如何使用CiteSpace进行文献的共被引分析——以恐怖主义研究为例

李杰

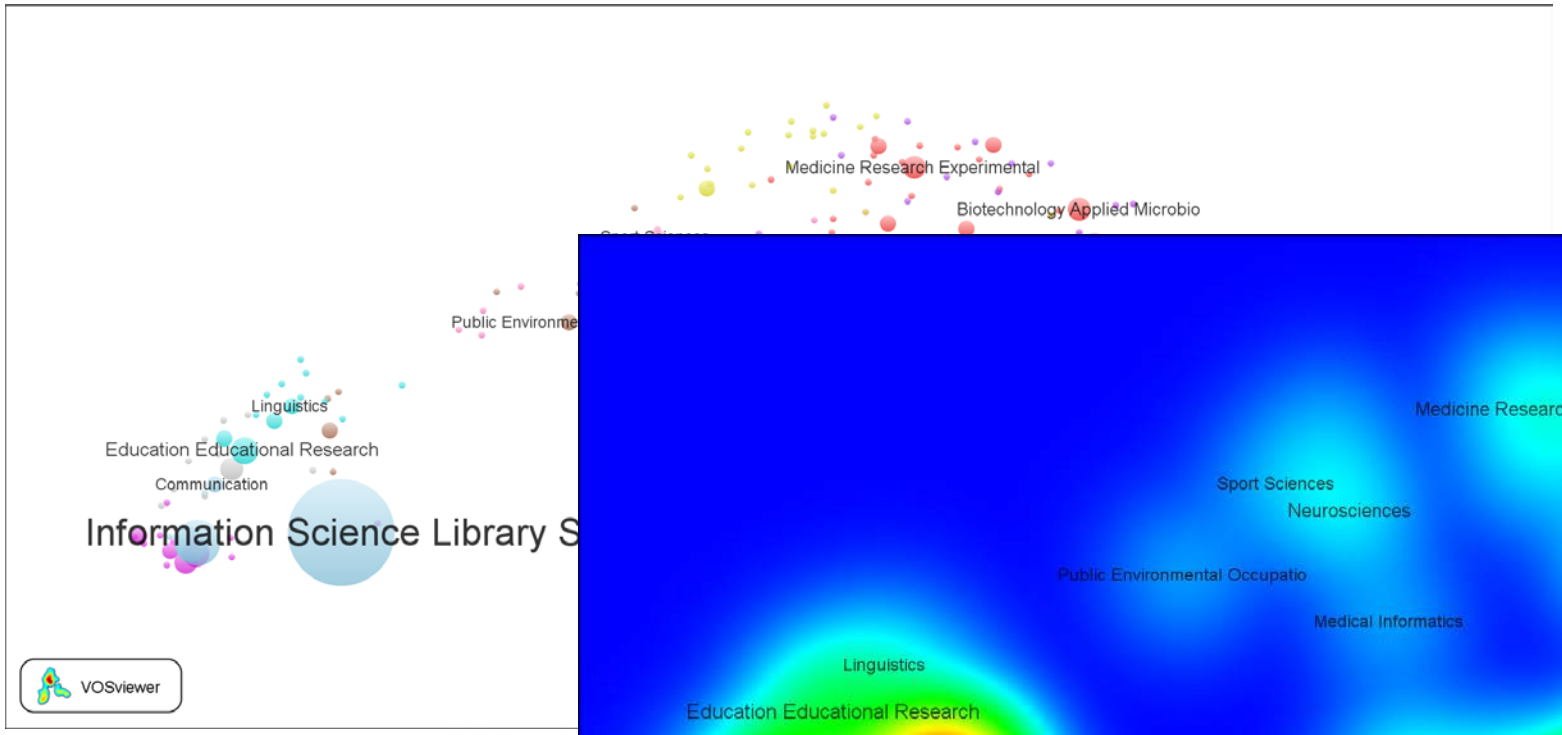
首都经济贸易大学 安全与环境工程学院

个人主页：http://blog.sciencenet.cn/u/jerrycueb

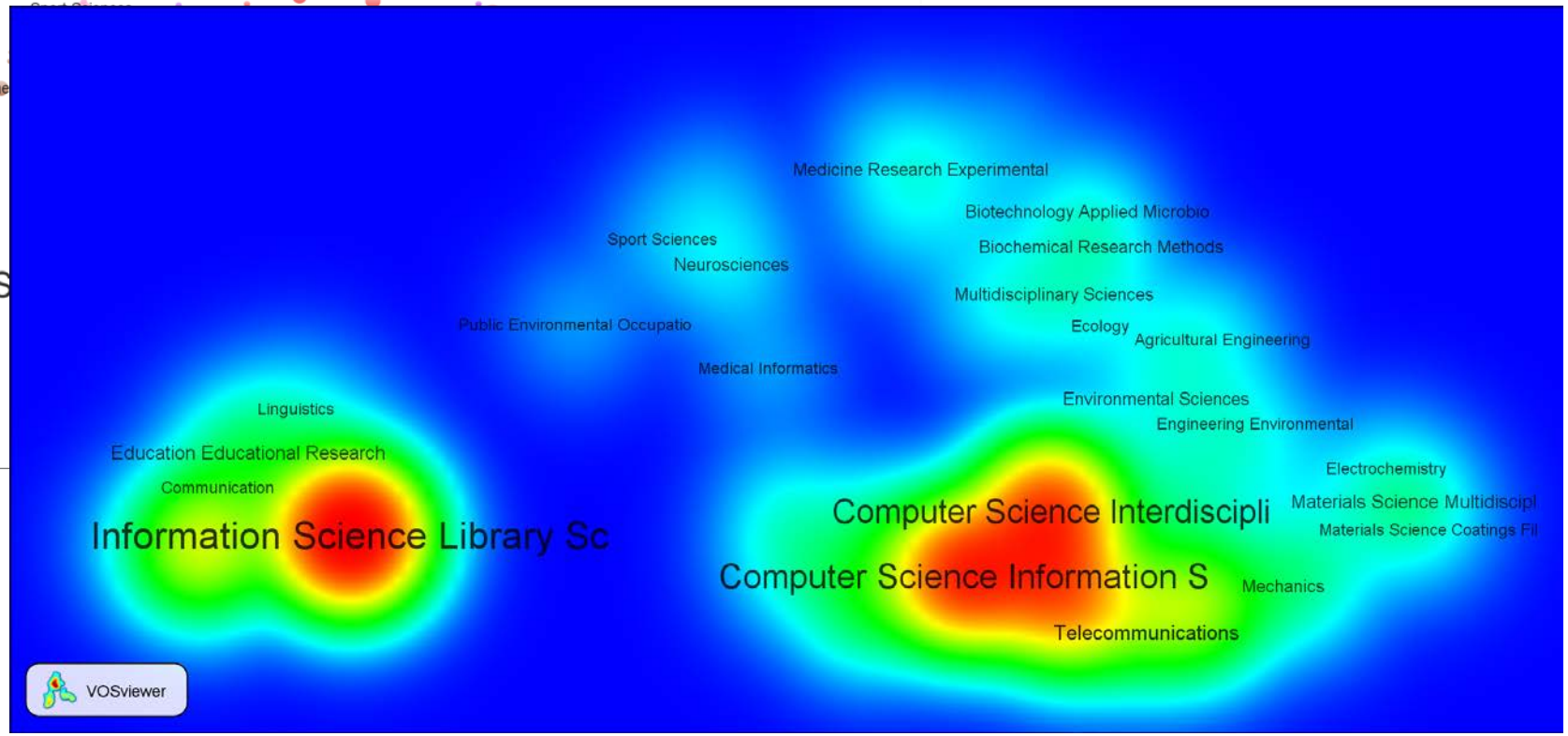
写在前面

- 当前文献图谱软件有十余种之多，各个软件都有其不同的优势。如VOSviewer在主题聚类方面清晰详细；SCI2在主题词分析时更加灵活；HistCite则对所下载数据集的文献互引按照时间以网路形式呈现（目前作者认为将很快被CitNetExplorer tool代替）；BibExcel则以数据集原始数据的Tag为依据，提供了多种多样的文献分析功能（需要外部的软件协助完成可视化，如Gephi,VOSviewer,Pajek等）；CiteSpace以其强大的文献共被引分析而知名（恐怖主义研究和生物大灭绝），且随着不断的发展算法和功能不断优化。
- 当前CiteSpace已经被广泛应用于计算机科学、信息科学以及医学等60多个领域（根据WoS分类统计，参见下页）

引用CiteSpace经典文献的施引文献的领域分布



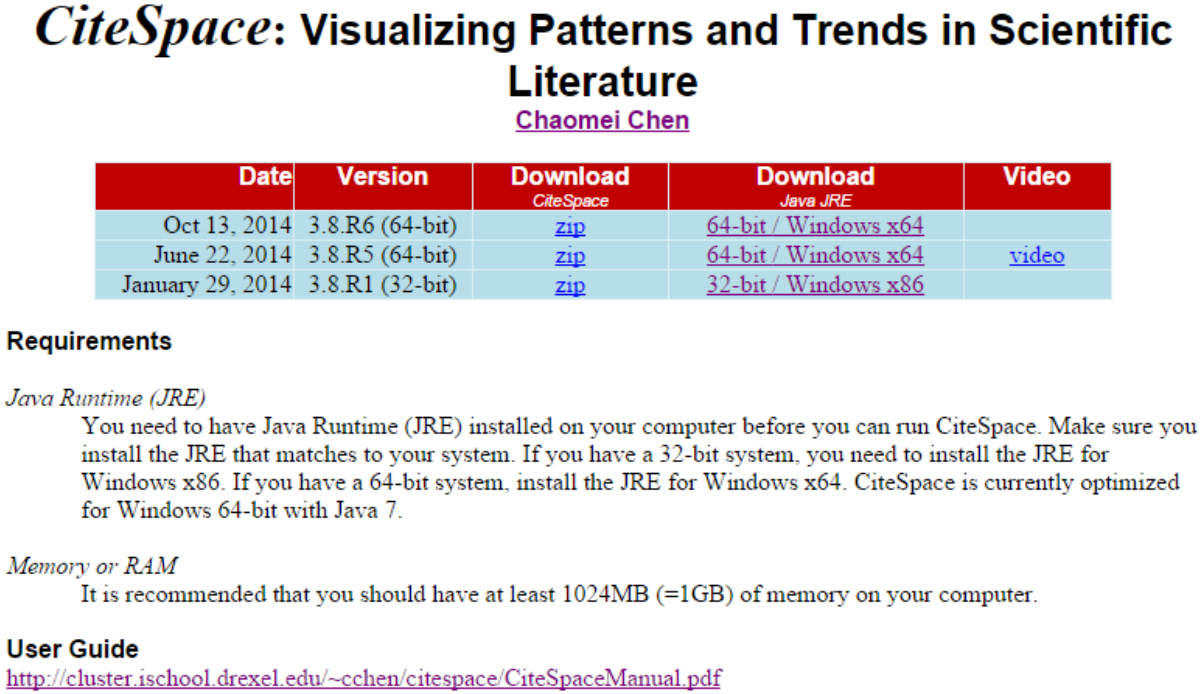
Label View



Density View

1. 准备软件

- 登陆CiteSpace主页下载软件
- http://cluster.ischool.drexel.edu/~cchen/citespace/download.html



***CiteSpace*: Visualizing Patterns and Trends in Scientific Literature**

Chaomei Chen

Date	Version	Download *CiteSpace*	Download *Java JRE*	Video
Oct 13, 2014	3.8.R6 (64-bit)	zip	64-bit / Windows x64	
June 22, 2014	3.8.R5 (64-bit)	zip	64-bit / Windows x64	video
January 29, 2014	3.8.R1 (32-bit)	zip	32-bit / Windows x86	

Requirements

Java Runtime (JRE)

You need to have Java Runtime (JRE) installed on your computer before you can run CiteSpace. Make sure you install the JRE that matches to your system. If you have a 32-bit system, you need to install the JRE for Windows x86. If you have a 64-bit system, install the JRE for Windows x64. CiteSpace is currently optimized for Windows 64-bit with Java 7.

Memory or RAM

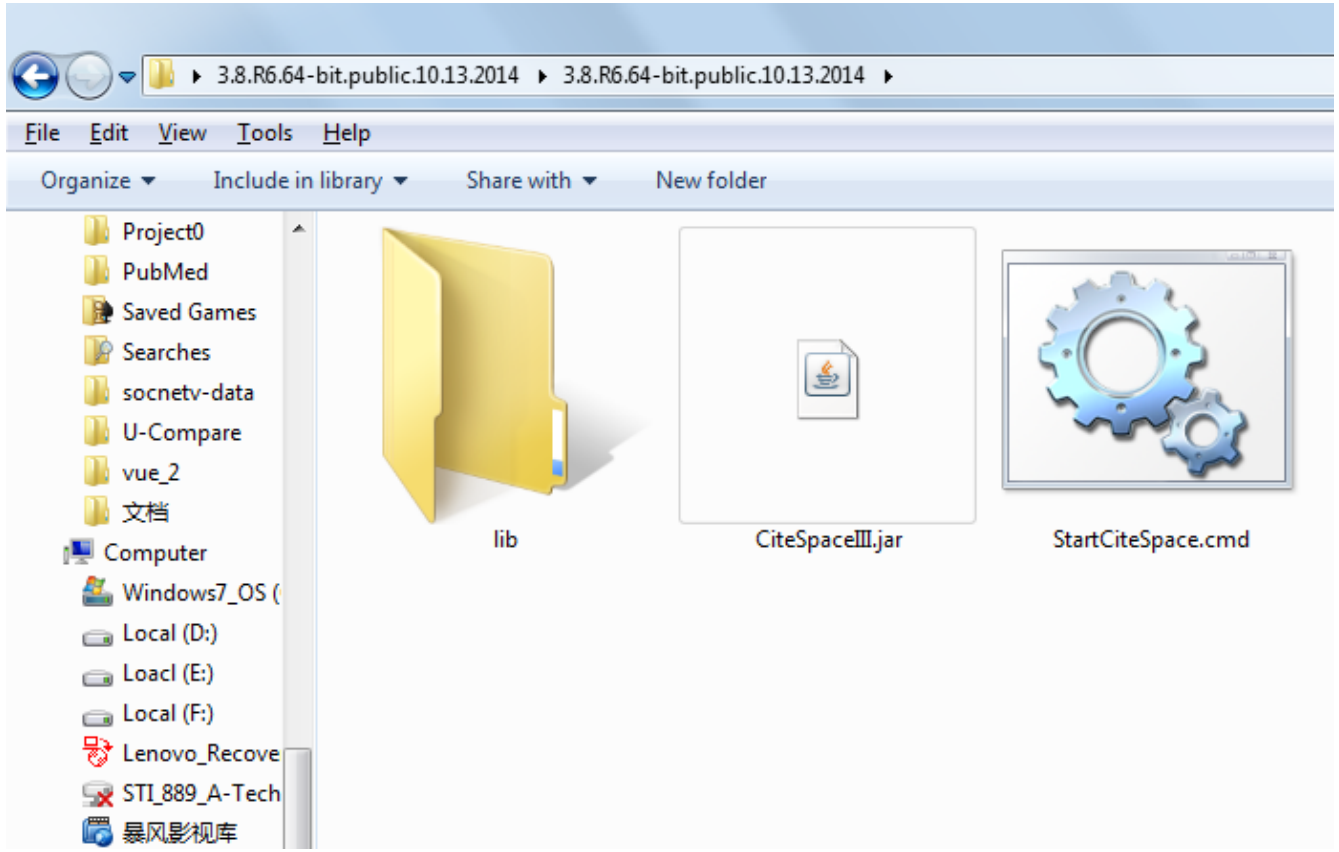
It is recommended that you should have at least 1024MB (=1GB) of memory on your computer.

User Guide

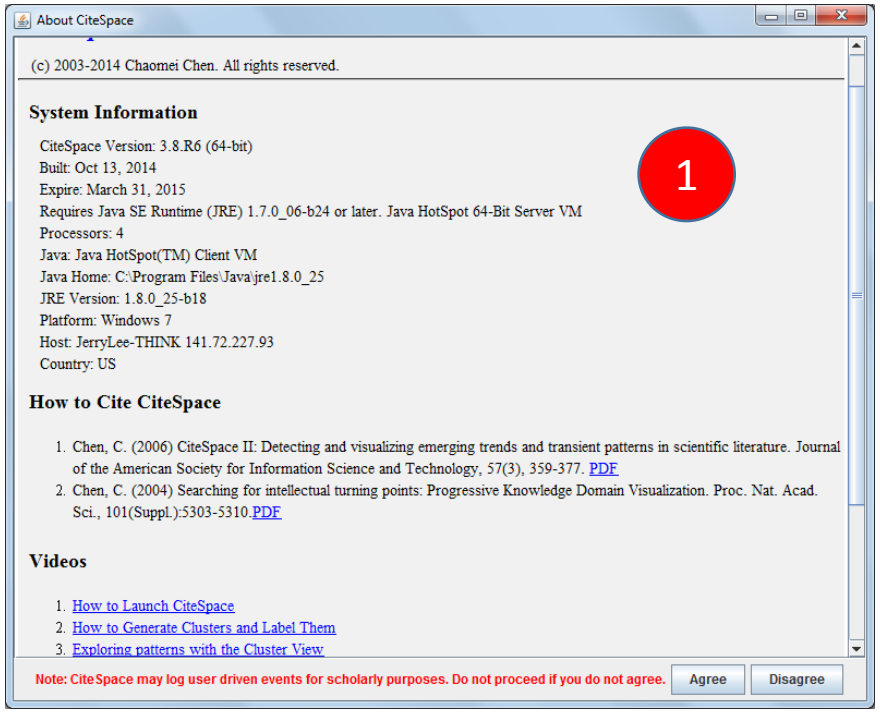
http://cluster.ischool.drexel.edu/~cchen/citespace/CiteSpaceManual.pdf

1. 准备软件

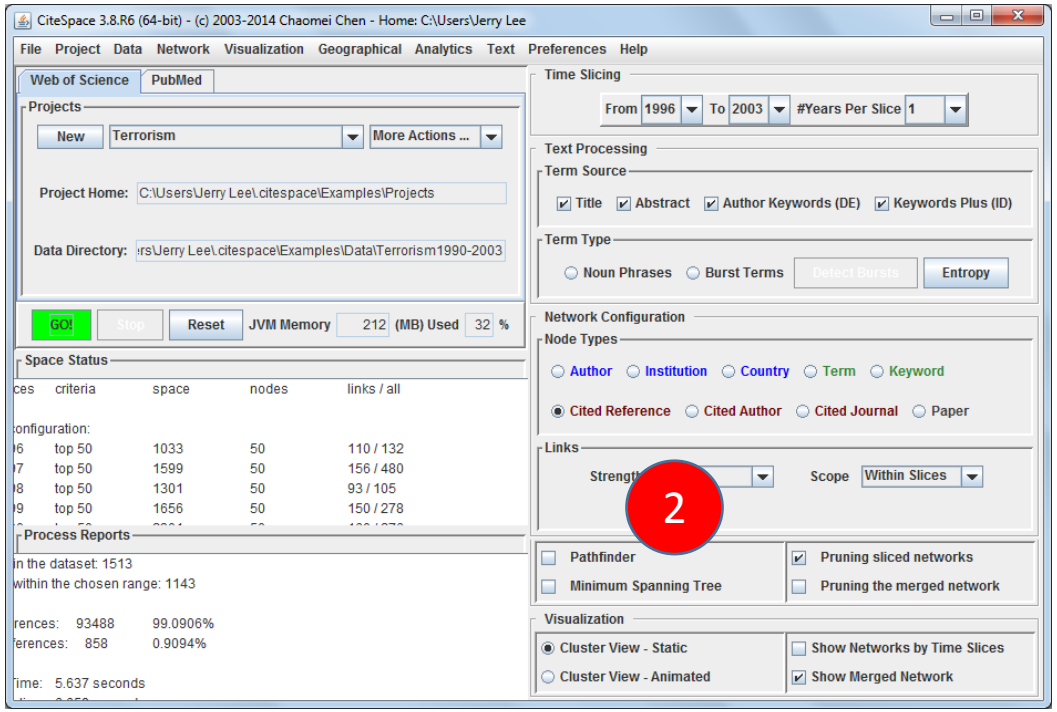
- 下载后解压，双击CiteSpaceIII.jar即可打开软件。



2.运行软件

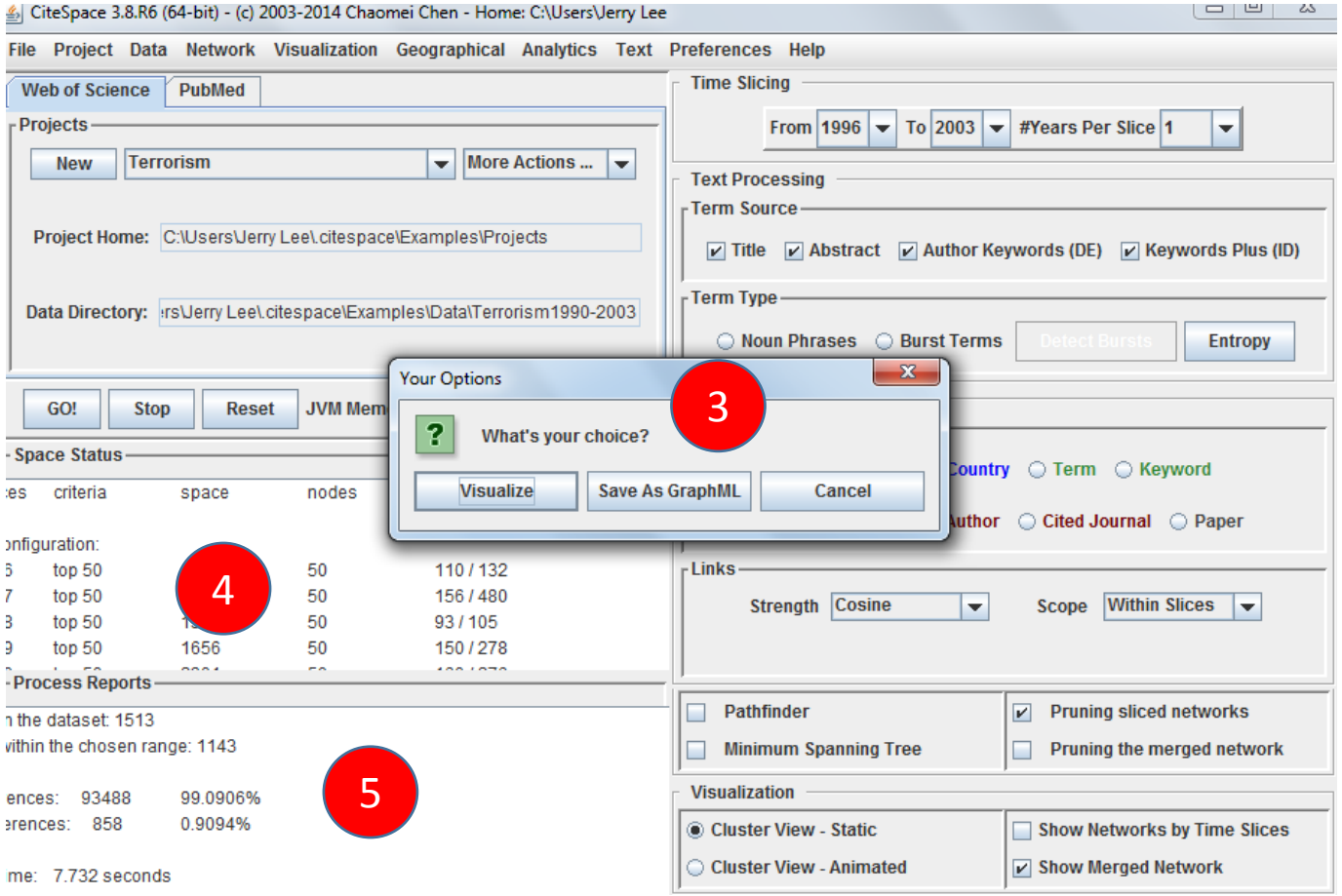


点击“Agree”进入软件界面



首次运行取消去网络进行的任何裁剪运算

3.点击“Go”进行运算

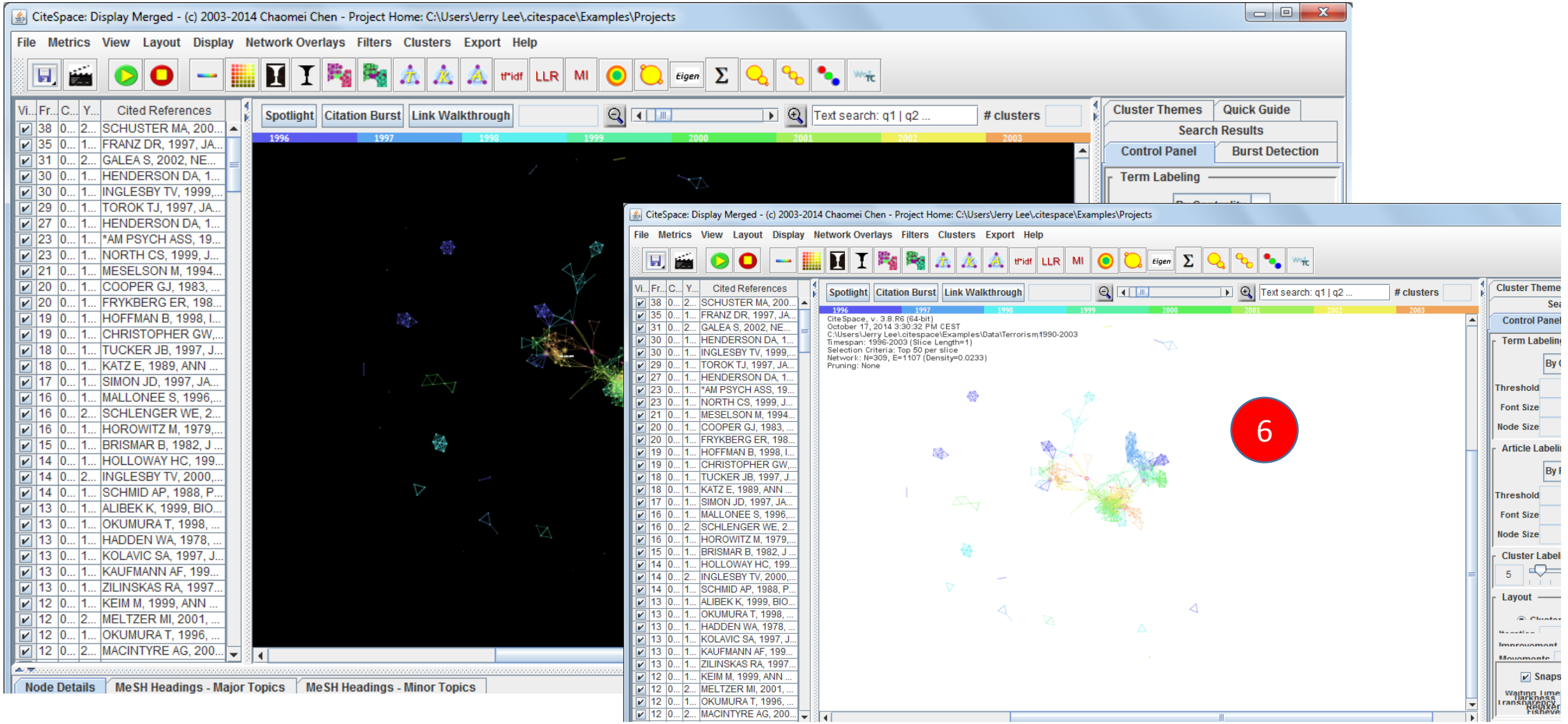


运行结束后点击“Visualize”进入可视化界面

4 结果按照预设条件运行的情况

5 网络基本参数及其运行基本参数

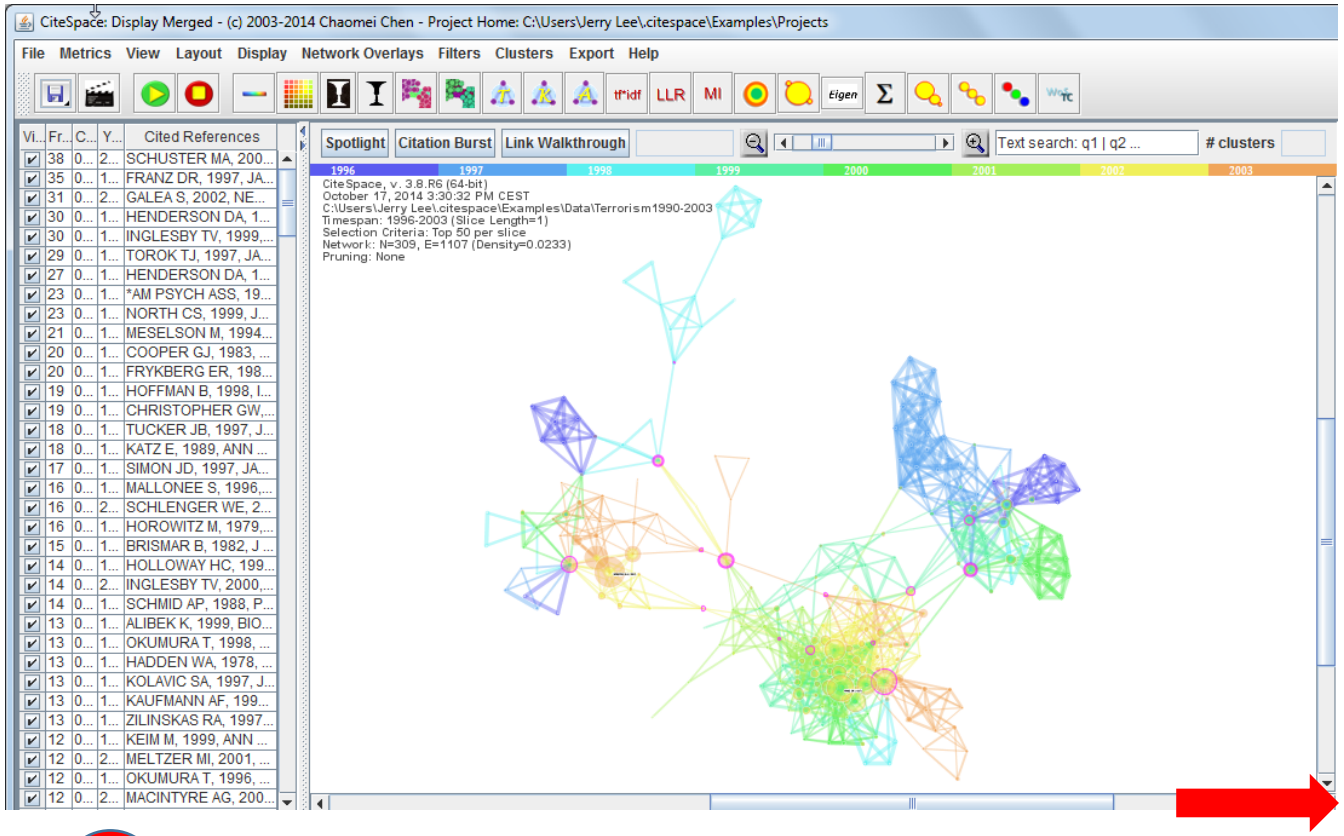
3.可视化结果

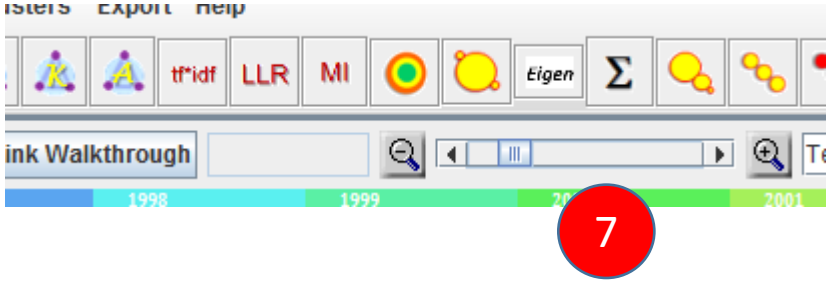


6 网络时时运动的，当网络布局稳定后背景会转成白色

4.可视化结果的调整

- 由于结果视图很小，因此要放大原图





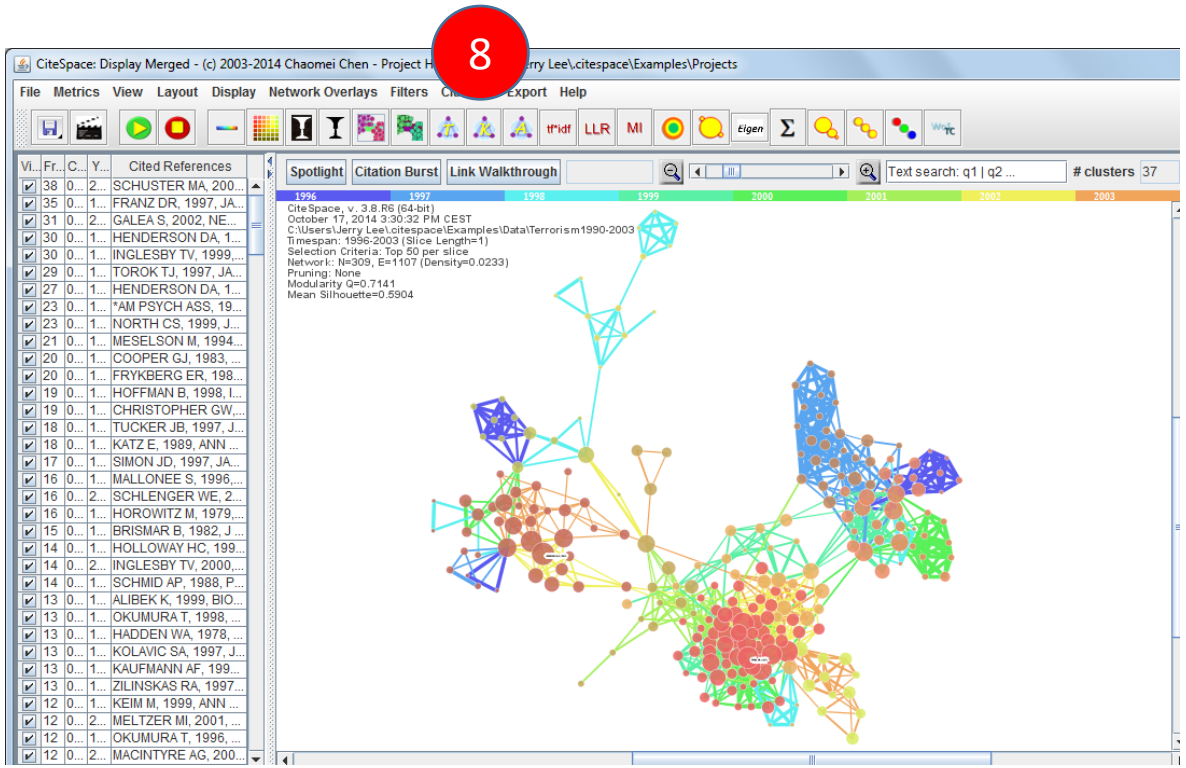
7

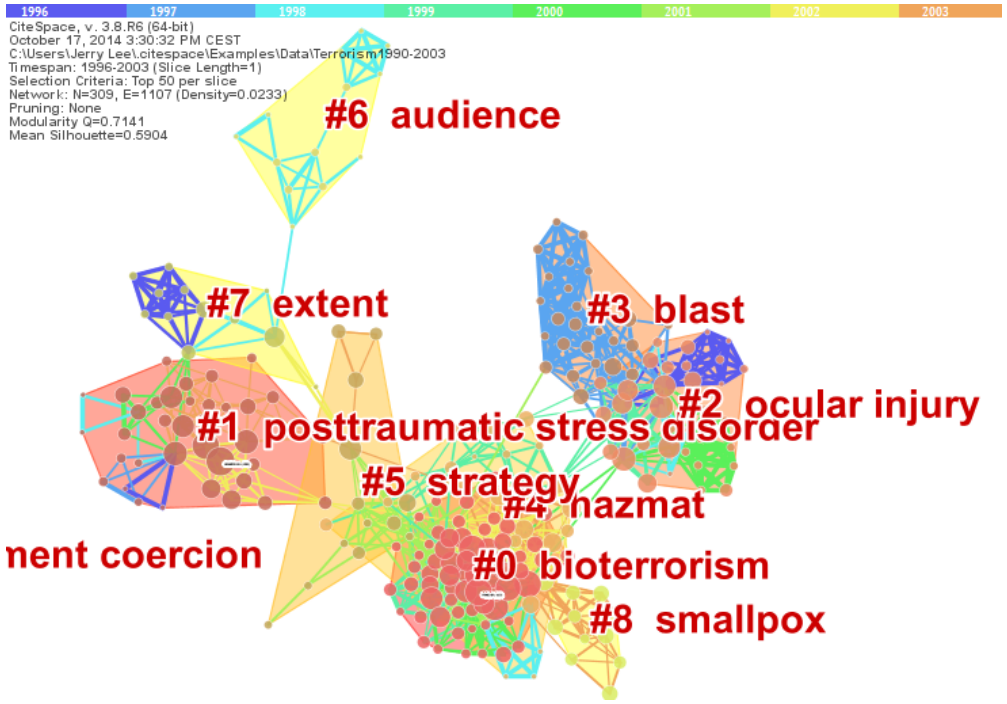
调整图形上下位置

调整图形左右位置

7 向右侧拖动放大工具条，并及时调整图形的二维位置

4.对共被引网络进行聚类



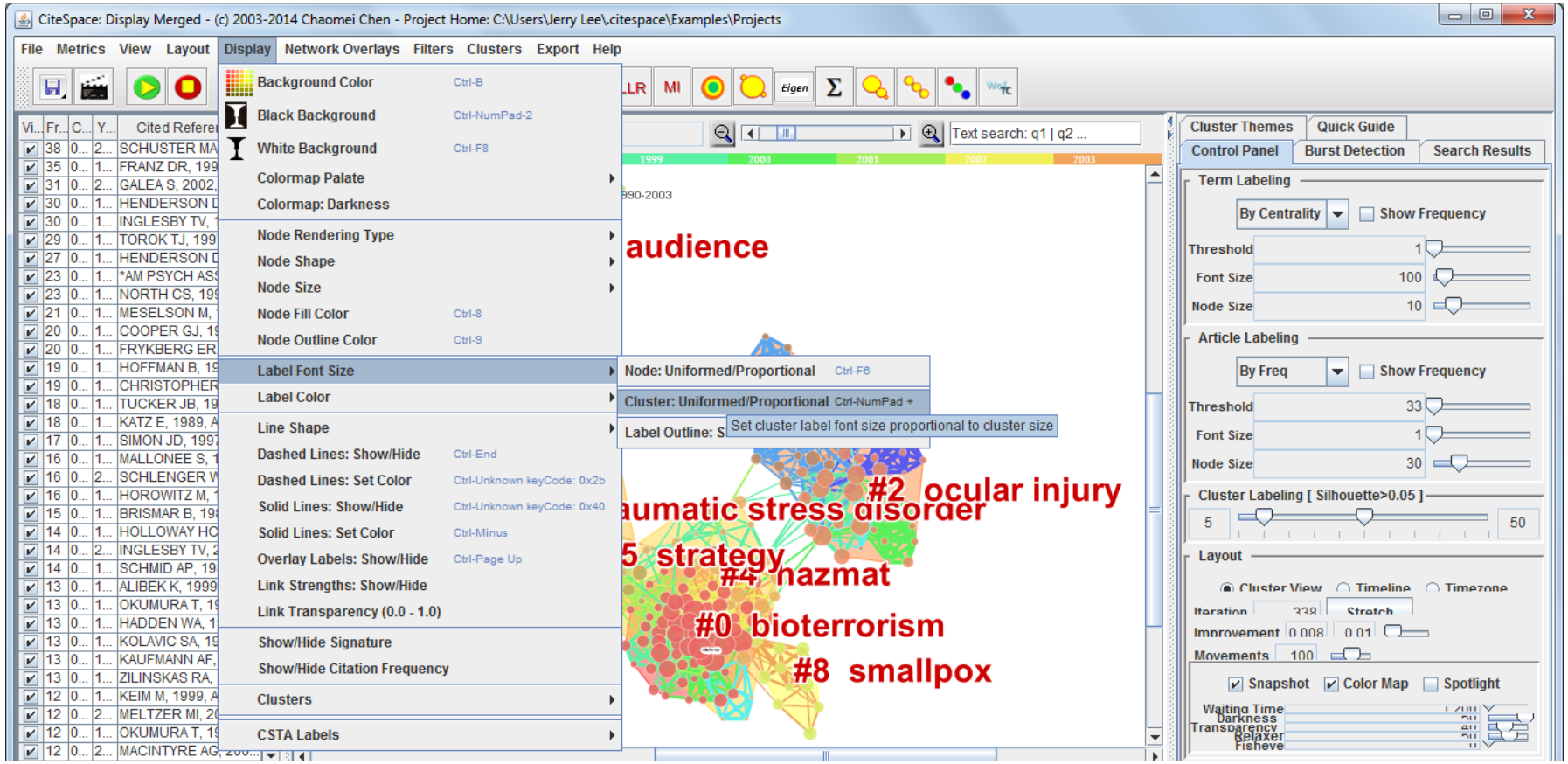


8 点击聚类图标  此时节点的属性将发生变化，聚类结果后点击  （从施引文献的标题中提取聚类命名术语）结果如右图。

为了增加结果的可读性，特别是显示重要的信息。可以进一步使用相关功能对图谱进行优化和调整。

5.对共被引网络聚类进行调整

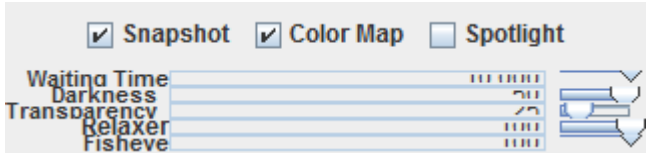
- 对聚类的标签进行调整（按照聚类规模进行显示）



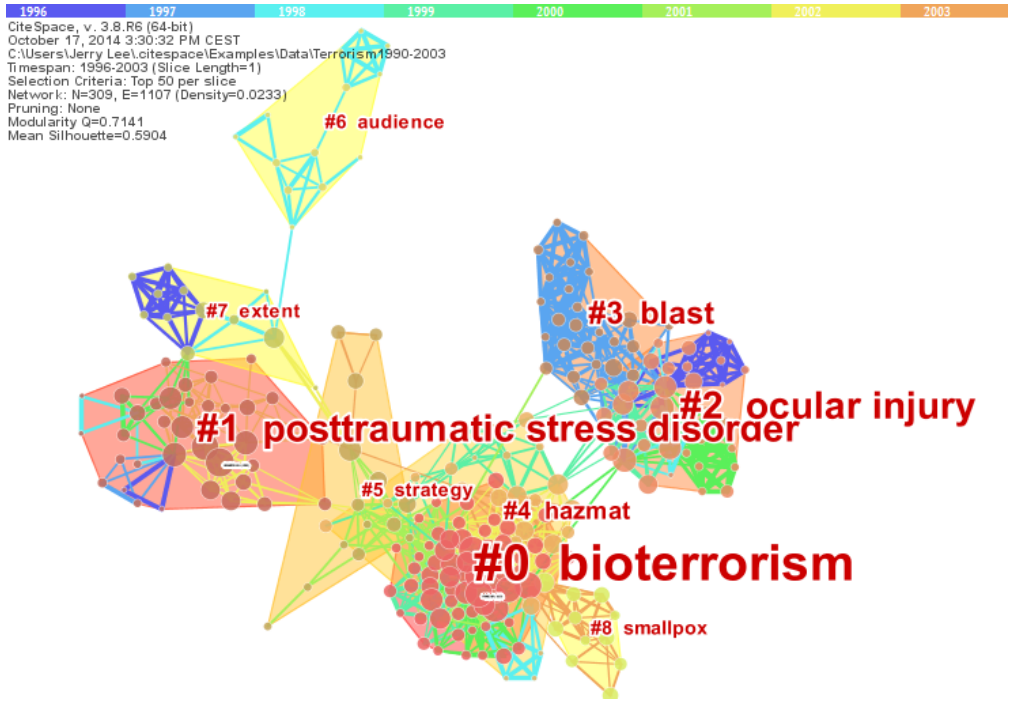
5.对共被引网络聚类进行调整

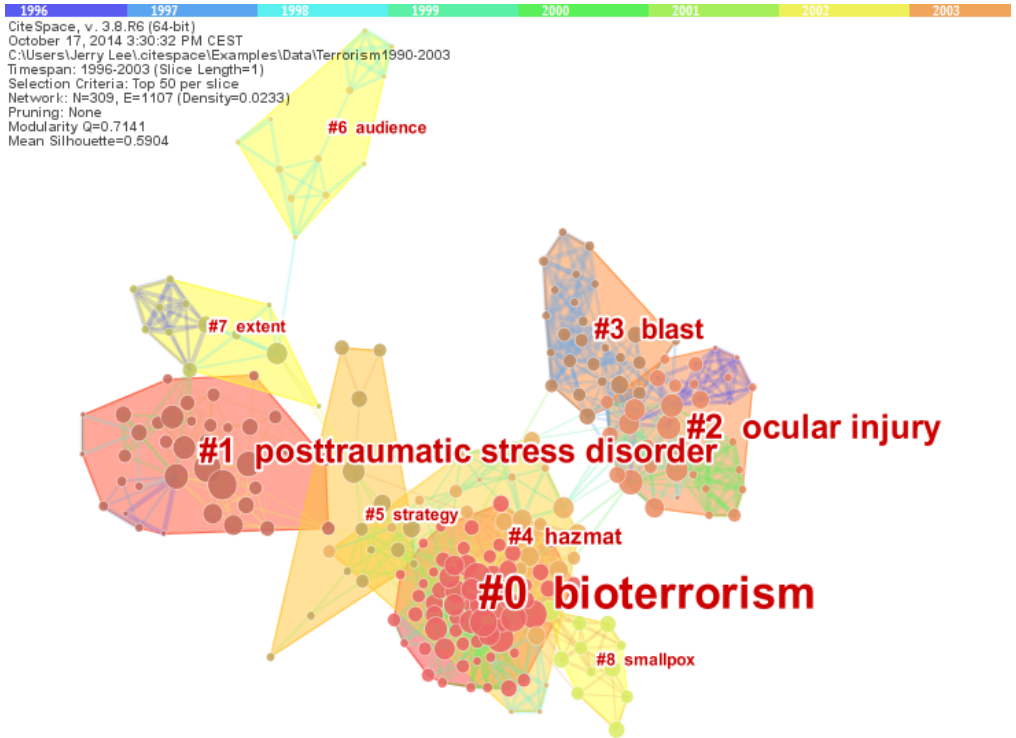
- 拖动 

对聚类标签进行调整

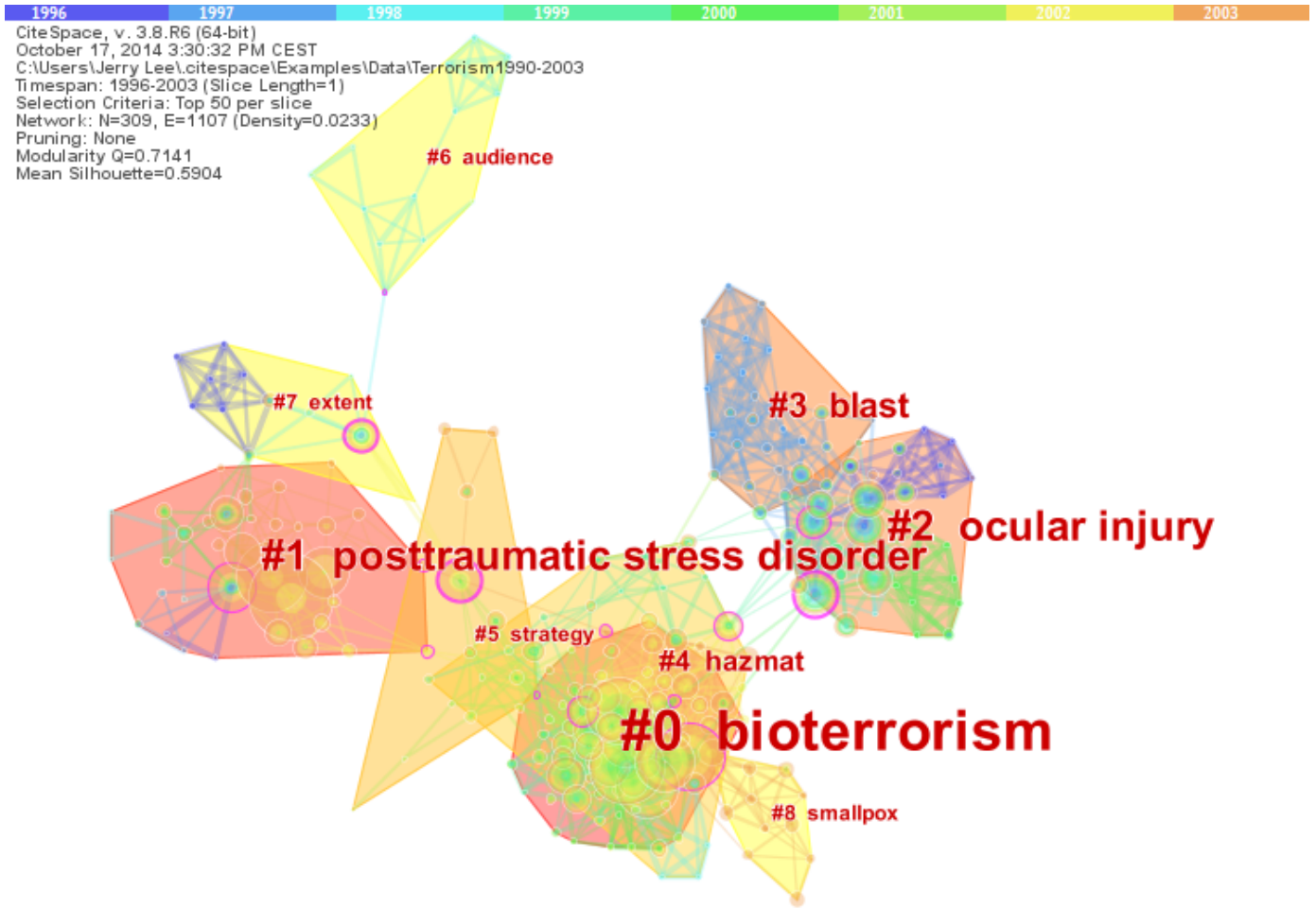
- 拖动 

- 对连线的透明度进行调整进行调整





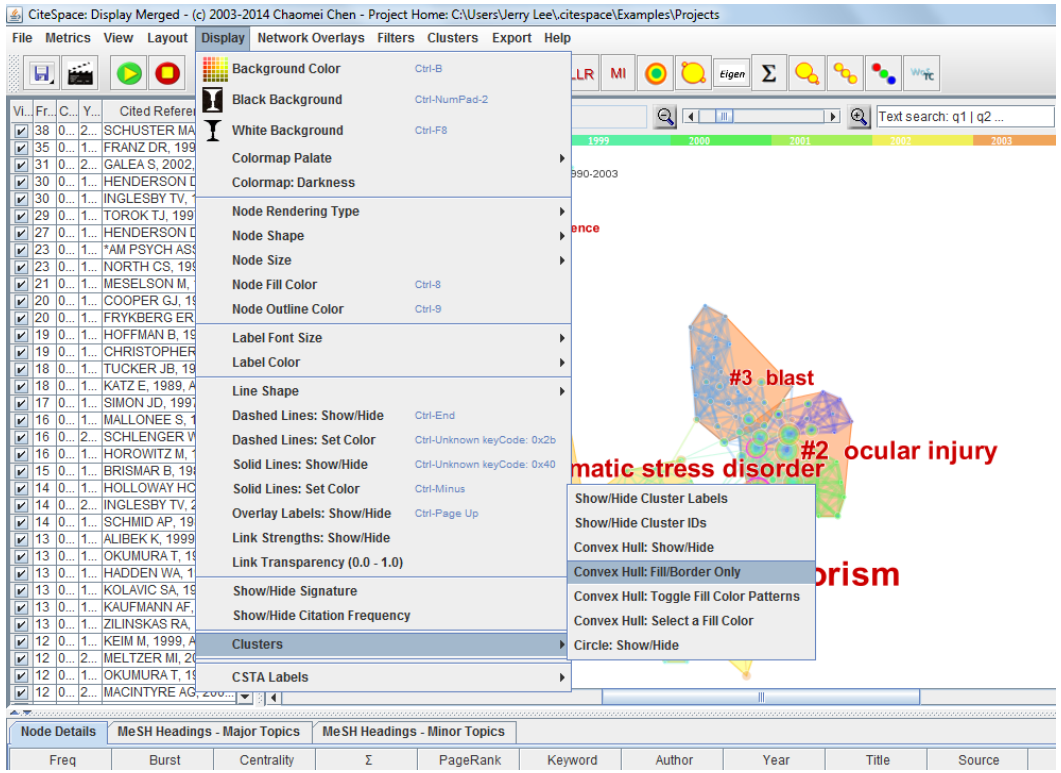
5.对共被引网络聚类进行调整

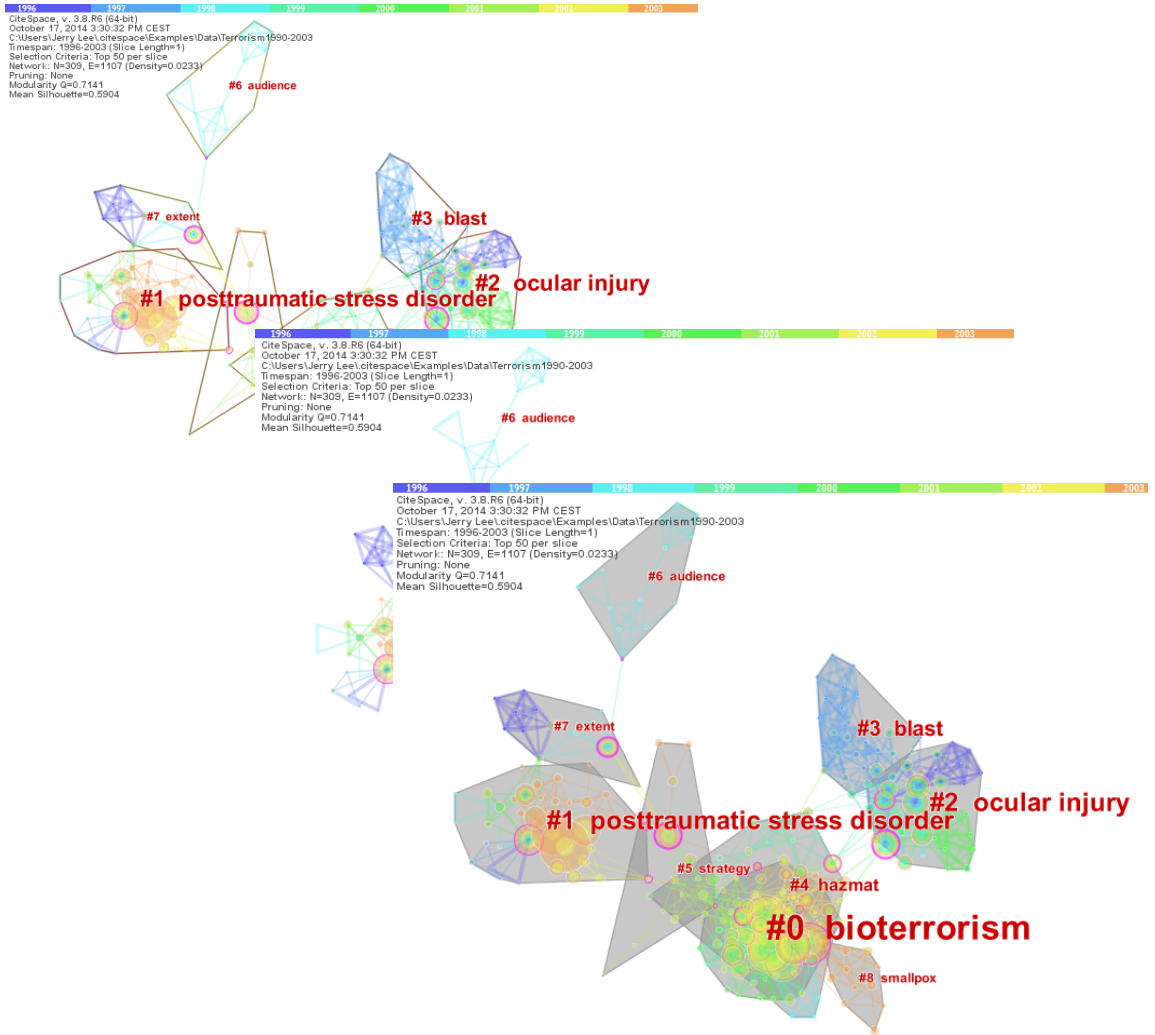


- 点击 

对节点属性进行调整

5.对共被引网络聚类进行调整

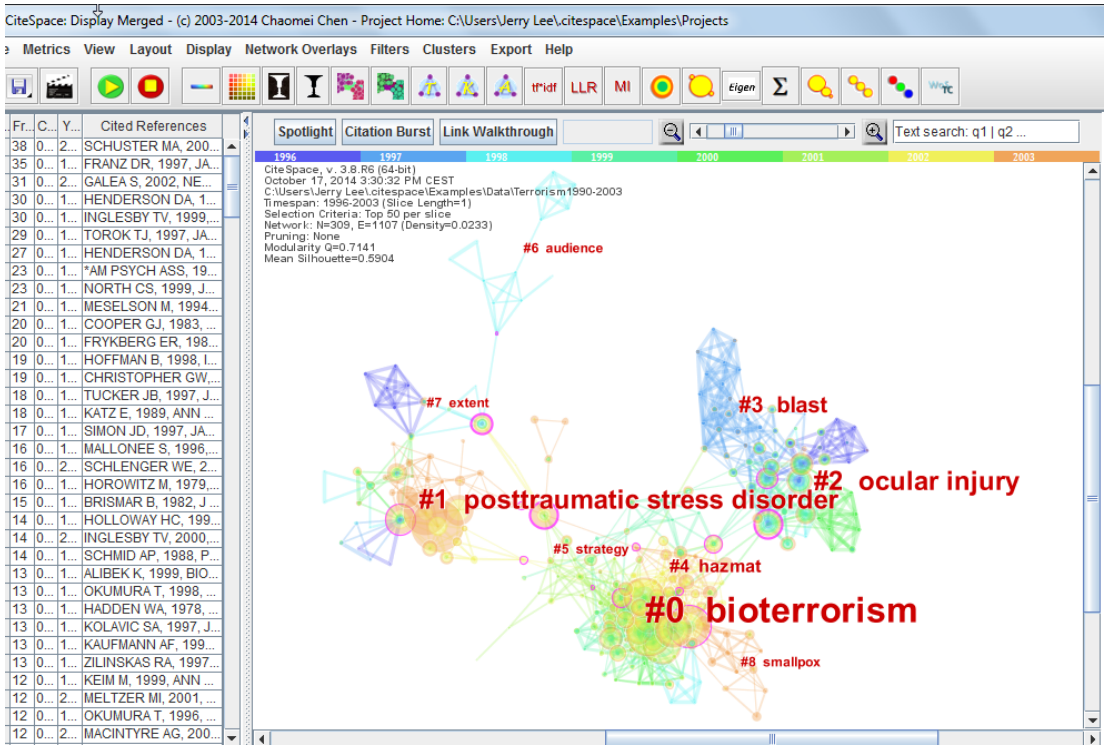


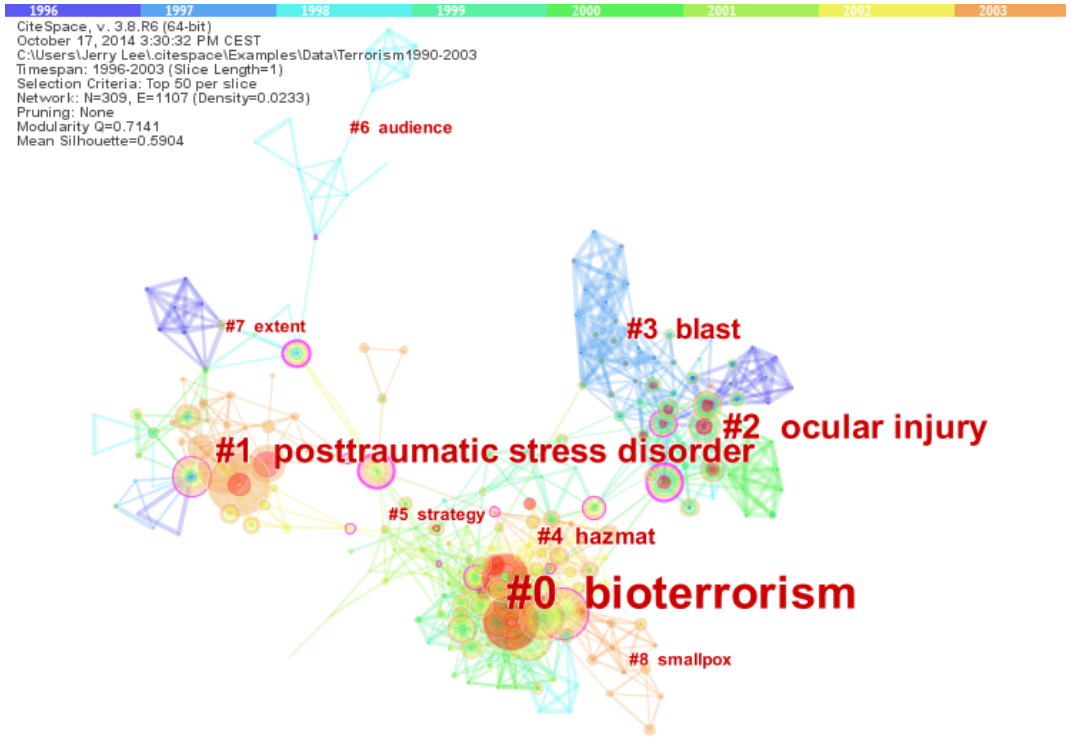


对聚类的轮廓显示进行调整

5.对共被引网络聚类进行调整

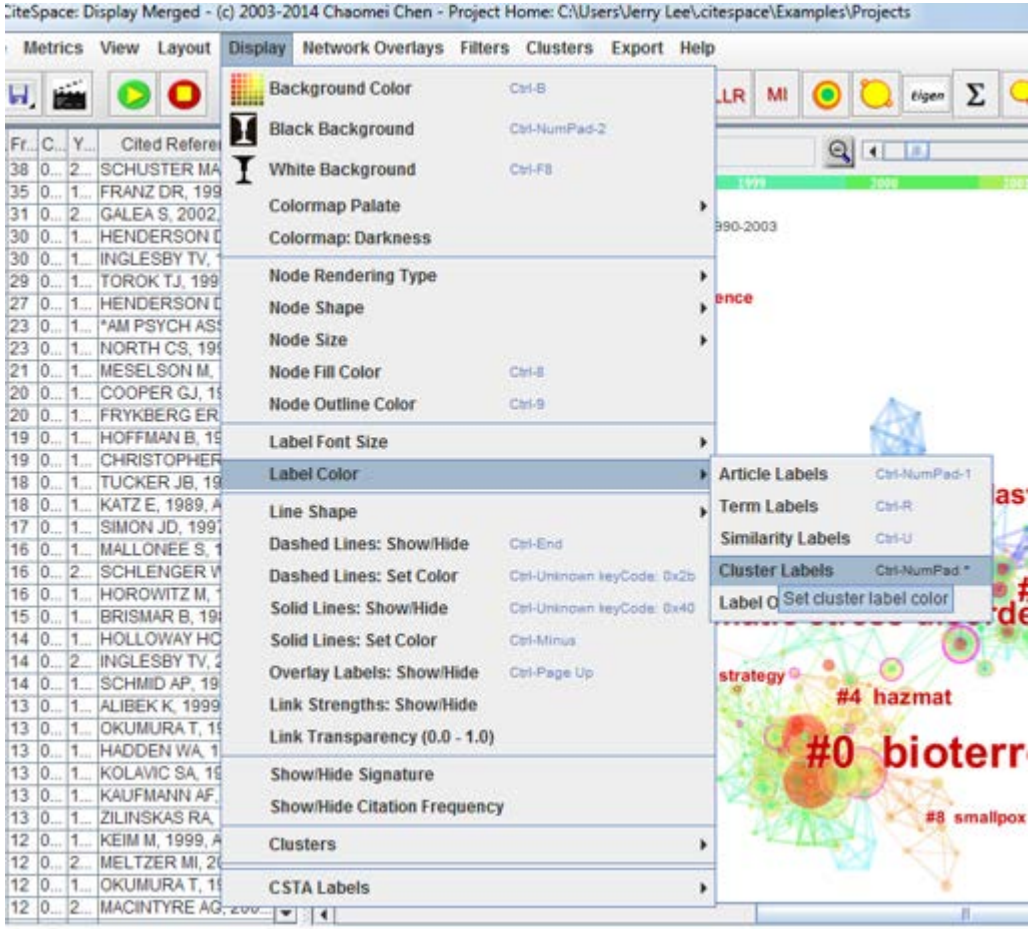
- 特别地，在合适的情况下可以对文献的突发性进行检测。

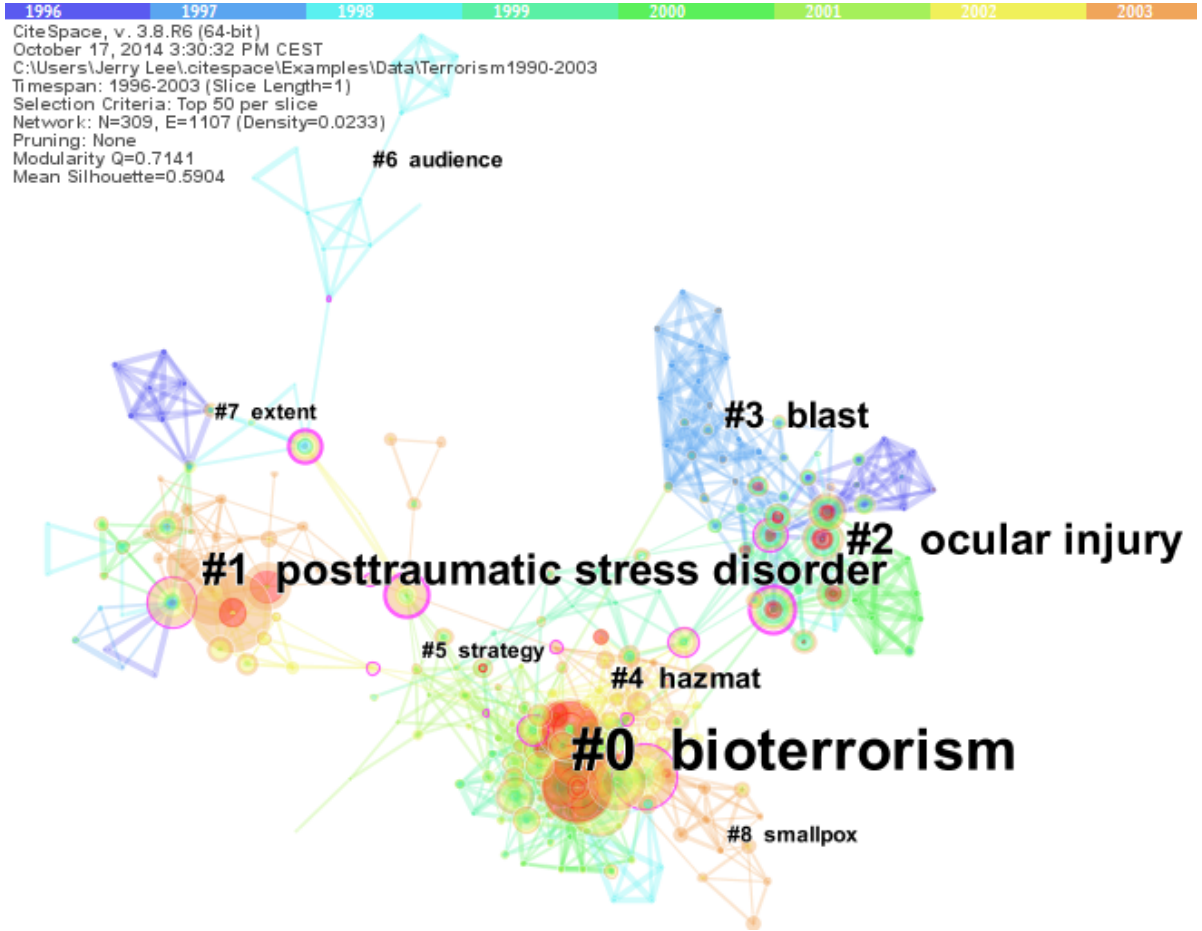




5.对共被引网络聚类进行调整

- 对聚类主题的颜色进行调整

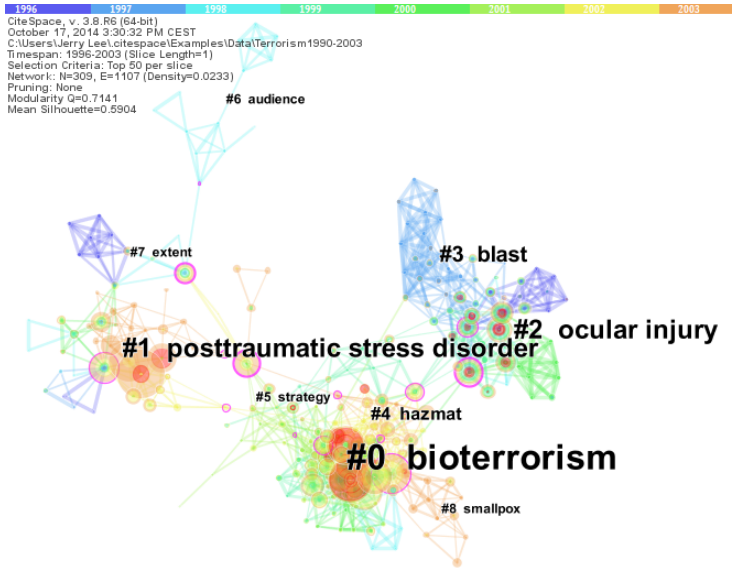




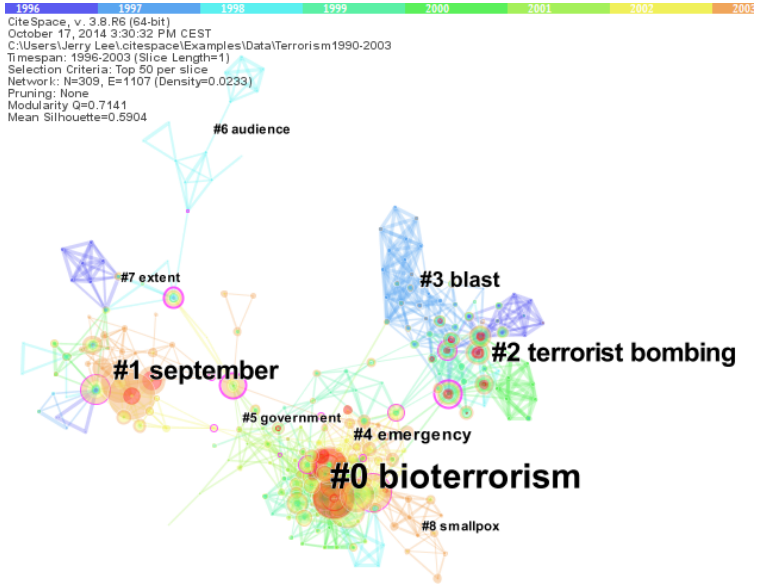
5.对共被引网络聚类进行调整

tf*idf | LLR | MI

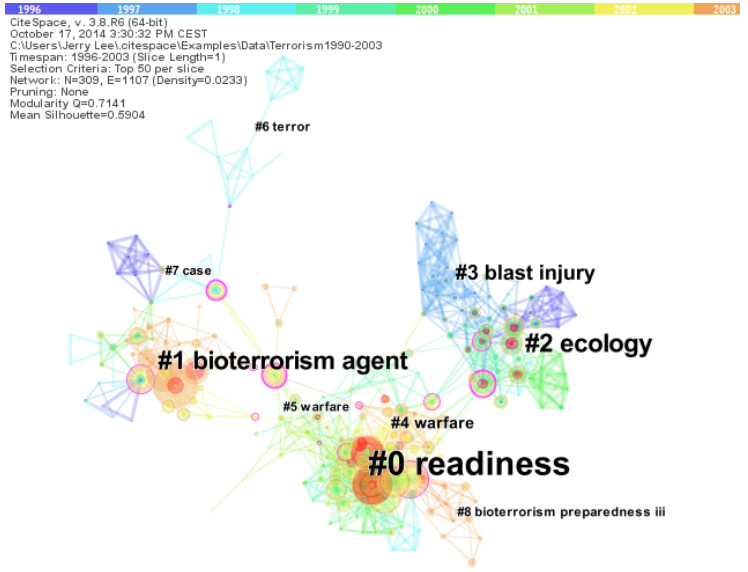
- 得到较为满意的图谱后，使用不同的方法对聚类进行命名。通常情况下陈教授推荐使用LLR算法得到的结果，下面给出三种算法得到结果的比较。



TF*IDF

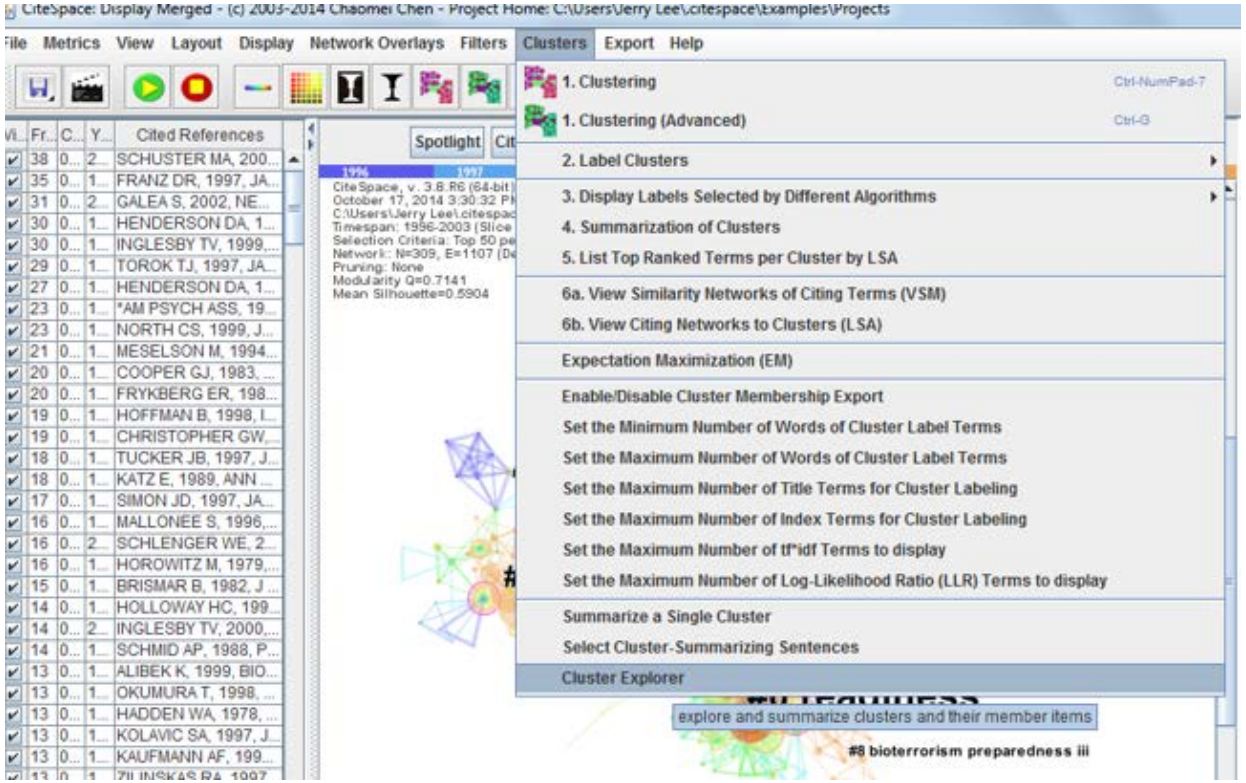


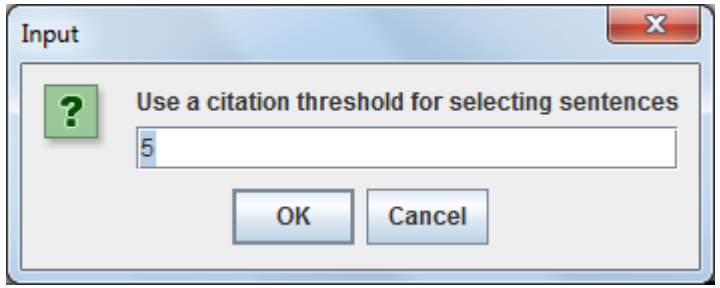
LLR

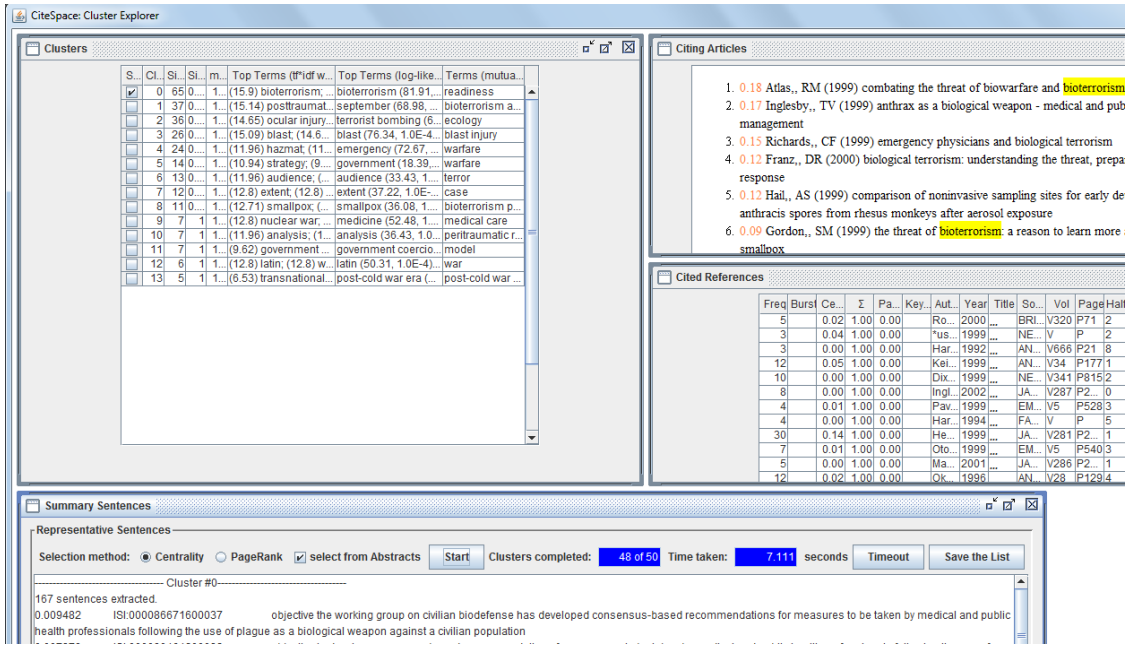


MI

6.对聚类详细信息的查询

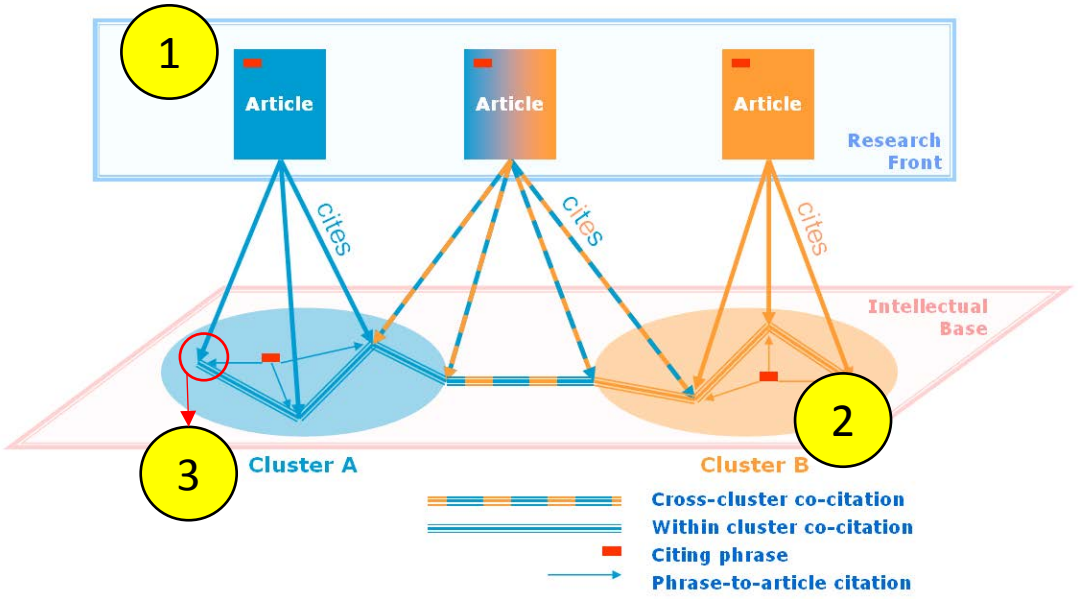


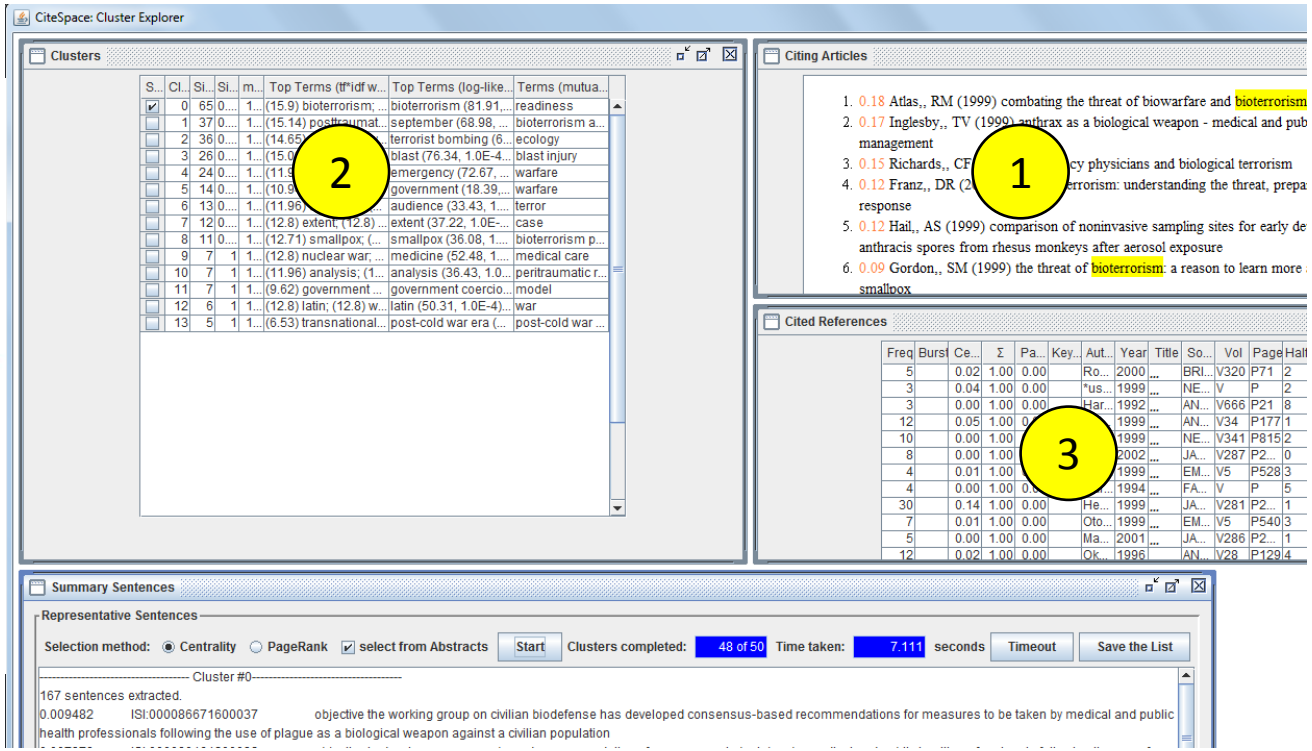




以下将对（右侧）三个重要的窗口进行解释

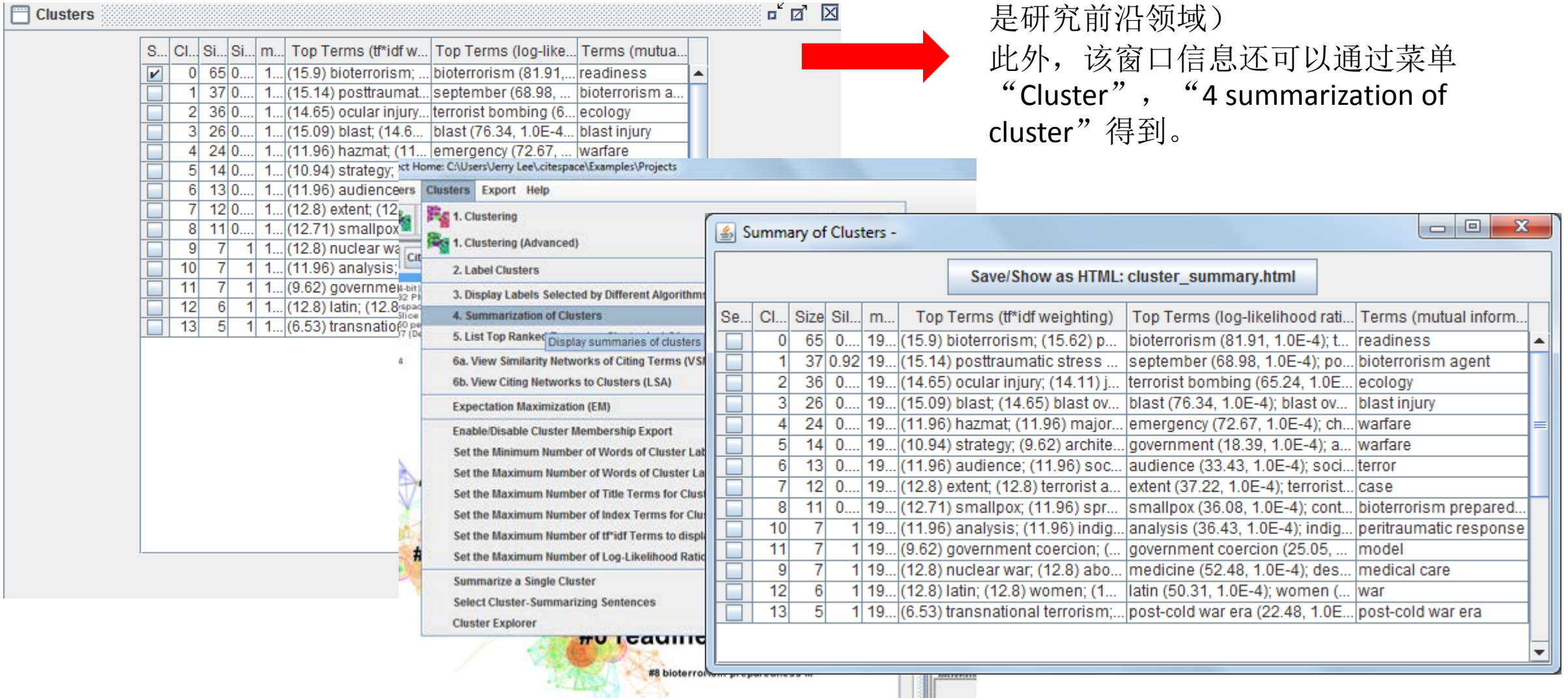
6.对聚类详细信息的查询





CiteSpace概念模型与软件提供的信息查询比较

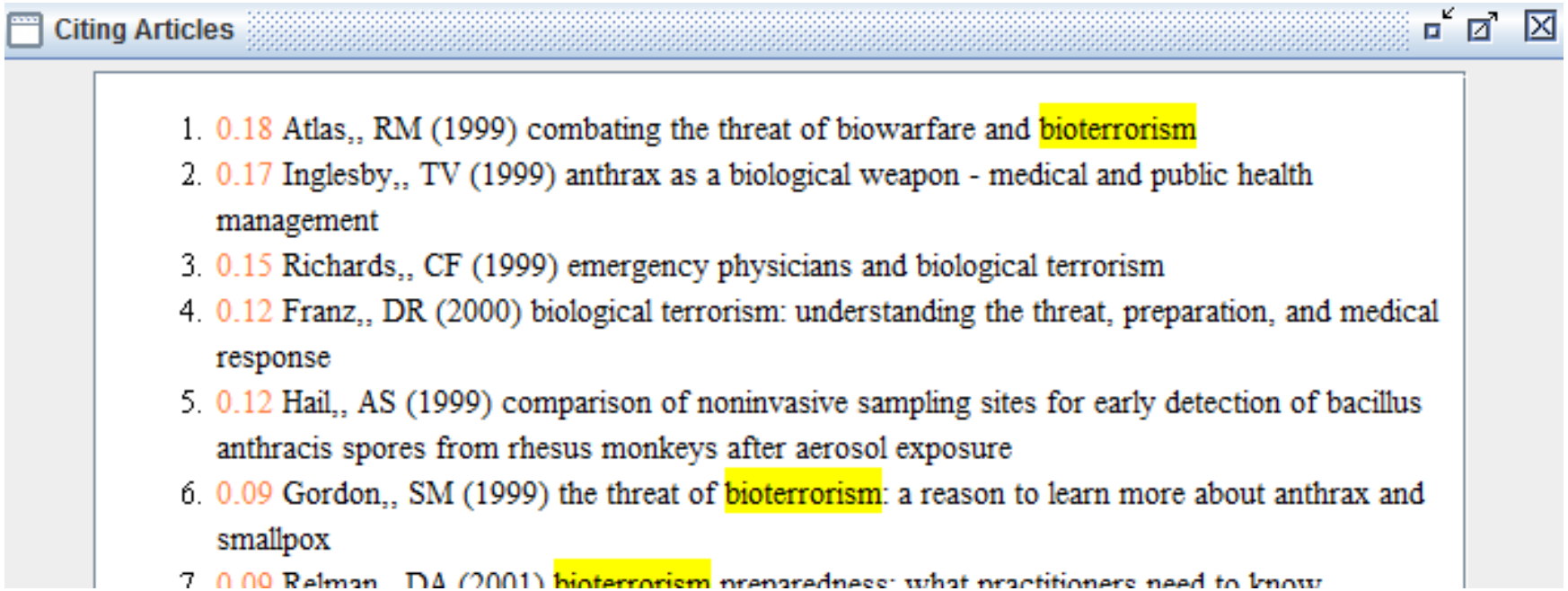
6.对聚类详细信息的解释



该窗口显示的是通过三种方法得到的聚类命名（笔者认为这些属于反映的是研究前沿领域）

此外，该窗口信息还可以通过菜单“Cluster”，“4 summarization of cluster”得到。

6.对聚类详细信息的解释



该窗口显示的施引文献（笔者认为这些文献代表了研究前沿）。标题中着重标识的词汇正是通过相关方法提取的聚类命名

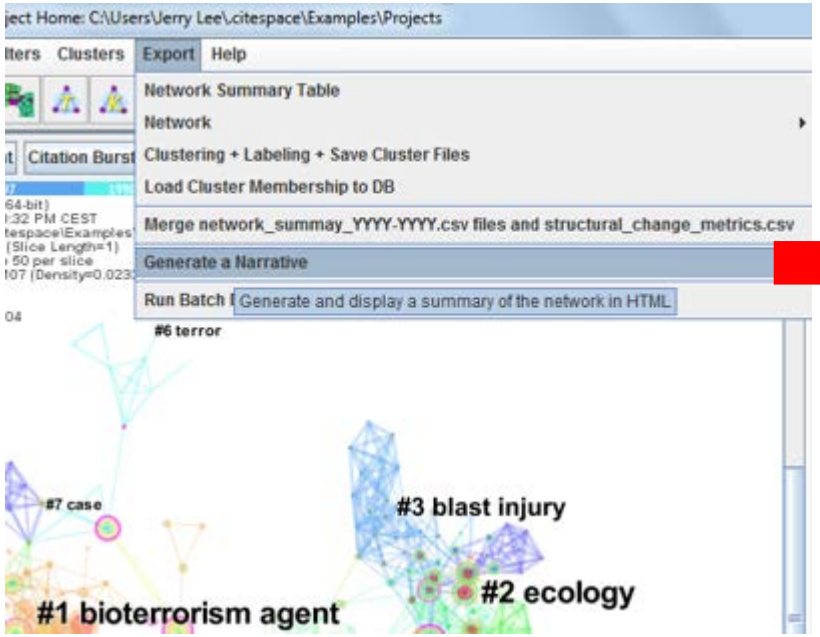
6.对聚类详细信息的解释

Cited References

Freq	Burst	Ce...	Σ	Pa...	Key...	Aut...	Year	Title	So...	Vol	Page	Half...	Clu...
5		0.02	1.00	0.00		Ro...	2000	...	BRI...	V320	P71	2	0
3		0.04	1.00	0.00		*us...	1999	...	NE...	V	P	2	0
3		0.00	1.00	0.00		Har...	1992	...	AN...	V666	P21	8	0
12		0.05	1.00	0.00		Kei...	1999	...	AN...	V34	P177	1	0
10		0.00	1.00	0.00		Dix...	1999	...	NE...	V341	P815	2	0
8		0.00	1.00	0.00		Ingl...	2002	...	JA...	V287	P2...	0	0
4		0.01	1.00	0.00		Pav...	1999	...	EM...	V5	P528	3	0
4		0.00	1.00	0.00		Har...	1994	...	FA...	V	P	5	0
30		0.14	1.00	0.00		He...	1999	...	JA...	V281	P2...	1	0
7		0.01	1.00	0.00		Oto...	1999	...	EM...	V5	P540	3	0
5		0.00	1.00	0.00		Ma...	2001	...	JA...	V286	P2...	1	0
12		0.02	1.00	0.00		Ok...	1996	...	AN...	V28	P129	4	0
6	2.78	0.00	1.00	0.00		Ingl...	2001	...	CLI...	V32	P436	1	0
3		0.00	1.00	0.00		*us...	1998	...	ME...	V	P	3	0
2		0.00	1.00	0.00		Ca...	1974	...	J I...	V12...	P	26	0

该窗口显示的是被引文献（笔者认为这些文献反映的是知识基础），
这些文献也是直接在图谱中显示的节点信息

7.其他功能：自动生成研究报告



AUTOMATICALLY GENERATED NARRATIVES

Time of creation: Fri Oct 17 16:31:42 CEST 2014

MAJOR CLUSTERS

The network is divided into **14** co-citation clusters. These clusters are labeled by index terms from their own citers. The largest **5** clusters are summarized.

Table 1. Summary of the largest 5 clusters.

ClusterID	Size	Silhouette	Label (TFIDF)	Label (LLR)	Label (MI)	mean(Citee Year)
0	65	0.651	(15.9) bioterrorism	bioterrorism (81.91, 1.0E-4)	readiness	1996
1	37	0.92	(15.14) posttraumatic stress disorder	september (68.98, 1.0E-4)	bioterrorism agent	1995
2	36	0.901	(14.65) ocular injury	terrorist bombing (65.24, 1.0E-4)	ecology	1987
3	26	0.818	(15.09) blast	blast (76.34, 1.0E-4)	blast injury	1982
4	24	0.815	(11.96) hazmat	emergency (72.67, 1.0E-4)	warfare	1995

The largest cluster (#0) has 65 members and a silhouette value of 0.651. It is labeled as ***bioterrorism*** by both LLR and TFIDF, and as *readiness* by MI. The most active citer (1999) combating the threat of biowarfare and bioterrorism.

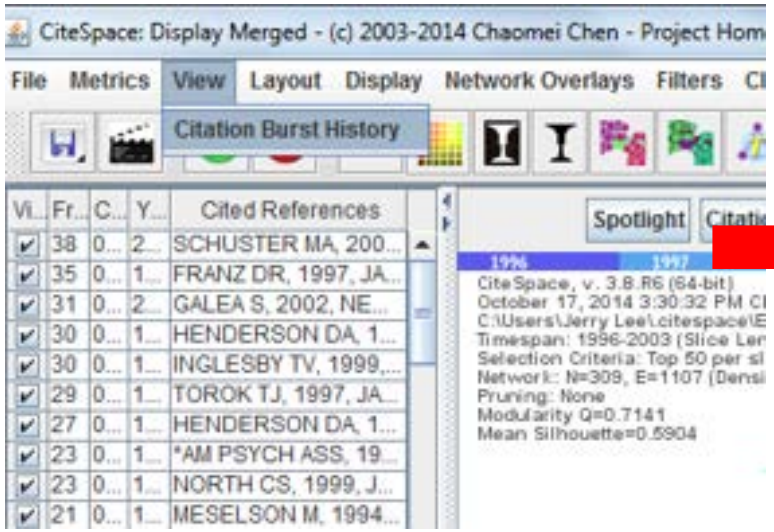
The second largest cluster (#1) has 37 members and a silhouette value of 0.92. It is labeled as *september* by LLR, *posttraumatic stress disorder* by TFIDF, and *bioterrorism* to the cluster is 0.16 Tucker,, P (2000) predictors of post-traumatic stress symptoms in oklahoma city: exposure, social support, peri-traumatic responses.

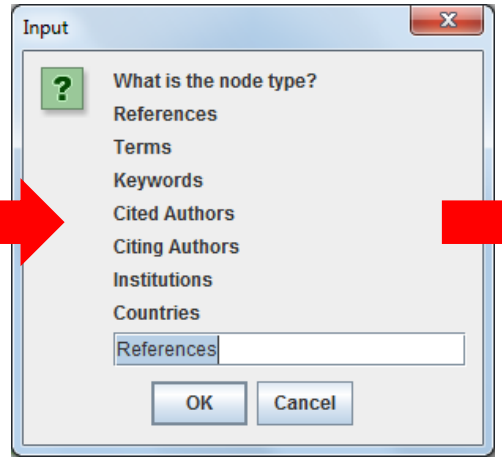
The third largest cluster (#2) has 36 members and a silhouette value of 0.901. It is labeled as *terrorist bombing* by LLR, *ocular injury* by TFIDF, and *ecology* by MI. The mc Stein,, M (1999) medical consequences of terrorism - the conventional weapon threat.

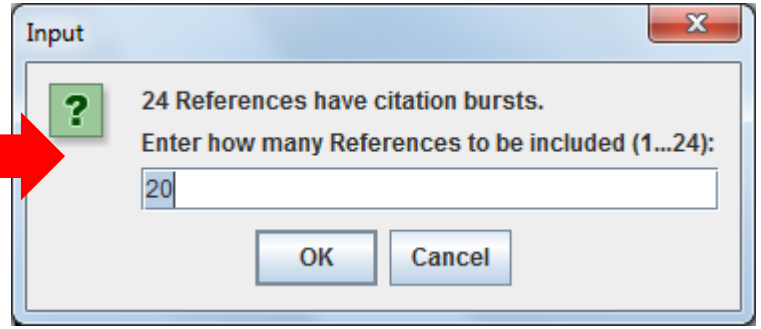
The 4th largest cluster (#3) has 26 members and a silhouette value of 0.818. It is labeled as ***blast*** by both LLR and TFIDF, and as *blast injury* by MI. The most active citer to (1997) toxicology of blast over-pressure.

The 5th largest cluster (#4) has 24 members and a silhouette value of 0.815. It is labeled as *emergency* by LLR, *hazmat* by TFIDF, and *warfare* by MI. The most active citer (1999) chemical warfare agents: emergency medical and emergency public health issues.

7.其他功能：突发性文献信息



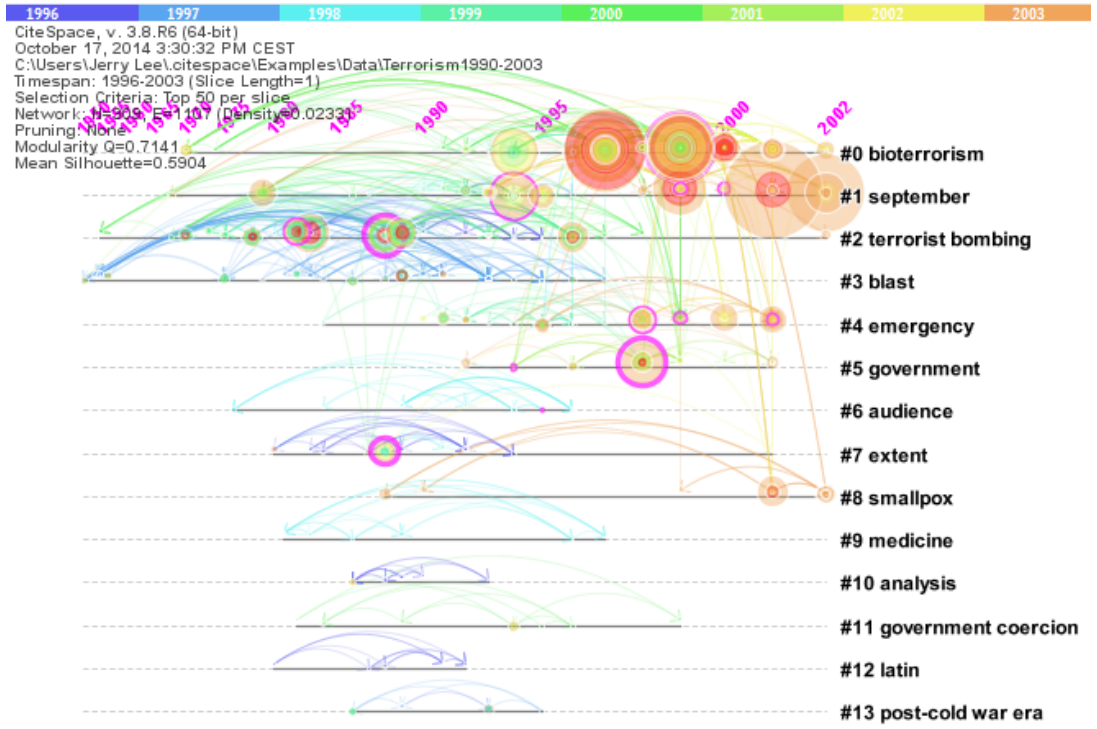


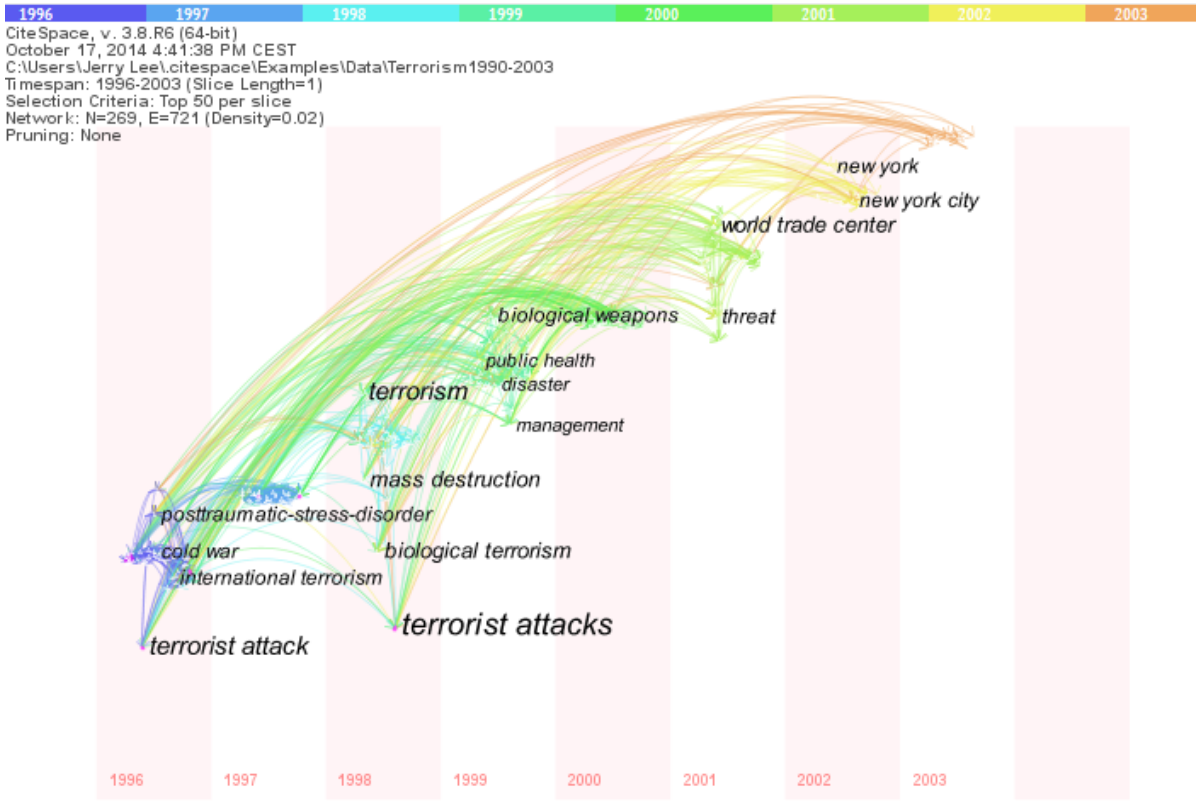


Top 20 References with Strongest Citation Bursts

References	Year	Strength	Begin	End	1996 - 2003
SCHUSTER MA, 2001, NEW ENGL J MED, V345, P1507	2001	**8.9432**	2002	2003	
COOPER GJ, 1983, J TRAUMA, V23, P955	1983	**4.6049**	1996	1999	
NORTH CS, 1999, JAMA-J AM MED ASSOC, V282, P755	1999	**4.4484**	2002	2003	
INGLESBY TV, 1999, JAMA-J AM MED ASSOC, V281, P1735	1999	**4.0214**	2001	2003	
FRYKBERG ER, 1988, ANN SURG, V208, P569	1988	**3.5106**	1997	1999	
HENDERSON DA, 1999, SCIENCE, V283, P1279	1999	**3.0777**	2001	2001	

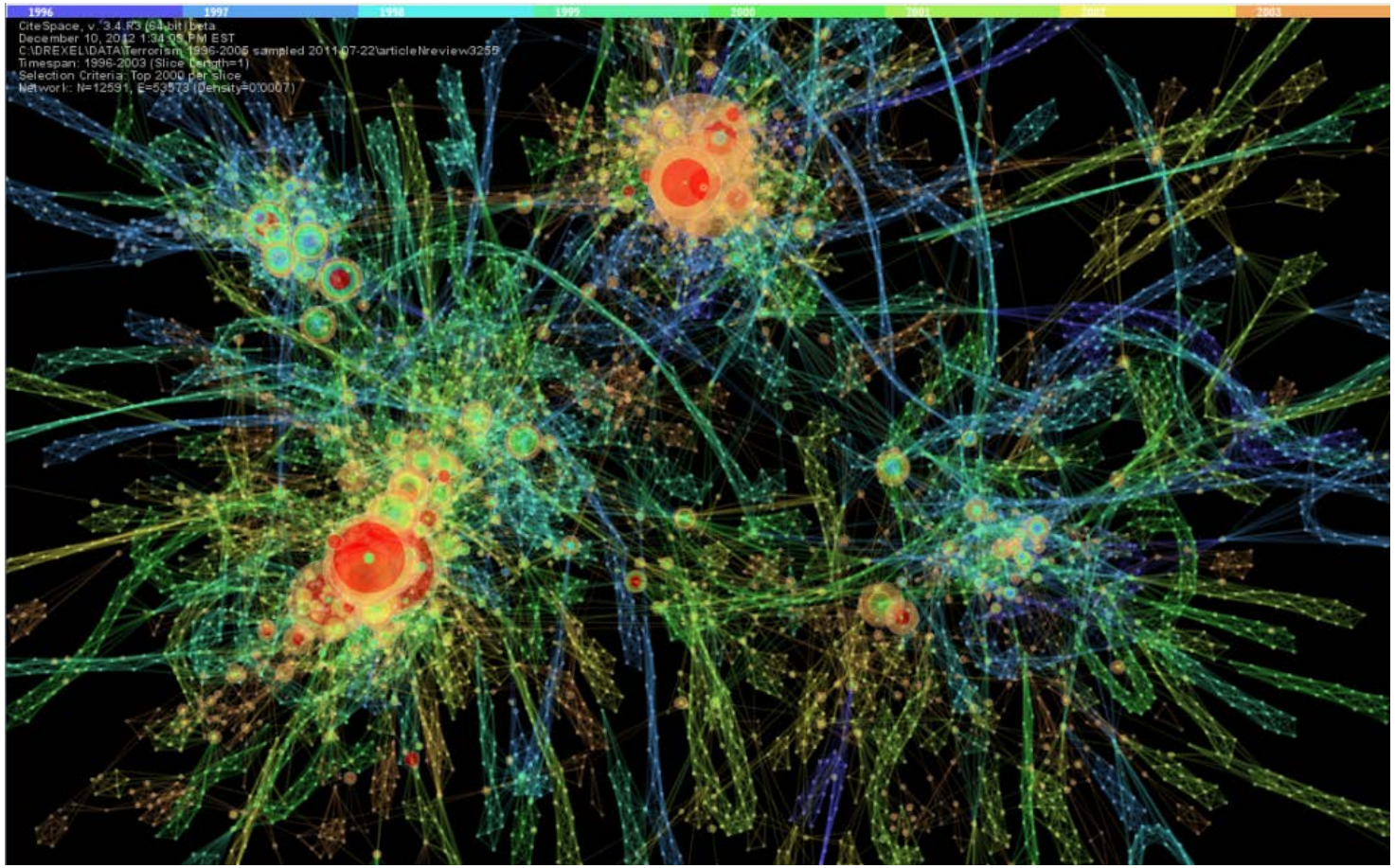
7.其他功能：不同的呈现方式





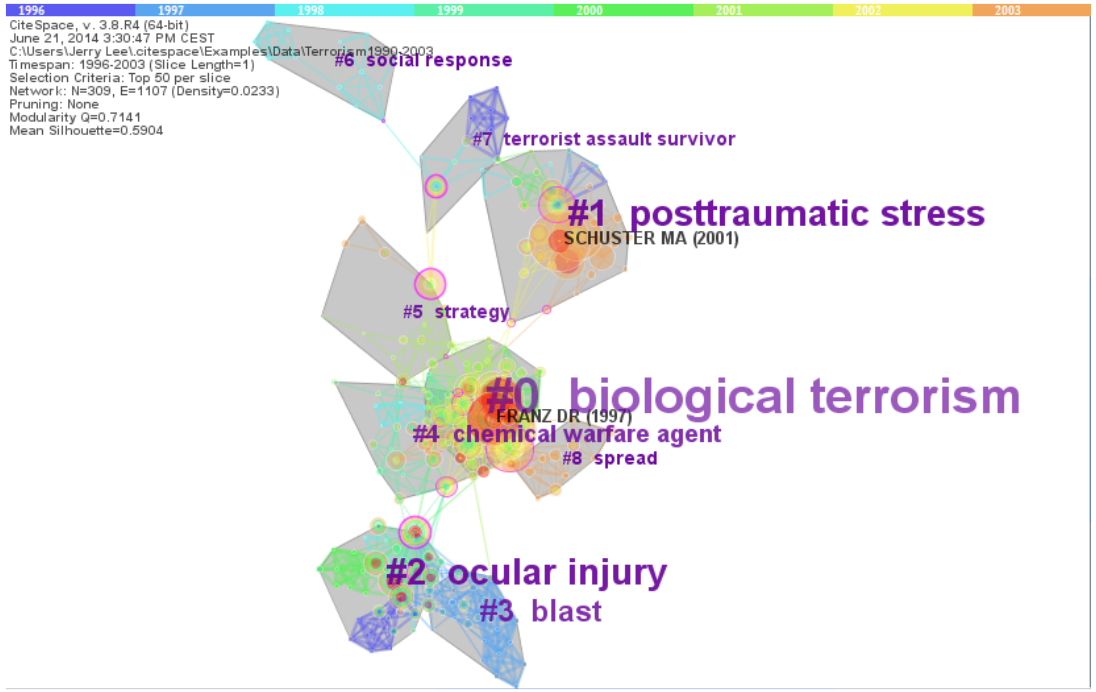
Timeline 呈现方式（共被引聚类仅仅可以选择两种呈现方式，主题词分析可以选择Time zone来呈现结果，如右图）

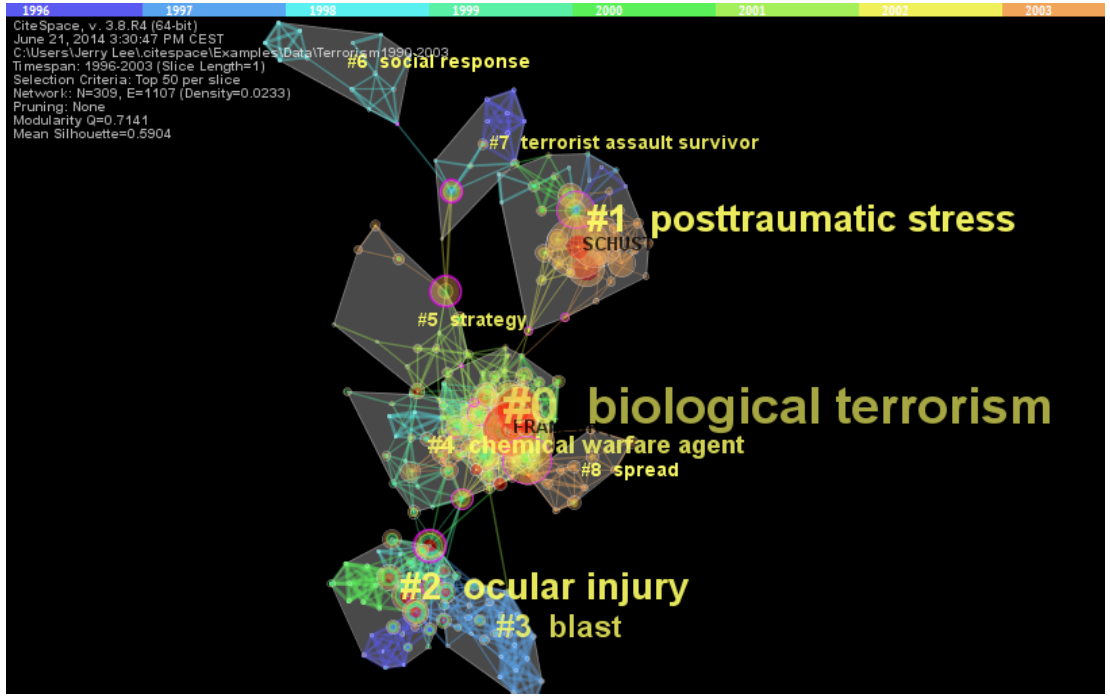
陈超美教授关于恐怖主义研究的补充图



来源：陈超美教授博客http://blog.sciencenet.cn/blog-496649-641014.html

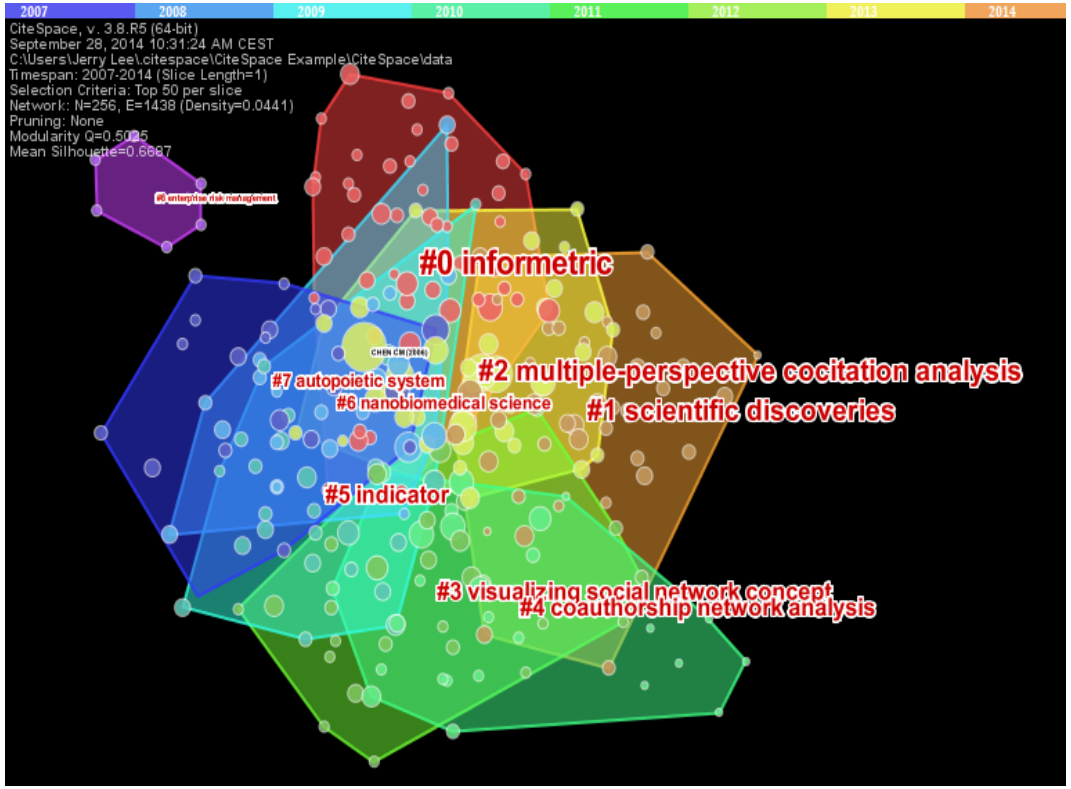
笔者完成的一些实例

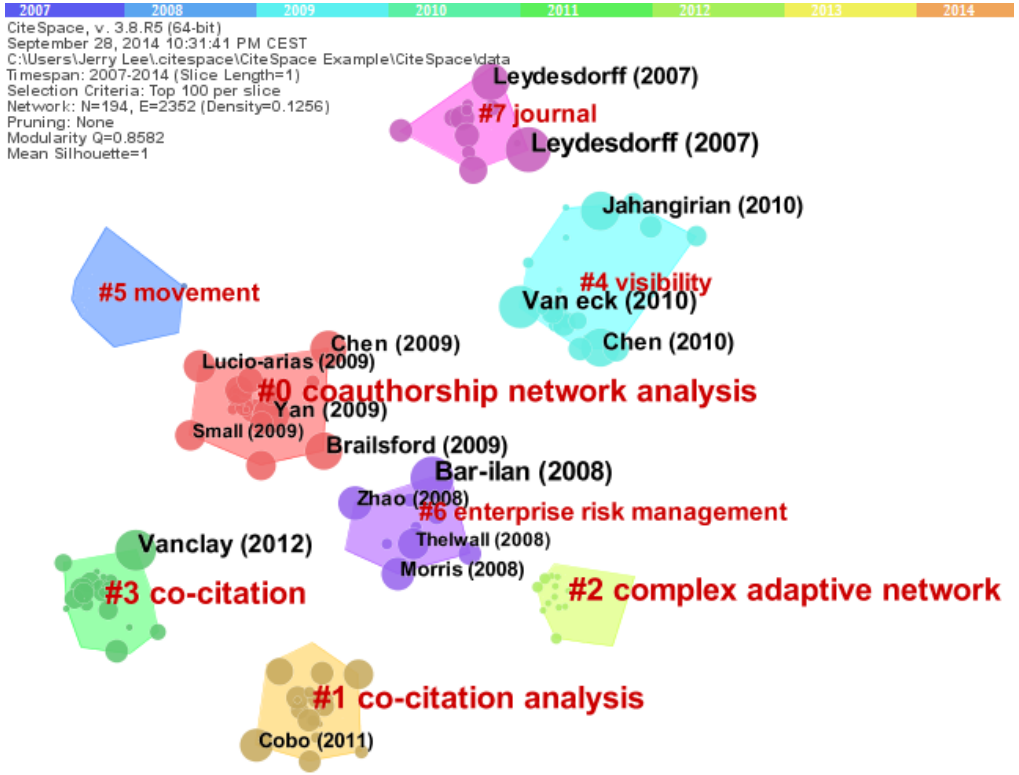




注释：根据CiteSpace提供的例子数据绘制，数据再首次运行时自带（文献共被引聚类）

笔者完成的一些实例





注释：本数据是笔者根据引用CiteSpace经典论文的施引文献绘制，左侧为文献共被引聚类结果，右侧为文献耦合聚类结果。更多结果参见：

CiteSpace经典文献的统计分析

http://blog.sciencenet.cn/home.php?mod=space&uid=554179&do=blog&id=831518

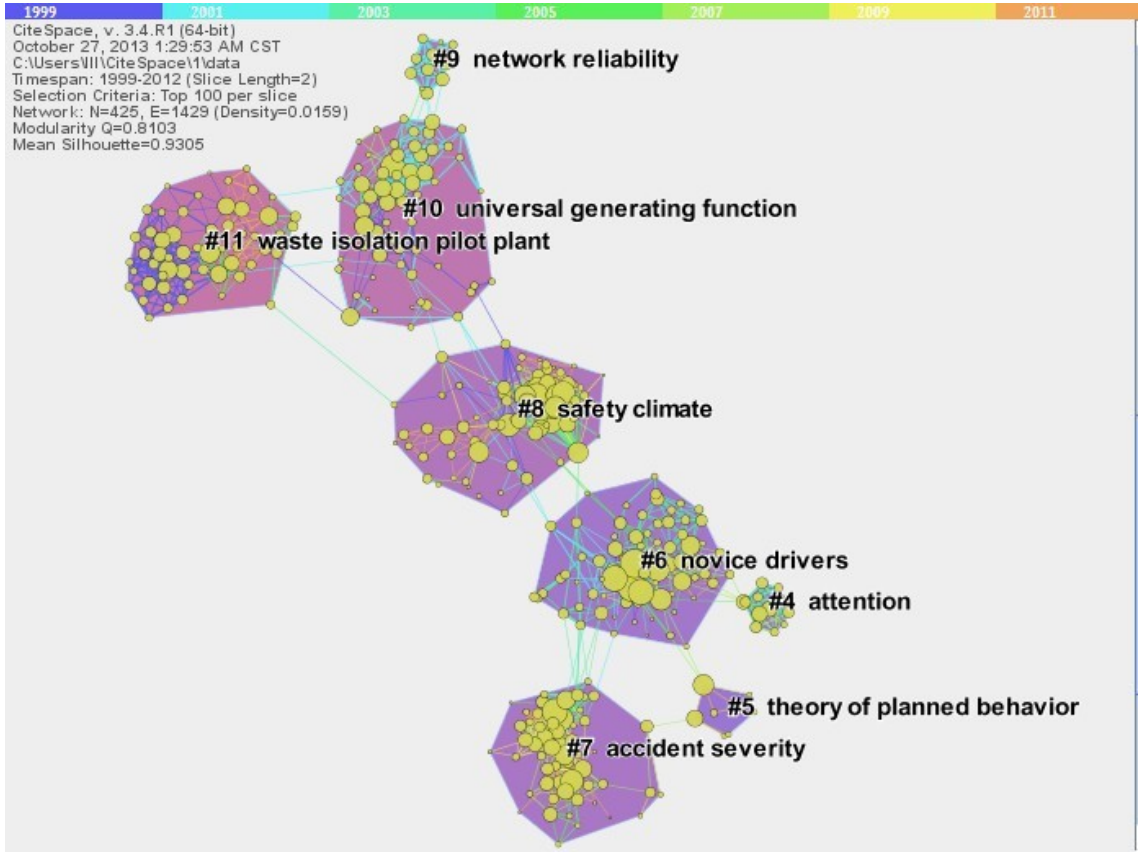
CiteSpace经典文献的统计分析（补充一：共被引）

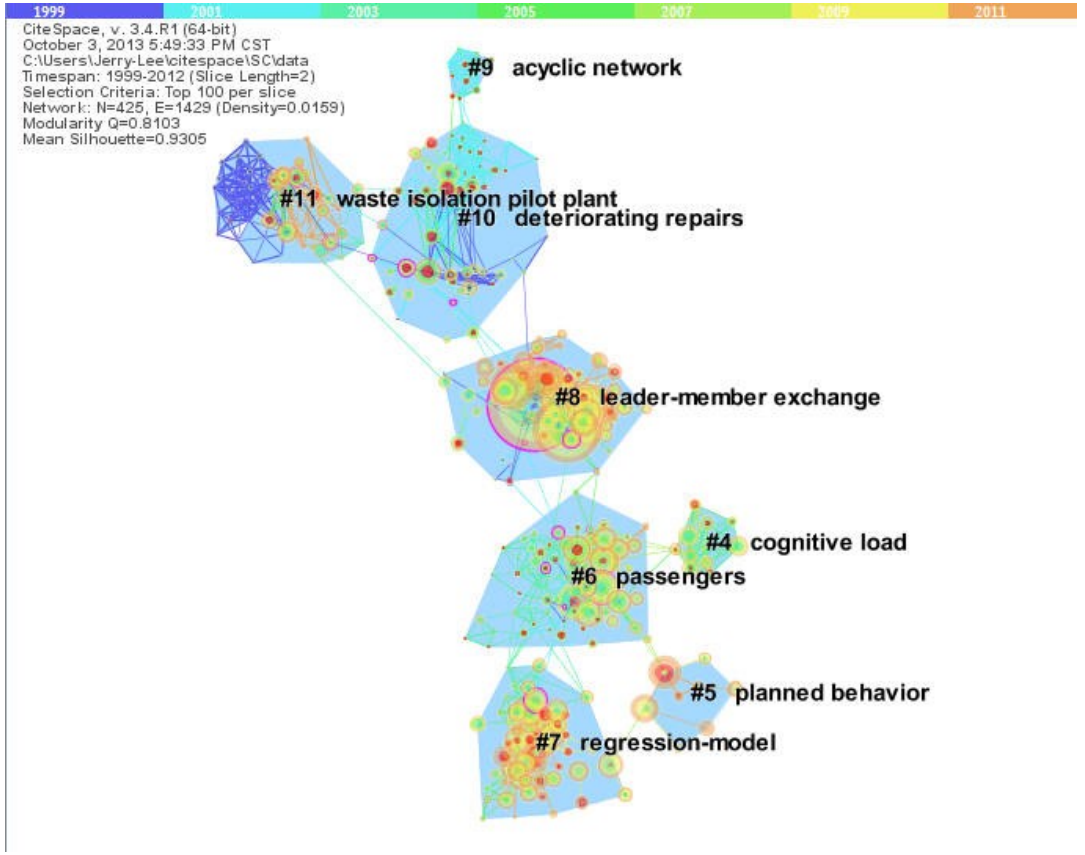
http://blog.sciencenet.cn/home.php?mod=space&uid=554179&do=blog&id=831634

CiteSpace经典文献的统计分析（补充二：文献耦合）

http://blog.sciencenet.cn/home.php?mod=space&uid=554179&do=blog&id=831638

笔者完成的一些实例

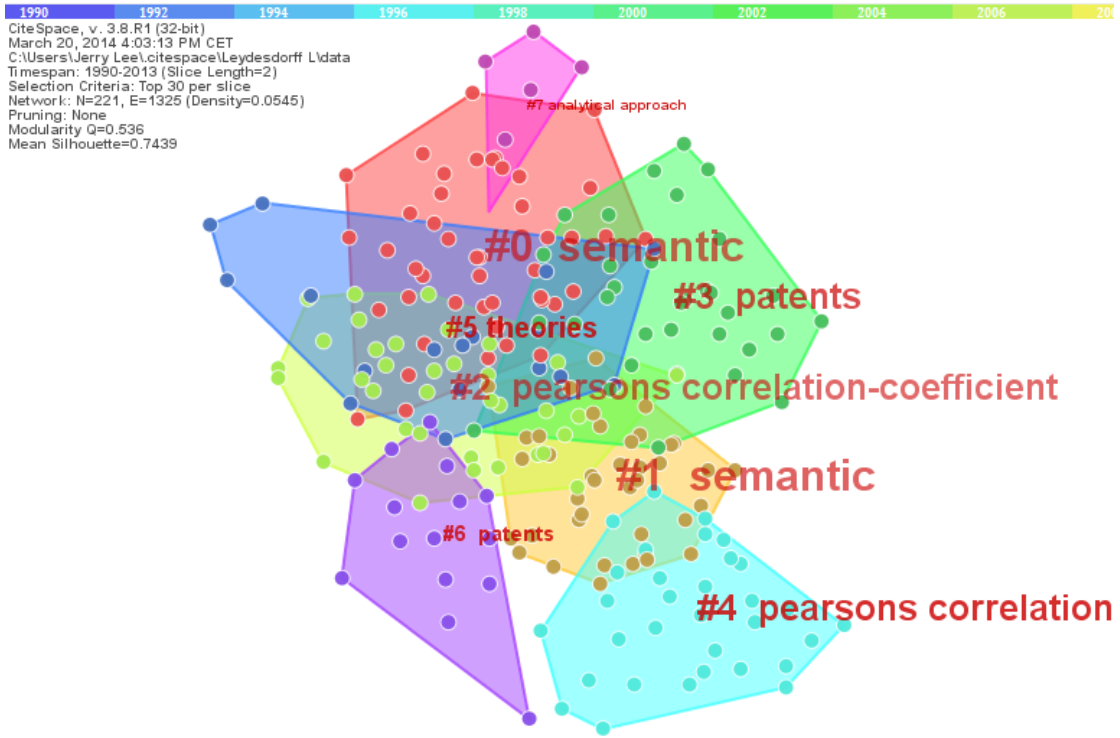


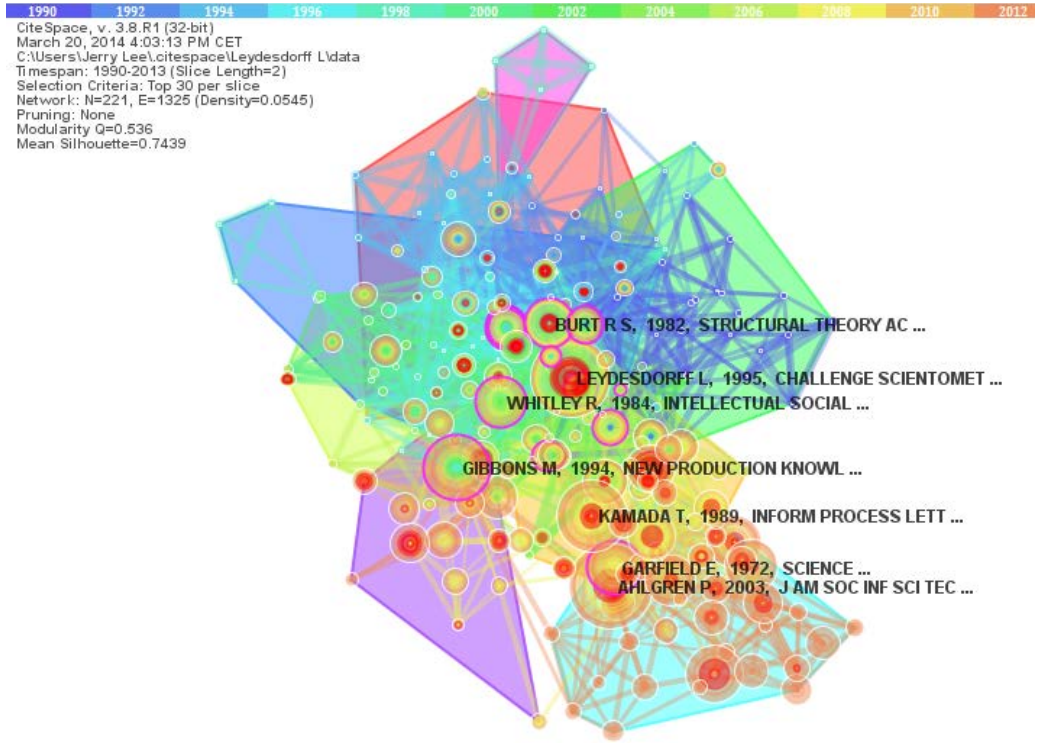


注释：本数据是笔者初次进行安全科学文献图谱分析时绘制（文献共被引聚类）

http://blog.sciencenet.cn/home.php?mod=space&uid=554179&do=blog&quickforward=1&id=729837

笔者完成的一些实例



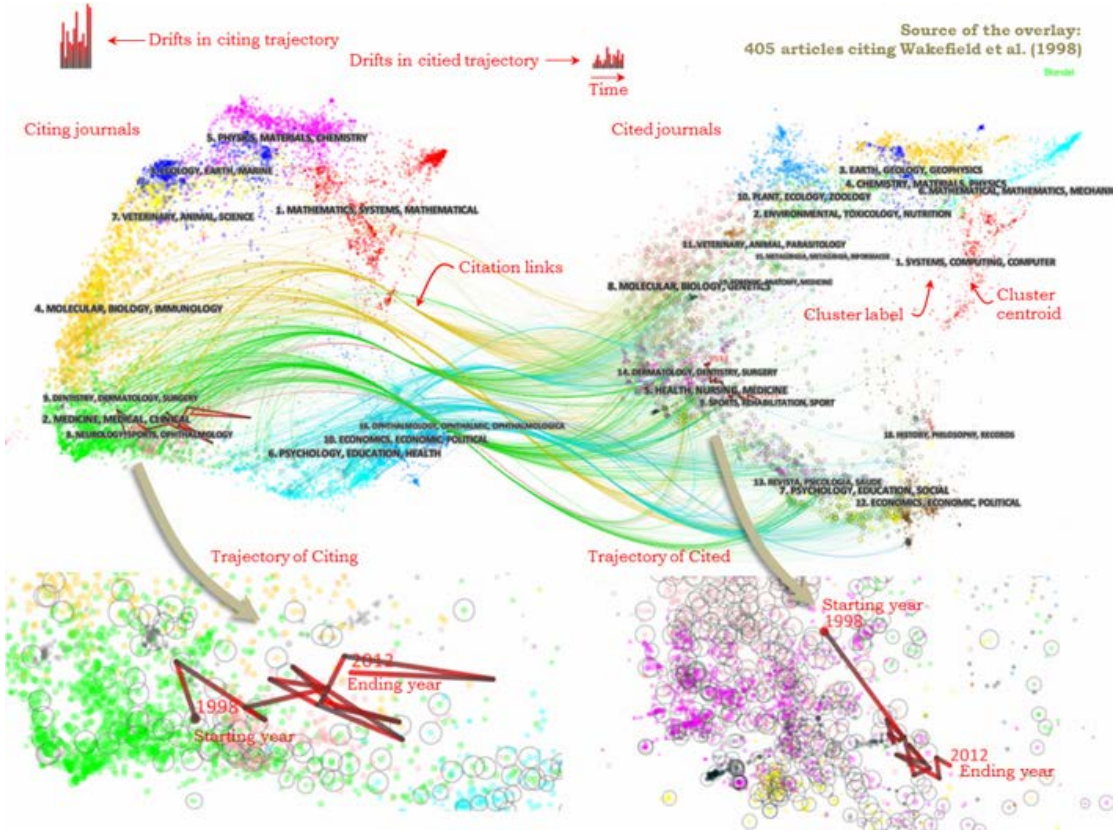


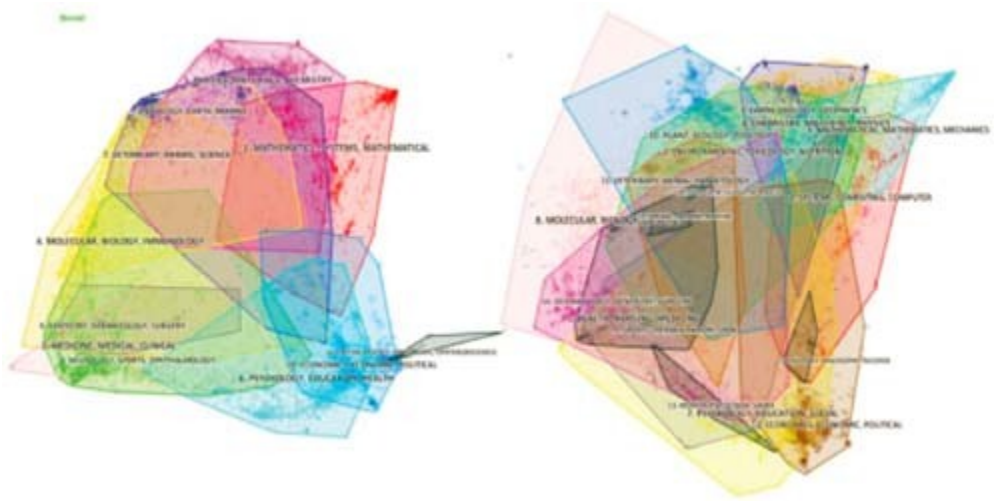
注释：本数据Loet Leydesdorff被Web of Science 收录的论文绘制（文献共被引）

更多结果参见Bibliometrics analysis of Loet Leydesdorff`s publication

http://blog.sciencenet.cn/home.php?mod=space&uid=554179&do=blog&id=777700

CiteSpace未来可能嵌入和开放的功能：双图叠加





全域科学图谱和叠加图 http://blog.sciencenet.cn/blog-528739-833807.html

常用CiteSpace功能备份

- （1）CiteSpace重要问题整理
- http://blog.sciencenet.cn/blog-554179-667300.html
- （2）CiteSpace 手册（英文）
- http://cluster.ischool.drexel.edu/~cchen/citespace/CiteSpaceManual.pdf
- （3）如何清除CiteSpace不同版本的Java WebStart app
- http://blog.sciencenet.cn/blog-496649-676029.html
- （4）CiteSpace中Pathfinder的作用
- http://blog.sciencenet.cn/blog-496649-712057.html

获取最新关于CiteSpace应用的论文

Chaomei Chen

Follow

Professor of Informatics, Col
University
Information visualization, vis
Verified email at cis.drexel.e

Email

Already following new articles

Already following new citations

CREATE ALERT Cancel List alerts

Title 1–100

引文空间分析原理与应用 CiteSpace实用指南(Principles and Applications of Analyzing a Citation Space) 2014
Y Chen, C Chen, Z Hu, X Wang
Science Press

3. Subjectivity of Evidence 2014
C Chen
The Fitness of Information: Quantitative Assessments of Critical Evidence ...

Supplemental Images 2014
C Chen
The Fitness of Information: Quantitative Assessments of Critical Evidence, a-p

5. Fitness Landscapes 2014
C Chen

Google Scholar

Citation indices	All	Since 2009
Citations	8451	4665
h-index	41	31
i10-index	84	61

Co-authors View all...

Katy Borner
Mary Czerwinski
Kevin W. Boyack
Robert Macredie
Xia Lin
Don Pellegrino

使用Google Scholar检索Chaomei Chen，然后点击“Follow”选中“Following new articles”和“ Following new Citations ”。这样就能得到陈超美教授最新发表的论文和其他研究人员引用其论文的情况。

更多例子参见

科学网博客http://blog.sciencenet.cn/u/jerrycueb

或者作者Flickr

https://www.flickr.com/photos/118823347@N08/