

g-index:节点选择的新方法

November 16, 2015 11:12:26 AM CET
C:\Users\Jerry Lee\citespace\Terrorism1990-2003\Data
Timespan: 1996-2003 (Slice Length=1)
Selection Criteria: g-index
Network: N=126, E=777 (Density=0.0987)
Pruning: None
Modularity Q=0.5615
Mean Silhouette=0.4682



李杰¹, 陈超美²

1. 首都经济贸易大学-安全与环境工程学院

2. Drexel University- College of Computing and Informatics

2015年11月16日

小提示

CiteSpace版本至少为:
CiteSpace 4.0.R3

配套教程: 李杰, 陈超美著. CiteSpace科技文本挖掘及可视化[M]. 首都经济贸易大学出版社. 2016.

作者博客: 李杰博客: <http://blog.sciencenet.cn/u/jerrycueb>; 陈超美博客: <http://blog.sciencenet.cn/u/ChaomeiChen>

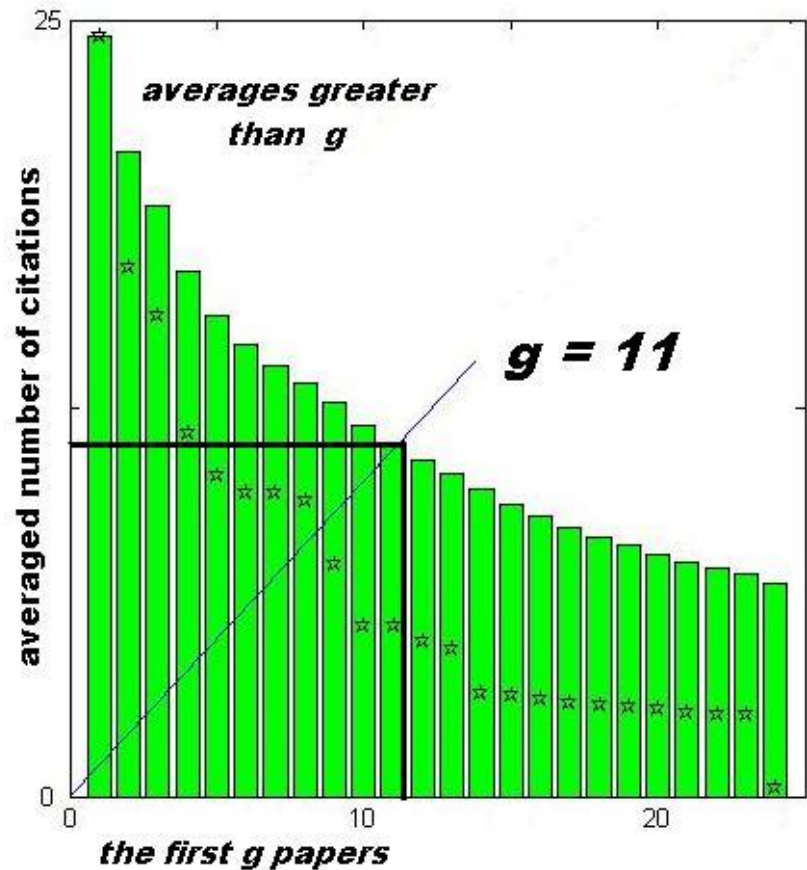
写在前面

在使用本功能之前请读者详细学习g-index的有关文献，这对于理解和认识CiteSpace中嵌入的功能很有益处。

推荐的学习文件如下：

1. Egghe, Leo (2006) Theory and practise of the g-index, *Scientometrics*, vol. 69, No 1, pp. 131–152. [doi:10.1007/s11192-006-0144-7](https://doi.org/10.1007/s11192-006-0144-7)
2. [Woeginger, G.J.](#) (2008) An axiomatic analysis of Egghe's g-index, *Journal of Informetrics*, vol. 2, pp. 364–368. [doi:10.1016/j.joi.2008.05.002](https://doi.org/10.1016/j.joi.2008.05.002)
3. Tol, R.S.J. (2008) A rational, successive g-index applied to economics departments in Ireland, *Journal of Informetrics*, vol. 2, pp. 149–155.

关于g-index



2006年, Egghe提出了g指数, g指数定义为:论文按被引次数排序后相对排前的累积被引至少 g^2 次的最大论文序次 g , 亦即第 $(g+1)$ 序次论文对应的累积引文数将小于 $(g+1)^2$ 。从定义可以看出, $g \geq h$, 而按被引量排序靠前的文章的被引次数越大, g指数越大。

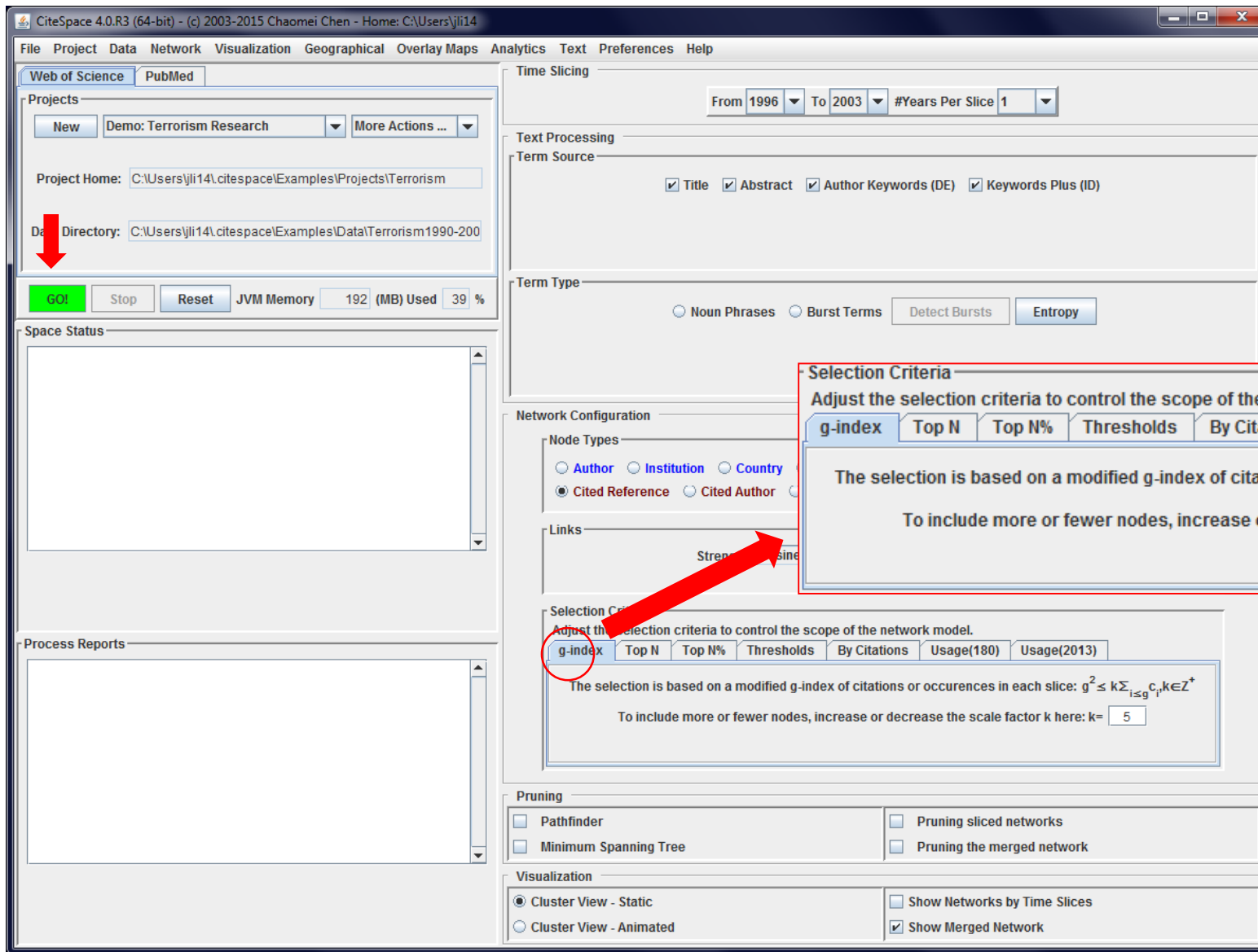
来源于百度百科:

<http://baike.baidu.com/link?url=CXn5jIBkZuTxoIT8-Xsc0Ht5tXxVetbk9HCzmOcEVU6KTyeXwUYDbLvLsWJJDICSCzg-PVE1xn1WBryySgrTKa>

An example of a g-index (the raw citation data, plotted with stars, allows the h-index to also be extracted for comparison).

来源于维基百科https://en.wikipedia.org/wiki/G-index#cite_note-Egghe-1

Step1:使用默认的“Terrorism”数据，在数据筛选中选择g-index，其他条件默认下点击GO。



Step2: 等待计算结束, 点击Visualize

The screenshot displays the CiteSpace 4.0.R3 software interface. The main window is titled 'CiteSpace 4.0.R3 (64-bit) - (c) 2003-2015 Chaomei Chen - Home: C:\Users\jli14'. The interface is divided into several panels:

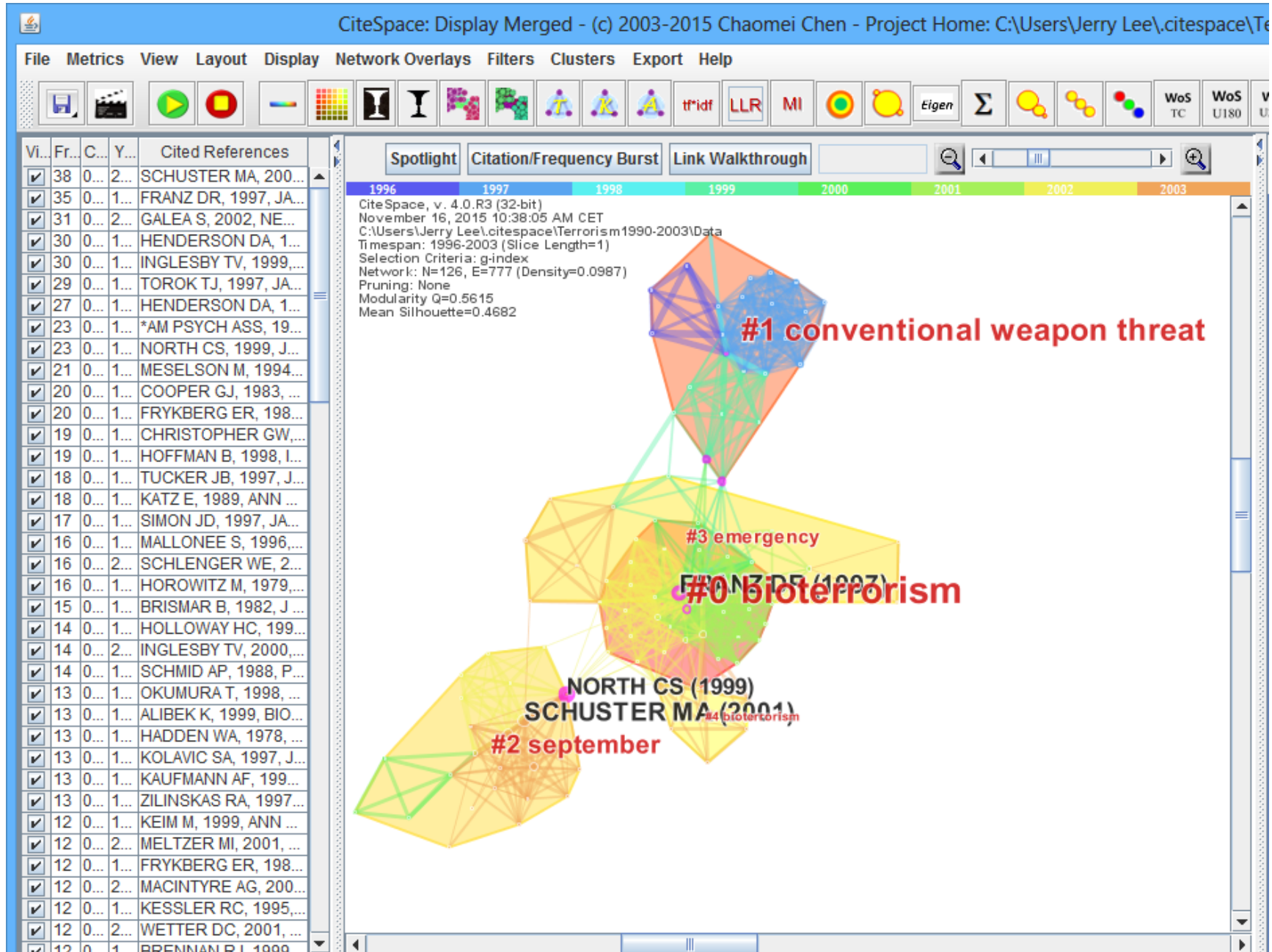
- Web of Science / PubMed:** Shows the project name 'Demo: Terrorism Research' and the data directory 'C:\Users\jli14\citespace\Examples\Data\Terrorism1990-200'. A red arrow points to the 'GO!' button.
- Time Slicing:** Set to 'From 1996' to 'To 2003' with '#Years Per Slice 1'.
- Text Processing:** Checkboxes for 'Title', 'Abstract', 'Author Keywords (DE)', and 'Keywords Plus (ID)' are all checked.
- Term Type:** Radio buttons for 'Noun Phrases', 'Burst Terms', and 'Entropy' are present.
- Network Configuration:** Includes 'Node Types' (Author, Institution, Country, Term, Keyword, Category, Journal, Paper, Grant) and 'Links' (Strength: Cosine, Scope: Within Slices).
- Space Status:** A table showing pruning configuration for 1-year slices from 1996-1996 to 2002-2002. The 'GO!' button is highlighted.
- Process Reports:** Shows 'Records in the dataset: 1513', 'Records within the chosen range: 1143', 'Valid references: 26990', 'Invalid references: 0', 'Parsing Time: 2.714 seconds', 'Total Run time: 3.494 seconds', and 'Merged network: Nodes=126, Links=777'.

A 'Your Options' dialog box is overlaid on the interface, asking 'What's your choice?' with three buttons: 'Visualize', 'Save As GraphML', and 'Cancel'. A red arrow points to the 'Visualize' button.

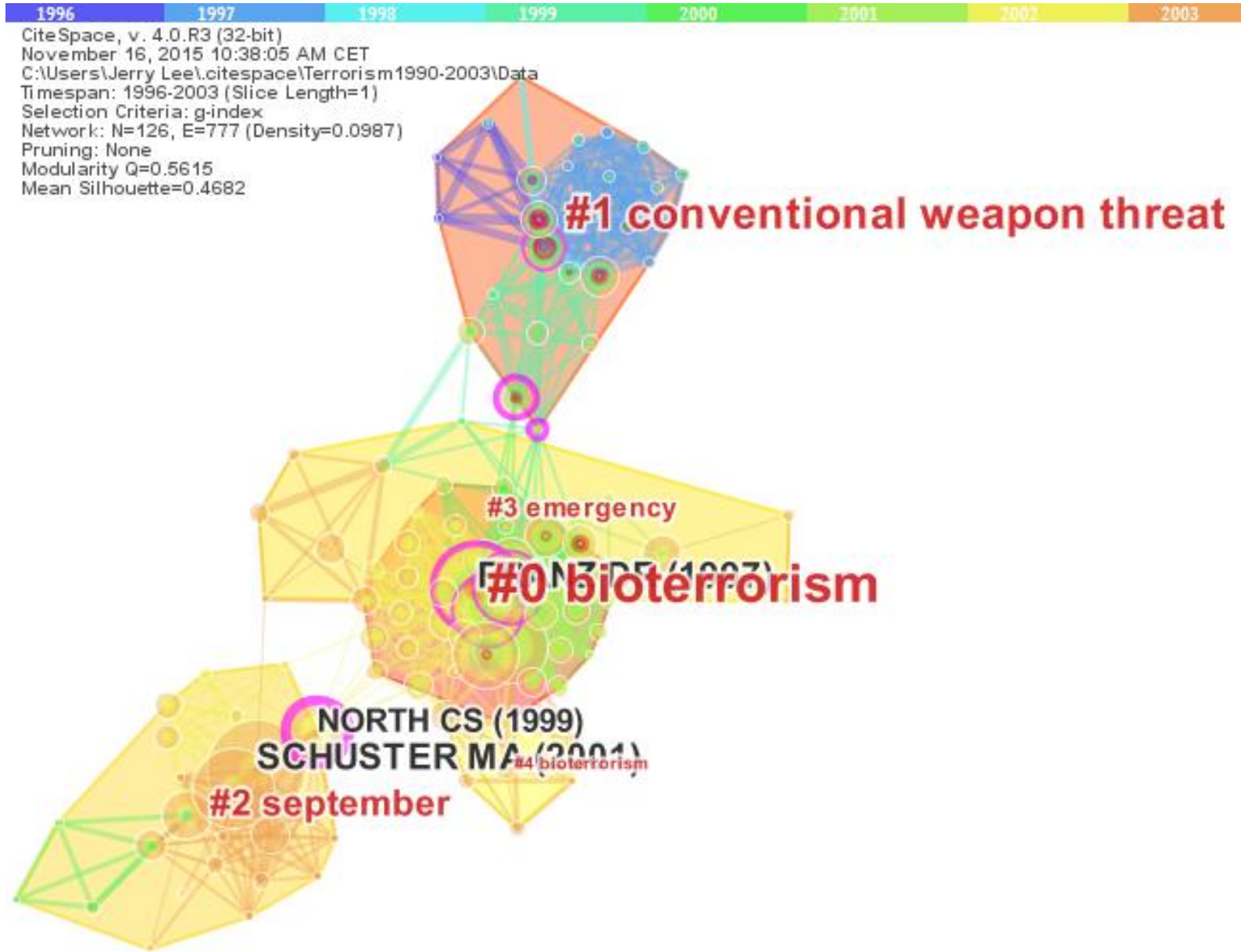
1-year slices	criteria	space	nodes	links / all
1996-1996	g=4, k=5	1033	12	18 / 18
1997-1997	g=6, k=5	1599	19	113 / 113
1998-1998	g=4, k=5	1301	13	17 / 17
1999-1999	g=6, k=5	1656	20	77 / 77
2000-2000	g=5, k=5	2301	16	40 / 40
2001-2001	g=9, k=5	3068	26	217 / 217
2002-2002	g=11, k=5	8561	35	250 / 250

1-year slices	criteria	space	nodes	links / all
1996-1996	g=4, k=5	1033	12	18 / 18
1997-1997	g=6, k=5	1599	19	113 / 113
1998-1998	g=4, k=5	1301	13	17 / 17
1999-1999	g=6, k=5	1656	20	77 / 77
2000-2000	g=5, k=5	2301	16	40 / 40
2001-2001	g=9, k=5	3068	26	217 / 217
2002-2002	g=11, k=5	8561	35	250 / 250
2003-2003	g=12, k=5	9410	34	113 / 113

Step3: 进入可视化界面, 并调整图形。

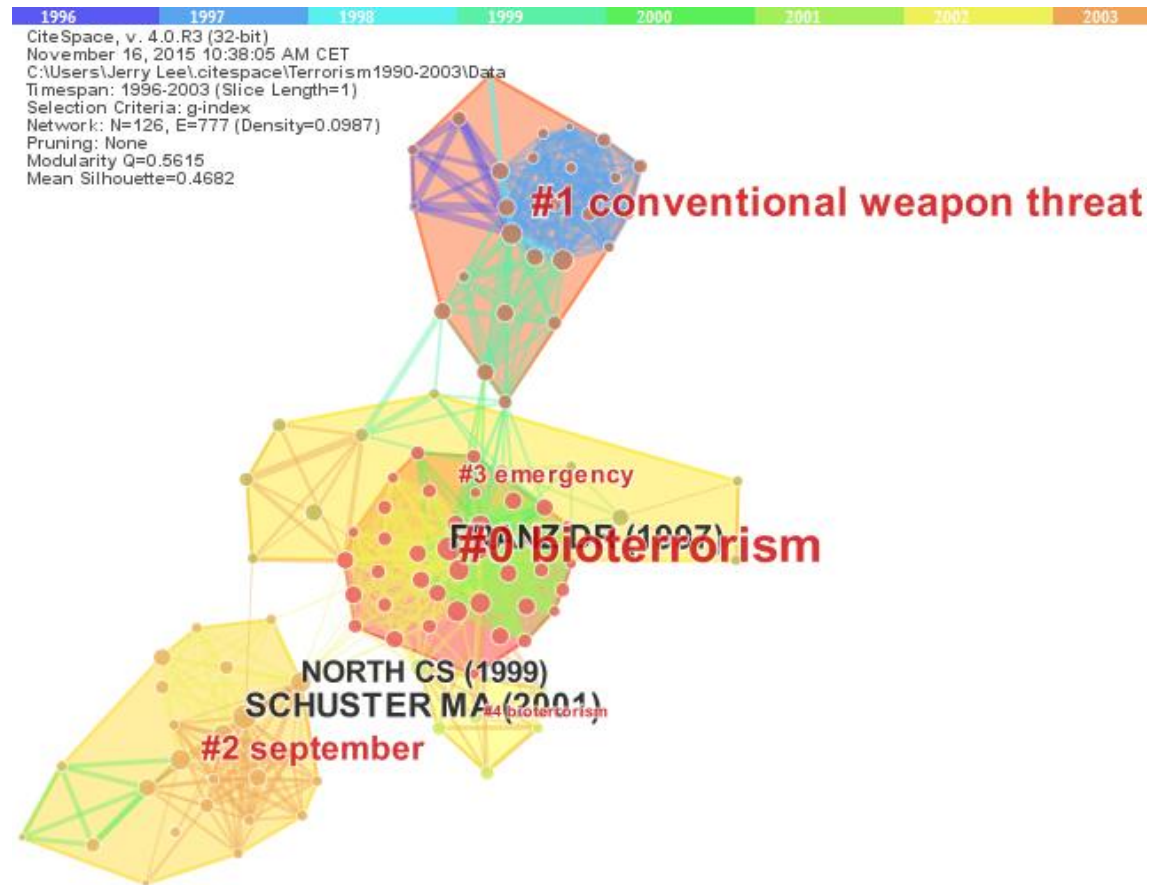
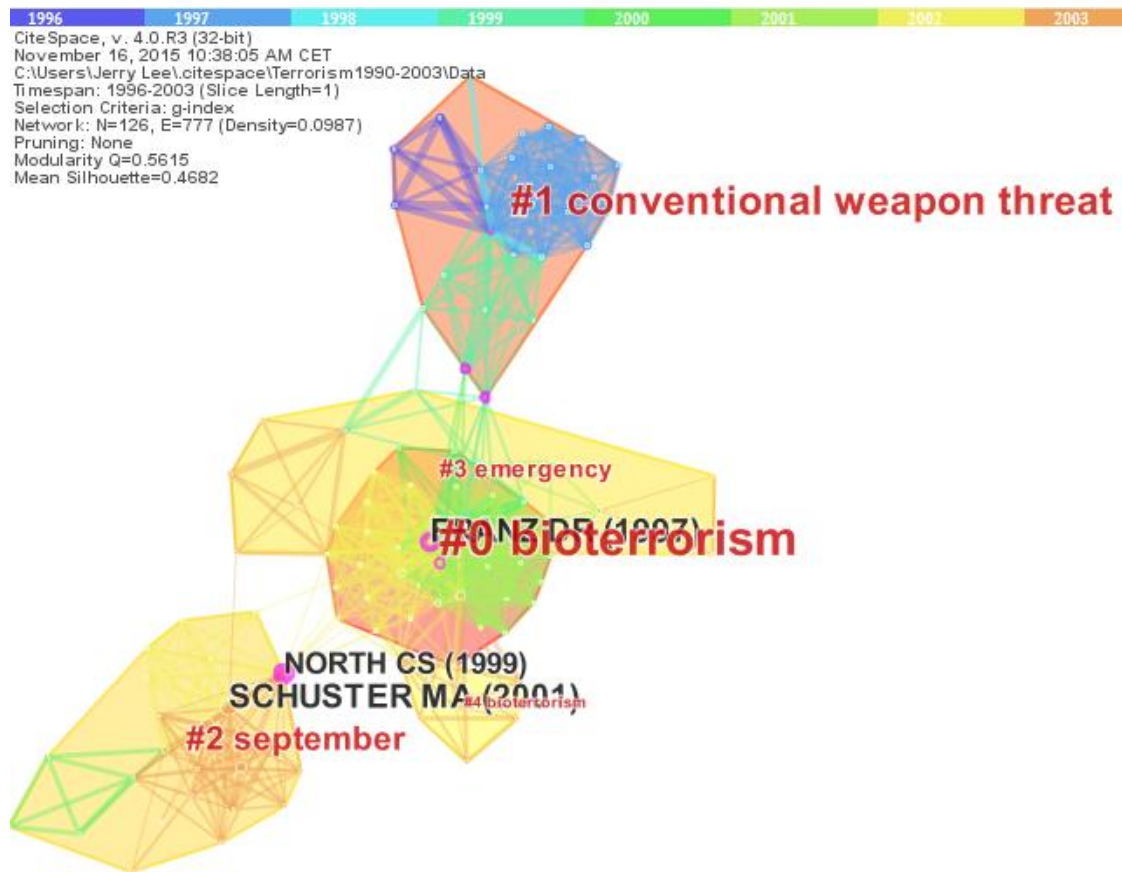


Step4: 得到最终的基于g-index进行节点筛选的结果



几种节点选择方法得到的文献共被引结果进行比较

g-index



几种节点选择方法得到的文献共被引结果进行比较

Top50

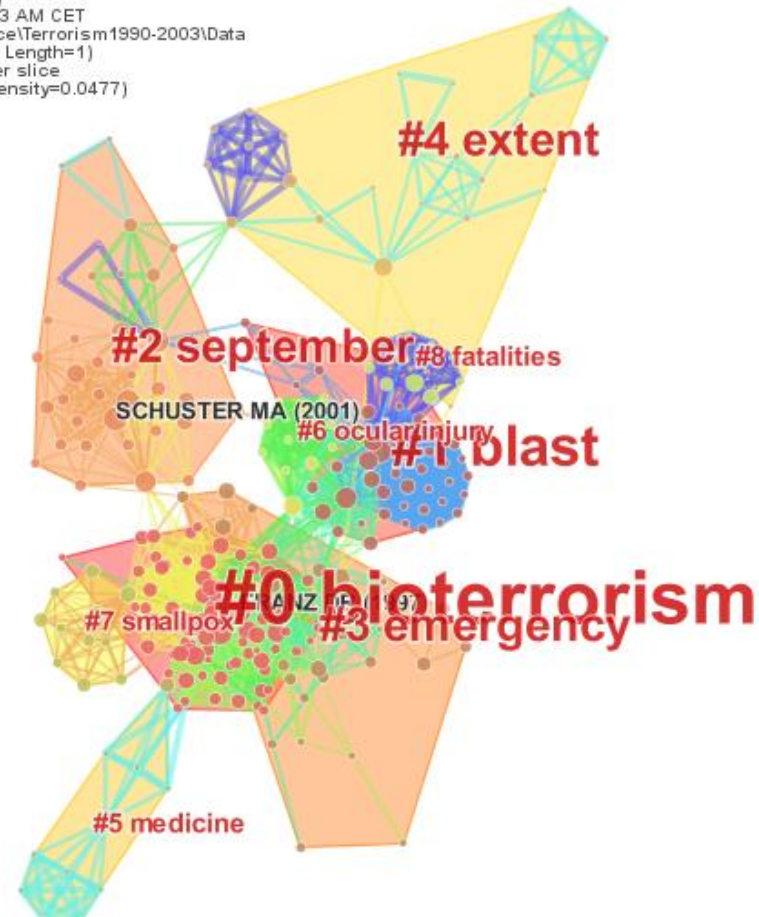
1996 1997 1998 1999 2000 2001 2002 2003

CiteSpace, v. 4.0.R3 (32-bit)
November 16, 2015 10:52:13 AM CET
C:\Users\Jerry Lee\.citespace\Terrorism1990-2003\Data
Timespan: 1996-2003 (Slice Length=1)
Selection Criteria: Top 50 per slice
Network: N=309, E=2270 (Density=0.0477)
Pruning: None
Modularity Q=0.6096
Mean Silhouette=0.5988



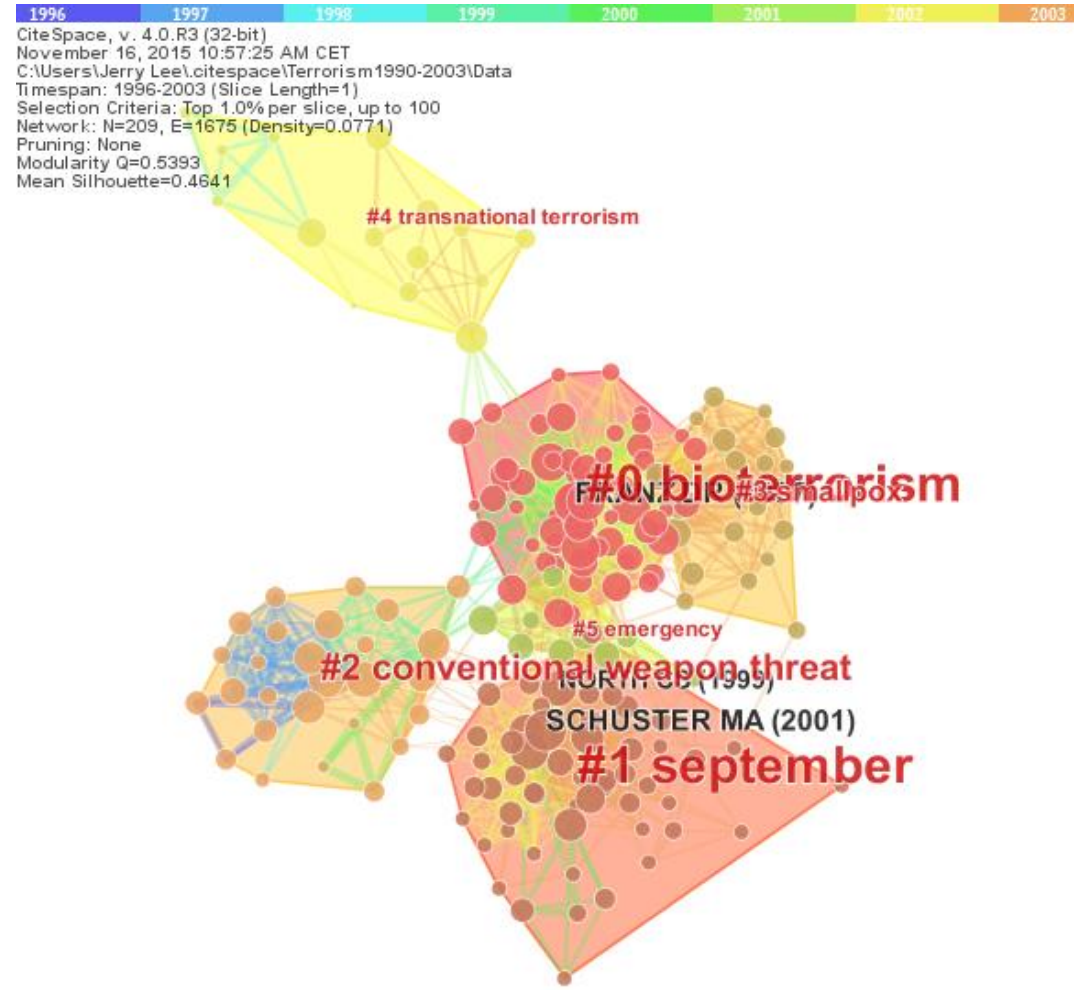
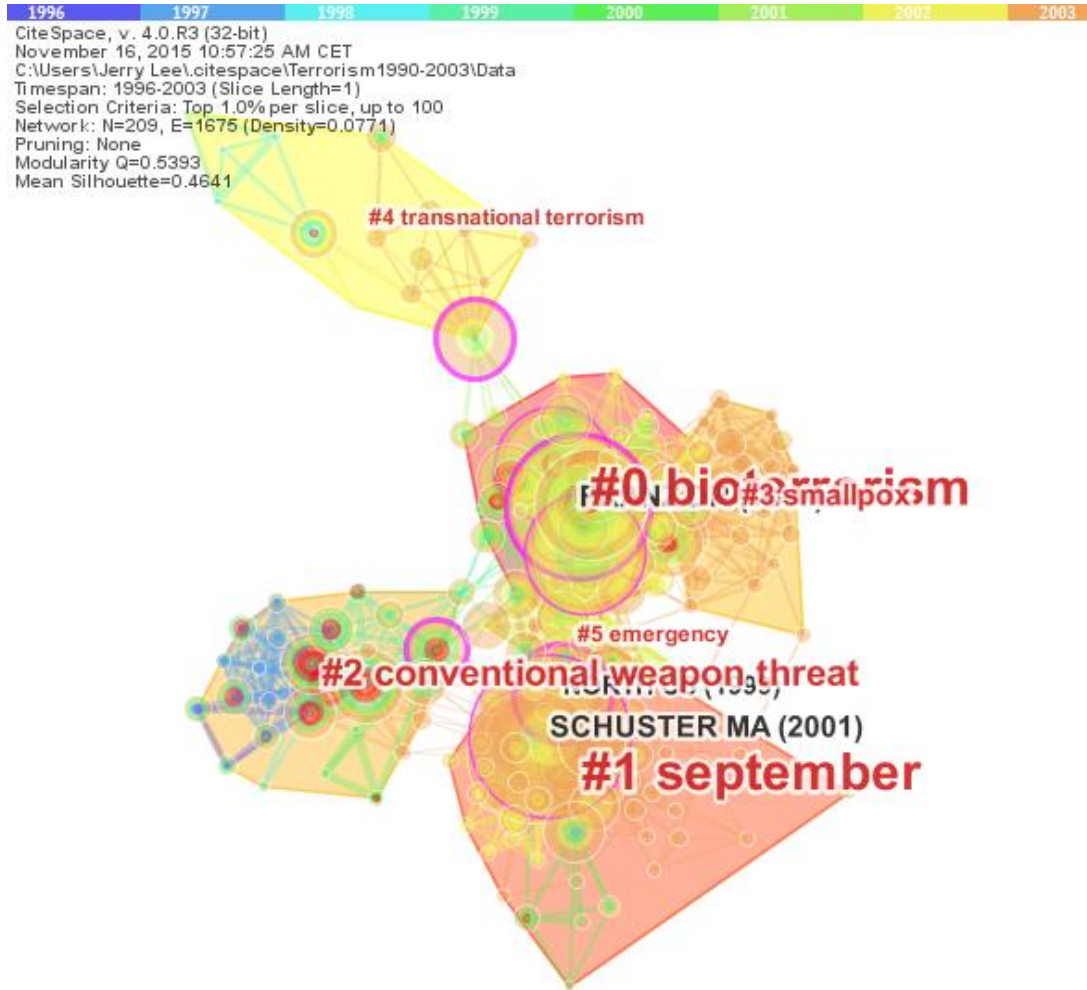
1996 1997 1998 1999 2000 2001 2002 2003

CiteSpace, v. 4.0.R3 (32-bit)
November 16, 2015 10:52:13 AM CET
C:\Users\Jerry Lee\.citespace\Terrorism1990-2003\Data
Timespan: 1996-2003 (Slice Length=1)
Selection Criteria: Top 50 per slice
Network: N=309, E=2270 (Density=0.0477)
Pruning: None
Modularity Q=0.6096
Mean Silhouette=0.5988



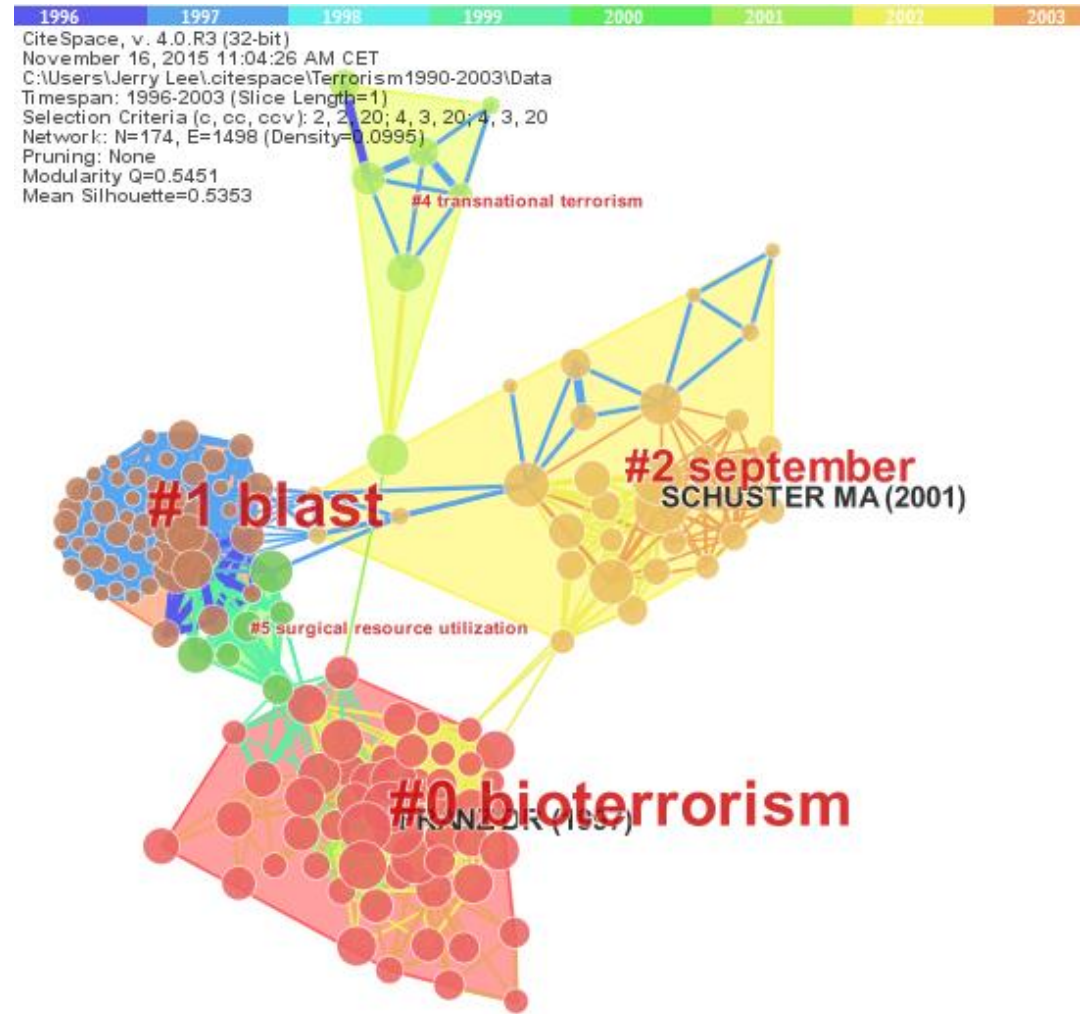
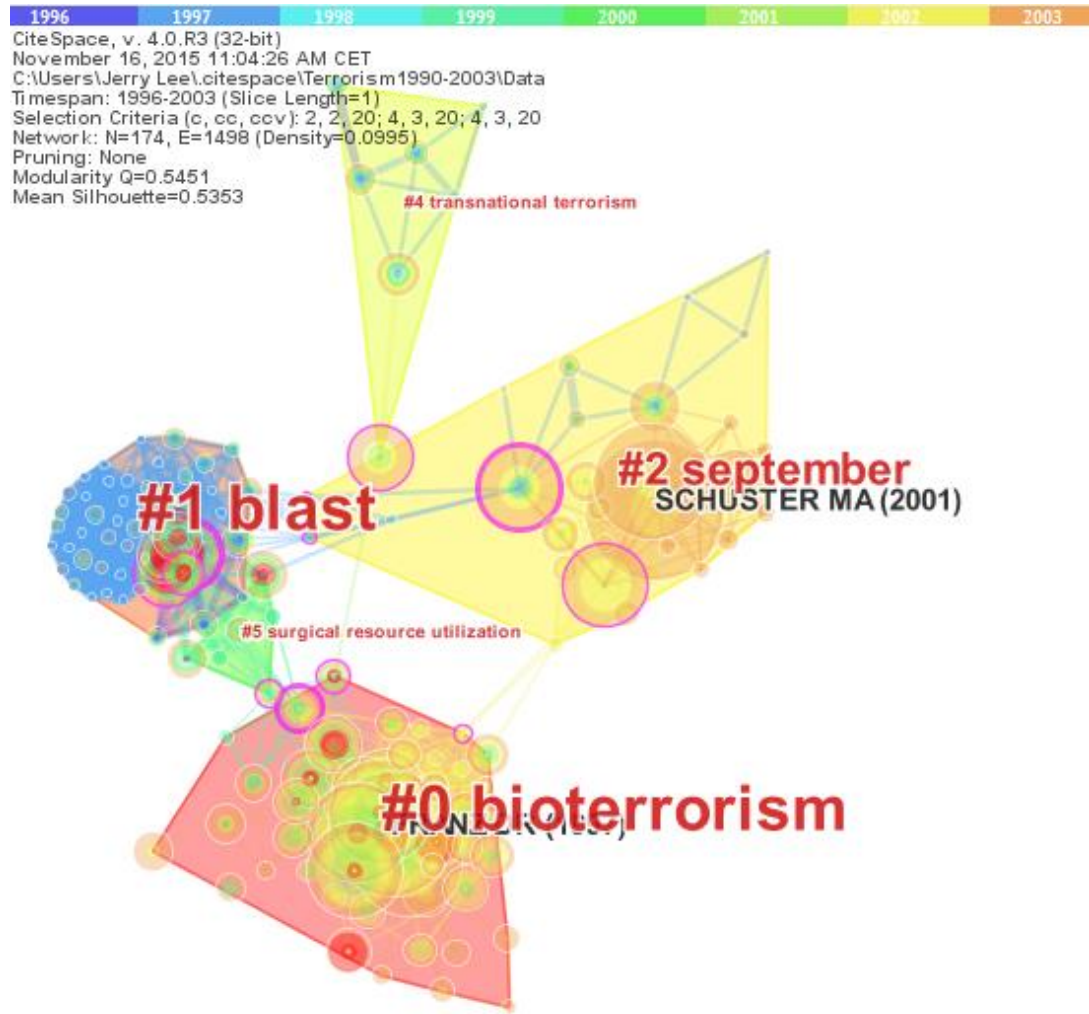
对几种节点选择方法得到的文献共被引结果进行比较

Top1%, 100



几种节点选择方法得到的文献共被引结果进行比较

Thresholds



关于Usage180和Usage300的选择请用户自行下载最新的Web of Science 核心数据集进行练习

或者参第16个PPT

<http://blog.sciencenet.cn/blog-496649-838067.html>

参考文献

1. Chen, C. (2014) The CiteSpace Manual. <https://leanpub.com/howtousecitespace>
2. 李杰, 陈超美. 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>
3. 李杰. CiteSpace中文指南. <http://blog.sciencenet.cn/blog-496649-886962.html>
4. 李杰, 陈超美. (2016) CiteSpace科技文本挖掘及可视化[M]. 首都经济贸易大学出版社.
5. 李杰,(2015)安全科学知识图谱导论[M].化学工业出版社.
6. 李杰等,(2014)安全科学技术信息检索基础[M].首都经济贸易大学出版社.
7. 陈悦, 陈超美等 (2014) 引文空间分析原理与应用[M].科学出版社.