

计算实验研究方法及应用

崔凯楠^{1,2} 郑晓龙^{2,3} 文丁⁴ 赵学亮²

摘 要 计算实验是一种研究复杂系统的新兴计算方法, 受到了国内外学者的广泛关注. 近年来, 随着相关研究的不断发展, 计算实验方法在多个领域显示出巨大的应用前景, 特别是复杂系统管理与控制相关的诸多重要领域, 如社会安全、电子商务、金融市场等. 本文将首先介绍计算实验的主要思想以及实验设计与计算机仿真等计算实验的研究基础; 其次, 我们将介绍计算实验的主要研究内容, 包括计算模型构建、计算实验设计以及计算实验执行; 最后, 我们将讨论计算实验的应用情况, 总结研究过程中面临的挑战并介绍潜在的研究方向.

关键词 计算实验, 人工社会, 基于代理的建模, 计算机仿真

引用格式 崔凯楠, 郑晓龙, 文丁, 赵学亮. 计算实验研究方法及应用. 自动化学报, 2013, **39**(8): 1157–1169

DOI 10.3724/SP.J.1004.2013.01157

Researches and Applications of Computational Experiments

CUI Kai-Nan^{1,2} ZHENG Xiao-Long^{2,3} WEN Ding⁴ ZHAO Xue-Liang²

Abstract Computational experiments are emerging computational methods for complex system research, which have drawn wide attentions from both domestic and foreign scholars. The rapid growth of computational experiments shows great application potential in research domains related to control and management of complex systems, such as social security, e-commerce, financial markets and so on. In this paper, we will firstly introduce the basic idea of computational experiments and review the research foundations of computational experiments including experimental design and computational simulation. Then we will mainly focus on presenting the main research contents of computational experiments including construction of models, design of computational experiments, and execution of computational experiments. Finally, we will discuss the applications and the existing problems in this field and present several potential research directions.

Key words Computational experiments, artificial societies, agent-based modeling, computer simulation

Citation Cui Kai-Nan, Zheng Xiao-Long, Wen Ding, Zhao Xue-Liang. Researches and applications of computational experiments. *Acta Automatica Sinica*, 2013, **39**(8): 1157–1169

计算实验是一种研究复杂系统的新兴计算方法,

近年来受到国内外学者的广泛关注, 产生了诸多新的应用. 计算实验方法是计算机仿真的天然扩展, 其核心思想是把计算机作为“生长培育”自然实际系统替代版本的“实验室”^[1–2]. 计算实验不要求计算模型完全再现实际系统的行为, 而把计算模型作为一种可能的存在.

在科学问题研究与探索过程中, 实验发挥着不可替代的重要作用. 为了区别起见, 本文中将在物理空间中进行的实验称为实物实验. 实物实验是理论与事实的桥梁, 它不仅促进着理论到技术方法的转换, 也使得理论研究更加具有目的性. 但是, 随着复杂系统研究的兴起, 传统的实物实验方法面临着种种挑战. 由于研究问题的复杂本质以及经济、法律、道德等多方面因素, 实物实验在许多涉及复杂系统的研究领域难以顺利开展^[3]. 在这种背景下, 计算机仿真方法受到了越来越多的重视, 然而复杂系统不断发展的特性使得计算机仿真方法面临着无真可仿的难题, 因此对计算机仿真方法进行扩展的需求

收稿日期 2012-12-27 录用日期 2013-04-15
Manuscript received December 27, 2012; accepted April 15, 2013

国家自然科学基金 (71103180, 91124001, 71025001, 91024030), 国家科技重大专项基金 (2012ZX10004801, 2013ZX10004218) 资助
Supported by National Natural Science Foundation of China (71103180, 91124001, 71025001, 91024030), and National Science and Technology Major Project (2012ZX10004801, 2013ZX10004218)

本文责任编辑 王占山

Recommended by WANG Zhan-Shan

1. 西安交通大学电子与信息工程学院 西安 710049 2. 中国科学院自动化研究所复杂系统管理与控制国家重点实验室 北京 100190 3. 中国科学院云计算产业技术创新与育成中心 东莞 523808 4. 国防科学技术大学军事计算实验与并行系统技术研究中心 长沙 410073

1. School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049 2. The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences (CASIA), Beijing 100190 3. Cloud Computing Industrial Technology Innovation and Incubation Center, Chinese Academy of Sciences, Dongguan 523808 4. Research Center for Military Computational Experiments and Parallel Systems Technology, National University of Defense Technology, Changsha 410073

日趋迫切.

本文将针对计算实验研究现状及相关应用进行阐述. 后续篇幅安排如下: 第 1 节将介绍计算实验的研究基础, 主要包括实物实验与计算机仿真; 第 2~4 节将从计算模型构建、计算实验设计、计算实验执行等三方面展示计算实验的主要研究内容与研究现状; 第 5 节将讨论计算实验应用的特点并重点阐述交通领域中的应用; 最后, 第 6 节将对全文进行总结.

1 计算实验研究基础

计算实验的研究基础包括实物实验与计算机仿真. 本节将分别介绍实物实验与计算机仿真的主要研究内容, 讨论实物实验与计算实验、计算机仿真与计算实验之间的区别.

1.1 实物实验

实物实验是指在物理空间中进行的实验, 通常在实验室、工厂或农场里进行. 实验设计 (Design of experiments) 是研究如何科学安排实验的学科, 经典的实验设计起源于 20 世纪 20 年代, 主要应用于实物实验, 下文称实物实验设计. 计算实验则将计算模型视为“生长培育”实际系统替代版本的“实验室”, 研究人员通过分析计算模型开展实验. 本文中面向计算实验的实验设计方法称为计算实验设计. 实物实验设计是计算实验设计的基础. 我们可将实物实验抽象为如下形式:

$$\mathbf{y} = f(\mathbf{x}) + \boldsymbol{\varepsilon} \quad (1)$$

其中, $\mathbf{y}$ 是实验的观测结果, 下文中将其称为响应. $f(\cdot)$ 表示实物实验过程隐含的数学关系, $\boldsymbol{\varepsilon}$ 表示实验误差. 实物实验设计的主要研究内容包括元模型选择和实验取值策略, 由于 $f(\cdot)$ 通常是未知的, 我们可以使用已知的数学函数 $g(\cdot)$ 对 $f(\cdot)$ 进行拟合, 该函数称为元模型, 元模型是一个通称, 泛指代表实验过程的数学抽象表示, 因此元模型并不局限于线性模型或统计模型等特定模型. 由于不同实验间的差异, 元模型的选择范围十分广泛, 而且随着其发展不断变化. 实验的输入组合由实验取值策略决定, 实验取值策略与元模型之间密切相关. 实物实验设计的发展大致可以分为三个阶段^[4].

1) 农业实验. 实物实验设计始于 Fisher 将析因设计和方差分析等方法应用到农业实验中^[5]. 农业实验中假设实验观测数据相互独立且服从正态分布, 主要选择线性统计模型为元模型, 例如效应模型 (Effect model):

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \quad (2)$$

其中, $i = 1, \dots, a$; $j = 1, \dots, n$. y_{ij} 为实验的响应, μ 为总体均值, τ_i 为第 i 个因子的效应, ε_{ij} 是代表误差的随机变量. 针对上述模型, 可以使用完全随机设计 (Completely randomized design) 作为实验取值策略, 并使用单因素方差分析处理所得数据.

2) 工业实验. 由于工业实验的立即性与连续性^[6], Box^[7] 等针对工业实验的两个特性提出了响应曲面法 (Response surface methodology, RSM). 响应曲面法中通常使用二阶多项式等更加复杂的元模型, 并使用最优化的方法分析元模型的响应曲面.

3) 健壮设计. 由于在制造业中需要保证工业过程的稳定性, Taguchi 等^[8-11] 提出了健壮设计的概念. 健壮设计是在优化响应变量的同时降低其差异性. 虽然健壮设计在工业界得到了认可, 但是有许多统计学者认为田口玄一的工程概念和目标缺乏理论支撑.

1.2 计算机仿真

广义而言, 计算机仿真属于计算实验的范畴, 计算机仿真主要面向简单系统, 而计算实验研究不仅面向简单系统, 也以复杂系统作为研究目标. 因而计算实验是计算机仿真的自然扩展, 可以理解为计算机仿真的超集. 计算机仿真是一个成熟的研究领域, 存在大量的研究工作与成果^[12-15]. 最早的计算机仿真应用可追溯到 20 世纪 40 年代的曼哈顿计划. 由于计算机仿真与计算实验之间的紧密关系, 计算机仿真与计算实验的主要研究问题之间具有很强的关联性, 本小节将介绍计算机仿真的主要研究内容.

1.2.1 仿真模型构建

本文将计算机仿真中的模型建立过程称为仿真模型构建. 仿真模型存在多种不同的分类方法, 如果按照模型的随机性划分, 则可分为确定性仿真和随机性仿真. 如果以计算模型的时间流机制划分, 可将仿真模型分为连续仿真模型和离散仿真模型^[16-17]. 若按照仿真目的划分, 可将计算机仿真分为分析仿真和虚拟环境两类, 分析仿真通常试图获取被仿真系统的详细定量数据. 虚拟环境仿真的目标是给用户提供置身于被建模系统中的视觉和感觉^[18].

1.2.2 仿真模型验证

仿真模型验证是计算机仿真研究中的重要组成部分^[19-20]. 仿真模型验证的目的是使计算机模型尽可能逼近于现实中的系统. Sargent^[21] 讨论了仿真模型验证的研究范式, 并将仿真模型验证划分为数据有效性 (Data validity)、概念模型验证 (Conceptual model validation)、程序检验 (Computerized model verification) 以及操作有效性 (Operational

validity) 等几个主要问题. 读者可通过参考相关综述 [21–22] 了解仿真模型验证中的具体方法. 虽然这个领域中已经存在大量研究工作, 但是研究人员仍然质疑仿真模型验证的有效性^[23].

1.2.3 仿真优化分析

仿真优化^[24–25] 是基于计算机仿真的目标优化问题. 优化是指非枚举地从输入空间中找到最佳输入变量值^[26], 使得输出结果为最优解或满意解的过程. 如果按输入变量类型划分, 可将仿真优化方法分为针对连续输入变量的方法与针对离散输入变量的方法.

1) 连续输入变量. 基于梯度的优化方法是一种适用于连续型输入变量的方法, 算法的主要目标为寻找梯度的零值点. 这类方法需要估计响应函数的梯度 ∇f , Fu^[27] 总结了这种方法在仿真优化中的应用.

2) 离散输入变量. 除基于梯度的优化方法外, 文献 [28] 中总结了针对离散输入变量的仿真优化算法, 可行解数量较少的问题 (2~200 个) 主要使用基于统计的优化方法, 如排序和选择方法 (Ranking and selection) 以及多重比较方法 (Multiple comparison procedures). 关于选择排序和多重比较的详细介绍可参见综述 [29]. 当可行解的数量很大时 (一般指大于 200 个), 研究人员多采用随机搜索和启发式搜索方法.

1.3 计算实验的提出

随着问题复杂性的增加, 基于代理的建模方法因其如下优势: 1) 可以描述涌现现象; 2) 是对系统的一种自然描述; 3) 高度柔性的, 为复杂系统的建模提供了新的研究思路^[30], 极大地扩展了计算模型构建方法的自由度与灵活度^[31–33]. 在此基础上, 为了解决复杂系统的实验评估问题, 王飞跃^[2, 34–35] 于 2004 年正式提出计算实验方法 (Computational experiments), 其核心思想是将计算机视为“实验室”. 计算实验的主要研究工作包括计算模型构建、计算实验设计与计算实验执行. 此后一些研究工作中也表现出与计算实验非常类似的思想, 例如: 计算机实验 (Computer experiments)^[36–37]、仿真实验 (Simulation experiments)^[38].

实物实验设计方法与仿真模型构建方法为计算实验的研究奠定了坚实的基础. 然而实物实验与计算实验、计算机仿真与计算实验之间也存在着显著的区别, 本小节将总结这些区别.

1.3.1 计算实验与实物实验的区别

1) 计算实验的因子数量巨大. 实物实验中的因

子通常少于 10 个, 很少存在同时研究成大量因子的实物实验. 相比于实物实验, 计算实验通常含有更多的考察因子^[24, 39].

2) 计算实验的目标多样化. 计算实验的目标多种多样, 包括获得对系统的初步认识、识别健壮方案、比较系统优劣等^[39]. 此外, 以复杂系统为研究对象的计算实验通常需要研究多个响应变量^[24], 而实物实验则多关注单响应变量.

3) 计算实验数据收集方式更加灵活. 一般而言, 实物实验的时间花销较大, 例如: 农业实验中需要一个作物生长周期才能完成一次数据收集. 因此实物实验数据收集多采用并行的方式完成. 计算实验的运行时间较短, 可以采用灵活的收集方式, 但是获取的数据也会表现出自相关性^[24].

4) 计算模型的其他影响因素. 相对于实物实验而言, 计算实验中不存在真正意义上的不可控因素^[24]. 此外, 实验的不同设置也会影响实验结果, 例如: 计算实验的运行时间选择^[40–41]、计算实验样本量选择^[42] 等.

除上述普遍的区别之外, 一些具体实物实验方法与计算实验方法的区别与联系也值得探讨, 例如在复杂社会系统或经济学研究中, “真人 (Human-subject) 实验” 有着广泛的应用. 真人实验中通过创造与实际相似的实验室环境并通过给予一定物质报酬, 使得预先挑选的受试对象按照一定规则完成实验^[43]. 真人实验属于实物实验的范畴, 以真人作为受试对象, 而计算实验中是以计算模型作为受试对象. 真人实验中计算机起辅助作用, 而计算实验中则将计算机视为实验室. 与计算实验相比, 真人实验的优势在于涉及人类行为的微观现象的研究与解释, 然而其实验规模受成本限制. 在复杂系统研究领域, 计算实验与真人实验都面临着实验是否能够反映现实的质疑.

1.3.2 计算实验与计算机仿真的区别

计算实验与计算机仿真的主要区别体现在计算机仿真主要面向简单系统, 而计算实验则同时面向简单系统与复杂系统^[3]. 简单系统具有良好的定义和明确的结构, 通常采用自顶向下的建模方法, 通过数学方程或者离散事件模型描述系统的行为. 这些建模方法的灵活度差, 难以满足复杂系统建模的要求^[2, 44]. 具体而言, 由于缺乏充分可用的理论和先验知识, 自顶向下的建模方法难于对复杂系统进行准确描述并深入分析系统行为.

不同于计算机仿真中要求仿真模型不断逼近实际系统, 计算实验不要求计算模型完全再现实际系统的行为, 然而如何证明计算模型的有效性与等价

性却没有取得共识^[45]. 一种较为流行解决方法是通过比较计算模型输出与实际系统输出间的统计特征证明计算模型的等价性^[46-47]. 另一种解决方法则放弃验证计算模型的有效性, 在假设计算模型无法验证的前提下进行分析, 尝试挖掘计算模型所隐含的潜在结论^[48-49]. 这种方法的应用较少, 但是其思路值得借鉴.

2 计算模型构建

在计算实验的研究中, 基于代理的建模方法是一种最流行的计算模型构建方法, 被广泛地应用于自然科学与社会科学中的诸多研究领域^[50-53]. 在自然科学领域中, 计算模型常用于描述分子的自组行为 (Self-assembly)、纳米材料的自组织过程以及复杂的生物系统等, 对于这些复杂系统的计算模型构建方法可参考相关综述^[51, 54], 本文主要总结了一些涉及人与社会复杂性的复杂系统模型. 在涉及人与社会复杂性的计算实验模型构建中, 主要包括对个体行为的建模、对组织结构的建模与对环境的建模三个部分, 本节将分别介绍这些建模方法.

2.1 个体行为建模

个体行为的建模方法可以大致分为基于概率的模型、基于规则的模型与基于博弈演化的模型. 也有一些模型同时具备多种类型的特点.

1) 基于概率的建模方法. 这类模型使用概率分布描述代理的决策与交互过程. 首先将代理划分为若干种不同的状态, 代理的行为最终表现为状态的迁移. 例如在经典传染病模型中, 个体被划分为以下几种状态: S (Susceptible, 易感状态)、I (Infected, 感染状态)、R (Removed, 被移除状态). 不同模型间的区别为状态集的不同. SI 模型中易感染状态被感染后不再恢复健康; SIR 模型中易感染状态被感染后会恢复健康并具有免疫性; 如果易感人群被感染又返回到易感染状态, 则称之为 SIS 模型. 独立级联模型 (Independent cascade model)^[55-56] 借鉴了 SI 模型的状态划分, 该模型认为代理 i 以概率 p_{ij} 来影响其邻居代理 j , 被影响的代理无法恢复为未影响状态.

2) 基于规则的建模方法. 该类建模方法范围十分广泛. 根据规则的复杂程度, 可将基于规则的建模方法分为基于逻辑推理的模型与基于简单规则的模型. 基于逻辑推理的模型有利于表达人类思维过程, 可以很好地实现基于代理心理属性、社会属性与行为规则的推理过程, 从而多被应用于实现代理间的沟通与多代理系统中的决策制定^[57]. 模态逻辑 (Modal logic) 是一种被广泛用于形式化知识、

信念、意图、目标等代理心理属性的建模方法^[58]. Kripke 结构为模态逻辑提供了良好的语义基础^[59]. 时态逻辑 (Temporal logic) 扩展了逻辑推理模型在计算环境中的表现能力, 是另一种被广泛应用的逻辑模型^[60], 如: BDI (Belief-desire-intent) 建模框架中代理的行为选择中就结合了模态逻辑与时态逻辑^[61].

虽然逻辑推理模型的建模表现能力强, 但是其复杂度高且计算负荷大. 许多研究问题并不需要如此复杂的模型, 一些简单规则就可以满足研究问题的要求, 例如信息扩散中的线性阈值模型^[62-63]. 该模型认为每个代理均存在一个阈值 $h_i \in [0, 1]$, 网络中的每条边 e_{ij} 都有一个权值 w_{ij} 并且满足 $\sum_{j \in N_i} w_{ij} = 1$, 其中 N_i 表示代理 i 的邻居节点. 当代理 i 满足如下条件时被影响, 其中 S 代表被影响的代理集合:

$$\sum_{j \in N_i, j \in S} w_{ji} \geq h_i \quad (3)$$

舆论演化研究中也存在大量基于简单规则的模型, 如 Sznajd 模型^[64-66] 中代理观察两个或多个邻居是否具有相同的意见, 若邻居具有相同的意见则改变自己的意见. 在 Deffuant^[67-68] 模型中当两者的意见差异小于阈值时才会发生意见交换.

3) 演化博弈建模^[69-70]. 采用演化博弈论的思想, 模型中通过多次博弈过程得到演化稳定策略. 每个时刻, 代理都会根据自己的总体收益改变自己的选择. 代理收益的形式化表示如下:

$$\sum_{j \in N_i} w_{ij} u(x_i, x_j) \quad (4)$$

其中, w_{ij} 代表边 e_{ij} 的权值, x_i 表示代理 i 的选择, N_i 代表代理 i 的邻居集合, $u(x_i, x_j)$ 表示收益函数. 对于两个邻接的代理, 相同选择的收益值大于不同选择的收益. 每一时刻代理都会根据收益最优化准则调整自己的选择.

除上述三种建模方法类型以外, 也有一些混合模型同时应用两种或三种方法, 例如: Voter 模型^[71-72] 中代理随机接受某一个邻居的意见, Majority rule 模型^[73] 中个体中接受随机小组中的多数意见等.

2.2 组织结构建模

早期的计算模型多建立在简单拓扑结构之上, 例如: 方形栅格、规则网络、完全网络等. 随着复杂网络研究的兴起, 基于复杂网络的社会网络建模可以更好地描述个体之间的相互影响^[74-75]. 例如 Sznajd 模型应用复杂网络模型之后, 更好地重现了

实证研究中的统计特性^[76]. 有些研究专注于网络拓扑结构对意见演化的影响, 网络的拓扑结构会显著的影响整个意见的演化过程^[77]. 例如: 网络的小世界特性加速了信息和意见在社会中的传播速度. 在文献 [78] 中发现 Voter 模型在小世界网络中产生了有趣的亚稳定状态. 在无标度网络中, 中心节点的意见可以决定网络能否达成共识. 此外, 网络的度分布直接影响达到共识的速度^[79].

2.3 环境建模

代理与环境的复杂交互是基于代理建模技术的重要组成部分, 我们可将计算模型中的代理和环境划分为设计型与分析型两种. 其中设计型基于研究人员的假设, 而分析型则基于实际数据. 基于上述划分, 可将基于代理的模型划分为四类^[80], 如表 1 所示.

随着地理信息系统 (Geographical information system, GIS) 的发展, 集成地理空间信息的计算实验逐渐成为研究的热点^[81-85]. 如何表示计算模型中的地理空间是该领域的重要研究问题. 如果从计算建模的角度出发, 表现地理空间的方式主要有按照单元格的划分方法 (Cell-based method) 以及按照地区的划分 (Zone-based method)^[80]. 也有研究人员采用两者的结合, 例如, 在文献 [81] 中先将地球表面划分为单元格网, 然后根据每个单元格与附近机场的距离, 将该单元格分配给最近的机场, 其划分情况如图 1 所示.

除地理空间信息外, 气候信息也是计算实验中一种重要的环境因素. 在涉及天气与气候信息的计算实验中, 往往集成已有的计算模型, 如 Nüssle

等^[86] 在计算模型中集成了火灾的模型, Tanigawa 等^[87] 在计算实验中加入洪水模型. Carley 等^[88] 则考虑了风对于疾病传播的影响.

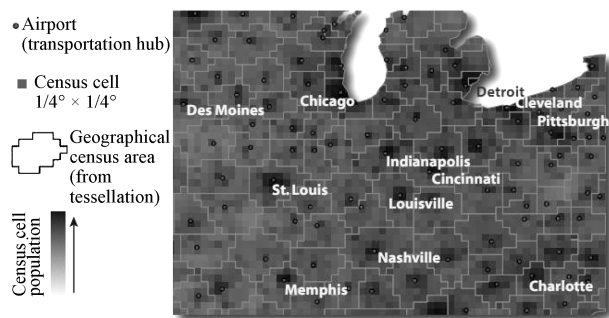


图 1 空间划分方法示意图^[81]

Fig. 1 Geospatial division method^[81]

3 计算实验设计

计算实验设计与实物实验设计具有相同的研究目标与研究框架, 主要包括元模型选取和实验取值策略. 然而由于实物实验与计算实验之间的区别, 许多实物实验中的假设与边界条件在计算实验中并不满足. 实物实验中通常假定误差是独立同分布的, 有些实验还认为实验误差满足正态分布, 然而计算实验中通常不能满足这些假定, 虽然使用不同的伪随机序列可以保证误差的独立性, 然而同分布这一性质仍然难于保证. 因此许多经典的实物实验设计方法不能直接应用于计算实验中. 例如: 部分析因设计^[89] 中需要假定因素间不存在交互效应或仅存在低阶交互效应, 然而对于复杂的计算模型而言, 这一

表 1 环境类型模型分类^[80]

Table 1 Model types based on environments^[80]

		代理类型	
		设计型	分析型
环境类型	研究类型	抽象研究	实验研究
	研究目的	关系类型发现; 理论验证	开展实验室实验
	验证策略	理论比较	检验设计完备度
	所需工具	易于实现与使用的计算模型	友好用户界面的计算模型
环境类型	研究类型	历史研究	实证研究
	研究目的	解释现象	解释、预测、场景分析
	研究策略	拟合度检验	拟合度检验
	所需工具	支持地理信息系统的计算模型	底层编程语言

假设很难满足. 由于计算模型表现的复杂非线性关系, 计算实验设计中需要表现能力更强的元模型与相对应的取值策略, 本节中将着重介绍计算实验中使用的元模型与取值策略.

3.1 计算实验中的元模型

元模型是对计算实验过程的数学抽象. 由于计算实验中的实验过程是由计算模型表示, 因此元模型是对计算模型的二次建模. 随着计算模型复杂程度的不断增加, 为了刻画输入因子与响应变量之间的复杂非线性关系, 计算实验设计中的元模型需要更高的数学表现能力. 一些其他研究领域中使用的空间统计模型及机器学习中的方法被应用于计算实验中. 这些模型在拥有更强数学表现能力的同时, 模型参数估计的难度增加, 对数据质量的要求也更高.

3.1.1 统计模型

1) Kriging 方法^[90-91]. 该方法也被称为空间相关建模, 可以用于处理复杂的响应曲面结构. 该模型假定多维输入空间中的点之间存在着某种空间相关性. 并利用这些相关性进行预测. 基本的 Kriging 模型可表示为如下形式:

$$w(d) = \mu + \delta(d) \quad (5)$$

其中, μ 为整个实验区域内的均值, $\delta(d)$ 为附加噪声, 从而构成了均值为 0 的协方差平稳过程. Kriging 模型的主要缺点为参数获取过程中需要巨大计算量, 并且假设的空间相关性难于验证.

2) 无参数回归模型^[37]. 一维无参数回归模型的形式化表示如下:

$$Y = m(X) + \varepsilon \quad (6)$$

其中, ε 是随机误差, $m(X) = E(Y|X)$ 被称为回归函数, 这里仅仅假设该函数为连续函数而无需引入其他参数, 相比于多项式回归, 无参数回归模型的优点在于其连续性.

3) 样条方法 (Spline method)^[92]. 其主要思想是通过一个逐步前向算法选择变量, 然后使用一个后向过程修剪模型. 该方法的思想是寻找模型的“节”点. 我们可以使用如下等式描述模型:

$$g(x) = \beta_0 + \sum_{m=1}^M \beta_m B_m(x) \quad (7)$$

其中被称为基础函数的 $B_m(x)$ 可表示为如下形式:

$$B_m(x) = \prod_{l=1}^{L_m} [S_{l,m} \cdot (x_v(l, m) - k_{l,m})] \quad (8)$$

其中, L_m 是第 m 个基础函数中修剪的线性函数数量, $x_v(l, m)$ 是第 l 个修剪函数中的输入变量, $k_{l,m}$ 是相应的节点值. $S_{l,m}$ 是 +1 或者 -1. 该方法比较灵活且易于实现, 但是该模型是线性模型, 无法对复杂度高的实验进行模拟.

4) 贝叶斯方法^[93]. 如果从贝叶斯理论的角度出发, 可以将未知的数学函数假设为下面所示的高斯过程:

$$Y = \{Y(x), x \in T\} \quad (9)$$

因此在得到响应观测 $(y_1, \dots, y_n)$ 后, 可对 $g(x)$ 进行预测. 贝叶斯方法的优点在于得到响应变量的同时可以获取一阶偏导数的信息.

3.1.2 机器学习模型

相比于统计方法, 机器学习方法对数据的要求更高. 主要的机器学习模型包括: 人工神经网络 (Artificial neural network)^[94] 和支持向量机 (Support vector machine, SVM)^[95].

1) 人工神经网络^[94]. 常用的前馈神经网络一般形式如下:

$$g(\mathbf{x}) = b_2 \left(\sum_{h=1}^H w_{hk} \cdot b_1 \left(\sum_{j=1}^p v_{jh} x_j + \theta_h \right) + \gamma_k \right) \quad (10)$$

其中, H 是隐藏节点的数量, p 是输入变量的数量. 神经网络是非线性统计模型, 主要优点为建模灵活度高, 但计算需求大, 而且容易出现过拟合现象 (Overfit the data), 此外大量很难解释没有物理意义的参数也是人工神经网络的一个缺点.

2) 支持向量机^[95]. 是另一种经典的机器学习模型. 以二分类问题为例, SVM 模型可以表现成如下形式:

$$h_{w,b}(\mathbf{x}) = k(w^T \mathbf{x} + b) \quad (11)$$

其中, $z \geq 0$ 时 $k(z) = 1$, 其他情况时 $k(z) = -1$.

3.2 计算实验取值策略

在实物实验设计文献中存在多种实验取值策略, 例如析因分析和部分析因分析、中心复合设计等^[89]. 析因分析易于理解和实施, 然而在含有大量因子的计算实验中, 所需的实验次数呈指数型增长, 无法广泛应用. 除实验次数这一指标外, 实验取值策略还有多个衡量指标^[36], 其中模型相关性 (Model free) 在计算实验中非常重要. 由于计算模型多描述复杂非线性关系, 许多实物实验中的经典假设并不成立. 模型无关的取值策略不依赖于模型的假设, 因此被广泛的应用于计算实验中. 此外, 由于计算实验执行灵活, 多阶段的实验取值策略也被应用到计算实验中,

本小节将重点介绍模型无关取值策略与多阶段实验取值策略。

3.2.1 模型无关取值策略

模型无关实验取值策略并不针对某一种特定的元模型, 而是关注由标准实验输入定义的 k 维单位立方体构成的设计空间。主要的模型无关设计包括超拉丁方设计 (Latin hypercube sampling)^[96] 与均匀设计 (Uniform design)^[97]。

1) 超拉丁方设计方法^[96, 98]是空间填充设计的一种类型, 该方法根据某种先验分布在输入空间上采样。超拉丁方设计关注的是一个由因子的各水平构成的正方形, 每个因子的不同水平在每一行和每一列仅出现一次。首先 x_j 被划分为 N 个等长区间, 然后这些区间被分配到一个 $N \times p$ 设计矩阵 X 中, 如图 2 所示为一个 4 维、9 水平的超拉丁方及其二维映射。

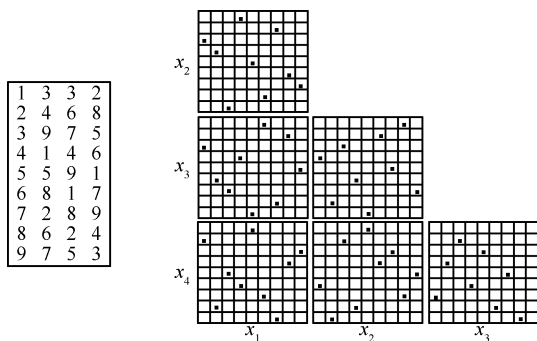


图 2 超拉丁方采样^[99]

Fig. 2 Latin hypercube samples methods^[99]

2) 均匀设计^[97, 100]是假设实验点在实验范围内均匀散布的一种多因素、多水平的实验设计方法, 在正交设计的基础上放弃正交表的整齐可比性, 进一步提高实验点的“均匀分散性”^[101]。其思想是从所有实验点中挑选出部分具有代表性的实验点, 使得 n 次实验所取的 n 个具有代表性的点在所考察的范围内尽可能地充分均衡分散而仍能反映主要特征, 并根据点在空间的散布程度提供均匀设计表。

均匀设计最大的优点是可以选取较大的因素水平数, 而实验次数又最节省, 这是绝大多数实验设计方法所不具备的。然而均匀设计只考虑实验点在实验范围内充分地均匀散布而不考虑整齐可比, 因此对实验结果进行统计分析比较复杂。

3.2.2 多阶段实验取值策略

相对于实物实验, 计算实验的执行更加灵活, 可根据已有的实验结果制定下一阶段的取值策略, 如顺序分支设计 (Sequential bifurcation)^[102-103] 和序

贯实验设计 (Sequential experimental design)^[99]。

1) 顺序分支设计^[102-103]。由于计算实验中通常涉及大量因子, 因此如何避免维度灾难, 寻找重要的因子是十分重要的。顺序分支方法通过将所有因子聚合为因子组, 检测因子组是否具有显著性, 并对具有显著性的因子组分支检验。基于 r 次重复的组效应简单估计可以表示为

$$B_{j'-j} = \frac{w_{(j);r} - w_{(j'-1);r}}{2} \quad (12)$$

其中, $w_{(j);r}$ 表示第 r 次重复中将因子 $1 \sim j$ 置于高水平而其他因子置于低水平所获得的响应值。

2) 序贯实验设计^[99]。在该设计中, 实验因子水平选择是根据以往水平选择与相应输出决定的, 即 x_{N+1} 是由 $(x_1, \dots, x_N)$ 与对应的 $(y_1, \dots, y_N)$ 决定。如下所示为随机近似方法的形式化表示:

$$x_{N+1} = x_N - a_N(y_N - \alpha) \quad (13)$$

其中, $\{a_N\}$ 是预设定的常量序列, α 为预先设定的常量。

4 计算实验执行

随着计算实验规模的不断增大, 对计算实验执行的要求不断增高。计算实验执行过程中的主要计算方法包括: 计算实验的自动化执行, 计算实验的高性能计算与计算实验的平行执行。本节中将分别介绍计算实验执行中的主要计算方法。

1) 计算实验的自动化执行。Fomel 等^[104]强调计算实验的可重复性。作者介绍了一个帮助计算实验重复执行的软件包, 然而该研究只考虑了计算的重复性而并没有对计算流程进行优化。Gil 等^[105-106]则将计算实验抽象为工作流 (Workflow), 提出了工作流的调度与映射框架, 通过将计算实验中的任务自动分配至分布式节点中执行, 从而提高运行效率。该工作已经应用于地震相关的分析与研究中^[107-108]。

2) 计算实验的高性能计算。目前已有的计算实验建模工具 (如: Swarm、Repast 和 Netlogo 等) 多限于小规模单机运行。为了降低计算实验所需的计算资源, 可通过减少不必要的计算, 适应新的硬件环境和计算架构等方式提高计算实验的运行效率。一些研究者通过优化模型的策略, 提高模型的计算效率。例如: Page 等^[109]尝试只更新那些状态发生变化的代理, 从而避免遍历所有代理的时间花销。一些工作则研究如何使用 GPU 提高计算模型的运行效率^[110-111]。

3) 计算实验的平行执行。目前计算实验研究多侧重于计算模型的构建和对计算模型的分析,

较少强调计算模型与真实系统之间的平行互动关系, 忽视了平行系统对现实系统的影响和指导作用. “人工系统 + 计算实验 + 平行执行” (Artificial systems, computational experiments, and parallel execution, ACP) 方法是在人工社会、计算实验等方法的基础上采用平行执行的方式实现对复杂系统的控制与管理^[1, 35, 112–113].

ACP 的基本思想框架如图 3 所示. 随着大数据时代的来临^[114], 以数据为驱动的平行执行与平行控制研究得到了广泛的应用, 目前已经具有多种具体应用形式^[115–117], 如: 平行交通^[118–119]、安全领域中对动态网民群体进行分析与预测^[120–121]、大型企业生产的平行管理系统等^[122–123].

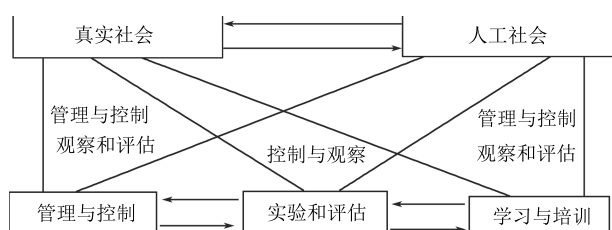


图 3 ACP 方法框架图^[1]

Fig. 3 ACP method framework^[1]

5 计算实验应用

计算实验的应用十分广泛, 包括金融、生态学、生物和医学、物理、地理、交通、军事等领域^[51, 54, 124–127]. 为了细致地描述计算实验应用, 本节将首先介绍计算实验应用的特点并主要讨论计算实验在交通领域中的应用.

5.1 计算实验应用的特点

在计算实验的诸多应用中, 计算模型的异质性与整体性特点凸显. 异质性是指计算模型中个体行为之间的差异. 在传统的金融市场研究中, 通常假定经济个体是完全理性的、同质的^[50], 而在基于计算实验的金融市场研究中, 可研究个体差异对市场波动的影响; 计算模型的整体性是指从全局的角度对个体间的互动以及环境与个体间的互动进行描述, 例如传染病研究中对个体交互行为及地理因素的考虑^[128–131]. 这些特性有助于在微观行为与宏观现象间建立连接, 利于对人与社会复杂行为的描述^[132–133], 使得计算实验表现出很强的跨学科性.

5.2 交通领域的计算实验应用

交通问题为计算实验方法应用于复杂系统和复杂性科学的研究提供了一个十分具体且具有科学和社会影响的案例. 针对交通系统中的三个要素—

人、路、环境, 可从不同层面开展计算实验的研究工作.

人的因素在交通系统的研究中非常重要. 人口老龄化、出生性别比异常、流动人口的活动等人口现象对交通系统的影响都可在计算实验中加以研究^[134]. 缪青海^[135]通过建立不同的人工人口规则研究群体性社会经济活动对交通的影响. 文献^[136]研究了人口规模、人口年龄结构以及人口分布对交通需求的影响, 分析了增长型、静止型、缩减型三种人口发展趋势下的交通需求变化.

在道路网络的研究中, 吕宜生等^[137]通过计算实验评估路网交通运行状况与路网通行能力. 汤淑明^[138]通过在路网内的两个交叉口之间增设路段, 将单行道改建成双向行驶车道, 考察增加建设道路对交通状况的影响. 在文献^[139]中, 作者通过计算实验评估交通信号控制系统对于路网通行能力的影响.

影响交通系统的环境因素庞杂, 例如天气、交通事故、突发事件等. 在文献^[140]中, 作者通过计算实验分析事故与天气因素对交通系统的影响. 文章通过观察选定路段的交通流量、平均车速、平均延误时间、尾气排放的变化, 分析发生交通事故后的涌现行为; 天气方面, 文章研究了中等程度降雨 (降雨量大于 10 毫米) 天气下, 发生在下班高峰时段的交通拥堵对路网流通性的影响, 观察降雨天气对车辆整体涌现行为的影响.

6 总结与展望

本文对计算实验研究方法与应用进展进行了综述, 分别从计算实验的研究基础、计算实验研究现状以及相关应用领域进行了阐述. 计算实验的研究基础主要介绍了实验设计、计算机仿真等传统方法. 计算实验的主要研究内容包括计算模型构建、计算实验设计、计算实验执行等. 最后本文分析了计算实验应用的特点, 并主要介绍了计算实验在交通领域中的应用.

计算实验的提出为复杂系统与分析提供了新的思路, 尽管相关研究已经取得了很大进步. 但是由于应用领域范围广泛, 在计算模型验证, 计算实验设计方面等还存在着很多问题与挑战. 例如: 计算实验中的模型种类多样, 计算模型之间很难相互比较与验证; 计算实验的设计方法尚缺乏通用性, 许多经典的实验设计方法不适用于计算环境中; 计算实验的开销不断增加, 需要可扩展的软件与硬件平台支持等. 本质上, 计算实验的研究面临着如何寻求建模灵活性与结论可信性之间的平衡问题. 我们的思路就是

针对不同的应用领域, 寻找计算实验之间的共性并抽象化, 设计更加通用的实验设计与分析方法, 突破传统方法中关于计算机仿真的理念与框架, 为复杂系统的管理与控制提供可靠的工具支持。

References

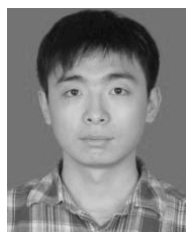
- Wang Fei-Yue. Artificial societies, computational experiments, and parallel systems: a discussion on computational theory of complex social-economic systems. *Complex Systems and Complexity Science*, 2004, **1**(4): 25–35
(王飞跃. 人工社会、计算实验、平行系统 — 关于复杂社会经济系统计算研究的讨论. 复杂系统与复杂性科学, 2004, **1**(4): 25–35)
- Wang Fei-Yue. Computational experiments for behavior analysis and decision evaluation of complex systems. *Journal of System Simulation*, 2004, **16**(5): 893–897
(王飞跃. 计算实验方法与复杂系统行为分析和决策评估. 系统仿真学报, 2004, **16**(5): 893–897)
- Wang Fei-Yue. Computational theory and method on complex system. *China Basic Science*, 2004, **6**(5): 3–10
(王飞跃. 关于复杂系统研究的计算理论与方法. 中国基础科学, 2004, **6**(5): 3–10)
- Montgomery D C. *Design and Analysis of Experiments*. New York: John Wiley and Sons Inc., 2008
- Fisher R A. *The Design of Experiments*. New York: Macmillan Pub. Co., 1971
- Box G E P, Wilson K B. On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society. Series B (Methodological)*, 1951, **13**(1): 1–45
- Box G E P. Statistics as a catalyst to learning by scientific method, Part II — A discussion. *Journal of Quality Technology*, 1999, **31**(1): 16–29
- Taguchi G, Wu Y. *Introduction to Off-line Quality Control*. Nagoya: Central Japan Quality Control Assoc, 1979
- Kackar R N. Off-line quality control, parameter design, and the Taguchi method. *Journal of Quality Technology*, 1985, **17**: 176–188
- Taguchi G. *Introduction to Quality Engineering: Designing Quality into Products and Processes*. Quality Resources, 1986
- Wolf A, Henes D, Bogdanski S, Lutz T, Krämer E. Statistical analysis of parameter variations using the Taguchi method. In: Proceedings of the 2013 Management and Minimisation of Uncertainties and Errors in Numerical Aerodynamics. Berlin, Heidelberg: Springer, 2013. 247–264
- Law A M, Kelton D W. *Simulation Modeling and Analysis*. New York: McGraw-Hill, 2000
- Fishwick P A. *Simulation Model Design and Execution: Building Digital Worlds*. Upper Saddle River, NJ: Prentice Hall, 1995
- Rubinstein R Y, Kroese D P. *Simulation and the Monte Carlo Method*. New York: Wiley-Interscience, 2011
- Mielczarek B, Uzia ko-Mydlikowska J. Application of computer simulation modeling in the health care sector: a survey. *SIMULATION*, 2012, **88**(2): 197–216
- Teichroew D, Lubin J F. Computer simulation-discussion of the technique and comparison of Languages. *Communications of the ACM*, 1966, **9**(10): 723–741
- Balci O. A life cycle for modeling and simulation. *Simulation*, 2012, **88**(7): 870–883
- Fujimoto R M. *Parallel and Distributed Simulation Systems*. New York: Wiley, 2000
- Naylor T H. *Computer Simulation Techniques*. New York: Wiley, 1966
- Kleijnen J P C. Experimental design for sensitivity analysis, optimization, and validation of simulation models. *Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice*. Hoboken, NJ: John Wiley & Sons, 2007. 173–223
- Sargent R G. Verification and validation of simulation models. In: Proceedings of the 37th Conference on Winter Simulation. Orlando, Florida: Winter Simulation Conference, 2005. 130–143
- Kleijnen J P C. Verification and validation of simulation models. *European Journal of Operational Research*, 1995, **82**(1): 145–162
- Fagiolo G, Moneta A, Windrum P. A critical guide to empirical validation of agent-based models in economics: methodologies, procedures, and open problems. *Computational Economics*, 2007, **30**(3): 195–226
- Burdick D S, Naylor T H. Design of computer simulation experiments for industrial systems. *Communications of the ACM*, 1966, **9**(5): 329–339
- Lee L H, Chew E P, Frazier P I, Jia Q S, Chen C H. Advances in simulation optimization and its applications. *IIE Transactions*, 2013, **45**(7): 683–684
- Carson Y, Maria A. Simulation optimization: methods and applications. In: Proceedings of the 29th Conference on Winter Simulation. Atlanta, Georgia: Winter Simulation Conference, 1997. 118–126
- Fu M C. Optimization via simulation: a review. *Annals of Operations Research*, 1994, **53**(1): 199–247
- Swisher J R, Hyden P D, Jacobson S H, Schruben L W. A survey of recent advances in discrete input parameter discrete-event simulation optimization. *IIE Transactions*, 2004, **36**(6): 591–600
- Swisher J R, Jacobson S H, Yücesan E. Discrete-event simulation optimization using ranking, selection, and multiple comparison procedures: A survey. *ACM Transactions on Modeling and Computer Simulation*, 2003, **13**(2): 134–154
- Wang Fei-Yue. On the modeling, analysis, control and management of complex systems. *Complex Systems and Complexity Science*, 2006, **3**(2): 26–34
(王飞跃. 关于复杂系统的建模、分析、控制和管理. 复杂系统与复杂性科学, 2006, **3**(2): 26–34)
- Kohler T A, Gumerman G G. *Dynamics in Human and Primate Societies: Agent-based Modeling of Social and Spatial Processes*. Oxford: Oxford University Press, 2000.
- Lansing J S. Artificial societies and the social sciences. *Artificial Life*, 2002, **8**(3): 279–292
- Prietula M, Carley K, Gasser L. *Simulating Organizations: Computational Models of Institutions and Groups*. Cambridge: AAAI Press/MIT Press, 1998
- Wang Fei-Yue, Lansing J-Stephen. From artificial life to artificial societies — new methods for studies of complex social systems. *Complex Systems and Complexity Science*, 2004, **1**(1): 33–41
(王飞跃, 兰森史帝夫. 从人工生命到人工社会 — 复杂社会系统研究的现状和展望. 复杂系统与复杂性科学, 2004, **1**(1): 33–41)

- 35 Wang Fei-Yue. Parallel system methods for management and control of complex systems. *Control and Decision*, 2004, **19**(5): 485–489, 541
(王飞跃. 平行系统方法与复杂系统的管理和控制. 控制与决策, 2004, **19**(5): 485–489, 541)
- 36 Chen V C P, Tsui K L, Barton R R, Meckesheimer M. A review on design, modeling and applications of computer experiments. *IIE Transactions*, 2006, **38**(4): 273–291
- 37 Fang K T, Li R Z, Sudjianto A. *Design and Modeling for Computer Experiments*. London: Chapman and Hall/CRC, 2006
- 38 Kleijnen J P C [Author], Zhang Lie-Gang, Zhang Jian-Kang, Liu Xing-Ke [Translator]. *Design and Analysis of Simulation Experiments*. Beijing: Electronic Industry Press, 2010
(Kleijnen J P C [著], 张列刚, 张建康, 刘兴科 [译]. 仿真实验设计与分析. 北京: 电子工业出版社, 2010)
- 39 Kleijnen J P C, Sanchez S M, Lucas T W, Cioppa T M. State-of-the-art review: a user's guide to the brave new world of designing simulation experiments. *INFORMS Journal on Computing*, 2005, **17**(3): 263–289
- 40 Steiger N M, Lada E K, Wilson J R, Joines J A, Alexopoulos C, Goldsman D. ASAP3: a batch means procedure for steady-state simulation analysis. *ACM Transactions on Modeling and Computer Simulation*, 2005, **15**(1): 39–73
- 41 Alexopoulos C, Goldsman D. To batch or not to batch? *ACM Transactions on Modeling and Computer Simulation*, 2004, **14**(1): 76–114
- 42 Loeppky J L, Sacks J, Welch W J. Choosing the sample size of a computer experiment: a practical guide. *Technometrics*, 2009, **51**(4): 366–376
- 43 Dong Zhi-Yong. *Experimental Economics*. Beijing: Peiking University Press, 2008
(董志勇. 实验经济学. 北京: 北京大学出版社, 2008)
- 44 Wang Fei-Yue. Web social media in disaster reduction and emergence management. *Science and Technology Review*, 2008, **26**(10): 30–31
(王飞跃. 万维社会媒体在防灾应急中的作用. 科技导报, 2008, **26**(10): 30–31)
- 45 Louie M A, Carley K M. Balancing the criticisms: validating multi-agent models of social systems. *Simulation Modelling Practice and Theory*, 2008, **16**(2): 242–256
- 46 Bernardes A T, Stauffer D, Kertész J. Election results and the Sznajd model on Barabasi network. *The European Physical Journal B*, 2002, **25**(1): 123–127
- 47 Weng L, Flammini A, Vespignani A, Menczer F. Competition among memes in a world with limited attention. *Scientific Reports*, 2012, **2**: 335 [Online], available: <http://www.nature.com/srep/2012/120329/srep00335/full/srep00335.html>, July 29, 2013
- 48 Hodges J S. Six (or so) things you can do with a bad model. *Operations Research*, 1991, **39**(3): 355–365
- 49 Dewar J A, Banks S C, Hodges J S, Lucas T W, Saunders-Newton D K, Vye P. Credible Uses of the Distributed Interactive Simulation (DIS) System [Online], available: <http://www.dtic.mil/docs/citations/ADA306986>, July 29, 2013
- 50 Macal C M, North M J. Tutorial on agent-based modelling and simulation. *Journal of Simulation*, 2010, **4**(3): 151–162
- 51 Chan W K V, Young-Jun S, Macal C M. Agent-based simulation tutorial — simulation of emergent behavior and differences between agent-based simulation and discrete-event simulation. In: *Proceedings of the 2010 Winter Simulation Conference*. Baltimore, Maryland: Winter Simulation Conference, 2010. 135–150
- 52 Metzger M, Polakow G. A survey on applications of agent technology in industrial process control. *IEEE Transactions on Industrial Informatics*, 2011, **7**(4): 570–581
- 53 Jonker C M, Treur J. A formal approach to building compositional agent-based simulations. *Simulating Social Complexity*. Berlin, Heidelberg: Springer, 2013. 57–94
- 54 Sotomayor M, Schulten K. Single-molecule experiments in vitro and in silico. *Science*, 2007, **316**(5828): 1144–1148
- 55 Goldenberg J, Libai B, Muller E. Talk of the network: a complex systems look at the underlying process of word-of-mouth. *Marketing Letters*, 2001, **12**(3): 211–223
- 56 Kempe D, Kleinberg J, Tardos E. Maximizing the spread of influence through a social network. In: *Proceedings of the 9th ACM International Conference on Knowledge Discovery and Data Mining*. Washington, USA: ACM, 2003. 137–146
- 57 Toni F, Bentahar J. Computational logic-based agents. *Autonomous Agents and Multi-Agent Systems*, 2008, **16**(3): 211–213
- 58 Su K L, Sattar A, Lin H, Reynolds M. A modal logic for beliefs and pro attitudes. In: *Proceedings of the 22nd National Conference on Artificial Intelligence*, Volume 1. Vancouver, British Columbia, Canada: AAAI, 2007. 496–501
- 59 Wooldridge M J. *The Logical Modelling of Computational Multi-agent Systems* [Ph.D. dissertation], University of Mancheste, Oxford, UK, 1992
- 60 Fisher M. Temporal development methods for agent-based. *Autonomous Agents and Multi-Agent Systems*, 2005, **10**(1): 41–66
- 61 Macal C M, North M J. Tutorial on agent-based modelling and simulation. *Journal of Simulation*, 2010, **4**(3): 151–162
- 62 Granovetter M, Soong R. Threshold models of diffusion and collective behavior. *The Journal of Mathematical Sociology*, 1983, **9**(3): 165–179
- 63 Hasan S, Ukkusuri S V. A threshold model of social contagion process for evacuation decision making. *Transportation Research, Part B: Methodological*, 2011, **45**(10): 1590–1605
- 64 Sznajd-Weron K, Sznajd J. Opinion evolution in closed community. *International Journal of Modern Physics C*, 2000, **11**(6): 1157–1165
- 65 Timpanaro A M, Prado C P C. Connections between the Sznajd model with general confidence rules and graph theory. *Physical Review E*, 2012, **86**(4): 046109
- 66 Crokidakis N. Effects of mass media on opinion spreading in the Sznajd sociophysics model. *Physica A: Statistical Mechanics and Its Applications*, 2012, **391**(4): 1729–1734
- 67 Deffuant G, Neau D, Amblard F, Weisbuch G. Mixing beliefs among interacting agents. *Advances in Complex Systems*, 2000, **3**(1–4): 87–98
- 68 Grauwin S, Jensen P. Opinion group formation and dynamics: structures that last from nonlasting entities. *Physical Review E*, 2012, **85**(6): 066113
- 69 Young H P. The dynamics of social innovation. *Proceedings of the National Academy of Sciences of the United States of America*, 2011, **108**(Supplement 4): 21285–21291

- 70 Dixit A K, Meyerson Milgrom E M, Milgrom P R. Dynamics of social, political, and economic institutions. *Proceedings of the National Academy of Sciences of the United States of America*, 2011, **108**(Supplement 4): 21283–21284
- 71 Redner S. *A Guide to First-passage Processes*. Cambridge: Cambridge University Press, 2001
- 72 Durrett R, Gleeson J P, Lloyd A L, Mucha P J, Shi F, Sivakoff D, Socolar J E S, Varghese C. Graph fission in an evolving voter model. *Proceedings of the National Academy of Sciences of the United States of America*, 2012, **109**(10): 3682–3687
- 73 Galam S. Minority opinion spreading in random geometry. *The European Physical Journal B — Condensed Matter and Complex Systems*, 2002, **25**(4): 403–406
- 74 Zheng X L, Zhong Y G, Zeng D, Wang F Y. Social influence and spread dynamics in social networks. *Frontiers of Computer Science*, 2012, **6**(5): 611–620
- 75 Zhang Q P, Wang F Y, Zeng D, Wang T. Understanding crowd-powered search groups: a social network perspective. *PLoS ONE*, 2012, **7**(6): e39749 [Online], available: <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0039749>, July 29, 2013
- 76 Bernardes A T, Stauffer D, Kertész J. Election results and the Sznajd model on Barabasi network. *The European Physical Journal B — Condensed Matter and Complex Systems*, 2002, **25**(1): 123–127
- 77 Suchecki K, Eguíluz V M, Miguel S M. Voter model dynamics in complex networks: Role of dimensionality, disorder, and degree distribution. *Physical Review E*, 2005, **72**(3): 036132
- 78 Castellano C, Vilone D, Vespignani A. Incomplete ordering of the voter model on small-world networks. *Europhysics Letters*, 2003, **63**(1): 153–158
- 79 Sood V, Redner S. Voter model on heterogeneous graphs. *Physical Review Letters*, 2005, **94**(17): 178701
- 80 De Smith M J, Goodchild M F, Longley P A. *Geospatial Analysis: A Comprehensive Guide to Principles, Techniques and Software Tools*. Leicester: Troubador Publishing, 2009
- 81 Balcan D, Colizza V, Goncalves B, Hu H, Ramasco J J, Vespignani A. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proceedings of the National Academy of Sciences of the United States of America*, 2009, **106**(51): 21484–21489
- 82 Guo D, Ren B, Wang C. Integrated agent-based modeling with GIS for large scale emergency simulation. *Advances in Computation and Intelligence*, 2008, **5370**: 618–625
- 83 Robinson C D, Brown D E. First responder information flow simulation: a tool for technology assessment. In: Proceedings of the 37th Conference on Winter Simulation. Orlando, Florida: Winter Simulation Conference, 2005. 919–925
- 84 Taillandier P, Vo D A, Amouroux E, Drogoul A. GAMA: a simulation platform that integrates geographical information data, agent-based modeling and multi-scale control. In: Proceedings of the 2012 Principles and Practice of Multi-Agent Systems. Berlin, Heidelberg: Springer, 2012. 242–258
- 85 Li Q, Yang T, Zhao E, Xia X, Han Z. The impacts of information-sharing mechanisms on spatial market formation based on agent-based modeling. *PLoS ONE*, 2013, **8**(3): e58270 [Online], available: <http://www.plosone.org/article/info:doi/10.1371/journal.pone.0058270>, July 29, 2013
- 86 Nüsse T, Kleiner A, Brenner M. Approaching urban disaster reality: the ResQ firesimulator. In: Proceedings of the 2004 RoboCup: Robot Soccer World Cup VIII. Berlin, Heidelberg: Springer, 2005. 474–482
- 87 Tanigawa M, Takahashi T, Koto T, Takeuchi I, Noda I. Urban flood simulation as a component of integrated earthquake disaster simulation. In: Proceedings of the 2005 IEEE International Safety, Security and Rescue Robotics Workshop. Kobe, Japan: IEEE, 2005. 248–252
- 88 Carley K M, Altman N, Kaminsky B, Nave D, Yahja A. BioWar: A City-Scale Multi-Agent Network Model of Weaponized Biological Attacks, DTIC Document, CMU-ISRI-04-101, Carnegie-Mellon University, USA, 2004
- 89 Sanchez S M. Work smarter, not harder: guidelines for designing simulation experiments. In: Proceedings of the 39th Conference on Winter Simulation. Piscataway, NJ: IEEE, 2007. 84–94
- 90 Koehler J R, Owen A B. Computer experiments. *Handbook of Statistics*. New York: Elsevier Science, 1996. 261–308
- 91 Kleijnen J C, Beers W, Nieuwenhuysen I. Expected improvement in efficient global optimization through bootstrapped kriging. *Journal of Global Optimization*, 2012, **54**(1): 59–73
- 92 Jin R, Chen W, Simpson T W. Comparative studies of metamodelling techniques under multiple modelling criteria. *Structural and Multidisciplinary Optimization*, 2001, **23**(1): 1–13
- 93 Morris M D, Mitchell T J, Ylvisaker D. Bayesian design and analysis of computer experiments: use of derivatives in surface prediction. *Technometrics*, 1993, **35**(3): 243–255
- 94 Simpson T W, Poplinski J D, Koch P N, Allen J K. Metamodels for computer-based engineering design: survey and recommendations. *Engineering with Computers*, 2001, **17**(2): 129–150
- 95 Staelin C. Parameter Selection for Support Vector Machines, Technical Report HPL-2002-354R1, Hewlett-Packard Company, 2003
- 96 McKay M D, Beckman R J, Conover W J. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 2000, **42**(1): 55–61
- 97 Fang K T, Lin D K J, Winker P, Zhang Y. Uniform design: theory and application. *Technometrics*, 2000, **42**(3): 237–248
- 98 Deutsch J L, Deutsch C V. Latin hypercube sampling with multidimensional uniformity. *Journal of Statistical Planning and Inference*, 2012, **142**(3): 763–772
- 99 Morgan J P, Deng X W. Experimental design. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2012, **2**(2): 164–172
- 100 Ning J H, Zhou Y D, Fang K T. Discrepancy for uniform design of experiments with mixtures. *Journal of Statistical Planning and Inference*, 2011, **141**(4): 1487–1496
- 101 Lee E [Author], Chen Jia-Ding, Dai Zhong-Wei [Translator]. *Statistical Methods for Survival Data Analysis*. Beijing: China Statistics Press, 1998
(Lee E [著], 陈家鼎, 戴中维 [译]. 生存数据分析的统计方法. 北京: 中国统计出版社, 1998)
- 102 Bettonvil B, Kleijnen J P C. Searching for important factors in simulation models with many factors: sequential bifurcation. *European Journal of Operational Research*, 1997, **96**(1): 180–194

- 103 Kleijnen J P C, van Beers W C M. Application-driven sequential designs for simulation experiments: Kriging meta-modelling. *Journal of the Operational Research Society*, 2004, **55**(8): 876–883
- 104 Fomel S, Hennenfent G. Reproducible computational experiments using SCons. In: *Proceedings of the 2007 IEEE International Conference on Acoustics, Speech and Signal Processing*. Austin, Texas: IEEE, 2007. 1257–1260
- 105 Gil Y, Ratnakar V, Deelman E, Mehta G, Kim J. Wings for pegasus: creating large-scale scientific applications using semantic representations of computational workflows. In: *Proceedings of the 19th National Conference on Artificial Intelligence*. Menlo Park, CA: AAAI Press, 2007. 1767–1774
- 106 Gil Y, Ratnakar V, Kim J, González-Calero P A, Groth P T, Moody J, Deelman E. Wings: intelligent workflow-based design of computational experiments. *IEEE Intelligent Systems*, 2011, **26**(1): 62–72
- 107 Maechling P, Chalupsky H, Dougherty M, Deelman E, Gil Y, Gullapalli S, Gupta V, Kesselman C, Kim J, Mehta G, Mendenhall B, Russ T, Singh G, Spraragen M, Staples G, Vahi K. Simplifying construction of complex workflows for non-expert users of the Southern California Earthquake Center Community Modeling Environment. *ACM SIGMOD Record*, 2005, **34**(3): 24–30
- 108 Deelman E, Callaghan S, Field E, Francoeur H, Graves R, Gupta V, Jordan T H, Kesselman C, Maechling P, Mehta G, Okaya D, Vahi K, Zhao L. Managing large-scale workflow execution from resource provisioning to provenance tracking: The cybershake example. In: *Proceedings of the 2nd IEEE International Conference on e-Science and Grid Computing*. Washington D.C., USA: IEEE, 2006. 4–6
- 109 Page B, Knaak N, Kruse A. A discrete event simulation framework for agent-based modelling of logistic systems. In: *Proceedings of the 2007 GI Jahrestagung*. Bremen, German, 2007. 397–404
- 110 Lysenko M, D'Souza R M. A framework for megascale agent based model simulations on graphics processing units. *Journal of Artificial Societies and Social Simulation*, 2008, **11**(4): 10. [Online], available: <http://jasss.soc.surrey.ac.uk/11/4/10.html>, July 29, 2013
- 111 Wang K, Shen Z. A GPU-based parallel genetic algorithm for generating daily activity plans. *IEEE Transactions on Intelligent Transportation Systems*, 2012, **13**(3): 1474–1480
- 112 Wang F Y. Toward a paradigm shift in social computing: the ACP approach. *IEEE Intelligent Systems*, 2007, **22**(5): 65–67
- 113 Lun Shu-Xian. Research on the classification of parallel execution modes of ACP theory. *Acta Automatica Sinica*, 2012, **38**(10): 1602–1608
(伦淑娴. ACP 理论的并行执行方式分类研究. *自动化学报*, 2012, **38**(10): 1602–1608)
- 114 Wang F Y. A big-data perspective on AI: Newton, Merton, and analytics intelligence. *IEEE Intelligent Systems*, 2012, **27**(5): 2–4
- 115 Wang Fei-Yue. Parallel control: a method for data-driven and computational control. *Acta Automatica Sinica*, 2013, **39**(4): 293–302
(王飞跃. 平行控制: 数据驱动的计算控制方法. *自动化学报*, 2013, **39**(4): 293–302)
- 116 Wang Peng, Chen Sen. A parallel control approach to optimization of information security management measures. *Acta Automatica Sinica*, 2011, **37**(11): 1351–1355
(王鹏, 陈森. 基于平行控制的信息安全管理措施仿真与优化. *自动化学报*, 2011, **37**(11): 1351–1355)
- 117 Huang Wen-De, Wang Wei, Xu Xin, Xi Xiao-Ning. Computational experiments for abort planning of manned lunar landing mission based on ACP approach. *Acta Automatica Sinica*, 2012, **38**(11): 1794–1803
(黄文德, 王威, 徐昕, 郝晓宁. 基于 ACP 方法的载人登月中止规划的实验研究. *自动化学报*, 2012, **38**(11): 1794–1803)
- 118 Xiong G, Dong X S, Fan D, Zhu F H, Wang K F, Lv Y S. Parallel traffic management system and its application to the 2010 Asian Games. *IEEE Transactions on Intelligent Transportation Systems*, 2013, **14**(1): 225–235
- 119 Zhu F H, Wen D, Chen S H. Computational traffic experiments based on artificial transportation systems: an application of ACP approach. *IEEE Transactions on Intelligent Transportation Systems*, 2013, **14**(1): 189–198
- 120 Wang Fei-Yue. Study on cyber-enabled social movement organizations based on social computing and parallel systems. *Journal of University of Shanghai for Science and Technology*, 2011, **33**(1): 8–17
(王飞跃. 基于社会计算和平行系统的动态网民群体研究. *上海理工大学学报*, 2011, **33**(1): 8–17)
- 121 Zhou Yun, Qiao Hai-Quan, Qiu Xiao-Gang, Huang Ke-Di, Hu De-Wen. Research and implement on dynamic simulation engine for society computing experiments. *Journal of National University of Defense Technology*, 2011, **33**(4): 163–167
(周云, 乔海泉, 邱晓刚, 黄柯棣, 胡德文. 社会计算实验动态仿真引擎的研究与实现. *国防科技大学学报*, 2011, **33**(4): 163–167)
- 122 Cheng Chang-Jian, Cui Feng, Li Le-Fei, Xiong Gang, Zou Yu-Min, Liao Chang-Yong. Parallel management systems for complex productions systems: methods and cases. *Complex Systems and Complexity Science*, 2010, **7**(1): 24–32
(程长建, 崔峰, 李乐飞, 熊刚, 邹余敏, 廖昌勇. 复杂生产系统的平行管理方法与案例. *复杂系统与复杂性科学*, 2010, **7**(1): 24–32)
- 123 Xue Xiao, Wang Yang. ACP-based research on complexity of cluster supply chain. *Computer Engineering and Design*. 2011, **32**(12): 4030–4034
(薛霄, 王杨. 基于 ACP 的集群式供应链复杂性研究. *计算机工程与设计*, 2011, **32**(12): 4030–4034)
- 124 Zheng X L, Ke G Y, Zeng D D, Ram S, Hao L. Next-generation team-science platform for scientific collaboration. *IEEE Intelligent Systems*, 2011, **26**(6): 72–76
- 125 Zeng D, Wang F Y, Zheng X L, Yuan Y, Chen G Q, Chen J. Intelligent-commerce research in China. *IEEE Intelligent Systems*, 2008, **23**(6): 14–18
- 126 Zheng Xiao-Long, Zhong Yong-Guang, Wang Fei-Yue, Zeng Da-Jun, Zhang Qing-Peng, Cui Kai-Nan. Social dynamics research based on web information. *Complex Systems and Complexity Science*, 2011, **8**(3): 1–12
(郑晓龙, 钟永光, 王飞跃, 曾大军, 张清鹏, 崔凯楠. 基于网络信息的社会动力学研究. *复杂系统与复杂性科学*, 2011, **8**(3): 1–12)
- 127 Sanchez S M, Lucas T W, Sanchez P J, Nannini C J, Wan H. Designs for large-scale simulation experiments, with applications to defense and homeland security. *Design and Analysis of Experiments: Special Designs and Applications, Volume 3*. Hoboken, NJ, USA: John Wiley and Sons, Inc, 2012. 413–441

- 128 Epstein J M, Goedecke D M, Yu F, Morris R J, Wagener D K, Bobashev G V. Controlling pandemic flu: the value of international air travel restrictions. *PLoS ONE*, 2007, **2**(5): e401. [Online], available: <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0000401>, July 29, 2013
- 129 Wang Y Z, Zeng D, Cao Z D, Wang Y, Song H B, Zheng X L. The impact of community structure of social contact network on epidemic outbreak and effectiveness of non-pharmaceutical interventions. In: *Proceedings of the 6th Pacific Asia Conference on Intelligence and Security Informatics*. Berlin, Heidelberg: Springer, 2011. 108–120
- 130 Duan W, Cao Z D, Ge Y Z, Qiu X G. Modeling and simulation for the spread of H1N1 influenza in school using artificial societies. In: *Proceedings of the 6th Pacific Asia Conference on Intelligence and Security Informatics*. Berlin, Heidelberg: Springer, 2011. 121–129
- 131 Duan W, Cao Z, Cui K, Zheng X, Qiu X. Heterogeneous and stochastic agent-based models for characteristic analysis of super spreaders in infectious diseases. *IEEE Intelligent Systems*, [Online], available: <http://doi.ieeecomputersociety.org/10.1109/MIS.2013.29>, July 29, 2013
- 132 Tan Z W, Mao W J, Zeng D, Li X C, Bao X G. Acquiring netizen group's opinions for modeling food safety events. In: *Proceedings of the 2012 IEEE International Conference on Intelligence and Security Informatics*. Arlington, VA: IEEE, 2012. 114–119
- 133 Chen W Y, Li X, Zeng D. Estimating collective belief in fixed odds betting. In: *Proceedings of the 6th Pacific Asia Conference on Intelligence and Security Informatics*. Berlin, Heidelberg: Springer, 2011. 54–63
- 134 Wang Fei-Yue, Jiang Zheng-Hua, Dai Ru-Wei. Population studies and artificial societies: a discussion of artificial population systems and their applications. *Complex Systems and Complexity Science*, 2005, **2**(1): 1–9
(王飞跃, 蒋正华, 戴汝为. 人口问题与人工社会方法: 人工人口系统的设想与应用. *复杂系统与复杂性科学*, 2005, **2**(1): 1–9)
- 135 Miao Qing-Hai. Design of Artificial Transportation Systems Based on Srtience [Ph.D. dissertation], Institute of Automation, Chinese Academy of Sciences, China, 2007
(缪青海. 基于 Artience 的人工交通系统研究与设计 [博士学位论文], 中国科学院自动化研究所, 中国, 2007)
- 136 Zhao Hong-Xia. A Study on Computational Experiments of Relationship between Artificial Population and Travel Demand [Ph.D. dissertation], Institute of Automation, Chinese Academy of Sciences, China, 2009
(赵红霞. 人工人口与交通需求关系计算实验研究 [博士学位论文], 中国科学院自动化研究所, 中国, 2009)
- 137 Lv Yi-Sheng, Ou Yan, Tang Shu-Ming, Zhu Feng-Hua, Zhao Hong-Xia. Computational experiments of evaluating road network traffic conditions based on artificial transportation systems. *Journal of Jilin University (Engineering and Technology Edition)*, 2009, **39**(S2): 87–90
(吕宜生, 欧彦, 汤淑明, 朱风华, 赵红霞. 基于人工交通系统的路网交通运行状况评估的计算实验. *吉林大学学报 (工学版)*, 2009, **39**(S2): 87–90)
- 138 Tang Shu-Ming. A Preliminary Study for Basic Approaches in Artificial Transportation Systems [Ph.D. dissertation], Institute of Automation, Chinese Academy of Sciences, China, 2005
(汤淑明. 人工交通系统的基本方法研究 [博士学位论文], 中国科学院自动化研究所, 中国, 2005)
- 139 Zhu F H, Li G X, Li Z J, Chen C, Wen D. A case study of evaluating traffic signal control systems using computational experiments. *IEEE Transactions on Intelligent Transportation Systems*, 2011, **12**(4): 1220–1226
- 140 Zhu Feng-Hua. A Study on the Evaluation of Urban Traffic Signal Control System Based on Artificial Transportation Systems [Ph.D. dissertation], Institute of Automation, Chinese Academy of Sciences, China, 2008
(朱风华. 基于人工交通系统的城市交通信号控制系统评价研究 [博士学位论文], 中国科学院自动化研究所, 中国, 2008)



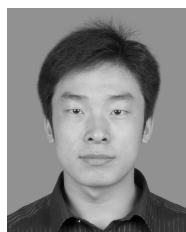
崔凯楠 西安交通大学电信学院博士研究生. 主要研究方向为计算实验与社会媒体分析. E-mail: kainan.cui@live.cn
(**CUI Kai-Nan** Ph.D. candidate at the School of Electronic and Information Engineering, Xi'an Jiaotong University. His research interest covers computational experiments and social media analytics.)



郑晓龙 中国科学院自动化研究所复杂系统管理与控制国家重点实验室副研究员. 主要研究方向为社会计算与复杂网络. 本文通信作者.
E-mail: xiaolong.zheng@ia.ac.cn
(**ZHENG Xiao-Long** Associate professor at the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. His research interest covers social computing and complex network. Corresponding author of this paper.)

文丁 国防科学技术大学军事计算实验与平行系统技术研究中心教授. 主要研究方向为行为运作管理, 人力资源管理, 管理信息系统, 智能系统.
E-mail: wending2010@gmail.com

(**WEN Ding** Professor at the Research Center for Military Computational Experiments and Parallel Systems Technology, National University of Defense Technology. His research interest covers behavioral operation management, human resource management, management information system, and intelligent systems.)



赵学亮 中国科学院自动化研究所复杂系统管理与控制国家重点实验室工程师. 主要研究方向为社会计算, 智能信息处理. E-mail: xueliang.zhao@ia.ac.cn
(**ZHAO Xue-Liang** Engineer at the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. His research interest covers social computing and intelligent information processing.)