

Principles and  
Applications of  
Analyzing a Citation Space

引文空间蕴含着巨大的知识宝藏等你来挖掘

# 引文空间分析原理 与应用

CiteSpace实用指南

陈悦 陈超美 胡志刚 王贤文 等/著



科学出版社



Principles and  
Applications of  
Analyzing a Citation Space

---

# 引文空间分析原理 与应用

CiteSpace实用指南

陈悦 陈超美 胡志刚 王贤文 等/著

科学出版社

北京

图书在版编目(CIP)数据

引文空间分析原理与应用: CiteSpace 实用指南/陈悦等著.  
—北京: 科学出版社, 2014. 9

ISBN 978-7-03-041752-7

I. ①引… II. ①陈… III. ①引文分析-图谱 IV. ①G353.1-64

中国版本图书馆 CIP 数据核字 (2014) 第 199975 号

责任编辑: 邹 聪 王首艳 / 责任校对: 彭 涛  
责任印制: 赵德静 / 封面设计: 无极书装  
编辑部电话: 010-64035853  
E-mail: houjunlin@mail. sciencep. com



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>

中国科学院印刷厂印刷

科学出版社发行 各地新华书店经销

\*

2014 年 9 月第 一 版 开本: 720×1000 1/16

2014 年 9 月第一次印刷 印张: 10 3/4

字数: 193 000

定价: 78.00 元

(如有印装质量问题, 我社负责调换)

# 目 录

CiteSpace 的精美图谱	1
绪论 CiteSpace 的关注度	5
<b>1 概述</b>	<b>11</b>
1.1 设计理念	13
1.2 概念模型	14
1.2.1 基于引文分析的学科基础	15
1.2.2 基于信息可视化的技术基础	16
1.3 理论基础	17
1.3.1 库恩的科学发展模式理论	17
1.3.2 普赖斯的科学前沿理论	18
1.3.3 博特的结构洞理论	18
1.3.4 信息觅食理论	19
1.4 系统要求	19
1.5 工作窗口	19
1.6 应用流程	20
1.6.1 数据采集及处理	22
1.6.2 参数功能选择	22
1.6.3 数据可视化	23
1.6.4 图谱解读	24
<b>2 初级启动</b>	<b>27</b>
2.1 启动 CiteSpace	28
2.2 运行内置数据的案例演示	31
2.3 数据下载	34
2.4 参数设定	38

2.4.1	设定时间分区	38
2.4.2	选取分析对象	39
2.4.3	分析对象的连接强度	39
2.4.4	分析对象的数据筛选	40
2.4.5	图谱修剪	40
2.4.6	图谱可视化方式	41
2.5	数据可视化	41
2.6	聚类生成	41
2.7	聚类标注	43
2.8	导出	44
2.8.1	可视化视图存储	44
2.8.2	可视化数据导出	44
<b>3</b>	<b>数据处理</b>	<b>45</b>
3.1	数据资源	46
3.1.1	WoS 数据库	46
3.1.2	PubMed 数据库	47
3.1.3	Derwent 数据库	47
3.1.4	CNKI 数据库	47
3.1.5	CSSCI 数据库	48
3.1.6	Scopus 数据库	48
3.1.7	Google Scholar 数据库	49
3.1.8	arXiv.org 数据库	49
3.1.9	ADS 数据库	49
3.1.10	NSF Awards	50
3.2	数据预处理	50
3.3	数据转换器和内置数据库	51
3.3.1	数据转换器	51
3.3.2	内置数据库	60
<b>4</b>	<b>网络模型</b>	<b>63</b>
4.1	分析对象的选择	64

4.1.1	施引文献	64
4.1.2	被引文献	65
4.2	阈值设置	66
4.2.1	时间阈值	66
4.2.2	连线阈值	66
4.2.3	节点阈值	67
4.3	网络修剪	68
4.3.1	最小生成树算法	69
4.3.2	寻径网络算法	69
<b>5</b>	<b>可视化交互</b>	<b>73</b>
5.1	可视化视图	74
5.1.1	聚类视图	75
5.1.2	时间线视图	76
5.1.3	时区视图	76
5.2	视图布局	77
5.3	标签调节	78
5.3.1	文献节点标签调节	78
5.3.2	词节点标签调节	78
5.3.3	聚类标签调节	80
<b>6</b>	<b>可视化设计</b>	<b>81</b>
6.1	节点显示	82
6.2	聚类显示	84
6.3	标签显示	86
6.4	背景显示	86
6.5	连线显示	86
6.6	过滤显示	86
6.6.1	查找	87
6.6.2	引用突现	88
6.6.3	分时连线	88
6.6.4	关键路径	89

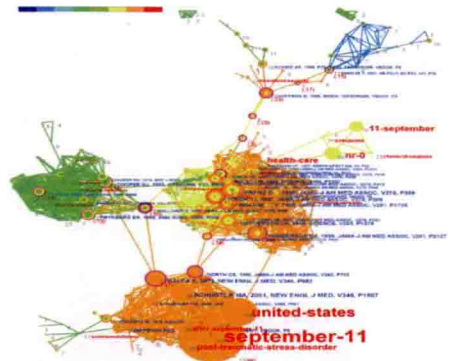
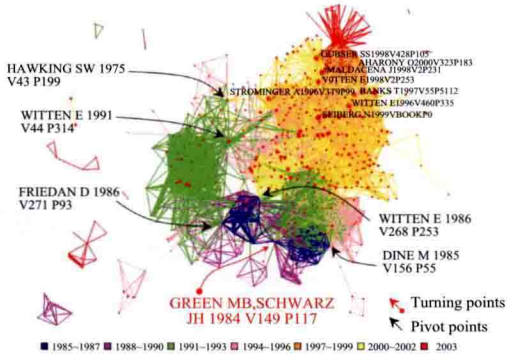
6.7	图示缩放	89
<b>7</b>	<b>关于节点和聚类的更多信息</b>	<b>91</b>
7.1	节点信息	93
7.1.1	引用历史	93
7.1.2	检索信息	94
7.1.3	隐藏/恢复节点	97
7.1.4	节点修正	98
7.1.5	节点汇总表	99
7.2	聚类信息	100
7.2.1	聚类信息列表	100
7.2.2	聚类施引文献信息	100
7.2.3	聚类汇总表	101
7.2.4	聚类浏览器	102
7.3	文本	103
7.3.1	向量空间模型	103
7.3.2	潜在语义分析	104
7.3.3	自动生成报表	104
<b>8</b>	<b>其他功能</b>	<b>109</b>
8.1	双图叠加	110
8.2	鱼眼视图	112
8.3	基于 Google Earth 的知识图谱	112
8.4	CiteSpace 的数据处理功能	113
8.5	概念树	115
<b>9</b>	<b>经典案例</b>	<b>117</b>
9.1	恐怖主义	118
9.1.1	研究目的	118
9.1.2	数据来源	118
9.1.3	图谱展示	119
9.1.4	案例总结	121

9.2	大规模灭绝 .....	121
9.2.1	研究目的 .....	122
9.2.2	数据来源 .....	122
9.2.3	图谱展示 .....	122
9.2.4	案例总结 .....	128
9.3	再生医学 .....	128
9.3.1	研究目的 .....	129
9.3.2	数据来源 .....	129
9.3.3	图谱展示 .....	129
9.3.4	信息显示 .....	132
9.3.5	案例总结 .....	137
9.4	科学学 30 年 .....	138
9.4.1	研究目的 .....	138
9.4.2	数据来源 .....	138
9.4.3	图谱展示 .....	139
9.4.4	案例总结 .....	140
<b>10</b>	<b>CiteSpace 的应用情况与常见问题</b> .....	<b>141</b>
10.1	555 篇国内使用 CiteSpace 工具的情况调查 .....	142
10.2	常见问题 .....	145
10.2.1	关于启动 .....	145
10.2.2	关于数据 .....	146
10.2.3	关于信息显示 .....	147
10.2.4	关于可视化 .....	151
	<b>参考文献</b> .....	<b>157</b>
	<b>常用词汇英汉对照</b> .....	<b>159</b>
	<b>致谢</b> .....	<b>163</b>



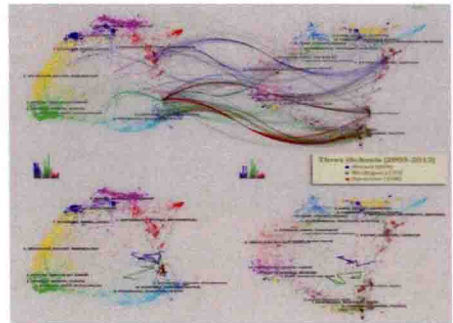
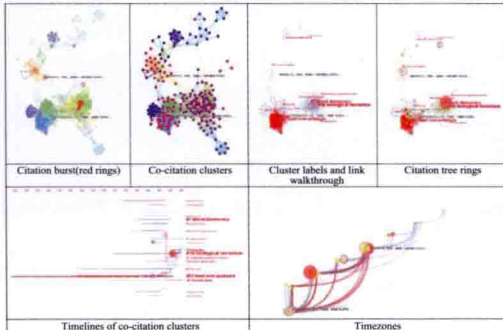


# CiteSpace 的精美图谱



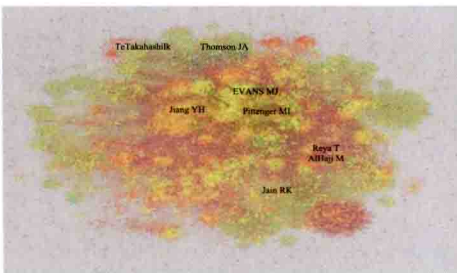
Chen C. CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, 2006, 57 (3): 359-377.

Chen C. Searching for intellectual turning points; progressive knowledge domain visualization. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 2004, 101 (1): 5303-5310.



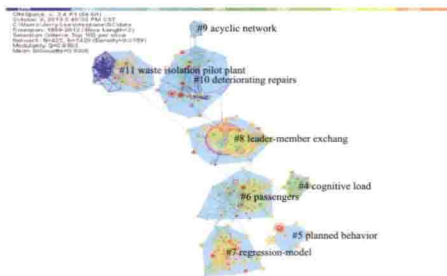
<http://cluster.ischool.drexel.edu/~cchen/citespace/gallery/gallery.pdf>

Chen C, Leydesdorff L. Patterns of connections and movements in dual-map overlays; a new method of publication portfolio analysis. *Journal of the Association for Information Science and Technology*, 2014, 65 (2): 334-351.

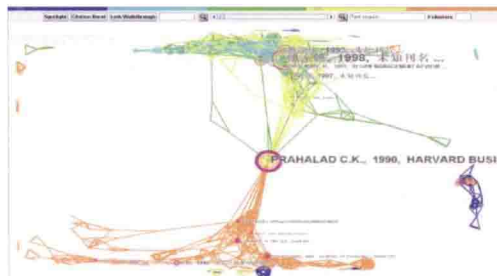


Chen C, Hu Z, Liu S, et al. Emerging trends in regenerative medicine: a scientometric analysis in CiteSpace. *Expert Opinions on Biological Therapy*, 2012, 12 (5): 593-608.

刘则渊, 胡志刚, 王贤文. 30年中国科学学历程的知识图谱展现. *科学学与科学技术管理*, 2010, 31 (5): 17-23.



李杰. <http://blog.sciencenet.cn/blog-554179-729837.html>.



陈悦. 中国创新管理论文的文獻共被引可視化图谱 (1998~2002)。

<http://facebook.com/pages/CiteSpace/276625072366558?ref=profile>





# 绪论 CiteSpace 的关注度

2004年,在作者共被引分析(author cocited analysis, ACA)的发源地——美国德雷克塞尔大学信息科学与技术学院(现为 College of Computing and Informatics, Drexel University),陈超美(Chaomei Chen)基于引文分析理论将 ACA 应用范围进行了扩展,应用 Java 计算机编程语言开发了 Information Visualization - CiteSpace 信息可视化软件,这是近年来在信息分析中最具有特色和影响力的信息可视化软件。软件开发于 2004 年 9 月,其版本不断升级和更新,当前版本为 CiteSpace 3.8.R3 (2014 年 5 月)。CiteSpace 软件被至少超过 60 个国家上万次启用(表 0-1,图 0-1),它已经成为科学计量学普遍采用的新工具。在大连理工大学 WISE 实验室的推广下,中国已成为最大的使用国,尤其是 2009 年由 WISE 实验室举办了首届全国培训班后,CiteSpace 在中国的使用范围迅速扩大,使用数量急剧增长。

表 0-1 CiteSpace 网站访问情况 (2008~2010 年)

排名	2008 年			2009 年			2010 年 8~12 月		
	国家/地区	访问量/次	占总访问量(7430 次)的百分比/%	国家/地区	访问量/次	占总访问量(11 737 次)的百分比/%	国家/地区	访问量/次	占总访问量(12 065 次)的百分比/%
1	美国	2 403	32.34	中国内地	4 494	<b>38.29</b>	中国内地	7 323	<b>60.70</b>
2	中国内地	1 594	<b>21.45</b>	美国	2 583	22.01	美国	1 507	12.49
3	德国	519	6.99	德国	846	7.21	德国	679	5.63
4	西班牙	401	5.40	法国	383	3.26	法国	360	2.98
5	瑞典	255	3.43	意大利	311	2.65	意大利	236	1.96
6	英国	251	3.38	英国	291	2.48	英国	169	1.40
7	韩国	234	3.15	西班牙	286	2.44	西班牙	161	1.33
8	法国	228	3.07	巴西	279	2.38	巴西	143	1.19
9	加拿大	172	2.31	韩国	209	1.78	韩国	121	1.00
10	芬兰	165	2.22	中国台湾	198	1.69	中国台湾	101	0.84



图 0-1 CiteSpace 用户在全球的分布示意图 (2013 年 8 月~2014 年 3 月)

绿点表示 1~10 个用户；黄点表示 10~100 个用户；红点表示 100~1000 个用户；大红点表示多于 1000 个用户。

陈超美致力于信息可视化的研究，在科学发现和科学知识图谱（mapping knowledge domains, MKD）领域获得了丰硕的成果，这些研究成果积淀成了今天的 CiteSpace，并产生了极其深远的影响（图 0-2）。

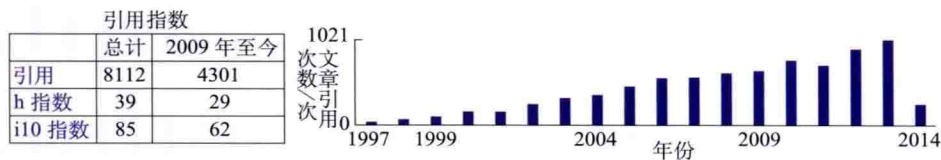


图 0-2 学术 Google 中体现的陈超美研究成果的影响力（2014-05-20）

关于 CiteSpace 的核心文献有以下 7 篇（其中一篇为中文译版），表 0-2 列出了这 7 篇文献在学术 Google 中的被引情况（2014-05-20），显示出学界对 CiteSpace 的关注持续增长：

A. Chen C. Mapping Scientific Frontiers [M]. London: Springer, 2003.

B. Chen C. Searching for intellectual turning points: progressive knowledge domain visualization [J]. Proceedings of the National Academy of Sciences, 2004, 101 (suppl 1): 5303-5310.

C. Chen C. Information Visualization: Beyond the Horizon [M]. Springer Science & Business, 2006.

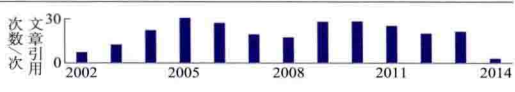

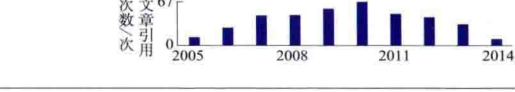
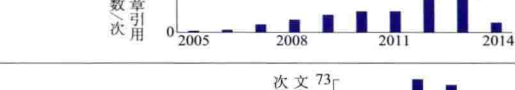
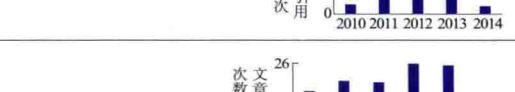
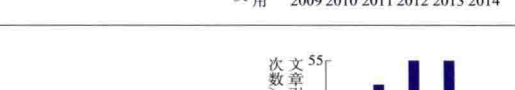

D. Chen C. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature [J]. Journal of the American Society for information Science and Technology, 2006, 57 (3): 359-377.

陈超美著，陈悦，侯剑华等译. CiteSpace II: 科学文献中新趋势与新动态的识别与可视化 [J]. 情报学报, 2009 (3): 401-421.

E. Chen C, Chen Y, Horowitz M, et al. Towards an explanatory and computational theory of scientific discovery [J]. Journal of Informetrics, 2009, 3 (3): 191-209.

F. Chen C, Ibekwe-SanJuan F, Hou J. The structure and dynamics of cocitation clusters: a multiple-perspective cocitation analysis [J]. Journal of the American Society for Information Science and Technology, 2010, 61 (7): 1386-1409.

表 0-2 CiteSpace 核心文献的影响力

CiteSpace 的核心文献	被引频次	主要内容	年度被引
A. Chen C-2003	268	探索科学前沿的历史和哲学背景及先驱者的工作	
B. Chen C-2004	395	CiteSpace I	
C. Chen C-2006	397	信息可视化	
D. Chen C-2006	807	CiteSpace II 2006 英文版	
陈超美-2009	184	CiteSpace 2009 中译版	
E. Chen C-2009	104	科学发现的可解释性计算理论	
F. Chen C-2010	157	CiteSpace 的进一步扩展: 自动标识聚类主题辅助图谱的可解读性	

陈超美于 2010 年 10 月在科学网上建立了专门发布 CiteSpace 相关信息的个人网站 (<http://blog.sciencenet.cn/home.php?mod=space&uid=496649>), 截至 2014 年 5 月 20 日, 他发表了 60 篇博文, 拥有 434 位好友, 活跃度为 110 150, 已有 248 681 人来访过。

CiteSpace 在中国产生了极大的影响。大连理工大学分别在 2009 年 9 月、2010 年 4 月和 2013 年 8 月举办了 3 次关于“科学知识图谱原理、方法与应用”的培训班, 培训人员合计 200 余人。中国图书馆学会、中国科学院文献情报中心、上海市科技情报学会等也举行过多次涉及 CiteSpace 使用的培训班。大连理工大学从 2012 年为全校的硕士研究生和博士研究生开设了“学术规范与写作”(16 学时) 公共课, 其中有 4 个学时用来讲解 CiteSpace 的基本操作和如何应用 CiteSpace 做文献综述。2012 年, CiteSpace 的爱好者们自发形成了第一个 QQ 群 (261664017), 并很快满员。该 QQ 群由大连理工大学刘盛博管理, 他曾在



德雷克塞尔大学在陈超美的指导下完成了博士生联合培养项目（2010~2012年）。2012年11月，CiteSpace爱好者又自发创建了第二个QQ群（281555946），目前也已满员。两个群规模都是500人，目前共1000人。这两个QQ群的主要功能是CiteSpace使用交流，讨论使用CiteSpace软件时遇到的困惑，群中讨论无法解决的问题统一提交给陈超美帮助解决。群共享中定期整理群中讨论的CiteSpace常见问题及解决方案，为大家学习CiteSpace提供便利条件。

其他常用的科学知识图谱绘制软件工具包括VOSviewer、Sci<sup>2</sup>、CoPalRed、INSPIRE、VantagePoint、Bibexcel等。VOSviewer (<http://www.vosviewer.com/download>)是由荷兰莱顿大学的艾克(Nees Jan van Eck)博士和瓦尔特曼(Ludo Waltman)博士合作开发的一款可视化软件，它提供了4种可视化方案，即标签视图(Label View)、密度视图(Density View)、聚类密度视图(Cluster Density View)和散点视图(Scatter View)。图0-3是利用VOSviewer软件自带的数据库绘制的期刊共被引密度视图。Sci<sup>2</sup> (<https://sci2.cns.iu.edu/user/index.php>)是美国印第安纳大学布卢明顿分校图书馆与信息科学学院信息基础设施网络科学中心及信息可视化实验室开发的一个专门为科学研究而设计的软件，它是一个模块化的工具集，支持对微观(个体)、中观(局部)及宏观(全局)三个层面的数据进行基本统计、时间、空间、主题和网络层面的分析以及可视化(图0-4)。CoPalRed是由西班牙格拉纳达大学的EC研究团队开发的商业软件，使用共同单元来分析文献，它包括结构分析、战略分析和动态分析三个方面(图0-5)。INSPIRE (<http://in-spire.pnnl.gov>)是由美国西北太平洋国家实验室开发的，它提供Galaxies星空分布和ThemeScape地貌图两种可视化图谱，也提供系列工具来发现隐藏的知识，如Time Slicer(图0-6)。美国搜索技术公司(Search Technology Inc.)开发的免费文本挖掘和可视化软件VantagePoint (<https://www.thevantagepoint.com/login.html>)，可从大量的机构数据中发现模式和关系，快速确定知识内容(图0-7)。Bibexcel (<http://www8.umu.se/inforsk/Bibexcel>)是由瑞典于莫奥大学的佩尔森(Olle Persson)教授开发的一款文献计量学软件。该软件可以对从Web of Science(WoS)、Scopus等数据库下载的文献题录数据进行处理，实现一些基础的文献计量学功能，如频次分布、共现分析等(图0-8)。