报告题目

报告题目：图神经网络中的黑盒与逃逸攻击研究

报告人：周潘

报告时间：2021 年 4 月 9 日，星期五，10：00-11：30

报告地点：四川大学江安校区网络靶场会议室

报告内容：

神经网络的兴起与应用成功推动了模式识别和数据挖掘的研究。尽管传统的深度学习方法被应用在提取欧氏空间数据的特征方面取得了巨大的成功，但许多实际应用场景中的数据是从非欧式空间生成的，传统的深度学习方法在处理非欧式空间数据上的表现却仍难以使人满意。近年来，“图神经网络（Graph Neural Networks，GNN）”已成为一个新的研究热点，其在诸如点云分类、动作识别等计算机视觉领域和推荐系统，以及推断蛋白质结构等领域有许多杀手级应用。然而，图神经网络在敌手环境下无法抵抗一些轻微的扰动，其健壮性仍有待提高。如在信用评分系统中，欺诈者可以伪造与几个高信用客户的联系，以逃避欺诈检测模型。因此，对图神经网络在对抗攻击下的健壮性研究具有极其重要的研究价值。在图神经网络的对抗攻击中，黑盒攻击和逃逸攻击只允许攻击者查询模型从而获得查询结果，攻击所掌握的信息程度最小，其相较于白盒攻击而言更具有实际意义。本次讲座主要介绍针对图神经网络的黑盒攻击（Black-box attack）以及逃逸攻击（Evasion attack）等相关内容，并介绍我们在这个方向上的三个工作：基于影响函数的图神经网络逃逸攻击、基于强化学习的黑盒逃逸攻击算法以及基于 bandit 的图神经网络黑盒攻击结构扰动攻击算法。

报告人简介：

周潘，华中科技大学网络空间安全学院教授/博士生导师，武汉国家网安基地存储与大数据安全研究所所长，IEEE 高级会员、ACM 会员。2002 年-2008 年就读于华中科技大学，先后获得学士和硕士学位。2011 年获美国佐治亚理工学院，电子与计算机工程学院哲学博士学位。2011 年至 2013 年，担任美国甲骨文公司波士顿地区 Oracle 云平台大数据分析与刷新算法研究员。自 2013 年 6 月起任教于华中科技大学，主要从事大数据分析与隐私保护、人工智能安全、边缘计算与物联网安全方向的研究。主持国家自然科学基金项目 2 项，参与国家 863 重大项目一项。2017 年获华中科技大学“科技新星”。2020 年，获 25 届国际模式识别大会 25th International Conference on Pattern Recognition (ICPR2020)“最佳科学论文奖”。在 IEEE TDSC, IEEE TIFS, IEEE TIT，IEEE TON, IEEE INFOCOM,

IEEE CVPR, IEEE ICCV、IEEE ICDE 等重要期刊与会议上发表学术论 150 余篇。自 2020 年 9 月起，担任 IEEE Transactions on Network Science and Engineering 副编辑。

欢迎广大师生踊跃参加！

网络空间安全学院

2021 年 04 月 07 日