人工智能在智慧城市领域应用的安全性分析

Security Analysis of Artificial Intelligence Application in Smart City

姚树为，陈　龙，沙默泉，王　悦（中国联通智能城市研究院，北京 100048）
Yao Shuwei，Chen Long，Sha Moquan，Wang Yue（China Unicom Smart City Research Institute，Beijing 100048，China）

摘　要：
人工智能技术在智慧城市的众多场景有广泛的应用，主要包括交通、物流、环境、政务、安防。在智慧城市的建设过程中，人工智能技术的运用可有效提升城市信息的全面感知和协调能力，与此同时，数据安全、网络安全、算法安全等一系列安全性问题是智慧城市建设面临的严峻挑战。建立相应的规范机制，采取相应的应对措施，形成开放共享的行业共识，既可促使人工智能技术更好地服务于智慧城市建设，又能为我国新型智慧城市规划与发展奠定基础。

关键词：
人工智能；安全性分析；智慧城市
doi：10.12045/j.issn.1007-3043.2023.01.003
文章编号：1007-3043（2023）01-0011-04
中图分类号：TN918
文献标识码：A
开放科学（资源服务）标识码（OSID）：

Abstract：
AI technology is extensively apply to many spectacle of smart cities，mainly including transportation，logistics，environment，government affairs and security. During the development of smart city，the application of AI technology can effectively improve the synthetic perspective and collaboration skills of city informatization. In the same time，a sequence of safety problems such as data security，network security，algorithm security and so on make the construction of smart city face forbidding crunchs. Building-up a corresponding normative mechanism，taking corresponding countermeasures，and shaping an open and shared industry consensus can not only accelerate artificial intelligence technology to better give service to the development of smart cities，but also lay the foundation for the planning and developing of new smart cities in China.

Keywords：
Artificial intelligence；Safety analysis；Smart city

引用格式：姚树为，陈龙，沙默泉，等. 人工智能在智慧城市领域应用的安全性分析[J]. 邮电设计技术，2023(1)：11-14.

0 引言

智慧城市是城市信息化建设的新阶段，是城市智慧化发展和进步与技术融合创新的新时代产物。人工智能在智慧城市建设过程中扮演着重要角色，实现了信息动态感知，决策智能高效[1]，为社会进步和发展注入活力，提升城市发展精细化水平，提升人民生活质量。充分利用人工智能技术，将现代信息技术融入到城市建设过程中，满足智慧城市领域各场景需求，有效保证智慧城市规划发展的效果和质量。以数据为驱动，整合物联侧数据、社会侧数据、政府侧数据，提供持续高效的城市运营支撑，是智慧城市领域人工智能技术应用的核心所在[2]。智慧城市作为一种新的时代产物、新的发展模式，必然面临着安全机制不完备、政策法规不健全、数据共享不充分的问题。本文结合人工智能在智慧城市领域的应用现状，分析人工智能在这些领域应用时面临的安全性问题，并给出对应的防范措施，有助于增强智慧城市从业者安全意识，进而提升智慧城市规划和发展水平。

收稿日期：2022-12-20

1 人工智能在智慧城市领域应用现状

人工智能技术的不断成熟，有效提升了城市运营

管理水平。城市管理数字化逐渐向智能化转型升级，深度学习为图像识别技术提供动力，在城市安全和城市管理方面有着举足轻重的作用。例如智能视频识别应用，在交通领域可以对人流和车流的聚集进行检测，对通过指定界限或者区域的车或者人进行自动识别、登记、统计，也可以实现在人流或者车流达到既定阈值时进行报警；在城市运营管理方面，通过对实时视频流的观测分析，对监控范围内出现的垃圾乱堆乱放、车辆乱停、工程违规作业、占道晾晒等现象进行及时的上报处理，全过程跟踪处置；在治安管理方面，对监控视频进行分析，或者检测视频中对象的异常行为（打架斗殴、人员非法聚集、拥挤踩踏），为治安管理提供强有力的支撑；在城市居民生活方面，自动识别技术也广泛引入各类身份认证应用，人脸支付、免身份证进站等场景是人脸识别技术普及最为广泛，应用频次比较高的场景。自动识别技术有效提升了办事效率，降低重复性工作，提升了城市居民生活水平。

a）安防。在园区、社区、楼宇、政府领域等场所，通过摄像头进行数据采集，以视觉识别技术为支撑。视频监控、智能门禁、智能梯控、消防预警的应用都将传统防御式的安防逐步向主动判断、可分析预警安防发展。可以提供实时的可视化能力、智能化的解决方案。随着边缘计算的蓬勃发展，可以使算力不再聚集于云端，每个设备都有数据处理的能力。但新型智能化设备的高成本限制了安防行业智能化的进程。

b）政务。智能预审，让老百姓少跑路，提升政府服务能力，降低工作强度和复杂度。图像识别技术对通过网络途径提交的材料进行审核，对于其中明显不符合规范的地方进行修正或者提出相应的修改意见。网上预约，通过图像识别录入人像信息，办事人员在规定的时间段到达服务大厅，可以通过人脸识别进行身份认证自动取号。辅助决策，可以对服务满意度低、办件时间长、老百姓跑腿次数多的情况进行统计分析，不断优化政府服务手段，提高政府办事效率。

c）交通。交通情况实时监控，能够及时完整地将交通环境与各交通参与者的相关信息监测记录下来，中央控制中心高效控制监控系统，机器学习等技术的引入，也使得监控系统具备高度的自适应能力和自我学习能力；智能交通仿真，可以模拟交通路面环境以及周边的基础设施，摸清交通运行规律，提高智慧交通科学决策水平；应急管理，可以根据道路、天气及周边的基础服务设施情况，构建高效统一的应急管理平台，对事件发生、跟踪、复盘进行全生命周期的管理。

d）环境保护。人工智能技术应用于环境保护，融入环境监测、应急管理。环境传感器可以收集温湿度、雨量、光照、风速风向等数据。监测摄像头可以监测土壤、大气、固废、土壤污染情况。环境监测系统主要由数据感知流程、网络传递流程、应用分析流程组成，人工智能技术是构建环境监测系统的重要技术支撑。各类传感器协调工作，对环境中各项指标进行监测，可以根据信息来源选择不同的路由策略，进而选择合理的数据信息传递通道。应用层获取到的数据，一方面可以直接通过可视化平台进行实时的展示，另一方面对其中出现的异常数据也可以进行统计分析，辅助决策者做出合理化的调整。针对智能设备获取到的数据，应用人工智能技术做数据处理和分析，比如数据挖掘、深度学习，为环境保护领域的智慧化决策提供支撑，加速城市智慧转型。

e）物流。产品服务开发、营销和销售、服务运营和供应链管理是人工智能在智慧物流场景下的主要业务功能。无人仓，现在的主流快递服务公司通过自动标识扫描，由仓库机器人来负责货物的分拣、存储、运输，有效降低人工成本，提升工作效率；动态定价，根据产品供应量、竞品信息、客户数据等数据，利用机器学习算法分析客户历史行为记录，动态调整产品价格，应对不同时期的需求波动；智慧配送，随着无人驾驶技术的不断完善，可以实现实时跟进道路信息，做好配送路径选择，节约配送时间，保证配送过程的及时性和准确性。

2 安全问题分析

针对人工智能在智慧城市领域的应用现状，本章从数据、算法、网络3个层面分析人工智能在不同场景下应用可能出现的安全性问题。

2.1 数据安全

人工智能在智慧城市领域的应用依赖大量的业务数据，这些业务数据就成为人工智能发挥作用的数据源，人工智能也依赖这些数据，运用机器学习算法而变得更加智能。引发数据安全问题主要因素包括如下几个方面（见图1）。

a）模型偏斜。攻击者在得知系统使用的模型类别之后，尝试在模型中输入数据，这些数据会逐渐降低系统识别的准确性，导致后续的攻击不再被人工智能系统识别为异常，通过控制数据源来达到控制人工

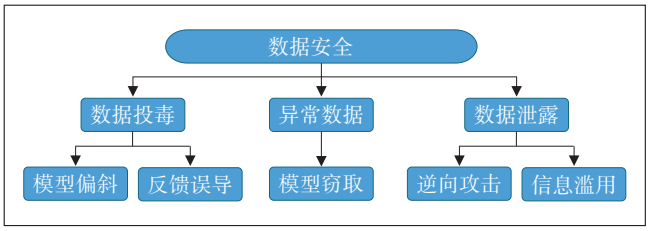

图1　数据安全

智能系统的目的。比如在公路上正常行驶的无人驾驶车辆可能突然撞向迎面而来的车辆，生产系统可能发生业务流程混乱，破坏生产活动。

b）反馈误导。则是将有缺陷的数据代入对业务大数据进行处理的算法，干扰预测处理结果，达到控制人工智能系统的目的。

c）模型窃取。通过迁移学习机制、对抗样本进行模型窃取的攻击技术，当数据原因导致系统瘫痪后，可以实现对处理数据的还原。

d）逆向攻击。能够仅依靠分析和发问响应，用目标AI的输出来训练攻击者自己的AI，攻击模型算法内部，获取数据。

e）信息滥用。医疗数据等与个人强相关的数据，信息滥用或者泄露将造成严重后果。

2.2 算法安全

算法安全主要包括如下3个方面。

a）鲁棒性。外界环境发生变化时，算法应当具有一定的韧性。健壮的算法模型拥有很高的精度，这也是机器学习中对模型算法的最低要求。模型假设偏差较小的情况下，对算法模型做出的决策影响较小，比如最常见的噪点。模型假设偏差较大的情况下，对算法模型的决策和算法模型的影响是灾难性的，比如最常见的离群点。

b）可靠性。当异常数据作为算法模型的输入时，算法不会崩溃。当前深度学习、机器学习算法的发展还处于初级阶段，在业务系统中应用低可靠性的算法，算法失效会产生一些负面作用。

c）可解释性。编写算法的人员应当明确算法具体是如何做决策的。但是一些算法的决策过程不透明，除算法编写人员外其他人无法解释算法是如何根据不同的输入数据做出不同的决策。

针对算法安全的几个方面，人工智能技术在智慧城市的应用也出现了相关的安全问题。比如图像识别技术多用于智慧安防场景中，但是有时会出现图像欺骗情况，比如利用二维照片就可以通过需要人脸动态认证的门禁系统，说明识别算法的可靠性需要有标准的测试和评估流程。在应用人工智能算法时，需要注意提升算法的鲁棒性，对抗潜在攻击。比如被恶意修改的模型文件，对于同样的处理过程，可以输出不同结果，自动驾驶领域的一些实验已经印证了这一算法漏洞。算法设计不合理、算法隐藏不平等的规则、训练集数据误差范围太大，都容易导致最终的训练结果不在预期范围内。

2.3 网络安全

人工智能设备与互联网设备通过网络进行连接，针对这些智能设备的网络攻击会带来一定的物理伤害。人工智能设备收集的数据，并不是全部传输到云端进行数据处理，有时也需要利用边缘计算将算力靠近设备侧进行相应的处理，这样也给网络攻击提供了合适的土壤。在应用的登录页面，经常会遇到除账号密码之外的验证码识别功能，以此来防止机器人恶意使用网站服务的情况。但是机器学习可以使这一验证功能失效。人工智能技术在智慧城市领域应用，而接入网络则有可能将应用置于危险环境。人工智能技术本身也可以通过接入网络窃取数据，对特征库目标进行检索查找，找出可以攻击的对象，进行数据勒索。通过机器学习、自然语言处理等技术，可以产生混淆视听的虚假情报，对系统进行骚扰攻击。

3 安全措施

基于人工智能在智慧城市领域应用可能出现的安全性问题，本文从提升数据安全治理水平、完善算法完全防护体系、强化网络安全防护、健全法律法规体系4个角度采取相应的安全措施，助力城市智慧化建设，为智慧城市各领域应用的实践拓展保驾护航。

3.1 数据安全治理

一是做好数据资产保护，主要从2个层面入手。

a）分类分级，按照数据内容进行分类分级，公共安全数据目前共分为4个等级，敏感→较敏感→低敏感→不敏感，数据的分类分级有助于提升数据开放共享能力，提高数据的利用率。比如智慧政务领域的数据可以通过建立政务数据分类分级平台，提升政府服务治理能力。

b）数据质量管理，利用深度学习算法，做好数据质量审查工作，形成知识库，提升数据质量。比如在智慧医疗行业，可以重塑源数据，输出结构化数据，有效推动智慧医疗行业的基础建设。

二是做好数据流转过程的防护，主要包括3个方面：做好数据加密、防止数据泄露、数据脱敏工作。针对智慧城市领域数据的多源异构性，应当用相应的数据处理框架进行处理。存储数据时，应当做好数据加密工作。传输数据时，防止网络侧的数据窃取攻击，避免出现数据泄露。比如在智慧政务领域的一些数据，是关于一些组织或者机构的枢纽位置、组织架构、重要人员的信息。这些数据在数据流转过程中，必须做好数据脱敏处理工作。

三是构建统一的人工智能模型评判标准，制定合理的数据安全策略，辅助决策者科学、快速做出决策。

3.2 算法安全维护

智慧城市领域应用人工智能技术，针对其中应用的一些学习型算法，可以从以下4个方面进行维护。

a）数据全生命周期，制定严格规范的安全协议，并严格遵守执行。做好加密算法模型、加密模型文件的工作。

b）做好数据训练集训练过程中的行为记录，做到行为可追踪、调整有依据。数据训练集通过算法模型进行处理，在智慧城市不同的应用场景存在一定泛化误差，输出决策也不是最优的。做好训练行为跟踪，当训练结果出现偏差时，及时对算法模型进行调整。

c）在模型资源接入层面，加强访问控制，按照合理的加密协议，做好身份认证和权限校验，防止出现恶意修改算法的情况。

d）建立完备的风险评估机制和合理的应急处置流程。当算法对训练集进行处理时，根据风险评估机制对处理过程可能出现的风险做定性和定量分析。当算法失效时，应急处置流程发挥效用，采取及时的调整措施，保证算法的失效不会影响软件系统的正常运行。

3.3 网络安全防护

人工智能在智慧城市领域应用的网络风险，可以从以下几个方面采取相应的措施。

a）在智能终端侧增加硬件的端点安全连接，硬件设备是通过众多网络节点连接在一起。防止因为互联的设备太多，一旦单一节点设备被网络攻破，大量数据都将被窃取。

b）对终端侧设备进行有效的加密，在用户操作设备结束后擦除行为记录和用户凭证信息，避免网络侧对用户凭证和设备行为记录的攻击，逆向窃取数据。

c）对数据资源的存储介质进行加密，防止数据在传输过程中接入不可信的网络，导致数据遭到破坏。

d）做好设备与人之间的交互行为记录和跟踪。

e）按需设置局域网，在复杂环境中做网络隔断。

f）IPv6新增的安全机制，可抗重放攻击，支持源发认证，提高系统的保密性，使网络安全有较大提升。因此在智慧城市领域的各应用场景下，IPv6与新兴技术的融合，必将助力城市智慧化转型和信息化建设。

3.4 健全法律法规体系

欧盟委员会于2021年4月提出AI新法规，以保证企业及个人安全为初衷，希望以法律框架为约束，建立安全可信的AI中心。我国首部应用于人工智能的地方性法规于2021年6月在深圳市提交审议，旨在统一人工智能领域的分类统计标准。目前国家层面在人工智能领域的法律法规还是缺失的，健全法制体系，促进技术研究创新，明确人工智能研究及应用过程中的行为准则，增强统一协调性，使人工智能更好地服务于智慧城市建设。

4 结束语

随着智慧城市的发展，人工智能在政务、医疗、教育等领域有着广泛的应用，业务系统的数量和业务数据也在快速增加。在城市数据的全生命周期，隐藏着可能影响城市发展质量的安全性问题。因此，在智慧城市建设发展的过程中，提升数据安全保障，健全算法安全治理，丰富网络安全策略就成为城市发展决策者需要着重注意的问题。通过对人工智能在智慧城市领域多个场景的应用现状研究，分析归纳应用过程中容易出现的安全性问题，提出相应的应对策略，有助于发挥人工智能的正面效应，对智慧城市规划建设过程提供安全有效的帮助。

参考文献：

[1] 陈群，陈肇强，侯博议，等．人工智能风险分析技术研究进展[J]．大数据，2020，6(1)：47-59.

[2] 刘雅，刘昕彤，张春茜，等．人工智能在智慧城市发展中的应用[J]．集成电路应用，2021，38(8)：162-163.

作者简介：
姚树为，毕业于吉林大学，工程师，硕士，主要从事智慧城市领域相关技术创新与业务开发工作；陈龙，毕业于大连海事大学，工程师，主要从事Web前端开发、移动端开发、小程序开发等工作；沙默泉，毕业于北京师范大学，高级工程师，硕士，主要从事3S技术应用、智慧城市创新产品研发工作；王悦，毕业于北京邮电大学，工程师，硕士，主要从事智慧城市领域相关的产品研发工作。