

# 基于复杂网络的港口物流领域研究热点实证分析<sup>①</sup>

王绍卜

(浙江万里学院 现代物流学院, 宁波 315100)

**摘要:** 为获取国内港口物流领域研究热点, 对中国期刊 CNKI 全文数据库中收录的 2913 篇样本文献的关键词信息进行实证分析. 利用复杂网络理论, 构建一个包含 50 个节点、982 条边(关键词共现)的无向加权网络. 使用社会网络分析工具 Pajek 进行度、路径、聚类系数等统计指标的计算, 验证该网络所具有小世界、无标度特性. 基于网络的点度中心性和介数中心性, 离析出研究热点, 研究热点相互联系紧密, 但关注度差异性大.

**关键词:** 复杂网络; 港口物流; 研究热点

## Empirical Analysis of the Research Hotspot in the Field of Port Logistics Based on Complex Network

WANG Shao-Bu

(Modern Logistics School, Zhejiang Wanli University, Ningbo 315100, China)

**Abstract:** In order to obtain the research hotspot of port logistics field in China, the empirical analysis of the Chinese journal full text database of CNKI in 2913 samples of keywords in the literature information. This paper uses the theory of complex network to construct a contains 50 nodes, 982 edges (keyword co-occurrence) undirected weighted network. Calculation analysis tool Pajek statistical index, path, clustering coefficient using social network, the network has small world, scale-free property. The network degree are centrality and betweenness centrality based on separate, research focus, research focus are closely linked, but the attention degree of difference.

**Key words:** complex network; port logistics; research focus

### 1 引言

经济全球化的浪潮已经把世界融为一个整体, 港口作为对外贸易进出口货物的集散中心, 成为国际物流供应链的重要节点和物流通道的枢纽, 港口物流领域的研究热潮不减. 在港口物流信息技术方面, 侯贵宾通过 3G 无线技术在港口物流中的实践, 将 3G 无线通信技术 TD-SCDMA 在港口进行研究和应用, 形成了示范效应<sup>[1]</sup>; 贾腾飞研究了一种基于 ZigBee 和 GPRS 技术港口门机安全监控系统, 利用远程 GPRS 无线传输技术和近距离的无线传感器网络 ZigBee 技术结合的方法, 对港口门座式起重机的安全进行监控, 确保门机运行安全<sup>[2]</sup>. 在港口物流经济方面, 沈秦伟等利用格兰杰因果检验、向量自回归模型及相应的脉冲响应函数分析了大连港口物流与城市经济增长的动

态相关性和因果关系<sup>[3]</sup>; 徐良婧采用定性、定量研究相结合的方法, 通过比较来研究港口物流经济带动效应<sup>[4]</sup>. 在港口物流信息系统方面, 李云峰分析了信息化电子平台的应用在港口物流中的作用, 设计了集装箱出口单证自动核对模块和自动配载模块<sup>[5]</sup>; 蔡芸、霍永忠针对集装箱港口物流系统研究中采用的仿真技术和优化技术, 分析了研究现状和存在问题, 并对研究的方向进行了评述<sup>[6]</sup>.

以上可见, 研究者主要是在微观层面上针对相关的技术、经济、系统形态开展了深入的研究, 但缺乏在宏观层面上对港口物流的研究热点分布进行定量分析. 主要原因是研究热点分布网络错综复杂, 数据量大. 随着计算机存储能力和处理数据能力的增强, 以及一些大规模系统数据库的建立, 复杂网络理论在社

<sup>①</sup> 基金项目:浙江省临港现代服务业与创意文化重点研究基地课题(12JDLG01Z)

收稿时间:2013-11-15;收到修改稿时间:2014-01-26

会网络分析中得到了广泛的应用,文章将基于复杂网络理论与方法进行港口物流领域研究热点关键词的选取、筛选,构建相应网络,并对网络拓扑结构、研究热点实证分析.

## 2 数据来源及研究方法

### 2.1 数据来源

文章选取中国期刊 CNKI 全文数据库,文献类型为期刊论文、博硕士学位论文和会议论文,以“港口”、“物流”为篇名检索词进行检索,检索条件为“发表时间 between (2000-01-01,2013-07-16 and 题名=港口 and 题名=物流 ) (精确匹配)”,共检索出 2932 篇港口物流领域相关文献,排除没有关键词的文章,共得到 2913 篇样本文献,文献的时间分布情况如下:

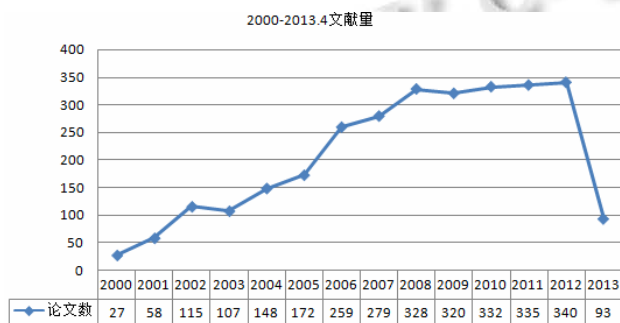


图 1 2000-2013 年样本文献统计  
(注: 2013 年最新文献日期是 4 月)

### 2.2 研究方法

#### 2.2.1 共词分析法

共词分析法属于内容分析方法的一种,通过分析在同一个文本中的款目对共同出现的形式,确定主题间的关系. 即当两个表达某一学科领域研究主题或研究方向的关键词出现在同一篇文献中,表明这两个词之间具有一定的内在联系,并且出现的次数越多,表明关系越密切、距离越近. 共词分析的核心内容有三种: 因子分析、聚类分析和多维尺度分析.

#### 2.2.2 复杂网络理论

复杂网络是对复杂系统的抽象和描述方式,任何包含大量组成单元(或子系统)的复杂系统,当把构成单元抽象成节点、单元之间的相互关系抽象为边时,都可以当作复杂网络来研究. 复杂网络是研究复杂系统的一种角度和方法,它关注系统中个体相互关联作

用的拓扑结构,是理解复杂系统性质和功能的基础<sup>[7]</sup>.

## 3 港口物流领域研究热点网络构建

### 3.1 基本算法

定义网络  $GW=(V,E,W)$ , 其中  $V=\{v_1,v_2,v_3,\dots,v_n\}$  表示以样本文献的关键词为顶点的集合;  $E=\{e_1,e_2,e_3,\dots,e_k\}$  表示边的集合(若两关键词在同一篇文献中出现,则认为它们间存在联系,定义为边);  $W=\{w_1,w_2,w_3,\dots,w_k\}$  表示边的权值集合(以共词关系在文献中再现的频率定义为边的权值),下面对算法进行举例说明.

假设有 3 篇文献,其关键词如表 1 所示.

表 1 关键词列表

文献	关键词
港口企业构建物流联盟的博弈分析	港口、港口物流、博弈分析
港口物流发展优势及对策的探讨	港口物流、优势、对策
基于全球供应链管理模式的港口物流	港口、港口物流、对策

三篇文献共有五个不同的关键词,分别是港口、港口物流、博弈分析、优势、对策,定义为节点  $v_1, v_2, v_3, v_4, v_5$ . 其对应的共现关系(定义为边)及频数(定义为权值)如表 2, 由这些节点和边构建的无向加权网络拓扑模型如图 2.

表 2 边及权值

边(共现关系)	权值
港口-港口物流	2
港口-博弈分析	1
港口物流-博弈分析	1
港口物流-优势	1
港口物流-对策	2
优势-对策	1
港口-对策	1

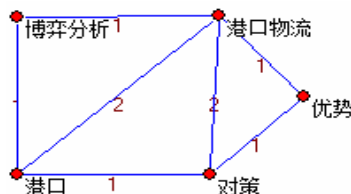


图 2 加权网络模型

### 3.2 节点的选取(研究热点关键词的挖掘)

根据基本算法, 利用 JAVA 开发的软件首先对 2913 篇样本文献的关键词进行统计, 共得到 24576 个关键词. 统计过程中, 对关键词进行适当的筛选, 合

并“港口”、“港口企业”、“中国港口”、“港区”、“现代港口”等类型的同义词, 删除“市委书记”、“副市长”等职位关键词. 遴选出前 50 个高频词(出现的频次大于 25, 百分比大于 0.1%)定义为节点.

表 3 高频关键词统计表

序号	关键字段	频次	百分比%	序号	关键字段	频次	百分比%	序号	关键字段	频次	百分比%
1	港口物流	1241	5.0957	18	物流产业	###	0.4270	35	保税区	46	0.1889
2	港口	1136	4.6645	19	临港产业	92	0.3778	36	北部湾地区	46	0.1889
3	物流	946	3.8844	20	港口功能	89	0.3654	37	枢纽港	41	0.1684
4	发展策略	910	3.7366	21	物流体系	83	0.3408	38	经济区	40	0.1642
5	集疏运	410	1.6835	22	天津港	82	0.3367	39	宁波港	39	0.1601
6	物流服务	352	1.4453	23	上海港	68	0.2792	40	厦门港	38	0.1560
7	吞吐量	333	1.3673	24	港口城市	61	0.2505	41	现代服务业	37	0.1519
8	物流中心	312	1.2811	25	物流基地	61	0.2505	42	物流成本	36	0.1478
9	港口物流企业	276	1.1333	26	临港工业	59	0.2423	43	港口资源	34	0.1396
10	港口物流园区	267	1.0963	27	仓储	58	0.2382	44	青岛港	33	0.1355
11	物流系统	176	0.7227	28	第三方物流	58	0.2382	45	转型升级	32	0.1314
12	供应链	175	0.7186	29	港口群	55	0.2258	46	产业集群	31	0.1273
13	港口经济	171	0.7021	30	珠三角	52	0.2135	47	深圳港	30	0.1232
14	集装箱码头	147	0.6036	31	集装箱	50	0.2053	48	舟山港	29	0.1191
15	多式联运	132	0.5420	32	连云港	50	0.2053	49	SWOT分析	28	0.1150
16	港口优势	113	0.4640	33	国际航运中心	48	0.1971	50	大连港	25	0.1027
17	港口建设	112	0.4599	34	物流管理	46	0.1889				

从高频关键词的分布可以发现, 可以分为以下几类:

1)区域港口分析: 相关关键词有天津港、上海港、连云港、厦门港、大连港、珠三角、宁波港、深圳港、青岛港、舟山港及北部湾地区、枢纽港、港口群. 这些港口的吞吐量和生产能力都是名列我国沿海港口前列.

2)物流服务研究: 相关关键词有仓储、集装箱、集装箱码头、现代服务业、保税区、物流基地、集疏运、多式联运、国际航运中心、物流服务.

3)物流经济研究: 相关的关键词有临港产业、临港工业、港口经济、港口优势、港口建设、港口功能、经济区、产业集群、港口城市、物流产业、物流管理.

4)发展策略研究: 相关关键词有港口资源、发展策略、转型升级、SWOT 分析.

5)物流理论研究: 相关关键词有吞吐量、港口物流、物流、物流中心、物流体系、第三方物流、港口物流企业、港口物流园区、供应链、物流成本、物流

系统<sup>[8]</sup>.

### 3.3 边的确定(共现关系分析)

统计样本(50 个高频关键词)在文献中的共现关系, 以此构建一个 50\*50 的邻接矩阵 N.

$$N = \begin{bmatrix} 0 & a_{1,2} & \dots & a_{1,50} \\ a_{2,1} & 0 & \dots & a_{2,50} \\ \dots & \dots & \dots & \dots \\ a_{50,1} & a_{50,2} & \dots & 0 \end{bmatrix}$$

其中, (i,j=1,2,3,4 50)表示第 i 个关键词与第 j 个关键词是否共同出现在同一篇文献中, 0 表示没有, 其它数值表示在多篇文献中出现的频数, 表示权值, 邻接矩阵部分数据如下.

0	349	277	489	189	152	127	113	...
349	0	558	412	300	196	196	170	...
277	558	0	378	195	188	159	156	...
489	412	378	0	167	190	100	138	...
189	300	195	167	0	81	159	71	...
152	196	188	190	81	0	27	81	...
127	196	159	100	159	27	0	43	...
113	170	156	138	71	81	43	0	...
...	...	...	...	...	...	...	...	...

图 3 邻接矩阵数据





这两个节点的最短路径上的边数. 任意两个节点间距离的最大数称为网络的直径, 记为  $D$ , 即:

$$D = \max_{i,j} d_{ij}$$

网络的平均路径长度  $L$  定义为任意两个节点之间距离的平均值, 即( $N$  为网络节点数):

$$L = \frac{1}{N(N-1)} \sum d_{ij}$$

#### 4.2.2 集聚系数

集聚系数按需求的不同分为整体集聚系数、局部集聚系数和平均集聚系数.

整体集聚系数是建立在闭三点组(邻近三点组)之上, 定义为一个图中所有闭三点组的数量与所有连通三点组(无论开还是闭)的总量之比(也有定义为这个值的三倍, 使得在完全图中的整体集聚系数等于 1, 罗伯特·邓肯·路斯和阿尔伯特·D·佩里).

局部集聚系数(即节点的聚类系数) $C_i$  定义为与该节点相连的点之间实际存在的边数  $E_i$  和总的可能边数  $k_i(k_i-1)/2$  之比, 即:  $C_i = 2E_i / (k_i(k_i-1))$ . 平均聚类系数为所有节点聚类系数  $C_i$  的平均值, 为整个网络的聚类系数,  $0 \leq C \leq 1$  (邓肯·J·瓦兹与斯蒂芬·斯特罗加茨).

由 Pajek 计算出研究热点网络的平均路径长度及平均聚类系数, 并与相同规模(节点数目、边数相同)的随机网络比较.

复杂网络的小世界特征主要表现在平均路径长度和聚类系数两个指标上, 比较上面计算结果可知: 研究热点网络相比相同规模的随机网络具有较小的路径长度和较大的聚类系数, 具有明显的小世界网络特征.

表 4 平均路径长度、平均聚类系数对比表

项目	节点数	边的总数	平均路径长度	聚类系数
研究热点网络	50	982	1.19837	0.616
随机网络	50	982	1.59918	0.354

#### 4.3 中心性分析(研究热点分析)

中心性分析是社会网络分析中常见的研究节点重要性方法, 常用的中心性分析指标有点度中心性和介数中心性.

##### 4.3.1 点度中心性

中心度指单个节点在网络中的位置, 社会网络理论认为, 节点的度越高, 说明节点与很多其它节点有直接的联系, 该节点就处于中心位置, 信息资源通达它的速度也越快, 可供调用的信息资源就越多, 拥有较大的“权利”, 意味着这个节点在某种意义上越“重要”.

计算得: 研究热点网络节点度值情况,  $K_{max}=49$ ,  $K_{min}=9$ , Average Degree = 39.28.

表 5 节点度排名前 10 的关键词

排序	关键词	点度	排序	关键词	点度
1	港口物流	49	6	吞吐量	48
2	港口	49	7	港口物流企业	48
3	物流	49	8	港口物流园区	48
4	发展策略	49	9	物流中心	47
5	集疏运	48	10	物流系统	47

与同规模网络(50 个节点、982 条边的随机网络)比较如下:

表 6 节点度对比图

项目	节点数	边的总数	最大度	最小度	网络平均度
研究热点网络	50	982	49	9	39.28
随机网络	50	982	42	25	32.32

比较可知: 研究热点(关键词)网络的平均度高于同规模的随机网络, 说明网络的连通性好, 研究热点问题联系紧密. 除去“港口物流”、“港口”、“物流”三个与检索条件直接相关度较高的节点外, “发展策略”、“集疏运”、“吞吐量”三个关键词点度较高, 与其它研究主题联系紧密.

##### 4.3.2 介数中心性

介数中心性反映的是相应的节点或边在整个网络中作为媒介的能力, 分为边介数和节点介数两种, 节点介数定义为网络中所有最短路径中经过该节点的路径数目占最短路径总数的比例. 点度高的节点不一定比点度低的节点更好的处于网络中心位置, 也有可能局部范围内介数值.

由 Pajek 计算出网络节点的介数值, 排名前十的关键词如表 7:

表 7 介数排名前十的关键词

排序	关键词	介数集中性	排序	关键词	介数集中性
1	港口物流	0.01188	6	吞吐量	0.00695
2	港口	0.01188	7	港口物流企业	0.00695
3	物流	0.01188	8	港口物流园区	0.00695
4	发展策略	0.01188	9	青岛港	0.00654
5	集疏运	0.00695	10	物流中心	0.00599

对比表 3、表 5、表 7，大多数频次高的关键词，其对应的度和介数也较高，但并不完全一致。如“物流服务”关键词频次高，但其度与介数却较低；“青岛港”关键词频次较低，但其介数较大；“物流系统”关键词度数大，介数小。节点的度较高说明在区域内该节点处于中心位置，介数高说明经过此节点的最短路径多，只有度与介数都较高的节点，才是研究热点，如“发展策略”、“集疏运”、“吞吐量”。

## 5 结论

文章应用复杂网络的理论构建了港口物流领域研究热点网络模型，使用 Pajek 软件进行仿真得到网络拓扑结构的可视化全景图，通过计算节点的度值、度分布、平均路长度、介数、聚类系数对网络的拓扑结构进行分析，结果表明该网络具有小世界网络特征，其度分布呈幂律分布，是典型的无标度网络。综合节点的频次、度值、介数可知：港口物流领域的研究热点是以上海港、天津港等大港为研究对象，主要针对港口物流的发展策略、集疏运、物流服务、吞吐量和信息服务体系等内容开展的，涵盖港口物流经济、物流服务、物流体系全方位，且相互间联系紧密，连通性

好，不存在孤立研究现象。

## 参考文献

- 侯贵宾. 3G 技术在港口物流中的应用研究. 中国港口, 2010(6):54-56.
- 贾腾飞. 基于 ZigBee 和 GPRS 技术的港口门座起重机安全监控系统研究. 武汉: 武汉理工大学, 2012.
- 沈秦伟, 韩增林, 郭建科. 港口物流与城市经济增长的关系研究—以大连为例. 地理与地理信息科学, 2013(1):70-73.
- 徐良婧. 港口物流经济带动效应国际比较研究. 杭州: 浙江大学, 2013.
- 李云峰. 港口物流信息管理平台核心模块的设计与研究. 大连: 大连海事大学, 2011.
- 蔡芸, 霍永忠. 集装箱港口物流系统仿真和优化研究综述. 系统仿真学报, 2009(8):2119-2124.
- 戴诚, 成全. 近十年我国社会网络研究热点透析. 现代情报, 2013(1):160-167.
- 皇甫青红, 华薇娜, 等. 国际数字图书馆领域研究热点及作者团体分析. 情报杂志, 2012(9):198-223.
- 韩秋明. 基于复杂网络的国内信息行为研究热点及衍化路径分析. 图书情报工作, 2012(7):94-101.