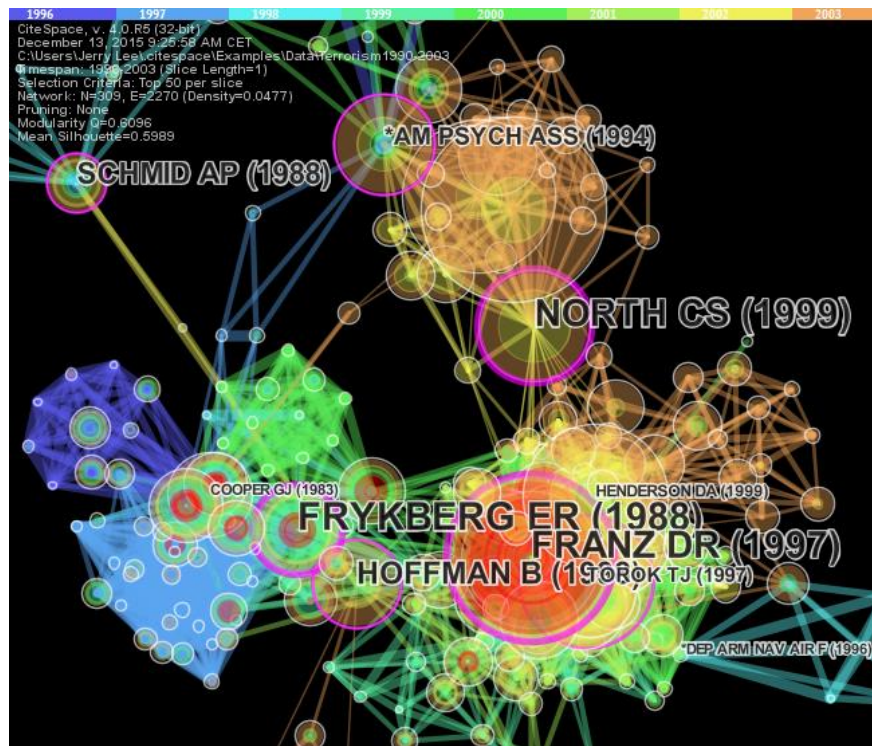


CiteSpace增加显示主要聚类功能



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12月12日

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>

说明

更新手记：4.0 R5 12/12/2015 Display: Clusters: Show/Hide Small Clusters

该功能可以自动隐藏分析得到的一些无关和小的聚类。

软件要求：CiteSpace最低版本要求为：4.0.R5.64-bit.public.12.12.2015

<http://cluster.ischool.drexel.edu/~cchen/citespace/current/4.0.R5.64-bit.public.12.12.2015.7z>

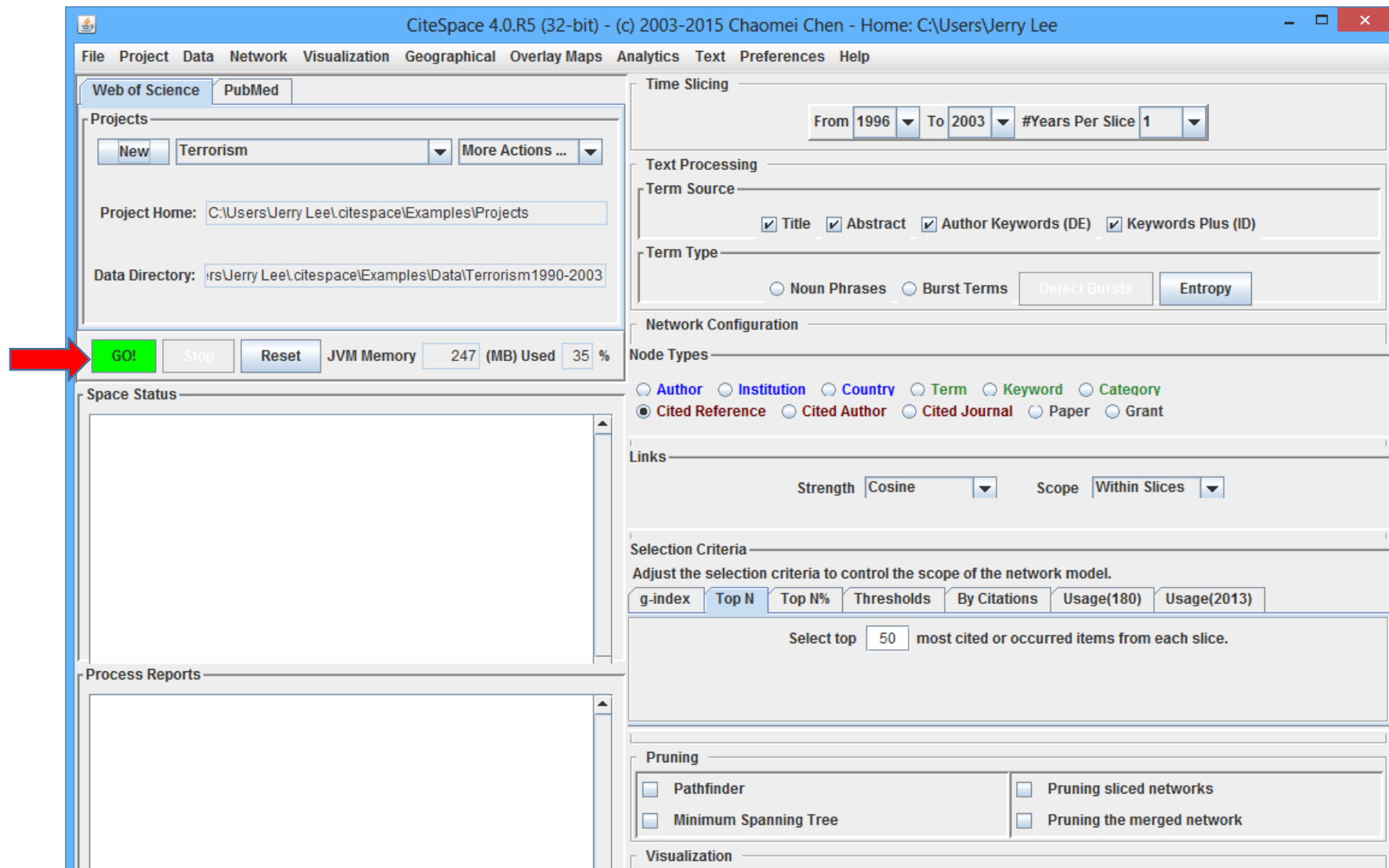


网络存在很多的游离的小类

隐藏这些小类

在可视化界面中Display→Clusters: Show/Hide Small Clusters

Step1. 运行案例数据



Step1. 运行案例数据

The screenshot displays the CiteSpace 4.0.R5 (32-bit) software interface. The main window is titled "CiteSpace 4.0.R5 (32-bit) - (c) 2003-2015 Chaomei Chen - Home: C:\Users\Jerry Lee". The interface is divided into several panels:

- Web of Science / PubMed**: The "Web of Science" tab is selected. The "Projects" section shows a dropdown menu with "Terrorism" selected. The "Project Home" is "C:\Users\Jerry Lee\citespace\Examples\Projects" and the "Data Directory" is "C:\Users\Jerry Lee\citespace\Examples\Data\Terrorism1990-2003".
- Time Slicing**: The "From" year is 1996, "To" year is 2003, and "#Years Per Slice" is 1.
- Text Processing**: The "Term Source" section has checkboxes for "Title", "Abstract", "Author Keywords (DE)", and "Keywords Plus (ID)", all of which are checked. The "Term Type" section has radio buttons for "Noun Phrases" and "Burst Terms", with "Noun Phrases" selected. There are buttons for "Detect Bursts" and "Entropy".
- Network Configuration**: The "Node Types" section has radio buttons for "Author", "Institution", "Country", "Term", "Keyword", "Category", "Journal", "Paper", and "Grant". The "Scope" is set to "Within Slices".
- Space Status**: A table showing the results of the pruning configuration. A red arrow points to the "Visualize" button in the "Your Options" dialog box.
- Process Reports**: Shows "Records in the dataset: 1513" and "Records within the chosen range: 1143". It also displays "Valid references: 26990 100.0000%" and "Invalid references: 0 0.0000%". The "Parsing Time" is 7.881 seconds.

The "Your Options" dialog box is a blue window with a red "X" button. It contains a green question mark icon and the text "What's your choice?". Below the text are three buttons: "Visualize", "Save As GraphML", and "Cancel".

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	8561	50
2003-2003	top 50	9410	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	8561	50
2003-2003	top 50	9410	50

Adjust the selection criteria to control the scope of the network model.

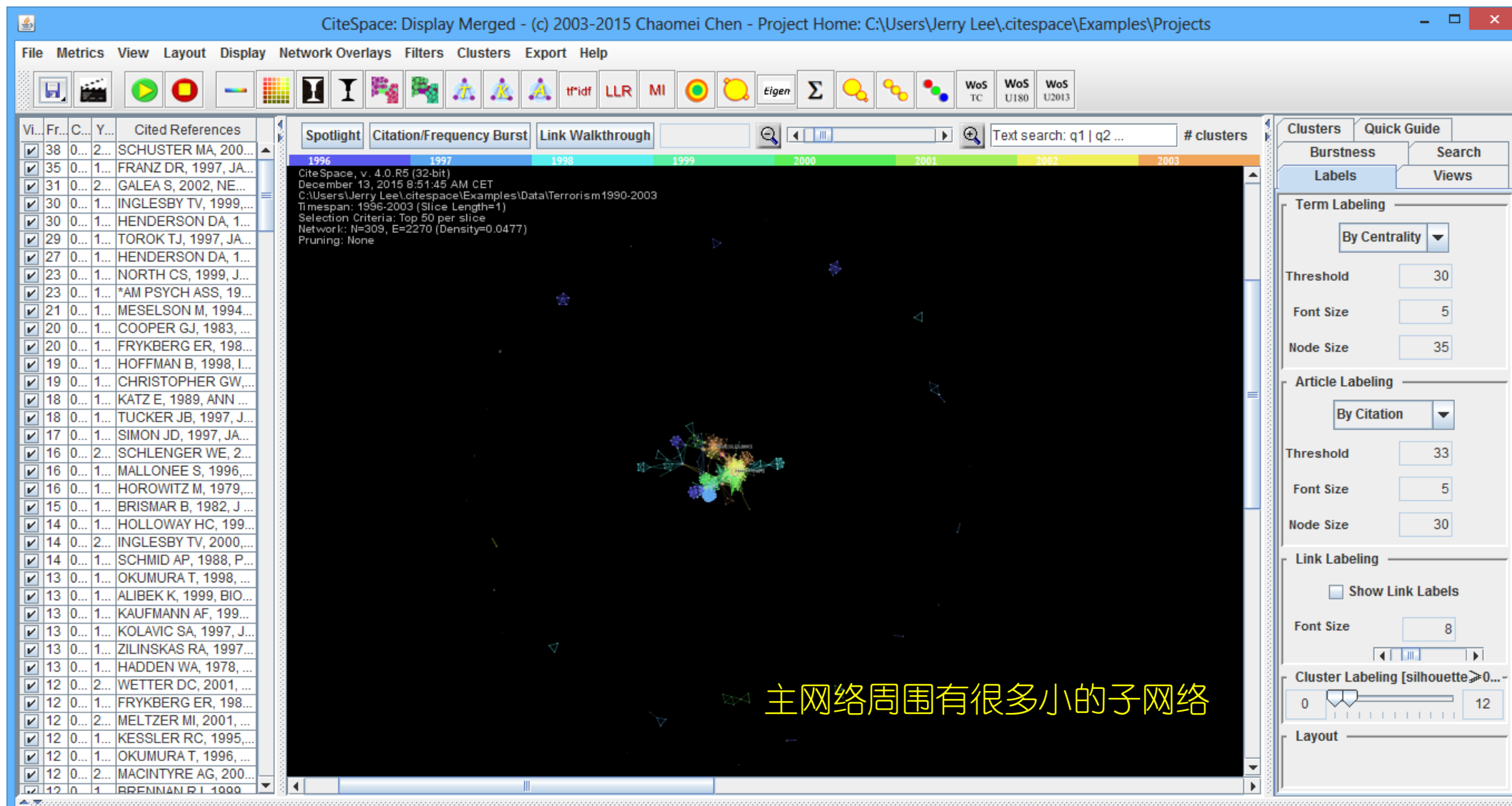
g-index Top N Top N% Thresholds By Citations Usage(180) Usage(2013)

Select top 50 most cited or occurred items from each slice.

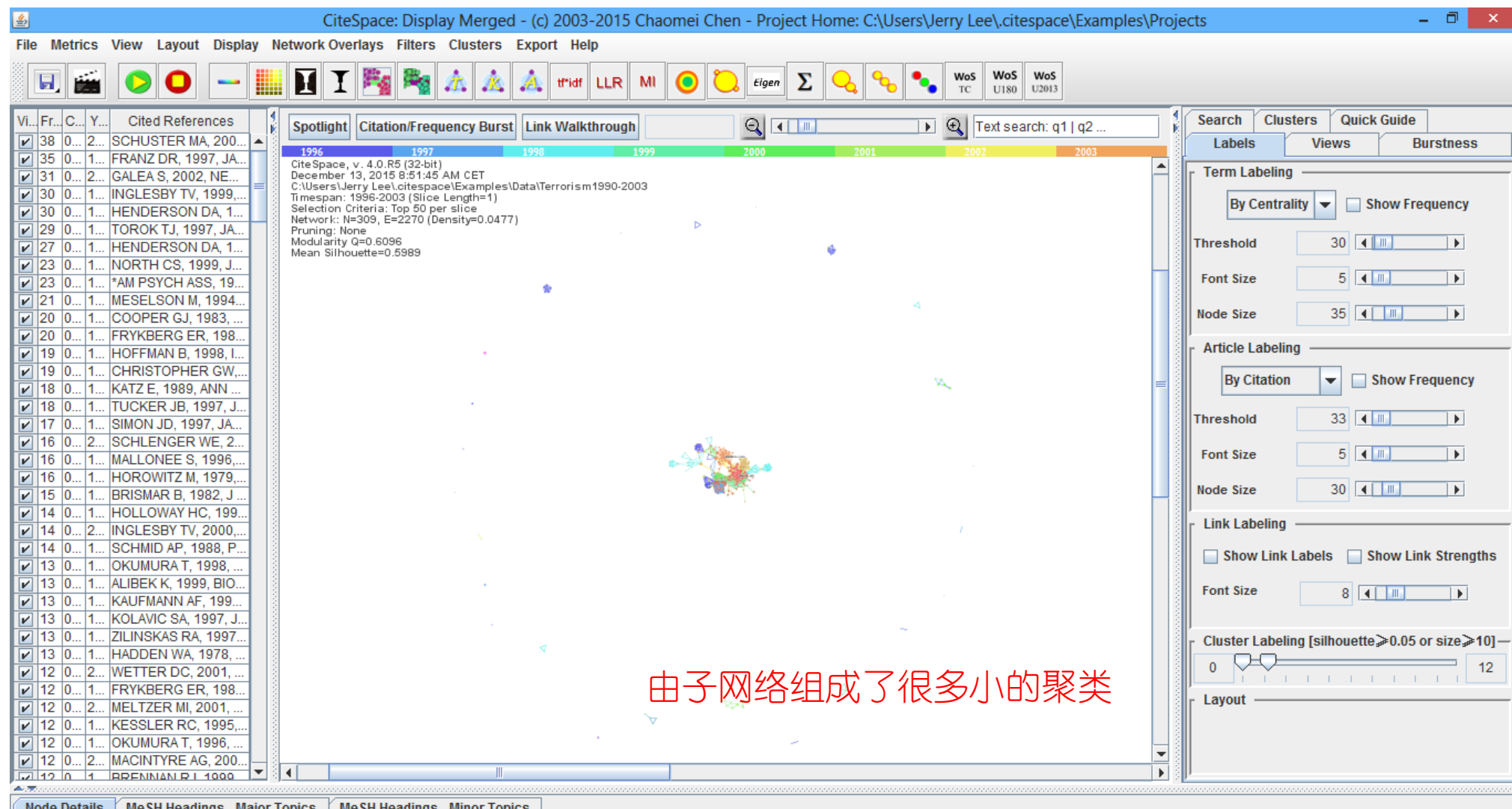
Pruning

- ☐ Pathfinder
- ☐ Minimum Spanning Tree
- ☐ Pruning sliced networks
- ☐ Pruning the merged network

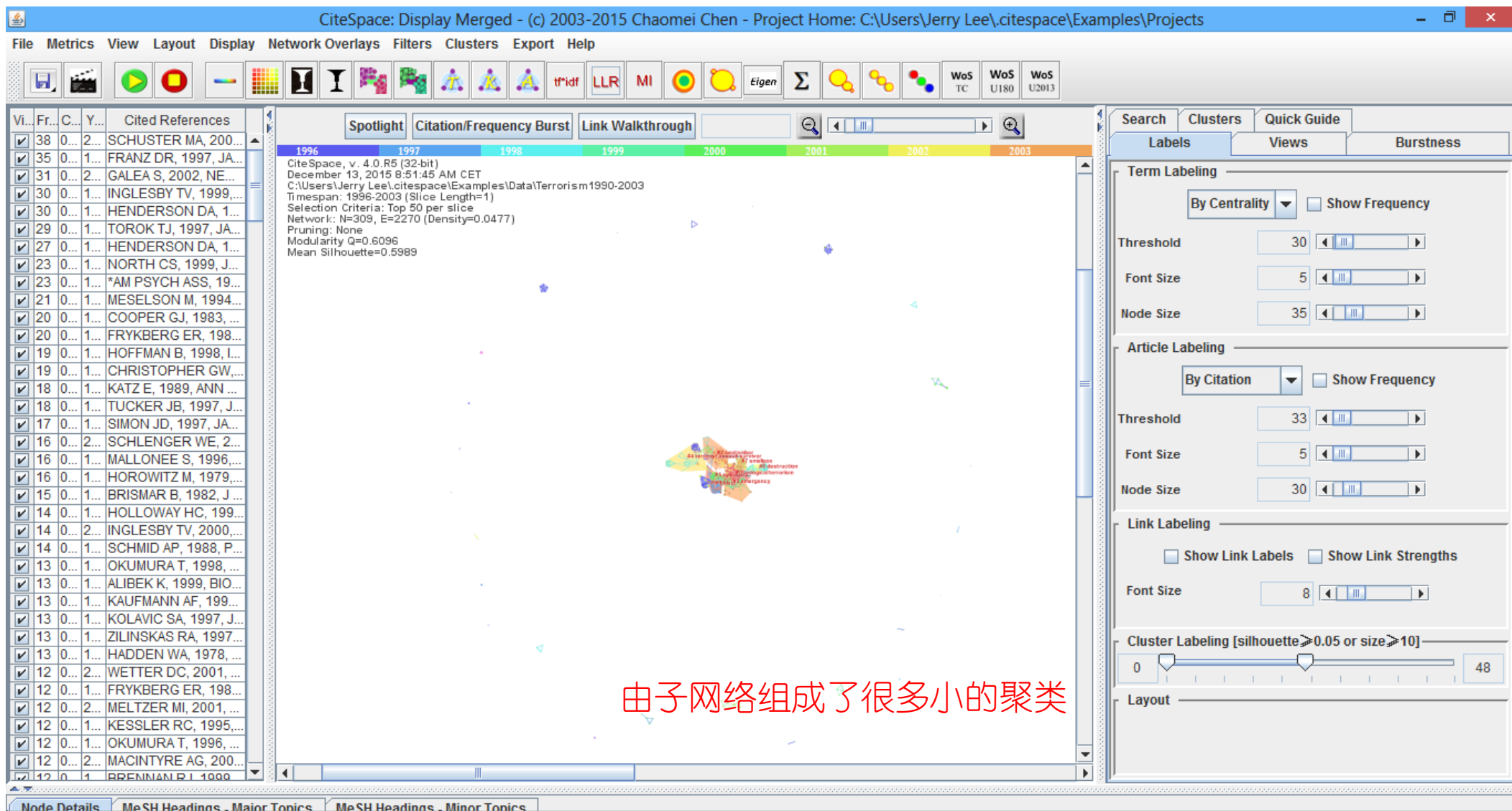
Step2. 得到文献共被引网络



Step3.对网络进行聚类



Step4.对聚类网络进行命名



Step5.隐藏很小的聚类， 仅仅展示主要聚类

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

File Metrics View Layout Display Network Overlays Filters Clusters Export Help

Background Color Ctrl-B
Black Background Ctrl-NumPad-2
White Background Ctrl-F8
Colormap Palette
Node Rendering Type
Node Shape
Node Size
Node Fill Color Ctrl-8
Node Outline Color Ctrl-9
Label Font Size
Label Color
Line Shape
Dashed Lines: Show/Hide Ctrl-End
Dashed Lines: Set Color Ctrl-Unknown keyCode: 0x2b
Solid Lines: Show/Hide Ctrl-Unknown keyCode: 0x40
Solid Lines: Set Color Ctrl-Minus
Overlay Labels: Show/Hide Ctrl-Page Up
Link Labels: Show/Hide
Link Strengths: Show/Hide
Link Transparency (0.0 - 1.0)
Show/Hide Signature
Show/Hide Citation Frequency
Clusters
Timeline View

Link Walkthrough
1999 2000 2001 2002 2003
1990-2003

默认是显示小的聚类的，我们在这里可以隐藏这些小的聚类。

Show/Hide Cluster Labels
Show/Hide Cluster IDs
Convex Hull: Show/Hide
Convex Hull: Fill/Border Only
Convex Hull: Toggle Fill Color Patterns
Convex Hull: Select a Fill Color
Circle: Show/Hide
Circle Packing
Show/Hide Small Clusters
Show/Hide Small Clusters in view

Search Clusters Quick Guide
Labels Views Burstness

Term Labeling
By Centrality Show Frequency
Threshold 30
Font Size 5
Node Size 35

Article Labeling
By Citation Show Frequency
Threshold 33
Font Size 5
Node Size 30

Link Labeling
Show Link Labels Show Link Strengths
Font Size 8

Cluster Labeling [silhouette ≥ 0.05 or size ≥ 10]
0 48

Layout

Step5.隐藏很小的聚类， 仅仅展示主要聚类

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

File Metrics View Layout Display Network Overlays Filters Clusters Export Help

Spotlight Citation/Frequency Burst Link Walkthrough

1996 1997 1998 1999 2000 2001 2002 2003

CiteSpace, v. 4.0.R5 (32-bit)
December 13, 2015 8:51:45 AM CET
C:\Users\Jerry Lee\citespace\Examples\Data\Terrorism1990-2003
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=0.5989

Vi... Fr... C... Y... Cited References

38	0	2	SCHUSTER MA, 200...
35	0	1	FRANZ DR, 1997, JA...
31	0	2	GALEA S, 2002, NE...
30	0	1	INGLESBY TV, 1999, ...
30	0	1	HENDERSON DA, 1...
29	0	1	TOROK TJ, 1997, JA...
27	0	1	HENDERSON DA, 1...
23	0	1	NORTH CS, 1999, J...
23	0	1	*AM PSYCH ASS, 19...
21	0	1	MESELSON M, 1994...
20	0	1	COOPER GJ, 1983, ...
20	0	1	FRYKBERG ER, 198...
19	0	1	HOFFMAN B, 1998, I...
19	0	1	CHRISTOPHER GW, ...
18	0	1	KATZ E, 1989, ANN ...
18	0	1	TUCKER JB, 1997, J...
17	0	1	SIMON JD, 1997, JA...
16	0	2	SCHLENGER WE, 2...
16	0	1	MALLONEE S, 1996, ...
16	0	1	HOROWITZ M, 1979, ...
15	0	1	BRISMA B, 1982, J ...
14	0	1	HOLLOWAY HC, 199...
14	0	2	INGLESBY TV, 2000, ...
14	0	1	SCHMID AP, 1988, P...
13	0	1	OKUMURA T, 1998, ...
13	0	1	ALIBEK K, 1999, BIO...
13	0	1	KAUFMANN AF, 199...
13	0	1	KOLAVIC SA, 1997, J...
13	0	1	ZILINSKAS RA, 1997...
13	0	1	HADDEN WA, 1978, ...
12	0	2	WETTER DC, 2001, ...
12	0	1	FRYKBERG ER, 198...
12	0	2	MELTZER MI, 2001, ...
12	0	1	KESSLER RC, 1995, ...
12	0	1	OKUMURA T, 1996, ...
12	0	2	MACINTYRE AG, 200...
12	0	1	BRENNAN R, 1999, ...

Term Labeling

By Centrality ☐ Show Frequency

Threshold 30

Font Size 5

Node Size 35

Article Labeling

By Citation ☐ Show Frequency

Threshold 33

Font Size 5

Node Size 30

Link Labeling

☐ Show Link Labels ☐ Show Link Strengths

Font Size 8

Cluster Labeling [silhouette >= 0.05 or size >= 10]

0 48

Layout

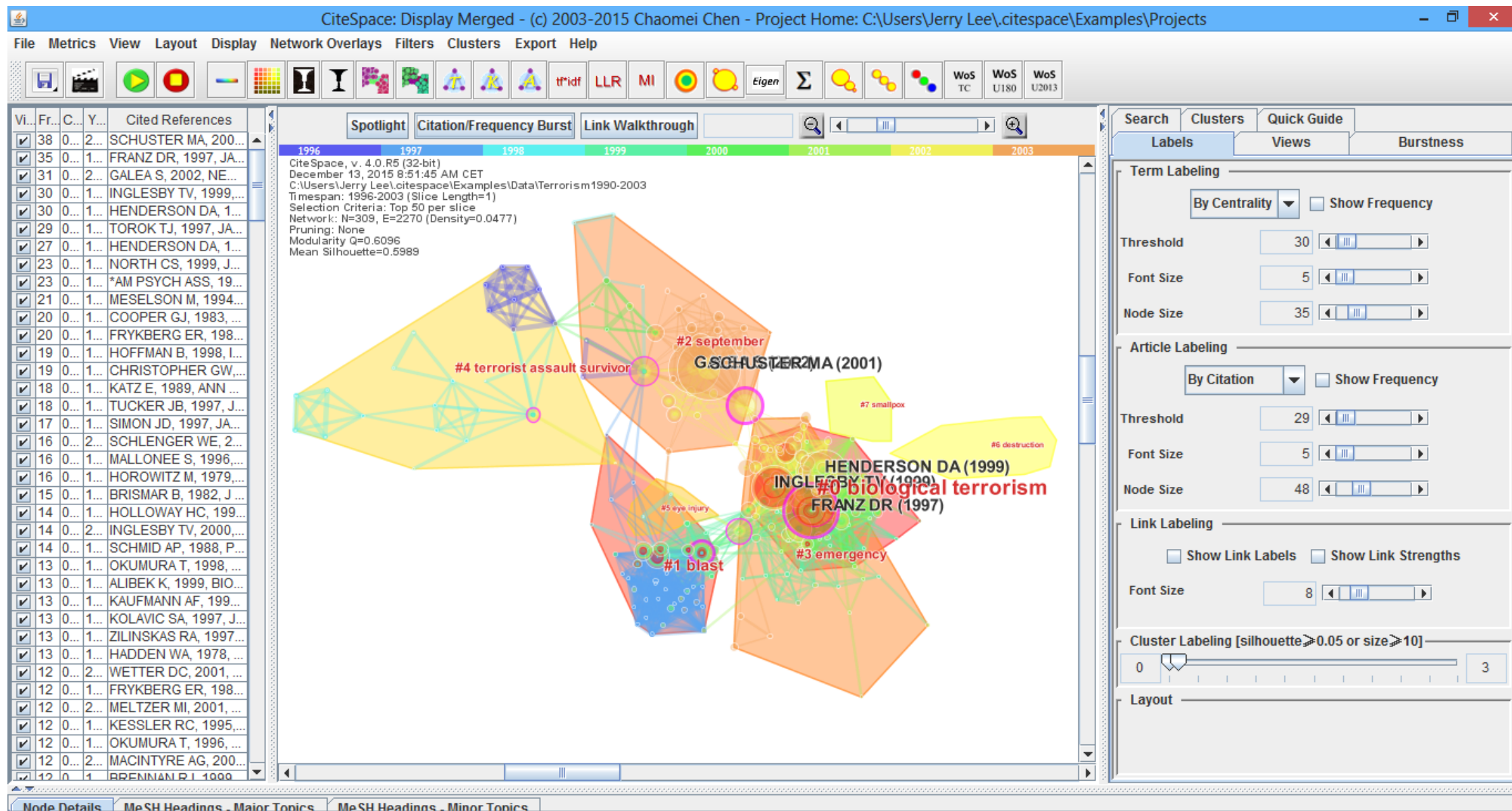
Node Details MeSH Headings - Major Topics MeSH Headings - Minor Topics

这里仅仅显示了主要网络的聚类
下面我们需要对得到的可视化图形进行调整。

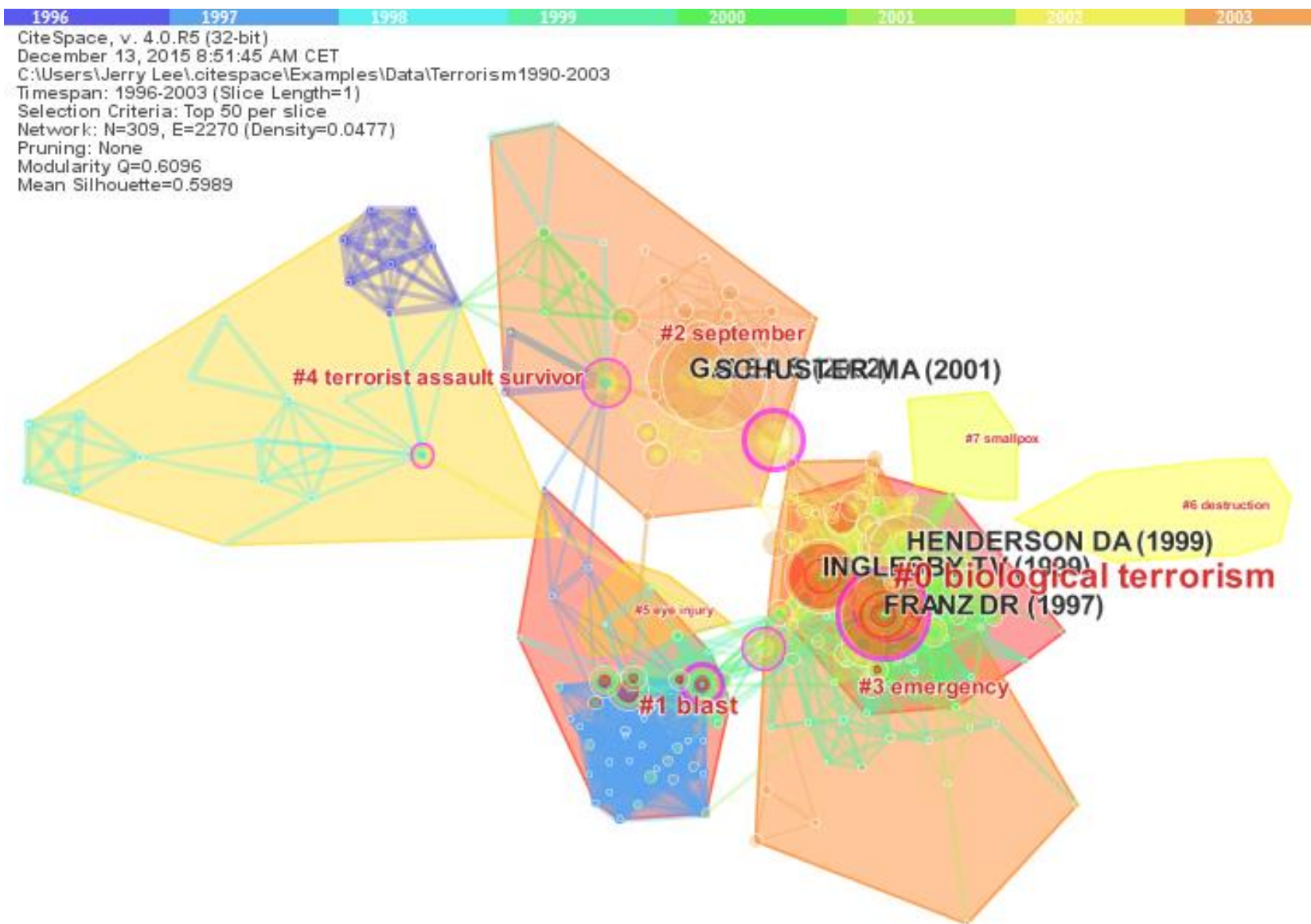
提示

在该版本中，用于可以使用鼠标左键了
整体移动图形在画布上的位置
Display: drag with the left mouse button to
pane the graphics

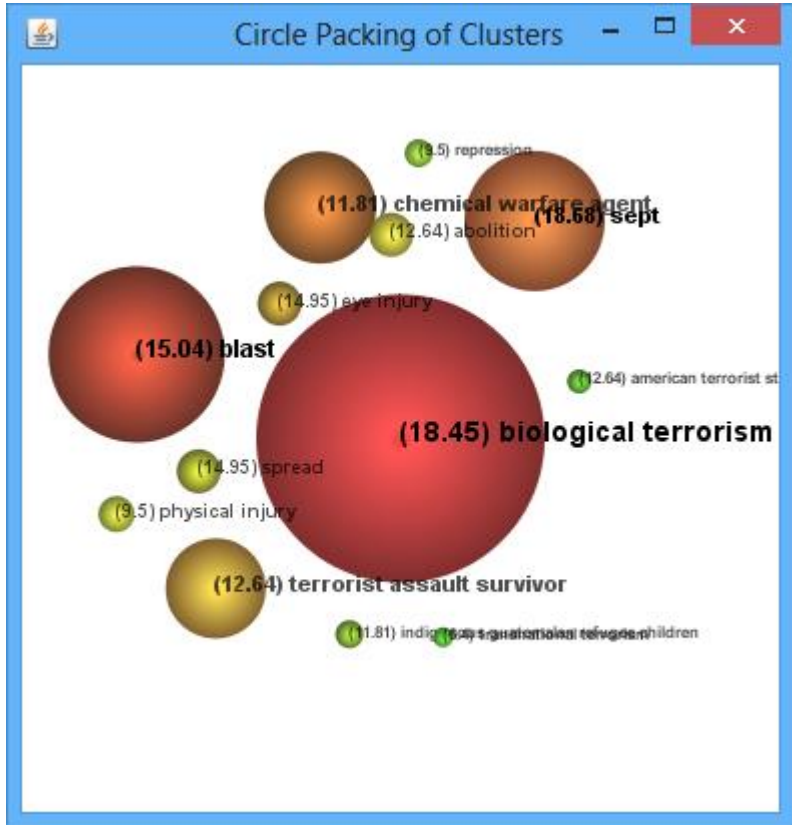
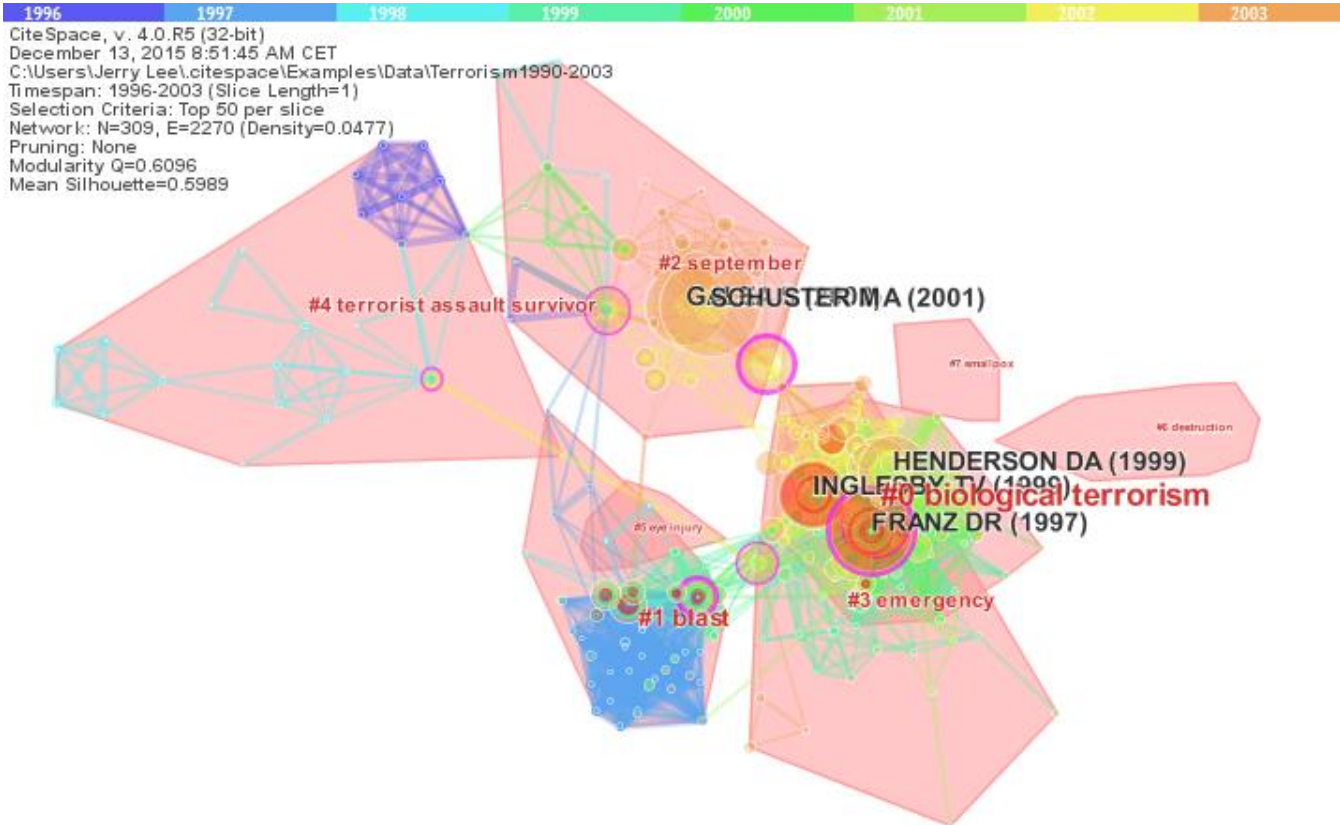
Step6.图形的放大和左右调整



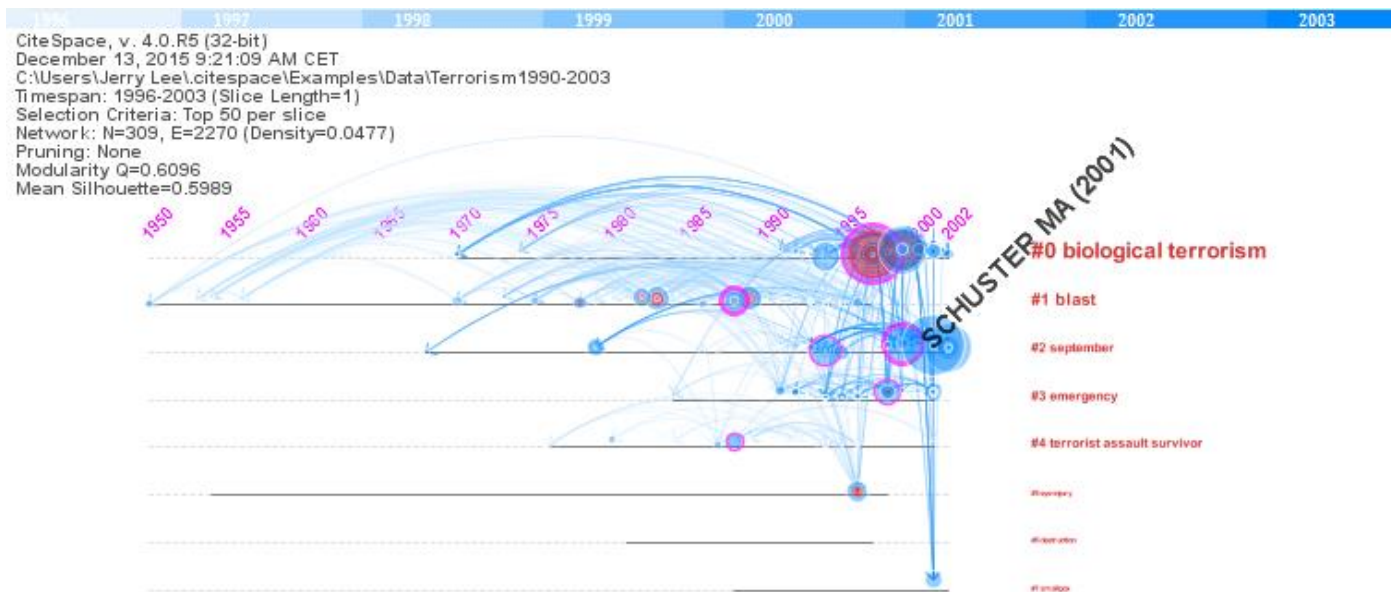
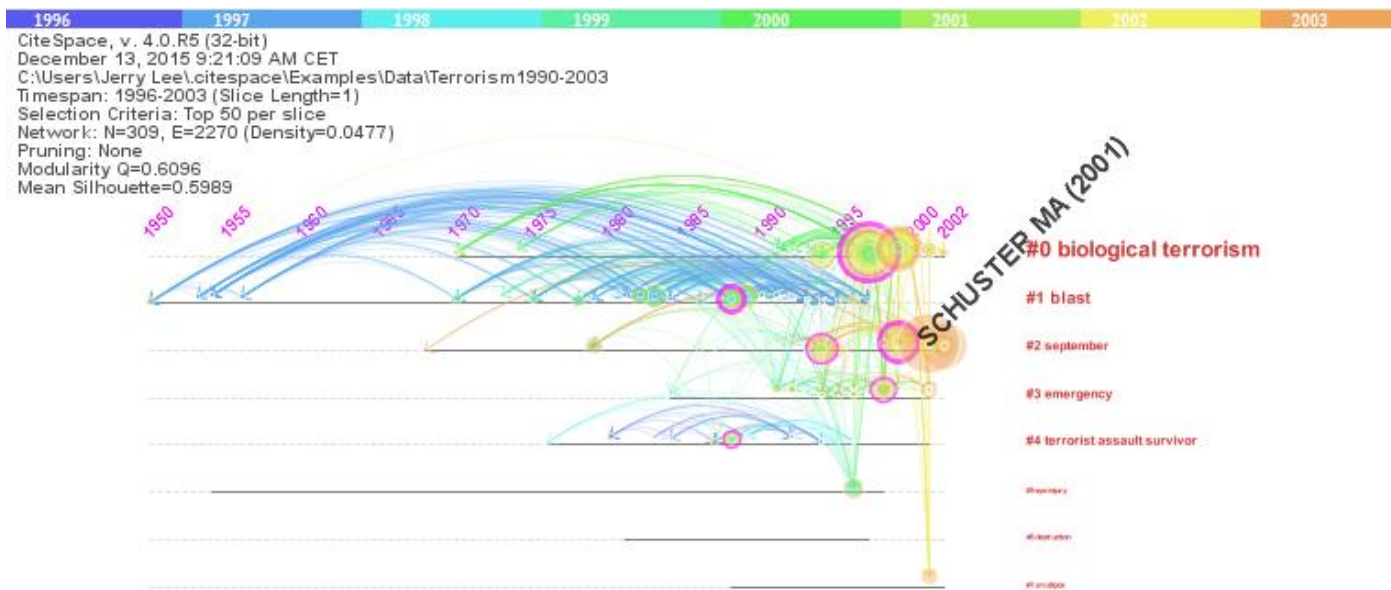
Step7.结果的保存



Step7.结果的保存



Step7.结果的保存



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.