

# 人工智能时代计算机信息安全与防护研究

郭鹏飞

中华通信系统有限责任公司河北分公司 河北 石家庄 050000

**摘要:** 本文围绕人工智能时代的计算机信息安全展开, 首先指出信息安全的重要性及其防护的紧迫性。随后, 深入分析当前面临的数据泄露、算法模型漏洞、新型网络攻击手段以及智能设备的物理安全和伦理风险等挑战。文章最后提出了一系列防护策略, 包括强化数据加密与访问控制、提升算法模型鲁棒性、有效应对新型网络攻击、加强物理安全与伦理法规建设, 以期为保障信息安全、促进AI技术健康发展提供参考和指导。

**关键词:** 人工智能时代; 计算机信息; 安全防护

引言在数字化与人工智能快速发展的时代背景下, 计算机信息安全与防护的重要性日益凸显。信息泄露、算法模型安全漏洞、新型网络攻击以及物理安全与伦理风险等问题频发, 对个人隐私、企业安全乃至国家安全构成了严重威胁。因此, 本文旨在探讨人工智能时代计算机信息安全面临的挑战及防护策略, 以期构建安全、可靠的数字化环境提供理论参考和实践指导。

## 1 计算机信息安全与防护的重要性

在数字化时代, 计算机信息安全与防护的重要性愈发凸显。信息泄露会带来严重后果。个人层面, 大量个人信息存储于计算机, 如身份证号、银行卡信息等, 一旦泄露, 可能导致个人隐私被侵犯, 遭遇诈骗、盗刷等风险。企业方面, 商业机密关乎其生存与发展, 客户数据、核心技术等信息泄露, 会使企业声誉受损, 甚至可能在市场竞争中一败涂地。计算机系统若遭受攻击导致瘫痪, 将严重影响正常运转。对于医院, 这可能致使医疗设备无法正常工作, 危及患者生命; 对于交通部门, 会造成交通信号混乱, 引发交通事故; 对于金融机构, 交易系统瘫痪会导致资金交易无法进行, 扰乱金融秩序。关键信息基础设施如能源、通信等领域的计算机系统若被攻击, 将威胁我国主权、安全和发展利益。因此, 加强计算机信息安全与防护刻不容缓。安装杀毒软件、防火墙, 定期更新系统和软件, 设置复杂密码等, 都是简单有效的防护手段。企业和机构更要建立完善的信息安全管理体系, 提升员工安全意识。只有做好计算机信息安全与防护, 才能保障个人、企业和我国在数字化浪潮中稳健前行, 免受信息安全威胁带来的巨大损失。

## 2 人工智能时代计算机信息安全面临的问题

### 2.1 数据安全问题

在当今数字化时代, AI技术的飞速发展带来了前所未有的数据安全挑战。数据作为AI技术的核心要素, 其

全生命周期, 包括收集、处理、存储和传输等环节, 都存在着潜在的安全风险。(1) 在数据收集阶段, 随着物联网设备的普及, 各类设备广泛采集个人和企业数据。如果安全防护措施不到位, 攻击者便能轻松窃取这些数据。一些恶意软件可以通过网络漏洞潜入数据收集终端, 将收集到的原始数据传输给不法分子, 从而导致个人隐私泄露和企业信息失窃。(2) 在数据处理环节, 恶意数据注入的威胁尤为严重。攻击者可能会精心构造虚假数据, 并将其注入到AI模型的训练过程中, 导致模型训练出现偏差, 最终输出错误的结果。在金融风控模型中, 如果注入了恶意数据, 可能会导致对贷款风险的误判, 给金融机构带来巨额的经济损失。(3) 数据存储同样面临严峻的考验。数据库一旦被黑客攻破, 大量敏感数据将面临泄露的风险。此外, 虽然云存储为数据管理提供了便利, 但在云部署模式下, 数据所有权与控制权的分离使得数据的安全性和隐私保护面临更大的挑战。(4) 在数据传输过程中, 如果未采用加密等安全措施, 数据在网络传输过程中极易被截取和篡改。这种情况可能导致敏感信息泄露, 甚至引发连锁反应, 影响到整个数据生态系统的安全。

### 2.2 算法模型安全

(1) AI算法模型, 尤其是深度神经网络, 由于其非线性、大规模的特性, 使得其行为在理论上难以被完全证明。这些模型的决策过程犹如一个“黑箱”, 外界难以洞察其内部运作机制。当模型做出错误决策时, 由于缺乏透明度, 很难精准追溯原因并加以修正。在医疗影像诊断领域, AI模型可能会出现误诊, 但无法明确是数据偏差、算法缺陷还是其他因素导致的。这种不确定性不仅会延误患者的治疗, 还可能造成严重的医疗后果, 影响患者的生命安全。(2) 算法的逻辑依赖性也可能导致AI生成看似合理却不符合常理的内容, 从而产生知

识偏见与误导。在新闻推荐算法中,如果算法过度依赖某些数据特征,可能会忽略信息的多样性和平衡性,导致用户接收到片面、有偏见的信息。这种信息偏食现象会影响用户对事件的全面认知,甚至加剧社会分裂和偏见。(3) AI框架作为算法模型开发和运行的基础环境,其安全性至关重要。由于AI框架的复杂性,它们可能存在安全漏洞。一旦这些漏洞被恶意攻击者发现和利用,后果不堪设想。攻击者可能会篡改模型参数,窃取模型的知识产权,甚至利用模型进行恶意攻击,如生成深度伪造内容、发动网络攻击等,从而引发严重的安全事件。

### 2.3 新型网络攻击

随着AI技术的广泛应用,网络安全领域面临前所未有的挑战。新型网络攻击手段层出不穷,使得网络安全防御变得愈发艰难。(1) 基于AI的深度伪造技术,尤其是AI换脸和换声,已经成为不法分子实施诈骗的新工具。这些技术能够逼真地模仿他人的形象和声音,使得欺诈案件呈爆发式增长。犯罪分子通过伪造公司高管的声音,成功欺骗财务人员进行大额转账,给企业带来了巨大的经济损失。(2) AI技术也被用于强化传统网络攻击的效力和效果。利用AI自动化生成独特的恶意软件脚本和多阶段载荷,这些恶意软件不仅难以检测,而且防御难度大幅提升。传统的安全防护系统往往无法识别这些高度定制化的攻击手段,导致安全防线被轻易突破。

(3) AI技术在钓鱼攻击中的应用也日益增多。通过AI创建的大量极具个性化且极具说服力的钓鱼内容,能够精准地针对特定目标人群。这些钓鱼邮件或信息看起来非常真实,诱使目标人群点击恶意链接,从而泄露敏感信息,如个人银行账户、密码和社会安全号码等。

### 2.4 物理安全与伦理风险

AI技术的发展无疑为我们的生活带来了极大的便利,从智能家居到智能交通,从在线教育到医疗诊断,AI的应用无处不在。然而,正如一枚硬币有两面,AI技术在给我们带来便捷的同时,也暴露出了一系列物理安全和伦理道德方面的潜在威胁。在物联网环境中,无人机、自动驾驶汽车等智能设备依赖大量传感器和AI技术进行信息解析和自动操作决策。这些设备在提高效率、节省人力物力的同时,也存在安全隐患。一旦AI系统的数据分析和决策过程遭受黑客攻击或恶意干扰,后果不堪设想。自动驾驶汽车的AI系统若被黑客篡改行驶指令,可能导致车辆失控,引发严重的交通事故,威胁乘客的生命安全。在伦理道德方面,AI技术的滥用问题同样不容忽视。人脸识别技术在提高社会治安、便捷生活的同时,也存在隐私泄露的风险。如果数据收集和使

用不当,可能会侵犯个人隐私,甚至被用于非法监控。一些AI决策系统在招聘、贷款审批等场景中,可能因算法偏见而导致不公平的决策结果。这种偏见可能无意中损害特定群体的利益,加剧社会不公。

## 3 人工智能时代计算机信息安全防护策略

### 3.1 加强数据加密与访问控制

当今数字化浪潮的推动下,数据已经成为企业与组织的核心资产,其安全性直接关系到企业的生存与发展。为了确保这些宝贵的数据资产安全无虞,采用先进的数据加密技术成为了必不可少的措施。(1) 对称加密算法,如AES(高级加密标准),是数据传输过程中的常用加密手段。在这种加密方式中,发送方使用一个密钥对数据进行加密,而接收方则使用相同的密钥进行解密。即使数据在传输途中被截获,没有密钥的攻击者也很难将其破解,从而保证了数据的安全。对称加密算法因其高效性和相对简单的实现方式,在许多应用场景中得到了广泛的应用。(2) 在数据存储环节,非对称加密技术发挥着关键作用。非对称加密使用一对密钥,即公钥和私钥。公钥用于加密数据,而私钥则用于解密。这种加密方式特别适用于需要高度安全性的场合,如银行对客户账户信息的存储。通过公钥加密的数据,只有对应的私钥才能解密,大大提高了数据的安全性。(3) 除了加密技术,严格的访问控制也是守护数据安全的重要防线。基于角色的访问控制(RBAC)模型是一种常用的访问控制方法。在这种模型中,用户根据其组织内的角色被分配相应的权限。在企业中,财务人员可以访问财务数据,而研发人员则对研发资料有相应的访问权限。这种权限分配机制避免了普通员工因权限过高而随意访问敏感数据,有效防止了内部人员未经授权的访问与操作,从而降低了数据泄露的风险<sup>[1]</sup>。

### 3.2 提升算法模型的安全性

在人工智能领域,算法模型的鲁棒性是抵御各种恶意攻击的关键因素。随着技术的发展,对抗性攻击手段日益多样化和复杂化,因此增强算法模型的鲁棒性变得尤为重要。对抗训练是一种有效的提升模型鲁棒性的方法,它通过在训练过程中故意引入对抗样本,使模型学习如何识别并抵御这些攻击,从而提高模型对恶意输入的抗性,防止输出错误结果。在图像识别模型的训练中,对抗样本的加入可以帮助模型更好地理解数据的真实分布,而不是仅仅记住训练数据。这种方法使得模型在遇到未曾见过的攻击时,也能保持较高的识别准确率和稳定性。通过在训练集中添加经过微小扰动的图像,模型能够学习到如何忽略这些扰动,专注于图像的本质

特征,从而在测试时能够正确识别出未被篡改的图像。引入模型解释性技术也是提高模型安全性和可信度的重要途径。LIME(局部可解释模型无关解释)技术是一种流行的解释性方法,它能够为模型的决策提供局部解释。LIME通过近似原始模型的行为,为特定的预测生成一个可解释的模型,将复杂模型的决策过程转化为人类可理解的形式。在医疗诊断模型中,这种解释性技术尤为关键。它允许医生理解模型为何得出某种诊断结果,这不仅有助于医生信任模型,还能帮助发现模型可能存在的潜在安全漏洞<sup>[2]</sup>。

### 3.3 应对新型网络攻击

在当今网络环境日益复杂的背景下,部署智能防御系统已经成为企业和社会组织的当务之急。(1)基于AI的入侵检测系统是智能防御体系的重要组成部分,它能够通过分析网络流量的模式、频率等特征,实时监测网络状态。这种系统具有高度的自适应性和学习能力,能够迅速识别出异常流量,如短时间内的大量恶意请求,并及时采取措施阻断攻击,有效应对自动化恶意软件脚本等新型网络攻击。(2)除了技术层面的防御,员工安全意识培训也是防范网络攻击的重要环节。企业应定期开展网络安全培训,向员工介绍最新的网络攻击手段,如深度伪造欺诈案例,以及如何识别和防范这些攻击。通过模拟钓鱼邮件等实际操作演练,可以提高员工的防范意识和能力,减少因人为错误导致的网络安全事件。

(3)建立一套完善的应急响应机制对于应对网络安全事件至关重要。企业需要制定详细的应急响应计划,明确各部门在安全事件发生时的职责和处理流程。一旦遭遇数据泄露,应急响应计划能够迅速启动,及时通知受影响的用户,采取数据恢复措施,以降低损失并尽快恢复系统的正常运行<sup>[3]</sup>。

### 3.4 加强物理安全与伦理监管

(1)在智能技术飞速发展的今天,智能设备的物理安全同样是一个不容忽视的重要议题。以无人机为例,为了防止其被恶意控制或造成意外伤害,必须设置多重

安全防护机制。这包括安装防撞系统以确保无人机在飞行中能够自动避开障碍物,以及采用加密通信链路来防止信号被截获或篡改。同样,对于自动驾驶汽车,硬件加密芯片的应用保障了车载系统的安全,防止黑客入侵并篡改行驶指令,确保乘客和行人的安全。(2)建立AI伦理制度框架也迫在眉睫。需要通过立规来规范AI的数据收集和使用行为,明确侵犯个人隐私的制度责任,以保护公民的个人信息不被滥用。在招聘、贷款审批等场景中,应当规定算法必须遵循公平公正的原则,防止算法偏见导致对特定群体的不公平决策,维护社会的公正和谐。(3)加强国际合作在应对AI带来的挑战方面具有重要意义。各国应当共享AI安全威胁情报,共同制定全球性的AI安全标准,以协同应对跨国网络攻击和其他安全威胁。通过国际合作,可以更有效地解决AI带来的全球性风险与挑战,推动AI技术的健康发展,同时确保技术的进步不会损害人类的根本利益<sup>[4]</sup>。

结语:综上所述,人工智能时代的计算机信息安全与防护面临着前所未有的挑战,从数据安全、算法模型安全到新型网络攻击和物理安全及伦理风险,每个方面都亟需加强。未来,随着技术的不断进步,我们必须持续探索和完善信息安全策略,同时促进国际合作,制定统一的安全标准,确保AI技术健康发展的同时,保护个人隐私和社会稳定,为数字化时代的信息安全提供坚实保障。

### 参考文献

- [1]童辛迪.人工智能时代计算机信息安全与防护研究[J].科学与信息化,2024(19):50-52.
- [2]安仲源.基于人工智能的计算机信息安全与防护研究[J].信息记录材料,2024,25(6):161-163.
- [3]刘仪.人工智能时代计算机信息安全与防护研究[J].网络安全技术与应用,2022(4):173-174.
- [4]苏彤.人工智能时代计算机信息安全与防护研究[J].数字化用户,2022,28(27):43-45.