

复杂网络分析方法在情报学中的应用研究综述

陈迪, 吴亚平, 赵怡然, 姚玲苗, 王继民
(北京大学信息管理系, 北京 100871)

摘要: 复杂网络分析方法为情报学应用研究提供新的方法与视角, 文章首先对复杂网络的背景和分析方法进行系统地梳理和归纳, 然后从科研合作、知识管理、信息传播、信息资源管理、竞争情报等5个方面系统地总结归纳复杂网络方法在情报学中的应用研究现状, 以期对相关研究提供参考。

关键词: 复杂网络; 情报学; 科研合作; 竞争情报; 信息传播

中图分类号: G353.1

DOI: 10.3772/j.issn.1673-2286.2015.09.007

1 复杂网络概述

复杂网络是社会生活中广泛存在的一种现象, 其理论基础是20世纪末出现在系统科学领域的复杂网络理论, 而其根源则可以追溯到20世纪60年代提出的随机图理论。钱学森认为复杂网络是具有自组织、自相似、吸引子、小世界、无标度(无尺度)中部分或全部性质的网络^[1]。复杂网络在人类社会中大量存在, 现实世界的网络既不是有着规则结构的网络, 也不是随机状态下产生的网络, 它具有小世界网络和无标度网络的特性, 与传统的规则网络和随机网络的统计特性有显著的差别。1998年康奈尔大学的Watts和其导师Strogatz发表了关于集体行为动力学研究的文章, 构建了一个小世界网络模型。该模型反映出社会网络的一种特性, 即社会关系网络具有较短的平均路径长度和较大的聚类系数, 这个特性就是“小世界”效应。1999年, Barabasi发表了“随机网络中标度的涌现”, 揭示了复杂网络的无标度性质。复杂网络的小世界效应说明大规模的网络中任意两个节点之间存在最短路径, 从而使得整个网络成为一个可以互相通达的小世界; 无标度特性则揭示了真实网络结构中节点的度值符合幂率分布的特征^[1]。

国际上从1998年左右开始广泛关注复杂网络理论及其应用, 我国在2002年之后关于复杂网络理论的研究文章逐渐增多, 主要是关于数学和物理学的建模问

题、网络性质及其推理论证等^[1]。此后, 有关复杂网络的研究取得了较多的进展, 借助计算机和先进的网络技术, 关于复杂网络的研究范围和角度也从最开始的网络结构及其物理特征发展到如今对复杂网络动力学研究、社团结构与闭包现象、复杂网络中的控制等内容的探讨, 产生了大量的有价值的研究成果。与传统的社会网络分析方法不同, 复杂网络侧重对网络动态过程的研究, 反映的是网络结构的演化过程、网络结构与网络行为之间的交互规律、网络动力学特征等^[2]。因此, 复杂网络比社会网络具有更广的应用范畴, 并逐渐向其它学科领域渗透。

2 复杂网络分析维度与主要指标

节点和边是复杂网络中的基本要素。其中, 节点可以代表具有一定意义的个体或事物, 如科研合作网络中的作者、机构、科研项目等。边代表节点之间的某种关系, 包括有向边和无向边, 其权重代表关系的强弱。对于复杂网络的分析可以从两个维度展开, 一个是以个体为中心的网络分析法, 主要从节点的中心度指标(包括点度中心度、中介中心度、接近中心度), 关系指标(包括关系方向、关系强度), 关键节点及其特性(如结构洞)、聚集系数等角度来分析个体网络特征。另一个是以整体网络为中心的网络分析法。主要从网络的

表1 复杂网络主要分析指标及其意义

主要指标		概念	概念解读
以个体为中心	中心度指标	点度中心度	网络中与某节点之间有关系的其它节点的数目, 定义为该点的点度中心度。该指标通常用来描述某个节点在网络中拥有的权力。
		中介中心度	如果一个节点处于许多其他两个节点之间的最短路径上, 它就具有控制其他两个节点之间交往的能力。通常用来衡量行动者对资源的控制程度。
		接近中心度	一个节点到许多其他节点最短距离之和, 定义为该节点的接近中心度。接近中心度越小, 表明该点在网络中越处于核心地位。
	关系指标	关系方向	在有向图中表示节点间的指向, 例如网页间的链接关系的链入和链出代表了两种关系方向。
		关系强度	根据节点之间的关系强弱, 可以区分为强关系和弱关系。
	关键节点及其特性	结构洞	代表网络中的“中间人”, 掌握多方面的资源, 信息冗余度低。
		聚集系数	与某个节点相连的节点之间关系密度大小, 用来衡量某个节点的“内聚”能力。
以整体为中心	整体结构	网络密度	密度指图中各点之间关系的紧密程度, 密度越高说明该网络中节点之间联系密切。
	局部结构	核心-边缘结构	识别一个网络中处于核心地位的节点和处于边缘地位的节点。
		网络社区	又称“小团体”, 在网络中小团体内的成员节点关系十分密切, 内聚力高。不同“小团体”之间节点关系稀疏。

密度、核心-边缘结构以及凝聚子群等多方面来分析整体网络特征, 如表1所示。

3 复杂网络在情报学中的具体应用

复杂网络分析方法除了本身理论的延伸以外, 还可为其他学科的研究提供了新的方法与视角。情报学是一个具有交叉性质的综合性应用学科, 在其发展过程中也吸收了大量的其他领域的理论和方法, 由于情报学中的诸多问题都可以归结为由各类节点和边组成的网络, 并且网络的规模在不断扩大, 因此复杂网络在情报学学科中也得到了较为丰富的应用。综合分析国内外相关研究论文, 我们梳理并总结了复杂网络在情报学中的应用研究情况, 主要可归纳为以下几个方面。

3.1 科研合作

1963年, 普赖斯^[3]在《小科学, 大科学》一书中对无形学院的科研合作问题进行了研究, 由此开创了科研合作研究的先河, 并指出科研任务与工作量的增大将带来的一系列新的需求, 科学计量学家Glanzel等人的研究也表明, 科学合作势不可挡。目前学界对于科研合作

的概念还没有统一的认识, 张洋等从科研人员、科研过程和科研成果三方面将科研合作的内涵归纳为学者网络、引证网络、合著网络^[4]。然而, 科研合作中产生的网络不仅限于此, 随着学术任务复杂度的提高、跨学科合作需求的增强以及学术交流活动带来的学者之间的交流频繁度的增加, 使得合作现象越来越明显。目前, 科研合作更是体现出宽领域、多类型、高频次等特点, 其中体现出的合作关系也在不断演化, 可以用复杂网络方法来对合作关系进行研究, 来更准确地识别核心作者、核心团队、核心期刊, 分析跨学科趋势、主题变迁等, 从而更好地把握领域的发展方向。具体应用表现在以下几个方面。

3.1.1 合著关系网络

合著关系网可表现为作者合著网络、机构合著网络、学科间的交叉合作网络、区域间的合作网络等。研究合著网络一般从两个维度进行, 一个是从单个科学家的角度, 分析其在整个合著网络中的中心度、聚集系数、平均最短路径等来确定学者的影响程度大小以及与其它科学家的关系强弱^[5]。其中通过分析单个科学家的平均最短路径长度可以找出每年的核心作者。另一个是从团队角度出发, 即根据凝聚子群、子连通图等特征来划分研

究团队,并分析他们的研究兴趣,得出学科研究的主要内容。李纲等人^[6]将科研团队在结构上分为团队领导人、团队核心成员和非核心成员,并通过合著网络整体的中心性分析发现团队领导人,再通过派系划分为若干团队以及核心成员,通过滚雪球的方式发现非核心成员。

3.1.2 引文关系网络

基于论文引用关系的网络分析包含对作者引用关系、机构引用关系、期刊引用关系、学科引用关系等的分析。其中网络中节点的介数、中心度、聚集系数等特征可以衡量学者与论文的影响力,凝聚子群可以分析得出关系较为紧密的期刊或机构。整体网络的密度可以反映科学家之间合作频繁程度和学术交流活跃程度。通过引文网络特征来追溯科技发展史、识别学科新的发展趋势和学术共同体、核心期刊以及作者等^[7]。宋歌等人^[8]通过期刊引用网络将期刊分为“知识源”、“中转站”和“储备库”等三种“角色”。文章^[9]中通过分析期刊引用关系并找到期刊所属的学科大类,探索跨学科现象;通过“出口”和“进口”比例衡量学科影响力以及对其他学科的依赖程度。

3.1.3 共词关系网络

共词网络是关键词共现构建的网络,在同一篇文献中共同出现的关键词之间存在一个边,一个研究领域的共词网络通过复杂网络方法分析可以揭示某一领域的知识结构、研究主题演化、主题扩散等趋势^[10]。Zhu, DH等人^[11]从Scopus数据库中抽取图书情报领域11万个关键词,构筑关键词网络,使用复杂网络分析方法,揭示了关键词网络的小世界效应,并在此基础上发现了学科研究热点。除此之外,通过论文间的共词网络也可以识别论文间的关系强弱,进而识别社区、小团体等。将科研团队成员的关键词与学科领域研究主题相结合,识别科研团队的研究主题^[12]。通过共词网络识别随时间变化的词的更迭,发现“上升词”、“下降词”、“短周期词”、“怀旧词”、“时髦词”等,来反映学科的研究重点与热点的变迁等。

3.2 知识管理

社会网络分析在20世纪90年代以后逐渐应用于其

他领域的实践中,后来被引入管理学领域,应用于战略管理、知识管理和组织行为的研究中。进入信息社会以来,现实世界中知识爆炸式增长,传统的社会网络分析方法无法处理规模庞大的知识,复杂网络分析方法此时发挥了更大的作用。人类社会网络中不仅蕴含着人际关系,还将知识和信息嵌入其中,使得流动性知识作为社会资本通过社会网络进行传递。人际关系有助于信息和知识的获取,知识主体之间的交流可以促进新知识的产生和知识商业化的发展。因此,复杂网络分析方法被引入知识管理的实践中,用于分析网络的规模、强度、模式对知识获取、存储、转移、共享、创新的影响,探讨组织中知识动态演化机制,从而促进组织内知识管理战略的顺利实施。

3.2.1 组织知识管理与扩散

复杂网络分析能够直观地对组织内外部的关系进行量化,找出存在问题。它在组织知识管理中最常见的实践就是通过识别中心人物与边缘人物来发现并监控组织中知识的传递方向,发现组织中知识转移过程中的薄弱环节,从而调整组织结构来做出积极应对。基于关系的研究一直在知识管理研究领域占有重要地位,复杂网络中的社会资本、关系强度、结构洞、信任和互惠等概念是目前探讨知识转移与共享、基于不同关系模式的网络演进等问题的主流观点^[13]。张树人等^[14]将社会网络分析在组织管理中的应用分为多个层次:从组织内外部进行社会网络分析,发现组织中存在的问题,明确组织的发展定位;为组织网络建立动力学模型,预见组织演化的趋势;对引进的管理措施进行效用评估等。

文献^[15]运用小世界网络模型模拟了知识在组织中扩散的过程,发现了知识扩散中存在的小世界特征,得出了“知识在小世界网络中扩散后的平均知识水平最高,知识差异最低”的结论。

3.2.2 知识推送

知识推送技术是复杂网络分析方法在知识管理领域内的另一个应用。这类技术是数字图书馆信息服务的重要组成部分。近年来,随着用户知识需求愈加多元化,主流的知识推送技术并不能很好的满足用户的个性化需求,引入复杂网络分析方法,可以有效解决知识

推送服务中隐性知识推送不足的问题。通过对复杂网络中知识群体的关系和群落进行划分,实现对知识需求之间的相关性判断,识别基于用户关系距离的凝聚子群,采用用户子群划分、用户中心性计算、子群知识共享等策略,可为用户提供满意的隐性知识推送服务^[16]。

3.2.3 知识地图

隐性知识地图的构建是知识管理领域里的一个研究主题。隐性知识地图是一种指示隐性知识来源与去向的图表,也称为专家知识地图^[17]。隐性知识的主体是知识共享网络中的社会资本拥有者,他们在组织中的共享行为可以促进组织内部知识的转移与转换,实现隐性知识的流动和外显化。处在结构洞位置的主体具有较少的冗余知识,掌握较高的社会资本,对其隐性知识进行挖掘并构建知识地图,对于组织内部知识的流通具有重要的作用。

3.3 信息传播

3.3.1 发现传播模式

建立在互联网之上的网络虚拟社区是现代网络技术和信息传播关系的集合体。现代网络平台允许用户发布、评论、转发信息,促使网络社区成员之间以非正式的方式熟悉和了解。从现实社会中衍生的社会网络分析的研究范式同样可以应用在网络虚拟社区中。传统的媒体本位时代下,信息单向流动,而在虚拟社区中的信息传播是双向互动的,是社区内成员关系的一种凝聚^[18]。通过测量复杂网络分析中的各项评价指标,可以获得网络社区中各个成员节点的特性,发现虚拟社区内的组织结构,从而找出社区内信息传播的关键节点和规律,构建虚拟社区中信息传播的一般模式。Liu等人^[19]通过计算最大k核数值下目标节点与节点之间的距离,提出了一种改进的评估节点传播影响力的方法,这种方法可以更好地识别网络中节点的传播效应,对于促进网络中信息流的传播具有重要意义。

3.3.2 舆情监控和预警

网络舆情是通过网络空间传播的由特定事件激发而产生的公众态度、情感、观点、意见等集合^[20]。在过

去,分析舆情走势时,往往比较关注网民的言论,而忽视有多少人支持此意见;往往重视解读文字内容,却忽略了网民互动的社会关系网络^[21]。大数据环境下,零碎信息组成了庞大的复杂网络,网络中用户之间的关注、点赞、转发和评论都会形成不同的关系。整个社交网络舆论场可以构成一个巨大的有向关系网络,每个用户都是一个节点,所有人际关系最终全景式地展示了整个社交网络。个体之间频繁联系,相互影响,一个节点的异动可能会引起整个网络的波动,由关键节点引发的蝴蝶效应甚至会带来整个社会的剧烈反应。由于社交网络平台的开放性以及信息传递的快捷广泛等特征,用户可以自由地参与到热点事件或者国家重大问题的讨论上。一些重大或者突发事件往往会通过社交网络迅速成为网络平台讨论的热点,一些具有影响力的人会成为网络舆论的“意见领袖”,能够引导整个舆论的导向。识别出具有影响力的关键节点,对社会重大舆情事件进行监控和预警,积极做出应对,合理妥善处理矛盾,对于组织或者政府来说都具有重要的意义。

3.4 信息资源管理

传统的社会网络分析方法已经无法应对来势汹涌的大数据浪潮,复杂网络分析应运而生。复杂网络理论涉及到图论中的连通度、节点间的距离等属性,这些属性与网站链接、网站排名算法、信息检索、资源配置等有着非常密切的关联。

3.4.1 网站链接分析

随着互联网信息的爆炸式增长,复杂网络分析方法开始在互联网上广泛应用,搜索引擎公司为了找出能够与用户提问式匹配的最佳页面,将每个网页视为社会网络中的一个节点,而节点之间存在指向与被指向的关系,从而形成一个巨大的有向网络图。这些由链接构成的聚簇或者核心节点就是整个网络中的核心资源。目前互联网搜索引擎巨头Google公司采用的PageRank算法就是基于网络分析方法改进而成的网络链接排序方法。裴雷等^[22]探讨了社会网络分析在Web中的应用,针对使用Google的网络链接排序、超链引导的主题搜索、页面信息嵌入和过滤等方法进行网站链接分析和网站排名的不同效果,得出网络链接算法的效果要好

于超链引导的主题搜索方法。由于传统情报学领域中的信息离散现象具有“马太效应”的特征，Web资源上也具有相类似的现象，因而裴雷认为可以利用Web上的主题分布有一定短范围的特征，发现局部相关的其他资源。

3.4.2 信息资源挖掘与分配

复杂网络分析在Web资源挖掘方面也有现实的应用意义。大数据环境下互联网页面中有着十分丰富的信息，网页之间相互应用，链接关系错综复杂，利用复杂网络分析方法将这个庞大的网络视为有向图，找出核心节点和关键聚簇，从网络中进行相关的主题挖掘，对于Web资源挖掘具有很强的现实意义。

在分配有限资源时，如何充分使用现有资源是经常需要考虑的问题。使用复杂网络分析方法，找出在资源使用中具有决定作用的关键节点，赋予其较多的资源，对于剩余的节点按照其点度中心度高低进行资源分配，从而能够使得整个网络处在一种高效利用资源的状态，实现对现有资源的合理分配。

3.5 竞争情报

王知津等^[23]将竞争情报定义为：为达到竞争目标，合法而合乎职业伦理地搜集竞争对手和竞争环境的信息，并转化为情报的连续系统化过程。正可谓“知己知彼，百战不殆”，情报对于企业极其重要，竞争情报是企业的战略，是企业逐鹿市场、夺取商机、积极应变、以智取胜的基本战略。情报是经过加工后的知识，经常是在人脑中存在的隐性知识，竞争情报的搜集、加工、分析都离不开人的参与，竞争情报中的社会网络指的是人际情报网络。包昌火将人际情报网络定义为应情报活动的需要而构架的一种人际网络，是情报从业者获取、分析和传播非公开信息和隐性知识的重要平台。人际情报网络在竞争中的应用可以归纳为以下三方面：

3.5.1 整体行业分析，识别竞争对手

通过整体分析行业内部的核心-边缘结构、核心企业等来识别强劲的、潜在的竞争对手，并深入分析其经营模式、合作伙伴等。此外除了本行业，也要分析有密切合作和制约性的其它行业或组织，如供应链的上方

和下方企业，政府政策环境等，把握整体行业特征和环境带来的机遇与挑战，并识别关键强劲对手、潜在对手和利益相关者。

3.5.2 情报源挖掘

确定好竞争对手后就要寻找情报源，传统的情报源挖掘中主要通过文献调研和信息检索技术去获取情报，具有滞后性、零碎性和低可信度等特点。只强调了“物化”载体，忽视了人的作用，限制了情报源的范围。人际情报网络可以拓展传统竞争情报分析的视角，突破“物”的范围，从人的角度寻找情报源，更能够提高获得情报的可信度。曹超等人^[24]系统地将竞争对手情报源的挖掘划分为三个角度，基于整体网络、基于个人网络、基于第三方网络。通过分析对手企业组织中的核心-边缘结构、凝聚子群等来识别关键的情报源，分析员工间的工作网络、友谊网络等来拓宽情报搜集渠道，分析有紧密关联的第三方，从而拓宽情报源的范围。

3.5.3 情报战略制定

通过分析自身企业组织内部的人际情报网络，可以识别情报源的分布是否合理，情报的传播途径是否有泄漏隐患，分析出关键“节点”对其情报权利进行控制，减少“结构洞”节点，对内部“小团体”的效应进行控制，降低情报泄露的几率，从而优化人际情报网络结构。在情报泄露时也可以找关键的节点进行有效控制。

4 结论与展望

信息的爆炸式增长带来了复杂的数据网络关系，复杂网络分析方法能够量化大型网络的基本特征，并识别网络演化的机制和信息扩散的模式，在科研合作、知识管理、信息传播、信息资源管理、竞争情报中已得到初步应用，并取得较好成果。复杂网络方法对于优化信息资源配置、促进知识传播与共享、促进信息资源开发与利用有较大意义，相信会在未来会有更大的发展空间。

参考文献

[1] 李晓辉, 徐跃权. 复杂网络理论的情报学应用研究[J]. 情报资料工作,

- 2007(3): 9-13.
- [2] 罗家德. 社会网分析讲义[M]. 北京: 社会科学文献出版社, 2005: 202.
- [3] Price D. Little Science, BigScience and Beyond[M]. New York: Columbia University Press, 1963: 301.
- [4] 张洋, 谢齐. 基于社会网络分析的机构科研合作关系研究[J]. 图书情报知识, 2014(2): 84-94.
- [5] Alan S, Gary K, Cathal G, et al. Analysis of Papers from Twenty-Five Years of SIGIR Conferences: What Have We Been Doing for the Last Quarter of a Century 2003[J]. ACM SIGIR Forum, 2003, 37(1).
- [6] 李纲, 李春雅, 李翔. 基于社会网络分析的科研团队发现研究[J]. 图书情报工作, 2014(7): 63-70.
- [7] 宋歌. 社会网络分析在引文评价中的应用研究[J]. 图书情报工作, 2010(14): 16-19.
- [8] 宋歌, 叶继元. 基于SNA的图书情报学期刊互引网络结构分析[J]. 中国图书馆学报, 2009(5): 27-34.
- [9] Larivière V, Sugimoto CR, Cronin B. A bibliometric chronicling of Library and Information Science's first hundred years[J]. Journal of the American Society for Information Science and Technology, 2012, 63(5): 997-1016.
- [10] 张斌. 共词网络的结构与演化: 概念与理论进展[J]. 情报杂志, 2014(7): 103-108.
- [11] Zhu D, Wang D, Hassan S U, et al. Small-world phenomenon of keywords network based on complex network[J]. Scientometrics, 2013, 97(2): 435-442.
- [12] Hiemstra D, Hauff C, Jong F D, et al. PAPER SIGIR's 30th anniversary: an analysis of trends in IR research and the topology of its community[J]. AcmSigir Forum, 2007, 41(2): 18-24.
- [13] 向希尧, 蔡虹. 基于专利引用的社会网络分析在知识管理研究中的应用[J]. 管理学报, 2012, 9(4): 562-569.
- [14] 张树人, 刘颖, 陈禹. 社会网络分析在组织管理中的应用[J]. 中国人民大学学报, 2006(3): 74-80.
- [15] 胡峰, 张黎. 知识扩散网络模型及其启示[J]. 情报学报, 2006(1): 109-114.
- [16] 黄微, 高俊峰, 王晨, 等. 基于社会网络分析的隐性知识推送服务方法研究[J]. 现代图书情报技术, 2014, 30(2): 48-54.
- [17] 吴才唤. 社会网络分析在隐性知识地图构建中的应用[J]. 图书馆, 2010(1): 48-51.
- [18] 翟延祥. 基于社会网络分析的网络社区信息传播模式研究[D]. 南京: 南京航空航天大学, 2011.
- [19] Liu J G, Ren Z M, Guo Q. Ranking the spreading influence in complex networks[J]. Physica A Statistical Mechanics & Its Applications, 2013, 392(18): 4154-4159.
- [20] 王国华, 冯伟, 王雅蕾. 基于网络舆情分类的舆情应对研究[J]. 情报杂志, 2013, 32(5): 1-4.
- [21] 唐涛. 基于大数据的网络舆情分析方法研究[J]. 现代情报, 2014, 24(3): 3-6.
- [22] 裴雷, 马费成. 社会网络分析在情报学中的应用和发展[J]. 图书馆论坛, 2007, 26(6): 40-45.
- [23] 王知津, 樊振佳. 基于社会网络分析的企业竞争情报战略[J]. 图书情报知识, 2007(6): 5-10.
- [24] 曹超, 盛小平. 社会网络在竞争情报源挖掘中的应用研究[J]. 情报理论与实践, 2009, 32(3): 60-62.

作者简介

陈迪, 女, 1992年生, 硕士, 研究方向: 科学评价、Web数据挖掘。

吴亚平, 女, 1991年生, 硕士, 研究方向: 科学评价、Web数据挖掘。

赵怡然, 女, 1993年生, 本科, 研究方向: 统计与数据分析。

姚玲苗, 女, 1994年生, 本科, 研究方向: 统计与数据分析。

王继民, 男, 1966年生, 博士, 北京大学信息管理系教授, 研究方向: 搜索引擎与Web数据挖掘、科学评价、社会网络分析等, 通讯作者, E-mail: wjm@pku.edu.cn。

A Review of the Application of Complex Network Analysis Method in Information Science

CHEN Di, WU YaPing, ZHAO YiRan, YAO LingMiao, WANG JiMin
(Department of Information Management, Peking University, Beijing 100871, China)

Abstract: Complex network analysis methods provide new approaches and perspective for information science. This paper summarizes the background and methods of complex network, then shows its application in research cooperation, knowledge management, dissemination of information, information resource management and competitive intelligence, hoping to provide a reference for the relevant research.

Keywords: Complex Network; Information Science; Research Cooperation; Competitive Intelligence; Dissemination of Information

(收稿日期: 2015-07-14; 编辑: 王立学)