新闻热线：010—58884062
E-mail：erin74mail@163.com
■责编 翟冬冬
科技日报 2019年4月15日 星期一

戴上“白帽子” 人工智能投身网络安全攻防战

本报记者 **刘园园**

面对计算机系统和网络的缺陷和漏洞，黑客们找准机会实施攻击，白帽黑客则利用黑客技术来测试网络和系统的性能以判定它们能够承受入侵的强弱程度。短短几年时间，人工智能已进驻多个行业，落地无数场景。其中一些行业和场景已为大家所熟知，还有一些正在进入我们的视野，在网络安全领域，AI“白帽”正成为网络安全工程师的得力助手。

据报道，近日美国市场调研公司CB Insights发布报告预测了2019年人工智能行业的发展趋势，其中一个趋势便是用人工智能发现网络威胁。

正在赋能网络安全

“人工智能技术的蓬勃发展，为网络安全攻防带来的，不仅有机遇，也有挑战。”北京理工大学网络攻防对抗技术研究所所长闫怀志接受科技日报记者采访时说。

先说好的一面。360安全研究院邹权臣博士告诉科技日报记者，目前人工智能已经应用于恶意代码检测、恶意流量检测、威胁情报收集、软件漏洞挖掘等网络安全领域。

“例如在恶意代码检测方面，人工智能通过对恶意程序的API调用序列、系统CPU利用率、收发的数据包等信息，自动识别恶意代码的特征，进而判定分类。”邹权臣介绍，相比于传统的基于动静态分析的特征检测、启发式检测技术，人工智能可以大幅度提升检测的准确率。

具备明显“过人之处”

与传统的应对网络安全的方式相比，人工智能确实展示了其“过人之处”。

在闫怀志看来，人工智能方法在解决人力所不及的安全大数据统计和抽取规律方面具备天然优势，它能够全面提高威胁攻击的识别、响应和反制速度，提升风险防范的预见性和准确性。特别是在异常行为检测等应用场景模糊的非精确识别和匹配方面，更是如此。

“人工智能针对未知威胁和攻击的检测也更出色。因为传统的特征匹配方法对未知威胁几乎无能为力，而人工智能方法有时不需要先验知识，对未知威胁的检测能力较强。”闫怀志说。

360安全研究院研究员张德岳介绍，在软件漏洞挖掘方面，采用人工智能技术从漏洞相关的数据中提取经验和知识，并用训练好的模型提高漏洞挖掘的精度和效率，可以缓解当前该领域研究遇到的一些瓶颈问题，具体应用场景包括漏洞程序筛选、源代码漏洞点预测等。

“人工智能在网络安全领域的应用日益广泛，运用人工智能赋能网络空间安全，主要体现在主动防御、威胁分析、策略生成、态势感知、攻防对抗等诸多方面。”闫怀志说，其中包括采用人工神经网络技术，来检测入侵行为、蠕虫病毒等安全风险源；采用专家系统技术，进行安全规划、安全运行中心管理等；此外人工智能方法还有助于网络空间安全环境的治理，比如打击网络诈骗。

不得不说，人工智能系统还具备成本效益优势。闫怀志认为，人工智能可以在第一时间发现和识别预防威胁，并立即启动应急响应，高效的智能检测流程有助于减少人工参与、简化流程、降低成本、减小损失。

“传统的应对网络安全的方法依赖于人工硬编码定义、提取特征的方式完成相关任务，而人工智能可以直接对原始数据进行训练，从大量的数据中提取特征，自动完成分类判定的工作。”张德岳说，如此一来后者既可以提高网络安全中预测、防范、检测、销控等各个风险环节的自动化和智能化程度，又能提升响应速度和判定的准确率。

不能靠它包打天下

“虽然人工智能搅动了网络安全领域的一池春水，但是应该理性看待人工智能在应对网络安全方面的优缺点，不能指望全靠人工智能来包打天下。”闫怀志说，人工智能在应对网络安全问题时，也有较强的局限性。

邹权臣分析，这一方面受限于人工智能算法本身的能力。因为传统的机器学习技术依赖特征提取，而算法的效果和性能又依赖识别和提取特征的准确性。深度学习具有在高维数据中自动提取特征的能力，同时面临着持续学习、数据饥饿、可解释性等问题。

“另一方面机器学习、特别是深度学习过分依赖数据，但在恶意代码检测、软件漏洞挖掘等领域，目前仍然存在数据收集困难的问题，缺少较好的数据集用于训练，影响对相关领域的研究。”邹权臣补充说，另外人工智能严重依赖于耗费计算资源，复杂的深度学习网络需要同时计算成百上千万次的计算，需要强大的人工智能芯片计算力的支撑。

闫怀志则从不同方面总结了人工智能的不足。比如，易于忽视或者抛弃人类专家在网络安全领域的知识和经验积累，对网络安全的复杂应用场景考虑不足，对于已知威胁的检测效率远低于传统的精确特征识别方法等。

“使用神经网络和深度学习等算法，能够较好地识别出未知攻击威胁风险，达到‘知其然’的目的，但是这些算法通常无法揭示产生这种安全风险的基本机理，也就是‘不知其所以然’，从而为从源头防御这种攻击风险带来极大障碍。”闫怀志说。

脆弱面带来安全风险

人工智能在应对网络安全问题时，有时甚至会展现出脆弱的一面。

“一个真实环境中的人工智能系统，会面临数据安全、模型/算法安全、实现安全等多方面的安全威胁。”张德岳告诉科技日报记者。

张德岳举例说，在数据安全方面，在数据收集与标注时出现错误或注入恶意数据，将导致数据污染攻击；在模型/算法安全方面，针对人工智能算法存在黑盒和白盒对抗样本攻击，可导致识别系统出现混乱；在实现安全方面，除了人工智能系统本身的代码实现，其所基于的人工智能框架以及所依赖的第三方软件库中的软件实现漏洞，也都可能导致严重安全问题。

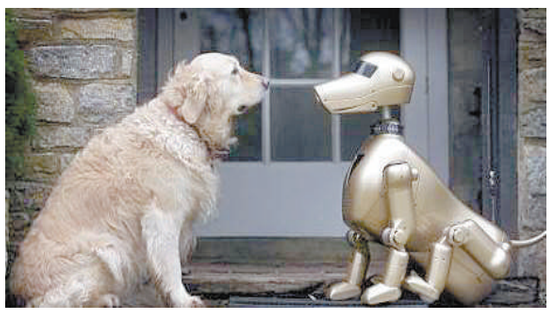
“人工智能对现有网络安全格局的影响，离不开算法、数据和计算能力3个方面，其容易遭受攻击的弱点也来自于此。”闫怀志总结说。

对于防范人工智能的脆弱性所带来的安全风险，闫怀志指出：首先要从体系架构、系统算法容错容侵设计、漏洞检测和修复、安全配置等方面来增强人工智能系统自身的安全性；其次，要用其所长，尽量减小其暴露给外界的潜在攻击面；最后要构建网络空间安全综合防御体系，从安全技术和安全管理等层面来协同防范安全攻击，间接减缓攻击者直接针对人工智能系统发起攻击以及攻击成功的可能性。

来自360安全研究院的专家也给出多个建议，其中包括：在数据获取过程中，要加强对数据来源的控制与过滤，在一定程度上保证数据安全可靠；在数据传输过程中，要使用更加安全的传输协议与加密算法；在人工智能系统的实现中，要保证代码质量并进行完善的测试，此外还要及时更新或修补框架或依赖库中存在的漏洞等。

新鲜事

认知环境能力赶超动物 人工智能才称得上“厉害”



无论下棋还是玩电子游戏，人工智能的表现似乎都强于人类。然而，近日，英国研究人员认为，只有能够像鸡、老鼠等动物一样具备灵活的认知环境的能力，人工智能才称得上“厉害”。

据新华社报道，研究人员因此举办“动物人工智能奥林匹克”竞赛，让各种人工智能设备在指定“赛场”内完成认知任务，例如，向人工智能展示某种食物并将食物挪走，看它能否重新找到食物。研究人员说，向鸡展示某种食物并当面将食物移走，鸡大概知道这种食物被拿到了哪里，并能够克服障碍重新找到食物。

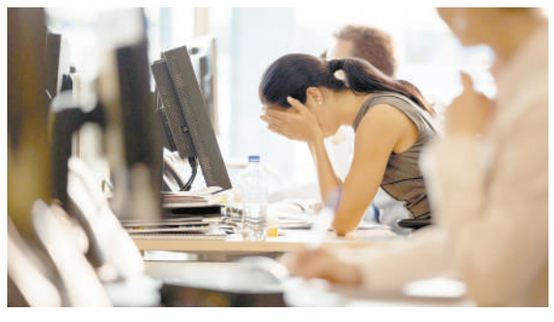
英国《泰晤士报》援引研究参与者、帝国理工学院研究人员马修·克罗斯比的话报道：“人工智能在电子游戏领域进展显著，已经能够打败人类。令人工智能熟悉环境并理解其物理构成，这是我们真正需要取得进步的领域。”

研究人员计划于4月底开放赛场，供参赛队伍训练其人工智能设备，6月开始正式竞技，让参赛人工智能设备解决一些动物日常遇到的小问题。取得优胜的参赛队伍将获得1万美元（约合6.7万元人民币）奖金。

克罗斯比说，只有当人工智能展现出与一只略微机灵的老鼠相当的认知灵活度时，才是为机器人取代人类的前景担忧的时候。

情报所

英企利用人工智能监测员工活动



何时发送邮件，何时编辑和访问哪些文件，在什么时间和哪些人见面……在你工作的同时一双“眼睛”可能正在默默注视着你。近日，据《卫报》报道，英国数十家企业正在利用一款叫做Isaak的人工智能程序监测员工的工作状态。

这款软件可以按照贡献度将员工排名，并且将收集到的数据和员工档案、销售业绩等进行比较，以找出员工动态和工作产出的关系。目前该系统已经收集了13万员工的10亿行动数据。

对此，有工会警告，这种做法可能增加员工压力并导致信任危机。还有批评者认为利用人工智能监测员工行为有可能对员工的心理健康构成威胁。赫特福德大学教授乌苏拉·胡斯表示：“如果他们将手从键盘上移开五分钟，将被视为不工作。但是他们可能在思考，而这些指标并没有得到衡量。这对于需要大量创新的工作将产生什么影响？”

但也有支持者表示，基于人工智能的动态监测有助于增加工作效率并消除工作中因社会偏见而导致的不平等因素。

欧盟发布人工智能伦理准则

欧盟委员会近日发布人工智能伦理准则，以提升人们对人工智能产业的信任。欧盟委员会同时宣布启动人工智能伦理准则的试行阶段，邀请工商企业、研究机构和政府机构对该准则进行测试。

据外媒报道，该准则由欧盟人工智能高级别专家组起草，列出了“可信赖人工智能”的7个关键条件——人的能动性和监督能力、安全性、隐私数据管理、透明度、包容性、社会福祉、问责机制，以确保人工智能足够安全可靠。欧盟将“人工智能”定义为“显示智能行为的系统”，它可以分析环境，并行使一定的自主权来执行任务。

根据官方解释，“可信赖的人工智能”有两个必要的组成部分：一是应尊重基本人权、规章制度、核心原则及价值观；二是应在技术上安全可靠，避免因技术不足而造成无意的伤害。

对于这份准则，欧委会副主席兼欧盟单一数字市场战略副总裁安德鲁斯·安西普表示：“符合伦理标准的人工智能将带来双赢，可以成为欧洲的竞争优势，欧洲可以成为人们信任的、以人为本的人工智能领导者。”

（本版图片来源于网络）

扫一扫
欢迎关注
AI瞭望站
微信公众号

避免同案不同判，海南来了位AI“法官”

第二看台

王祝华 本报记者 **江东洲 刘 昊**

“量刑规范化智能辅助办案系统很受基层法官的欢迎，这是一个可以给年轻法官提供经验，给年长法官提供保护的最新科技成果。”海南省琼海市人民法院副院长王春豹说。

王春豹所说的量刑规范化智能辅助办案系统，是海南省高院引入大数据和人工智能技术、自主研发的覆盖量刑规范化改革的全方位智能量刑系统，它不仅能辅助法官办案，还破解了量刑公正难题。这也使得海南量刑规范化工作一直走在全国前列，多次受到最高法院的肯定，一些经验在全国推广。

当前，我国正在大力发展人工智能技术，各地政府部门均纷纷尝试引入人工智能手段进行辅助决策，在法律领域人工智能也开始崭露头角。

解决“案多人少”的矛盾

“人工智能技术在法律领域的应用正在逐步深入，海南省高院走在全国前列，最大限度地运用人工智能技术来辅助法官办案。”海南省高级人民法院刑事审判庭副庭长吴向东介绍。他曾在量刑规范化智能辅助办案系统开发过程中给出了不少建议。

海南省的量刑规范化智能辅助办案系统综合了大数据、自然语言处理、知识图谱、深度学习等大数据及人工智能技术，模拟法官审理量刑规范化案件过程，自动识别智能提取并回填案件要素，智能分析并运用历史量刑数据，智能生成裁判文书和法律文书等。该系统自主研发的法律知识图谱及法律NLP平台等获得了10多项发明专利。

吴向东说，系统根据各单位量刑实施细则的具体差异以及法律法规调整变化的情况进行实时更新，整个系统的智能性不仅仅是因为使用了AI技术而智能，而是因为AI与法律业务的深度融合，在合适的业务场景使用了合适的技术。

北京智慧正安公司总经理李正才告诉科技日报记者，系统开发过程中，海南多名专业法官亲自参与大量用于机器学习的文书标注工作，创新性地将贝叶斯算法进行灵活应用，在不改变法官思考、工作习惯的前提下，大幅度提升刑事量刑效率和准确率；系统还会根据法官的使用习惯，不断地进行自我学习，提升智能识别提取案件要素的准确性。

“一个上午，我在琼海法院就开了三个庭并随案作出了判决书，这样的效率在以前是不可想象的。”王春豹说。

量刑规范化智能辅助办案系统已经在海南全省三级法院使用，根据使用单位汇总的数据显示，在法官员额制改革后，人手紧张的情况下，海南法官办理量刑规范化案件的时间减少约50%，制作裁判文书的时间缩短约70%，制作程序性法律文书的时间减少近90%，大幅度减轻了法官量刑办案的工作量，进一步解决“案多人少”的矛盾，有效避免同案不同判现象，提高刑事法官精准办案效率。这是“人工智能+大数据+智慧法院”的典型应用。

不能承担政策制定的责任

在感受到人工智能实实在在的便利之后，法律界人士也纷纷探讨：未来人工智能会不会成为法律主体，抢了律师甚至是法官的“饭碗”呢？

中山大学逻辑与认知研究所副所长熊明辉教授认为，法律界深度使用人工智能的原因有两个：一是案多人少的形势日益严峻，促使政法机关不得不采取更加高效率的手段去应对。二是公正司法的需要。通过人工智能建立一种标准化的裁判基准或裁判模型，将其运用到案件处理当中，会使裁判结果更加公正。

在海南外经律师事务所律师郭俊看来，现在谈人工智能是否为法律主体为时过早，毕竟人工智能只是一种趋势，而非普遍存在。人工智能作为高智能工具，是人类的好帮手，可以因保护需要成为法律的客体，但无法成为法律的主体。“尽管如此，我们仍需谨慎界定人机之间的关系。”

国务院在《新一代人工智能发展规划》中提出，“建立人工智能法律法规、伦理规范和政策体系，形成人工智能安全评估和管控能力”。为智能社会划出法律和伦理道德的边界，让人工智能服务人类社会。这也是世界范围内的一项共识。

李正才认为，人工智能技术在专业法律领域的最大隐患是由人工智能技术替代法官进行了价值判断，这是绝对不行的，最终的价值判断还是必须由法官做出。

“在公共领域，由于当前我国基层法律服务极度欠缺，人工智能技术恰好可以发挥其快速甄别、智能联想等优势，以达到一个初、中级律师的能力水平，让基层群众24小时全天候随手获取法律服务，人工智能技术在法律界这个方向最能发挥其优势。技术是把双刃剑，若提供服务的企业不能很好地将AI技术与法律业务进行深度融合的话，可能会带来错误的咨询意见或错误的引导。”李正才说。

清华大学电子工程系博士谢耘也向记者表示，人工智能在法律界或者说政府公共决策上只能做辅助作用，不能承担政策制定的责任，这是原则。

谢耘认为，人工智能的隐忧在于人对它的盲目信任，以及我们对人工智能的可信边界不清楚，或者说缺少经验积累。人工智能应该用科学技术的方法来分析，而不是使用艺术的方法来想象，不应强加给人工智能根本不具备的特征、性质、属性等。