

艾滋病传播

智能模拟与地理空间可视化

杨 昆 彭双云 朱彦辉 王加胜 许泉立 / 著

AIDS Spreading: Intelligent Simulation
and Geo-Spatial Visualization



科学出版社

艾滋病传播智能模拟 与地理空间可视化

杨 昆 彭双云 朱彦辉 王加胜 许泉立 著

科学出版社

北 京

内 容 简 介

艾滋病对人类的健康和社会发展构成了巨大威胁,是一个全球的公共卫生问题和社会问题。预防和控制艾滋病是一项刻不容缓且需要长期努力的艰巨任务。我国目前正处在防治艾滋病的关键时期,日益紧迫的艾滋病预防、控制、治疗和关怀任务,急需我们从科学的角度重新审视它。本书基于多智能体的复杂系统建模理论和技术,结合作者基于 GIS 技术开展流行病研究的长期积累,通过与昆明市疾病预防控制中心等相关单位的紧密合作,研究 HIV/AIDS 的传播机制,建立 HIV/AIDS 的多智能体模型;应用 GIS 技术在空间分析方面的优势,构建一个基于多智能体与 GIS 技术的艾滋病传播智能体模型,并对艾滋病时空传播趋势进行可视化的、动态的模拟和预测;最后,开发一个能够为艾滋病防控相关部门及单位提供信息查询统计、空间可视化、疫情辅助决策、疫情传播预测和预防控制等相关功能的艾滋病传播模拟与决策支持信息平台。

本书可供从事流行病防控、智能体模拟、GIS 等研究领域的科技人员、高等院校师生及管理者参考。

图书在版编目(CIP)数据

艾滋病传播智能模拟与地理空间可视化 / 杨昆等著. —北京:科学出版社, 2018.1

ISBN 978-7-03-056135-0

I. ①艾… II. ①杨… III. ①地理信息系统-应用-获得性免疫缺陷综合征-防治-研究 IV. ①R512.91-39

中国版本图书馆 CIP 数据核字 (2017) 第 317809 号

责任编辑: 张 展 唐 梅 / 责任校对: 韩雨舟

责任印制: 罗 科 / 封面设计: 墨创文化

科学出版社出版

北京东黄城根北街16号

邮政编码: 100717

<http://www.sciencep.com>

成都锦瑞印刷有限责任公司印刷

科学出版社发行 各地新华书店经销

*

2018年1月第一版 开本: 787×1092 1/16

2018年1月第一次印刷 印张: 10 1/4

字数: 250千字

定价: 108.00元

(如有印装质量问题, 我社负责调换)

前 言

艾滋病是一种严重危害人类健康、阻碍社会经济发展、传播速度快、目前尚无有效治愈办法、病死率极高的流行传染病。它在全世界的广泛流行已成为严重的公共卫生问题、社会问题,甚至是政治问题。据《中国艾滋病防治联合评估报告》的调查表明,艾滋病仍然处在全国低流行和局部地区及特定人群高流行并存的态势,特别在云南、新疆、四川和河南某些地区的部分人群中,该疾病已呈现较高的流行水平,且艾滋病正由高危人群向一般人群传播。同时,云南省作为艾滋病重灾区,疫情一直不容乐观,感染者数量逐年上升,现已呈现出快速度、多渠道、多层面流行的态势。据云南省卫生厅和有关部门的专门调查显示,云南省的艾滋病感染者总数实际已超过 8 万人,成为全国艾滋病疫情严重的地区之一,云南省的艾滋病防治工作迫在眉睫。

面对这样严峻的疫情形势,艾滋病防治在迫切需要全国和国际社会的持续关注和帮助的同时,也需要更为有效的艾滋病疫情防治手段,而结合“3S”技术的现代地理信息技术正是有效提高艾滋病防治能力的信息技术手段之一。本书的价值主要在于:①为艾滋病防治工作提供信息化的新思路、新技术和新手段;②探索地理空间信息技术在艾滋病防治中的应用,研究空间信息技术在艾滋病防治领域的基础理论和专业分析模型,推进我国空间信息技术应用于艾滋病防治研究的深度和广度;③为云南省乃至其他省份的艾滋病区域防治与监控工作提供科学的决策依据与信息技术支持手段,研究形成的空间型艾滋病监测与管理信息系统软件具有较为广阔的推广应用前景;④基于本书的研究结论和成果,开发可应用于省、市和县级疾病预防与控制中心(CDC)的通用地理信息平台软件;⑤推动以云南省为区位的、包括东南亚地区的艾滋病防治国际科技合作,提高国际区域防治的整体效果。

作者在总结前人研究成果的基础上,基于智能体技术,从微观的角度模拟艾滋病传播过程,通过图表、虚拟环境、文件输出等多种方式展现艾滋病宏观的、完整的传播过程;通过 ABM 技术与 GIS 技术的集成,把个体与现实环境相对应,探求艾滋病传播的时空规律,并对未来艾滋病的流行趋势进行预测;通过自主开发的艾滋病传播模拟与决策支持信息平台实现艾滋病空间数据、属性数据及动态模拟数据的一体化管理,实现艾滋病信息查询统计、空间可视化、疫情辅助决策、预防控制、艾滋病疫情传播模拟、艾滋病预警分析等功能,从而为公共卫生管理部门提供一个集信息管理、查询、分析和辅助决策多种功能为一体的信息化平台。

全书共有 7 章。第 1 章基于国内外相关研究文献,在对艾滋病及其病毒流行原因、流行态势、我国时空分布特点及防控策略分析的基础上,总结归纳艾滋病传播研究领域的发展历程;第 2 章从智能体模型的概念、智能体模型的特点、智能体模型设计中的关键内容、智能体模型的设计和实现以及智能体模型的验证方面介绍智能体模型的基本原

理和方法；第3章在归纳总结智能体模型与GIS基础技术、方法的基础上，提出本书所采用的模型集成方法，并通过实例证明该方法是合理可行的；第4章探索艾滋病的时空信息表达机理，研究艾滋病如何从现实世界映射到概念世界，并利用GIS技术实现对艾滋病时空信息机理的表达；第5章提出基于多Agent技术的艾滋病概念模型、智能体模型构建和模型模拟及验证的思路，并利用Repast J 3.1仿真平台和Java语言编程实现了该模型，完成多Agent模型与GIS技术的集成；第6章阐述GIS功能在艾滋病防控中的应用，专题查询主要是用于对HIV/AIDS的高危人群聚居点、重灾区、监测哨点和医疗设施的查询，疫情分析则主要是利用GIS的空间查询和空间分析功能，基于相关地区历史病例数据对疫情进行空间预警分析和预案分析，从而为艾滋病的防控提供决策上的支持；第7章就如何设计、制作艾滋病疫情空间数据库、模拟与决策支持系统的总体结构、模型与GIS集成模型以及系统的各个功能模块的作用和操作进行详细介绍。

第1章由杨昆、许泉立撰写；第2章由杨昆、王加胜撰写；第3章由杨昆、彭双云撰写；第4章由杨昆、陈克伟撰写；第5章由杨昆、王加胜、熊建红撰写；第6章由杨昆、彭双云、许泉立、朱彦辉、王加胜、李允良撰写；第7章由杨昆、彭双云、许泉立、朱彦辉、王加胜、王少勇撰写。全书由杨昆总体设计并拟定章节内容，由杨昆、彭双云、董雯、李鸣蝉统稿。

本书研究成果受作者承担的国家863计划项目“艾滋病传播的智能体模型及其与GIS的集成技术研究”(2007AA12Z231)、国家自然科学基金项目“云南省艾滋病传播时空分析模型及其地理信息技术决策支持平台研究”(30560135)等研究项目的支持与资助。

由于作者水平有限，在撰写过程中难免有不当之处，敬请读者和有关专家给予批评指正。

作 者
2017 年冬

目 录

第 1 章 绪论	1
1.1 艾滋病及其流行现状	1
1.1.1 艾滋病与艾滋病病毒	1
1.1.2 艾滋病流行的时空分布	2
1.1.3 艾滋病流行的原因分析	2
1.1.4 艾滋病防控的主要策略	3
1.1.5 我国艾滋病流行的总体态势	5
1.2 艾滋病传播模型	5
1.2.1 数学模型阶段	5
1.2.2 空间分析模型阶段	6
1.2.3 智能体模型阶段	7
1.3 智能体建模技术	7
1.3.1 复杂适应系统理论(CAS)与智能体(Agent)	7
1.3.2 多 Agent 技术与智能体建模	9
1.3.3 智能体建模(ABM)的方法论	11
1.4 公共卫生地理信息技术	13
1.4.1 GIS 技术在公共卫生安全中的作用	13
1.4.2 GIS 在公共卫生领域中的应用范围	14
1.4.3 公共卫生 GIS 技术的发展趋势	15
1.5 本章小结	16
参考文献	17
第 2 章 智能体建模基础	19
2.1 智能体模型的基本原理	19
2.1.1 计算机模型	20
2.1.2 智能体	20
2.1.3 模型	21
2.1.4 实验	22
2.1.5 环境	23
2.2 智能体模型实例:捕食者-被捕食者模型	24
2.3 智能体模型的特点	26
2.3.1 本体一致性	26
2.3.2 异质的智能体	26

2.3.3	环境的表达	26
2.3.4	智能体间的交互	26
2.3.5	有限的理性	27
2.3.6	学习能力	27
2.4	智能体模型的运行环境和时间尺度	27
2.4.1	智能体	27
2.4.2	环境	32
2.4.3	随机数	32
2.4.4	时间尺度	33
2.5	智能体模型的设计	34
2.5.1	基于智能体建模的一般流程	34
2.5.2	智能体模型的设计准则	37
2.6	智能体模型的实现	38
2.6.1	智能体模型模拟工具	38
2.6.2	使用 Repast J 创建智能体模型	40
2.7	智能体模型的验证	44
2.7.1	模型程序正确性验证	44
2.7.2	模型结果有效性验证	45
2.8	本章小结	48
	参考文献	49
第3章	智能体模型和 GIS 集成的原理与方法	51
3.1	集成必要性分析	51
3.1.1	GIS 与智能体模型集成是 GIS 空间分析发展的重要方向	52
3.1.2	GIS 是构建艾滋病传播智能体模型的重要辅助手段	52
3.1.3	GIS 与智能体模型集成有利于发挥各自的优势	52
3.2	集成方法	52
3.2.1	松散耦合型	53
3.2.2	紧密耦合型	53
3.2.3	GIS 中嵌入模型	54
3.2.4	模型中嵌入 GIS	54
3.3	艾滋病传播智能体模型与 GIS 集成探索	56
3.3.1	集成框架	56
3.3.2	集成案例分析	58
3.4	本章小结	60
	参考文献	60
第4章	艾滋病传播的时空信息机理	62
4.1	引言	62
4.2	艾滋病的监测	63
4.3	艾滋病在云南传播特点	66

4.4 艾滋病传播的时空特征	69
4.4.1 艾滋病性传播的时空特征	69
4.4.2 艾滋病吸毒传播的时空特点	70
4.5 基于 GIS 的艾滋病传播时空信息机理	72
4.5.1 从现实世界到概念世界	74
4.5.2 从概念世界到 GIS 世界和规则库世界	76
4.5.3 从 GIS 世界和规则库到空间变化	78
4.6 本章小结	79
参考文献	79
第 5 章 艾滋病传播智能体模型研究	80
5.1 艾滋病传播概念模型构建	80
5.1.1 性传播途径	81
5.1.2 吸毒传播途径	84
5.1.3 综合模型构建	84
5.2 艾滋病传播的智能体模型构建	86
5.2.1 性传播途径	86
5.2.2 吸毒传播途径	89
5.2.3 综合途径智能体模型	90
5.2.4 模型实现	95
5.3 艾滋病传播智能体模型的模拟及分析	97
5.3.1 艾滋病传播智能体模型模拟结果	97
5.3.2 敏感性分析	98
5.4 艾滋病传播智能体模型验证	100
5.4.1 Z 值检验	100
5.4.2 线性拟合	102
5.5 本章小结	103
参考文献	104
第 6 章 GIS 技术在艾滋病防控中的应用研究	105
6.1 GIS 技术与艾滋病	105
6.1.1 GIS 技术在艾滋病研究中的应用	105
6.1.2 GIS 技术在艾滋病防控中面临的问题及优势	107
6.2 基于 GIS 的艾滋病信息查询	108
6.2.1 查询方式的分类	108
6.2.2 专题查询	109
6.3 基于 GIS 的艾滋病疫情分析	110
6.3.1 疫情分析	110
6.3.2 空间预警分析	111
6.3.3 疫情预案分析	112
6.4 基于 GIS 的艾滋病防控决策支持	113

6.4.1 决策的类型	113
6.4.2 决策支持系统	114
6.4.3 决策支持系统基本结构	114
6.4.4 GIS 在 HIV/AIDS 防控中的决策支持应用	115
6.5 本章小结	116
参考文献	116
第 7 章 艾滋病传播的时空模拟及其决策支持系统——以昆明市为例	117
7.1 艾滋病疫情空间数据库设计	117
7.1.1 疫情数据获取、制作及分析	117
7.1.2 疫情数据库模型设计	118
7.1.3 疫情数据库结构设计	118
7.2 昆明市艾滋病时空模拟及其决策支持系统数据库实现	120
7.2.1 SQL Server 2005 数据库及 SDE 存储空间数据的系统结构	120
7.2.2 存储文件和表空间的命名原则	122
7.2.3 数据表和视图的命名原则	123
7.2.4 数据库组成	124
7.3 艾滋病传播模拟与决策支持信息平台系统	126
7.4 艾滋病传播时空模拟模型和 GIS 集成	128
7.5 艾滋病传播的时空模拟和决策支持平台开发	129
7.5.1 功能模块	130
7.5.2 平台功能说明	136
7.6 系统应用效果	138
7.6.1 数据操作模块	138
7.6.2 GPS 数据处理模块	140
7.6.3 图形操作模块	141
7.6.4 查询模块	141
7.6.5 地图输出	143
7.6.6 地形地貌分析	146
7.6.7 专题查询统计模块	147
7.6.8 工具模块	149
7.6.9 地图渲染模块	150
7.6.10 疫情分析	151
7.6.11 系统特点	152
7.7 本系统的作用和影响	152
7.7.1 对空间流行病学发展的作用和影响	152
7.7.2 对艾滋病传播的时空模型研究的作用和影响	153
7.7.3 模拟模型与 GIS 集成技术的作用和影响	153
7.7.4 进一步的工作和前景分析	154
7.8 本章小结	155
参考文献	155

第 1 章 绪 论

1.1 艾滋病及其流行现状

1.1.1 艾滋病与艾滋病病毒

艾滋病的医学全名为“获得性免疫缺陷综合征”(Acquired Immune Deficiency Syndrome, AIDS),它是由于感染了人类免疫缺陷病毒(Human Immunodeficiency Virus, HIV)后引起的一种致死性传染病。获得性免疫缺陷综合征是指艾滋病毒感染的最后阶段,出现 20 余种机会性感染中的任一种感染或与艾滋病毒相关的任一种肿瘤,即可定义为获得性免疫缺陷综合征(WHO, 2007)。

艾滋病病原体即 HIV 病毒,医学全名为“人类免疫缺陷病毒”,是一种能攻击人体免疫系统的病毒。它把人体免疫系统中最重要 T₄ 淋巴细胞作为攻击目标,大量吞噬、破坏 T₄ 淋巴细胞,从而破坏人的免疫系统,最终使免疫系统崩溃,使机体逐渐丧失防卫能力而不能抵抗外界的各种病原体(宝福凯, 1994)。即使一个对一般健康人来说微不足道的感染,如小伤口或普通感冒,也可以致艾滋病人于死地。HIV 属于反转录病毒科,慢病毒属,灵长类免疫缺陷亚属,其显著特征是基因具有高度变异性。现已证实 HIV 分为两型: HIV-1 型和 HIV-2 型,它们又有各自的亚型(从 A 到 H 及 O 型)。不同地区流行的亚型不同,同一亚型在不同地区也存在一定差异,但都会引起相同的艾滋病。

HIV 病毒对外界环境的抵抗力较弱,离开人体后,常温下只可生存数小时至数天。高温、干燥以及常用消毒剂都可以杀灭这种病毒。HIV 病毒在人体内的潜伏期平均为 12~13 年。在潜伏期内 HIV 感染者没有任何症状和体征,他们能和正常人一样工作生活。一个人从感染 HIV 到死亡都具有传染性,是终生的艾滋病传染源(王岚等, 2011)。HIV 感染后导致人体免疫机能缺陷,从而使患者发生机会性感染等一系列临床综合征,病死率几乎达 100%。

艾滋病临床上分为 3 期(杨春燕, 2004): ① 艾滋病病毒感染期。病人感染艾滋病病毒后没有任何临床症状,仅血液中的艾滋病病毒抗体检测结果为阳性。② 艾滋病相关综合征。在淋巴器官中大多可检测出病毒。③ 艾滋病阶段。病人突出的表现是致病性感染、恶性肿瘤的发生以及找不到原因的细胞免疫缺陷,病人一般在 6~24 月内死亡。

自 1985 年发现首例艾滋病开始,中国在近 30 年的防治过程中,与许多发展中国家一样,面临控制效果差、疫情不断扩散的严峻局面。对艾滋病的预防与控制已不仅仅是

公共卫生领域要面对的问题,更是一种社会问题和政治问题,因此进行艾滋病传播的模拟研究具有重要的现实意义(国务院防治艾滋病工作委员会,2006;中国疾病预防控制中心和性病艾滋病预防控制中心,2009)。

1.1.2 艾滋病流行的时空分布

据报道,截至2014年12月,全球共有艾滋病患者约3690万人,2014年新增200万人,死亡120万人(数据来源:联合国儿童基金会和联合国艾滋病规划署2015年数据)^①。截至2014年10月底,我国累计报告艾滋病病毒感染者和艾滋病病人500679例,其中艾滋病病人204683例,死亡报告158743例。中国卫生部^②、联合国艾滋病规划署和世界卫生组织联合^③对2014年中国艾滋病疫情进行了估计,结果显示:截至2014年底,中国现存艾滋病病毒感染者和病人295996例,当年新发艾滋病病毒感染者103501人,死亡20650例。艾滋病已经成为中国乃至全人类最具威胁的健康杀手之一。云南省是中国艾滋病的高发区之一,自1989年云南省首次发现艾滋病感染者以来,截至2014年,云南省累计现存活艾滋病病毒感染者和病人79915例,其中艾滋病病毒感染者50263例,艾滋病病人29652例,死亡23316例。全省16个州(市)中,3个处于高度流行期,11个处于中度流行期,2个处于低度流行期。

1.1.3 艾滋病流行的原因分析

艾滋病流行的原因可以从不同角度进行分析。从大的方面来讲,其流行有社会、政治、文化、科技等诸多因素的驱动;从小的方面来讲,主要体现在艾滋病的传播途径、艾滋病的病理特征以及艾滋病的社会特征等方面(程茂金,2003;王立秋,2003)。当前,按照国家卫生部对流行病的分类,艾滋病属于乙类传染疾病,它的传播条件具有传染病的一般特点,即:排除、存活、进入和足量。此外,它还具有潜伏期长、隐蔽性强、传播途径多样化、病死率高、目前既无能治愈的药物,也无有效的疫苗等特点,给防治工作带来了诸多的困难和问题,这些是促使艾滋病流行的最根本原因(钟福华,2007)。

1. 艾滋病传播途径的特殊性

目前,国际公认的艾滋病传播途径主要有三种,即:性传播、血液传播和母婴(垂直)传播(吐尔逊·尤努斯和王珏,2008)。

(1)性传播: HIV是一种传播性疾病。位于黏膜上皮表面的LangerHans细胞有助于其传染,且该细胞易被感染。而其他能引起黏膜溃疡和炎症的性传染疾病,也有助于HIV的感染。HIV有嗜 CD_4^+ T淋巴细胞的特性, CD_4^+ T淋巴细胞表面的受体,可促进

① <http://www.unaids.org/en/>

② <http://www.chinacdc.net.cn>

③ <http://www.who.int/features/en/>

HIV 与其结合并进入细胞内。当与艾滋病带菌/患者发生不安全的性行为,将接触到受感染的血液、精液或阴道分泌物,从而导致感染。

(2)血液传播:血液传播是感染概率最高的途径之一,一般包括以下几种形式:①输入被病毒污染的血液或血液制品,如输血感染;②使用了被血液污染而又未经严格消毒的注射器、针灸针、拔牙工具;③与艾滋病病毒感染者共用一只未消毒的注射器,如静脉注射吸毒感染,而蚊虫叮咬(通常认为病毒不“足量”)则不会传播;④共用牙刷和剃须刀也有可能造成艾滋病感染。

(3)母婴(垂直)传播:如果母亲是艾滋病感染者,那么她可能会在以下三种情况传染艾滋病给小孩:①怀孕,病毒传染给子宫内的胎儿;②分娩过程;③母乳喂养。

从其传播途径可以看出,不论是哪种传播途径,其实质都是被感染者与感染者有足量的体液或者血液交换过程。虽然这些交换过程并不像流感或者 SARS 等传染病那样容易发生,但都普遍存在于人类一般的行为范畴之内,因此,艾滋病的传播途径虽然隐蔽,但却非常普及,这是导致其在全世界范围流行的一个重要原因之一(俞志贤,1991)。

2. 艾滋病具有较强的易感性

虽然艾滋病有其相对显著的易感人群或者环境,但是对于接触或者经历了其传播途径的人群而言,感染艾滋病病毒的概率相对较高。这主要是因为:①艾滋病病毒一旦感染,以目前的医疗水平尚不能将其杀死于人体内,所以一次感染,终身携带;②艾滋病病毒目前没有有效的免疫疫苗,使得人体缺乏抵御艾滋病病毒入侵的抵抗力。因此,艾滋病的这种易感性也是促其流行的重要原因之一。

3. 艾滋病疫情监控相对困难

从艾滋病的传播途径可以了解到,艾滋病的传播过程极具隐蔽性,常规的流行病监测方法对于艾滋病而言很难奏效。艾滋病疫情的这种难以监测性直接导致艾滋病防控的复杂性和艰苦性,使得艾滋病疫情能在较短时间内呈现不可控制的爆发性,从而导致了艾滋病在特定区域或者时间段上的流行性(柏涌海等,2014)。

1.1.4 艾滋病防控的主要策略

近几年来,在各级防艾组织和管理部门的不懈努力下,国际社会已经清晰地认识到了艾滋病的危害性,不管是政府还是个人,都在观念上和行为上有了较强的艾滋病防控意识,在艾滋病疫情监测方法以及模拟技术方面也有了长足的进步。综合近十年来人类在抵御艾滋病危害方面的成果,当前艾滋病防控的主要策略有如下几种。

1. 防艾宣传

历史的经验告诉我们,永远不要低估了宣传的作用。广泛而深入的防艾宣传,可以让更多的人了解艾滋病的病理特征、传播途径及其危害性,从而唤醒人们对于艾滋病的抵御意识,并且指导人们如何对艾滋病进行有效的预防和治疗(陈琦等,2001)。目前,

防艾宣传是国际社会公认的最能从意识形态上根本消除艾滋病的且最行之有效的防控措施。

2. 医学防艾

通过研制艾滋病疫苗和抗艾药物,拟从生物学和医学的角度寻求消灭或者控制艾滋病的一种防控策略(胡国龄,2004)。如果说防艾宣传是从人类的意识层面来达到消减艾滋病的目的,那么,医学防艾则更加侧重于艾滋病的物理和生物特征,是有望从根本上消除艾滋病的防控措施。但是令人遗憾的是,到目前为止,还没有任何一种药物或者疫苗可以达到这种理想效果。因此,虽然医学防艾是一种根本解决途径,然任重而道远,短时间内很奏效。

3. 加强监测

当前艾滋病疫情的监测主要有三大类,即主动监测(active surveillance)、被动监测(passive surveillance)和行为监测(behavior surveillance survey)(吴尊友,2009)。主动监测主要是指防艾部门主动通过一些监测方法来获取和甄别艾滋病个案,这包括哨点监测、血清横断面调查和常规筛查等方法;被动监测主要是指案例报告,即常规医疗机构(非哨点)在监测到艾滋病个案后主动将该案例报告给地方或中央艾滋病预防与控制部门(比如疾病预防控制中心或者艾滋病防治办公室等政府部门);行为监测是一种混合监测模式,其主要目的是监测艾滋病高危人群的疫情变化,常用的方法是抽样统计学。

4. 信息防艾

长期以来,各领域的专家学者一直都致力于艾滋病防治的研究,期望能够找出一条有效的途径来模拟艾滋病的传播,为疾病的预防和控制提供依据。医学家从病理学、病原学、临床医学等角度出发,试图找寻阻隔艾滋病传播的药物;社会学家采用队列研究、抽样调查等方法,从人口学、民族学、伦理学等方向,探究不同经济、文化和社会背景对艾滋病传播的影响,并号召人文关怀、减少对艾滋病人的歧视,让艾滋病人得到更好的治疗;生物学家从动物模型、基因重组等入手,探寻防治艾滋病病毒方法;公共卫生专家从疾病预防入手,深入一线掌握艾滋病流行的一手资料。而信息科学,作为一个新兴的学科,在数据管理、分析、模拟、表达上,有着自己独特的优势,信息防艾也越来越受到人们的重视和青睐。信息防艾的目标之一便是通过引入信息化技术,分析和模拟艾滋病的传播过程,预测艾滋病疫情的发展趋势,从而为艾滋病防控部门提供科学的决策依据,辅助政府决策机构科学而合理地进行艾滋病的预防与控制(刘黎等,2010)。信息防艾的方法也是百花齐放,且具有交叉性,可以从纯数理统计的角度来研究,也可以从数学模拟的角度来研究,还可以从系统论和信息论的角度加以研究(杜长慧等,2010)。值得指出的是,当前基于地理信息系统(Geographic Information System)技术的时空模拟技术已然成为国内外信息防艾的热点研究领域(威晓鹏,2004;杨昆等,2008),而这也恰恰正是本书所要着重论述的内容。

1.1.5 我国艾滋病流行的总体态势

1. 我国报告的感染者和病人人数增加

截至2014年10月底,我国累计报告艾滋病病毒感染者和艾滋病病人500679例,其中艾滋病病人204683例,死亡报告158743例。

2. 三种传播途径并存

2014年1~10月报告的已开展个案流调的感染者中,异性性传播68719例(66.4%),同性性传播26746例(25.8%),注射毒品传播5783例(5.6%),母婴传播731例(0.7%)。

3. 经性途径感染呈上升趋势

2014年新报告的8.7万病例,性途径传播占91.5%,其中异性性传播占66%,同性性传播占25%。

吸毒人群感染率呈下降趋势,从2005年的7.5%下降到2014年的平均3.3%,男同性恋呈上升趋势,虽然采取了措施,但是疫情还是在上升,从2005年的1.4%上升到2014年的7.7%。

4. 艾滋病疫情进一步蔓延的危险因素仍然存在

哨点监测数据显示,目前暗娼人群坚持使用安全套的比例为38.7%,注射吸毒人群共用针具吸毒的比例为50.8%。另外,感染者的流动、大量人口流动和性病疫情上升等是造成艾滋病蔓延的重要因素。

1.2 艾滋病传播模型

从艾滋病流行的现状分析不难看出,艾滋病疫情防控的重点还是在于控制艾滋病传播的速度与范围,即艾滋病应该实施“预防重于治疗”的策略。同时,随着信息技术在艾滋病防控中应用的深入,出现了以模拟艾滋病传播过程为主要目标的信息化防艾新技术,这就是艾滋病传播模型。

截至目前,研究艾滋病传播的主要模型包括数学模型、艾滋病的动力学模型、基于个体的元胞自动机模型和基于多智能体的传播模型等,这些模型较好地反映了艾滋病传播模型从宏观到微观、从静态到动态、从自动化到智能化的发展历程(杨昆等,2008)。

1.2.1 数学模型阶段

数学是建模的基本工具之一。1981年,当发现首例艾滋病病例的时候,就有数学模

型应用在艾滋病领域。当时的研究主要关注点在于模拟艾滋病病毒 HIV 在机体细胞内部怎样进行传播扩散(井竹君, 1991)。数学模型包括有基于 SIR 思想的改进型微分模型和 Workbook 模型、AEM(Asia Epidemic Model)模型等, 前者把系统人群分为三类, 即易感染者、已感染者和移出系统者(死亡者), 根据研究区的历史数据建立微分方程以期模拟研究区的艾滋病传播情况; 后者是联合国艾滋病署、亚太艾滋病署等国际机构, 根据亚洲主要流行特点, 利用已有的艾滋病病毒(HIV)相关流行病学和行为学资料, 对 HIV 流行和防治进行估算和预测的半经验化过程模型(喻达等, 2008)。系统动力学模型也是一种数学预测模型, 它有离散型和连续型, 其基本思想是把人群分为易感人群(如 FSW、MSM 等)、感染人群以及患病人群。由于系统 $t+1$ 时刻易感者人数等于 t 时刻易感者人数加上在这段时间内新进入的人数, 减去被感染的人数和死亡的人数, 然后建立一个推替方程组。在获取某一地区历年来 HIV 感染者和 AIDS 患者的人数以及每年新增加的高危人群数 N , 就可以估算出感染率 $\lambda(t)$ 及 γ , 进而可以预测出未来年份的 HIV 感染者和 AIDS 患者的人数(蔡雨阳等, 2010)。

数学模型的建立主要依赖于研究者对研究对象的熟悉, 对不同影响因子的权重的确定是关键, 通过多次反复的实验来选择和确定较为合适的权重比例。运算的简单化、影响因子的不确定性、影响因子权重的确定等都是数学模型是否有用的关键所在, 然而这些关键因子的确定往往需要很准确的历史信息, 这对于具有隐蔽性和隐私性的艾滋病而言简直是天方夜谭, 可想而知其参数的率定和完善过程是多么的艰辛和漫长。所以, 数学模型是有用的, 但它的复杂性也决定了它的可操作性、可使用性较低。

1.2.2 空间分析模型阶段

数学模型除了建模困难之外还有一个明显的缺陷, 那就是很难将艾滋病传播过程的空间特征考虑进去, 而艾滋病传播本身具有很强的空间特性, 如何将空间特征纳入艾滋病传播模型一直以来都是本领域研究的热点(钱莎莎, 2014)。当前发展迅速的 GIS 技术是关于空间信息采集、管理、处理和分析的新兴信息学科, 能够较好地解决艾滋病传播模型中的空间因子的获取、确定和计算问题, 是当前艾滋病传播模型研究的前沿领域, 具有广阔的发展前景(钱莎莎等, 2014)。

通过研究发现, GIS 在流行病传播模型的应用研究已然兴起, 这些研究成果和经验对于研究艾滋病传播空间分析模型具有重要的参考价值和借鉴意义。云南师范大学的杨昆教授于 2005 年承担的国家自然科学基金项目“云南省艾滋病传播时空分析模型及其地理信息技术支持平台研究”为艾滋病传播的空间分析模型的研究写下了浓墨重彩的一页, 是第一个将 GIS 技术完整应用于艾滋病传播过程的国家级项目, 极大地促进了艾滋病传播空间分析模型的发展(陈洪萍, 2007; 杨昆等, 2008; 姚信英, 2008; 谭骏翔, 2009; 崔庆雄, 2009)。

但是, 单纯性的研究艾滋病传播的空间分析模型也存在着弊端, 那就是如何将传播过程的时间特征进行计算和表达, 这也是该类模型今后需要重点解决的问题。

1.2.3 智能体模型阶段

智能体模型(ABM)理论是建立在复杂系统理论和人工智能的发展之上,对元胞自动机模型(CA)、计算机仿真实理论的一种拓展(Daniel et al., 2005),它将每一个智能体放在一个开放的环境中去学习、适应和生存,在很大程度上符合人类活动的现实,是个体、空间和时间相结合的最佳切入点。结合 GIS 技术,将现实世界中的研究对象作为一个智能体,即从现实世界中去获取属性数据,通过规则库的建立,经由模拟平台在时间上的动态模拟,来实现对现实世界的最大程度上的还原,最后再通过决策支持和应急响应系统的完善,完成对研究对象的表达和描述(潘志方等, 2011)。因此,我们认为智能体模拟是对空间分析方法的一种完善和突破,它在很大程度上解决了传统 GIS 在时间表达上的缺陷,是一种真正的时空分析模型,是目前可用来模拟艾滋病传播的最适宜的模式。

当前艾滋病传播的智能体模型研究是国际上本领域的热点和重点,代表了艾滋病传播模型的发展方向和趋势,本书作者有幸于 2007 年承担了国家 863 计划“艾滋病传播的智能体模型及其与 GIS 的集成技术研究”项目,深入探究了艾滋病传播的智能体模型的构建原理、方法和验证相关问题,取得了一系列相关成果,本书也是该项目成果之一。

1.3 智能体建模技术

1.3.1 复杂适应系统理论(CAS)与智能体(Agent)

1. 复杂适应系统(CAS)概述

复杂性科学是 20 世纪 80 年代兴起的一个研究范畴,是主要研究复杂系统和复杂性的科学。它是系统科学和非线性科学的进一步发展、充实和深化,是系统科学研究的最新、最前沿的领域,是人类面临的一个崭新的认识领域。复杂性科学被称为 21 世纪的科学,它的主要目的就是要揭示复杂系统的一些难以用现有科学方法解释的动力学行为(Fleming and Sorenson, 2001)。图 1-2 描述了复杂系统与其他系统的区别。

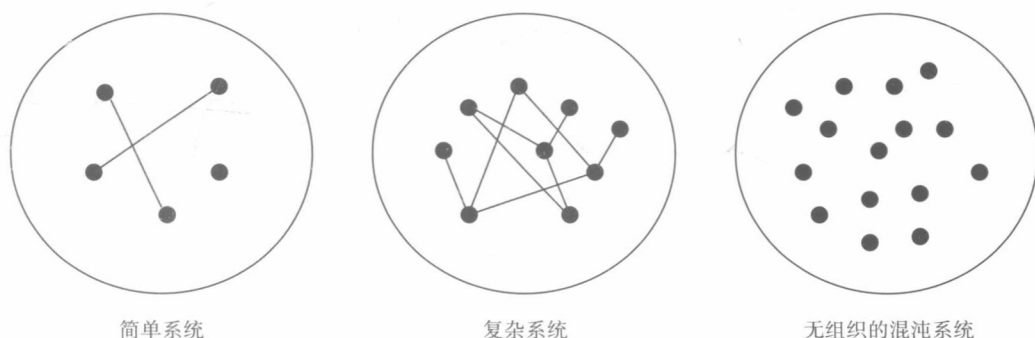


图 1-2 简单系统、复杂系统与混沌系统

目前研究复杂系统较有影响力的有三个学派,即欧洲学派、以钱学森的“开放复杂巨系统”理论为代表的中国学派和美国圣塔菲研究所(Santa Fe Institute, SFI)的“复杂适应系统(Complex Adaptive System, CAS)”理论为代表的美国学派(戴金海等, 2003)。

CAS理论是遗传算法的提出者——霍兰(Holland)教授于1994年在SFI成立十周年时正式提出的,CAS理论受到来自于生物系统及对于遗传算法研究的启发,并借鉴了经济学中的代理人(Agent)概念,提出了适应性主体(Adaptive Agent)的概念。其基本思想可以概括如下:系统中的成员称为具有适应能力的主体(Adaptive Agent),简称为主体或者智能体(Agent)。Agent是具有自身目的性与主动性,有活力和适应性的个体。Agent可以在持续不断地与环境以及其他主体的交互作用中“学习”和“积累经验”,并且根据学到的“经验”改变自身的结构和行为方式,正是这种主动性及Agent与环境的、其他Agent的相互作用,不断改变着它们自身,同时也改变着环境,才是系统发展和进化的基本动因。整个系统的演变或进化,包括新层次的产生、分化和多样性的出现,新的聚合而成的、更大的主体的出现等,都是在这个基础上派生出来的。

2. 复杂适应系统(CAS)理论

单独用Agent这一概念,无法完全表达CAS理论的丰富内容。围绕Agent这一核心概念,霍兰进一步提出了适应和演化过程研究中特别重要的7个有关概念(Holland, 2006),它们包括4个特性(聚集、非线性、流、多样性)和3个机制(标识、内部模型、构件)。

(1)聚集(Aggregation):主要用于个体通过“黏合(Adhesion)”形成多主体的聚合体(Aggregation Agent)。聚合体具有新的功能和属性,在系统中像一个单独主体一样行动。

(2)非线性(Nonlinearity):指个体及它们的属性在发生变化时,并非遵从简单的线性关系。特别是与系统或环境的反复交互作用中,这一点表现得更为明显。

(3)流(Flow):在主体与环境之间、主体与主体之间存在着物质流、能量流和信息流。这些流的渠道是否通畅、周转迅速到什么程度,都直接影响着系统的演化过程。

(4)多样性(Diversity):在适应过程中,由于种种原因,个体之间的差别会发展与扩大,最终形成分化,这是CAS的一个显著特点。CAS理论强调,正是相互作用和不断适应的过程,造成个体向不同方向的发展变化,从而形成了个体类型的多样性。而从整体来看,这事实上是一种分工。如果和前面提到的聚合结合起来看,就是宏观上看到的“结构”的“涌现”,即所谓“自组织现象”的出现。

(5)标识(Tagging):为了相互识别和选择,主体的标识在主体与环境的相互作用中是非常重要的。无论在建模中,还是在实际系统中,标识的功能与效率是必须认真考虑的因素。标识的意义在于提出了主体在环境中搜索和接收信息的具体实现方法。

(6)内部模型(Internal Models):在CAS中不同层次的个体都有预期未来的能力,每个个体都是有复杂的内部机制的。对于整个系统来说,这就统称为内部模型。

(7)构件(Building Block):复杂系统常常是相对简单的一些部分通过改变组合方式而形成的。因此,事实上的复杂性往往不在于构件的多少和大小,而在于原有构件的重