



人工智能时代的网络空间安全专题简介

李琦^{1*}, 任奎², 季新生³, 陶小峰⁴, N. ASOKAN⁵,
Elisa BERTINO⁶, Zhengkai LIANG⁷

1. 清华大学, 北京 100084, 中国
2. 浙江大学, 杭州 310030, 中国
3. 紫金山实验室, 南京 211111, 中国
4. 北京邮电大学, 北京 100876, 中国
5. University of Waterloo, Waterloo N2J 3G1, Canada
6. Purdue University, West Lafayette IN 47907, USA
7. National University of Singapore, Singapore 119077, Singapore

* 通信作者. E-mail: qli01@tsinghua.edu.cn

近年来,人工智能和机器学习的迅速发展,既给网络空间安全带来了机遇也带来了挑战:大量的人工智能和机器学习算法应用在攻击流量识别、入侵检测系统和恶意软件识别等领域,有效提高了网络空间的安全;同时,人工智能和机器学习算法也存在着不可解释性的缺陷,简单构造的对抗样本有可能导致机器学习算法输出错误的预测结果等.应对这些挑战需要研究安全鲁棒的人工智能和机器学习算法及相关应用.因此, *SCIENCE CHINA Information Science* 在2022年第65卷第7期组织出版了“人工智能时代的网络空间安全专题”(Special Focus on Cyber Security in the Era of Artificial Intelligence).经过严格的同行评审,本专题共收录8篇文章.

综述论文(Intelligent networking in adversarial environment: challenges and opportunities)总结分析了在对抗场景下神经网络所面临的机遇和挑战,并探讨了对抗机器学习在对抗场景的应用.5篇研究论文分别探讨了机器学习在不同典型场景中的应用,包括:物联网场景下基于联邦学习的纹理表示问题(Non-IID federated learning via random exchange of local feature maps for textile IIoT secure computing)、基于机器学习网络入侵系统的后门攻击问题(VulnerGAN: a backdoor at-

tack through vulnerability amplification against machine learning-based network intrusion detection systems)、基于机器学习的电磁频率指纹认证问题(Reliable resource allocation with RF fingerprinting authentication in secure IoT networks)、基于深度强化学习的网络探测攻击防御问题(Defensive deception framework against reconnaissance attacks in the cloud with deep reinforcement learning)、基于随机平滑的补丁攻击可验证防御问题(Certified defense against patch attacks via mask-guided randomized smoothing).此外,专题也研究探讨了机器学习的隐私保护问题,包括隐私保护的联邦学习(ACCEL: an efficient and privacy-preserving federated logistic regression scheme over vertically partitioned data)以及基于机器学习的安全数据交易(Post quantum secure fair data trading with deterability based on machine learning).

综上,本专题关注人工智能和机器学习对网络空间安全的机遇和挑战,旨在促进人工智能和机器学习的相关安全研究.在此,我们衷心感谢投稿本专题的所有作者,并向所有匿名审稿人给予的及时细致的评审工作致以诚挚的谢意.

引用格式: 李琦, 任奎, 季新生, 等. 人工智能时代的网络空间安全专题简介. 中国科学: 信息科学, 2022, 52: 1362, doi: 10.1360/SSI-2022-0276