

● 徐媛媛, 朱庆华 (南京大学 信息管理系, 江苏 南京 210093)

# 社会网络分析法在引文分析中的实证研究<sup>\*</sup>

**摘要:** 本文以参考咨询领域内的 32 名高被引作者为研究对象, 运用社会网络分析法, 结合 UCINET 软件分析作者之间的引文关系, 通过网络密度分析个体之间互引的频度, 通过点度中心度分析核心研究者, 通过中间中心度分析知识传播过程中承上启下的人物, 通过凝聚子群分析找出关系紧密的研究小团体。

**关键词:** 社会网络分析法; 参考咨询; 引文; 方法论; 应用研究

**Abstract:** Using the method of Social Network Analysis and the software of UCINET, this paper discusses the citation relationship of 32 high rate cited authors in the field of reference service such as frequency of inter citation by network density analysis, important researchers by degree centrality analysis, connecting authors in knowledge transmission by betweenness centrality analysis and research subgroups by cohesive subgroup analysis.

**Keywords:** social network analysis method; reference consultation; citation methodology; application study

引文分析法是情报学领域常见的、应用广泛的方法, 通过这种方法可以统计出高被引文献、高被引作者、核心研究机构等, 从而便于发掘学科热点、把握学科发展方向。邱均平教授利用引文分析法, 从引文量、高被引作者和高被引论文数量及被引次数、共被引次数 3 个方面分析了近年来“参考咨询”研究的现状, 得出具有突出贡献的作者和文献、研究热点等结论<sup>[1]</sup>。但是, 这种方法只是以个体为研究对象, 对其进行数量上的比较和排序, 主体内部例如作者之间、研究机构之间千丝万缕的联系却无从得知。为此, 可以用社会网络分析法解决, 它是一种研究社会结构和社会关系的分析方法, 适用于研究小整群体的人际关系与群体结构。目前, 社会网络分析法广泛应用于社会、经济、心理等领域, 例如郭立亚、朱瑜将社会网络分析法应用在运动队结构和人际特征分析中<sup>[2]</sup>, 彭小川、毛晓丹运用社会网络分析法分析 BBS 的群体特征<sup>[3]</sup>, 杨锐、黄国安也曾将其运用到杭州手机产业集群的分析中<sup>[4]</sup>。但是却鲜有人将社会网络分析法运用到引文分析领域中来。本文试图以邱均平教授提出的张晓林等 32 名高被引作者组成的群体为研究对象, 通过社会网络分析法, 运用 UCINET 软件, 进一步分析参考咨询领域内的作者引文关系网络的群体特征。通过网络密度分析个体之间互引的频度, 通过点度中心性分析核心研究者, 通过中间中心性分析知识传播过程中承上启下的人物, 通过凝聚子群分

析找出关系紧密的研究小团体。

## 1 研究方法

### 1.1 研究对象

本研究的数据来源为南京大学中国社会科学研究评价中心的数据库“中文社会科学引文索引”(CSSCI)。在 CSSCI 的“被引文献”数据库中, 依次搜索被引文献篇名中含有“参考咨询”, 并且被引文献作者是上述 32 名作者之一的论文(排除自引), 得到每个作者“参考咨询”方面的论文被其他 31 名作者引用的情况。

将张晓林等 32 名高被引作者用①~⑳号表示, 名字与序号的对应关系见表 1。根据数据搜集的结果绘制出社群图(见图 1)。成员之间的被引关系用带箭头的连接线表示, 箭头指向被引者, 双箭头表示相互引用, 通过社群图可以直观地揭示出群体的引用状况与结构。由图 1 可以发现, 这个群体网络结构比较紧凑, 其中以 1 号、3 号、5 号、18 号、19 号的网络关系最为密集, 同时 20 号、24 号、29 号、31 号孤立在外。

表 1 作者姓名序号对照表

1	张晓林	9	刘静	17	王纯	25	范爱红
2	詹德优	10	李晓芸	18	文庭孝	26	韩丽风
3	初景利	11	张海涛	19	郭晶	27	花芳
4	张鹰	12	黄解明	20	伍宪	28	黄敏
5	莫少强	13	马远良	21	曹海峰	29	吕佳
6	夏侯炳	14	赵媛	22	陈淑莲	30	邵敏
7	林曦	15	宋永香	23	郭劲赤	31	王群
8	韩志萍	16	肖珑	24	刘陆先	32	肖冬梅

\* 本文为南京大学研究生院核心课程建设项目“情报学理论与方法”的研究成果之一。

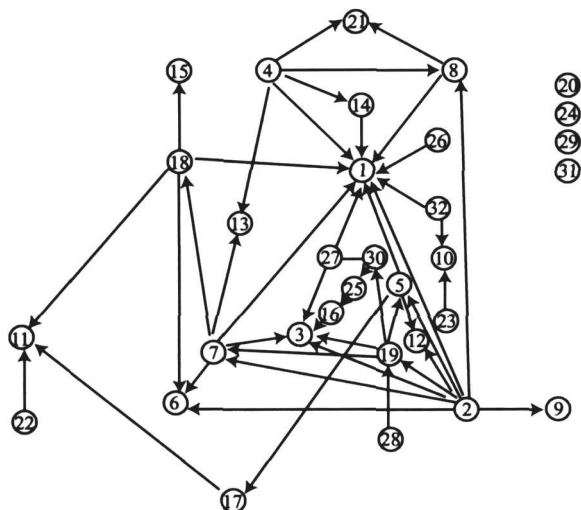


图 1 社群图

根据数据搜集的结果绘制出矩阵：行和列分别对应 32 名作者，以行 列交叉处的格值  $Z_{ij}$  表示对应两点之间的连接关系， $Z_{ij}=0$  表示没有直接连接， $Z_{ij}=1$  表示有直接连接。本文中的连接线有方向，则以行对应点为起点，列对应点为终点，从而得出  $32 \times 32$  的矩阵，本文截取其中 6 行 6 列，如图 2 所示。例如第 2 行与第 1 列交叉处的格值为 1，这表示 2 号作者詹德优引用过 1 号作者张晓林的文献，而第 1 行与第 2 列交叉处的格值为 0，表示 1 号作者张晓林没有引用过 2 号作者詹德优的文献。通过矩阵分析，研究者可以看出每个作者引用其他作者文章以及文章被其他作者引用的情况，大致了解群体在“参考咨询”领域内学术交流的情况。

$$K = \begin{matrix} & \begin{matrix} ① & ② & ③ & ④ & ⑤ & ⑥ \end{matrix} \\ \begin{matrix} ① \\ ② \\ ③ \\ ④ \\ ⑤ \\ ⑥ \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix}$$

图 2 矩阵图

### 1.2 研究角度

社会网络分析法已经比较成熟，可以从多个不同的角度构建社会网络进行分析。刘军在《社会网络分析导论》中比较全面地介绍了社会网络分析的基本概念和方法<sup>[5]</sup>，笔者引用的概念都来自这本书，并从密度、点度中心度、中间中心度和凝聚子群分析 4 个角度进行。

#### 1.2.1 密度

密度是社会网络分析中最常用的一种测度，它指的是一个图中各个点之间联络的紧密程度。这个概念是为了汇总各个线的总分布，以便测量该分布与完备的差距有多大，固定规模的点之间的连线越多，该图的密度就

越大。

**1.2.2 点度中心度** 在一个社会网络中，如果一个行动者与很多其他行动者之间存在直接联系，那么该行动者就居于中心地位，在该网络中拥有较大的“权力”。在这种思路的指导下，网络中一个点的点度中心度，就可以用网络中与该点有直接联系的点的数目来衡量，这就是点度中心度。

点度中心度又分为绝对点度中心度和相对点度中心度。所谓绝对点度中心度，就是将某一点的点度中心度等同于该点的度数，即与该点有直接联系的点的个数，而相对点度中心度是点的绝对中心度与网络中点的最大可能的度数之比。

中心度刻画的是单个行动者在网络中所处的核心位置，有时研究者关心的不是点，而是整个图，研究不同的图是否有不同的中心趋势，为此可以用中心势来刻画网络图的整体中心性。

**1.2.3 中间中心度** 如果一个行动者处于许多交往网络路径上，可以认为此人居于重要地位，因为他具有控制其他两人之间的交往的能力。研究者用中间中心度来刻画行动者个体中心度，它测量的是行动者对资源控制的程度。如果一个点处于许多其他点对的测地线（最短的途径）上，表示该点具有较高的中间中心度，它起到沟通各个他者的桥梁作用。

**1.2.4 凝聚子群** 通俗地讲，凝聚子群就是团体中的一小群人关系特别紧密，以至于结合成一个小团体。从不同角度分析，凝聚子群的概念有多种，主要包括建立在互惠性基础上的“派系”，建立在可达性和直径基础上的“ $n$ 派系”、“ $n$ 宗派”，建立在点度数基础上的“ $k$ 丛”，以及根据子群内外关系建立起来的“成分”。本文涉及的主要是“成分”及“ $k$ 丛”。

如果一个图可以分为几个部分，每个部分内部任何两点都可以通过一定的途径相连，但是，各个部分之间没有任何关联，就把这些部分称为成分。一个  $k$  丛就是满足下列条件的一个凝聚子群，即在这样一个子群中，每个点都至少与除了  $k$  个点之外的其他点直接相连。也就是说，如果一个凝聚子群的规模为  $n$ ，那么只有当该子群中的任何点的度数都不小于  $n-k$  这个值的时候，才称之为  $k$  丛。

### 1.3 研究工具

为了使社会网络分析更简单易行，很多学者都致力于开发社会网络分析的软件或程序。社会网络分析软件有多种，在国际社会网络分析网（NSNA）的网站上列出的可用于社会网络分析的程序软件多达 54 种<sup>[9]</sup>。其中一些是可视化软件，用于对社会网络进行可视化；一些具有分析的功能。中国社会科学院社会学研究所主办的中国社会学

BLOCK DENSITIES OR AVERAGES

Input dataset: D:\大四下\yy1

Relation: Page 1

Density (matrix average) = 0.0454  
Standard deviation = 0.2081

Use MATRIX>TRANSFORM>DICHOTOMIZE procedure to get binary image matrix.  
Density table(s) saved as dataset Density  
Standard deviations saved as dataset DensitySD  
Actor-by-actor pre-image matrix saved as dataset DensityModel

Running time: 00:00:01  
Output generated: 28 四月 07 15:31:24  
Copyright (c) 1999-2005 Analytic Technologies

图3 密度计算结果

网 (<http://www.sociology.cass.cn/>) 的“研究工具”专栏列出了一些常用的社会网络分析软件, 其中主要包括: KrackPlot, SIRUCTION, GEGOPY, Pajek, Stocne和 UCNET等。其中, UCNET是研究者使用较多的社会网络分析软件。

UCNET是由 Borgatti, Everett和 Freeman开发的<sup>[7]</sup>。他们3人在社会网络分析领域都取得了很大的成就。UCNET是一个用来处理社会网络数据的软件包, 能读取多种不同形式的文件, 如文本文件 (Text Files) 和 Excel文件, 它能处理 32 767个网络节点。当然, 从实际操作来看, 当节点数在 5 000 ~ 10 000 之间时, 一些程序的运行就会很慢。社会网络分析法包括中心性分析、子群分析、角色分析和基于置换的统计分析等。另外, 该软件包有很强的矩阵分析功能, 如矩阵代数和多元统计分析。

罗家德在其《社会网分析讲义》中具体介绍了 UCNET软件的使用方法<sup>[8]</sup>, 本文以此为参考, 运用 UCNET完成本文的数据分析。

## 2 结果及讨论

### 2.1 密度

在 UCNET软件中, 沿着 Data-Spreadsheet Editor的路径, 就会跳出 EXCEL的格式, 输入 32×32的矩阵, 输入完毕后, 按 File-Save as 键入文件名 yy1, UCNET会自动在预存取的档案夹中创造两个档案, 分别是 yy1.#和 yy1.#.d

然后沿着 Network-Cohesion-Density的路径可以计算密度。结果如图3所示。

此社会网的密度是 0.0454 与一般社会网相比密度偏低。这种情况在搜集数据阶段已有初步估计, 这个高被引作者群体的文章多数被其他研究者所引用, 而他们之间

文章互引的情况不多。笔者探索性思考其原因, 高被引作者都是在参考咨询领域内颇有建树的学者, 也就是学科带头人, 他们的研究对学科有突出贡献, 所以他们的文章被很多后来研究者引用。但是他们分属不同的研究方向, 所以内部之间互相引用的情况较少。

### 2.2 点度中心度

同样, 沿着 Network-Centrality-Degree进行点度中心度的计算。结果如图4所示。

	1 OutDegree	2 InDegree	3 NrmOutDeg	4 NrmInDeg
2 詹德优	9.000	0.000	29.032	0.000
4 张鹰	6.000	0.000	19.355	0.000
7 林曦	5.000	2.000	16.129	6.452
19 郭晶	4.000	2.000	12.903	6.452
18 文庭孝	4.000	1.000	12.903	3.226
5 莫少强	2.000	2.000	6.452	6.452
27 花芳	2.000	1.000	6.452	3.226
8 韩志萍	2.000	2.000	6.452	6.452
30 邵敏	2.000	1.000	6.452	3.226
32 肖冬梅	2.000	0.000	6.452	0.000
17 王纯	1.000	1.000	3.226	3.226
25 范爱红	1.000	1.000	3.226	3.226
23 郭劲赤	1.000	0.000	3.226	0.000
22 陈淑莲	1.000	0.000	3.226	0.000
28 黄敏	1.000	0.000	3.226	0.000
26 韩丽风	1.000	0.000	3.226	0.000
16 肖珑	1.000	1.000	3.226	3.226
13 马远良	0.000	2.000	0.000	6.452
1 张晓林	0.000	9.000	0.000	29.032

.....  
Network Centralization(Outdegree)=25.286%  
Network Centralization(Indegree)=25.286%

图4 点度中心性计算结果

通过以上结果, 可以找出哪些作者在文献交流过程中处于核心地位。不同的作者表现出不同的点出度和点入度。其中, 詹德优的点出度最大为9次, 即引用其他作者的次数最多, 说明他在文献交流过程中比较活跃, 积极学习并引进他人的优秀研究成果为自己所用, 笔者定义他为“活跃型”。其他点出度较高的作者还有张鹰、林曦、郭晶、文庭孝。而张晓林、初景利、夏侯炳、张海涛4人的

点入度较高,说明他们善于提出新颖的、有建设性的观点并得到同行认可,在知识传播过程中处于源头或者是上流位置。他们是“参考咨询”领域内的专家级人物,控制着整个领域的研究方向。笔者定义他们为“权威型”。另外,笔者还发现伍宪、吕佳、王群、刘陆先4人的点入度和点出度都为0,也就是说他们在这个群体网中既不引用别人的文章,他们的文章也未被别人引用,笔者定义为“孤立型”。

整个网络的标准化点入度中心势和点出度中心势均为25%,这说明该群体引用与被引用关系比较对称。中心势越接近1,说明网络越具有集中趋势。从上述点度中心势来看,该群体网络的集中趋势不是很大。

### 2.3 中间中心度

沿着 Network—Centrality—Betweenness—Nodes 路径可以得到中心中间度,结果如图5所示。

		1 Betweenness	2 nBetweenness
19	郭晶	18.000	1.935
7	林曦	15.000	1.613
30	邵敏	9.000	0.968
18	文庭孝	8.000	0.860
25	范爱红	4.000	0.430
5	莫少强	3.000	0.323
27	花芳	2.000	0.215
8	韩志萍	1.000	0.108
16	肖珑	1.000	0.108
17	王纯	1.000	0.000
10	李晓芸	0.000	0.000
4	张鹰	0.000	0.000
13	马远良	0.000	0.000
6	夏侯炳	0.000	0.000
15	宋永香	0.000	0.000
.....			

Network Centralization Index=1.78%

图5 中间中心性计算结果

上述结果刻画出作者们在多大程度上处在其他任何两个作者之间的测地线上,也就是在多大程度上控制文献交流过程中的引用关系。与分析点度中心度不同的是,在分析中间中心度的时候,数据是被看成对称的。因此,引用次数多的和被引次数多的作者的中间中心度都很高,但是意义迥然不同。对于被引用者来说,其中心度越高,说明他越能控制文献的传播。对于引用者来说,其中心度越高,说明他在文献传播过程中越被他人控制。从分析结果看,郭晶、林曦、邵敏、文庭孝等人的中间中心度较高。由于他们既引用过别人的文章,也被别人引用过文章,所以难以判断他们在文献传播过程中起控制作用还是被别人控制。但是,毋庸置疑他们较频繁地处在其他作者引用关系的测地线上,在知识传播、观点一派相承的过程中起着桥梁的作用。就整个网络来说,中间中心势指数为1.78%,这说明整个网络的中心度较低。

### 2.4 凝聚子群分析

沿着路径 Network—Regions—Components—Simple Graphs 可以得到该群体网络中含有的成分。需要说明的是,软件运行过程中会出现一个界面需要操作者选择“强关联”(Strong)或者“弱关联”(Weak)。强关联指成分中任何两点都需要双向地相连,即能从某点到另外一点,也能反过来从另外一点到这个点。弱关联则无此要求,成分中任何两点只要能通过一定的途径相连即可。由于本文研究的网络比较稀疏,若使用强关联,32个个体就会出现32个成分,每个个体都是一个成分,研究没有意义。所以改选择弱关联,结果如图6所示。

#### 5 components found

##### Component Sizes

Component	Nodes	Proportion
1	28	0.875
2	1	0.031
3	1	0.031
4	1	0.031
5	1	0.031

##### NODE COMPONENT

张晓林	1
詹德优	1
初景利	1
张鹰	1
莫少强	1
夏侯炳	1
林曦	1
.....	
伍宪	2
曹海峰	1
陈淑莲	1
郭劲赤	1
刘陆先	3
范爱红	1
韩丽风	1
花芳	1
黄敏	1
吕佳	4
邵敏	1
王群	5
肖冬梅	1

##### Components with 3 or more members:

1:张晓林 詹德优 初景利 张鹰 莫少强 夏侯炳 林曦 韩志萍 黄晓明 马远良 赵媛 宋永香 肖珑 王纯 文庭孝 郭晶 曹海峰 陈淑莲 郭劲赤 范爱红 韩丽风 花芳 黄敏 邵敏 肖冬梅

图6 成分计算结果

由图6中的结果可知,该群体网络由5个成分组成。成分1由张晓林、詹德优等28名作者组成, {伍宪}, {刘陆先}, {吕佳}, {王群}分别为另外4个成分。也就是说,张晓林、詹德优等组成的小团体内部彼此之间或直接或间接出现过引用被引用关系,学术方面经历过借鉴、继承。而伍宪、刘陆、吕佳、王群与其他任何人的研究不存在交流,处于独立的状态。这与前面的社群图、点度中心性的研究结论一致。

沿着 Network—Subgroups—K-plex 路径可以找出建立在点度数基础上的小团体。这个过程不会一下就有明确的结果出来,需要反复试验、自行判断。UCINET会把具方向性社会网当作无方向性社会网来计算,凡是两个节点中有一方选择了对方,UCINET就当作两节点间有一条线。笔者选择K为2,最小的团体不得少于3个节点,得到

105个小团体，完全无法判断，表示设定的条件太宽（K值越小，最小的小团体节点数越大则条件越严格）。于是继续试验，若K为3，最小的团体不得少于5个节点，则得到54个小团体，仍然不好判断。最后，笔者选择K为2，最小的团体不得少于4个节点，得到如图7所示的17组团体。

17 k-plexes found.

- 1: 张晓林 詹德优 初景利 林曦
- 2: 张晓林 詹德优 初景利 花芳
- 3: 张晓林 詹德优 张鹰 韩志萍
- 4: 张晓林 詹德优 莫少强 林曦 郭晶
- 5: 张晓林 詹德优 莫少强 韩志萍
- 6: 张晓林 詹德优 莫少强 黄晓明
- 7: 张晓林 詹德优 夏侯炳 林曦 文庭孝
- 8: 张晓林 詹德优 林曦 韩志萍
- 9: 张晓林 初景利 林曦 花芳
- 10: 张晓林 张鹰 林曦 马远良
- 11: 张晓林 张鹰 韩志萍 曹海峰
- 12: 詹德优 初景利 莫少强 郭晶
- 13: 詹德优 初景利 夏侯炳 林曦
- 14: 詹德优 初景利 林曦 郭晶
- 15: 詹德优 莫少强 黄晓明 郭晶
- 16: 詹德优 夏侯炳 林曦 郭晶
- 17: 初景利 郭晶 花芳 邵敏

图7 k从计算结果

结合图8所示的聚类图，可以做出判断，张晓林、詹德优、林曦、初景利、郭晶、莫少强、夏侯炳、黄晓明是一个小团体，马远良、张鹰、韩志萍、曹海峰是一个小团体。

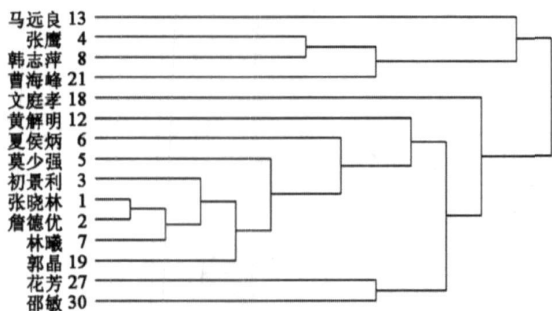


图8 树状聚类图

如果两个作者在同一个小团体内，说明他们在“参考咨询”领域内的学术关联较大，他们可能研究方向相同，可以互相交流、借鉴，所以内部出现较多引用与被引用的关系，联系较为紧密，而与外界联系稀疏。

### 3 结束语

运用社会网络分析法进行引文分析，与传统的引文分析法相比，两者还是存在着差异的。

- 1) 方法基础。前者采用统计学思想和逻辑思维方法，后者采用图论的思想。
- 2) 研究对象。前者侧重研究引文的数量特征，后者着力研究文献某个特征量例如作者、机构或者期刊组成的群体的网络关系。
- 3) 研究结论。通过引文分析法可以得出核心作者、

核心期刊、核心研究机构等结论，而社会网络分析法可以分析出某群体网是否紧密、哪些节点处于网络传递的中心地位、有哪些明显的小团体等。

4) 应用范围。前者仅限于文献分析领域，后者广泛应用于人类学、心理学、社会学、经济学等领域。另外，现在已出现若干个由社会网络分析专家开发的分析软件，使得社会网络分析变得简单易行。所以，相比于引文分析法，社会网络分析法具有它的优势和独特之处。但是，国内将社会网络分析法运用到文献分析领域内的研究还较少。引文分析法应与社会网络分析法互相补充，共同将文献分析推向新的高度。

笔者通过对“参考咨询”领域中高被引作者的引文网络分析可以发现，该群体网络的密度和中心度都很小，说明作者之间引文关系较少，处于核心地位的作者显著性不够突出。所以该领域发展还不够成熟，不同研究方向的学科带头人之间需要多多交流、集思广益以获得进一步的发展。

本文分析的是二值有向图，即两个主体之间只用1（有联系）和0（无联系）两个数值表示。而数据搜集过程中发现，有部分作者曾多次引用另外一个作者的文章，如詹德优曾3次引用初景利的文章。如果这种情况不能很好地反映到有向图中，会影响到后面的网络分析。这是本文的不足之处。一种改进的办法是采用赋值图进行研究，这将是今后的研究方向。□

### 参考文献

- [1] 邱均平, 李江. 基于引文分析法的国内“参考咨询”的研究现状分析 [J]. 图书情报知识, 2007 (1): 54-57
- [2] 郭立亚, 朱瑜. 社会网络分析法在运动队结构和人际特征分析中的应用研究 [J]. 中国体育科技, 2007, 41 (5): 10-13
- [3] 彭小川, 毛晓丹. BBS群体特征的社会网络分析 [J]. 青年研究, 2004 (4): 39-44
- [4] 杨锐, 黄国安. 网络位置和创新——杭州手机产业集群的社会网络分析. 工业技术经济, 2005 24 (7): 114-118
- [5] 刘军. 社会网络分析 [M]. 北京: 社会科学文献出版社, 2004
- [6] Computer Programs for social network analysis [EB/OL]. [2007-03-27]. [http://www.insna.org/NSNA/soft\\_inf.html](http://www.insna.org/NSNA/soft_inf.html)
- [7] Borgatti S P, Everett M G, Freeman L C. Ucinet for windows software for social network analysis [M]. Harvard MA: Analytic Technologies, 2002
- [8] 罗家德. 社会网分析讲义 [M]. 北京: 社会科学文献出版社, 2005

作者简介: 徐媛媛, 女, 1985年生, 硕士生。  
朱庆华, 男, 1963年生, 博士, 教授。  
收稿日期: 2007-10-15