

基于社会网络分析的数字图书馆资源聚合实证研究^{*}

贯君¹, 王雨¹, 吴海媛²

(1. 吉林大学管理学院, 长春 130022; 2. 大连大学图书馆, 大连 116622)

摘要: 分析数字图书馆资源的社会网络关系, 结合社会网络分析的指标, 按照资源聚合的模式, 利用中心性、凝聚子群、核心-边缘等测度分别验证资源聚合的结果及效果, 以优化资源的质量、提高资源的可用性并增强资源的关联性。

关键词: 社会网络分析; 数字图书馆; 资源聚合

中图分类号: G250.7

DOI: 10.3772/j.issn.1673—2286.2014.06.002

1 研究背景和相关工作

随着信息技术飞速地进步, 信息数量也随之发生了突飞猛进的增长, 探索新型的信息存储和传播方式作为信息时代无法抗拒的趋势, 给网络数字信息的发展带来巨大的影响。如何发现、捕获和挖掘有效的信息资源一直是信息科学和计算机科学领域的研究热点。为优化信息资源的加工、存储、检索、传输和利用, 20世纪90年代数字图书馆(Digital Library, DL)^[1]应运而生, 将信息资源利用数字技术存储, 成为“没有围墙”的知识网络系统^[2]。然而, 以图像、文本、音频、视频为载体的数字图书馆馆藏资源众多, 如何将数字资源聚合, 优化数字资源配置, 完善知识结构, 提高资源利用率, 成为目前数字图书馆建设面临的重要问题。通过分析资源特征间和资源利用过程中的社会网络关系, 对于解决上述问题, 实现数字图书馆资源的聚合具有重要意义。

对此, 本文从对数字图书馆资源聚合的宏观、中观、微观分析入手, 基于社会网络分析的视角, 将数字图书馆资源作为社会网络分析的核心要素, 对资源的内部特征、外部特征、效用特征进行深入的研究, 利用社

会网络分析, 建立起资源间及资源利用过程中的关联, 以达到资源聚合的目的。

2 数字资源聚合方式

数字图书馆资源聚合是为了实现各类资源之间的有效关联, 从而为用户提供优质、高效的知识服务。数字图书馆资源聚合建立在知识组织的基础之上, 以知识组织的原理为指导, 主题地图作为知识组织的一种有效工具, 能够对数字图书馆资源之间的关系进行有序的揭示, 并且能够有效地定位信息资源, 建立信息之间的关联路径^[3]。在数字资源加工处理的程度及技术方法方面, 数字图书馆资源聚合的三个层面为数据聚合、信息聚合和知识聚合, 对数字图书馆资源进行数据聚合可采用基于导航系统的网络资源聚合模式; 信息资源聚合通常都采用基于链接系统的网络资源聚合模式; 知识聚合可以利用基于知识本体的网络资源聚合模式。

数字图书馆资源聚合表现为统一检索与资源调度两个方面。前者指实现所有资源的跨库跨平台检索, 并将符合检索要求的信息一次性返回给读者, 它反映的是

^{*} 本研究得到国家社会科学基金重大项目“基于语义的馆藏资源深度聚合与可视化研究”(编号: 11&ZD152)和国家自然科学基金项目“语义网络环境下数字图书馆资源多维度聚合与可视化展示研究”(编号: 71273111)资助。

资源的统一性;后者指通过OpenURL技术实现异种资源的动态链接,并通过已知的信息点获取与之相关的其他资源,它反映的是资源的关联性^[4]。资源聚合的应用在各个领域越来越受重视,很多网站及图书馆均利用不同的技术和方法对资源进行资源聚合,推出知识搜索服务,多维度呈现知识内容。

目前,数字图书馆领域具有代表性的是,CALIS专门开发出了一组数字图书馆门户的聚合应用系统^[6]。王文清在《构建高效数字图书馆综合服务门户》一文中提到:将数据库中的内容数据通过筛选、复制、导入等技术聚合起来,建立联合资源仓储,实现资源的内容聚合。CALIS正在研制“联合资源仓储管理与服务系统”,具有多种资源的规划管理功能,提供对多种类型资源的跨库检索和多种分类体系的资源导航功能。CALIS的两个子项目《CALIS高校专题特色数据库》和《CALIS高校学位论文全文数据库》都属于资源聚合型门户。

3 基于社会网络分析的数字图书馆资源聚合实证分析

本文从资源特征关联维度、资源利用过程关联维度以及用户需求维度入手,对图书情报与数字图书馆领域的文献资源进行社会网络分析,利用作者互引关系、作者合作关系、作者-关键词共现关系等路径建立资源聚合模式,形成作者-关键词之间、作者之间、合作者之间的社会网络,通过社会网络关系可视化地展示,发现作者间引用情况,识别核心资源,挖掘作者间的相同的研究主题,揭示外在的合作群体。

3.1 社会网络分析方法的应用

使用Ucinet软件进行社会网络分析时,主要是围绕准备数据、数据处理和数据分析三个步骤进行。为更直观地了解分析步骤,笔者绘制了主要操作步骤图,如图1所示。

3.1.1 样本选取

知识节点间的关联程度就是对资源聚合的表征,聚合的群体便是连接较强的节点的集合。聚合群体的阈值要求越高,聚合深度就越高,聚合群体内部的知

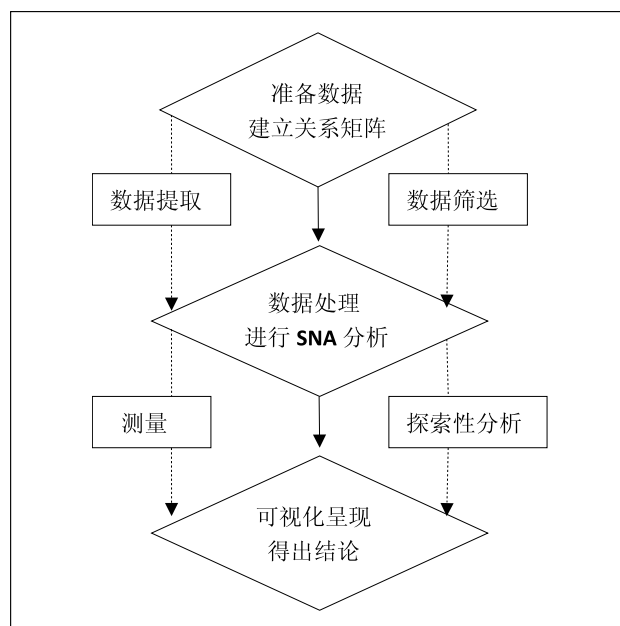


图1 社会网络分析流程图

识节点就越有聚合的可能。所以聚合深度就是结合具体研究领域,将知识节点聚合为一个群体的聚合强度阈值。

本文研究样本来自于“中国知网(CNKI)”数据库,选定该数据库中“图书情报与数字图书馆”下属的“情报学、情报工作”学科领域。为动态地了解该领域近5年的发文以及近5年该领域的主要研究方向、内容,本文选取样本的发表时间为2008-2012年(由于中国知网收录的文章在时间上会存在一定的滞后性,所以笔者在2013年获取数据时,为保证数据的准确性,并未选取2013年的数据作为研究样本)。综合以上检索条件,共检索到论文总数为12,815篇,剔除评论性报道文章、索引文章并对论文去重后,共获取论文11288篇。

3.1.2 数据处理

根据上述获取的11288篇论文,通过CNKI“导出/参考文献”功能,提取作者、题名、文献来源、年、关键词、机构等字段,保存到Excel文件中。统计过程中发现,多数文章存在作者合著的现象,对此,笔者运用Access数据库及SQL Server数据库软件,将作者字段按照第一作者、第二作者、第三作者分别提取、统计,最后将统计数据合并,得出共有10011位作者。同时,也对关键词按照相同方法进行了统计,将统计数据清洗、去重合并后,得出共有23613个关键词。

3.1.3 核心作者及高频关键词选取

普赖斯在其代表名著《小科学,大科学》一书中指出:在同一主题中,半数的论文为一群高生产能力作者所撰,这一作者集合的数量约等于全部作者总数的平方根^[7],这就是著名的普赖斯定律。根据普赖斯定律,可获得发文在17篇以上的高产作者有48位。同时,根据数据处理结果,选取了出现次数 ≥ 24 次的80个高频关键词用以对作者-关键词关系、共词关系的社会网络进行分析,辅助学科主题、学科领域的定位。依据上述统计结果,选取发文在17篇以上的48位高产作者作为互引关系的研究对象。国内图书情报学领域中发文量最多的是邱均平、王知津、毕强、武夷山等人,他们在该领域具有很强的学术影响力。除此之外,筛选和清理选取的关键词,除去“发展”、“相关性”等实际意义较宽泛的关键词,合并意义相同的关键词,最终确定80个高频关键词,用以构建作者和关键词的共现矩阵。

3.1.4 构建关系数据矩阵

(1) 核心作者互引矩阵

根据确定的核心作者群,通过中国知网的高级检索平台进行检索,在输入内容检索条件处选择参考文献,输入作者B,在输入检索控制条件处选择作者,输入作者A。汇总48位核心作者的互引情况,通过检索确定核心作者相互引用的频次,在Excel中建立核心作者互引矩阵,进而根据互引矩阵分析核心作者互引网络。

同理,构建作者合作关系矩阵,在输入检索控制条件处选择作者A“且”作者B。通过检索获得48位核心作者合作的关系,生成的矩阵为2值矩阵,存在合作关系即为“1”,不存在合作关系即为“0”。由于合作关系不存在方向性,所以所构建的矩阵为无向对称矩阵。

与此同时,构建作者-关键词2-模关系矩阵,在“输入内容检索条件”中输入关键词,并在“输入检索控制条件”中输入作者检索,即可在页面左边导航栏找到“关键词”这一项,便可知相对应的关键词出现次数。在此基础上,构建多值矩阵。作者-关键词关系形成的是2-模网络,故其关系矩阵是2-模长方形矩阵。在后续分析过程中须利用Ucinet中Transform>Bipartite路径将其转换为2-模二部方阵,即作者-作者矩阵和关键词-

关键词矩阵。

3.2 构建社会网络关系图谱

3.2.1 作者互引关系网络图谱

根据核心作者互引矩阵值,构造作者互引矩阵图,并通过NetDraw将网络关系可视化地表现出来。如图2所示,其中每一个圆形的节点代表核心作者,有向的连线表示作者间的引用关系,箭头指引的方向即是从引用者指向被引用者。从作者互引关系网络图谱不难发现,引用次数多的作者连线较粗,箭头较大,证明该学者处于互引关系的核心位置,其影响力较大;节点大的说明该学者中心度较高,自身控制能力较强。

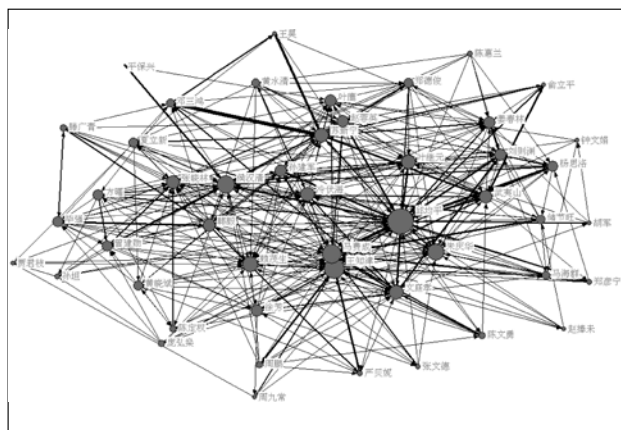


图2 作者互引关系图

3.2.2 作者合作关系网络图谱

根据构造的作者合作关系矩阵值,通过NetDraw将网络关系可视化地表现出来。如图3所示,其中每一个圆形的节点代表核心作者,节点之间的连线表示作者间的合作关系,因作者间的合作关系是双向的,故箭头的指向是双向的。从作者合作关系网络图谱不难发现,合作发文次数多的作者间连线较粗,反之亦然。其中,武夷山和俞立平间的连线最粗,表明这2位作者间的合作关系最强,从分析数据可知,两者共合作17次。作者间的合作关系表明哪些作者的写作方向是一致的,合作强度决定着在该学科领域资源聚合中是否此类合作者的发文可以代表该学科的研究方向和学科发展的分支。

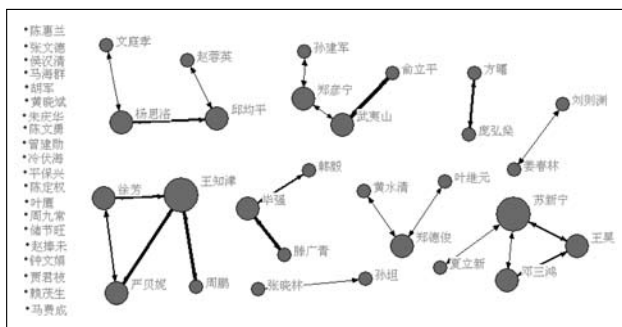


图3 作者合作关系图

从作者合作关系图可以清晰地看出,该网络由48位核心作者构成,形成合作关系的子网有9个,其他没有关联的作者代表单独发文,彼此间没有合作关系。对于子网分析可分为四个类型:单点型、双核型、核心型、桥梁型^[9]。本文核心作者共有20位作者属于单点型,独立发文,彼此间无合作关系。有6位作者是两两合作发文,属于双核型。处于核心位置的王知津、苏新宁形成的是核心型子网,二者处于核心位置,与该子网的其他节点均有连接。在本文的作者合作关系中,并未有桥梁型的子网,表明作者间的合作关系不够紧密。

3.2.3 作者-关键词关系网络图谱

在Ucinet中,通过Transform>Bipartite路径,将2-模长方形矩阵转换为双向矩阵,以便于下一步形成可视化知识图谱。在此基础上,通过“NetDraw>File>Open>UCINET dataset>2-Mode Network”路径导入双向矩阵,生成2-模图。

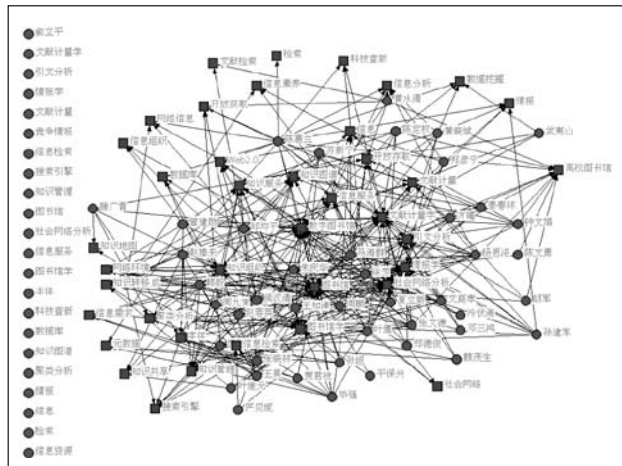


图4 作者-关键词关系图

从图4可以看出,作者和关键词之间的关系表现出一定的聚类性。图左侧一列体现的是不具备关系的作者和关键词,关系图中作者通过关键词聚类,有相同关键词的作者均指向同一关键词。

3.3 数字图书馆资源社会网络关系的拓展分析

3.3.1 作者互引网络分析

利用Ucinet分析社会网络关系时,为保证各项指标测度的准确性和一致性,均将关系矩阵转化为二值矩阵。White H.D在《*Authors as Citers over Time*》^[10]一文中将被引用群体分成两部分,以2次为分界线,通过“Ucinet/Transform/Dichotomize”路径转化成二值矩阵后,在弹出的窗口中输入原始数据集,阶段运算符选择“Greater Than or Equal”截断值处添入“2”,便可得到二值矩阵,用以对中心度、密度、距离和凝聚子群的分析。

(1) 中心度分析

“中心度”是社会网络分析的重点测量指标之一。个人或组织在其社会网络中具有怎样的权力,或者处于怎样的中心地位,通过对中心度的分析便可一目了然。个体的中心度(Centrality)测量个体处于网络中心的程度,反映了该点在网络中的重要程度^[1]。

① 点度中心度

在Ucinet中调入作者二值互引矩阵数据,按照以下路径Network>Centrality>Degree进行点度中心度分析。从结果可以看出,点出度较高的为1号“邱均平”、35号“文庭孝”、2号“王知津”,点出度均高于10,表明他们引用其他作者文章的程度较高;点入度较高的为1号“邱均平”(点入度为27)、45号“马费成”(点入度为18)、2号“王知津”(点入度为13)、32号“苏苏宁”(点入度为12),表明他们处于被引用的核心地位,发表的文章具有较高的参考价值及学术影响力。通过对点出度和入度的分析,作者之间的引用关系一目了然,为相关资源的聚合奠定了基础。基于引用关系的聚合,使得用户的检索结果不再是单一的资源,而是引用或被引用程度高的资源均能得以展现,实现资源检索的智能化。

② 中间中心度

通过获取图情领域作者互引关系的中间中心度,判断学者是否处于核心位置,以此定位此学者在图情

领域的学术影响力。1号“邱均平”、2号“王知津”、34号“苏新宁”、8号“侯汉清”是中间中心度比较高的。也就是说上述4个学者在互引关系中处于核心位置,其他学者获取信息对这4个学者的依赖性比较大。由此可知,1号、2号、34号、8号学者在图情领域学术圈中影响力比较大,通常他们的文章更能引起其他学者或用户的关注,亦是引用率较高的文章。网络的中间中心势为34.78%,比例较高,即在作者互引网络中大部分节点需要别的节点作连接,才可以获得所需资源。

③ 接近中心度

将上述“二值作者互引矩阵”输入到Ucinet软件中,根据Network> Centrality> Closeness路径进行接近中心度分析。

从接近中心度的结果可以看出,左列学者编码就是根据资源的引用到其他各点的便捷难易程度排序的,越靠前则越不容易受到他人控制,越独立。从获取信息的难易程度而言,其顺序略有变化,排在前几位的即32号“储节旺”、47号“黄水清”、43号“郑德俊”、7号“张文德”。越靠前的越容易引用他人,不易受人控制,独立性强。

(2) 凝聚子群分析

在Ucinet中,根据路径Network> Subgroups> cliques对作者互引情况进行派系分析,为能更鲜明地展示凝聚子群的效果,在Ucinet输入数据集时,将最小值设置为5,得出结果。从总体上看,互引网络中互动性强的作者并不多,6个派系包含的核心作者共有12位。

(3) 距离

分析社会网络关系,主要目的在于揭示各要素间的联系是否紧密,是否有利于资源间的聚合。描述社会网络特征的指标主要是密度和距离。其中,在资源聚合领域主要应用特征路径长度指标来刻画网络的小世界特征^[12]。相对于中心度分析而言,小世界分析更加深入和全面地描述了资源聚合的紧密程度。特征路径越短,小世界的程度越高。小世界的程度越明显,资源聚合的紧密程度越大。

特征路径长度指网络中所有节点之间的平均最短距离,表现了网络结构的全局特征。首先将关注网络矩阵输入到Ucinet软件中,沿着Network→Cohesion→Distance路径即可计算出48位成员的关心网络距离。各用户之间的平均距离 $L=2.364$,即在作者互引关系中,互引作者中最多经过3个中间人就可以到达下一位作者,平均经过不到1人就可以和其他作者产生联系^[13]。鉴于

此,图情领域的作者互引关系具有小世界效应,作者之间交流越紧密,信息传播越广泛。

3.3.2 作者合作网络分析

(1) 网络中心度分析

利用Ucinet软件,对作者合作网络的三种中心性进行测算,如表1所示。

表1 核心作者合作关系的中心度测度

排序	姓名	nCloseness	姓名	nCloseness	姓名	nCloseness
1	王知津	1.785	苏新宁	0.185	苏新宁	2.222
2	苏新宁	1.537	王知津	0.185	王知津	2.222
3	邱均平	1.204	邓三鸿	0.185	邓三鸿	2.221
4	苏新宁	1.036	邱均平	0.185	邱均平	2.221
5	毕强	0.968	徐芳	0.185	徐芳	2.221

从表1可以看出中心性靠前的是王知津、苏新宁、邱均平,三位在我国图书情报学领域的合作网络中处于核心地位,与其他作者连接最近,同时在该领域具有较大的影响力,处于中间位置,是图书情报学领域核心作者合作网络中的核心人物^[14]。

(2) 凝聚子群分析

通过作者合作关系知识图谱可以发现,该群体是由多个相互之间联系不太紧密的小群体构成。这些小群体便形成了建立在可达性基础上的n-派系。通过路径Network> Subgroups> N-cliques分析该无向网络中的n-派系,可发现共形成8个小团体。如表2所示。

从对作者合作网络的凝聚资源分析结果可以看出,共有8个派系,包含22位核心作者,表明作者间的合作

表2 凝聚子群群体

1	邱均平	杨思洛	文庭孝	
2	邱均平	杨思洛	赵蓉英	
3	王知津	徐芳	周鹏	严贝妮
4	毕强	韩毅	滕广青	
5	武夷山	孙建军	郑彦宁	
6	武夷山	郑彦宁	俞立平	
7	邓三鸿	苏新宁	王昊	夏立新
8	叶继元	郑德俊	黄水清	

性较强。从派系分布的分析,便可知哪些作者处于同一个子网络。而子网络作为整体网的构成,每个子网络的形成便意味着资源内部聚类,最终形成整体网络上资源的聚合。

(3) 距离和密度

在Ucinet中沿着“Network→Cohesion→Distance”路径测算出合作网络中核心作者间的距离,其中Average distance = 1.424, Distance-based cohesion = 0.024。即在作者合作网络关系中,合作作者中最多经过2个中间人就可以到达下一位作者,平均经过不到1人就可以和其他作者产生联系。鉴于此,图情领域的作者合作关系具有小世界效应,作者之间合作关系紧密。

同样,沿着“Network→Cohesion→Density”路径测算出图情领域核心作者合作网络的密度为0.1020。由此可以看出,核心作者整体之间的合作并不频繁。这并不利于资源的聚合,当用户按照核心作者的合作关系查找时,得到的结果可能是松散的、关联性弱的。

3.3.3 作者-关键词网络分析

(1) 2-模量表因子分析

因子分析最早由英国心理学家C.E.斯皮尔曼提出,是指研究从变量群中提取共性因子的统计技术^[15]。因子分析主要是为了发现多个变量中共享的一个或多个因子,将相同本质的变量归入一个因子,可减少变量的数目,还可检验变量间关系的假设。

通过结果可以看出,前2个因子解释掉总方差的32.3%,前3个因子解释掉总方差的43.7%。再观察旋转因子的负载值,因子一可以解释为“邱均平因子”,因子二可以解释为“王知津因子”,因子三可以解释为“毕强因子”。此外,每个关键词在各个关键词上的负载值亦都有所体现。例如,孙坦的发文中使用“文献计量学”的次数很多(负载值为3.700),毕强使用“文献计量学”的次数很少(负载值为0.090)。

(2) 2-模网络的核心-边缘分析

核心-边缘理论最早是在1966年J.R.弗里德曼提出的,将其定义为:“一种解释经济空间结构演变模式的理论,它试图解释区域如何由互不关联、孤立发展,再变成彼此联系、发展不平衡,最后变成相互关联的、平衡发展的区域系统^[16]。”在社会网络分析中,核心-边缘理论是一种理想化的结构模式。从某种程度上看,核心

区域便是作者和关键词聚合的区域。

在Ucinet中,沿着“Network > 2-Mode > Categorical Core/Periphery”路径,输入“作者-关键词”转化后的二值矩阵,构建作者-关键词2-模网络的核心-边缘分区。所得结果如图5所示。

```
2-MODE CATEGORICAL CORE/PERIPHERY MODEL
Starting fitness: 0.420
Final fitness: 0.429
Blocked Adjacency Matrix
Density matrix
  1  2
-----
1 0.429 0.172
2 0.243 0.069
```

图5 作者-关键词核心-边缘分析

从分析结果可以看出,初始拟合值(Starting fitness)为0.420,最终拟合值(Final fitness)为0.429。拟合值的测度是0-1,0表示拟合度低,1表示完全拟合。在作者-关键词关系中,拟合值还和理想值有一定的差距,但从块1.1的密度来看,密度较高,可以说明核心-边缘问题。在作者-关键词分块矩阵中,核心是由邱均平、王知津、储节旺等24位核心作者和文献计量、情报学、引文分析等12个高频关键词组成的,其他作者和关键词均处于边缘地位。通过作者-关键词核心-边缘关系的分析,针对馆藏资源即可按照核心-边缘将资源聚合,提高资源的利用价值和核心资源的定位。

(3) 2-模网络的分派分析

分派分析是指将网络进行分组,发现相互排斥的组系,从而提高群体内部的密度,降低群体之间的密度。在Ucinet中,沿着“Network > 2-Mode > 2-Mode Factions”路径,输入作者-关键词二值矩阵,便可得知作者和关键词的分组定位。

计算结果显示,该网络分为两组,“文献检索、网络环境、知识转移”等关键词和除了“邱均平、陈惠兰、杨思洛”等作者以外的所有其他作者构成了网络核心,其余关系处于网络的边缘。观察计算结果,可知初始拟合值是0.081,最终拟合值是0.225。与核心-边缘分析相似的是,两种分析方法得到的结果均是按派系分布,很容易将关系紧密的资源分为一派,实现资源的聚合。

2-MODE CATEGORICAL CORE/PERIPHERY MODEL

Starting fitness: 0.081 Final fitness:

0.225 Correlation to ideal: 0.225

Blocked Adjacency Matrix

Density matrix

1 2

1 0.090 0.145

2 0.094 0.307

图6 作者-关键词分派分析

4 结论与展望

随着大数据时代的到来,信息量急剧增加,尤其对于数字图书馆资源来说,资源的聚合已成为迫切解决的问题。聚合技术作为一种行之有效的内容汇聚手段,对于组织海量的资源、有效建立资源之间的联结起到重要的作用。

本文根据构建的聚合模型,利用Ucinet软件及矩阵的构建进行社会网络分析,主要围绕准备数据、数据处理和数据分析三个步骤进行聚合分析。选取图情学科领域,以核心作者群和高频关键词为研究样本,通过构建核心作者互引矩阵、作者合作关系矩阵、作者-关键词矩阵,绘制了各要素之间的关系网络图谱,在此基础上,结合社会网络分析的指标,从中心性、距离、密度、凝聚子群、2-模网络等测度对资源的社会网络关系进行了拓展分析,并最终获得资源聚合的结果。

对于今后的研究,笔者认为还需要从以下几个方面加以完善:

(1) 社会网络分析理论与方法应用的拓展。本文引入社会网络分析指标测度(中心性、密度、距离、凝聚子群、2-模网络)来衡量数字图书资源聚合程度,但文中所运用的指标测度并非全部,要真正实现社会网络分析在数字图书馆资源聚合中的契合应用,还需要引入更多的指标测度来对资源特征间的关联以及资源利用过程的关联进行分析,从而使社会网络分析方法的应用更具实际意义。

(2) 数字图书馆资源聚合维度的丰富。资源的聚合不但需要从某一学科层面挖掘资源聚合的学科深度,更需要考虑资源内在属性结构的组织维度、资源与人关联

的传递维度以及知识不断更新的发现维度等。我们必须明确数字图书馆资源聚合不是目的,只是手段。

(3) 数字图书馆资源聚合技术方法上的互补融合。在以往的研究中,大多数资源聚合的视角往往基于某一单一的技术方法。然而很多技术方法在资源聚合方面存在着必然联系和相似性,如社会网络分析方法、复杂网络分析方法、文献计量学等。因此,了解技术方法上的互补融合之处,是全面、透彻地聚合馆藏资源的关键所在。数字图书馆资源聚合的视角必定是多样的,而且是无法穷尽的。多种技术方法的互补融合,既是对馆藏资源关联全方位的掌控,更是对馆藏资源聚合维度的拓展。

参考文献

- [1] GREENSTEIN D, THORIN S E. The Digital Library: A Biography [M]. Washington: Digital Library Federation, 2002.
- [2] 赵伟.数字图书馆说略[J].情报资料工作,1998,(4):38-40.
- [3] 李冠楠.基于主题图的数字图书馆信息资源整合研究[D].华中师范大学,2011.
- [4] 袁小一,俞毅.浅析医学数字图书馆的资源整合[J].河南图书馆学刊,2006(6):93-95.
- [5] 百度知道推出知识搜索服务[EB/OL]. [2013-08-30]. <http://net.chinabyte.com/376/12603876.shtml>.
- [6] 构建高校数字图书馆综合服务门户[EB/OL]. [2013-08-30]. <http://www.chinalibs.net>.
- [7] 普赖斯 D,张季娅.洛特卡定律与普赖斯定律[J].科学学与科学技术管理,1984(9):17-22.
- [8] 孙海生,张曙光.情报学核心作者互引网络分析[J].情报杂志,2011(3):78-83.
- [9] 邱均平,王菲菲.基于SNA的国内竞争情报领域作者合作关系研究[J].图书馆论坛,2010(12):34-40.
- [10] WHITE H D. Authors as Citers over Time [J]. Journal of the American Society for Information Science and Technology, 2001, 52(2): 87-108.
- [11] 朱庆华,李亮.社会网络分析法及其在情报学中的应用[J].情报理论与实践,2008(2):179-183.
- [12] KLEINBERG J. The Small-World Phenomenon: An Algorithmic Perspective [J]. Cornell University, 2005(6): 1-12.
- [13] 李长玲,纪雪梅,支岭.基于社会网络分析的企业内部知识传播效率分析[J].情报理论与实践,2011(10):56-61.

- [14] 邱均平,陈木佩.我国计量学领域作者合作关系研究[J].情报理论与实践,2012(11):56-60.
- [15] 郭志刚.社会统计分析法:SPSS软件应用[M].北京:中国人民大学出版社,2007.
- [16] 王小玉.“核心-边缘”理论的国内外研究述评[J].湖北经济学院学报,2007(10):41-45.

作者简介

贯君, 男, 1982年生, 吉林大学管理学院博士研究生。

毕强, 男, 1954年生, 吉林大学管理学院教授。研究方向: 知识管理。通讯作者, E-mail: biqiang12345@163.com。

Empirical Study on Digital Library Resource Aggregation Based on SNA

GUAN Jun¹, WANG Yu¹, WU HaiYuan²

(1. School of Management, Jilin University, Changchun 130022, China; 2. Library, Dalian University, Dalian 116622, China)

Abstract: This paper analyzes the relationship between the digital library resources of social network, combined with the index of social network analysis, according to the resource aggregation model, using centrality, condensing subgroup, and the core-edge measure, and verifies the result of the resource aggregation effect respectively, to optimize the quality of the resources, improve the availability of resources, and enhance the relevance of the resources.

Keywords: SNA; Digital library; Resources aggregation

(收稿日期: 2014-04-30)

《汉语主题词表》(工程技术卷) 即将出版

《汉语主题词表》自1980年问世以来,并经1991年进行自然科学版修订,在我国图书情报界发挥了其应有的作用,曾经获得了国家科学技术进步二等奖。为了适应网络环境下知识组织与数据处理的需要,2009年由中国科学技术信息研究所主持,并联合全国图书情报界相关机构,完成了《汉语主题词表》(工程技术卷)的重新编制工作,即将由科学技术文献出版社正式出版。

《汉语主题词表》(工程技术卷)共收录优选词19.6万条,非优选词16.4万条,等同率0.84。在体系结构、词汇术语、词间关系等方面进行改进创新。为了方便工程技术领域不同专业用户使用,《汉语主题词表》(工程技术卷)按专业分13个分册出版。同时建立《汉语主题词表》服务系统,提供在线概念检索和辅助标引服务。

《汉语主题词表》(工程技术卷)作为网络环境下知识组织和数据处理的基础性工具,期待着全国图书情报界共同推进其持续更新建设,也期盼社会各界进行深层次应用和实践。