平台, 建立控制全身运动的"视觉—语言— 动作"模型库和动作基元库。目前,该 型机器人在完成工厂物流搬运、分拣 等作业中,展示出有效性和应用潜力。

## 模块化通用型机器人

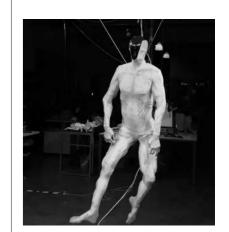


英国仿生科技公司 Humanoid 发 布了一款通用型人形机器人HMND-01。这款机器人采用模块化系统架构 设计,其硬件平台支持上下半身及末端 执行器的灵活组合配置。这种"模块化 机器人"理念不仅便于功能扩展,更能

该人形机器人身高175厘米、体 平均续航时间4小时,可满足工业场 景的常规作业需求。其整体造型模拟 人体形态,包含头部感知模块、仿生躯 、灵巧手臂及双足行走机构等。通 过配置41个高精度运动关节,该机器 人不仅实现了类人的自然步态,更展 现出优异的操作能力。

其技术团队表示,HMND-01当 前主要聚焦工业自动化领域,在高速 高扭矩执行器的驱动下,可完成物料 搬运、精密组装、包装分拣等任务。

## 双足肌肉骨骼机器人



波兰人形机器人研发企业 Clone Robotics推出了全球首款具备仿生骨 骼肌肉系统的双足人形机器 Protoclone V1。这款产品通过整合约200 个运动关节、1000组人工肌肉束及 500个高精度传感器,构建出高度仿 生的人体运动系统,以实现接近真实 人类的运动协调性。

研究团队表示,该机器人的核心创 新在于其仿生驱动架构——仿生水冷 散热系统模拟人体汗腺排热机制,可以 有效解决高密度人工肌肉群工作时的 过热问题,确保持续机体运动的稳定性。

作为全球首款突破肌肉骨骼协同 控制瓶颈的人形机器人, Protoclone V1 标志着仿生机器人技术从传统机 电驱动向生物启发式驱动的重要跨 越。其模块化架构设计不仅为后续产 品迭代奠定基础,也为医疗康复、特种 作业等领域提供了全新的技术范式。