



计算机科学

COMPUTER SCIENCE

面向超大规模社会系统仿真的概念模型

张明新

引用本文

张明新. [面向超大规模社会系统仿真的概念模型](#)[J]. 计算机科学, 2022, 49(4): 16-24.

ZHANG Ming-xin. [Conceptual Model for Large-scale Social Simulation](#)[J]. Computer Science, 2022, 49(4): 16-24.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于多 Agent 联合决策的队组协同攻击规划](#)

Team Cooperative Attack Planning Based on Multi-agent Joint Decision

计算机科学, 2021, 48(5): 301-307. <https://doi.org/10.11896/jsjcx.200800174>

[有适应力的分布式状态估计方法](#)

Resilient Distributed State Estimation Algorithm

计算机科学, 2021, 48(5): 308-312. <https://doi.org/10.11896/jsjcx.200300117>

[异构无人机编队防御及评估策略研究](#)

Study on Heterogeneous UAV Formation Defense and Evaluation Strategy

计算机科学, 2021, 48(2): 55-63. <https://doi.org/10.11896/jsjcx.191100053>

[多智能体强化学习综述](#)

Overview on Multi-agent Reinforcement Learning

计算机科学, 2019, 46(8): 1-8. <https://doi.org/10.11896/j.issn.1002-137X.2019.08.001>

[多智能体模态逻辑系统 \$KD45_n\$ 中的知识遗忘](#)

Knowledge Forgetting in Multi-agent Modal Logic System $KD45_n$

计算机科学, 2019, 46(7): 195-205. <https://doi.org/10.11896/j.issn.1002-137X.2019.07.030>

面向超大规模社会系统仿真的概念模型

张明新

国防大学政治学院 上海 200433

摘要 超大规模基于智能体的社会仿真正逐渐被证明是研究人类社会的一种有效方法,它可以为社会科学中的决策、计算机科学中的分布式人工智能和智能体技术、计算机仿真系统的理论和建模实践等领域作出贡献。然而,现有的研究实践在平衡模型复杂度和仿真性能方面存在一定的困难。针对目前存在的问题,提出了一种基于智能体和大数据驱动的超大规模社会仿真概念模型框架,提供了模型组件的参考实现,并以超大规模人工城市疫情预测与控制为例,说明了如何利用所提概念框架对具有复杂人类行为和社会交互的超大规模社会系统进行建模,同时也指出了在其他社会科学领域的潜在应用,如微观交通系统和城市疏散规划。

关键词: 社会仿真;智能体;概念模型;超大规模;社会交互

中图法分类号 TP391

Conceptual Model for Large-scale Social Simulation

ZHANG Ming-xin

College of Politics, National Defense University, Shanghai 200433, China

Abstract Large-scale agent-based social simulation is gradually proved to be an effective method for the study of human society. It can contribute to decision-making in social science, distributed artificial intelligence and agent technology in computer science, theory and modeling practice of computer simulation system, etc. However, the existing research practice has difficulties in balancing model complexity and simulation performance. In view of the existing problems, this paper proposes a conceptual model framework of large-scale social simulation based on agent and big data driving, and provides the reference implementation of model components. Taking the epidemic prediction and control in a large-scale artificial city as an example, it illustrates how to use the proposed conceptual framework to model the large-scale social system with complex human behavior and social interaction. It also points out the potential applications in other social science fields, such as micro transportation system and urban evacuation planning.

Keywords Social simulation, Agent, Conceptual model, Large scale, Social interaction

1 引言

新冠疫情的爆发使得对超大规模人类社会的突发事件(几千至上万、百万人参与或被影响)的研究分析和预测的重要性与日俱增。在现有的研究方法中,社会仿真(Social Simulation)是一种对人类社会中的社会现象进行建模、理解或预测的应用计算方法。Epstein等^[1]于1996年发表了《生长人工社会:从底向上的社会科学(Growing Artificial Societies: Social Science from the Bottom Up)》一文,奠定了基于智能体的社会仿真方法在人类社会分析与预测研究领域的核心地位。

当前,基于智能体的社会仿真方法对社会突发事件的研究成果不胜枚举,如新冠疫情爆发以来,在ScienceDirect上检索相关研究超过了1000篇,典型研究如评估居家隔离、戴

口罩等措施的作用等^[2]。在大量既有研究的基础上,本文针对当前超大规模复杂社会系统建模中存在的模型复杂度与计算效率之间的矛盾问题,提出一种基于智能体和大数据驱动的超大规模社会仿真概念模型框架,提供了模型组件的参考实现,给出了超大规模传染病的实践案例。实践案例仿真结果显示,本文提出的概念模型不仅克服了一般智能体概念模型对环境完备性的限制,同时较好地解决了模型精度与计算效率之间的平衡难题。

2 研究综述

近年来,随着人工智能和计算理论的发展,基于智能体的社会仿真(Agent-based Social Simulation, ABSS)研究不断兴起。且随着计算资源(如GPU计算、超级计算和云计算)的日益扩张,ABSS的研究热点正在逐步转向超大规模的社会

到稿日期:2021-09-16 返修日期:2022-01-18

基金项目:国家自然科学基金(17CGL047)

This work was supported by the National Social Science Foundation(17CGL047).

通信作者:张明新(116455436@qq.com)

系统。以大规模流行病为例,为更好地理解流行病的传播机制以及有效地辅助疾病爆发期间的政策制定,学界研发了针对不同疾病、不同干预措施和不同场景的若干大型 ABSS 模型^[3-4]。这些模型有一个共同的特点,即都是从零开始构建的,没有一个是基于现有的 ABSS 平台或工具开发的。

对于基于智能体的计算学界(Agent-based Computing)中现有的模型和平台来说,针对大规模流行病研究的困难可以解释为复杂智能体行为(如规划和推理)带来的计算复杂性限制了仿真研究中智能体的数量。为了降低计算复杂性,通常在如 Brahms^[5]等语言中将智能体实现为基于线程的对象。然而,仿真过程中大量线程之间的同步协作会大大降低系统的性能,反过来又限制了系统规模。

对于社会科学学界(Social Science)的工具和工具包,学者们借助高性能计算中间件,开发了大型 ABSS 平台,如 Re-Past HPC^[6]和 Distributed MASON^[7]。然而,这些平台在研究大规模疫情预测与控制方面也存在一些局限性。首先,为便于模型的开发,这些平台中的智能体模型通常是封装和简化的,如 HLA-GRID-REPAST^[8]将智能体的环境看作是离散单元/网格,辅助模型划分以减少不必要的节点间的通信消息。在理想情况下,当智能体的空间移动没有特定模式或非常受限时,这可能是一种明智的方法。然而,在目前基于智能体的流行病学研究中,脱离封闭房间的直接物理接触(如接触)或间接传播接触(如打喷嚏)也被认为是一种有效的疾病传播方法,特别是在人口稠密的地方,如公共交通上。因此,需要对智能体所处的环境进行精细的连续建模,而不是基于单元/网格的离散建模。其次,由于这些平台中封装了智能体的决策逻辑概念,即一般表示为决策规则用以简化智能体建模,这不足以对紧急情况下的智能体应对疾病传播的复杂行为进行建模。第三,基于时间步长的仿真推进机制是在诸如 Repast 和 NetLogo 等工具中普遍采用的时间推进方法。在构建大规模流行病模型时,该机制会遇到比较严峻的挑战,如在大跨度的流行病模型中,时间步长设置过小会大大降低系统的运行性能,而加大步长则可能导致关键事件缺失,从而导致错误的仿真结果。

对于计算机仿真领域(Computer Simulation)的现有平台来说,面向对象方法在处理大规模系统时具有先天优势,特别是在并行和分布式仿真技术已经得到了广泛研究的情况下。然而,现有平台中针对智能体的概念也被大大简化,并且在这些平台中实现的社会网络通常也是简化和固定的,因为如果一个大规模的模型以并行和分布式方式运行,智能体复杂的社会交互会显著增加同步消息的数量。

因此,尽管数学家、流行病学学家、计算机科学家和社会科学家在研究流行病传播方面有着共同的兴趣,并且都依赖于非常相似的模型(如 SIR 等)来描述疾病的扩散,但实际上目前学界并没有一个用于流行病预测和控制研究的通用的架构或平台。现有的大规模流行病 ABSS 模型一般都是基于面向对象语言从头构建的,且存在以下局限性。

首先,尽管采用分布式/并行机制保证了模型的运行性能,但在超大规模 ABSS 模型中采用分布式/并行机制的前提通常是需要简化模型组件,包括智能体模型本身。因此,这些智能体通常是建模成简单的反射性智能体。换句话说,从

智能体的架构和决策机制来看,用于流行病研究的智能体比较简单。这些简单智能体要么基于简单规则作出决策,要么按照事先设定的时间-事件表行事。然而,流行病仿真的结果实际上与个体作出的决策行为高度相关。具体而言,传染病爆发期间智能体决策的差异和多样性会显著影响疾病干预措施的效果。

其次,在现有的超大规模 ABSS 模型中,智能体的社会制约,包括法规制度等(如宵禁等措施),并没有在模型中体现。然而这些规范和制度显然在疾病传播中发挥着重要作用,因为在现实社会中,它们会在制度上影响人群的行为与决策。

最后,大规模地模拟复杂的社交行为仍然是一项具有挑战性的工作。目前超大规模的 ABSS 模型忽略或简化了现实生活中社会交往的人类互动行为,如集体活动中的协商行为。现有模型中的智能体只会执行其预设的行为,并基于随机或预定义的社会网络与他人交互,当然也就无法描述疾病爆发期间社会关系的动态演变过程。

总而言之,以超大规模流行病的预测和防控为例,现有的超大规模社会仿真的基础架构研究仍然不充分,核心问题在于:现存的超大规模的 ABSS 模型与研究特定的超大规模社会性突发事件(如流行病预测与控制)的要求存在差距。

3 通用概念模型框架

3.1 模型概念和概念模型框架

在本节中,我们将为基于智能体的大规模社会仿真提供一个概念模型,该模型借鉴了其他流行的智能体模型概念框架中的许多概念。

这个概念模型中包含 5 个概念,如图 1 所示,分别如下:

(1)智能体(Agent)。智能体代表现实世界中的人类。

(2)物理容器(Physical Container)。物理容器用于表示智能体所在的物理环境。物理容器一般是分层组织的,每个智能体在任何时候都必须至少停留在一个物理容器中。典型的物理容器是卧室、办公室、教室和公共汽车等。

(3)社会监管(Social Regulation)。社会监管是设计用于模拟诸如社会规范、机制等能够在全局范围内影响人类行为的概念。例如,“疫情期间关闭公共交通”的政策可以被建模为流行病预测和控制系统中的一种社会监管。

(4)功能实体(Functional Entity)。功能实体是系统中可以影响或直接改变智能体、物理容器或社会监管等概念属性的主体对象。例如,疾病本体可被建模为功能实体来改变智能体的健康状态;风暴可被建模为功能实体来改变房间(物理容器)的温度。

(5)社会网络(Social Network)。社会网络可将智能体的社会关系定义为不同的类别,如家庭成员、同学和朋友。

与一般基于智能体的概念模型中的模型概念相比,上述概念模型的主要区别在于智能体环境的概念被区分为物理容器、社会监管和功能实体 3 个部分。这种区分克服了其他智能体模型对于环境概念不够完整的限制,并增强了模拟不同系统场景的灵活性,此外环境概念本身也得到了完善。例如,本文将可移动的物理容器用来表示在一般智能体平台上难以实现的交通工具。人工智能领域中关于社会监管的理论和概念也可以很容易地实现并集成到

模型中,同时表现出较好的性能。

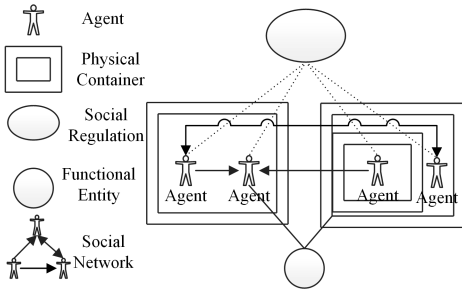


图1 超大规模社会仿真概念模型

Fig. 1 Conceptual model of large scale social simulation

本文提出的概念模型还借鉴了经典的面向对象仿真模型中的概念,即可以看作是一个混合概念模型。这种设计选择如图2所示。

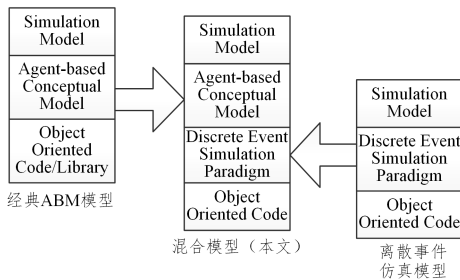


图2 概念模型的混合设计模式

Fig. 2 Hybrid design pattern of conceptual model

面向对象的离散事件仿真(图2右侧)中,UML兼容图应用广泛,它以3种模型表示所需的概念信息:对象模型、动态模型和功能模型。其中对象模型声明系统中的类,包括其属性和服务;动态模型指定了对象间的交互;功能模型则代表与对象状态变化相关联的信息。

受此启发,借鉴经典ABM(Agent-based Model)概念模型(图2左侧),我们将功能实体的概念引入大规模ABSS的概念模型中。这种概念的清晰分离,使得面向对象的范式更容易编程实现。后续实验结果也表明,采用这种概念模型的模型在系统性能和模型的高保真度两个方面的表现都较好。

3.2 智能体

与一般ABM概念模型中的智能体概念类似,大规模ABSS模型中的智能体概念在实现之前应设计好内部架构。

针对当前多层体系结构和BDI体系结构在应用于大规模社会系统时都不“简单”和不够“小”,不足以胜任超大规模的基于智能体的社会仿真的现状问题,本文研究设计了一种新型的基于智能体的大规模社会仿真智能体体系结构,该结构简化了传统的多级体系结构,但仍保持一定的决策能力。

大规模的社会系统通常包含数以百万计的智能体。这些智能体典型的行为调度的实现是基于活动的,且所有活动都在输入数据源中预定义或在仿真运行之前生成,这种做法会消耗大量内存。假设系统有大约2000万个智能体,每个智能体每天预设10个活动,那么4周仿真时长的活动总数达56亿。为了减少内存消耗,本文将基于活动模式的方法引入多层体系结构中,这是受Mossong等^[9]的研究启发。他们认为相同国家的人的行为模式具有相似性,并且这些模式与年龄高度相关。

基于活动模式的方法最早出现于20世纪70年代的出行规划领域。其核心在于将个体日常从事的活动和出行的序列定义为一天中个人的活动模式。受此方法的启发,我们构建了一个用于大规模智能体社会仿真的智能体模型体系结构(如图3所示),它是基于活动模式的方法和多层次体系结构的集成。该智能体体系结构由3个主要部分组成:1)智能体对象;2)活动模式;3)多级决策模块。

作为智能体体系结构的一部分,智能体对象是本文提出的智能体体系结构的基本组件。它被设计为智能体行为主体和状态指示器的集合体,而决策等具体任务则由其他组件承担。每个智能体对象都与一组属性变量和状态变量相关联。属性变量主要包括智能体的个人信息,如性别、年龄、活动模式、家庭位置和工作/学校位置等;状态变量主要包括智能体的当前行为和状态,如其当前所处的物理容器、当前的活动和当前活跃的社会网络。

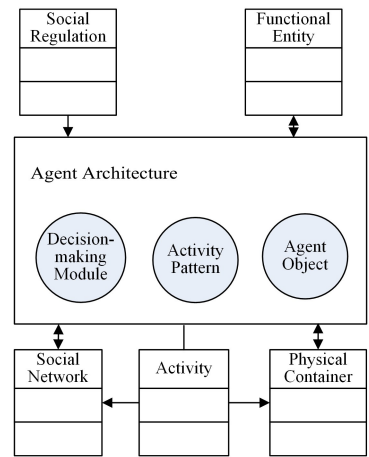


图3 智能体体系结构

Fig. 3 Agent architecture

对于给定的智能体,活动模式指定其将执行哪些行为计划。基于各自的活动模式,智能体将主要根据活动的初始预定义序列执行其活动。但是,活动模式不会指明这些活动发生的时间、地点以及和谁一起执行,因为这些信息会在仿真过程中由决策模块实时决定。

活动模式可以基于任意时间段(如每小时、每天和每周),具体选择取决于仿真研究的问题域和时间范围。例如,如果研究的是一个长期的流行病问题建模,一个大学生可以被分配一个大学生的周活动模式,而一个工人可以被分配一个工人的周活动模式。同时,一个模式可以由多个智能体共享,模式也可以分层组织。例如,一个周模式可以由7个日模式组成。对于一般在办公室工作的员工来说,周一至周五前5天的模式可以保持不变,最后两天可以是两个重复的周末模式。在退休人员的一周模式中,七日模式可以是相同的。

决策模块是智能体的大脑,它是专门针对一些实时决策问题而设计的,可被认为是对智能体活动模式的补充。本文将智能体设计为具有3级决策能力,主要通过将智能体环境的概念分解为3个部分来实现。

首先,虽然活动模式中的活动顺序是预先定义的,但是智能体仍然可以根据偏好在常规活动的执行过程中调整其活动位置和持续时间。本文研究将此过程定义为操作层面的决策。

活动地点的操作决策主要由物理容器支持。除了活动地点之外,活动持续时间的确定也是智能体操作级决策的一个方面。

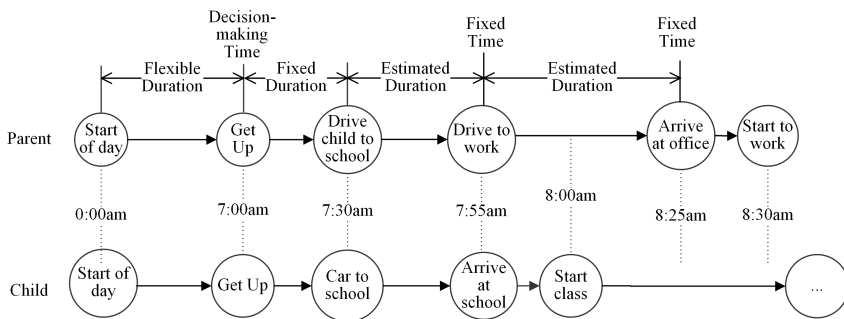


图4 父母与孩子活动模式之间的协调

Fig. 4 Coordination of activity patterns between parents and children

从图4中可以发现,父母必须在早上7点左右起床,尽管他/她的“睡眠”活动的持续时间被设置为灵活的。本文研究中的操作级决策模块是通过扫描预测决策过程来解决这一时间约束的。

当智能体收到其活动模式中未预先安排的社交活动邀请时,它可以自主决定加入社交活动或拒绝。这是本文研究设计的智能体在一天内重新规划解决社交互动问题的能力,智能体如何回应这些邀请被定义为战术层面的决策。为了实现这一功能,智能体需要与功能实体进行大量的协调。

此外,为了增加智能体的异质性,还必须考虑智能体的战略决策。战略层面的决策如智能体在违反规则或遇到某些紧急事件(如流行病爆发期间的强制性政策干预)时应考虑替换原本遵循的活动模式。这种能力是通过社会规则的调节来实现的。

使用3级决策并不是基于智能体建模的新思路。然而,要建立一个包含所有3个决策层次的综合智能体模型非常困难,尤其在战略层面,这可以通过进一步的研究来逐步实现。

3.3 智能体环境-物理容器

本文研究将智能体环境的概念分为物理容器、社会监管和功能实体。物理容器表示智能体所在的物理环境。典型的物理容器有学校、教室、办公室、卧室和火车等。每个物理容器属性包含ID、类型、GIS信息(经纬度)、总面积(平方米)、当前状态(如已满、已关闭等)、子物理容器和当前容器内智能体的列表。

类型属性用于对物理容器进行分类。在传染病模型中,传染病的类型可以是家庭、工作场所和医院。在交通模型中,交通类型可以是私家车、公共汽车、火车和公路车道。GIS信息提供了地图中物理容器的准确位置。

物理容器的总面积属性在这个定义中是唯一的,并且会用于生成物理子容器(如学校的教室)。

物理容器也是分层组织的。通过为每个物理容器提供物理子容器的属性列表,可以将每个物理容器划分为多个物理子容器,如学校的教室、购物中心的商店或工作场所的办公室。当智能体处于不同的物理容器层次结构时,它们就可以有不同的接触者。

每个物理容器都分配了一个状态变量,用于表明智能体是否可以进入或离开该容器。所有在当前物理容器中的智能体可以由变量智能体列表检索到。

以图4为例,早上需要送孩子上学的家长,其起床时间的计算应该以他/她和孩子上班/上学都不能迟到为前提约束条件。

除了基本属性之外,物理容器也可以像智能体一样对外显示出“行为”。其中一个重要的行为“计算距离”主要根据GIS坐标信息(纬度和经度)来计算两个物理容器之间的实际距离。由于计算两个物理容器之间距离的过程是许多社会系统不可或缺的步骤,这种“行为”在仿真运行期间比较常见。

其他常见的“行为”包括“查找最近的物理容器”和“查找一定距离内的物理容器”。第一个“行为”给出最近的物理容器,第二个“行为”给出与任何指定物理容器指定距离内的物理容器列表。智能体在仿真过程中会频繁询问这两种“行为”,因为人们在没有特别偏好的情况下,更愿意在附近进行某些活动,如购物、吃饭和休闲。

此外,我们特意设计了一个为大规模系统计算的三级缓存机制。该机制创造性地实现了CPU利用率和内存使用之间的平衡。第一个缓存是最近缓存,它将距当前物理容器最近的物理容器存储到缓存中,新的对象只有在首次计算之后才会被添加到此缓存中。第二个缓存是网格缓存,可以把整个地图划分为网格,并在网格中保留物理容器的索引(类似于二叉树)。第三个缓存是距离缓存,在最近缓存或网格缓存中找不到结果时使用;对于任何特定的物理容器,此缓存可以按距离排列附近的物理容器。

物理容器的概念与一般智能体环境概念的区别和分离使得在社会仿真模型中包含交通组件模块变得更加容易,这是通过将车辆视为模型中可移动的物理容器来实现的。我们在基于智能体的大规模社会仿真概念模型中引入了可移动的物理容器的概念。为了使这一概念切实可行,我们提出并在大规模社会仿真中实现了一个微观公共交通系统(包括地铁和公共汽车)的解决方案^[10]。

3.4 智能体环境-社会监管

在真实的人类社会中,社会法规通过限制直接威胁公共健康和安全福利的行为来保护公共利益。特别是在卫生保健领域,了解人类疾病和社会环境之间的复杂相互作用引起了很多学者的研究兴趣。

在ABM相关文献中,已经有一些研究者试图模拟社会规则的部分功能。其中一个例子是DIVAs^[11]平台,用户可以在运行时创建影响环境的事件,环境可以使用组合函数对智能体提出规则和约束。Okuyama等^[12]提出了一种集成环境和组织建模的方法,利用由规范对象和规范场所组成的规范基础设施在社会环境中分发规范信息。然而,这两种实践都

没有明确地将组织和规范与它们所处的环境模型联系起来。

另一方面,在多智能体系统中的典型方法是,通过对规范和制度等实体的定义,可以在个人层面上查看和遵守社会监管。同样,本文研究将社会监管的概念从一般智能体环境概念中分离出来,使得基于智能体的大规模规范社会仿真成为可能。

在基于智能体的大规模社会仿真中,一个参考性实现是,当目标智能体不符合当前任何标准时,切换智能体的活动模式。通过此过程,智能体可以在仿真运行期间响应不同的情况。例如,在疾病爆发过程中调节智能体的行为是建立大规模基于智能体的流行病预测模型必不可少的一部分。智能体如何应对疾病爆发是一个轻量级的战略决策过程,因为决策结果将对智能体的行为产生较大影响。

图 5 给出了智能体战略决策模块和社会规则之间的协调过程。

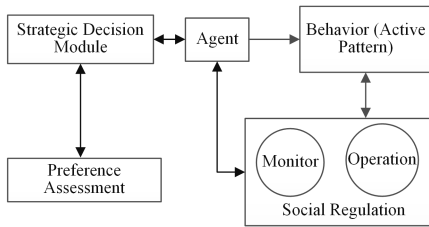


图 5 智能体战略决策模块和社会规则之间的协调

Fig. 5 Coordination between agent's strategic decision module and social rules

图 5 中的工作过程细节如下。

(1) 社会监管使用监视器来观察智能体的状态或系统状态,收集结果并将结果与预定义的标准进行比较。

(2) 当某个智能体或整个系统的状态不符合任何标准时,将通知该智能体并使用两个步骤触发该过程。

(3) 在社会监管触发操作之前,第一步通过智能体的决策模块进行偏好评估,以计算偏好的可能性,并指明智能体是否需要通过转移到另一个活动模式来改变其行为。

以疾病爆发期间智能体的反应为例,偏好评估可能基于概率 $P(A)$,如式(1)所示:

$$P(A) = 1 - e^{-\lambda \cdot A_p \cdot A_s \cdot |A_a - \epsilon|} \quad (1)$$

其中, A_p 表示当前活动模式对智能体 A 的重要性程度, A_s 表示智能体健康状态的严重程度, A_a 表示智能体的年龄, ϵ 是该病患者的平均感染年龄, λ 是权重系数。计算结果越接近 1,智能体转移到临时活动模式的可能性就越大。

(4) 第二步发生在社会监管触发操作之后,这意味着智能体必须决定是否遵守下发的规则。决策模块将根据偏好计算出符合的可能性。

采用与上述相同的示例,该评估可基于下面等式中的概率 $P(A)$ 。与式(1)相比,增加的参数为 P_c ,表示法规的强制性程度。

$$P(A) = 1 - e^{-\lambda \cdot A_p \cdot A_s \cdot P_c \cdot |A_a - \epsilon|} \quad (2)$$

3.5 智能体环境-功能实体

在基于智能体的社会仿真模型中,除了代表人的智能体外,还有许多其他实体也与之相关。在流行病模型中,研究人员表明,气候条件极大地影响着疫苗的有效性,温度也显著地

影响了流行病的动态,流感病毒在不同温度和湿度条件下的传播能力也各不相同。因此,通常的方法是在智能体环境中将它们建模为内置组件或属性,但这将限制超大型模型的可伸缩性和性能。

因此,我们将这些实体从智能体环境概念中分离出来,并将之定义为功能实体,即系统中可以影响或直接改变智能体、物理容器或社会监管属性的额外对象。例如,将风暴建模为功能实体,以改变房间(物理容器)的温度。

3.6 社会网络

社会网络是由一组社会行为体(如个人或组织)以及这些行为体之间的一组二元关系组成的社会结构。社会网络及其分析是一个从社会心理学、社会学、统计学和图论等学科发展而来的交叉学科。对于社会科学家来说,社会网络理论是非常有益的,它从心理学到经济学的多个学科角度解释了社会现象。

社会网络通常包含 3 种类型的关系:家庭、同事/同学和朋友。

家庭、同事和同学的关系可以很容易地从一个完整定义的拓扑结构中产生,该拓扑结构可以清楚地指定所有的关系连接。朋友作为最复杂的社会关系,相对难以界定。由于友谊关系的动态性,朋友关系的拓扑结构会随着时间的推移发生变化,且还会随着规模的提升变得更加复杂。

Wang 等^[13]认为,在社会网络中,人们的连接意愿不仅受到人格魅力的影响,还受到其偏好相似程度的强烈影响。他们提出,人在形成社会网络的过程中不断地在外在偏好和内在偏好之间取得平衡。MOCA^[14]平台提供了一个允许智能体根据其个人和集体的重复行为模式(分别称为角色和组织)来动态选择创建、加入或退出社会网络的理论模型。

在此研究的基础上,本文基于社交圈模型中的社会可达性概念和两个智能体之间的聚合效用理论,提出“社会相似性”的概念来动态地生成一种特殊类型的社会网络朋友圈。朋友的候选人有 3 种来源:邻居、同学/同事和随机选择。当智能体 A 计划社交活动时,生成好友的算法可以描述如下。

首先, A 的朋友数 N_f 服从幂律分布。

然后,由于 A 的朋友的来源组成可能不同于另一个智能体,根据均匀分布的组合来计算 A 的朋友不同来源的百分比。例如,智能体 A 可能喜欢与邻居交朋友,而智能体 B 可能喜欢在商店或餐馆之类的地方随机结交新朋友。

最后,随机选择一个候选人,计算候选人和智能体 A 是朋友的可能性。如果计算结果超过预定义的阈值(例如,初始设置为 0.25),则将候选人放入智能体 A 的朋友列表中;否则,选择一个新的候选对象并重复计算过程,直到生成 A 的所有好友。如果新的好友列表仍未满,可以减少阈值并再次重复计算过程。计算过程主要基于本文研究提出的“社会相似性”概念,计算两个智能体之间的相似性。所考虑的变量包括年龄、社会角色、家庭角色和朋友数量。在本文研究中,两个智能体 A 和 B 之间的相似性 $S(A, B)$ 通过加权欧氏距离进行评估,如式(3)所示:

$$S(A, B) = 1 - \sqrt{\sum_{i=a, s, f, n} \mu_i (A_i - B_i)^2} \quad (3)$$

其中, A 表示年龄, S 表示社会角色(转换为指数), f 表示家庭角色, n 表示智能体的朋友圈大小, μ 表示不同变量的权重。

以内)工作和活动的人;3) 活动范围平均离家 7 km 的人;4) 活动范围平均离家 10 km 的人;5) 活动距离 15 km 以上的人。在此基础上,本文首先将第 3)、第 4) 和第 5) 种类型相结合,然后分为公交换乘和私家车出行两种新的类型。私家车通勤类型人群又分为两种,即每天要拼车送孩子上学的人和无此需要的人。对于周末休息的上班族则设计了 4 种模式:1) 周末在家中活动的人;2) 在附近(3 km 以内)活动的人;3) 乘坐公共交通超过 3 km 活动的人;4) 驾车超过 3 km 活动的人。

对于退休人员,我们为他们设计了两种日模式。第一种类型喜欢待在家里,另一种则喜欢在外活动,这两种类型在工作日和周末都没有差异。对于学生,由于数据匮乏,根据学生上学的方式,设计了 3 种上学日模式。周末,设计了 4 种类似于上班族的日模式。由于学生的通勤方式与学校距离和父母出行的模式高度相关,因此,我们通过模型确定学生的模式比例分配。对于婴儿,我们假设只有一个典型的一日模式,即与他们的父母在家。由于该模型用于流行病预测,特此设计了一种住院人员日模式。

智能体在执行任意活动时都需要明确活动位置,在本文中研究中活动地点被建模为一个物理容器。这意味着智能体在仿真运行中的任何时候都有一个关联的物理容器,无论是在静止的地方执行活动(如在餐馆吃饭)还是在出行期间(步行或乘坐公共汽车)。物理容器是可以定位的,任意时刻物理容器或智能体都有一个地理参考坐标(经度和纬度),以便对其进行定位。这个描述对建模所需数据的完整性和一致性提出了严格要求。

利用统计信息生成的数据,我们在北京城市模型中对学校、餐馆、商店、医院等 800 万个静态物理容器进行了建模。目前静态物理容器共有 18 种,分为 6 类。显然,这些并不能涵盖北京所有的场所类型,例如,目前数据库中缺少属于“消费场所”的小商店和属于“娱乐场所”的电影院。下一步研究将考虑如何生成或收集这些缺失类型的真实数据,因为这些数据对疾病传播也很重要。

除了静态物理容器外,我们还建立了一个包含可移动物理容器的运输系统来执行智能体的出行活动。该公共交通系统是微观的,包含北京地铁和公交系统的所有线路和站点。截至 2015 年,北京的公共交通包括 17 条地铁线路、227 个地铁站和近 1000 条公交和无轨电车线路,是世界上最大的公共交通系统之一。

交通基础设施网络的地理信息是使用 Java 库 osmosis 从开源地区 OpenStreetMap 中获取的,同时利用 Java 库 jgrapht 建立了北京市公共交通网络的拓扑结构图。该图将公交地铁站点建模为节点(静态物理容器),将道路(地铁线路)建模为链接。对于道路上的车辆(汽车和出租车),我们没有对真实的道路进行建模,而是根据距离和拥堵的历史统计数据来计算预估的通行时间。

在本案例中,疾病爆发期间的政策干预措施被建模为模型的社会监管组件。如某政策描述为“当观察到‘感染’的智能体数量达到阈值时,处于‘有症状’阶段的智能体必须待在家中隔离。如果违反,则将被迫在医院隔离”。

本模型动态生成了 3 种类型的社会网络,即家庭、同事/

同学和朋友。家庭和同事/同学通过检查生活和工作在同一个物理容器中的其他智能体就能直接生成,朋友则根据之前介绍的算法生成。除了智能体之间规律性的交互外,本模型中的智能体还能够进行可协商的社交活动。这一能力是通过社交网络的动态构建和社交活动的动态执行来实现的。在执行临时社交活动时,模型会产生一个称为“群活动”的功能实体来组织和管理参与者。

除了“群活动”是功能实体外,病毒本身也被建模为功能实体,它可以改变智能体的健康状态。根据 Stroud 等^[18]的研究,我们在他们的疾病 H1N1 转移模型之上增加了一个“接种”阶段,可用于防疫政策建模。

5 仿真研究

针对现有的研究实践在平衡模型复杂度和仿真性能方面存在一定的困难,本文研究提出了一种基于智能体和大数据驱动的超大规模社会性突发事件应急建模与仿真推演概念模型框架,提供了模型组件的参考实现。为了检验该设计,我们主导了两个仿真实验:一是仿真系统运行性能,二是智能体的复杂行为能力(以动态社交活动为例)。

5.1 系统运行性能

基于智能体的大规模社会仿真中,对具有决策能力的智能体建模的一个挑战是如何保证系统的性能,因为复杂推理和决策过程中的通信信息数不胜数,而智能体大量的局部论证知识将大大降低仿真的效率和性能。

当前大规模 ABS 研究中仿真运行阶段的典型解决方案是引入通信中间件(如 MPI 或 RTI)和高性能计算,将模型划分为多个部分并在并行和分布式环境中运行不同的部分。

作为一种不同的科研尝试,本文研究尝试从模型设计和实现的角度出发,采用基于活动模式的智能体日常行为设计和多层次的决策机制来获得合理的系统性能。为了测试本设计的性能,我们进行了 3 组仿真实验,并记录了在不同情况下当智能体数量增加时实验的执行时间,结果如图 7 所示。仿真模型及实验运行在基于 Java 的仿真器 DSOL^[19]上,硬件环境为 1 台高性能笔记本电脑(Intel Core i7 10th, 16 GB RAM, 500 GB SSD)。

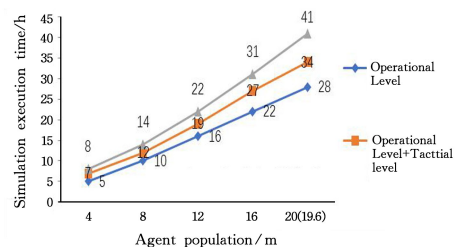


图 7 不同数量智能体决策能力的仿真性能对比
Fig. 7 Simulation performance comparison of decision-making ability of different number of agents

从图 7 可以看出,如果智能体只包含操作级决策过程,那么执行时间基本上是线性的。当在仿真中向智能体添加更高级别的决策过程时,执行时间略有增加。当所有具有完整决策能力的智能体(1960 万)都被仿真时(这是最极端的情况),一次仿真运行大约需要 40 h。我们认为这依然是可接受的性能,因为这个仿真时长为 30 天。

5.2 智能体行为能力

本文研究中,复杂智能体的代表性行为之一是能够基于动态生成的社会网络对非计划的社交活动作出合理反应。为了验证该设计对一般社会系统的影响,我们对大规模人工城市模型(无疾病爆发)进行了仿真,并进行了 30 天的仿真运行。

表 1 列出了智能体对接收到新的家庭社交活动后的决定的平均分布情况。

表 1 智能体参加家庭社会活动的决策分布

Table 1 Decision distribution of agents participating in family social activities

Decisions	Equation based	Decision Tree based
	Process	Process
Accept	0.78	0.67
Decline	0.22	0.33

从表 1 可以看出,33%的智能体在经过决策树过程后决定拒绝参加。

与表 1 类似,表 2 列出了智能体对接受到的同事/同学社交活动的决策的平均分布。两组数字之间最大的区别是,更多的智能体愿意参与同事/同学的社交活动,而不是家庭社交活动。这是因为同事/同学的社交活动通常安排在与智能体原先的计划没有重大冲突的时间。

表 2 智能体对同事/同学社交活动的决策的平均分布

Table 2 Average distribution of agents' decisions on social activities of colleagues and classmates

Decisions	Equation based	Decision Tree based
	Process	Process
Accept	0.88	0.75
Decline	0.12	0.25

表 3 列出了智能体对临时朋友社交活动的决定的平均分布。与上述两种情况相比,该分布的不寻常之处在于,参与的智能体数较少。这表明朋友圈网络中成员在日常日程安排上往往都是有冲突的。

表 3 智能体对临时朋友社交活动的决定的平均分布

Table 3 Average distribution of agents' decisions on temporary friends' social activities

Decisions	Equation based	Decision Tree based
	Process	Process
Accept	0.67	0.54
Decline	0.33	0.46

图 8 给出了一天不同社交活动中的智能体数量的演变情况。对于家庭社交活动,图 8 中有 3 个高峰,晚上为最高点,这表明人们更愿意与家人一起计划活动。对于同事或学生,上午和下午有 2 个高峰。这是工作人员往往需要在上午和下午参加集体会议、而学生在下午参加集体体育活动的模式造成的。对于朋友社交活动,看上去大多数朋友只会晚上见面,一起吃饭、逛街或去电影院。但是,由于目前还没有真实的数据,这些结果在本文研究中还不能得到很好的验证。

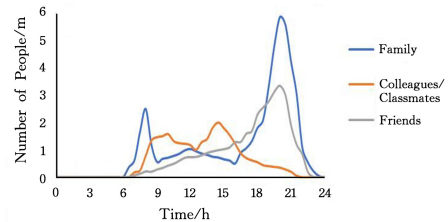


图 8 不同类型社交活动中的智能体数量

Fig. 8 Number of agents in different types of social activities

结束语 超大规模基于智能体的社会仿真正逐渐被证明是研究人类社会的一种有效方法,它可以为社会科学中的决策、计算机科学中的分布式人工智能和智能体技术、计算机仿真系统的理论和建模实践等领域作出贡献。同时,基于智能体的大规模社会仿真的应用领域也各不相同,从国家/城市层面的日常交通,到大规模的突发事件应急响应,再到社会演化预测和社会结构分析等。

然而,现有的研究实践在平衡模型复杂度和仿真性能方面存在一定的困难。一方面,分布式/并行机制在当前大规模基于智能体的社会仿真中的广泛应用,已被证明是实现系统性能和可扩展性的有效解决方案,但另一方面,模型中包括智能体行为、智能体环境以及社会网络和交互的严重简化。

针对目前存在的问题,本文提出了一种基于智能体和大数据驱动的超大规模社会仿真概念模型框架,提供了模型组件的参考实现,并以北京市疫情预测与控制为例进行了仿真研究。这个概念模型可以看作是一个混合模型,混合了基于一般智能体的概念模型和离散事件仿真范式。

针对概念模型中智能体的概念,本文提出了一种新的实现方法。智能体的参考实现由 3 个主要部分组成:1)智能体对象;2)活动模式;3)多级决策模块。通过这种设计,实现的智能体可以执行许多复杂的活动,并表现出不同的行为,同时又保持足够的“简单”和“小”以适应模型尺度的可伸缩性。

针对社会网络的概念,本文提出了一种新的动态生成社会网络的方法,用于在仿真运行期间模拟智能体之间的交互。本文借鉴了社交圈模型中“社会接触”的概念,提出了“社会相似性”的概念来生成一种特殊类型的社会网络,即朋友圈。利用生成的社会网络,模型中的智能体能够进行协商参与临时的社交活动。在执行临时社交活动时,模型会生成一个称为“活动群组”的功能实体来组织和管理参与者。

与一般 ABM 概念模型中智能体环境的概念相比,本文研究中的概念模型将智能体环境的概念分为物理容器、社会监管和功能实体 3 个部分,它克服了其他 ABM 模型对环境完备性的限制,并提供了模拟不同系统场景的灵活性。

利用这些概念的参考实现,我们构建了一个大型的人工北京的模型作为案例研究,以检验疾病在全部人口(1960 万)中传播的态势。这个案例研究可以作为一个例子来说明如何利用所提概念模型对具有复杂人类行为和社会交互的大型社会系统进行建模。该模型还显示了在其他社会科学领域的潜在应用领域,如微观交通系统分析和城市级疏散方案评估等。

参考文献

[1] EPSTEIN J M, AXTELL R L. Growing Artificial Societies: So-

- cial Science from the Bottom Up [M]. The Brookings Institution, 1996.
- [2] TATAPUDI H, DAS R, DAS T K. Impact assessment of full and partial stay-at-home orders, face mask usage, and contact tracing: An agent-based simulation study of COVID-19 for an urban region [J]. *Global Epidemiology*, 2020, 2: 100036.
- [3] MONTAGUD A, PONCE-DE-LEON M, VALENCIA A. Systems biology at the giga-scale: Large multiscale models of complex, heterogeneous multicellular systems [J]. *Current Opinion in Systems Biology*, 2021, 28: 100385.
- [4] FAZIO M, PLUCHINO A, INTURRI G, et al. Agent-Based Modelling of Mobility Restrictions at a Large Scale: Exploring the Impact on the COVID-19 Spreading in Italy [J]. *Journal of Transport & Health*, 2021, 22: 101195.
- [5] SIERHUIS M. *Brahms: A multi-agent modeling and simulation language for work system analysis and design* [D]. Amsterdam: University of Amsterdam, 2001.
- [6] COLLIER N, NORTH M. Parallel agent-based simulation with Repast for High Performance Computing [J]. *Simulation*, 2012, 89(10): 1215-1235.
- [7] CORDASCO G, SCARANO V, SPAGNUOLO C. Distributed MASON: A scalable distributed multi-agent simulation environment [J]. *Simulation Modelling Practice and Theory*, 2018, 89: 15-34.
- [8] MINSON R, THEODOROPOULOS G K. Distributing RePast agent-based simulations with HLA [J]. *Concurrency Computation Practice and Experience*, 2008, 20(10): 1225-1256.
- [9] MOSSONG J, HENS N, JIT M, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases [J]. *PLoS Medicine*, 2008, 5(3): 381-391.
- [10] ZHANG M X, MENG R Q, VERBRAECK A. Including public transportation into a large-scale agent-based model for epidemic prediction and control [C] // *Proceedings of the Conference on Summer Computer Simulation*, 2015: 1-8.
- [11] MILI R Z, STEINER R, OLADIMEJI E. DIVAs: illustrating an abstract architecture for agent-environment simulation systems [J]. *Multiagent and Grid Systems-An International Journal*, 2006, 2: 505-525.
- [12] OKUYAMA F Y, BORDINI R H, COSTA A C R. A distributed normative infrastructure for situated multi-agent organisations [M] // *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Heidelberg: Springer, 2009: 29-46.
- [13] WANG X, COLLINS A. Popularity or proclivity? Revisiting agent heterogeneity in network formation [C] // *Proceedings of the 2014 Winter Simulation Conference*. IEEE, 2014.
- [14] AMIGUET M, MULLER J P, BAEZ-BARRANCO J A, et al. The MOCA platform - Simulating the dynamics of social networks [C] // *Proceedings of the 3rd International Conference on Multi-Agent-Based Simulation II*. Springer-Verlag, 2002.
- [15] ALLEN L J S. Some discrete-time SI, SIR, and SIS epidemic models [J]. *Mathematical Biosciences*, 1994, 124(1): 83-105.
- [16] GE Y, MENG R, CAO Z, et al. Virtual city: An individual-based digital environment for human mobility and interactive behavior [J]. *Simulation*, 2014, 90(8): 917-935.
- [17] TA N, CHAI Y, GUAN M. Suburbanization, daily lifestyle and space-behavior interaction in Beijing (In Chinese) [J]. *Acta Geographica Sinica*, 2015, 70(8): 1271-1280.
- [18] STROUD P, VALLE S D. Spatial dynamics of pandemic influenza in a massive artificial society [J]. *Journal of Artificial Societies and Social Simulation*, 2007, 10(4): 1-18.
- [19] JACOBS P H M, LANG N A, VERBRAECK A. D-SOL: A Distributed JAVA based Discrete Event Simulation Architecture [C] // *Proceedings of the 2002 Winter Simulation Conference*. IEEE, 2002.



ZHANG Ming-xin, born in 1986, Ph.D., associate professor. His main research interests include modeling and simulation of complex social system.

(责任编辑:喻黎)