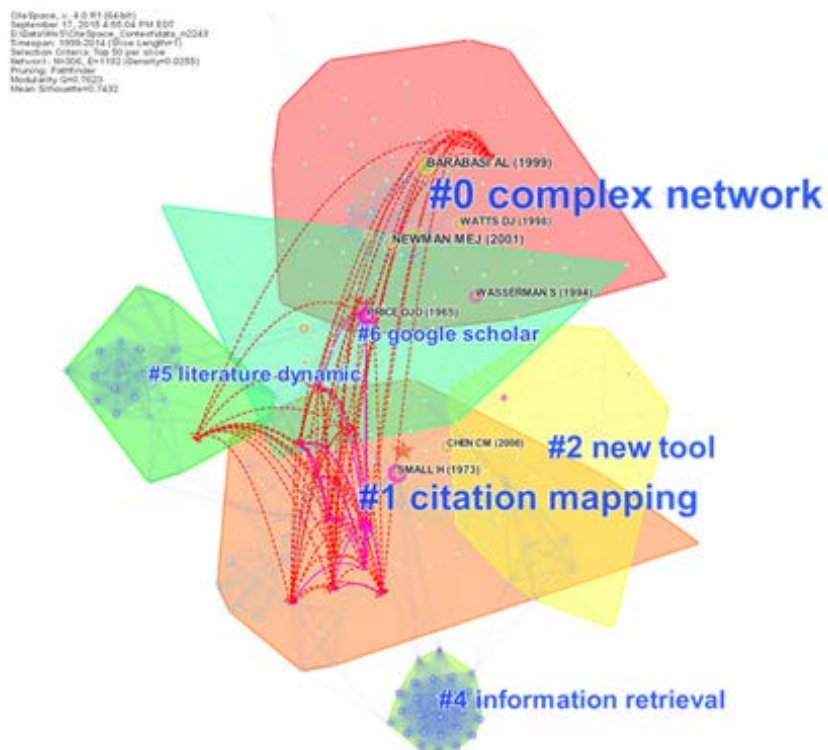


CiteSpace的结构变异分析功能

——如何使用CiteSpace的一组示范及常见问题解答之（十五）



李杰 陈超美

2015年9月18日

小提示

要求版本至少为4.0.R1.64-bit.
public.9.17.2015。下载地址为
<http://cluster.ischool.drexel.edu/~cchen/citespace/current/4.0.R1.64-bit.public.9.17.2015.zip>

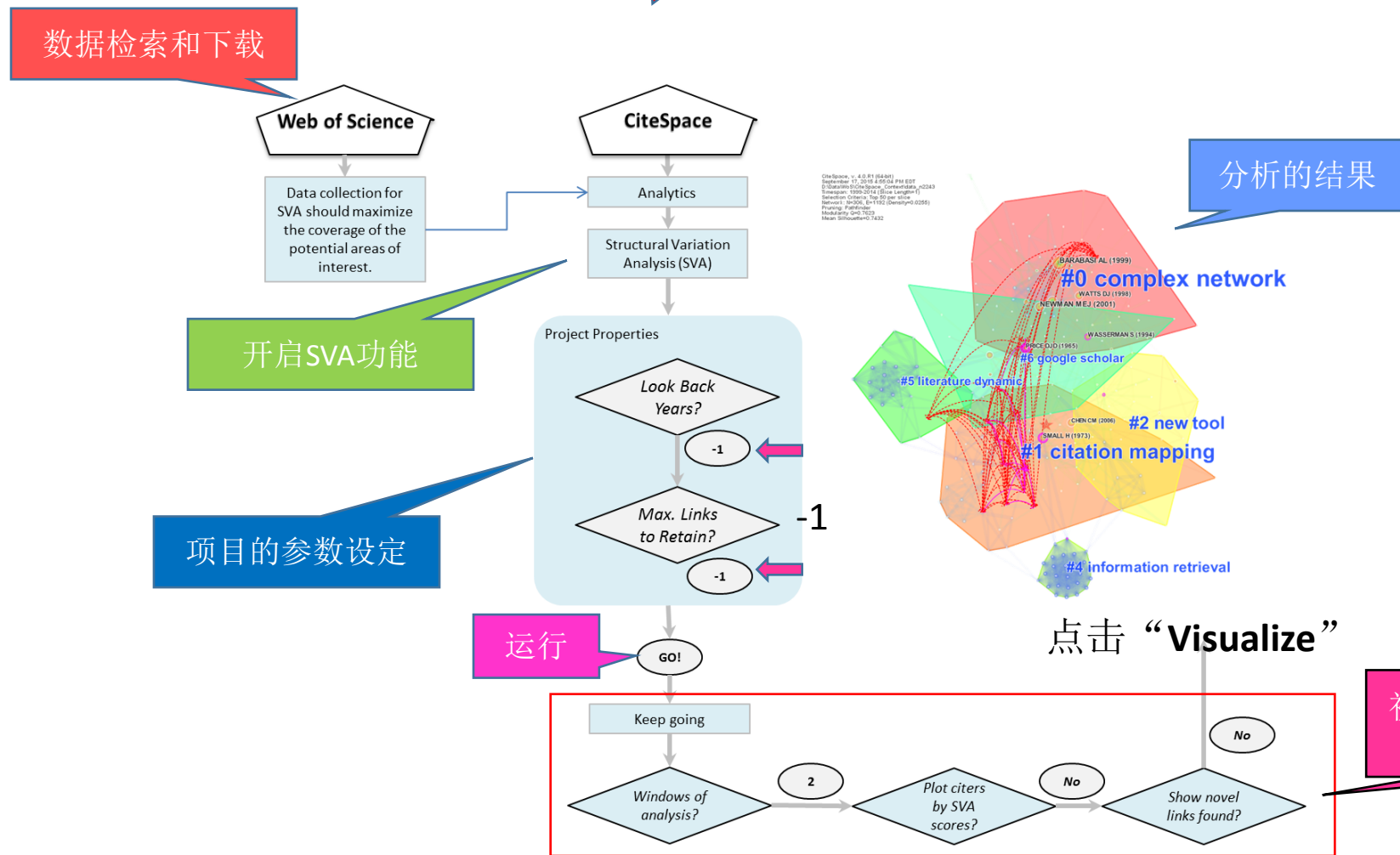
写在前面

- 关于结构变异理论（Theory of Structural Variation）的研究最初发表在2012年的一篇JASIST论文（见下面参考文献），其中也包括了几个案例研究，题为“[结构性变化预测被引次数的效果](#)”。其核心问题是如何能够量化新发表的论文的新颖性。
- 结构变异分析(Structural Variation Analysis - SVA)一方面基于科学创造方面的研究，尤其是新颖的重组在创造性思维中的作用和影响。这是基于以下观察：科学发现或创新在很大的程度上都具有一个共性，就是新思维能够容纳原本看似风马牛不相及的观念。换句话说，类似于在不同岛屿之间架起的一座新桥梁。第2个验证 条件就是这座新桥上是否确实吸引很快变得车水马龙。

参考文献

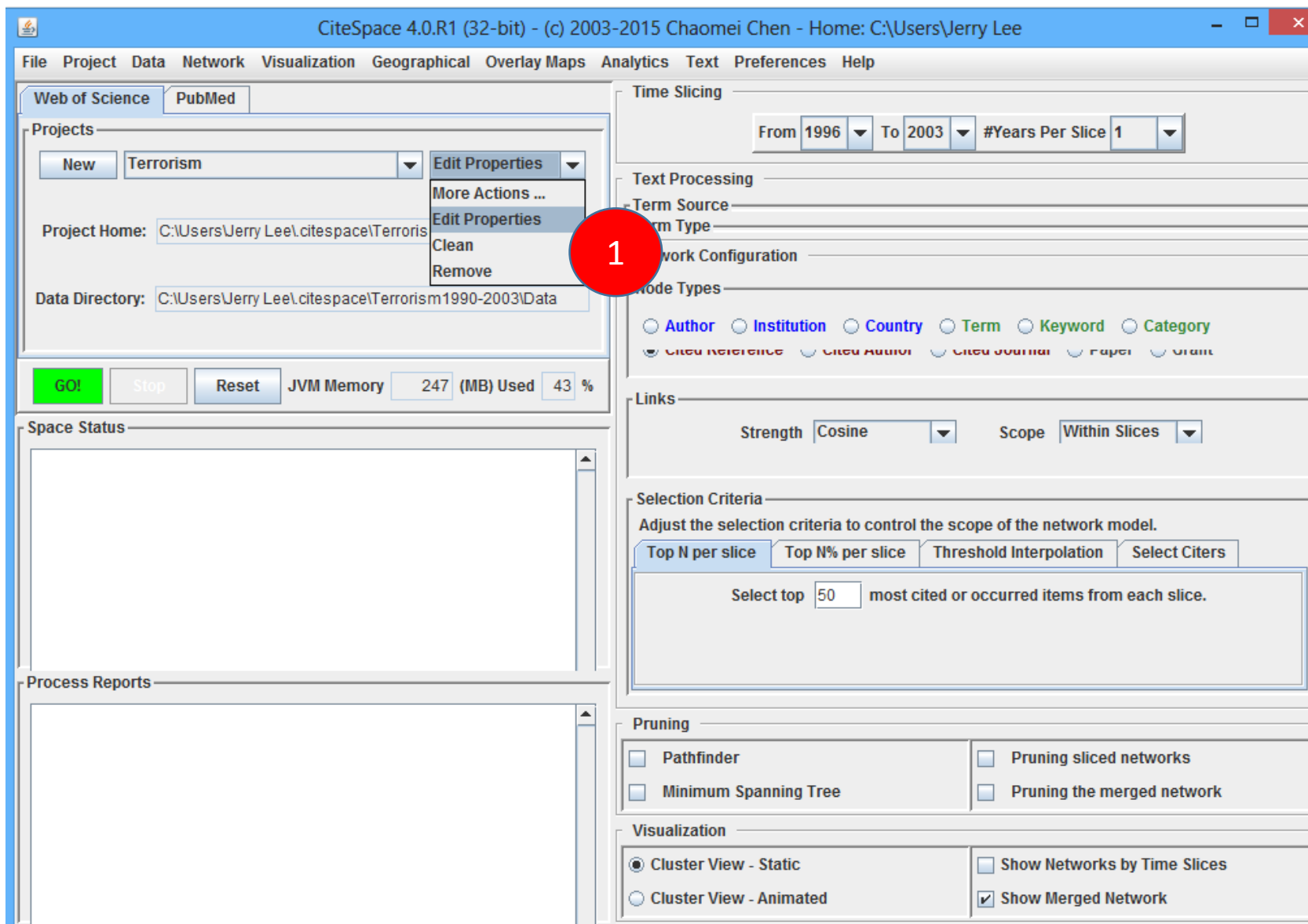
- Chen, C. (2012) [Predictive effects of structural variation on citation counts](#). *Journal of the American Society for Information Science and Technology*, **63**(3), 431-449.
- Chen, C. (2014) [The Fitness of Information: Quantitative Assessments of Critical Evidence](#). Wiley.

基本步骤

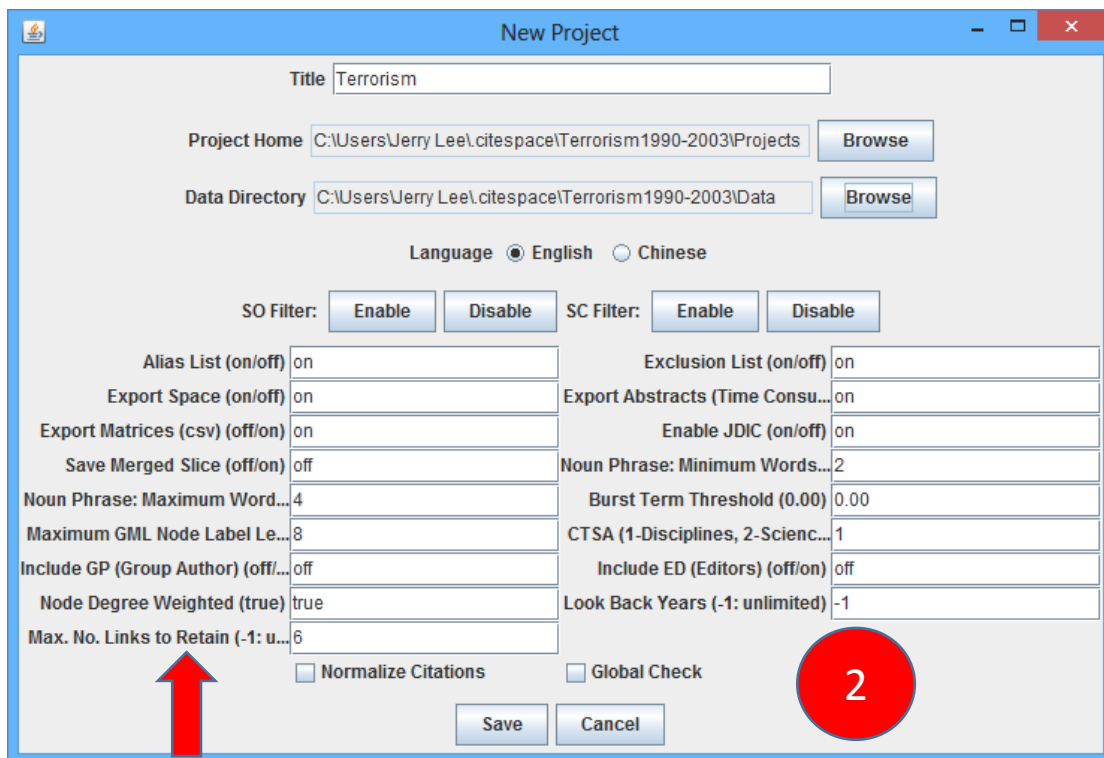


下面仅仅给出具体操作步骤，对结果的进一步分析解读请参照本PPT后面提供的原始文献。

建立好项目，对项目的参数进行进行调整。

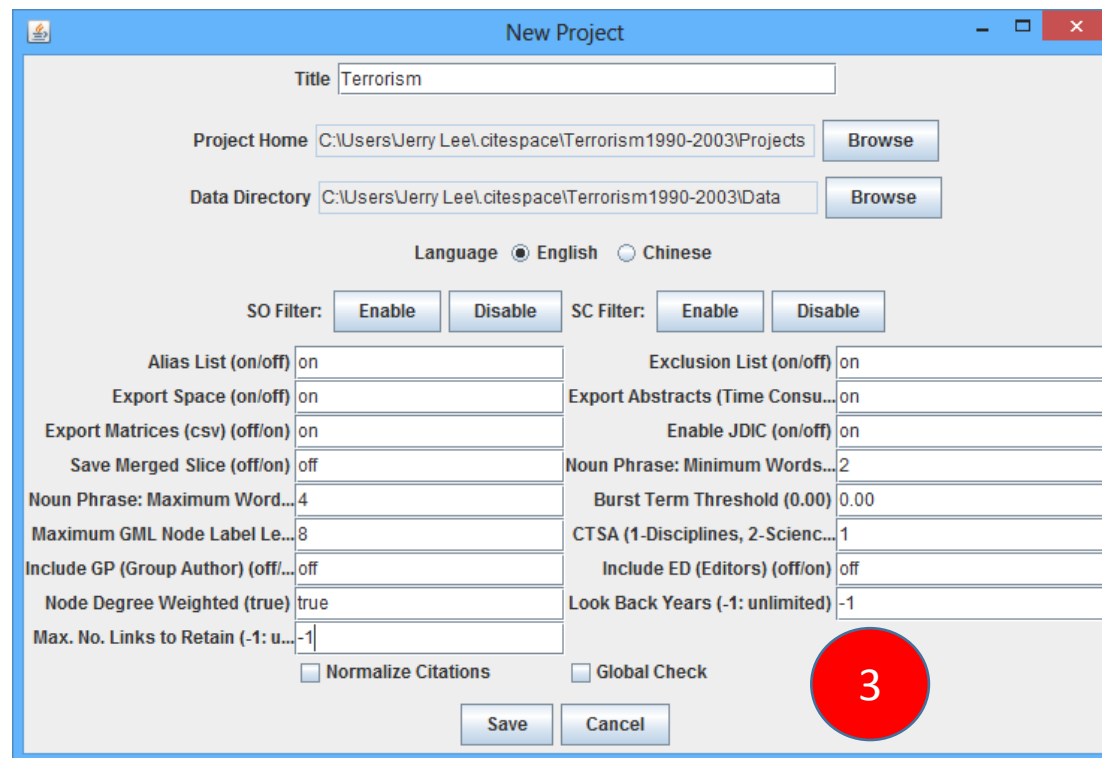


建立好项目，对项目的参数进行进行调整。



The 'New Project' dialog box shows the following settings:

- Title: Terrorism
- Project Home: C:\Users\Jerry Lee\citespace\Terrorism1990-2003\Projects
- Data Directory: C:\Users\Jerry Lee\citespace\Terrorism1990-2003\Data
- Language: ☒ English ☐ Chinese
- SO Filter: SC Filter:
- Alias List (on/off): on
- Exclusion List (on/off): on
- Export Space (on/off): on
- Export Abstracts (Time Consumption): on
- Export Matrices (csv) (off/on): on
- Enable JDIC (on/off): on
- Save Merged Slice (off/on): off
- Noun Phrase: Minimum Words: 2
- Noun Phrase: Maximum Word Length: 4
- Burst Term Threshold (0.00): 0.00
- Maximum GML Node Label Length: 8
- CTSA (1-Disciplines, 2-Scientific): 1
- Include GP (Group Author) (off/on): off
- Include ED (Editors) (off/on): off
- Node Degree Weighted (true): true
- Look Back Years (-1: unlimited): -1
- Max. No. Links to Retain (-1: unlimited): 6
- ☐ Normalize Citations ☐ Global Check
-



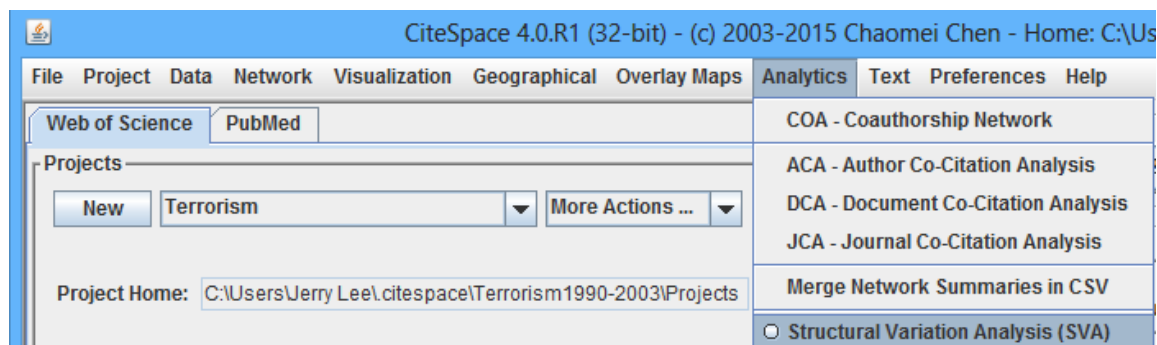
The 'New Project' dialog box shows the following settings:

- Title: Terrorism
- Project Home: C:\Users\Jerry Lee\citespace\Terrorism1990-2003\Projects
- Data Directory: C:\Users\Jerry Lee\citespace\Terrorism1990-2003\Data
- Language: ☒ English ☐ Chinese
- SO Filter: SC Filter:
- Alias List (on/off): on
- Exclusion List (on/off): on
- Export Space (on/off): on
- Export Abstracts (Time Consumption): on
- Export Matrices (csv) (off/on): on
- Enable JDIC (on/off): on
- Save Merged Slice (off/on): off
- Noun Phrase: Minimum Words: 2
- Noun Phrase: Maximum Word Length: 4
- Burst Term Threshold (0.00): 0.00
- Maximum GML Node Label Length: 8
- CTSA (1-Disciplines, 2-Scientific): 1
- Include GP (Group Author) (off/on): off
- Include ED (Editors) (off/on): off
- Node Degree Weighted (true): true
- Look Back Years (-1: unlimited): -1
- Max. No. Links to Retain (-1: unlimited): -1
- ☐ Normalize Citations ☐ Global Check
-

小提示

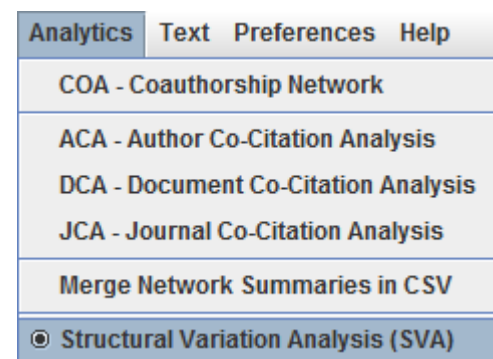
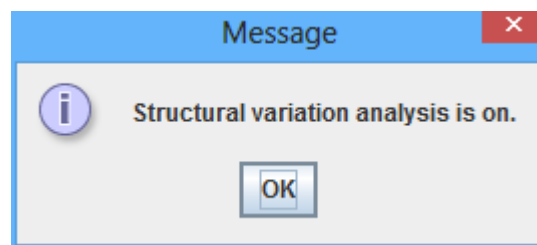
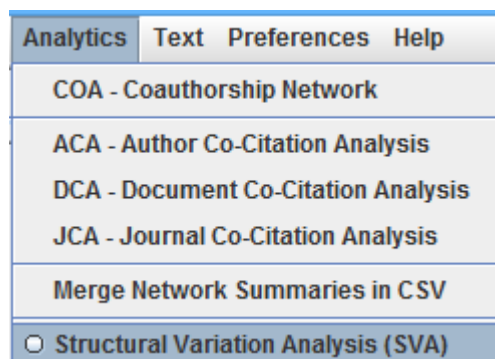
如果你不太清楚如何调整CiteSpace里project的参数，那么可以参见下面的博文：
<http://blog.sciencenet.cn/blog-496649-914950.html>

开启SVA功能



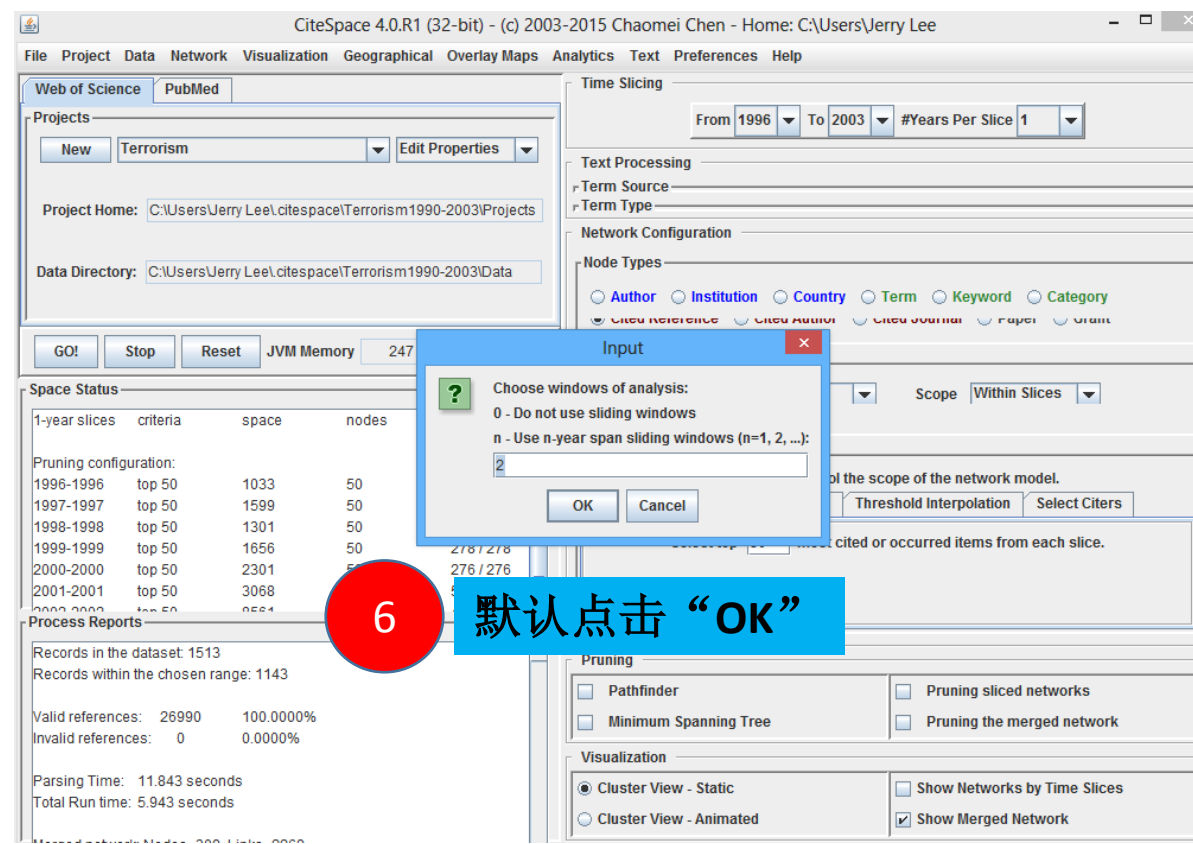
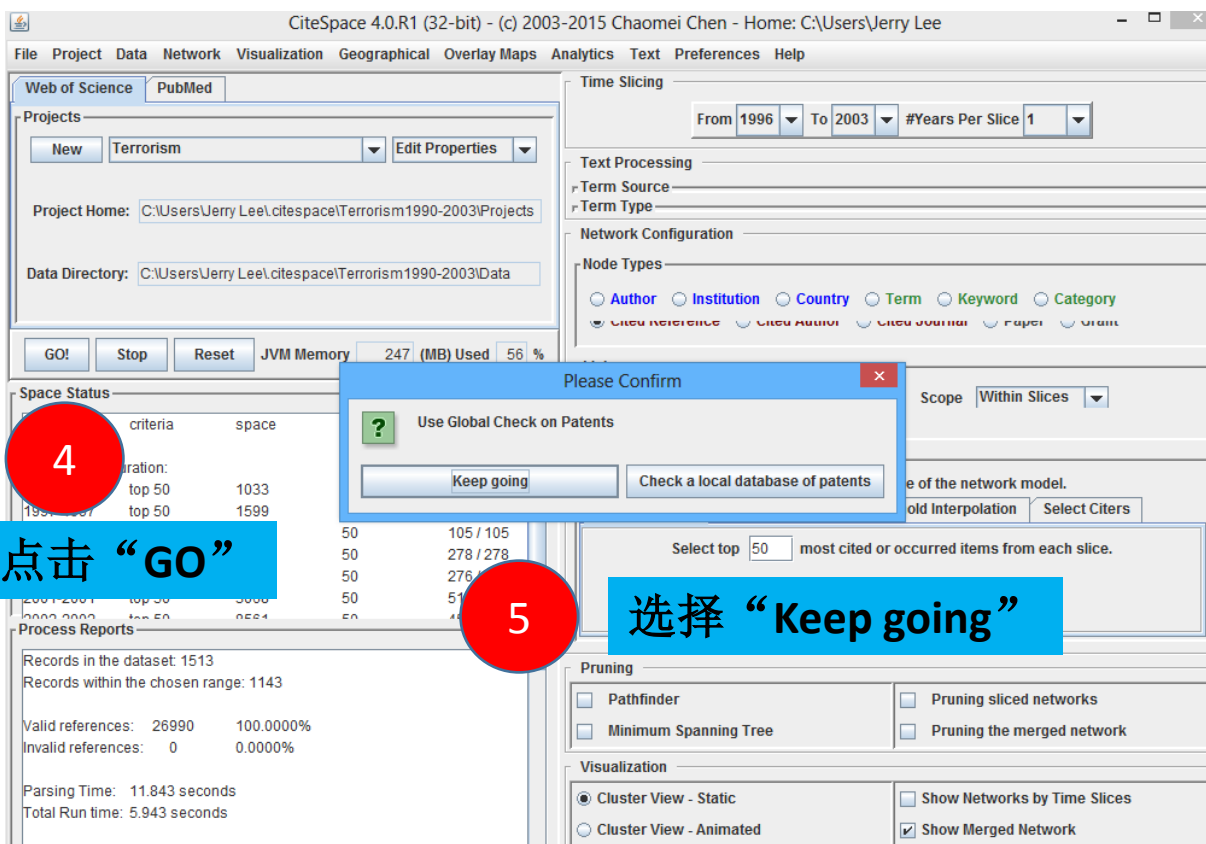
小提示

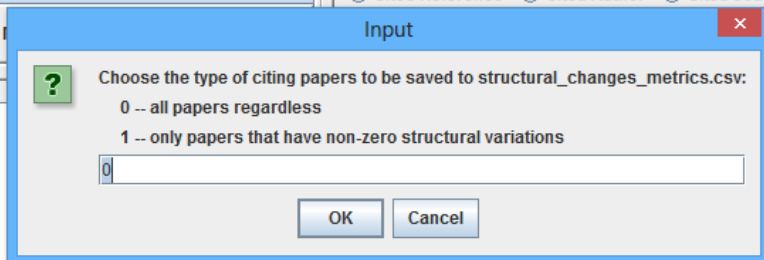
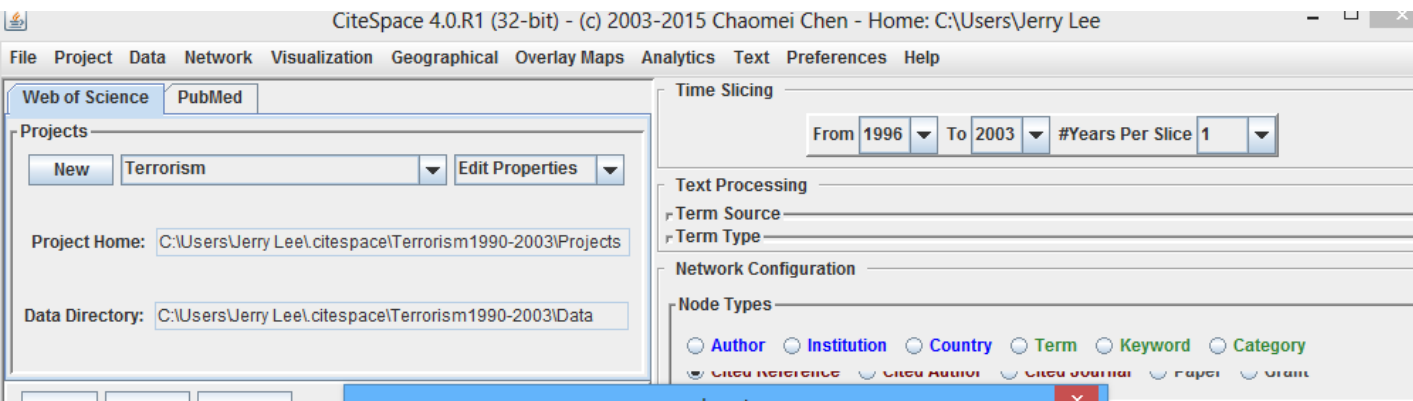
SVA的新功能就是加在了功能参数区的“Analytics”菜单中的“Structural Variation Analysis (SVA)”中。默认条件下该功能是关闭的，需要选择开启。



点击SVA后，在提示的信息框中点击“OK”即可开启SVA功能

在进行结果相关参数设置和开启SVA功能后，就可以点击“GO”运行分析结果。





默认点击“OK”

Space Status

1-year slices	criteria	space
1996-1996	top 50	1033
1997-1997	top 50	1599
1998-1998	top 50	1301
1999-1999	top 50	1656
2000-2000	top 50	2301
2001-2001	top 50	3068
2002-2002	top 50	4554

Pruning configuration:

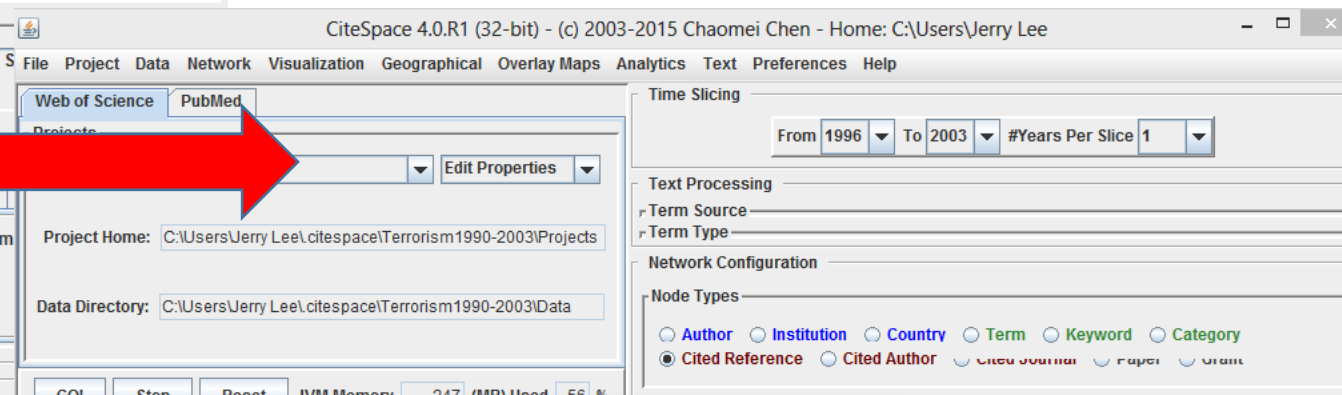
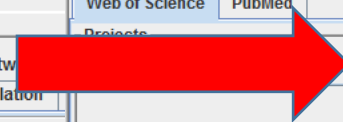
1-year slices	criteria	space	nodes
1996-1996	top 50	1033	50
1997-1997	top 50	1599	50
1998-1998	top 50	1301	50
1999-1999	top 50	1656	50
2000-2000	top 50	2301	50
2001-2001	top 50	3068	50
2002-2002	top 50	4554	50

Process Reports

Records in the dataset: 1513
Records within the chosen range: 1143

Valid references: 26990 100.0000%
Invalid references: 0 0.0000%

Parsing Time: 11.843 seconds
Total Run time: 5.943 seconds



默认点击“NO”

Space Status

1-year slices	criteria	space	nodes
1996-1996	top 50	1033	50
1997-1997	top 50	1599	50
1998-1998	top 50	1301	50
1999-1999	top 50	1656	50
2000-2000	top 50	2301	50
2001-2001	top 50	3068	50
2002-2002	top 50	4554	50

Pruning configuration:

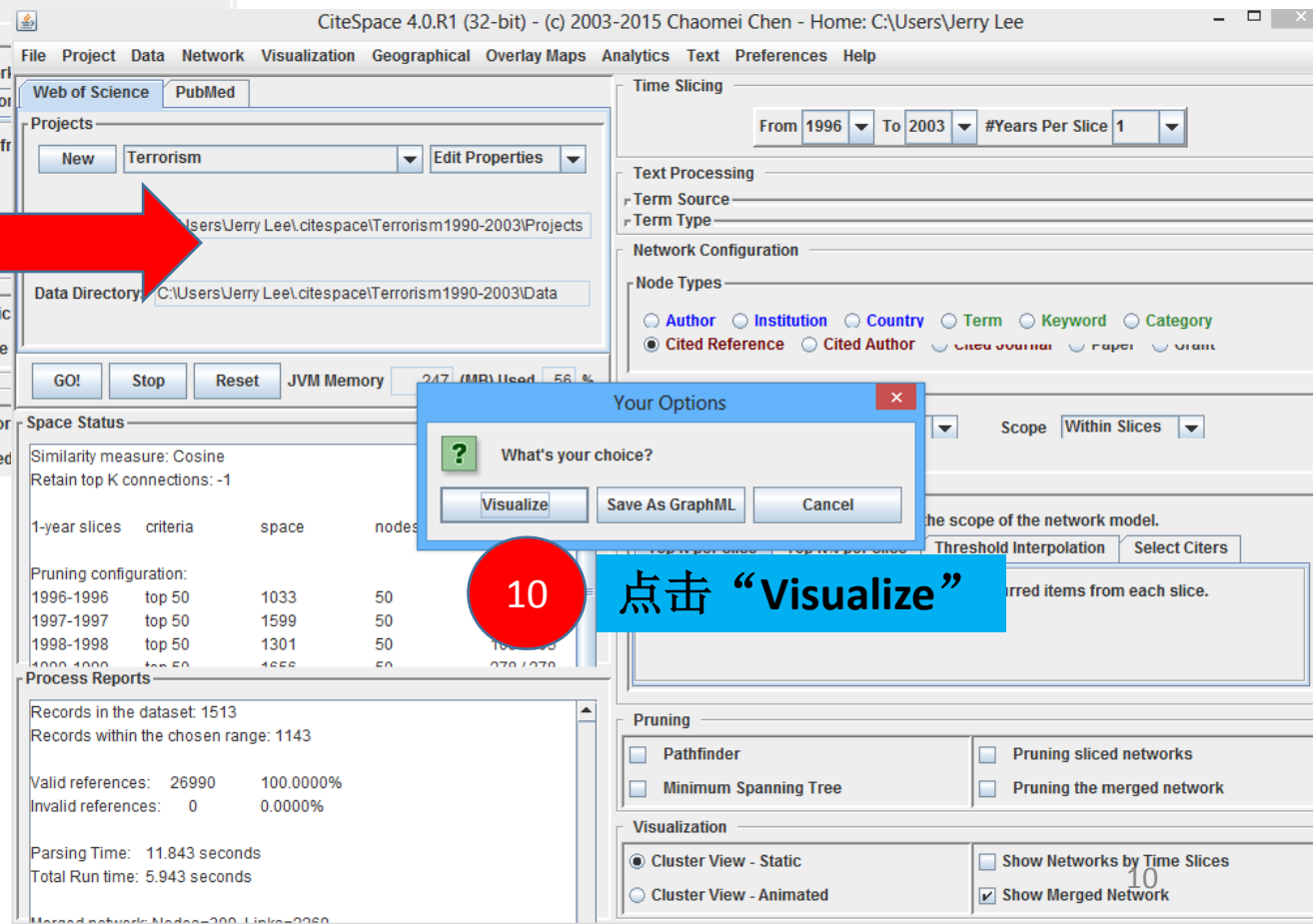
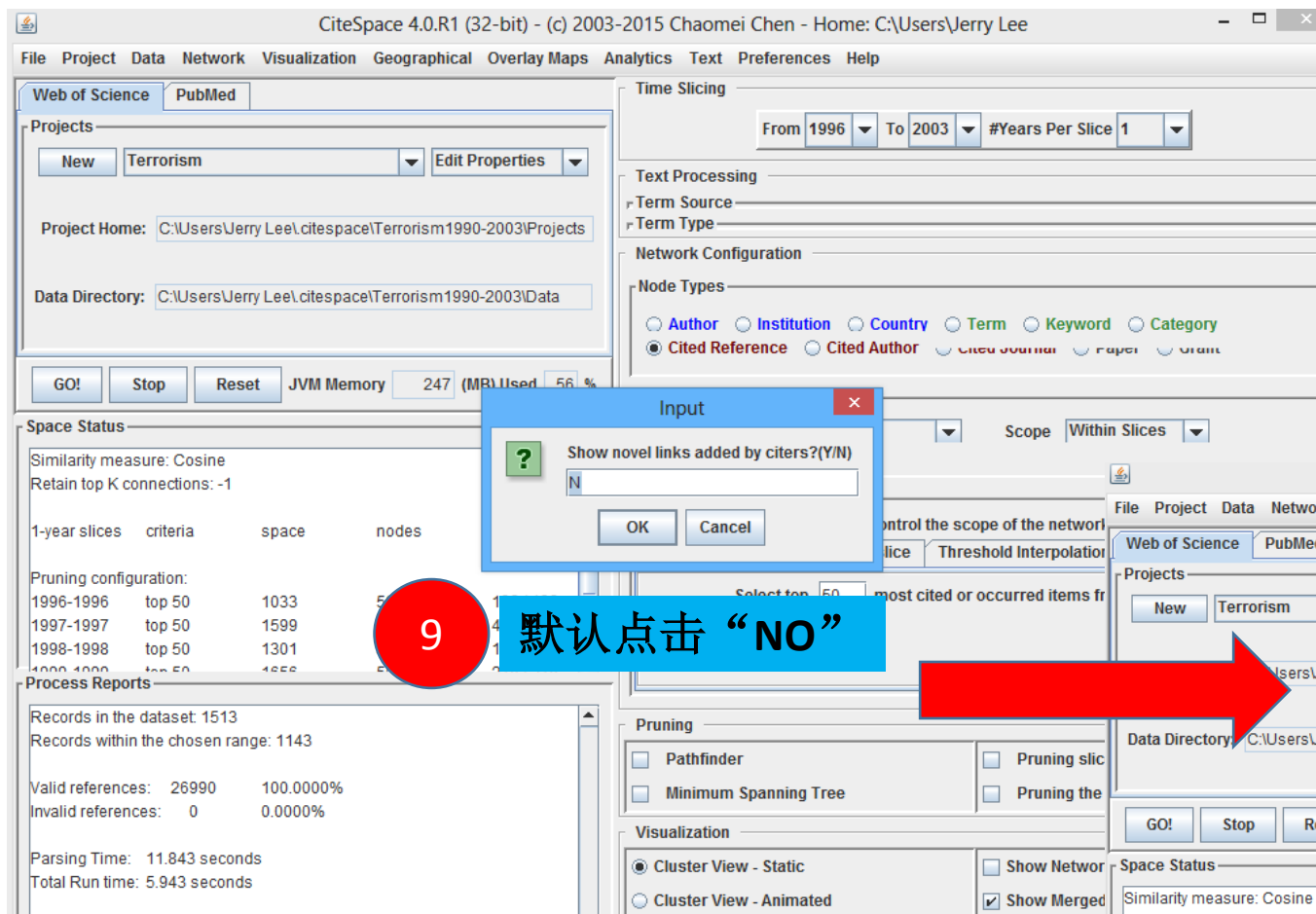
1-year slices	criteria	space	nodes
1996-1996	top 50	1033	50
1997-1997	top 50	1599	50
1998-1998	top 50	1301	50
1999-1999	top 50	1656	50
2000-2000	top 50	2301	50
2001-2001	top 50	3068	50
2002-2002	top 50	4554	50

Process Reports

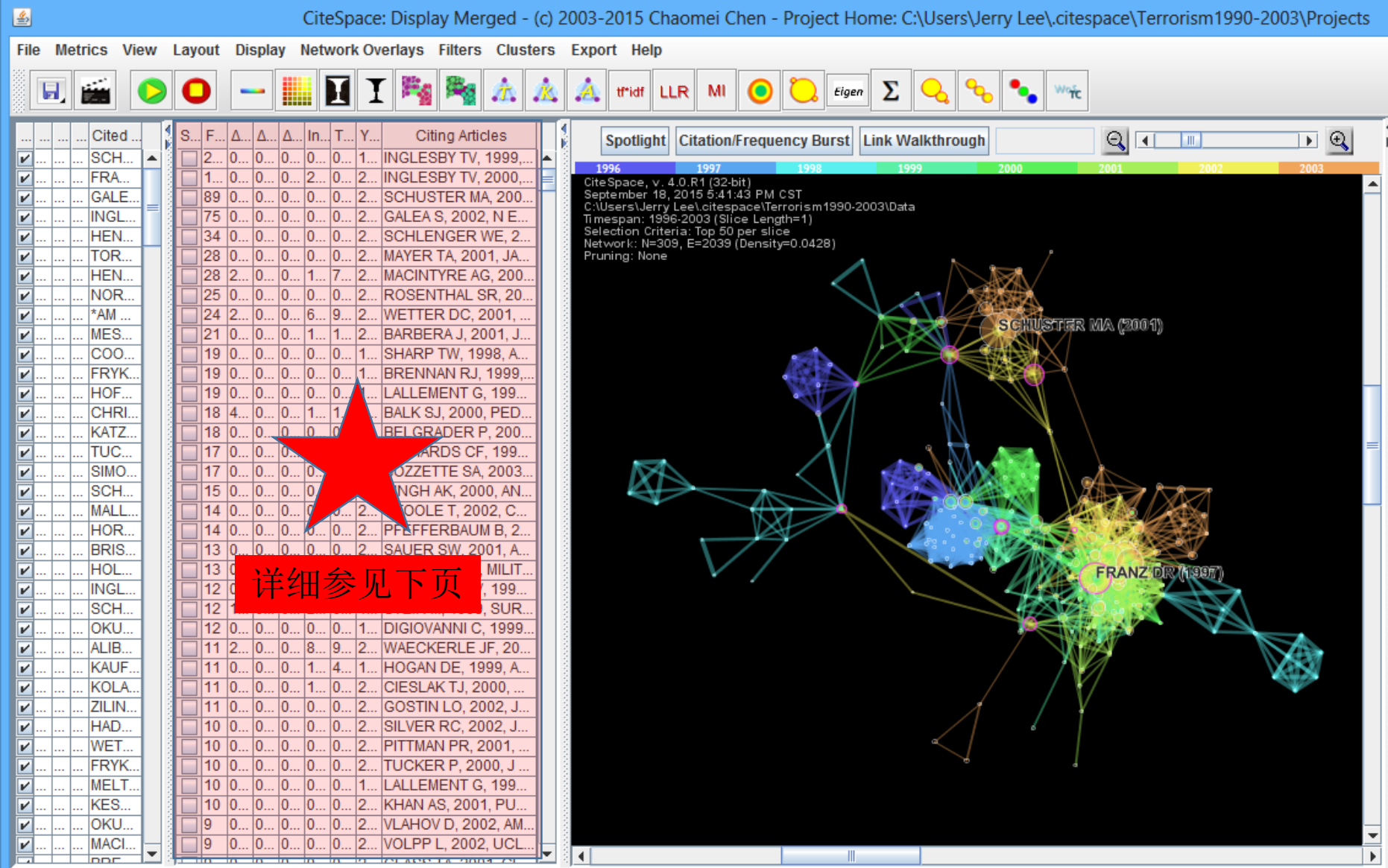
Records in the dataset: 1513
Records within the chosen range: 1143

Valid references: 26990 100.0000%
Invalid references: 0 0.0000%

Parsing Time: 11.843 seconds
Total Run time: 5.943 seconds



进入可视化界面后，首先可以发现现在可视化界面的左侧增加了一个信息列表栏。



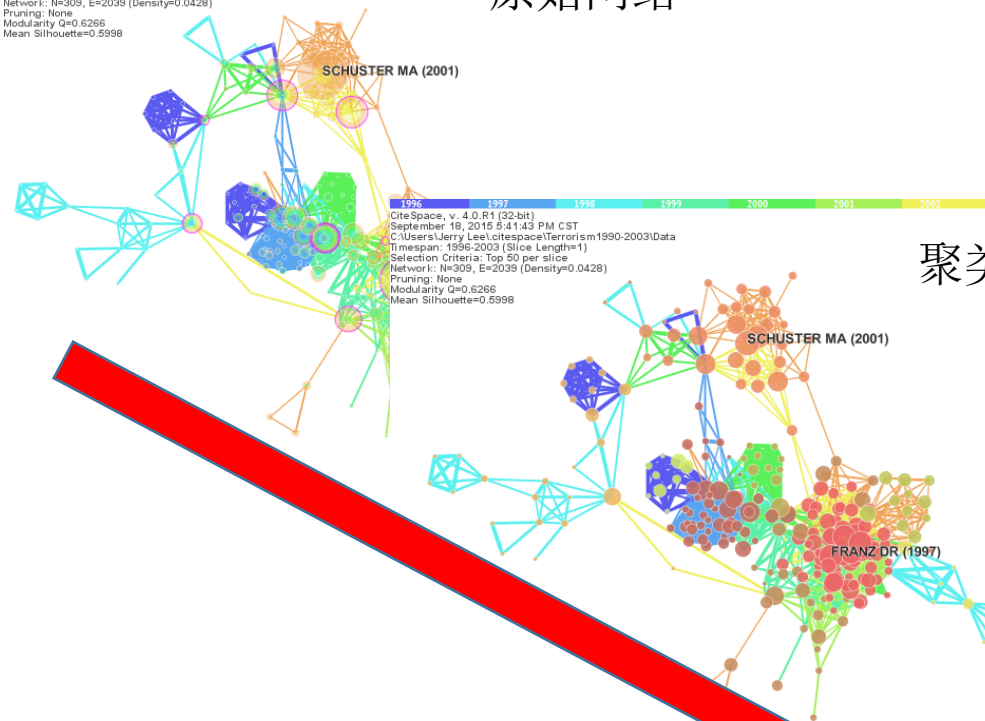
CiteSpace: Display Merged - (c) 2003-2015 Chaomei Chen - Project Home: C:\Users\Jerry Lee\citespace\Terrorism1990

File Metrics View Layout Display Network Overlays Filters Clusters Export Help

Show Links	Freq	Δ Modularity	Δ C-C Linkage	Δ Centrality	Incremental Links	Transformative ...	Year	Citing Articles
<input type="checkbox"/>	211	0.00	0.00	0.00	0.00	0.00	1999	INGLESBY TV, 1999, J...
<input type="checkbox"/>	106	0.10	0.00	0.00	2.00	0.00	2000	INGLESBY TV, 2000, J...
<input type="checkbox"/>	89	0.00	0.00	0.00	0.00	0.00	2001	SCHUSTER MA, 2001, ...
<input type="checkbox"/>	75	0.00	0.00	0.00	0.00	0.00	2002	GALEA S, 2002, N EN...
<input type="checkbox"/>	34	0.00	0.00	0.00	0.00	0.00	2002	SCHLENGER WE, 20...
<input type="checkbox"/>	28	2.63	0.01	0.13	1.00	7.00	2000	MACINTYRE AG, 2000, ...
<input type="checkbox"/>	28	0.00	0.00	0.00	0.00	0.00	2001	MAYER TA, 2001, JAM...
<input type="checkbox"/>	25	0.00	0.00	0.00	0.00	0.00	2001	ROSENTHAL SR, 200...
<input type="checkbox"/>	24	2.85	0.11	0.01	6.00	9.00	2001	WETTER DC, 2001, A...
<input type="checkbox"/>	21	0.30	0.00	0.00	1.00	1.00	2001	BARBERA J, 2001, JA...
<input type="checkbox"/>	19	0.00	0.00	0.00	0.00	0.00	1998	SHARP TW, 1998, AN...
<input type="checkbox"/>	19	0.00	0.00	0.00	0.00	0.00	1999	BRENNAN RJ, 1999, A...
<input type="checkbox"/>	19	0.00	0.00	0.00	0.00	0.00	1998	LALLEMENT G, 1998, ...
<input type="checkbox"/>	18	4.48	0.08	0.19	12.00	10.00	2000	BALK SJ, 2000, PEDIA...
<input type="checkbox"/>	18	0.00	0.00	0.00	0.00	0.00	2001	BELGRADER P, 2001, ...
<input type="checkbox"/>	17	0.00	0.00	0.00	0.00	0.00	1999	RICHARDS CF, 1999, ...
<input type="checkbox"/>	17	0.00	0.00	0.00	0.00	0.00	2003	BOZZETTE SA, 2003, ...
<input type="checkbox"/>	15	0.00	0.00	0.00	0.00	0.00	2000	SINGH AK, 2000, ANA...
<input type="checkbox"/>	14	0.00	0.00	0.00	0.00	0.00	2002	OTOOLE T, 2002, CLI...
<input type="checkbox"/>	14	0.00	0.00	0.00	0.00	0.00	2001	PFEFFERBAUM B, 20...
<input type="checkbox"/>	13	0.00	0.00	0.00	0.00	0.00	2001	SAUER SW, 2001, AN...
<input type="checkbox"/>	13	0.00	0.00	0.00	0.00	0.00	1999	HAIL AS, 1999, MILIT ...
<input type="checkbox"/>	12	1.00	0.03	0.00	2.00	6.00	1999	STEIN M, 1999, SURG...
<input type="checkbox"/>	12	0.00	0.00	0.00	0.00	0.00	1998	FELICIANO DV, 1998, ...
<input type="checkbox"/>	12	0.00	0.00	0.00	0.00	0.00	1999	DIGIOVANNI C, 1999, ...
<input type="checkbox"/>	11	2.85	0.09	0.01	8.00	9.00	2001	WAECKERLE JF, 200...
<input type="checkbox"/>	11	0.82	0.01	0.00	1.00	4.00	1999	HOGAN DE, 1999, AN...
<input type="checkbox"/>	11	0.05	0.00	0.00	1.00	0.00	2000	CIESLAK TJ, 2000, CL...
<input type="checkbox"/>	11	0.00	0.00	0.00	0.00	0.00	2002	GOSTIN LO, 2002, JA...
<input type="checkbox"/>	10	0.00	0.00	0.00	0.00	0.00	2002	SILVER RC, 2002, JA...
<input type="checkbox"/>	10	0.00	0.00	0.00	0.00	0.00	2001	PITTMAN PR, 2001, VA...
<input type="checkbox"/>	10	0.00	0.00	0.00	0.00	0.00	2000	TUCKER P, 2000, J B...
<input type="checkbox"/>	10	0.00	0.00	0.00	0.00	0.00	1999	LALLEMENT G, 1999, ...
<input type="checkbox"/>	10	0.00	0.00	0.00	0.00	0.00	2001	KHAN AS, 2001, PUBL...
<input type="checkbox"/>	9	0.00	0.00	0.00	0.00	0.00	2002	VLAHOV D, 2002, AME...
<input type="checkbox"/>	9	0.00	0.00	0.00	0.00	0.00	2002	VOLPP L, 2002, UCLA ...

CiteSpace, v. 4.0.R1 (32-bit)
September 18, 2015 5:41:43 PM CST
C:\Users\Jerry Lee\citespace\Terrorism1990-2003\Data
Timespan: 1996-2003 (Slice Length=1)
Selection Criteria: Top 50 per slice
Network: N=309, E=2039 (Density=0.0428)
Pruning: None
Modularity Q=0.6266
Mean Silhouette=0.5998

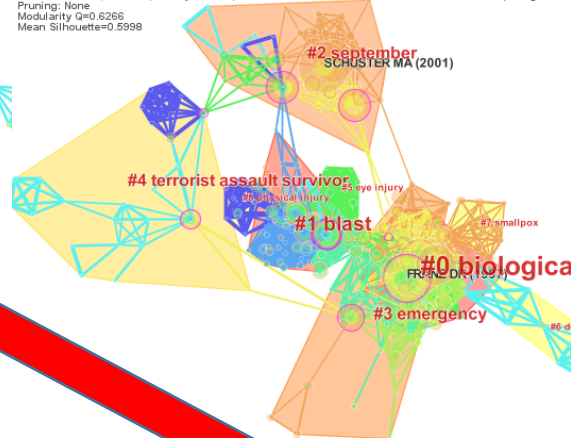
原始网络



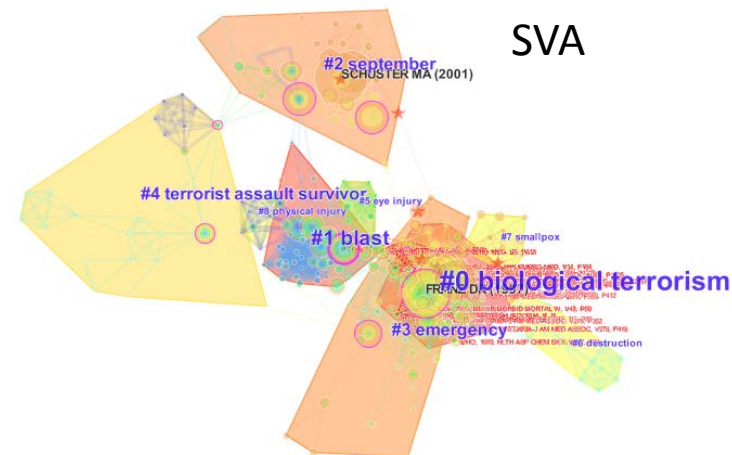
聚类

CiteSpace, v. 4.0.R1 (32-bit)
September 18, 2015 5:41:43 PM CST
C:\Users\Jerry Lee\citespace\Terrorism1990-2003\Data
Timespan: 1996-2003 (Slice Length=1)
Selection Criteria: Top 50 per slice
Network: N=309, E=2039 (Density=0.0428)
Pruning: None
Modularity Q=0.6266
Mean Silhouette=0.5998

聚类命名



聚类命名



接着我们这里依次完成网络的聚类、聚类命名以及图形的调整。

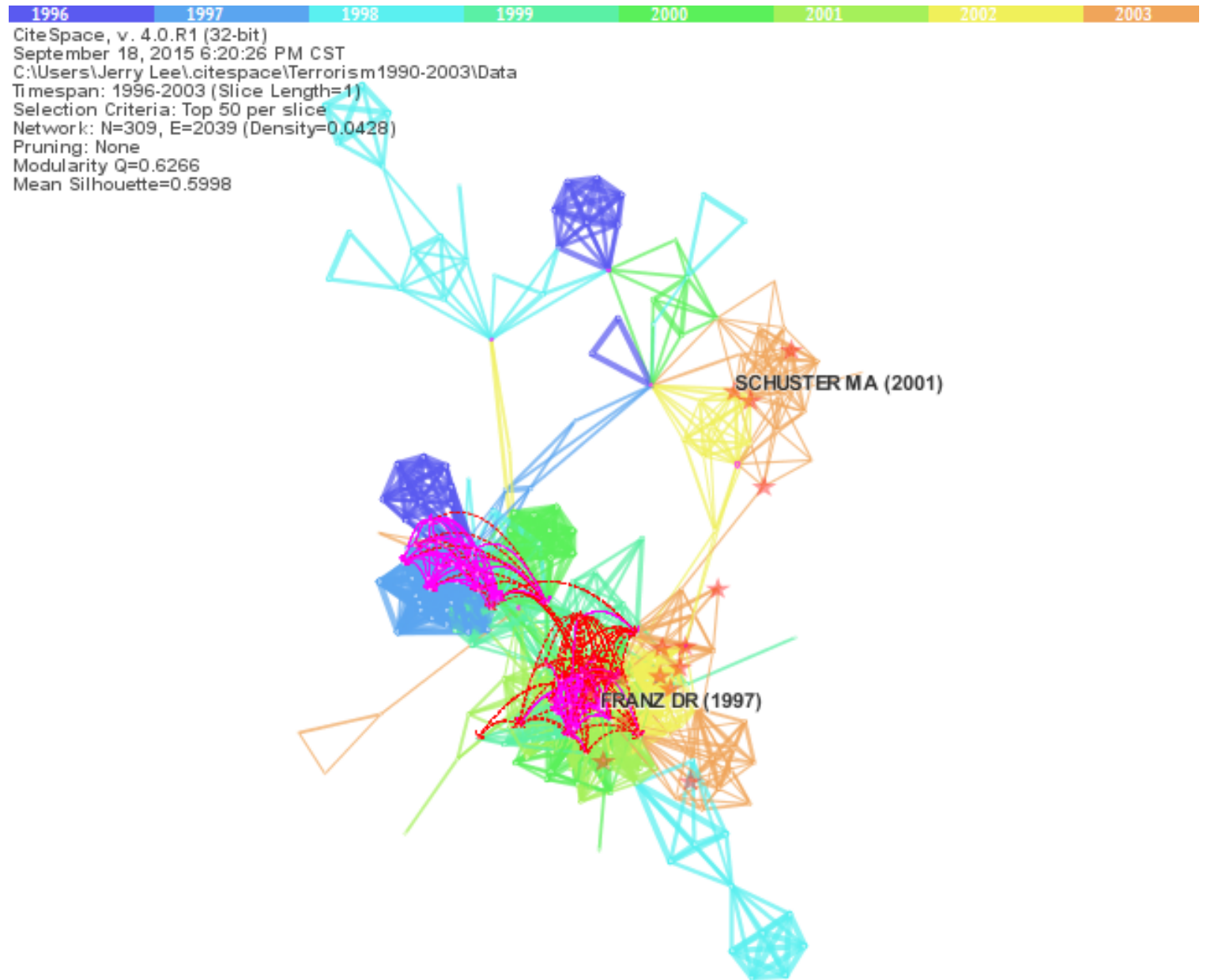
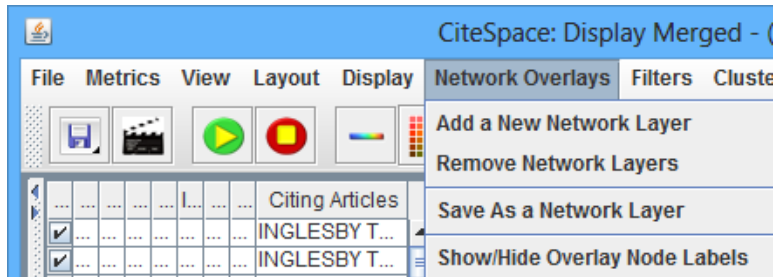
通过左侧新的列表信息来进行分析。

勾选需要关注的文献，会发现图上有对应的变化，显示了标签和红色虚线。

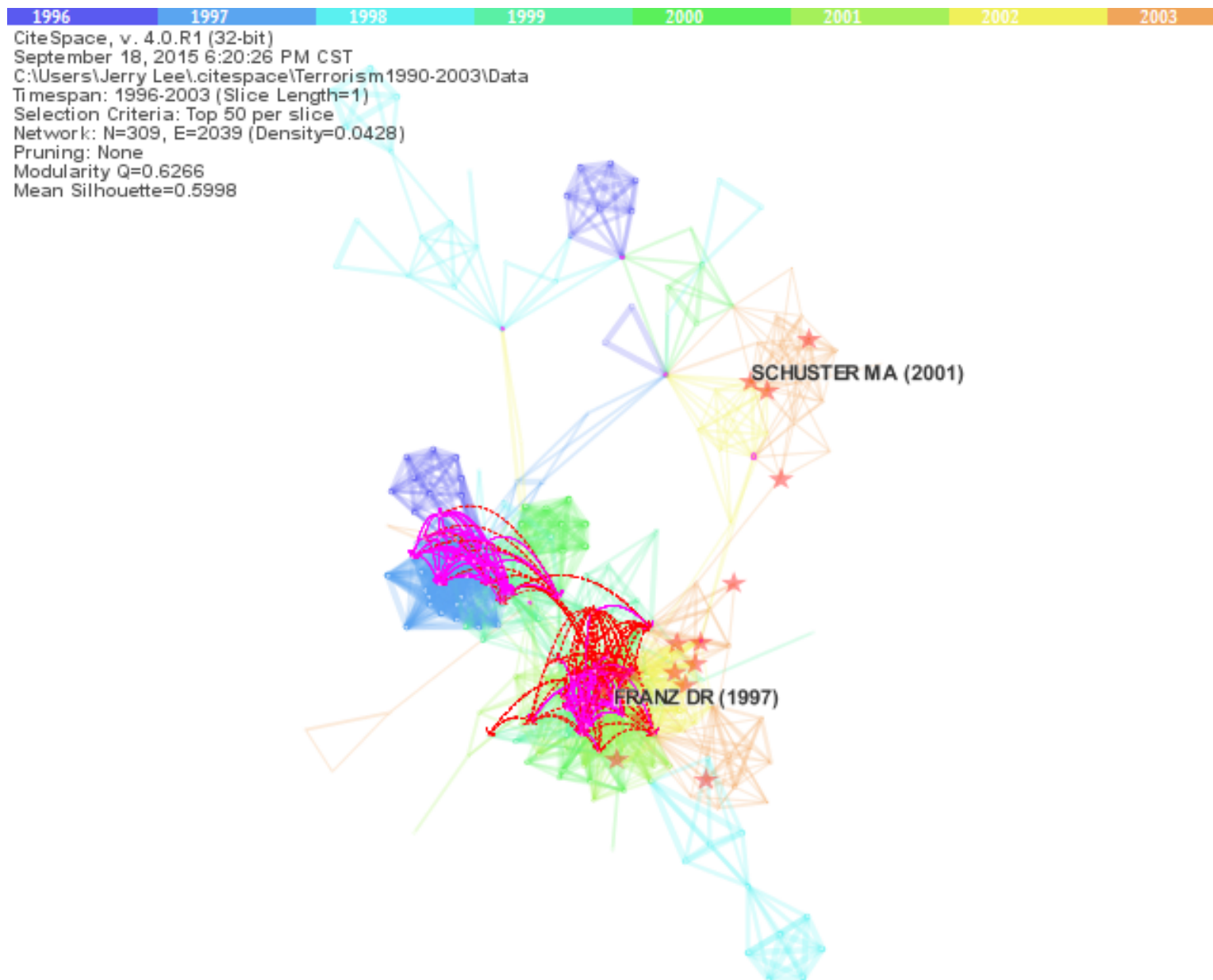
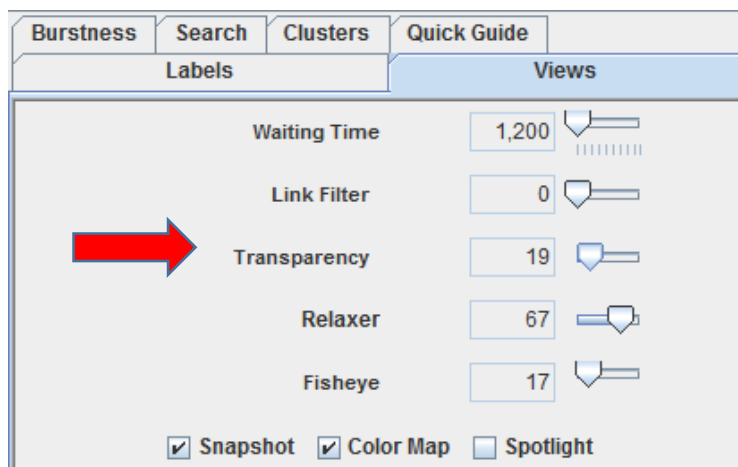
为了使得图形比较简洁，可以取消掉红色的标签。具体步骤如下：

Network overlay→Show/Hide

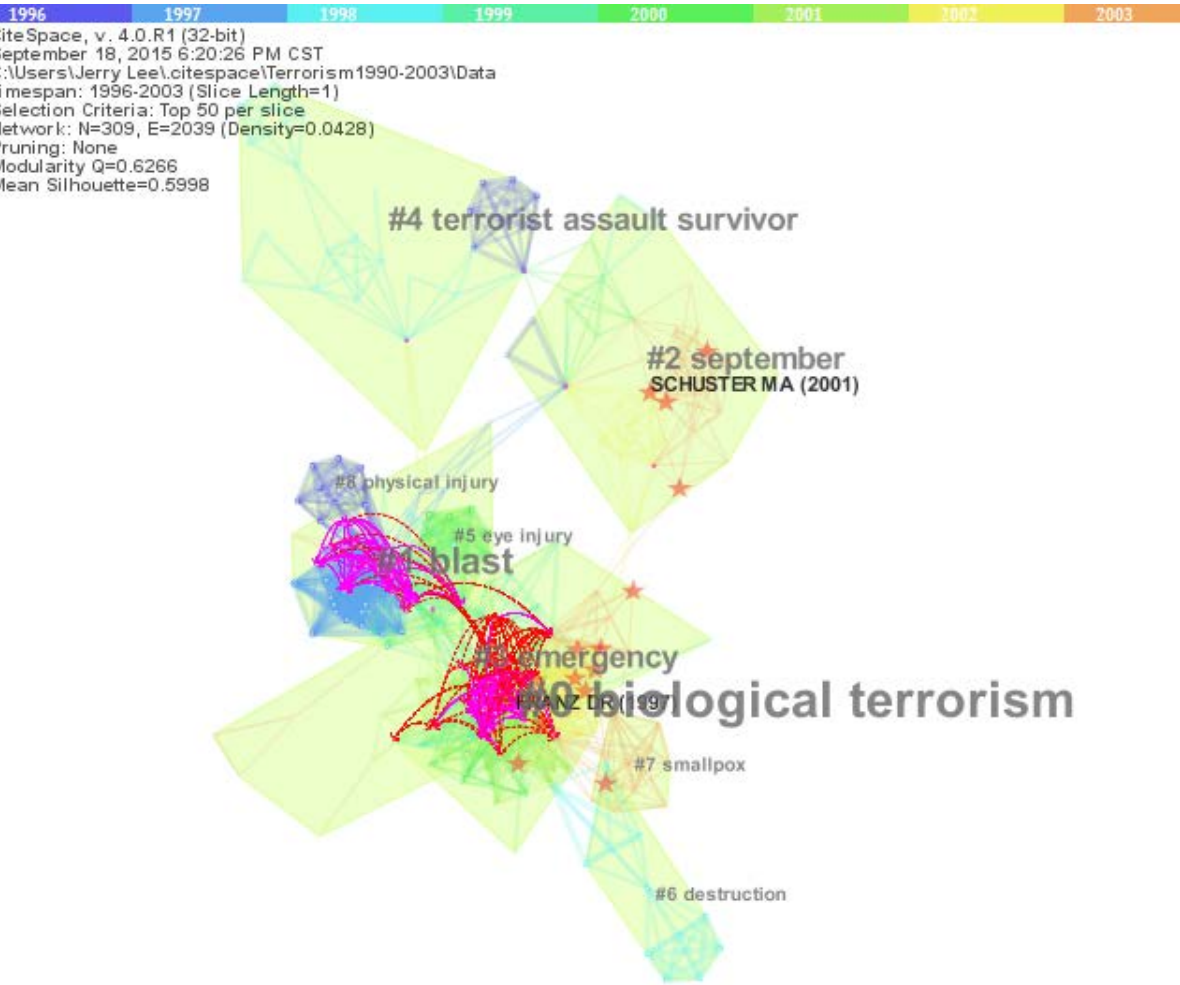
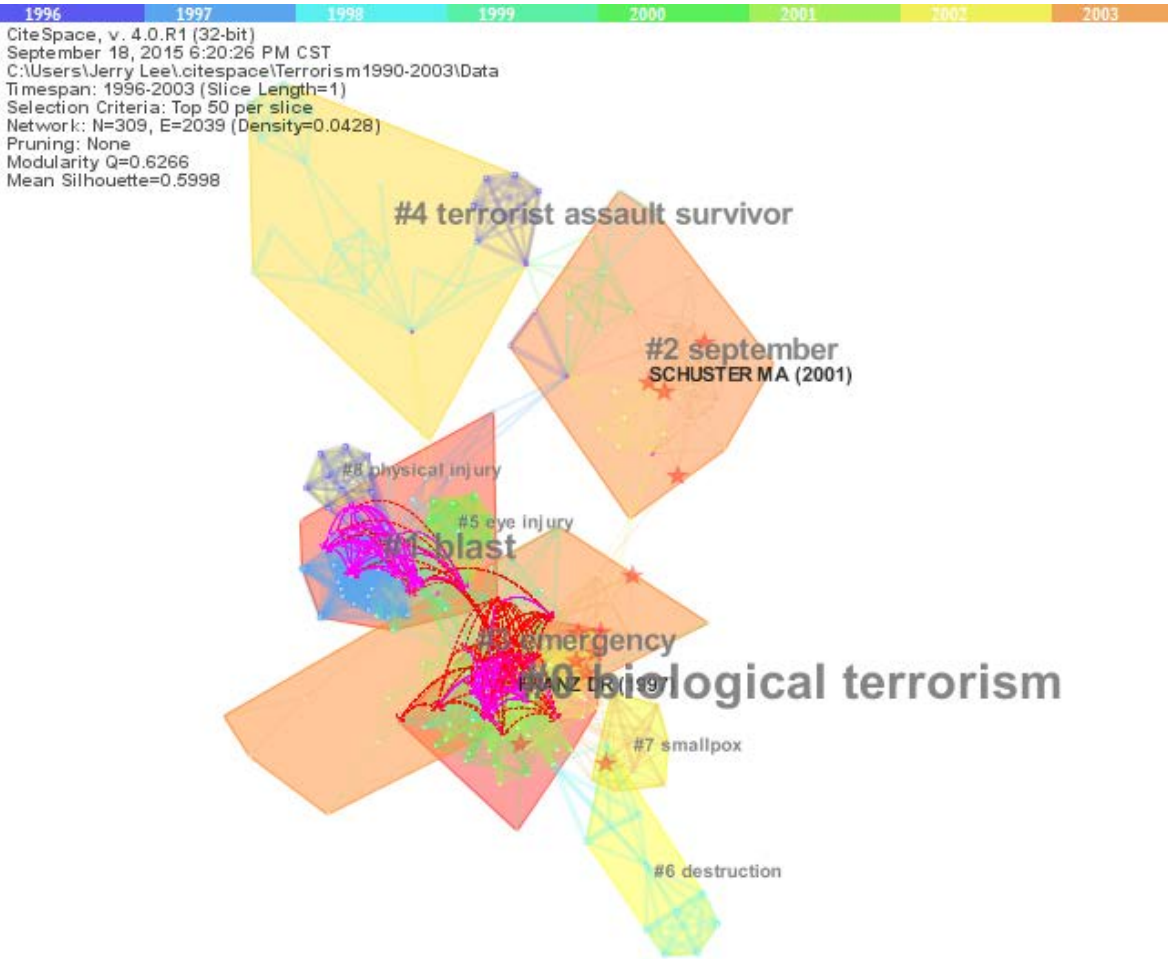
Overlay Node Labels。



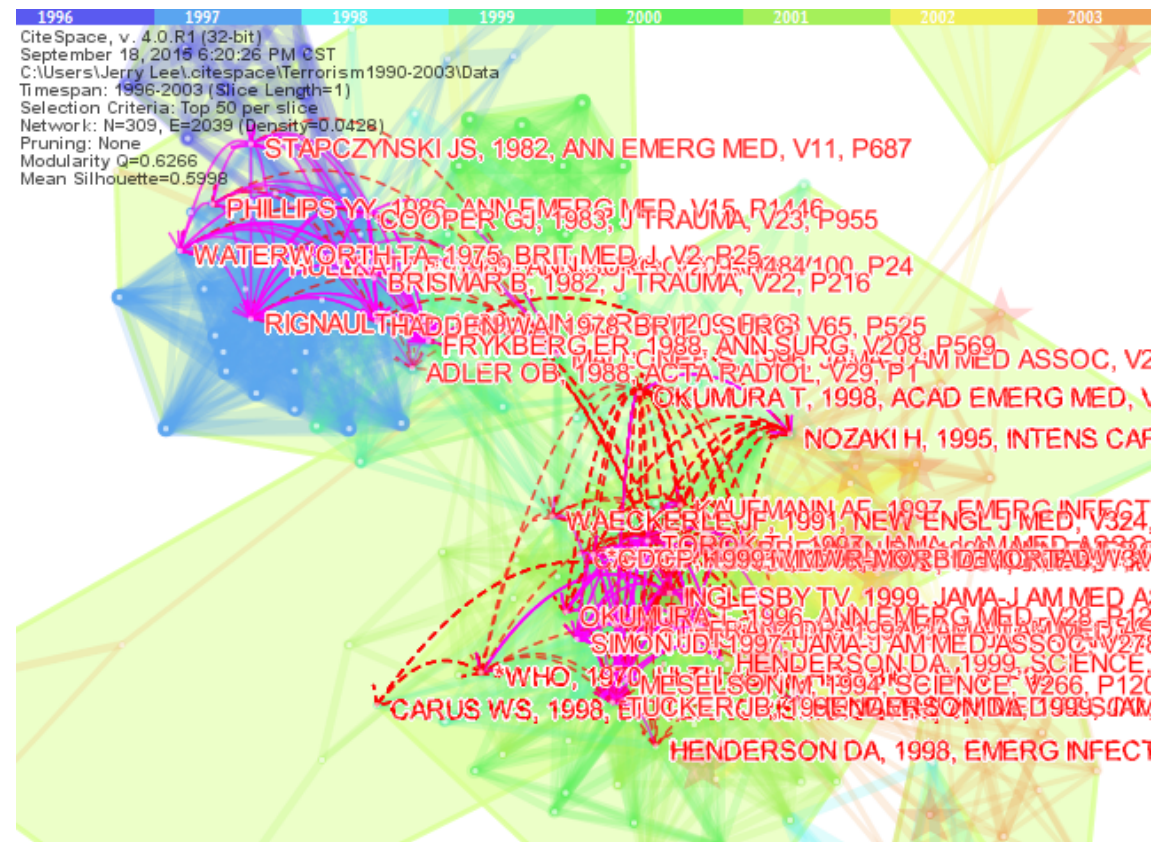
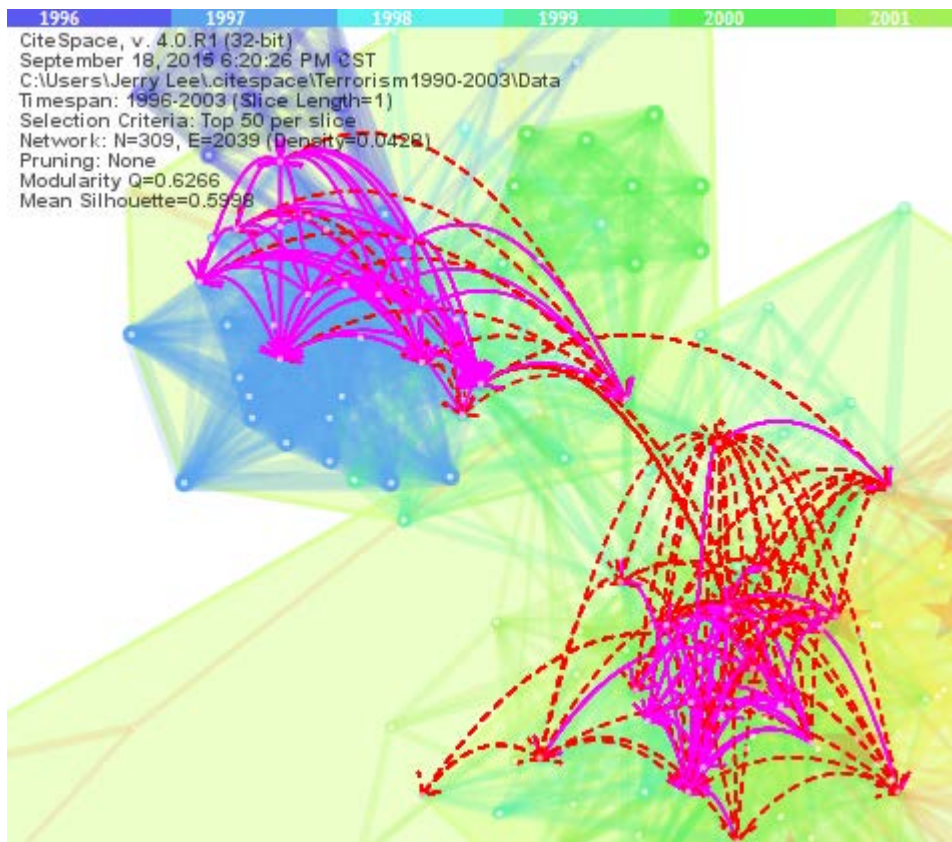
调整共被引连线的透明度，来凸显SVA分析的结果。



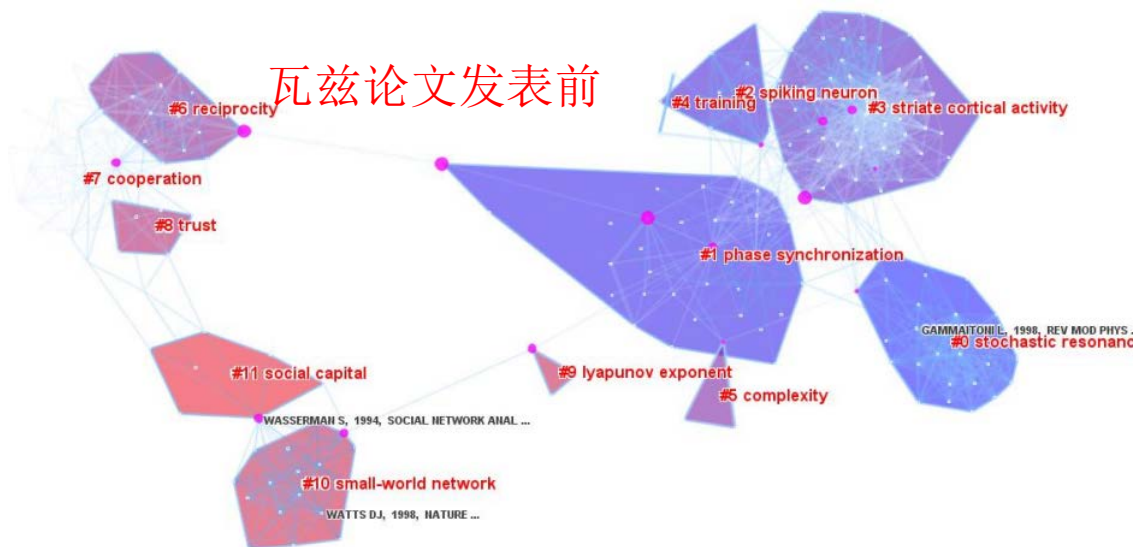
当然也可以把聚类信息加上，这样就能看到所关注的结果在哪些类里面。
另外为了显示的简洁，也可以将聚类的背景颜色调整为淡色。



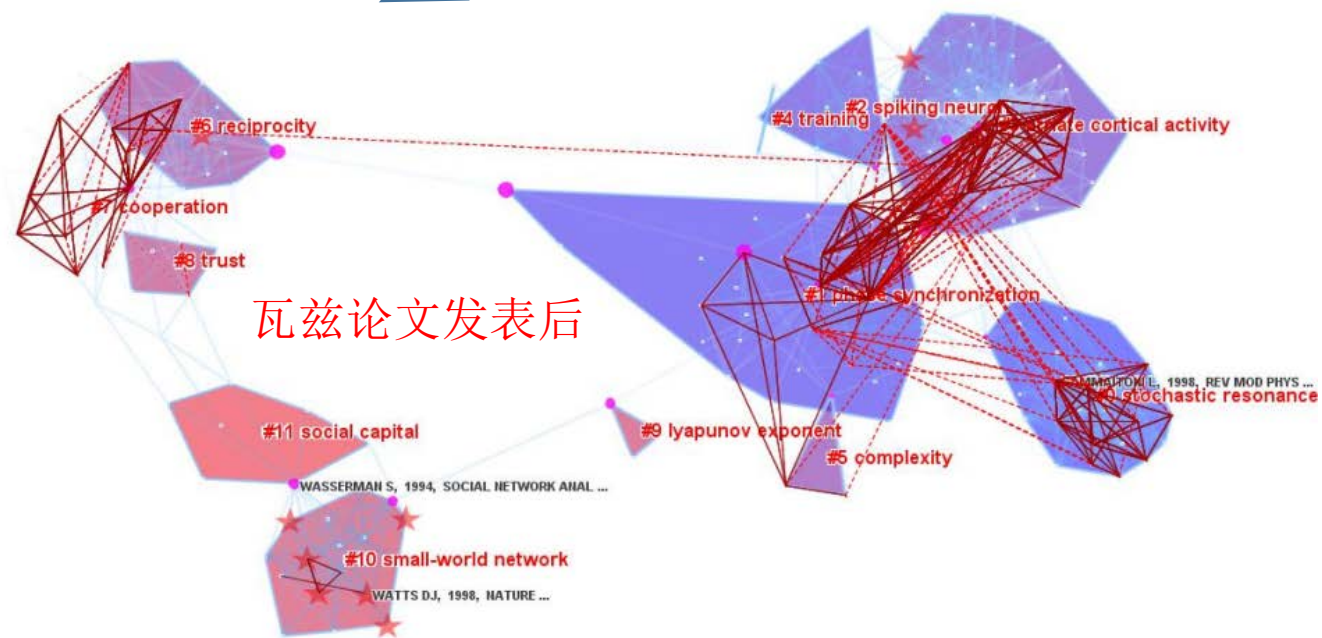
可以放大看看这里都是些什么内容。
实线为已有连接，虚线为新增连接。



更多的例子



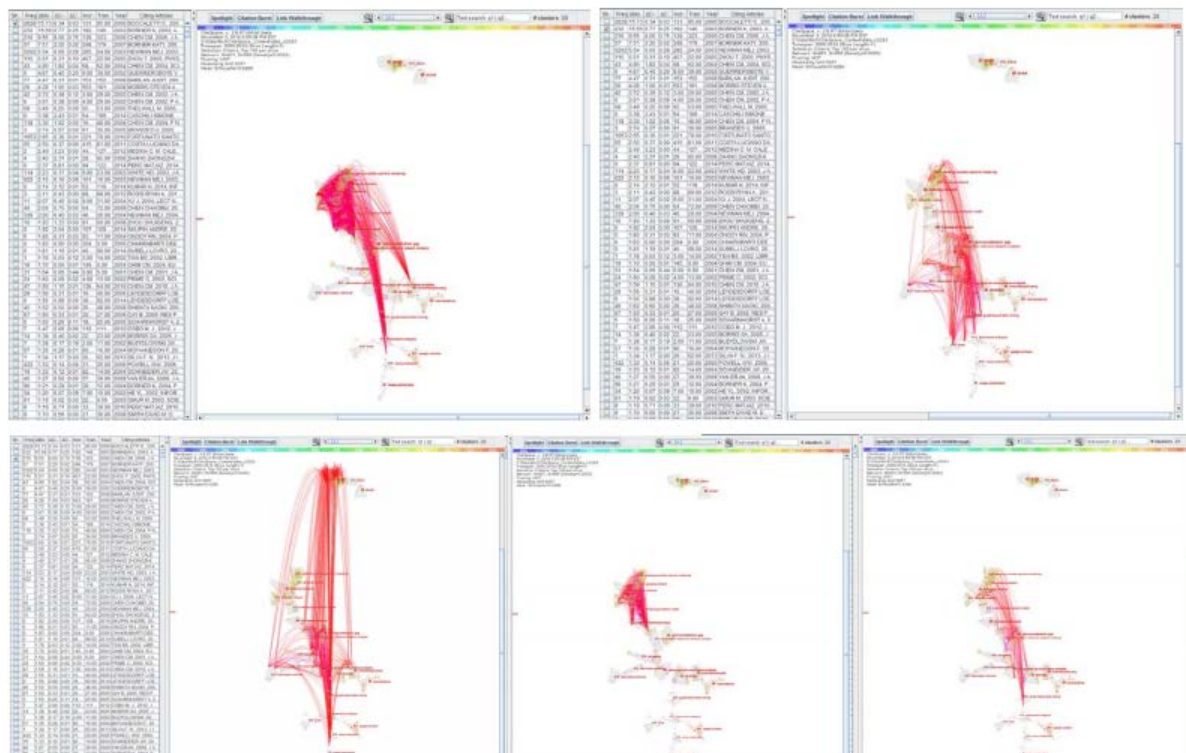
瓦兹论文发表前后的网络结构



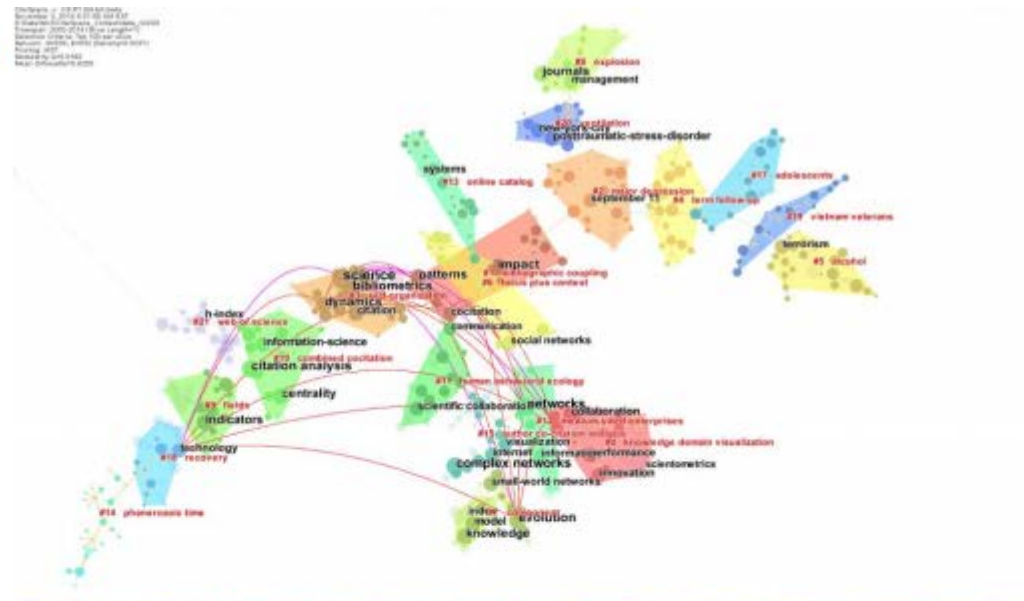
详见: Chen, C. (2014) *The Fitness of Information: Quantitative Assessments of Critical Evidence*. Wiley.

更多的例子

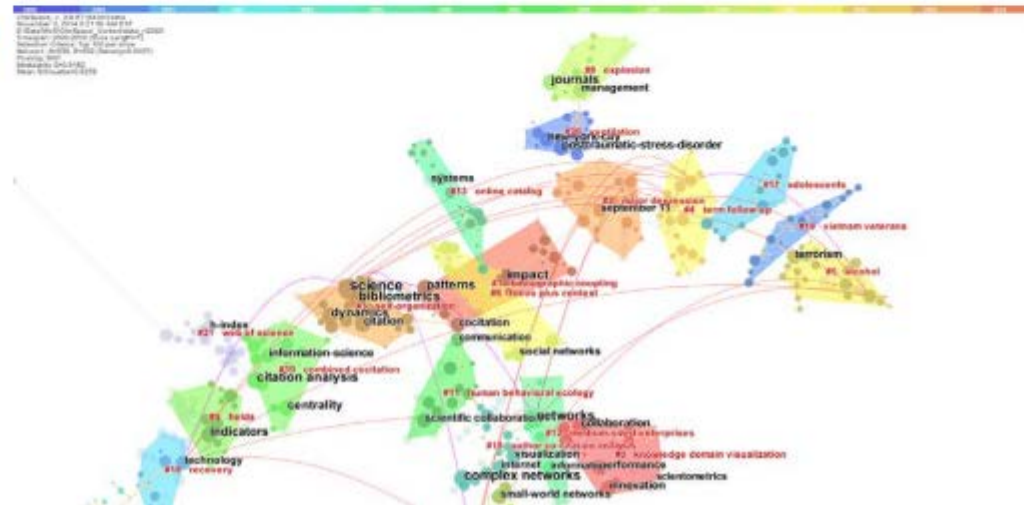
Modularity Change Rate



2004 PNAS



2006 JASIST



详见：Chen (2015) Boundary-Spanning Turning Points and Emerging Paradigms. <http://blog.sciencenet.cn/blog-496649-903442.html>

参考文献

有关该功能的原理及应用请读者参见下列文献：

1. Chen, C. (2012) [Predictive effects of structural variation on citation counts](#). Journal of the American Society for Information Science and Technology, 63(3), 431-449.
2. Chen, C. (2014) [The Fitness of Information: Quantitative Assessments of Critical Evidence](#). Wiley.
3. Chen (2015) Boundary-Spanning Turning Points and Emerging Paradigms. <http://blog.sciencenet.cn/blog-496649-903442.html>

相关介绍：

1. CiteSpace 4.0.R1增加结构变异分析功能. <http://blog.sciencenet.cn/blog-496649-921630.html>
2. Structural Variation Analysis (SVA). <https://sites.google.com/site/citespace101/11-advanced-topics/11-3structural-variation-analysis-sva>

更多Citespace学习资料：

1. 陈超美. [CiteSpaceManual](http://cluster.ischool.drexel.edu/~cchen/citespace/CiteSpaceManual.pdf)（英文）<http://cluster.ischool.drexel.edu/~cchen/citespace/CiteSpaceManual.pdf>
2. 陈悦，陈超美等. 引文空间分析原理与应用[M].科学出版社. 2014.
3. 陈超美著，陈悦等译. 科学前沿图谱 知识可视化探索[M].科学出版社. 2014.
4. 李杰. 如何使用CiteSpace的一组示范及常见问题解答<http://blog.sciencenet.cn/blog-496649-838067.html>
5. 李杰. CiteSpace中文指南. <http://blog.sciencenet.cn/blog-496649-886962.html>
6. 李杰，陈超美. CiteSpace科技文本挖掘及可视化[M]. 首都经济贸易大学出版社.2016.