

牛文元 / 主编

社会物理学系列 第6号

社会物理学 网络舆情安全

刘怡君 李倩倩 等 编著
王红兵 马 宁 王光辉



科学出版社

牛文元/主编

社会物理学系列 第6号

社会物理学 网络舆情安全

刘怡君 李倩倩 等 编著
王红兵 马宁 孟光辉



科学出版社

北京

内 容 简 介

本书是社会物理学系列丛书的第6号,在国际社会物理学研究前沿(第1号)、社会物理学理论与应用(第2号)、社会物理学社会动力学(第3号)、社会物理学社会管理学(第4号)、社会物理学社会治理(第5号)的基础上,本书从社会物理学与网络舆情安全研究视角,邀请国内外知名专家专门撰稿,旨在对国际国内网络舆情安全最新的概念方法、系统应用、实证分析等进行深入介绍。

作为自然与社会充分交叉的学科,本书具有重要的理论价值和应用前景,可为从事管理学、社会学、政策学和战略学等专业的科研人员和高校师生提供借鉴和参考。

图书在版编目(CIP)数据

社会物理学网络舆情安全 / 刘怡君等编著. —北京: 科学出版社, 2017.3
(社会物理学系列/牛文元主编; 6)

ISBN 978-7-03-051711-1

I. ①社… II. ①刘… III. ①互联网络-舆论-社会管理学-研究
IV. ①G219

中国版本图书馆 CIP 数据核字(2017)第 024884 号

责任编辑: 李 敏 杨逢渤 / 责任校对: 邹慧卿

责任印制: 张 伟 / 封面设计: 李姗姗

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>

北京建宏印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2017 年 3 月第 一 版 开本: 787×1092 1/16

2017 年 3 月第一次印刷 印张: 10 1/4 插页: 2

字数: 300 000

定价: 98.00 元

(如有印装质量问题, 我社负责调换)

序 言

根据世界发展进程的规律，在国家或地区的人均 GDP 处于 1000 ~ 3000 美元的发展阶段，往往对应着人口、资源、环境、效率、公平等社会矛盾的瓶颈约束最为严重的时期，也是“经济容易失调、社会容易失序、心理容易失衡、社会伦理需要调整重建”的关键时期。在“改革、发展、稳定”的关系中，社会稳定是维系国家系统有序运作的根本保证。未来的发展进程中，中国在“统一社会价值观念，整合社会有序能力，提高社会抗逆水平，健全社会道德约束”的同时，科学地、定量地、实时地诊断、监测社会稳定的总体状态变化，预测社会稳定的动态演化趋势，预警社会稳定的临界突破，提供社会稳定的处理对策，构建一个完整的识别国家稳定总体态势的指挥系统，将成为中央宏观决策的基础参考和得力工具。

2004 年，由中国科学院科技政策与管理科学研究所牛文元研究员（1939 ~ 2016 年）牵头，整合中国科学院的优势力量，6 个研究所的 20 多位专家在中共中央办公厅有关项目的支持下，共同组建了社会稳定预警研究组，正式启动研制“中国社会稳定预警系统”。中国社会稳定预警系统在对中国社会稳定态势定量判断的基础上，将为决策层提供可比、可测、可视和可控的有力手段和依据，从而从根本上提高对社会稳定状况的总体识别、定量诊断和预测预警的能力。

牛文元先生所带领的社会稳定预警研究组遵循着社会物理学的原理与方法，给出现代社会物理学的定义是：应用自然科学（以物理学为核心）的思路、概念、原理和方法，经过有效拓展、合理移植和理性修正，形成的一门用来揭示、模拟、解释和预测社会行为规律的充分交叉性学科。在可持续发展领域的两大核心主线中，社会物理学是认识“人与人之间关系”的基本理论基础。社会物理学已经被认为是社会稳定与社会和谐研究中的重要学科领域。

牛文元先生提出的社会物理学三大基本理论，解释了社会稳定状态与社会和谐水平，受到国内外社会学、传播学、系统学、心理学和计算机学等多学科专家学者的认同和应用。

- 一、社会燃烧理论：研究社会稳定发生与评价的基本机制。
- 二、社会激波理论：研究社会稳定状态的时空分布形式。
- 三、社会行为熵理论：研究社会群体行为规律的动力本源。

以社会燃烧理论为例，社会失衡、社会失序和社会失控，犹如自然界物质燃烧必须具备的三大基本条件一样，即需要“燃烧物质、助燃剂和点火温度”，与自然界燃烧的前条件类比，社会燃烧理论认为“人与自然”之间关系不协调和

“人与人”之间关系不和谐的积累总和，是引发社会不稳定的基本“社会燃烧物质”，构成了社会不稳定的基本材料，国内外的研究均指出：约 20 余项的“社会痛苦指数”的集合与积累可以表征社会燃烧物质的数量和规模；社会分配的不公、社会理性的缺失、社会舆论的挑动、事件的过分夸大、谣言的广泛传布、无中生有的挑衅、社会行为失衡的剧烈放大、单方面利益的贪婪追求等，这些因素共同引发的国民心理变化，构成了社会不稳定的“助燃剂”；具有一定强度和影响的突发事件，是引发社会不稳定的“点火温度”或“导火索”。

牛文元先生生前的 10 余年间一直致力于社会物理学的学科建设，并开展中国社会稳定预警系统研究，基于社会物理学及其三大理论，充分运用计算技术、网络技术、虚拟现实技术完成智能化、定量化和动态化的情景仿真，以期构成科学识别社会系统行为轨迹的现代指挥工具。2007 ~ 2015 年，作为主编出版了《社会物理学系列》（第 1 ~ 5 号），亲自为每本著作作序、撰文。在此将牛文元先生的原话作为激励和指导，以飨所有有志于从事社会物理学、社会稳定、社会和谐和社会治理等领域的有志青年：“科学是全人类的财富，更是全社会的公器，一向具有无私和共享的品格，只有在科学的海洋中洗涤和徜徉，才不仅仅只是为了幸福着别人，从根本上去说也是在幸福着自己。”

是以为序，并沉痛悼念 2016 年 9 月 28 日去世的牛文元先生。

刘怡君

2016 年 10 月 19 日

目 录

CONTENTS

系统工程与系统科学中与大数据相关联的十大进展.....	1
数据驱动的社会激励：从应用到理论	16
基于社会媒体大数据的网络舆情演化趋势预测	29
基于城市节点的微信传播网络社团探测结果分析	38
风险治理中公众信任的研究基础及创新空间	49
国家治理的系统认知	58
新陈代谢 3/4 幂律的权力结构解释	67
公共政策发布前网络舆情仿真研究	78
基于微博信息的舆情数据挖掘研究	87
赋权舆论传播网络中影响力节点发现算法	97
多视角下的舆论传播及演化研究综述.....	107
社会风险研究中的“正态分布陷阱”	123
失败公共政策的网络舆情研究.....	131
政策视角下的中国网络舆情治理（1997 ~ 2016 年）	147

系统工程与系统科学中与大数据相关联的十大进展

顾基发

摘要

大数据已经吸引了不少学者和实际工作者从不同角度去研究它，本文不是从数据和信息科学的角度去研究它，而是从系统的角度去研究。利用系统理论出发去研究与大数据相关的十大进展：大系统、大网络、大数据、大解析、大方法论、大心理、大行为、大综合集成、大研讨厅、大智慧。

1. 引言

大数据似乎已经被宣传的有些过火了。其实要处理的系统越来越大、越来越复杂，这才是我们更要关心的。本文企图从大系统连带出现的一些大数据现象及解决问题的方法着手来看看这些年的重要进展和突破。这里列举出十个，当然是讲其主要的进展，也许可以更多一些，也许可以少一点，其中的一些进展，作者在其他场合也曾有所涉及。

2. 大系统 (big system)

系统的名词早已有了，但是比较系统地去研究它还是 20 世纪 40 ~ 60 年代，之后也络绎不绝地出现了一些新的名词，如复杂系统、大规模系统、开放系统等，一直到 90 年代又出现了一些新的名字，并且受到更深的研究和更广的应用，涉及的规模都很大，这里只介绍其中三个比较著名的大系统：开放复杂巨系统 (open complex giant system)、复杂适应系统 (complex adaptive system) 和体系 (system of systems)。

2.1 开放复杂巨系统

钱学森在 1990 年提出一大类系统，如果其元素和子系统种类很多并有层次

结构,它们之间相互关系又很复杂,这就是复杂巨系统。如果这个系统又是开放的,就称作开放的复杂巨系统。所谓种类多其实就是巨大,而关系复杂是指相互作用是非线性的,而系统和环境之间不断交换物质和能量信息,即开放性。作为例子,钱学森曾提到脑系统、社会系统、军事系统、地理系统(含生态、环境系统)、人体系统、星系系统。

2.2 复杂适应系统

1994年,约翰·霍兰德(John Holland)在圣菲研究所做了名为“隐秩序”的著名演讲,而后出版了《隐秩序——适应性造就复杂性》一书。复杂适应系统(complex adaptive systems, CAS),也称复杂性科学(complexity science),是20世纪末兴起的前沿科学。复杂适应系统理论的核心思想——适应性造就复杂性。

复杂适应系统理论的基本思想可以概述如下:系统中的成员称为具有适应性的主体(adaptive agent),简称为主体。所谓具有适应性,就是指它能够与环境及其他主体进行交互作用。主体在这种持续不断的交互作用的过程中,不断地“学习”或“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。整个宏观系统的演变或进化,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,都是在这个基础上逐步派生出来的。

复杂适应系统理论把系统的成员看做是具有自身目的与主动性的、积极的主体。正是这种主动性及它与环境的反复的、相互的作用,才是系统发展和进化的基本动因。宏观的变化和个体分化都可以从个体的行为规律中找到根源。霍兰把个体与环境之间这种主动的、反复的交互作用用“适应”一词加以概括,适应产生复杂性,如蚁群、生态、胚胎、神经网络、人体免疫系统、计算机网络和全球经济系统。所有这些系统中,众多独立的要素在许多方面进行着相互作用。在每种情况下,这些无穷无尽的相互作用使每个复杂系统作为一个整体产生了自发性的自组织。

复杂适应系统理论的提出,正好解决了管理实践中如何才能增强组织适应力的问题;约翰·霍兰德认为将复杂适应系统理论应用到组织管理中,研究如何建立符合复杂适应系统特征的组织并进行有效管理,是组织管理的重要发展方向;并且,他提出了“复杂适应组织”的概念,它是指在组织成员具有自主判断和行为能力,具有与其他成员和环境交互信息和物质的能力,能够根据其他成员的行为和环境的变化不断调整行为规则,从而使自身及整个组织与环境相适应,如图1所示(Murray, 1994; Holland, 1992, 1999)。

2.3 体系

关于体系亦称系统的系统(system of systems),总宇宙是一个体系,各个星系是一个体系。社会是一个体系,人文是一个体系,宗教是一个体系,多军兵种合成作战是一个体系。大体系里含有无穷无尽的小体系,小体系里含有无尽无量的、可以无穷深入的更小的体系。

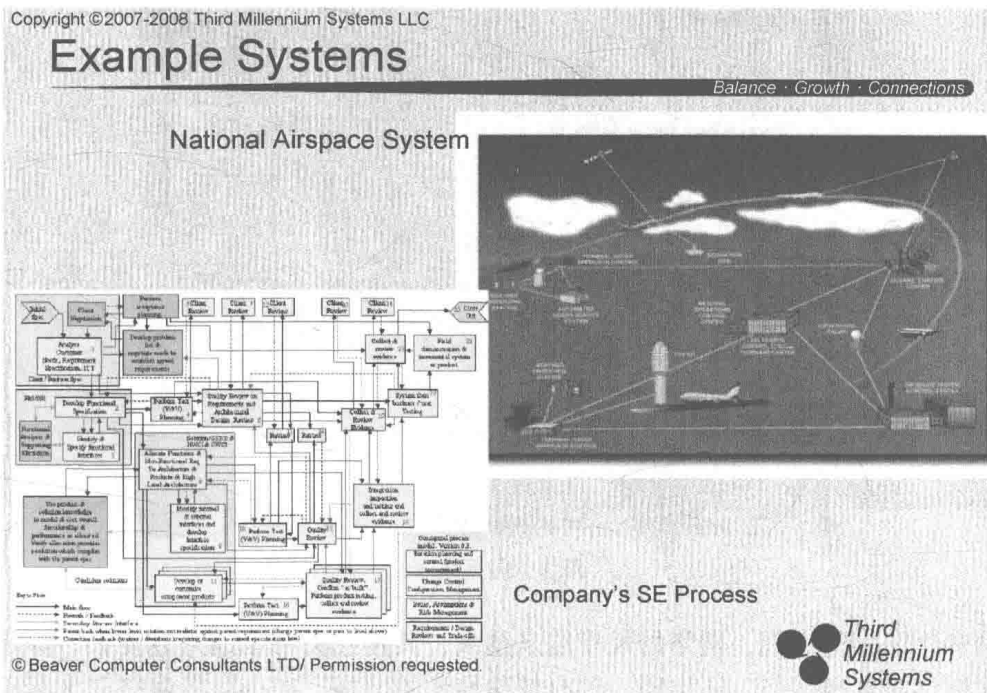


图1 复杂适应系统例子

体系是由一组复杂分系统构成的，它们有5个特点：①分系统操作是独立；②分系统管理是独立的；③系统和分系统都是在演化的；④总系统有突现的行为；⑤分系统在地理上是分散的（图2）。例如，军队组织多兵种大演习就是一个体系（图3）（Sheard, 2006, 2007；Mo, 2008）。

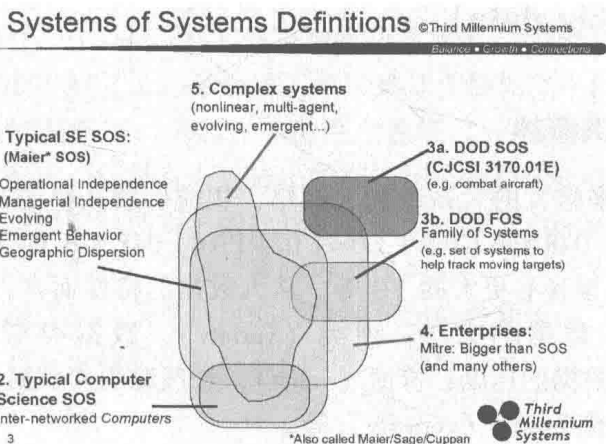


图2 体系的定义

3. 大网络 (big network)

生活中有不少系统会以网络形式出现，如交通网、通信网等。近年来由于互



图3 体系例子

联网和物联网等出现，网络的规模越来越大，其相互关系也越来越复杂，除出现了大网络和复杂网络，还出现一些有特殊性质的网络，如小世界网络和无标度网络等。从物理结构而言，有社会网络、生物网络、经济网络等。

从处理网络手段讲有复杂网络分析 (complex network analysis, CNA), 从时间演化看有动态网络分析 (dynamic network analysis, DNA), 从规模和因素复杂看还有网络的网络 (network of networks) 或超网络 (super network)。国内外应用较多的有社会舆情分析、寻找恐怖分子、供应链等。

4. 大数据 (big data)

4.1 大数据概述

由于出现越来越大的大系统和大网络，相应的就产生越来越大的数据，从 1 PB=1024 TB=1 048 576 GB、1 EB=1 024 PB=1 048 576 TB 到 1 ZB=1024 EB=1 048 576 PB，甚至还有更大的 YB 等。从大数据的特性而言，有所谓的“4V”：容量（volume），数据的大小；种类（variety），数据类型的多样性；速度（velocity），获得数据的速度；价值（value），合理运用大数据，也有认为是可变性（variability）或真实性（veracity）等。

大数据的出现使人们更关注它的存储、转换和加工等，并开发出各种类型的数据库和数据仓库等。另外为了人们能明白地理解它们，对数据的可视化也特别引起人们的关注。例如，美国国土安全部在开发 ADVISE 系统时，就专门开发了软件如 Starlight 系统、In-Spire 系统来显示得到的数据，并且用来找恐怖分子或其他对安全有威胁的人，也可用于找到有用的知识（图 4，图 5）。

而隐知识往往存在于专家脑中，以一种经验的形式存在。能否很好地运用隐知识也是一种智慧的表示，当然智慧更多是为了解决新问题而形成的新知识，即创新。

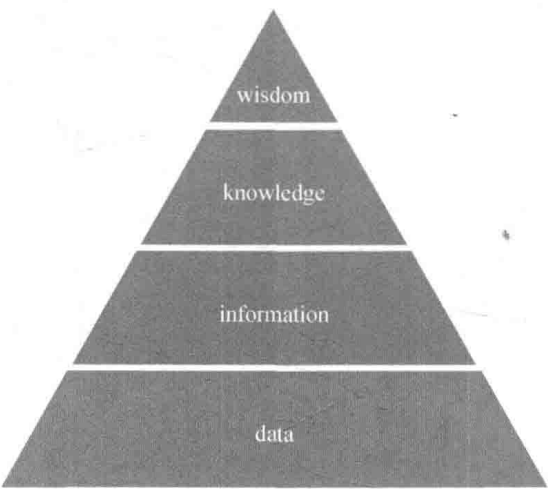


图6 DIKW

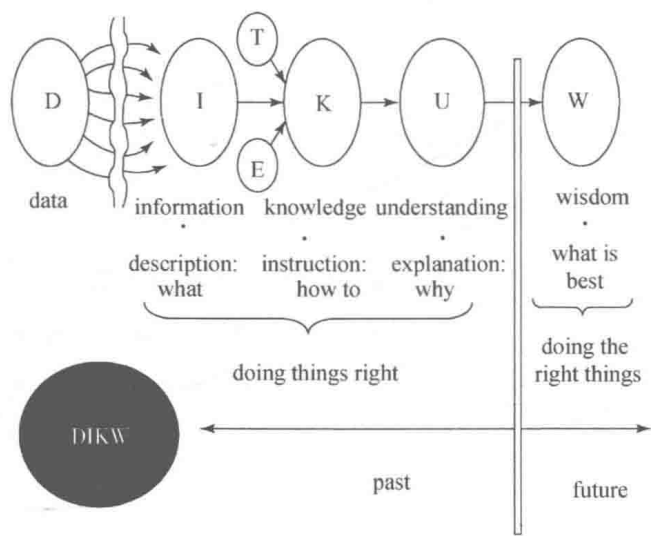


图7 DIKUW (Ackoff)

其实 ADVISE 系统很好地体现了从数据、信息到知识的全过程，里面的核心思想是知识网络图和数据科学。作者也是第一次从这个项目中学到数据科学的名词（图8）。

5. 大解析 (big analytics)

大数据要真有用处，离不开数字解析去深度分析。过去由于系统不大或数据

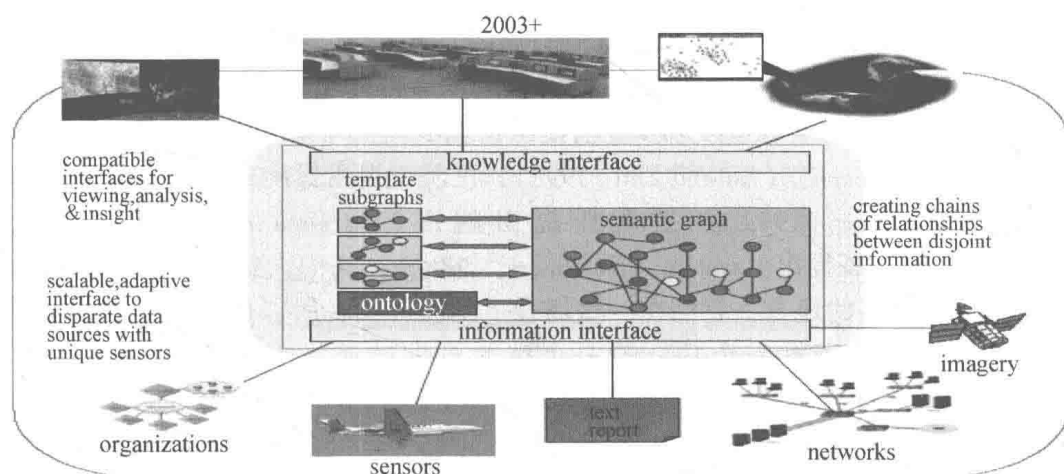


图8 ADVISE 系统的结构

不多往往用的解析工具也比较简单，而现在需要用到新的解析方法越来越多，我们将其分为七类：

第一类与计算方法有关，包括云计算、平行计算等；第二类与统计计算有关，包括多维尺度分析（multidimensional scaling）；第三类与网络计算有关，包括复杂网络分析、动态网络分析、超网络分析；第四类与半定性半定量的计算有关，如软计算、社会计算、超解析（meta-analytics）；第五类与系统科学计算有关，如混沌、分形、动力系统；第六类与优化计算有关，如遗传算法、模拟退火、蚁群优化、粒子群优化算法；第七类与仿真计算有关，包括元胞自动机、多主体仿真等。当然还有些专业方面的计算，如在金融方面，最近区块链的计算影响就很大。

6. 大方法论（big methodology）

做事、搞研究、思考问题都要用方法论，方法论中最大的分界是以不断将事物细化、分解的还原论和从整体出发的系统论。作为系统论的进一步代表是运筹学和系统工程的方法论，其中最著名的是在 20 世纪 50 ~ 70 年代人们常用到的霍尔系统工程方法论。进入 20 世纪 80 年代人们发现原来用的方法论有很大局限性和不足，于是出现了一批新的方法论，为了区别起见，人们把以前以主要处理物的系统为主的方法论叫做硬系统方法论，而处理具有更多人的因素在内的系统用软系统方法论称之，而且他们很快就出现了一大批，如生存系统模型（VSM），软系统方法论（SSM），战略假设表面化与检验（SAST），对话式计划（IP），战略选择发展与分析（SODA），批判式系统启发（CSH），战略选择（SC），问题结构建模（PSM），总体系统干预（TSI），超对策（hypergame），亚对策（metagame）（顾基发，唐锡晋，2006，顾基发，2009a）。

而进入 20 世纪 80 年代末和 90 年代初，由于日本、中国等一批系统界著名

学者也提出了带有东方色彩的系统方法论,如榎本义一提出西那雅卡系统方法论(shinayakana system approach)(Sawaragi, et al., 1987),钱学森等提出从定性到定量的综合集成系统方法论(meta-synthesis system approach)(Qian et al., 1990),受他们的影响顾基发和朱志昌提出物理事理人理系统方法论(wuli-shili-renli system approach)(Gu and Zhu, 1995),王浣尘提出螺旋式推进原理(spiral principle)(Wang, 1992),王众托提出超决策(meta decision)(Wang, 1991),中森义辉提出i-system(Nakamori, 2004)。这一大批方法论其实各有特色和自己应用的地方,它们在一些场合下是互补的,所以作者将其称为大系统方法论。另外还有一个方法论是“摸着石头过河”方法论,简称“Mo” approach。“摸着石头过河”其实是一种中国上层领导为了解决中国社会面临的新而复杂问题而提出的方法论,虽有待完善,但已经走出非常好的一步,现在中央又提出顶层设计与之相配使之更为完整。当然作为一个较为完整的系统方法论尚有不少工作要做,但是他的确是中国领导集体从实践中诞生的富有生命力的中国特色的方法论(辛鸣, 2013; 顾基发, 2014)。最近习近平总书记在B20峰会^①演讲中提到:我们“摸着石头过河”,不断深化改革开放,不断探索前进,开创和发展了中国特色社会主义。“摸着石头过河”其实也是一种试错法(trial and error),我们通过试点不断摸索,改革开放以来已经摸到不少有益的石头,当然也有的石头没有摸对,还需继续探索、完善,有时也不免会付出一些学费,但是我们取得的实际成果,远比一般的系统方法论要多得多(顾基发, 2014)。

7. 大心理 (big psychology)

作为大系统或大网络尤其是社会舆情网络中由于有人的参与,必然要考虑人的心理。不同的心理会影响人们对事物的认知和处理。但对于有很多人参与系统的运作时,单个人的心理也许不会起太大作用,但是对于群体的心理就会起很大作用,这时我们也可叫它大心理。美国就对各个州的幸福感作过大调查,结果发现夏威夷和拿帕是感到最幸福的州而路易桑拿是最不幸福的州。英国莱斯特大学做过在全世界出生在哪一个国家最幸福的调查。其中最前十名分别是丹麦、瑞士、奥地利、冰岛、巴哈马、芬兰、瑞典、不丹、文莱、加拿大,另外一些重要国家分别排名如美国(23)、德国(35)、英国(41)、法国(62)、中国(82)、日本(90)、印度(125)、俄罗斯(167)等(University of Leicester, 2006)。

在社会舆情研究中,对地区群体的心理的研究就显得十分重要。牛文元教授对制作心理版图就十分重视。例如,在研究2016年山东发生的“疫苗事件”中,牛文元小组就专门制作了中国的心理版图。根据网络信息挖掘平

^① B20 即 business 20, 译为 20 国集团工商峰会。

台，筛选 2012 年近 5000 万条社交网络微博的地理标签进行分析，统计各地区正面和负面词汇出现的频率，利用计算机统计分析，绘制出中国 31 个省的心理版图。

山东省“疫苗事件”发生在 2016 年 3 月。由于使用了过期疫苗引起群众的不满，要求政府做出及时的处理。群众在网上发表了不少意见，牛文元小组就舆论趋势与心理版图的对应画出图（图 9）。通过定量分析“疫苗事件”舆情的情绪传播，可将其分为以下三个阶段。

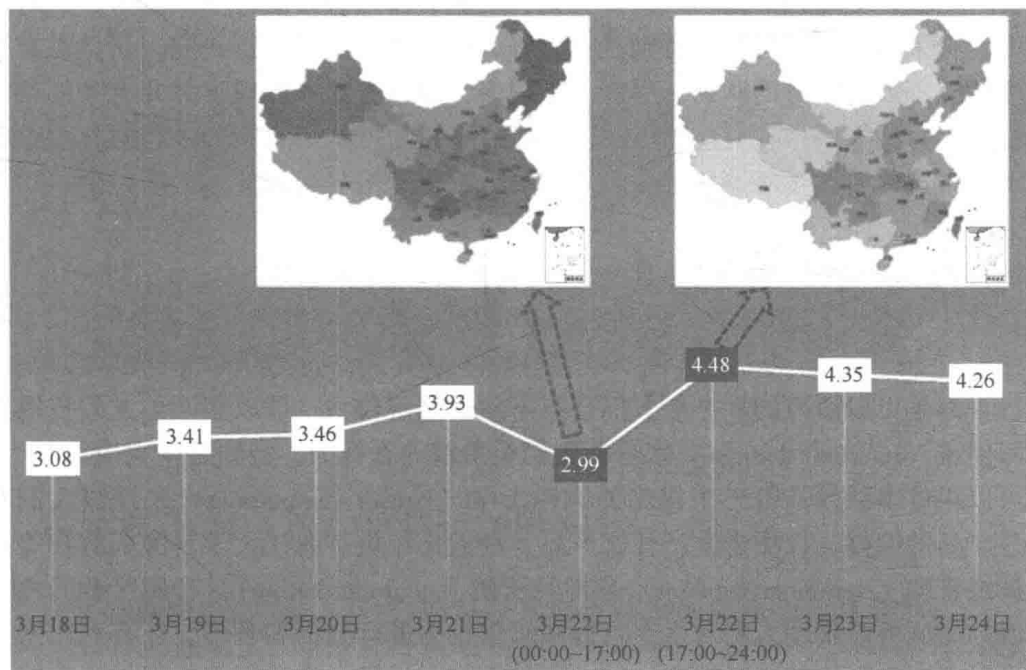


图9 2016年3月18~24日山东“疫苗事件”群众舆情指数与不同阶段的心理版图

第Ⅰ阶段：3月18日，引爆社会关注的“疫苗事件”。一经网络自媒体“澎湃新闻”披露（数亿元疫苗未冷藏流入 18 省份：或影响人命——山东广发协查函），立即引发舆情井喷，网络中怨愤情绪来势汹汹（指数为 3.08 ~ 3.93）。紧接着，舆情“涟漪效应”显现，公众从质疑非法疫苗流向转而矛头直指中国食品药品安全监管不力、国民素质低下、政府不作为。

第Ⅱ阶段：3月22日凌晨，一篇时隔三年的旧闻《疫苗之殇》通过偷换概念、模糊逻辑迅速激荡舆论场，引起恐慌性转发，全国网络社会情绪跌入谷底（指数为 2.99），尤其东北部、东部地区更为突出。然而，与以往舆情洪流不断深化恐慌情绪不同，反驳、对峙“疫苗之殇”的各种信息频出，以批判性思维，辨是非、摆事实，逻辑严密地理性抨击“殇”之误导，网络情绪迅速反转（指数为 4.48）。

第Ⅲ阶段：与此同时，政府相关职能部门纷纷发声，从而形成民间消极舆论场、民间积极舆论场和主流媒体三足鼎立的舆情空间。分化的舆论场伴随公众思辨意识的增强，民众开始回归理性，情绪开始趋稳（指数分别为 4.35 和 4.26）。

(“疫苗事件”舆情的情绪传播由刘怡君供稿)。

8. 大行为 (big behavior)

作为大系统或大网络由于有人的参与,必然要考虑人的心理,而进一步会引起人们表现出他们的行为,并采取一定的行动。例如,等待服务的排队顾客首先有等急了的心理反应,接着就会采取各种行为反应(图10),或者继续等待,或者自行离去,更糟糕的是当人们排了太长的队,而且有可能得不到服务,这时人们可能会采用激烈的行为。特别是一大群人的群体行为更值得研究。当然也有很有秩序的队,如阅兵时的队列,尽管人也很多,却整齐划一表现出井井有条的秩序(图11)。我们曾研究过2010年上海世界博览会的排队行为,如沙特阿拉伯馆面前等待参观的游客多时可达七八千人,中国馆由于事先可以预约,因此中国馆门前并没有什么队,但是进馆后,在馆内却排起了长队(图12)。我们最怕的是由于拥挤失控而发生践踏行为,如2010年的“德国音乐会事件”(有140万人到场,死了19人,伤300多人)和2014年的“上海外滩事件”(发生时约31万人,死了36人,伤49人)(图13,图14)。其实人的行为从个体到群体以至社会行为都是可以研究的。一般人们会有习惯做某些行为,可以用游伯龙教授提出的习惯域(habitual domain)来分析人们会采取什么行为,特别是研究决策行为,也可以用日本教授长町三生提出的感性工学(kansei engineering)来评测人们对一些产品的偏好,以决定生产什么产品。最近几年更有浙江大学马国庆教授等提出神经管理(neuromanagement)和神经营销(neuromarketing),利用传感元件对脑波的测试可以帮助人们对生产行为和销售行为进行分析和管理的。还有利用网上信息可以预测人们的实际行为,如中国科学院大学吕本富和彭赓教授就利用网上的查询信息去预测第二天参观上海世界博览会的游客人数,精度还很高。此外国内外有不少专家利用行人动力学对行人走路的行为进行研究,所以尽管人们的行为很复杂但是也是有规律可循的。



图10 表现各异的队

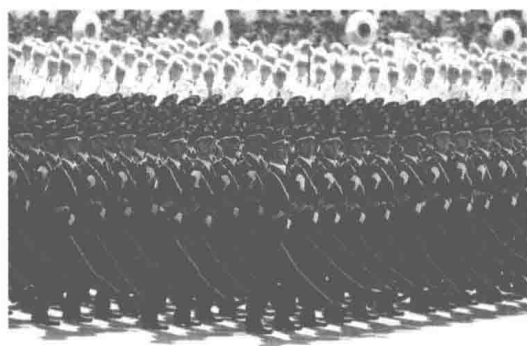


图11 整齐的受阅部队方队

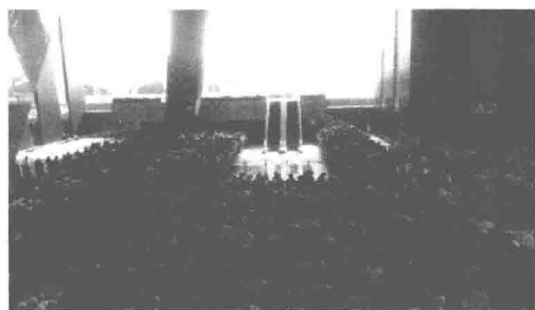


图 12 中国馆馆内的长队



图 13 2010 年 7 月 25 日德国音乐会发生践踏意外



图 14 2014 年 12 月 31 日上海外滩发生踩踏事件

9. 大综合集成 (big meta-synthesis)

为了解决复杂问题人们一方面可以集成各种新的信息和计算技术, 另一方面也会综合各种意见、医案、文献来帮助认识 and 解决问题。西方学者曾提出各种汇总的方法, 其中比较著名的有荟萃分析 (meta-analysis), 在医学界甚为流行。钱学森曾经评论这个方法的不足, 而且进一步提出我们自己的方法, 那就是从定性到定量综合集成的方法, 英文用了 meta-synthesis (钱学森等, 1990)。有意思的是西方也发现荟萃分析的不足, 美国一些学者在踌躇了将近十年终于也走上了综合集成的道路, 只是前面还加上定性两个字, 即定性综合集成 (qualitative meta synthesis) (Sandelowski and Barroso, 2007)。欧盟这些年对专家意见的合成也非常重视, 他们用了一个专业名词——结构化专家判断 (structured expert judgment, SJE), 最近欧洲共同体内不断开展各种有关 SJE 的学术会议来宣传这个思想, 他们的核心思想是现在我们碰到很多问题, 其不确定因素 (uncertainty) 很大, 只好请有经验的专家来判断, 只是应该使用有结构化的专家判断, 以避免过分主观 (Cooke, 1991; Wittmann et al., 2014)。

钱学森等提出的从定性到定量综合集成方法的核心思想是要把数据、信息、知识、模型、专家经验和智慧都综合起来, 这是一个站得高、看得远、集思广的方法论, 但是也应看到缺少一些具体可操作的方法。除了上面提到的荟萃分析、结构化专家判断外, 作者近年来在专家挖掘和知识综合方面分别也作了一些工作 (顾基发, 2010; Gu, 2011; 顾基发等, 2014)。