

# 基于联邦学习的新冠肺炎疫情跨区域智慧防控技术

## ——以上海市实践为例

钱学胜<sup>1,2,3</sup>, 吴寰宇<sup>4\*</sup>, 陈诚<sup>5</sup>, 黄晓燕<sup>4</sup>, 童庆<sup>5</sup>, 戴伟辉<sup>1,2</sup>

1. 复旦大学智慧城市研究中心, 上海 200433
2. 复旦大学管理学院, 上海 200433
3. 澳门系统工程研究所, 中国澳门 999078
4. 上海市疾病预防控制中心, 上海 200336
5. 万达信息股份有限公司, 上海 201112

**摘要** 新冠肺炎病毒不断变异的毒株的系列新特征, 给传统的传染病防控方法与公共卫生防控体系带来了巨大的挑战。充分发挥数字技术在抗疫过程中的关键赋能价值, 全面构建精准的常态化监测预警及智慧防控体系, 系应对上述挑战的有效途径。通过对疫情智慧防控体系构成要素的剖析, 揭示联邦学习对建立跨区域、跨部门的智慧疫情防控体系的关键路径作用。并基于跨区域涉疫数据分类及防控工作要点, 研究设计了基于联邦学习的跨区域疫情智慧防控技术及其平台应用。该应用已在上海市以及长三角区域的抗疫实践中取得了显著成效。

**关键词** 疫情联防联控; 智慧防控; 智慧城市; 城市数字化转型

新型冠状病毒肺炎(新冠肺炎)疫情是新中国成立以来发生的传播速度最快、感染范围最广、防控难度最大的一次重大突发公共卫生事件<sup>[1]</sup>, 也是

近百年来对世界各国威胁最为严重的全球性重大公共卫生事件<sup>[2]</sup>。在抗疫实践中, 中国采取积极有力、综合性的非药物干预策略, 有效阻断新型冠状

收稿日期: 2021-09-02; 修回日期: 2021-09-30

基金项目: 国家重点研发计划项目(2018YFB2101100); 教育部哲学社会科学研究重大课题攻关项目(19JZD010); 国家自然科学基金面上项目(71971066); 教育部人文社会科学研究规划基金项目(18YJA630019); 上海市科技创新行动计划项目(20492420102); 上海市哲学社会科学规划课题(2019BGL031)

作者简介: 钱学胜, 高级研究员, 研究方向为数据智能与融合应用, 电子邮箱: qianxuesheng@fudan.edu.cn; 吴寰宇(通信作者), 主任医师, 研究方向为传染病防治与公共卫生管理, 电子邮箱: wuhuanYu@scdc.sh.cn

引用格式: 钱学胜, 吴寰宇, 陈诚, 等. 基于联邦学习的新冠肺炎疫情跨区域智慧防控技术——以上海市实践为例[J]. 科技导报, 2021, 39(24): 96-107; doi: 10.3981/j.issn.1000-7857.2021.24.011

病毒(新冠病毒)的传播和扩散,有力改变了病毒传播的进程,为全球疫情的防控争取了宝贵的时间<sup>[3]</sup>,取得了抗击新冠肺炎疫情斗争的重大战略成果<sup>[4]</sup>。

2020年10月,新冠病毒变异毒株德尔塔(Delta)在印度发现,在新冠病毒原本具有的高感染力、长潜伏期、广泛的易感人群和不受气候因子影响的特征基础上,又叠加了病毒载量更高<sup>[5]</sup>、感染力更强<sup>[6]</sup>、以及免疫逃逸等新特性<sup>[7]</sup>,被世界卫生组织(WHO)称为“传播最快和最有适应性的”病毒<sup>[6]</sup>。截至2021年8月,德尔塔毒株已波及全球超过130个国家和地区<sup>[8]</sup>,成为引发全球第二波新冠大流行的主要病毒株<sup>[9]</sup>,对世界各国本就不稳固的抗疫防线形成了新的重大挑战。2021年7月20日,德尔塔毒株由境外病例输入至中国,引发了“多点多源暴发”,截至8月14日,德尔塔疫情已波及多达18个省48个市,为已进入“疫情防控常态化阶段”的全国疫情防控和社会经济发展带来了新的冲击。

与此同时,新的更具威胁的新冠病毒变异毒株仍在不断生成。目前,世界卫生组织已列出了伊塔(Eta)、约塔(Iota)、卡帕(Kappa)、拉姆达(Lambda)等4种“关注变种”(VOI)<sup>[10]</sup>,其中拉姆达毒株已在拉美地区造成了新一轮的大流行,并且仍在向欧洲及北美地区持续扩散<sup>[11]</sup>。鉴于变异毒株所引发的巨大风险,WHO敦促各国对解除防疫限制保持高度谨慎<sup>[12]</sup>。从未来的发展趋势看,不断变异的病毒完全有可能以更强的感染力进一步突破现有疫苗的保护屏障,给现代社会生活、交通环境下的跨区域感染追踪和疫情防控带来更为复杂的困难,使世界各国现已疲惫不堪的防控体系面临更为严峻的挑战。

充分发挥数字技术在抗疫过程中的关键赋能价值,构建精准的常态化监测预警及智慧防控体系,系应对上述挑战的有效途径。在中国过去的疫情防控实践中,以“健康码”为代表的数字技术创新大放异彩,在中国的精准防控中发挥了重要作用,使得中国政府始终坚持统筹推进疫情防控和经济社会发展的战略<sup>[1]</sup>得以实现。随着中国的疫情防控模式从应急管控向常态化转变,如何以新技术手段对分布于不同条块的多源大数据进行高效整合利

用,同时依法有效保护个人隐私以及涉及到相关主体的敏感信息,是亟待思考与解决的问题,也是构建常态化智慧防控体系、保障社会经济发展和法治治理的关键基础。

上海作为中国最具代表性的超大城市,以及抗击新冠疫情最为成功的城市之一。研究其在应对新冠肺炎疫情过程中基于数字技术赋能的多学科参与、跨部门、跨区域智慧联防联控技术及模式,对于未来全国各省市有效抵御海外变种病毒侵袭、高效落实常态化防控措施、快速遏制疫情扩散无疑具有十分重要的现实价值和启示意义。

## 1 新冠肺炎疫情对疫情防控体系的挑战

重大公共卫生风险严重威胁人民生命安全和身体健康,严重影响社会经济系统正常运行。及时捕捉疫情的发生时间、发生地点和感染人员,进而有效进行干预,是公共卫生体系疫情管理领域的核心职能。自2003年SARS(严重急性呼吸综合征)疫情结束至今,中国政府已经形成了一套突发传染性疾病的医疗防控制度,建立了全球最大生物监测体系和覆盖全国的疫情防控网络<sup>[13]</sup>,为抗击新冠肺炎疫情奠定了坚实的物质基础。

然而,新冠病毒具有人际传播力强<sup>[14]</sup>、潜伏期及感染初期难以识别<sup>[15]</sup>、无症状感染者具有强传播力<sup>[16]</sup>等特性,为传统疫情防控体系带来了巨大挑战,致使包括中国在内的全球各国公共卫生体系在疫情早期的防控过程中集体陷入被动局面。尤其表现在疫情预警和疫情监测两个领域中。

### 1.1 预警领域中对数学建模方法的挑战

数学建模是流行病学预警领域中一个重要的既定工具<sup>[17]</sup>。在过去的30年中,数学模型方法和研究广泛应用在应对SARS、MERS(中东呼吸综合征)、Ebola(埃博拉出血热)等重大公共卫生事件中。在本次新冠肺炎疫情中,以SEIR(susceptible exposed infected recovered model)传播动力学模型及其扩展为代表的数学建模方法,被大量学者广泛应用于疫情传播计算<sup>[18-19]</sup>。

然而,数学建模方法的准确性受限于多个先决条件。Woolhouse等曾基于西非 Ebola 疫情的监测经验,系统讨论了数学建模方式的局限性,包括指数过程的流行病学模型对于参数过于敏感、人群活动随机性、难以估计控制措施引入的影响等<sup>[20]</sup>。由于本次新冠肺炎疫情传播的隐性化、区域化、广泛化特点,疫情传播受到物理空间、网络空间、社会心理空间等多空间耦合关联影响,使新冠肺炎疫情的传播形态呈现高度的复杂性特征,需采集各空间的多源大数据进行整合分析。但这些数据分散在不同区域、不同单位,涉及大量个人隐私和敏感信息,难以获取,更难以进行集中分析。数据的片面化、碎片化,导致了本次新冠肺炎疫情较难通过传统数据建模方法进行精准预测和分析。2020年1月英国学者 Read 等基于模型预测,到当年2月4日武汉市内将有19万人感染,至2月18日上海会有80万病例产生,是较具代表性的谬误案例<sup>[21]</sup>。通常更为精准的预测结果,需依赖回顾性的后验参数介入和事后模型拟合。在此基础上的实践应用中,当前国内疾控系统普遍采用的公共卫生统计分析方法多局限于基于本地数据在滞后一段时间后的疫情时空传播过程的切片式静态认知,而无法支持疫情传播的即时性研判<sup>[22]</sup>。由此可见,上述方法较难适应于新冠肺炎疫情新的传播特性下的预警研判要求,也难以满足日常防疫工作中对快速响应和及时决策的迫切需求。

### 1.2 监测领域中对传播链发现方法的挑战

监测是“出于公共卫生目的系统地不断收集、核对和分析数据,以供评估和采取必要的公共卫生应对措施”的活动<sup>[23]</sup>。西非地区 Ebola 疫情、中东地区 MERS 疫情等公共卫生危机的反思表明,尽早发现和快速应对,是避免传染病发展成公共危机的重要关键<sup>[20]</sup>。

从理论上讲,生物监测体系是公共卫生安全的重要基础和必要前提,但并非充分条件。庞大的传染病监测数据不会自动给出疫情态势风险,传统的公共卫生统计分析方法也不足以解析不同时空区域异质分化的疫情动态演化过程<sup>[22]</sup>。从抗疫实践上看,公共卫生系统为应对新冠病毒的独特传播特

性进行了针对性调整,在原先依赖各级医疗机构临床化验上报模式的基础上,实施了“关口前移”和“平行检测”等应对策略。然而从源头上,跨地域病患难以追溯、聚集性感染难以第一时间监测等关键手段的欠缺,仍是当前疫情监控中的重大隐患。有效监测手段的缺失,使得公卫系统仍然需要依靠例如公布感染者行程轨迹并向全社会公开征集潜在暴露者的方式,较为低效地进行感染排查。事实上,这种病例向上溯源及向下追踪监测方法领域的困境和局限,成为德尔塔毒株引起的疫情在南京、扬州、郑州等地触发较大范围本土传播的重要因素,在后续防疫实践中不能不引以为鉴。

## 2 重大科技热点事件的传播模式与特征

新冠肺炎疫情对人民生命健康以及经济社会发展带来了重大威胁,也为传统的传染病防控方法带来了系统性挑战。这激发了全球学者以一种有别于传统疾控体系思维的全新视角来重新审视突发公共卫生事件的系统性、复杂性等内在机理,并且思考应用新技术、新机制的应对策略。当前,“建立智慧化预警多点触发机制,健全多渠道监测预警机制,提高实时分析、集中研判的能力<sup>[24]</sup>”已成为全国公共卫生系统疫情监测预警和应急反应能力提升的工作重点。充分发挥数字技术在抗疫过程中的关键赋能价值也已成为全球范围的抗疫共识<sup>[25]</sup>。

构建精准的常态化监测预警及智慧防控体系,首先须厘清针对新冠肺炎疫情的智慧防控体系构成要素,并以此为基础确定智慧防控体系的技术路线和应用模式。根据理论与抗疫实践,多点触发、多渠道监测的疫情智慧防控体系,应包括数字技术深度参与、多学科交叉协作、多部门数据联合、跨区域联防联控等4个要素。

### 2.1 数字技术深度参与

人类战胜大灾大疫离不开科学发展和技术创新<sup>[26]</sup>。在新冠肺炎疫情全球范围传播期间,新数字技术正以更快速度、在更广范围融入社会生活。在疫情暴发后,数字技术也自然而然被广泛应用于全

全球范围的抗疫实践,在包括人口监测、病例识别、接触者追踪、公众沟通、干预措施评估在内的众多领域中,发挥着至关重要的作用。Budd等<sup>[25]</sup>基于对全球数字抗疫案例的全面总结,指出数字化是公共卫生的发展未来。当前,技术赋能创新正在形成一套以应用为基础的社会问题解决机制<sup>[27]</sup>,为更精细地疫情监测和防控决策提供可能。

在具体数字抗疫成功案例中,“健康码”的应用成效尤为显著。其通过有效归集手机GPS、蜂窝网络和WiFi接入点的位置数据,实现了实时性的人口流动跟踪,成为采取必要个别化公共卫生应对措施的重要依据和大规模公共卫生计算的数据来源,是中国疫情防控的重要技术创新和标志性品牌。截至2020年8月,继中国之后,全球47个国家先后推出了类似用于接触者追踪的疫情管理数字应用<sup>[28]</sup>。以健康码为代表的一批数字技术的深度参与,标志着数字技术作为常规公共卫生措施的有效补充手段,在长期抗击新冠肺炎疫情,以及应对未来公共卫生挑战中承担的重要作用。

## 2.2 多学科交叉协作

近年,基于“情景一应对”的应急管理新方法已逐步替代了传统“预测一应对”的应急方法<sup>[29-30]</sup>。科学思维底层的突破,促进了公共卫生应急管理领域的多学科交叉协作的发展。积极邀请传统临床医学和公共卫生领域之外的多学科参与,已成为此前应对Ebola等全球重大公共卫生危机中的抗疫共识<sup>[20]</sup>。当前,如何及时发现和阻断传播链,如何进行零号病人的及时溯源等防控关键技术仍是困扰全球公卫系统的共性难题。中国在进入疫情“常态化防控”阶段后,如何持续确保疫情防控和经济社会发展双重目标的有效达成,更是对疫情防控体系在内的政府综合治理能力提出了更高要求,亟需多学科的参与和投入。

在宏观的公共决策领域,多学科交叉探究能够为公共决策找到更科学、可靠的决策依据。例如,Chari等<sup>[31]</sup>基于优化后的经济-传染病传播模型(ECON-SIR)发现相较于“一刀切”式的隔离与检测政策,有针对性的精准检测和精准隔离将带来更高的社会福利,而更快锁定感染者(发展更有效的

感染者追踪技术)将是取得最佳社会福利的关键。在微观和中观的疫情传播态势预警领域,由于疫情传播受到物理空间、网络空间、社会心理空间的复杂交互耦合关联影响,呈现复杂系统特征,难以通过一般的数理模型或基于外部数据拟合的方式进行有效描述及预测。已有的研究及应用实践表明:采用人工社会建模仿真、多源大数据挖掘及知识图谱等多学科交叉协作的方法,可以更好地对真实环境下的疫情传播状况进行分析预测,并对各类干预措施的效果进行综合研判。在当前全球疫情仍在持续演变,外部环境更趋复杂严峻,国内经济恢复仍然不稳固、不均衡的局面下<sup>[32]</sup>,推进多学科的交叉协作,对于常态化疫情防控阶段坚持疫情防控和经济社会发展工作“两手抓、两手硬、两手赢”具有重要的现实意义。

## 2.3 多部门数据联合

当今社会运行与社会治理模式的复杂化,使统筹协调成为新时代应急管理的能力<sup>[33]</sup>。为有效抗击突如其来的新冠肺炎疫情,中国政府依据《中华人民共和国传染病防治法》和《国家突发公共卫生事件应急预案》,及时启动了纵向综合协调和横向联合响应的疫情应对体系。基于常态期各政府部门的角色职能,以及其能调动的各社会主体能力和资源,建立了包括综合协调指挥、医疗救护、卫生防疫、交通管制与运输、气象监测、复工与就业支持等在内的共17类应急响应功能<sup>[34]</sup>。

多元主体参与的协同效应离不开有效的信息流动。在尤为核心的疫情监控过程中,及时汇聚疾控、医疗、海关、疫苗、交通等多部门相关信息是有效识别未发现病例、快速完成接触者追踪和隔离的必要条件。所涉任一主体数据的疏漏或滞后,都可能导致整个防控体系的“破口”,从而造成巨大的社会经济代价。为应对重大突发的新冠肺炎疫情,当前各部门已通过临时调整应急响应的管理结构与功能以回应抗疫的现实需要。但从机制上看,政府多部门间协同机制尚处于实践先导、理论与制度滞后的摸索阶段。从技术手段和数据基础上看,各领域部门之间数据也存在采集重复、权限不明、安全及隐私保护不力等亟待解决的问题<sup>[35]</sup>。当前持续

而严峻的抗疫形势压力,成为各级政府努力探寻基于创新技术手段的全新数据联合机制以满足精细化疫情防控需要的重要动力。

#### 2.4 跨区域联防联控

在现代社会生活与交通环境下,跨区域的经济社会联系和人员流动日趋频繁。以重大传染病为代表的突发公共卫生事件,往往容易因人流、物流快速蔓延,形成跨区域甚至跨国家的快速传播扩散现象。针对此类突发公共卫生事件,区域各自为战,将极大影响应急管理的总体效能。

在抗疫背景下,国内学者分别从不同角度对跨区域、跨部门联防联控的协同理论、要素、机制进行了阐述<sup>[34,36]</sup>。但需要看到的是,在当前国内抗疫实践中,跨省、跨地区的联防联控仍然缺乏有效的基于底层打通的横向协作机制。从疫情开始之初“同省不同码”以及“省内码不互认”,到至今仍见于网络的相同响应等级下省内各地不同的防疫管制策略,都是其浅表现象。在更深层次上,构筑于双方函件往来的信息交互与汇聚模式之上的疫情研判预警和疫情处置决策,势必因数据及时性和完整性的缺失而影响整体防控成效,为以新冠肺炎疫情为代表的突发急性传染病疫情的防控工作埋下重大隐患,且这种隐患和缺陷难以依靠上级单位协调指令进行弥补。从造成原因来说,数据的隐私使用、属地化管理要求、以及跨地域数据的信任及权限问题都是当前较为核心的客观限制。在相关立法短期内无法快速跟进完善的背景下,实现跨区域数据层面的横向互通,进而进行联合监测预警的实践案例在国内仍十分鲜见。在经历德尔塔病毒多地本土传播的侵袭之后,通过智慧化疫情监测预警,实现跨区域涉疫数据的综合分析,显示出十分迫切的现实需要。另一方面,在国家加强区域高质量一体化发展的大趋势下,基于疫情联防联控的“强需求”,探索数字赋能跨区域事件态势集中研判的全新机制与创新方案也具有十分重大的战略意义。

### 3 隐私计算与联邦学习概述

如何在保证数据安全与隐私保护的前提下,尽

地满足抗疫工作中迫切的数据协作需求,成为建立智慧化抗疫系统亟待解决的核心难点。因此,可实现数据不对外泄露前提下分析应用的隐私计算(privacy preserving computing)成为破题关键。

隐私计算发源于隐私增强技术(privacy enhancing technologies, PETs),是一项在网络传输中隐藏个人身份及可识别个人信息的技术<sup>[37]</sup>。当前,隐私计算主要分为基于密码学技术的多方安全计算(secure multi-party computation, MPC)、基于可信硬件的可信执行环境(trusted execution environment, TEE)、以及联邦学习(federated learning, FL) 3个主要分支<sup>[38]</sup>。相较于较低算法适用性和较高实施门槛的前两者,具有更高跨组织通用性、以及扁平化和去中心化特点的联邦学习,能更贴近须多部门联合、跨区域协同的新冠肺炎疫情防控场景的应用需要。且因其能支持更多复杂算法的适用性空间,也能满足应对新冠肺炎疫情监测预警的算法复杂性需求。

联邦学习作为一种加密分布式机器学习技术,能使各参与的数据拥有方在不转移各自数据前提下,完成数据使用和机器学习建模,以满足用户隐私保护、数据安全和政策法规的要求。其基本流程如下:(1) 每个节点将通用的全局模型(如疫情监测预警模型)下载到本地,并基于节点本地数据对全局模型进行训练;(2) 将计算后的损失值(LOSS)、梯度信息(GD)等结果加密上载到系统端;(3) 在系统端实现上载模型信息的联邦平均(FedAvg)并更新全局模型;(4) 节点将更新后的全局模型取回并进行下一轮训练,直到模型达到期望性能或满足终止条件(图1)。根据各参与数据方不同的特征和样本分布情况,联邦学习还具有横向联邦学习(特征对齐模式)、纵向联邦学习(样本对齐模式)、联邦迁移学习3种不同计算模式。

由于联邦学习通过虚拟模型有效解决了不同数据拥有方在不交换数据情况下的协作问题,因此自2016年谷歌首次提出<sup>[39]</sup>后的短短数年,即在移动互联网、金融、医疗、工业工程等行业催生了大量应用案例。尤其在数据隐私高度敏感的相关领域以及对隐私立法要求较高的国家地区中显示出巨大

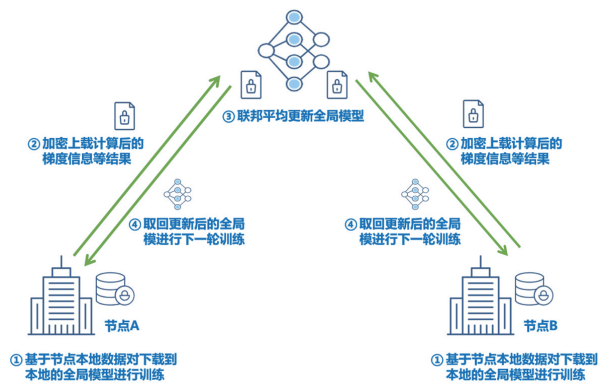


图1 联邦学习基本流程

潜力。Li等<sup>[40]</sup>详细总结了近年联邦学习在各个领域中的主要研究与应用。然而,在重大传染病预警监测、联防联控方面,目前全球尚未形成基于联邦学习的标志性案例。将联邦学习引入新冠肺炎疫情智慧防控体系,通过技术赋能和模式创新突破数据共享瓶颈,从而提高重大传染病跨区域的集中研判和精细化防控能力,是本文的重要贡献之一。

## 4 基于联邦学习的跨区域疫情智慧防控技术的上海实践

联邦学习能够基于跨区域疫情防控的工作特点,完整覆盖前述构建疫情智慧防控体系的4项要素,成为建立跨区域、跨部门的智慧化疫情监测预警机制的关键路径。在上海的抗疫实践中,基于联邦学习的跨区域联防联控平台的应用,为构建“主动发现、快速处置、精准有序”<sup>[41]</sup>为标志的疫情防控“上海模式”,提供了关键的数字支撑。

### 4.1 上海抗击新冠肺炎疫情的总体背景

本次新冠肺炎疫情是对国家治理体系和治理能力的一次大考<sup>[42]</sup>。上海拥有近2500万常住人口,是中国最具标志性的超大城市。研究显示,对于个体城市而言,更大的城市面积<sup>[43]</sup>和更发达的商业活动<sup>[44]</sup>都将导致更高的疫情扩散风险。在疫情暴发伊始,曾有海外学者预测,至2020年2月上海将超越武汉成为中国新冠肺炎病毒感染最多的城市<sup>[21]</sup>。

另一方面,上海所处的长三角区域也是中国经济发展最活跃、城镇集聚程度最高、人口和交通网

络最密集的地区。上海作为长三角区域的龙头和区域最重要的国际枢纽,还承载着区域内最为密集频繁的人流、物流往来。在前期疫情防控中,上海市内本土外省密接的近三成,以及收到密接协查函的约半数均来自于江浙皖三省,足以可见进入常态化防疫阶段后,区域活跃的社会经济活动对于疫情防控带来的风险与挑战,凸显出长三角区域公共卫生领域联防联控的迫切需要。人口基数大、国内外商贸往来频繁、疫情传播多样、社会形态复杂多元等特征,都使上海疫情的防控工作面临着举世公认的难度和无比巨大的压力。

### 4.2 跨区域不可共享类数据应用要求

在上海的疫情防控中,跨区域涉疫数据的综合利用是建立精准防控体系的关键。跨区域涉疫数据具有多样权属、多样权限、多样结构等特性。从信息共享角度区分,可分为可共享与不可共享2个类别。(1) 可共享类:可共享的涉疫信息包含地区确诊人数、地区疑似病例数、地区确诊人员姓名、暴露场所(公共交通、室内场所)等信息,可公开、可共享。(2) 不可共享类:不可共享涉疫信息主要包含了确诊人员、疑似、密接人员个人隐私数据,例如人员的身份证号、手机号、核酸检测报告、CT值、家庭住址、详细行为轨迹、流调信息等。

相较于可公开取得的可共享类信息,受限的不可共享类数据更为完整、精确、具有更高的数据分析价值,对于本地疫情态势的精准研判和潜在暴露者的主动发现起着至关重要的作用。在当前涉疫数据的管理规范中,不可共享类数据应用须满足以下2项基础要求。(1) 确保隐私数据的可用不可见。跨区域不可共享类数据的隐私保护是实现区域智慧化联防联控的先决条件和模式基础。(2) 确保数据高实时动态可用。常态化防控阶段,地区间人流物流往来频繁,疫情发展瞬息万变,须抢抓第一时间进行疫情的综合研判,为及时采取有效管控措施提供决策基础。

### 4.3 跨区域疫情智慧防控的联邦学习平台架构

上海市疾控中心以长三角科技联合攻关领域项目“重大突发性传染疾病的联防联控技术的开发及示范应用”开发和应用为支点,自2014年起便建

立了与苏浙皖三省疾控中心基于流感病毒的联防联控工作机制,研究跨区域不可共享类数据的综合利用技术,并以此为基础开展了目标可复制、可推广的跨区域重大传染病防控模式的探索。2020年,为应对来势汹汹的新冠肺炎疫情,上海市疾控中心专题开展“长三角区域重大公共卫生风险联合研判、预警和处置体系研究”,同时针对新冠肺炎疫情的精准化、智能化防控需要,研究设计了一个适用于跨区域疫情联防联控特性和满足于跨区域(兼顾跨部门)涉疫数据应用要求的疫情智慧防控的联邦学习平台(图2),并正式投入使用。基于联邦学习技术,通过分布式节点部署本地模型,确保“数据不出库、模型多跑路”,从而有效解决数据隐私和数据关联分析之间的矛盾;同时,采用高效的学习模型和算法实现了较高的计算性能,满足了常态化疫情精准防控与综合风险研判的实战需要。

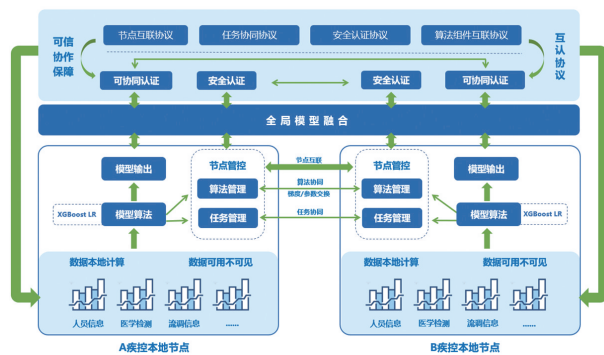


图2 面向跨区域疫情智慧防控的联邦学习平台架构

该平台技术架构以跨地域(跨部门)涉疫数据不可直接互联共享为前提,通过建立各参与单位多节点之间的可信、安全联合机制,形成包含安全认证协议、节点互联协议、算法组件互联协议、联合建模协同协议等协议要素为一体的多节点协同和互认协议体系,作为整个平台架构的可信协作保障,确保在节点并行计算过程中各参与方数据、模型、结果、任务的隐私安全。基于可信协作保障机制,算法将被下放到各地疾控中心的本地节点,实现数据在各地疾控内部可信环境下的本地计算。在这一过程中,各节点对算法和任务均进行独立管理,并基于可信协作系统进行任务协同和算法协同。

通过不断的模型本地计算和在线梯度与参数交换,驱动模型不断趋近算法的预期目标。其中,模型算法在此选用联邦学习系统中集成程度较好的XGBoost(极端梯度提升,eXtreme Gradient Boosting)和LR(逻辑回归,Logistic regression)。XGBoost源于梯度提升框架,通过对稀疏数据的有效处理以及并行计算等优化手段,实现了训练速度更快、更高效的梯度提升训练。LR是基于Logistic函数的常用二分类算法,具有实现简单、计算量小、速度快的特点。两者结合能够满足日常跨区域疫情态势综合研判任务中对于系统算法的及时性要求。当模型收敛达到预期目标时训练完成。此时每个节点保存本节点的子模型相关参数及数据信息。新的业务数据产生时,新样本将被下放个各地疾控的本地节点,各节点根据先前保存的模型信息对新的数据进行独立预测,预测结果基于预先制定的联邦平均算法进行融合,以形成最终的预测结果。

#### 4.4 跨区域疫情智慧防控分析算法框架

基于联邦学习技术架构,构建跨区域疫情智慧防控分析算法框架。按联邦学习平台架构,分析算法框架主要分为数据准备、数据清洗与抽取、样本数据预处理、联合建模、联合预测5个部分(图3)。



图3 跨区域疫情智慧防控分析算法框架

1) 数据准备:数据准备作为算法运行的启动环节,在进行多方数据交互之前,将可交互数据准备好并完成相关算法定义,如算法及协议的选择、多方交互变量的定义等。

2) 数据抽取与清洗:基于疾控专家对疫情防控各环节的经验判断及实际数据完整情况,从各地疾控中心抽取本地样本数据信息,如病例基本信息、发病治疗与转诊情况、流行病学史、样本采集与实验室检测等数据,构成建模特征空间 $X$ 。其主要

所选变量包括： $x_0$ (病例ID)、 $x_1$ (病例状态)、 $x_2$ (是否为医疗机构工作人员)、 $x_3$ (性别)、 $x_4$ (年龄)、 $x_5$ (首次发病时间)、 $x_6$ (最高体温)、 $x_7$ (白细胞(109/L))、 $x_8$ (淋巴细胞)、 $x_9$ (肺部影像学)、 $x_{10}$ (既往病史)、 $x_{11}$ (血(份))、 $x_{12}$ (咽拭子(份))、 $x_{13}$ (支气管冲洗液(份))、 $x_{14}$ (与其他疑似/确诊病例有无关系)等。完成数据抽取之后,基于业务规则及数据特征对数据进行清洗,如剔除重复患者、统一命名与日期格式等,以完成样本对齐。

3) 样本数据预处理:完成清洗后,数据已达成算法要求的多方交互基础条件。其后,对所选数据进行特征工程等精细化处理方法完成样本初步分析,如联邦相关系数、联邦IV值等,从而了解各节点各不同变量与预测变量之间的相关性,以及对评估不同变量对结果的贡献度。

4) 联合建模:经过数据预处理的特征工程之后,对数据进行建模训练。由于跨区域的疫情防控场景,符合特征重合较多、ID重合较少的情形。因此,在本文所述疫情分析算法框架中选择横向联邦学习的模式进行数据融合推理。模型算法采用XGboost算法进行关于疫情分析的多变量时间序列预测,实现对某一地区的疫情进行预测研判分析。其XGboost的目标函数如下:

$$Obj = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{i=1}^l \Omega(f_i) \quad (1)$$

式中, $\hat{y}_i$ 为结果预测值; $y_i$ 为真实值; $\sum_{i=1}^n l(y_i, \hat{y}_i)$ 代表 $n$ 个样本的损失函数; $\sum_{i=1}^l \Omega(f_i)$ 为正则化项,是对于全部 $l$ 棵树的复杂度的求和。

完成模型训练之后,本地区的疫情风险值 $h_y$ 函数为

$$h_y = \frac{\sum_{i=1}^n \hat{y}_i}{n} \quad (2)$$

式中, $h_y$ 为子模型最终的预测结果,表示本地区的风险值; $\hat{y}_i$ 为子模型预测的第 $i$ 天的确诊人数; $n$ 为预测天数的滑动窗口; $P$ 为本地区的常住人口数。

上述过程各子模型生成的模型损失值

(LOSS)、梯度(GD)均通过同态加密算法传输至模型发起方节点,而后准备进行下一轮学习过程。

5) 模型预测:最终模型预测结果为各地疾控中心子模型的预测结果在本地节点取滑动平均数后,汇总到联合模型处再通过联邦平均后取得。其数值范围为0~1之间,越接近1代表风险系数越大,说明该区域需要重点关注。而后将加密后加法计算的结果进行解密,得到最终的联合风险评估模型的预测数值:

$$h = Decry \left( \sum_{y=1}^m Encry(h_y) \right) \quad (3)$$

式中, $m$ 为参与计算的地区数。

在上海的实际抗疫过程中,以上述算法框架与计算过程的具体阐释如下。

1) 在平台架构和数据准备阶段,于上海疾控和参与该联邦学习平台的各省市疾控均各部署一个节点,为本地的模型训练进行准备。

2) 在数据抽取与清洗阶段,平台中各节点根据预先约定的数据规范和分析目标,在各自系统内部进行数据抽取和清洗,确保节点间的样本数据字段及格式完全对齐。

3) 在样本处理和联合建模阶段,通过模型训练方——上海疾控发起训练任务,将经过预处理的算法初始模型下发给各个节点,各个节点根据算法在本地训练子模型;在训练过程中,各子节点采用同样的加密算法,将计算后的迭代梯度发送至上海疾控,上海疾控进行联邦平均后,再将更新的模型下发给各个子节点进行进一步的训练,当模型收敛到预期目标时完成整个训练过程。

4) 在最终的模型预测阶段,达到预期目标的模型将进行保存。在实际防控应用中,上海疾控会每天将前一天的样本数据处理好之后,输入模型进行后预测。

通过上述方式,实现基于长三角整体疫情发展态势对上海市疫情风险的精准研判,为尽早采取必要防控措施,避免疫情大范围扩散奠定决策基础。

#### 4.5 跨区域疫情智慧防控平台实践成效

上海市疾控中心基于前期“长三角区域重大公

共卫生风险联合研判、预警和处置体系研究”,通过联邦学习技术汇聚长三角区域疾控、海关、医疗、交通等多个地区、多个相关部门的多源异构数据,同时通过融合了公卫科学、管理科学、数据科学等多学科计算模型优势。基于融合长三角整体疫情传播态势和疫情传播多空间耦合关联分析,实现了全面精准的常态化疫情研判和监测,充分发挥了数字技术在抗疫过程中的关键赋能价值,有力支撑了上海市精准化疫情防控各项工作的开展。

自上海市疾控为中心的跨区域联防联控平台上线以来,截至2021年8月31日,针对新冠防疫,已覆盖全部长三角确诊、疑似、死亡、无症状病例在内的“突发本土疫情应急响应”“病例报告与流调”“病例基本信息”“发病、诊疗与转归情况”“相关接触人员追踪情况”“防疫相关政策应对”等六大类数据,覆盖近90个关键信息字段,纳入涉疫样本总数据超过5万条、境外输入病例信息超过8000例。其中,疑似病例数、密接信息、流调信息、物品信息、传染路径、暴露场所等不可共享类数据占比约为25%。

基于联邦学习平台的隐私数据保障,使长三角区域开展跨区域的风险综合评估得以实现。在上海的实际应用中,基于此智慧防控技术能有效将传统需3~5 d才能完成的疫情风险研判,缩短至2 d以内,为上海市从严从实落实外防输入、切实提升疫情防范和早期发现能力起到了重要作用。助力上海在应对本轮突如其来的德尔塔变种病毒入侵过程中,在隔离仅止于小区粒度、每轮排查核酸不多于常住人口数3‰的情况下,即实现了对反扑的德尔塔疫情的有效遏制。保持了在“尽量不要因为防疫影响民众生活”的原则下,始终高效地“陶瓷店里抓老鼠”的高水平防疫实践。

同时,在长三角多个主要城市的疫情防控实战中,跨区域智慧防控技术也为各地公卫决策部门提高疫情传播链分析的精度和速度、缩短零号病例发现时间、科学研判不同区域的疫情风险和变化趋势提供了有力的技术支撑。在上海以及长三角区域内多起跨区域感染者的溯源和传播链追踪案例中提供了重要的决策参考。

## 5 上海疫情智慧防控技术的应用启示

自2020年初新冠肺炎疫情暴发以来,上海克服了自身作为超大城市及长三角区域中心在应对突发重大公共卫生事件中的种种不利因素。在至2021年8月近20个月的实际抗疫历程中,不管在疫情防控 and 经济社会发展领域,还是在防控模式与机制优化创新以及城市人文关怀方面,都交出了全国领先的优异答卷。诸如全国首创的确诊病例流调信息通报不披露个人信息及详细行动轨迹等<sup>[45]</sup>多项措施,均赢得了全国舆论的广泛赞誉<sup>[46]</sup>。

2021年9月17—18日,国务院应对新型冠状病毒肺炎疫情联防联控机制综合组在上海举办全国新冠肺炎疫情防控经验研讨会<sup>[47]</sup>,会议高度肯定了上海市在抗击新冠肺炎疫情中的各项经验,要求各地“深刻认识学习上海等地经验的重要现实意义”“准确把握学习上海等地经验的方向和内容”“以实事求是的态度学习借鉴上海等地经验”,并组织来自全国有关省份、省会城市、计划单列市的与会代表现场观摩了以智慧联防联控技术为重要特色的上海疫情防控指挥体系。

从上海的成功经验中不难发现:综合研判、主动发现、精准流调、快速应对对于城市疫情防控的关键作用。其背后是基于数字技术支撑下的灵敏高效、科学精准、联防联控的智慧化公共卫生防控体系对于持续高效抵御来自于境外、本土双向疫情波及的关键助力。由此可见,大力发展数字技术赋能下的主动化、快速化、精准化的防控体制和防控模式,是以最小成本获得最大防控成效的关键方式,也是平衡城市疫情防控和经济社会发展最为经济、有效的治理手段。

进一步剖析其技术应用的外部性价值,上海构建基于联邦学习技术的智慧疫情防控体系,带来了兼顾数据联合分析和数据隐私防护的全新思路,打破了当前跨条块数据无法有效共享的数字应用桎梏。通过跨区域数据联邦学习、协同运算,显著提升了公共决策的研判精度、优化了决策措施效力。更进一步,其提供了一种以技术创新应用解决当前治理领域区域协同难点,同时推动区域数据业务

深度应用,进而形成区域高质量一体化协同的珍贵案例。

## 6 结论

当前,全球疫情起伏反复,面对不断变异的病毒和海外持续延烧的疫情,使我们不得不加快公共卫生系统应急处置体系能力提升步伐,加速完善传染病疫情和突发公共卫生事件监测系统。在此过程中,数字技术的深度参与、多学科的交叉协作、多部门的数据联合、跨区域的联防联控对成功应对重大突发公共卫生事件具有关键作用。在上海的成功抗疫实践中,通过基于联邦学习技术的跨区域智慧防控技术及其数字化平台的实际应用,有机统整了上述4项要素,使之成为抗疫“上海模式”背后的模式支撑和突出亮点。该技术的研发和应用为上海市及长三角区域主要城市的抗疫实践提供了重要助力,对于未来全国各省市有效落实常态化防控措施、快速遏制疫情扩散具有至关重要的现实价值和启示意义。作为全国首个跨多个省级区域的数字化联防联控平台,也具有极高的模式研究和复制推广价值,为中国乃至世界区域疫情的集中研判和精细化防控提供了对标模板。

眼下数字化进程已成为城市现代化的必然选择,“有无完备的信息化技术与应用体系”已是评估一个国家或一座城市整体竞争力的重要标准之一<sup>[48]</sup>。遭遇新冠肺炎疫情冲击之后,使全球各国更加明确了稳健、高质量的健康公卫体系是城市发展与城市繁荣的关键支撑要素。在当前上海市全面推进城市数字化转型的大背景下<sup>[49]</sup>,数字技术对于城市重大突发公共卫生事件的新应用和新价值还在被不断探索和发掘。目前上海已正式提出:至2025年重大疫情和突发公共卫生事件的应对能力达到国际一流水准<sup>[50]</sup>。基于疫情防控的成功经验,上海将力争率先形成国内首个区域性公共卫生风险研判预警和处置信息平台的示范模板,对全国卫生应急信息平台的总体推进起到重要的示范作用,并持续激励上海向全球公共卫生最安全城市之一的目标加速前进。

## 参考文献 (References)

- [1] 习近平出席统筹推进新冠肺炎疫情防控和经济社会发展工作部署会议并发表重要讲话[EB/OL]. [2021-08-17]. [http://www.gov.cn/xinwen/2020-02/23/content\\_5482453.htm](http://www.gov.cn/xinwen/2020-02/23/content_5482453.htm).
- [2] World Health Organization. WHO Director-General's opening remarks at the media briefing on COVID-19-27 July 2020[EB/OL]. [2021-08-17]. <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19--27-july-2020>.
- [3] Lai S, Ruktanonchai N W, Zhou L, et al. Effect of non-pharmaceutical interventions to contain COVID-19 in China[J]. *Nature*, 2020, 585(7825): 410-413.
- [4] 全国抗击新冠肺炎疫情表彰大会在京隆重举行 习近平向国家勋章和国家荣誉称号获得者颁授勋章奖章并发表重要讲话[EB/OL]. [2021-08-17]. [http://www.gov.cn/xinwen/2020-09/08/content\\_5541722.htm](http://www.gov.cn/xinwen/2020-09/08/content_5541722.htm).
- [5] Li B, Deng A, Li K, et al. Viral infection and Transmission in a large well-traced outbreak caused by the Delta SARS-CoV-2 variant[J/OL]. *MedRxiv*, 2021, <https://doi.org/10.1101/2021.07.07.21260122>.
- [6] World Health Organization. WHO press conference on coronavirus disease (COVID-19)-21 June 2021[EB/OL]. [2021-08-17]. [https://www.who.int/multi-media/details/who-press-conference-on-coronavirus-disease-\(covid-19\)--21-june-2021](https://www.who.int/multi-media/details/who-press-conference-on-coronavirus-disease-(covid-19)--21-june-2021).
- [7] Liu C, Ginn H M, Dejnirattisai W, et al. Reduced neutralization of SARS-CoV-2 B. 1.617 by vaccine and convalescent serum[J]. *Cell*, 2021, 184: 4220-4236.
- [8] World Health Organization. WHO press conference on coronavirus disease (COVID-19)-30 July 2021[EB/OL]. [2021-08-17]. [https://www.who.int/multi-media/details/who-press-conference-on-coronavirus-disease-\(covid-19\)--30-july-2021](https://www.who.int/multi-media/details/who-press-conference-on-coronavirus-disease-(covid-19)--30-july-2021).
- [9] World Health Organization. WHO press conference on coronavirus disease (COVID-19)-18 June 2021[EB/OL]. [2021-08-17]. [https://www.who.int/multi-media/details/who-press-conference-on-coronavirus-disease-\(covid-19\)--18-june-2021](https://www.who.int/multi-media/details/who-press-conference-on-coronavirus-disease-(covid-19)--18-june-2021).
- [10] World Health Organization. Tracking SARS-CoV-2 variants[EB/OL]. [2021-08-17]. <https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/>.
- [11] GISAID. hCov-19 Variants-VOI Lambda GR/452Q.V1 (C. 37) first detected in Peru[DB/OL]. [2021-08-17].

- <https://www.gisaid.org/hcov19-variants/>.
- [12] World Health Organization. WHO press conference on coronavirus disease (COVID-19) - 7 July 2021[EB/OL]. [2021-08-17]. [https://www.who.int/multi-media/details/who-press-conference-on-coronavirus-disease-\(covid-19\)-7-july-2021](https://www.who.int/multi-media/details/who-press-conference-on-coronavirus-disease-(covid-19)-7-july-2021).
- [13] 董言, 姚华. 我国传染病网络直报的现状与发展[J]. 疾病预防控制中心通报, 2012, 27(1): 92-94.
- [14] Riou J, Althaus C L. Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020[J]. Eurosurveillance, 2020, 25(4): 2000058.
- [15] Li Q, Guan X, Wu P, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus - infected pneumonia[J]. New England Journal of Medicine, 2020, 382(13): 1199-1207.
- [16] Rasmussen A L, Popescu S V. SARS-CoV-2 transmission without symptoms[J]. Science, 2021, 371(6535): 1206-1207.
- [17] Anderson R M, May R M. Infectious diseases of humans: Dynamics and control[M]. New York: Oxford University Press, 1991.
- [18] 游光荣, 游翰霖, 赵得智, 等. 新冠肺炎疫情传播模型及防控干预措施的因果分析评估[J]. 科技导报, 2020, 38(6): 90-96.
- [19] 王健伟, 崔秩玮, 潘潇雄, 等. 基于广义SEIR模型的新冠肺炎传播机制及干预效果仿真[J]. 科技导报, 2020, 38(22): 130-138.
- [20] Woolhouse M E, Rambaut A, Kellam P. Lessons from Ebola: Improving infectious disease surveillance to inform outbreak management[J]. Science Translational Medicine, 2015, 7(307): 307.
- [21] Read J M, Bridgen J R E, Cummings D A T, et al. Novel coronavirus 2019-nCoV: Early estimation of epidemiological parameters and epidemic predictions[J]. Philosophical Transactions of The Royal Society B Biological Sciences, 2021, 376(1829): 20200265.
- [22] 曹志冬, 曾大军, 张清鹏, 等. 新冠肺炎疫情的复杂性特征与分析研判[J]. 中国科学基金, 2020, 34(6): 675-682.
- [23] World Health Organization. International health regulations (2005) 2nd ed[M]. Geneva: WHO Press, 2008.
- [24] 习近平. 构建起强大的公共卫生体系 为维护人民健康提供有力保障[J]. 求是, 2020(18): 4-11.
- [25] Budd J, Miller B S, Manning E M, et al. Digital technologies in the public-health response to COVID-19[J]. Nature Medicine, 2020, 26(8): 1183-1192.
- [26] 习近平总书记在专家学者座谈会上的重要讲话指明科研攻坚方向[EB/OL]. [2021-08-17]. [http://www.xinhuanet.com/politics/leaders/2020-06/04/c\\_1126074999.htm](http://www.xinhuanet.com/politics/leaders/2020-06/04/c_1126074999.htm).
- [27] 关婷, 薛澜, 赵静. 技术赋能的治理创新: 基于中国环境领域的实践案例[J]. 中国行政管理, 2019(4): 58-65.
- [28] Liang F. COVID-19 and health code: How digital platforms tackle the pandemic in China[J]. SAGE Public Health Emergency Collection, 2020, 6(3): 1-4.
- [29] 张辉, 刘奕. 基于“情景-应对”的国家应急平台体系基础科学问题与集成平台[J]. 系统工程理论与实践, 2012, 32(5): 947-953.
- [30] Xie T, Ni M, Zhang Z, et al. Parallel simulation decision-making method for a response to unconventional public health emergencies based on the scenario-response paradigm and discrete event system theory[J]. Disaster Medicine and Public Health Preparedness, 2019, 13(5-6): 1017-1027.
- [31] Chari V V, Kirpalani R, Phelan C. The Hammer and the Scalpel: On the Economics of Indiscriminate versus Targeted Isolation Policies during Pandemics[J]. Review of Economic Dynamics, 2020, doi: 10.1016/j.red.2020.11.004.
- [32] 习近平主持中共中央政治局会议 分析研究当前经济形势和经济工作[EB/OL]. [2021-08-17]. [http://www.gov.cn/xinwen/2021-07/30/content\\_5628481.htm](http://www.gov.cn/xinwen/2021-07/30/content_5628481.htm).
- [33] 王宏伟. 统筹协调——新时代应急管理的能力[J]. 中国安全生产, 2019, 14(2): 22-27.
- [34] 陶鹏, 张家俊. 新冠肺炎疫情联防联控机制的功能图景与预案优化[J]. 河海大学学报(哲学社会科学版), 2021, 23(3): 31-37.
- [35] 张成岗, 李佩. 科技支撑社会治理现代化: 内涵、挑战及机遇[J]. 科技导报, 2020, 38(14): 134-141.
- [36] 姜长云, 姜惠宸. 新冠肺炎疫情防控对国家应急管理体系和能力的检视[J]. 管理世界, 2020, 36(8): 8-18, 31, 19.
- [37] Goldberg I, Wagner D, Brewer E. Privacy-enhancing technologies for the internet[C]//Proceedings IEEE COMPCON 97. Digest of Papers. San Jose: IEEE Computer Socical Press, 1997: 103-109.
- [38] 闫树, 吕艾临. 隐私计算发展综述[J]. 信息通信技术与政策, 2021(6): 3-11.
- [39] McMahan B, Moore E, Ramage D, et al. Communication-efficient learning of deep networks from decentralized data[C]//Proceedings of the 20th International Conference on Artificial Intelligence and Statistics. Massachusetts: MIT Press, 2017.
- [40] Li L, Fan Y, Tse M, et al. A review of applications in

- federated learning[J]. *Computers & Industrial Engineering*, 2020, 149(5):106854.
- [41] 上海市市长: 本轮疫情已得到控制[EB/OL]. [2021-08-17]. <http://sh.people.com.cn/n2/2021/0127/c134768-34551230.html>.
- [42] 习近平. 全面提高依法防控依法治理能力 健全国家公共卫生应急管理体系[J]. *求是*, 2020(5): 4-8.
- [43] 屠鸿薇, 钟若曦, 肖建鹏, 等. 广东省新型冠状病毒肺炎分区分级防控策略研究[J]. *中国公共卫生*, 2020, 36(4): 486-492.
- [44] 漆翠芳, 杨力仁, 杨子轩, 等. 影响新型冠状病毒肺炎省际传播与发展的因素: 基于30个省市的数据分析[J]. *西安交通大学学报(医学版)*, 2020, 41(5): 757-763.
- [45] 为何上海流调报告值得点赞 不让隐私成谈资[EB/OL]. [2021-08-17]. [https://m.gmw.cn/2021-01/23/content\\_1302065235.htm](https://m.gmw.cn/2021-01/23/content_1302065235.htm).
- [46] 软硬兼施! 上海又交出了一份“模范作业”[EB/OL]. [2021-08-18]. <http://jl.people.com.cn/n2/2021/0818/c349771-34872402.html>.
- [47] 国家卫健委疾病预防控制局. 全国新冠肺炎疫情防控经验研讨会在上海举办[EB/OL]. [2021-09-27]. <http://www.nhc.gov.cn/xcs/yqfkdt/202109/d3bc7575e0b24e599ca3d4c8d8c61d5a.shtml>.
- [48] 汪光焘, 李芬. 推动新型智慧城市建设——新冠肺炎疫情对城市发展的影响和思考[J]. *中国科学院院刊*, 2020, 35(8): 1024-1031.
- [49] 钱学胜, 凌鸿, 黄丽华. 城市数字化转型 打造具有世界影响力的国际数字之都[J]. *上海信息化*, 2021(1): 6-12.
- [50] 上海市卫生健康委员会. 中共上海市委、上海市人民政府关于完善重大疫情防控体制机制健全公共卫生应急管理体系的若干意见[EB/OL]. [2021-08-17]. <http://ws-jkw.sh.gov.cn/sh1/20200408/bdc1a3241d214d76b69ae11980a7e64f.html>.

## Cross-region smart COVID-19 pandemic management and control based on federated learning: With Shanghai as an example

QIAN Xuesheng<sup>1,2,3</sup>, WU Huanyu<sup>4\*</sup>, CHEN Cheng<sup>5</sup>, HUANG Xiaoyan<sup>4</sup>, TONG Qing<sup>5</sup>, DAI Weihui<sup>1,2</sup>

1. Smart City Research Center, Fudan University, Shanghai 200433, China
2. School of Management, Fudan University, Shanghai 200433, China
3. Macau Institute of System Engineering, Macau University of Science and Technology, Macau 999078, China
4. Shanghai Center for Disease Control and Prevention, Shanghai 200336, China
5. Wonders Information Co., Ltd., Shanghai 201112, China

**Abstract** SARS-CoV-2 and its variants, the viruses that cause the COVID-19 pandemic, have some new characteristics, such as the high transmissibility, the long incubation period, the sweeping susceptible population, and the high environmental endurance, so a key question of the pandemic management and control is monitoring the asymptomatic transmission in daily life and socioeconomic activities, especially, the wide-range of cross-region spread. These features pose a great challenge to the traditional pandemic management and control methods and the global public health surveillance and control system. An effective way to tackle this challenge is making full use of the digital technology in the pandemic management and control, and building an accurate regular epidemiological surveillance and smart pandemic management and control system. By analyzing the essential factors of a smart pandemic management and control system, the critical role of the federated learning in the practice of the cross-region and cross-department smart pandemic management and control is shown. According to the classifications of the cross-region pandemic-related data and the pandemic management and control requirements, the technology and the application of the cross-region smart pandemic management and control based on the federated learning are explored. This method is successfully applied in the COVID-19 pandemic management and control in Shanghai and Yangtze River Delta, providing a new pathway for the integrated decision-making and targeted pandemic management and control in China. This method is also instrumental for other countries in the pandemic management and control.

**Keywords** joint pandemic prevention and control; smart pandemic management and control; smart city; urban digital transformation ●



(责任编辑 徐丽娇)