

Workshop Visualizing and Analyzing Scientific
Literature with *CiteSpace*

4

Chaomei Chen
College of Information Science and Technology
Drexel University
Email: chaomei.chen@cis.drexel.edu



Configuring CiteSpace
Time Slicing
Network Modeling and Pruning
Test Drive
Questions and Discussions

CREATE A PROJECT

CiteSpace 2.2.R1 - (c) 2003-2009 Chaomei Chen ~ Home: C:\Users\IBM

File Data Project Metrics Text Network Analyze Geospatial Maps Test Preferences Gallery Help

Web of Science PubMed CiteSeer

Projects: New Terrorism Edit

Project Create a new project, assign a project home and the data directory

Data Directory: C:\Users\IBM\citespace\Examples\Data\Terrorism1990-2003

GO! Stop Reset JVM Memory (K) 174080 Used (%) 37

- Space Status

1992-1992	top 30	511	30	121
1993-1993	top 30	818	30	48

New Project

Title ISSI 2009 Workshop

Project Home C:\Users\IBM Browse

Data Directory C:\Users\IBM Browse

Language English Chinese

Enable Alias List on

Enable Export Space on

Enable Export Matrices off

Save Merged Slice off

Noun Phrase: Maximum Words 4

Maximum GML Node Label Length 16

Enable Exclusion List on

Enable Export Abstracts (* Expensive) on

Enable JDIC on

Noun Phrase: Minimum Words 2

Burst Term Threshold 0.1

CTSA (1-Disciplines, 2-Sciences) 1

Save Cancel

Run time: 3978 milliseconds.

Merged network: Nodes=315, Links=1442

Exclusion List: 0

Network modeling ends at Sat Mar 07 16:48:03 EST 2009.

Time Slicing
From 1990 To 2003 #Years Per Slice 1 Entropy

Modeling

Node Types

Author Term Keyword Category Institution Country

Cited Reference Cited Author Cited Journal

Term Sources

Title Abstract Descriptors Identifiers

Minimum Spanning Tree Pruning the merged network

Visualization

Cluster View (Default) Show Networks by Time Slices

Time-zone View Show Merged Network

CiteSpace 2.2.R1 - (c) 2003-2009 Chaomei Chen ~ Home: C:\Users\IBM

File Data Project Metrics Text Network Analyze Geospatial Maps Test Preferences Gallery Help

Web of Science PubMed CiteSeer

Projects

New Complex Network Analysis 1980-2... More Actions ...

Project Home: C:\Users\IBM\...\Data\Social Networks 1980-2008\project

Data Directory: C:\Users\IBM\...\Data\Social Networks 1980-2008\data

GO! Stop Reset JVM Memory (K) 1/4080 Used (%) 37

Space Status

Process Reports

Time Slicing

From 1995 To 2008 #Years Per Slice 1 Entropy

Modeling

Node Types

Author Term Keyword Category Institution Country

Cited Reference Cited Author Cited Journal

Term Sources

Title Abstract Descriptors Identifiers

Term Selection

None Noun Phrases Burst Phrases [Find Burst Phrases](#)

Links

Strength Cosine Scope Within Slices

Top N per slice Top N% per slice Threshold Interpolation Select Citers

Use TC Filter 4 - 3460 Check TC Distribution Continue

0	2529	23.0601
1	1131	33.3728
2	847	41.096
3	622	46.7676
4	520	51.5091

Pruning

Pathfinder Pruning sliced networks

Minimum Spanning Tree Pruning the merged network

Visualization

Cluster View (Default) Show Networks by Time Slices

Time-zone View Show Merged Network

The screenshot shows the CiteSpace 2.2.R1 software interface. The main window has a menu bar with File, Data, Project, Metrics, Text, Network, Analyze, Geospatial Maps, Test, Preferences, Gallery, and Help. Below the menu is a tab bar with Web of Science, PubMed, and CiteSeer, where Web of Science is selected. The left side of the interface contains sections for Projects, Space Status, and Process Reports. The right side contains several configuration panels: Time Slicing (set to From 1995 to 2008, #Years Per Slice 1), Modeling (Node Types: Author, Term, Keyword, Category, Institution, Country; Cited Reference selected), Term Sources (Title, Abstract, Descriptors, Identifiers checked), Term Selection (None selected), Links (Strength: Cosine, Scope: Within Slices), Threshold Interpolation (Use TC Filter checked, value 4, 3460, Check TC Distribution button circled in red), Select Citers, Pruning (Pathfinder, Minimum Spanning Tree), and Visualization (Cluster View Default selected, Show Networks by Time Slices, Show Merged Network). The 'Check TC Distribution' button in the Threshold Interpolation panel is highlighted with a red oval.

CiteSpace 2.2.R1 - (c) 2003-2009 Chaomei Chen ~ Home: C:\Users\IBM - Collaborating Countries

File Data Project Metrics Text Network Analyze Geospatial Maps Test Preferences Gallery Help

Web of Science PubMed CiteSeer

Projects

New Complex Network Analysis 1980-2... More Actions ...

Project Home: C:\Users\IBM\...\Social Networks 1980-2008\project

Data Directory: C:\Users\IBM\...\Social Networks 1980-2008\data

GO! Stop Reset JVM Memory (K) 174080 Used (%) 37

Space Status

Process Reports

Time Slicing

From 1995 To 2008 #Years Per Slice 1 Entropy

Modeling

Node Types

Author Term Keyword Category Institution Country

Cited Reference Cited Author Cited Journal

Term Sources

Title Abstract Descriptors Identifiers

Term Selection

None Noun Phrases Burst Phrases

Links

Strength Cosine Scope Within Slices

Top N per slice Top N% per slice Threshold Interpolation Select Citers

Select top 30 most cited or occurred items from each slice.

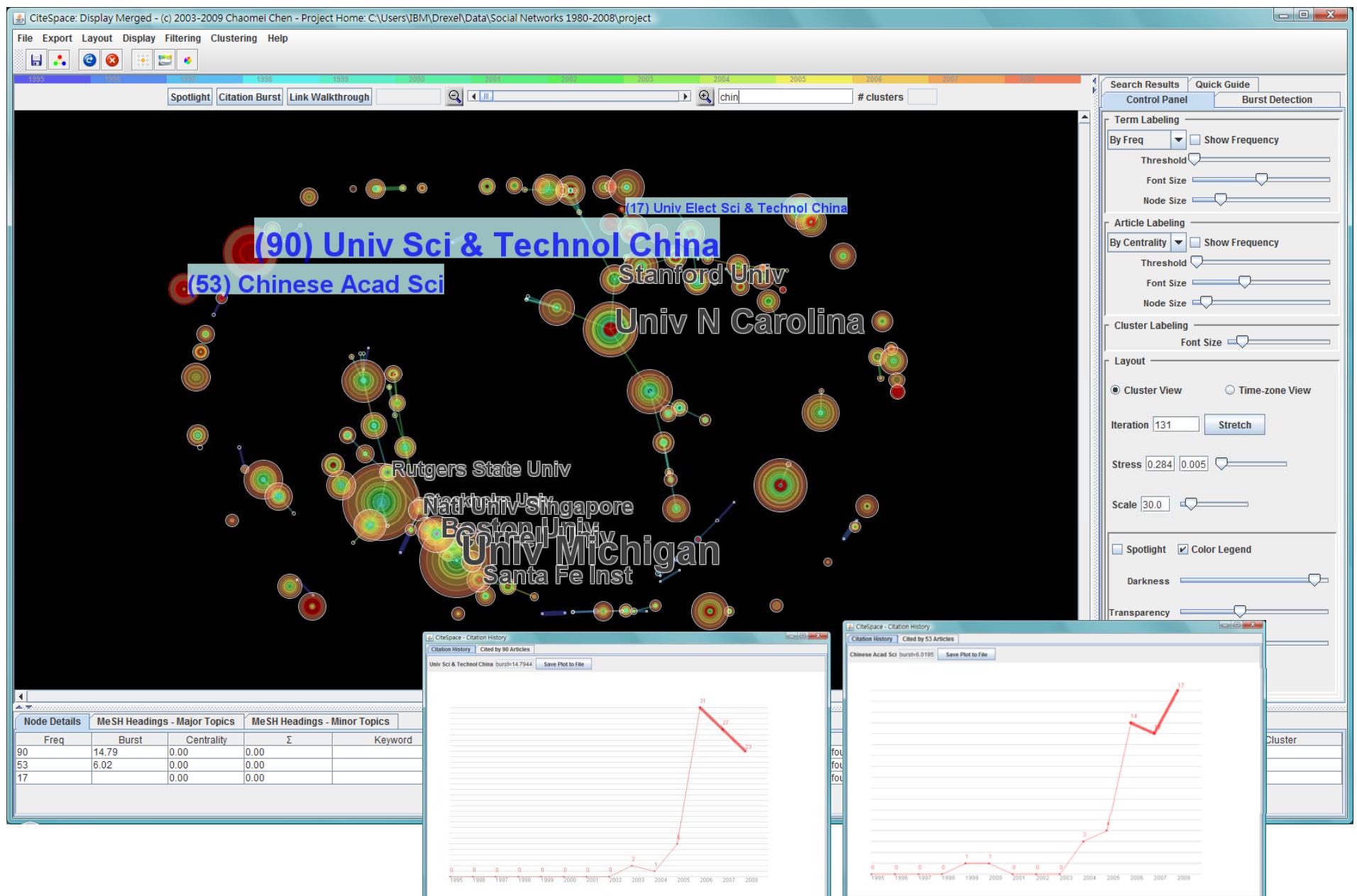
Pruning

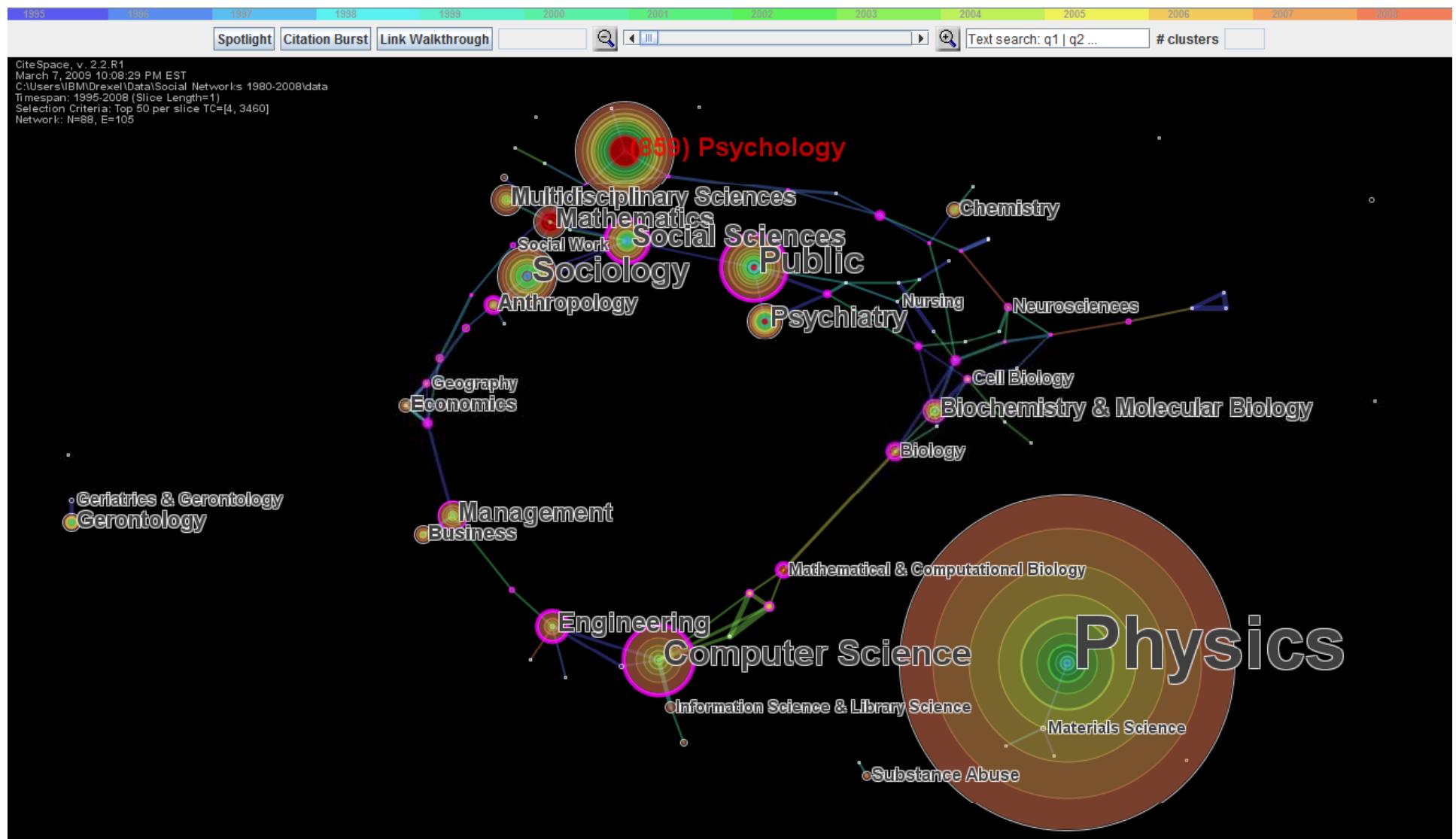
Pathfinder Pruning sliced networks
 Minimum Spanning Tree Pruning the merged network

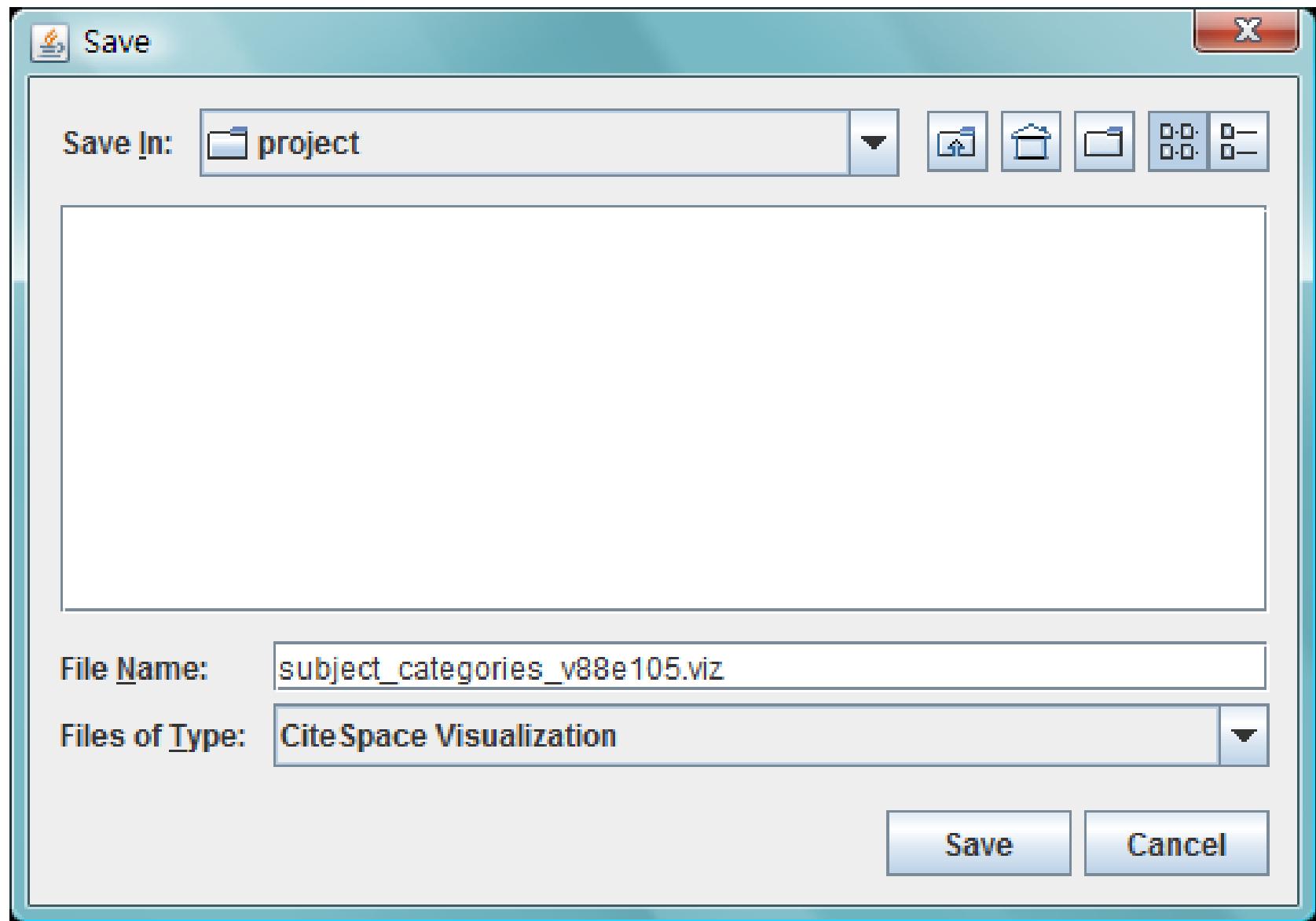
Visualization

Cluster View (Default) Show Networks by Time Slices
 Time-zone View Show Merged Network









CiteSpace 2.2.R1 - (c) 2003-2009 Chaomei Chen ~ Home: C:\Users\IBM - COA

File Data Project Metrics Text Network Analyze Geospatial Maps Test Preferences Gallery Help

Web of Science PubMed CiteSeer

Projects

New Complex Network Analysis 1980-2... More Actions ...

Project Home: C:\Users\IBM\Datasocial Networks 1980-2008\project

Data Directory: C:\Users\IBM\Datasocial Networks 1980-2008\data

GO! Stop Reset JVM Memory (K) 174080 Used (%) 37

Space Status

Process Reports

Time Slicing

From 1995 To 2008 #Years Per Slice 1 Entropy

Modeling

Node Types

Author Term Keyword Category Institution Country
 Cited Reference Cited Author Cited Journal

Term Sources

Title Abstract Descriptors Identifiers

Term Selection

None Noun Phrases Burst Phrases Find Burst Phrases

Links

Strength Cosine Scope Within Slices

Top N per slice Top N% per slice Threshold Interpolation Select Citers

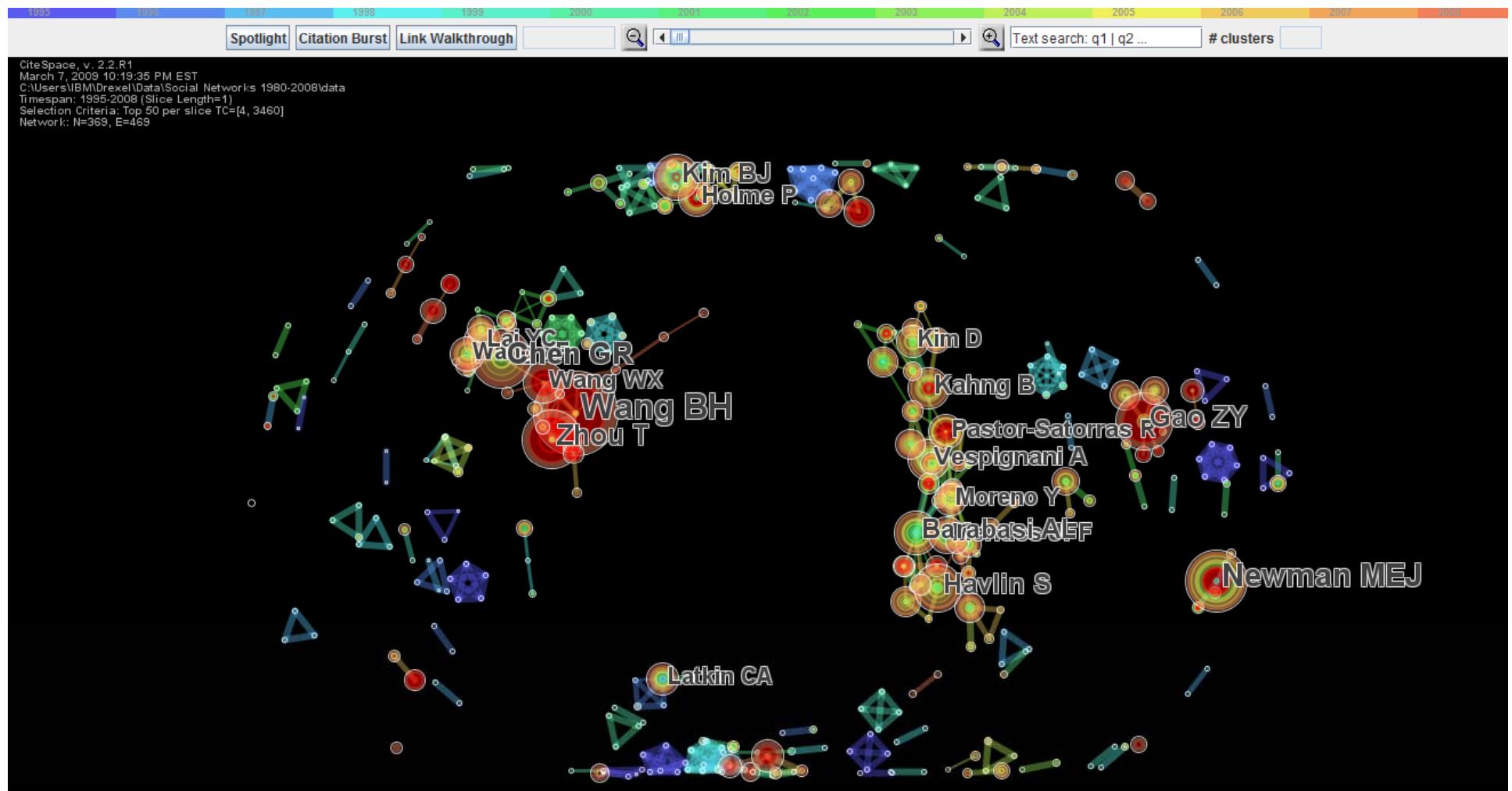
Select top 30 most cited or occurred items from each slice.

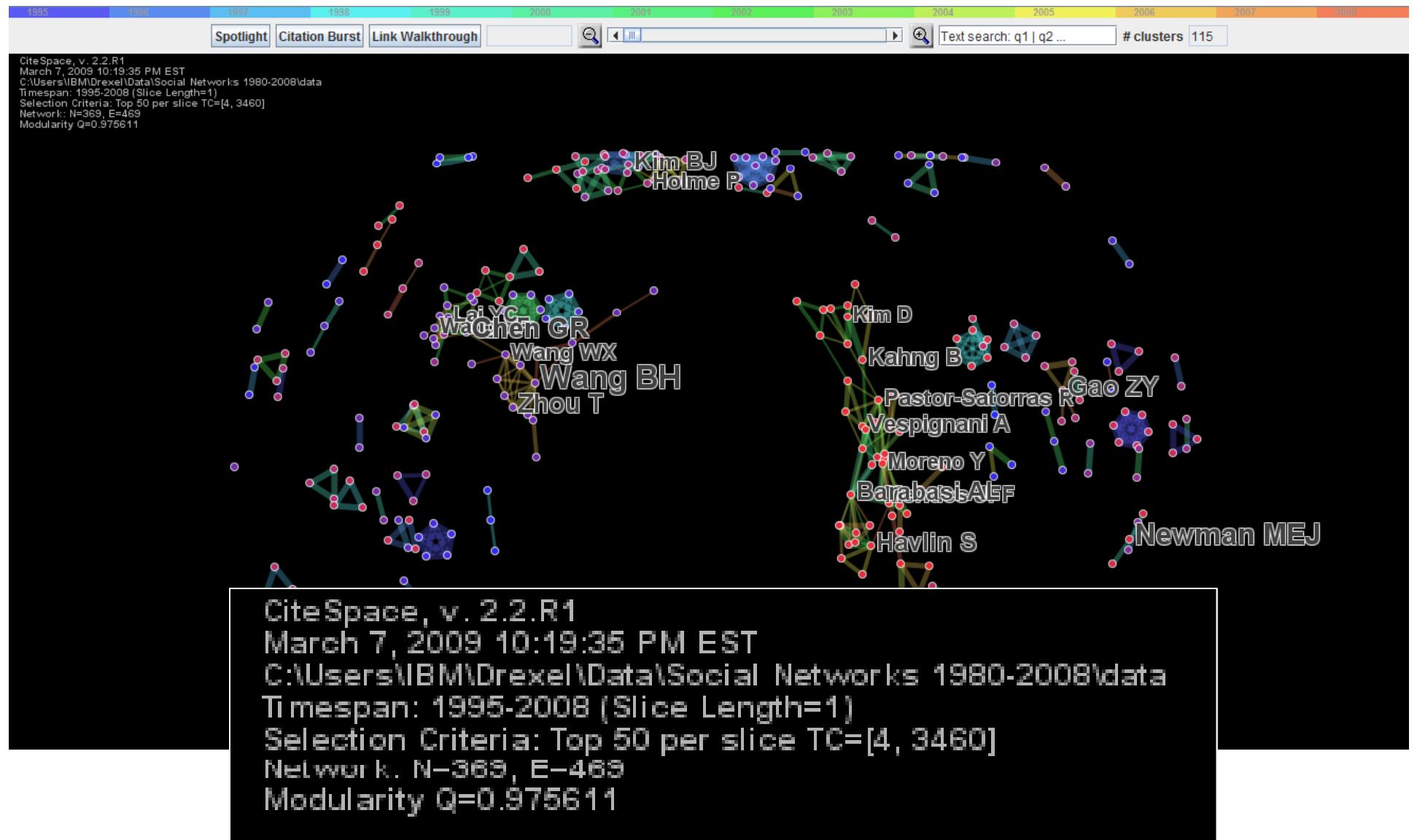
Pruning

Pathfinder Pruning sliced networks
 Minimum Spanning Tree Pruning the merged network

Visualization

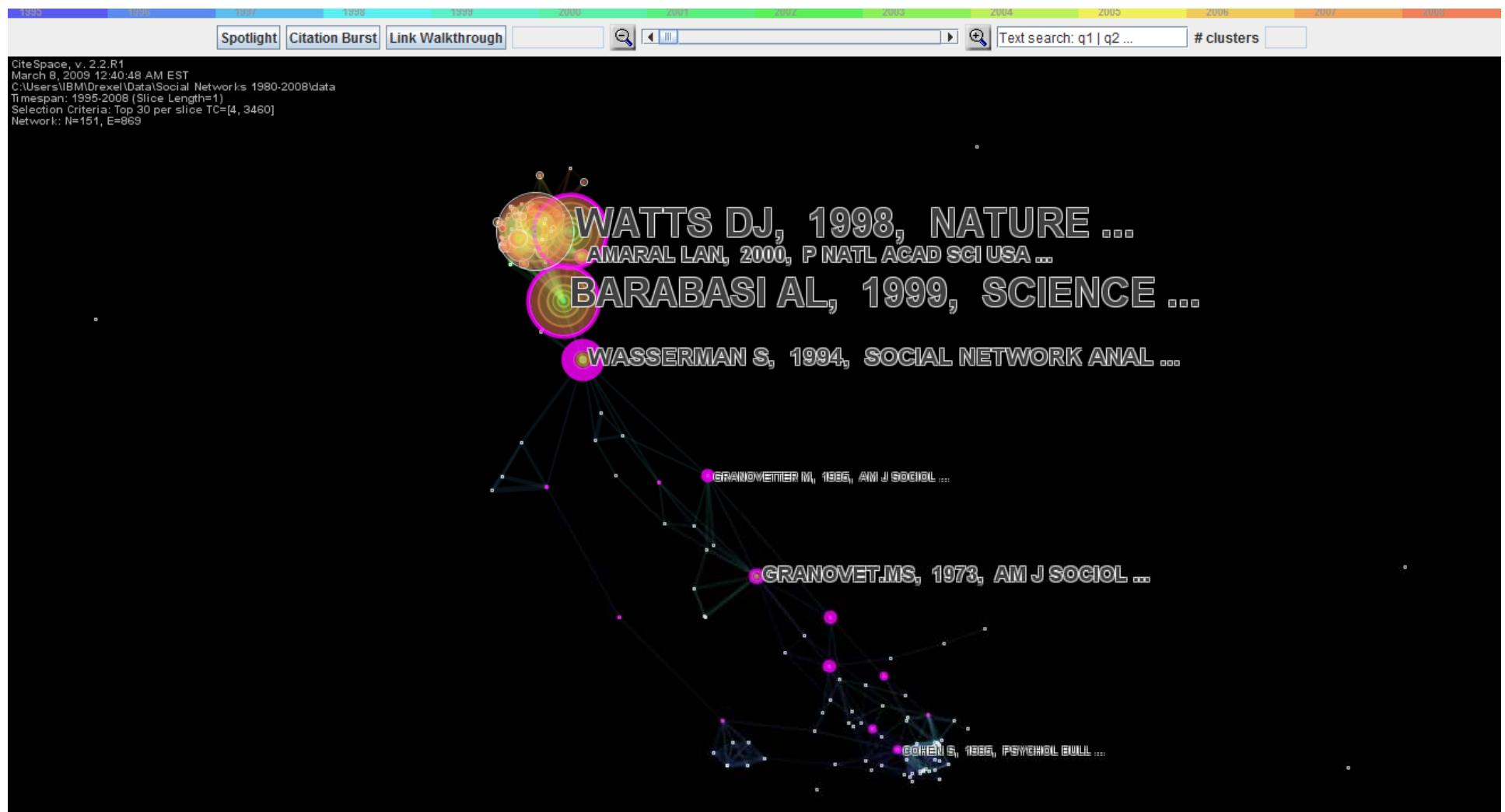
Cluster View (Default) Show Networks by Time Slices
 Time-zone View Show Merged Network





Summary of Clusters - terms from abstracts

Cluster	Size	Top 5 terms by tf*idf	Term by loglikelihood	Term by mutual information	Most frequent Citors
43	8	(303.68) endothelial cell; (...)	(24)biological consequence	...	(8)basic fibroblast growth factor (bfgf) is e...
111	8	(31.08) scale-free network...	(4)dynamical model	(6)optical gap	(3)we report optical gaps for ab initio gen...
39	8	(42.38) scale-free network...	(4)brief report	(4)complex network	(4)in this brief report, we investigate the c...
67	7	(137.6) coping strategy; (1...	(6)chronic psychotic patient	(9)social network	(6)background: to date, only few data are ...
31	6	(113.88) village communit...	(24)residential campuse	...	(6)the costs, nature, and benefits of resid...
92	6	(256.23) mucus secretion;...	(12)l galanin	...	(6)neurokinin a and b are present in neur...
100	6	(37.96) functional outcom...	(5)severe mental illness	...	(3)background social networks are impor...
96	6	(56.94) tough boy; (55.68) ...	(4)elementary classroom	...	(3)this study examined subtypes of popul...
37	6	(23.72) numerical comput...	(4)homogeneous network	(5)scale-free network	(2)a large computer program is typically d...
42	6	(18.56) different network to...	(2)act degree distribution	(4)complex network	(1)a model of epidemic spreading in a po...
44	6	(106.31) drug user; (102.5...	(12)personal social network	(12)social network	(6)few studies have examined the current...
14	5	(92.8) inner-city neighborh...	(5)hiv infection	...	(5)relational and attribute data were colle...
23	5	(47.45) subpopulation siz...	(5)power law response	...	(5)we asked respondents how many peo...
30	5	(111.36) sierpinsiki fractal; ...	(14)average path length	...	(4)we propose an extended local-world e...
104	5	(24.31) fraction p; (16.79) i...	(3)largests connected cluster	(4)scale-free network	(3)we study the tolerance of random netw...
34	5	(85.48) drug use; (56.94) ...	(6)alcohol consumption	(4)injection drug user	(2)purpose: the nature of competing soci...
36	5	(28.47) firing synchronizati...	(6)scale-free feature	(3)small-world network	(2)small-world networks are highly cluste...
73	5	(37.96) supply chain distri...	(6)rich node	(3)complex network	(3)a new model for network generation is ...
69	5	(23.72) construction comp...	(2)fad dynamic	(2)cooperation level	(2)we propose an information-based mo...
74	5	(70.00)



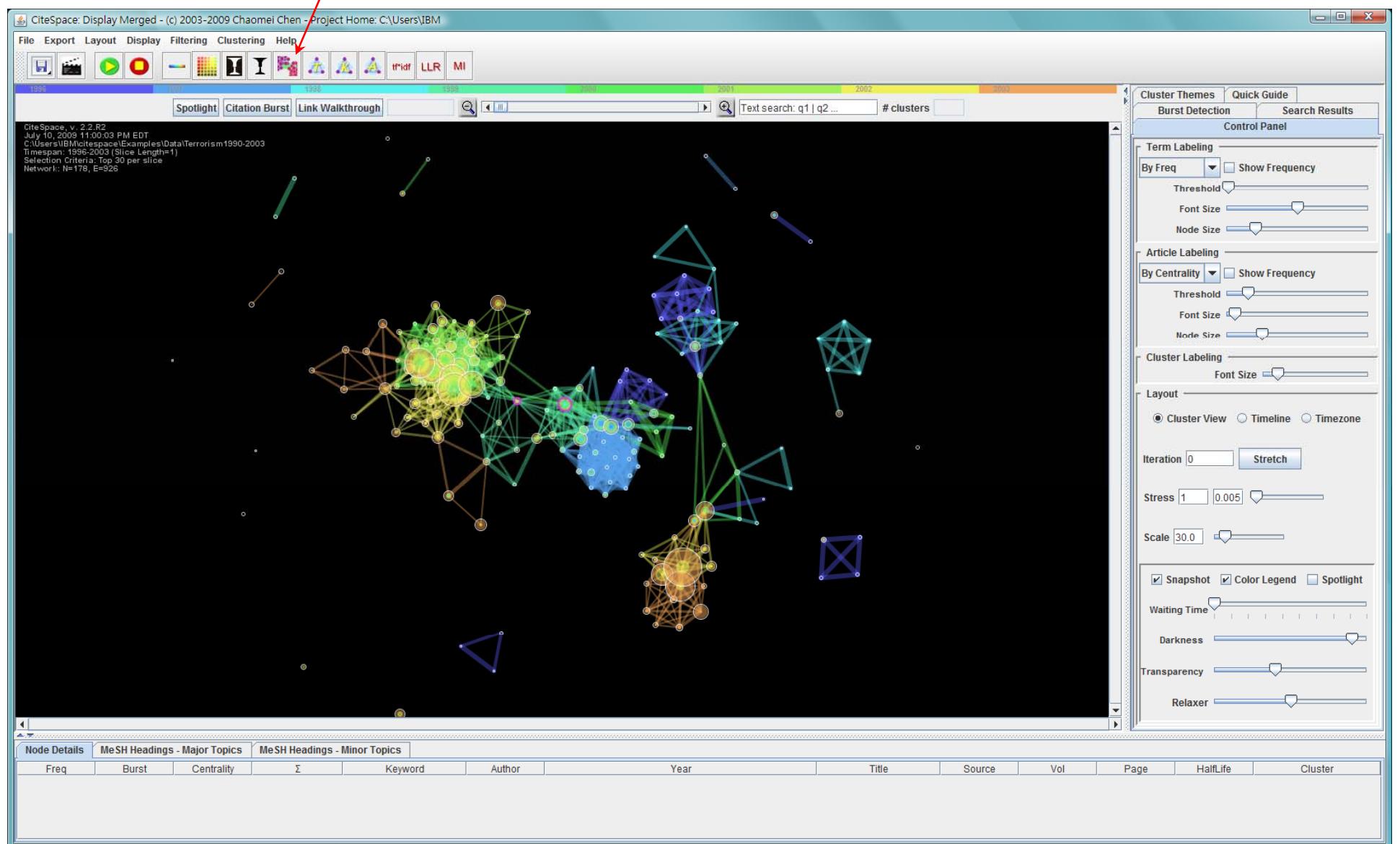
EXAMPLES

- 1. TERRORISM RESEARCH*
- 2. INFORMATION SCIENCE*
- 3. INFLUENZA PANDEMIC*
- 4. SLOAN DIGITAL SKY SURVEY (SDSS)*

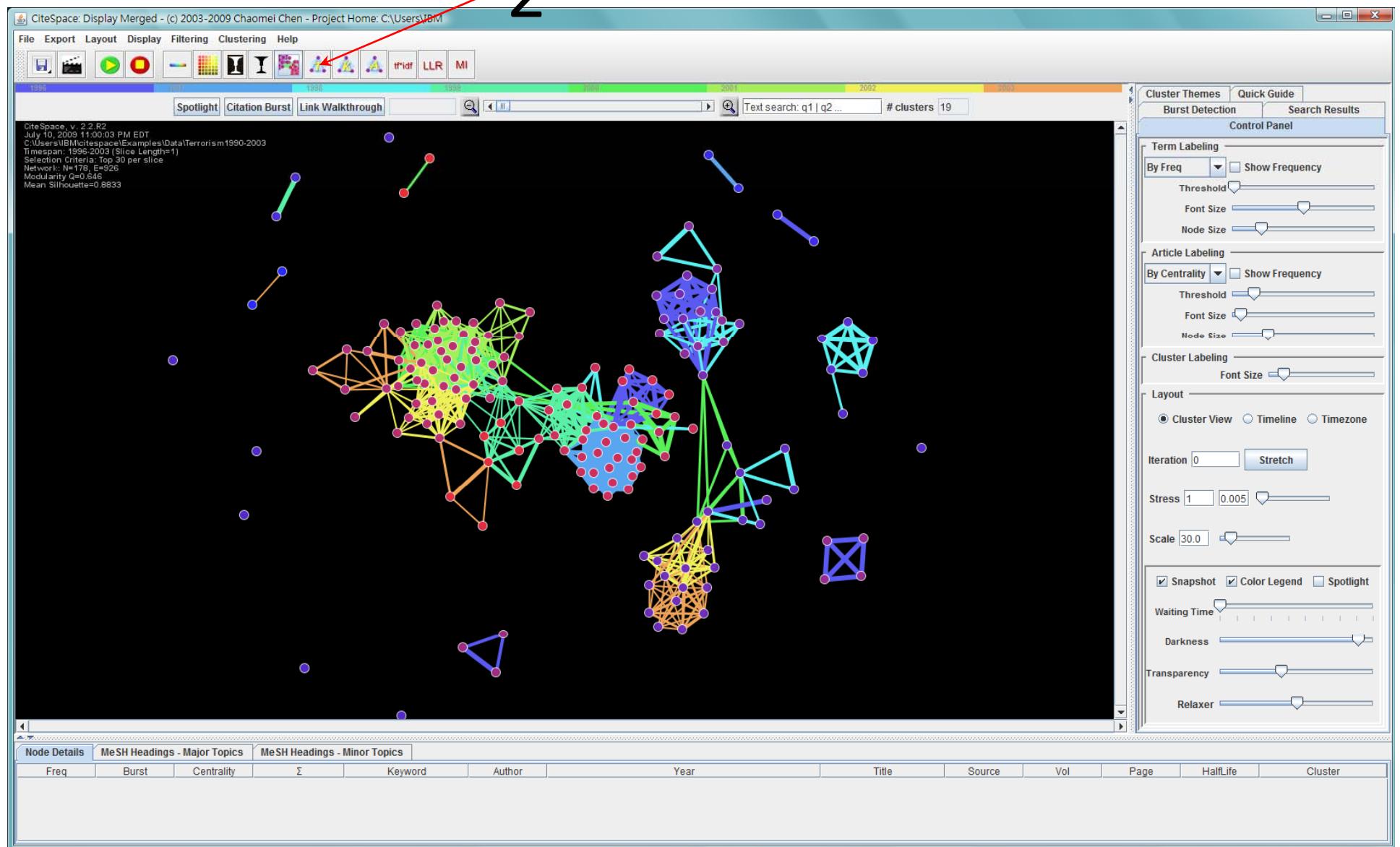
1. Terrorism

- What are the questions to ask?

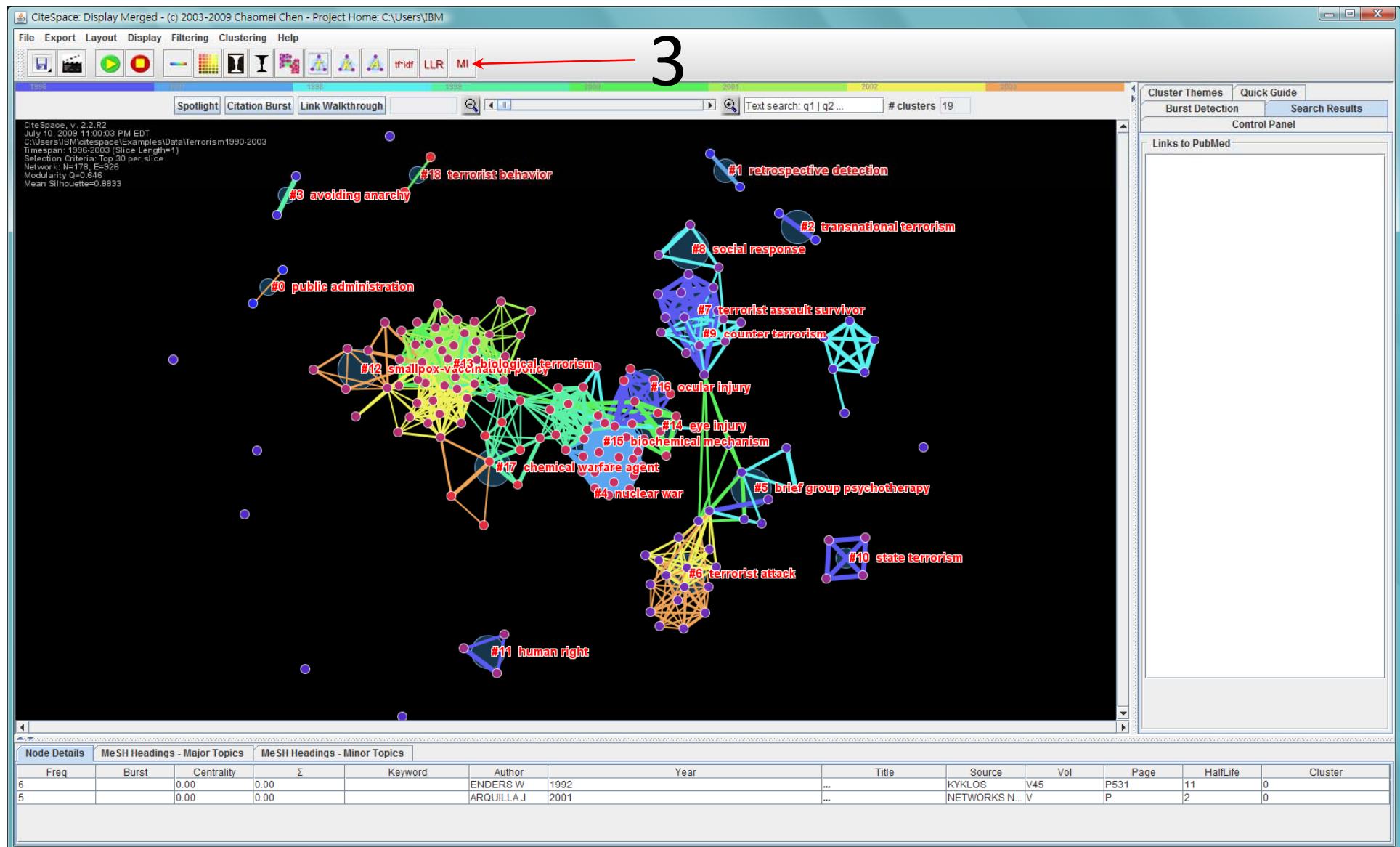
1

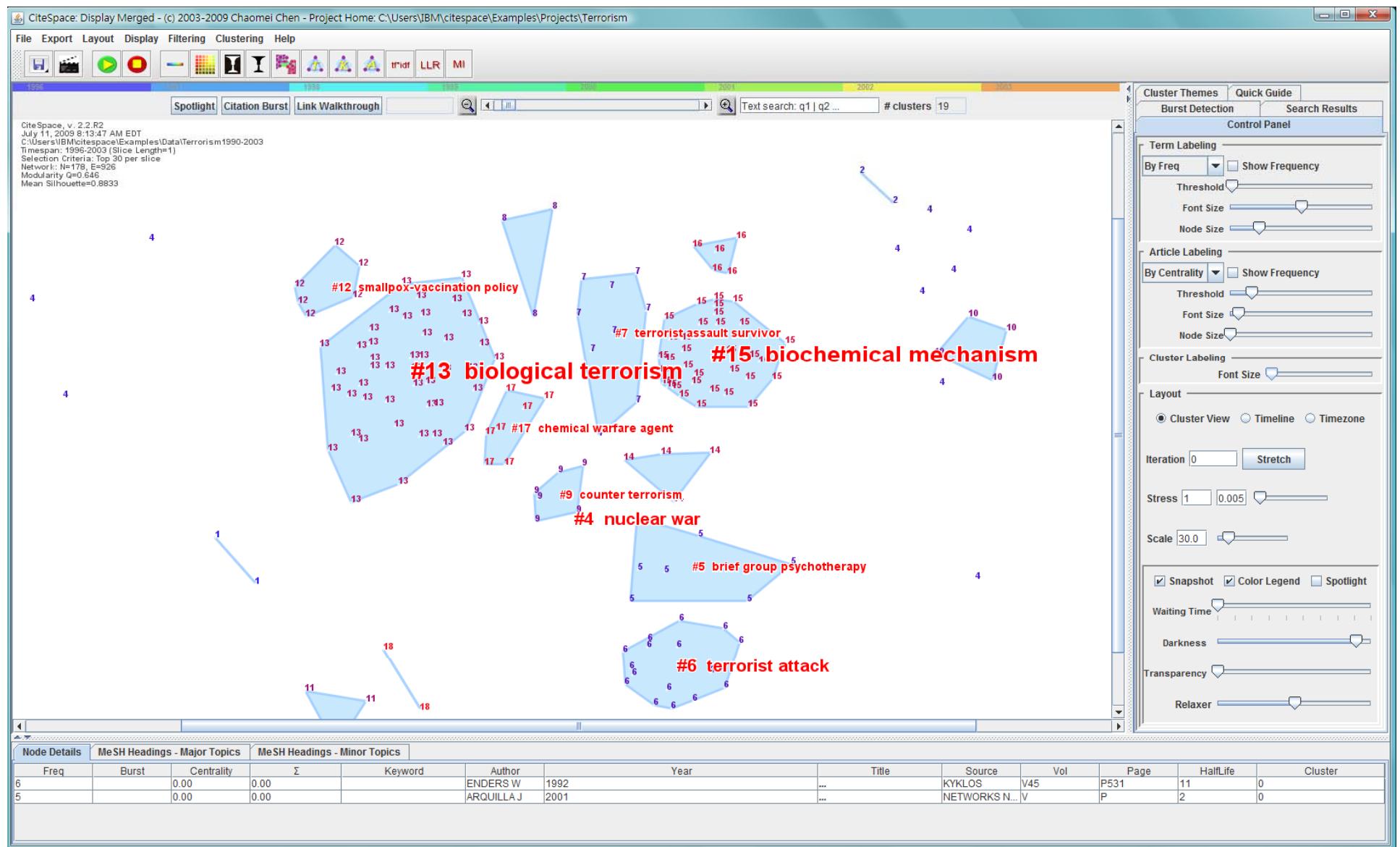


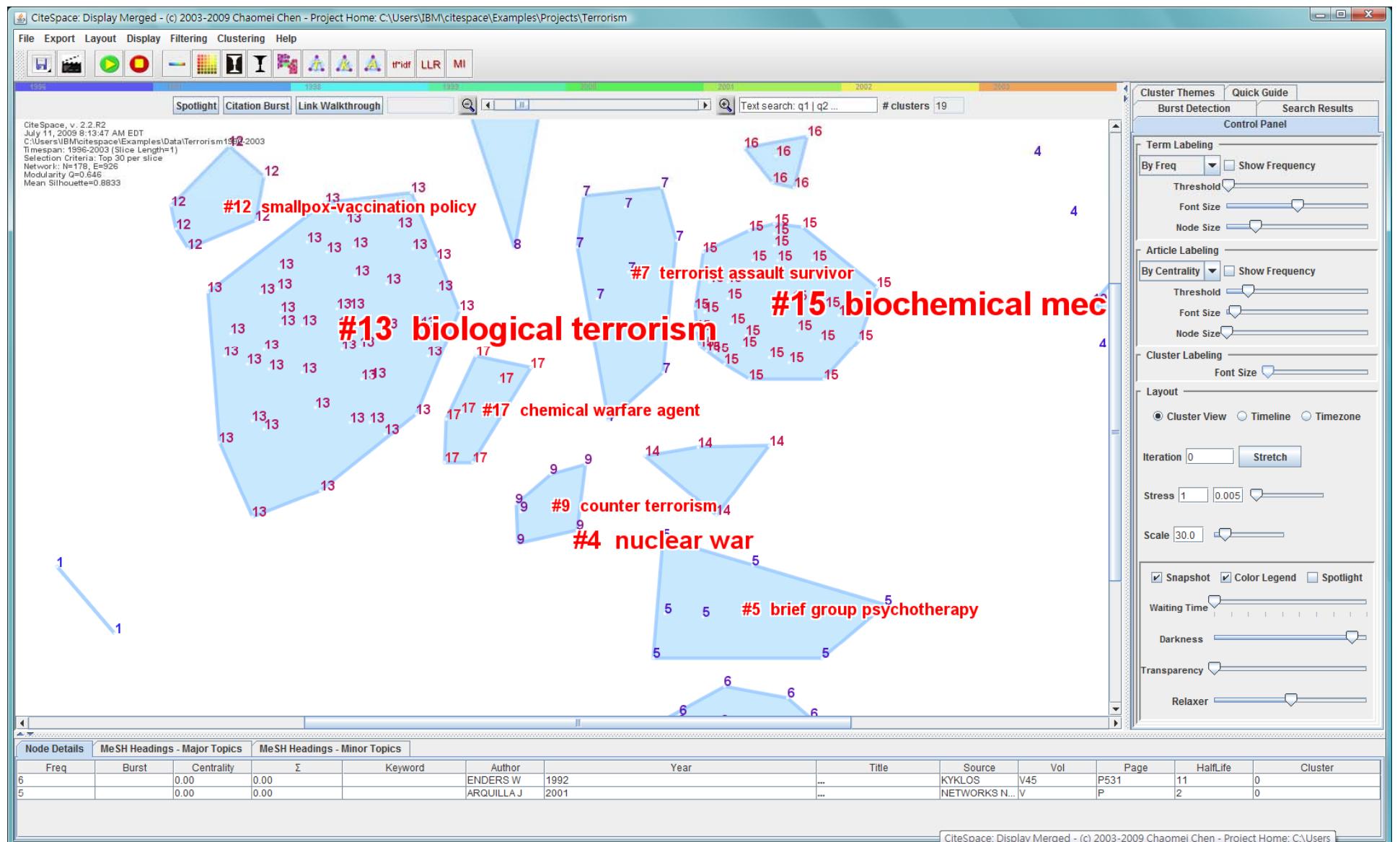
2



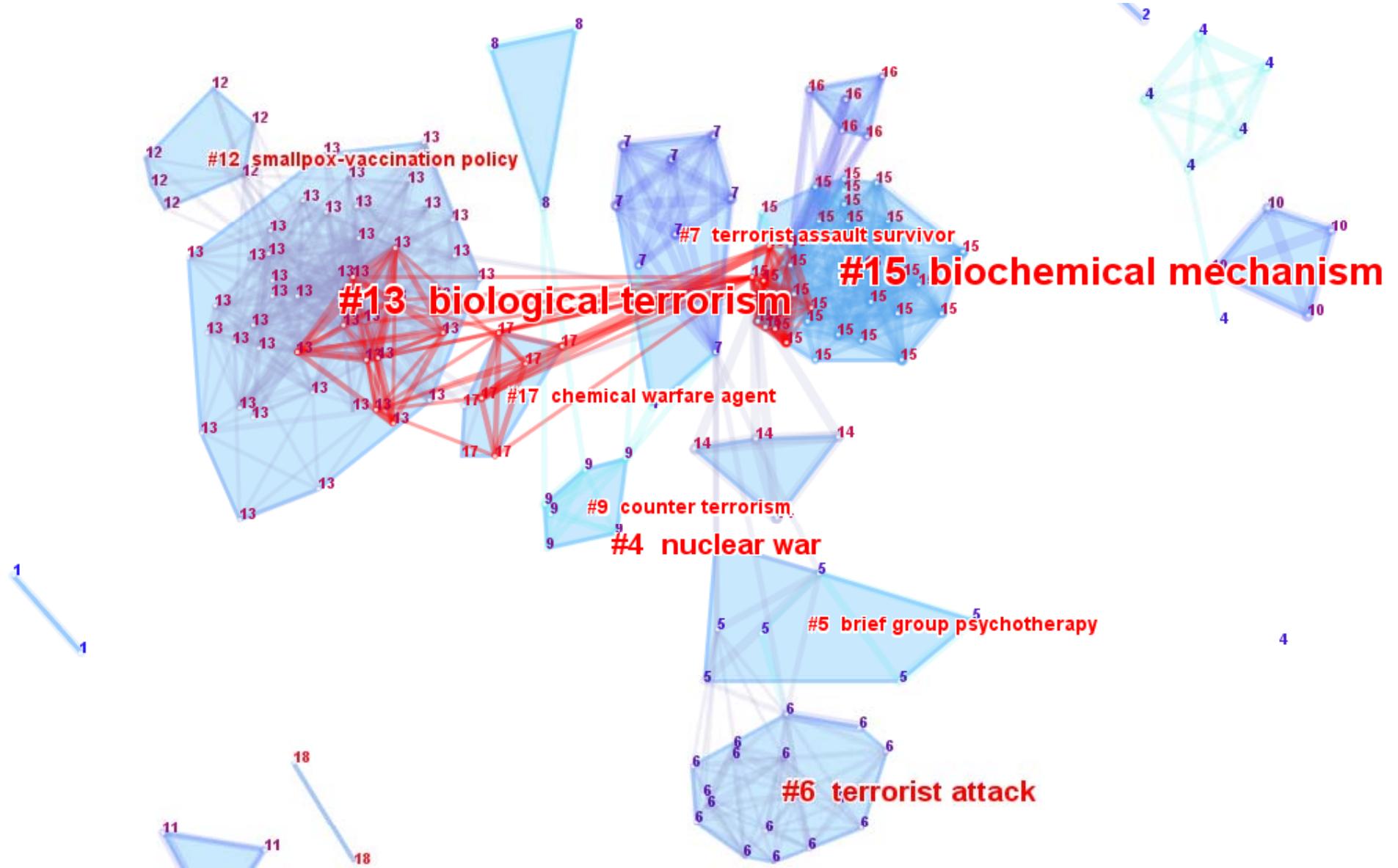
3



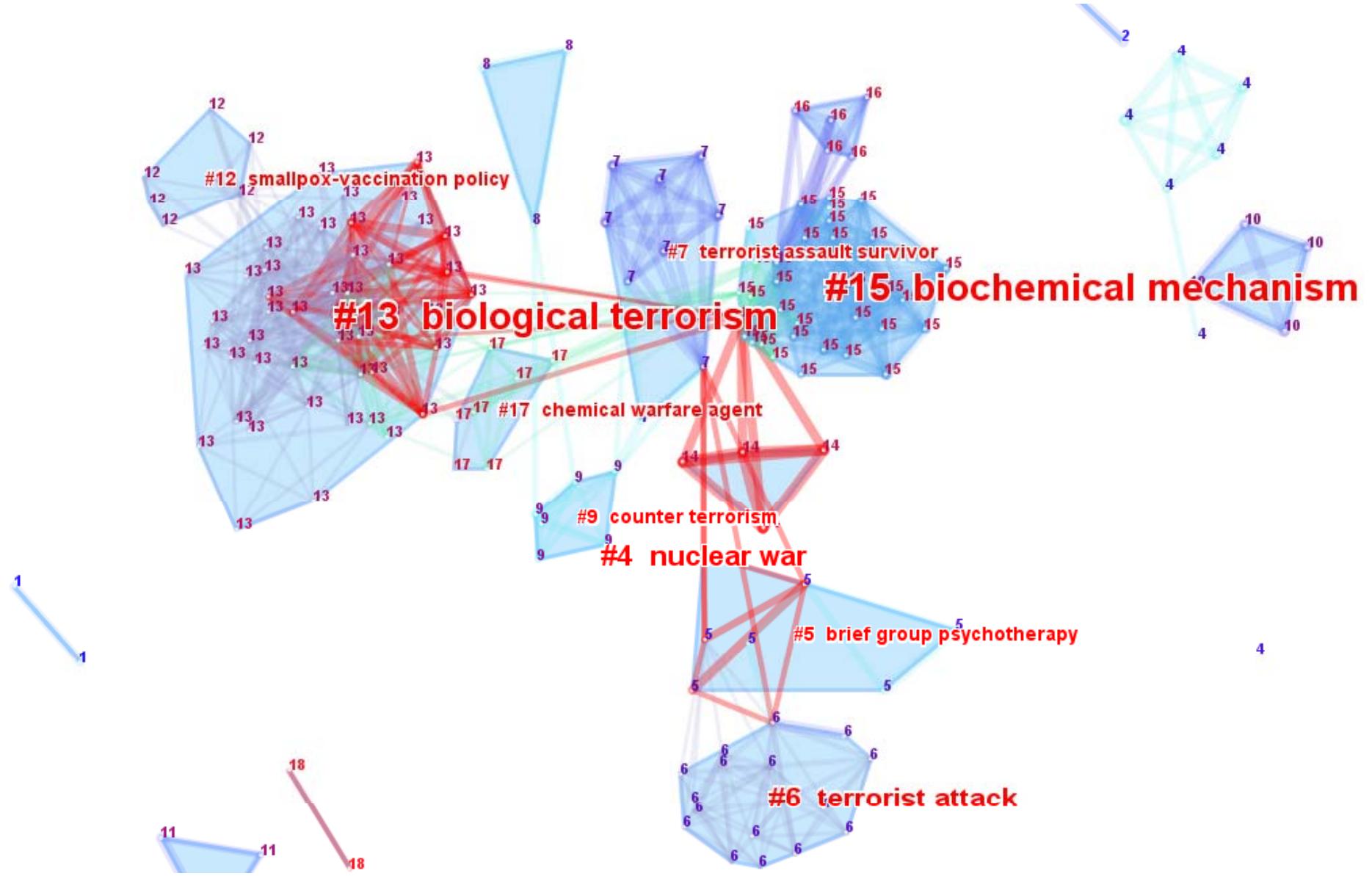


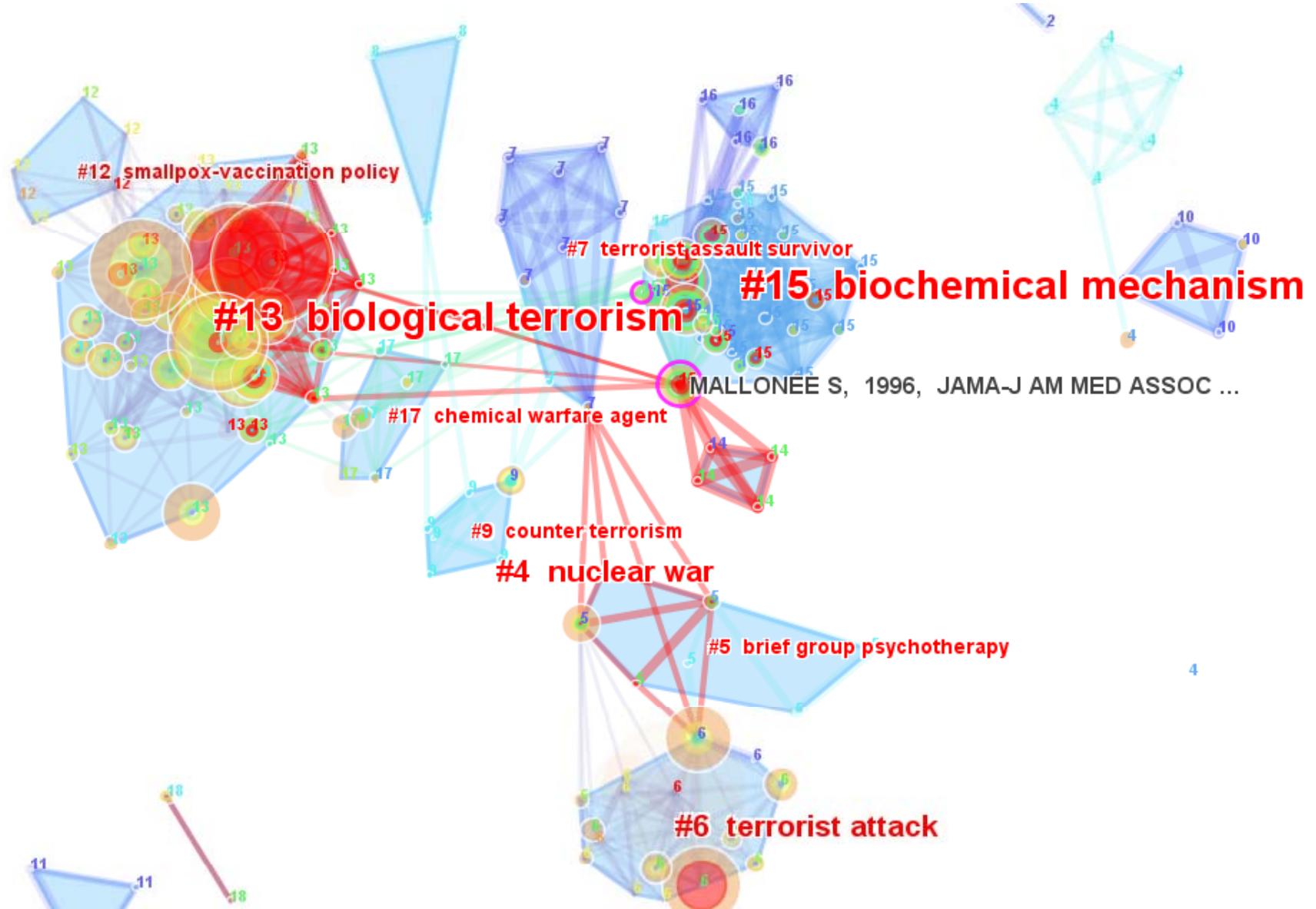


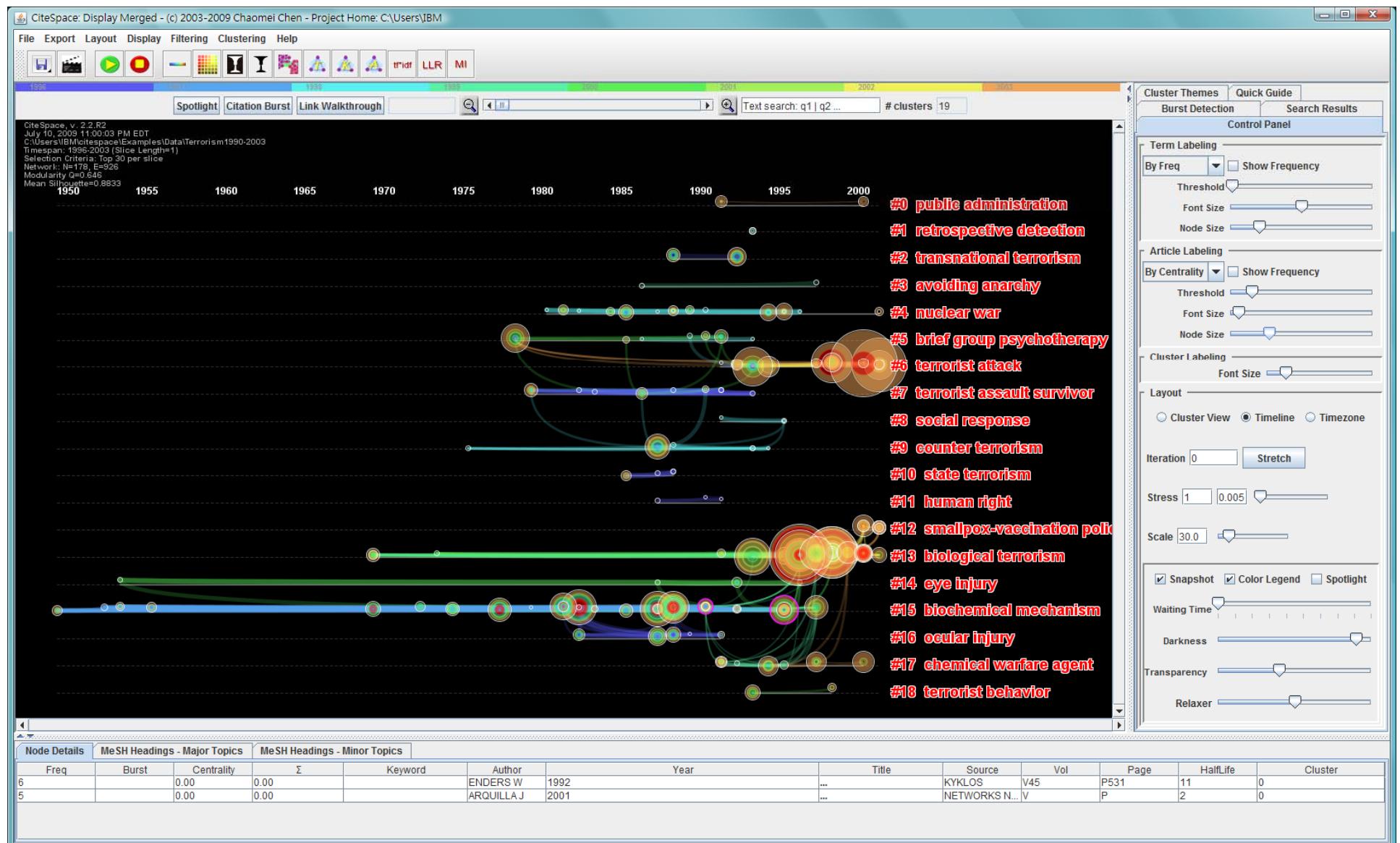
Link Walkthrough: 1999-1999



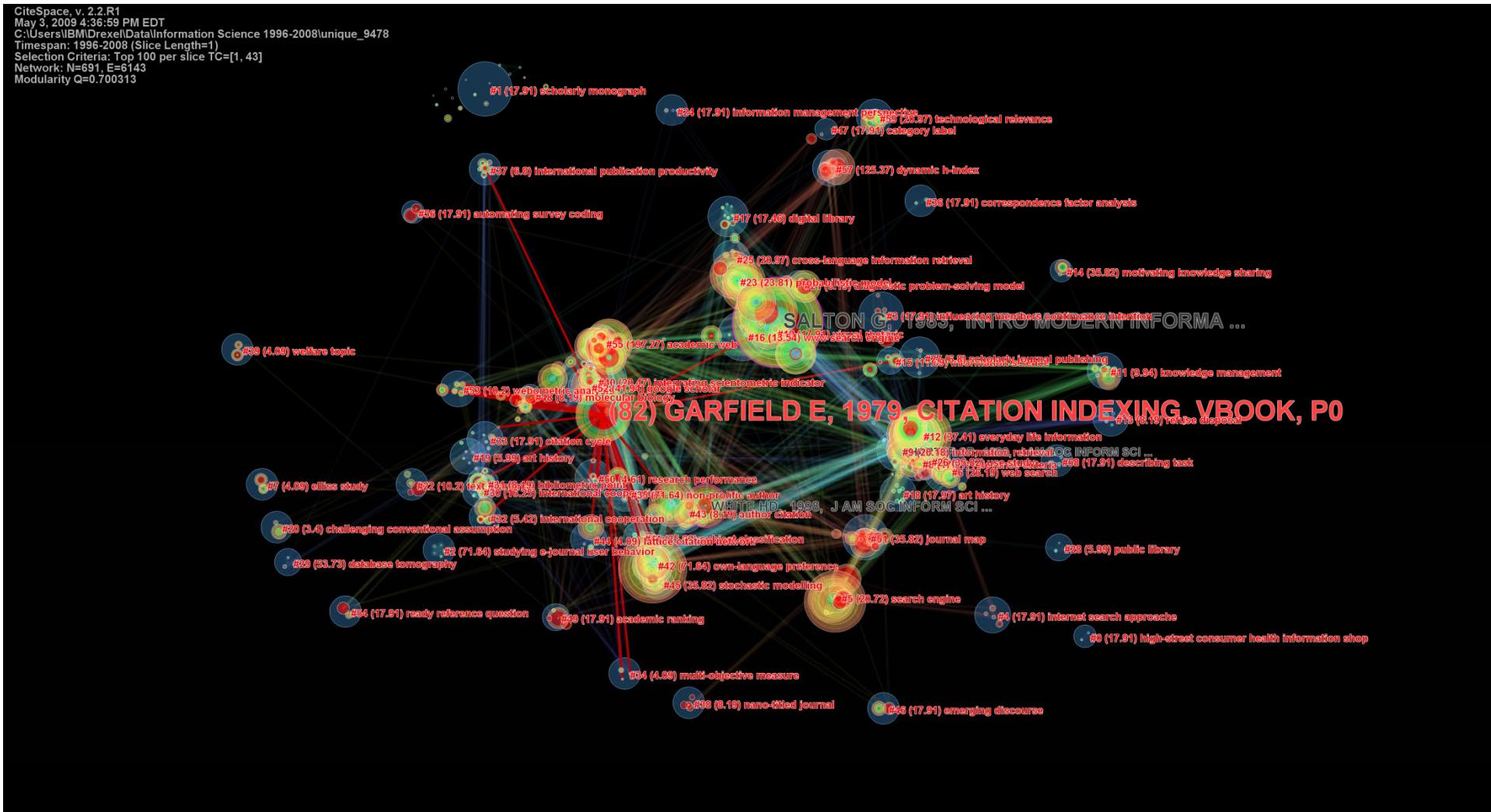
Link Walkthrough: 2000-2000



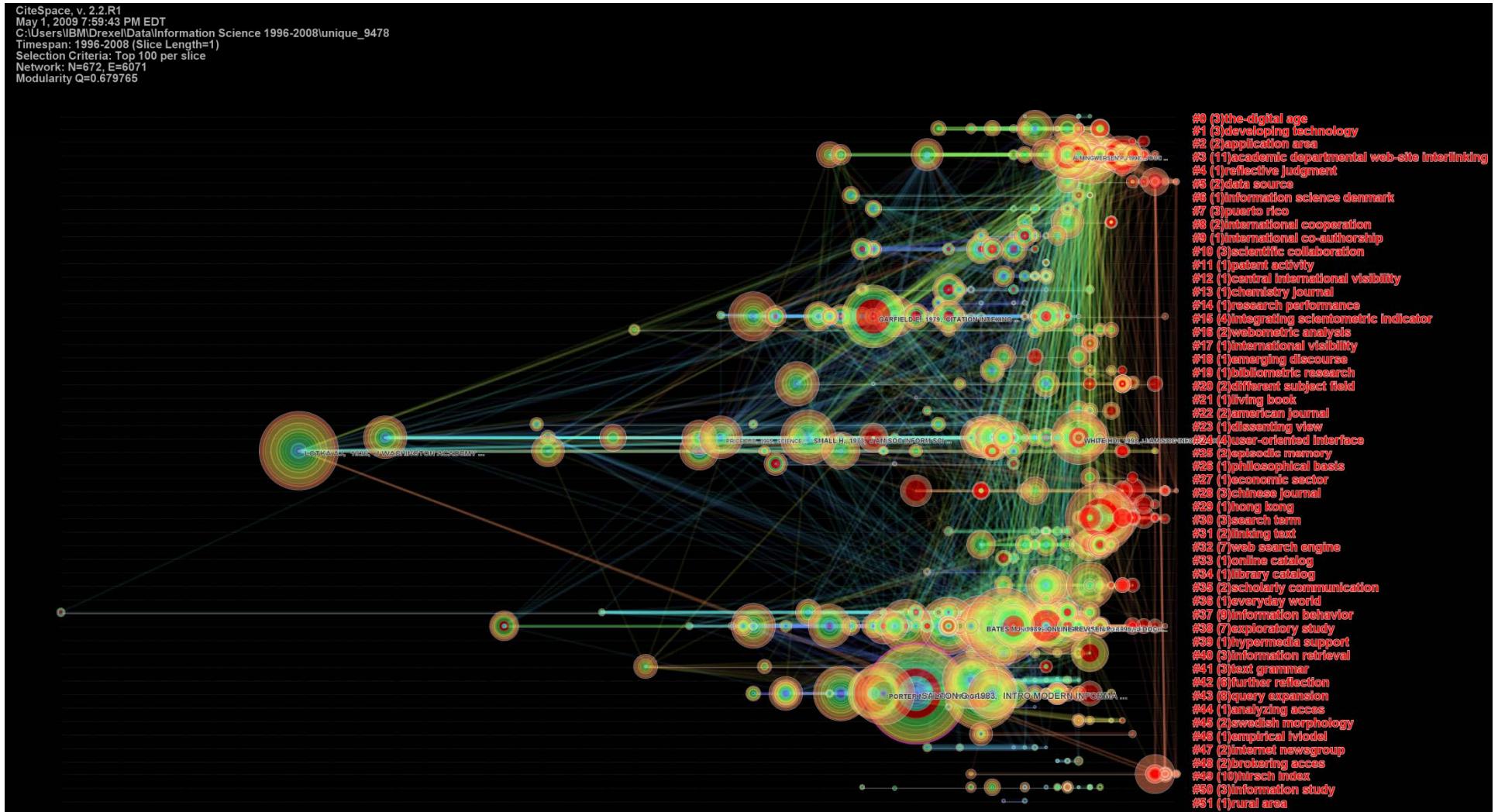




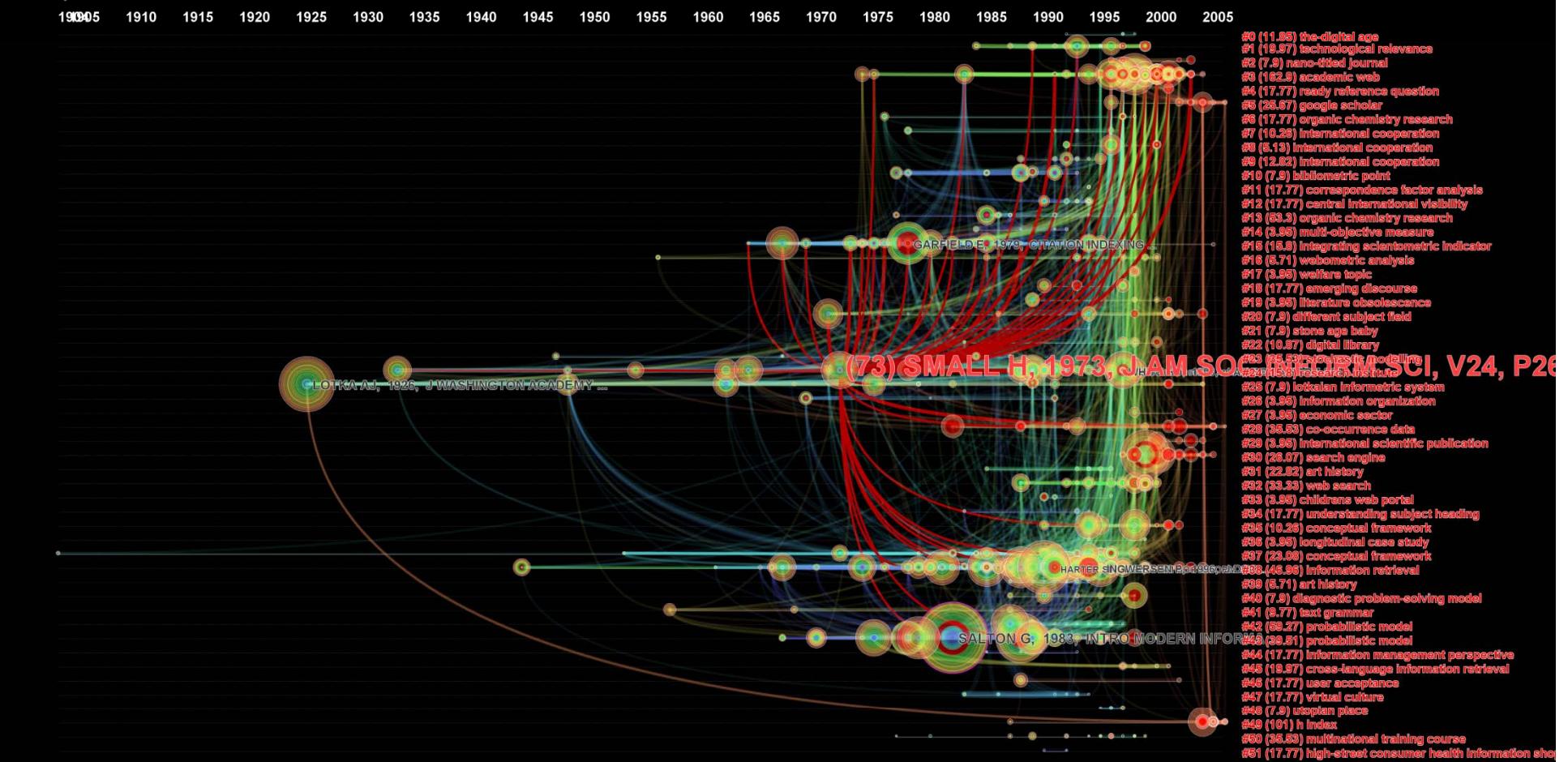
2. Information Science



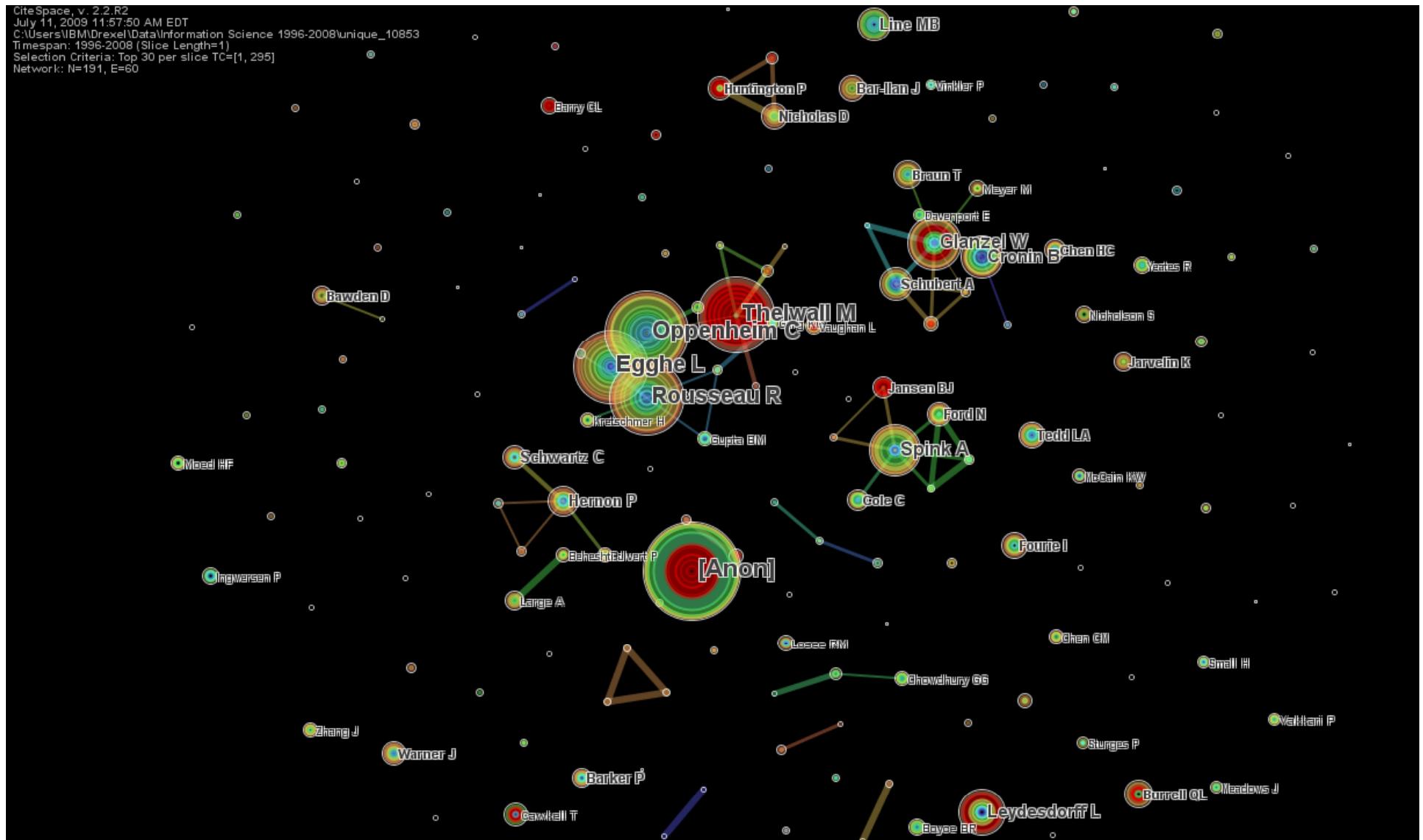
Information Science



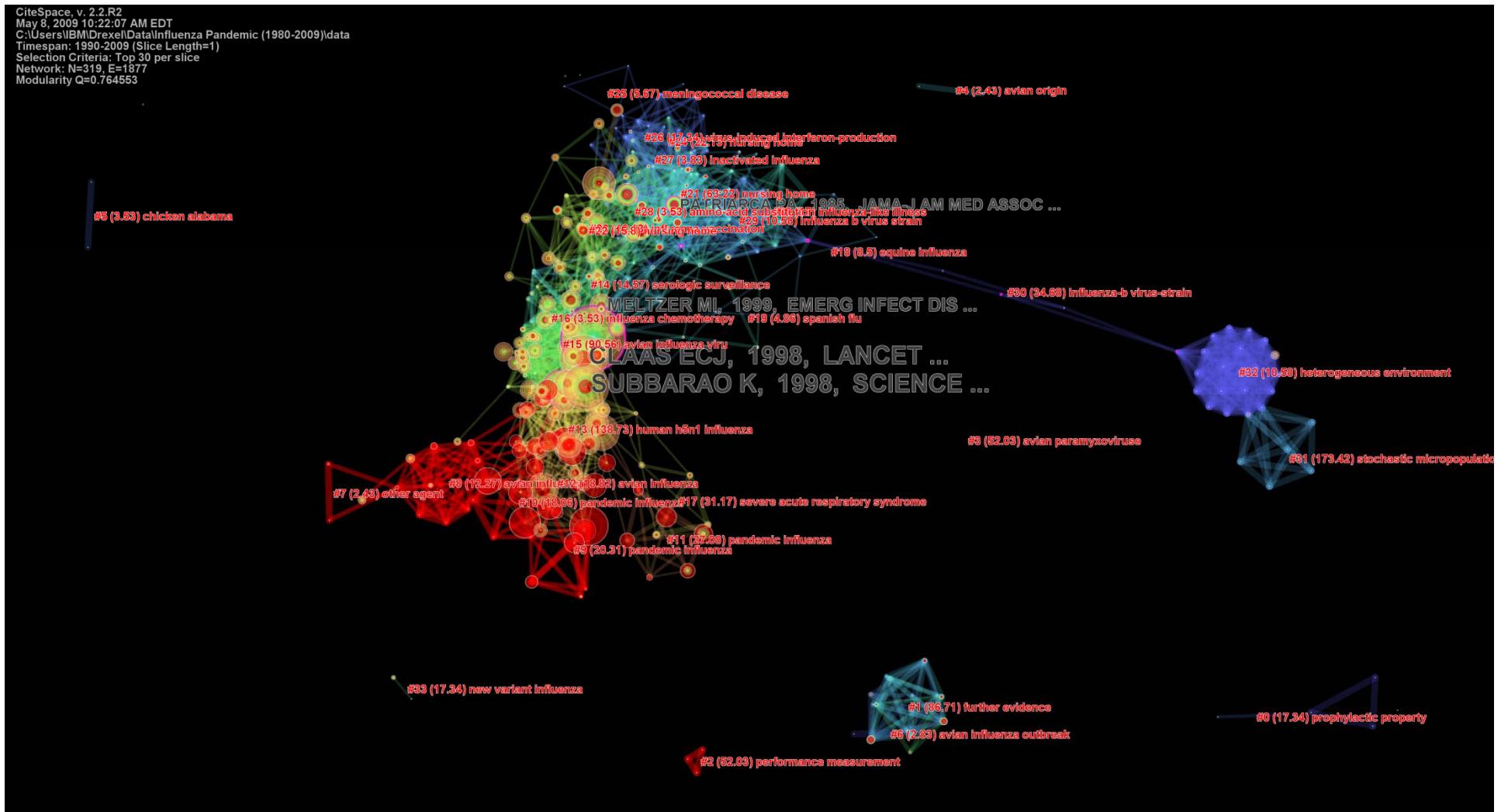
CiteSpace, v. 2.2.R1
 May 3, 2009 2:58:36 PM EDT
 C:\Users\BMDrexel\Data\Information Science 1996-2008\unique_9478
 Timespan: 1996-2008 (Slice Length=1)
 Selection Criteria: Top 100 per slice
 Network: N=672, E=6071
 Modularity Q=0.679765



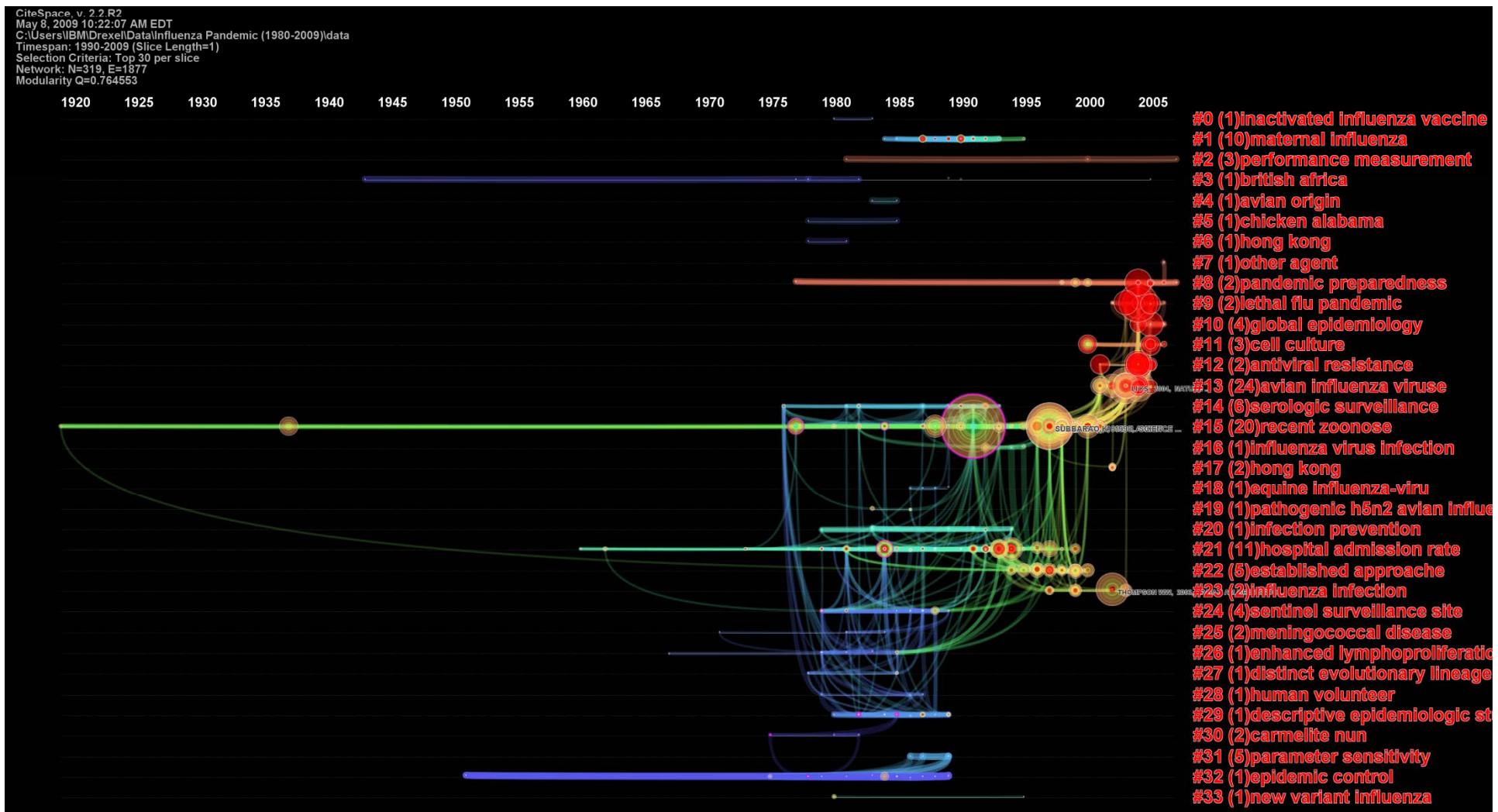
CiteSpace, v. 2.2.R2
July 11, 2009 11:57:50 AM EDT
C:\Users\IBM\Drexel\Data\Information Science 1996-2008\unique_10853
Timespan: 1996-2008 (Slice Length=1)
Selection Criteria: Top 30 per slice TC=[1, 295]
Network: N=191, E=60



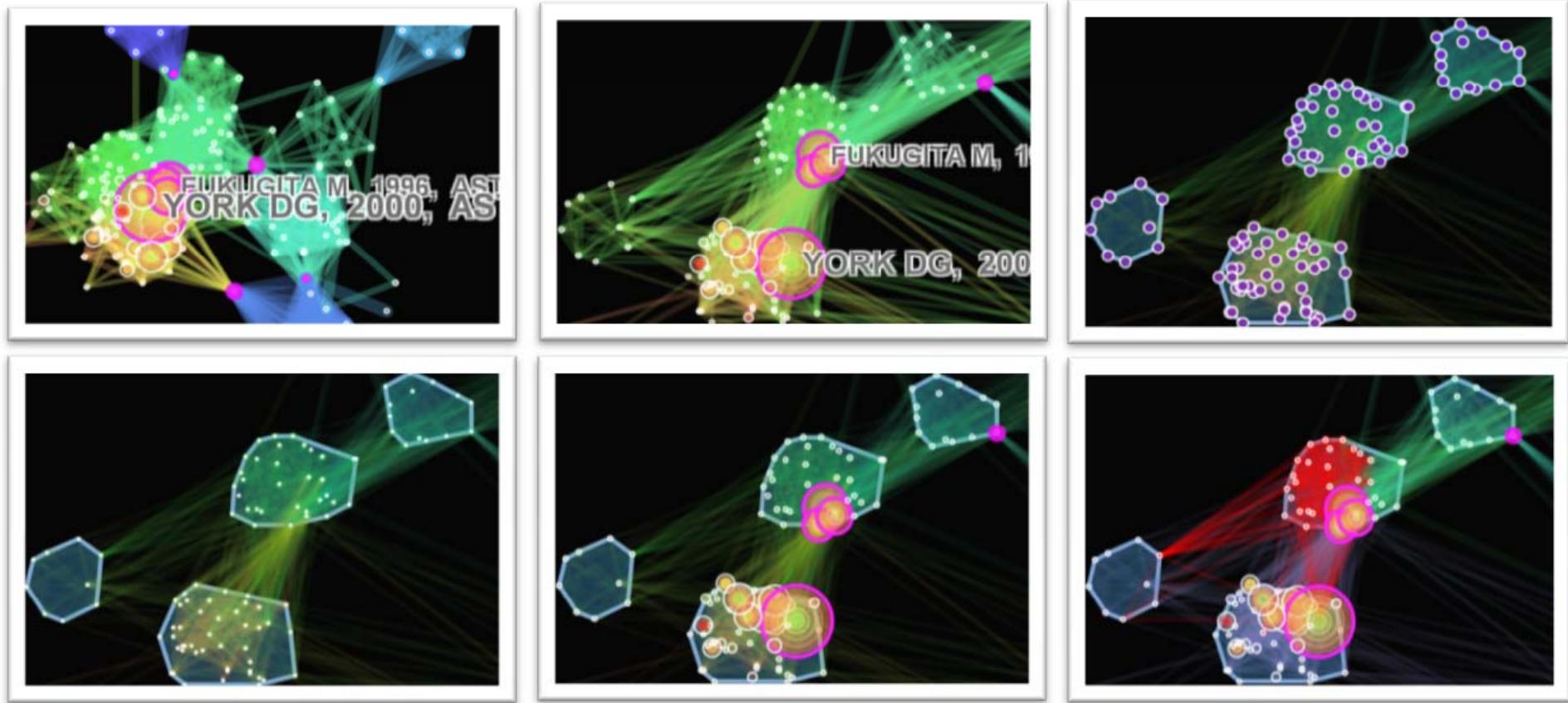
3. Influenza Pandemic



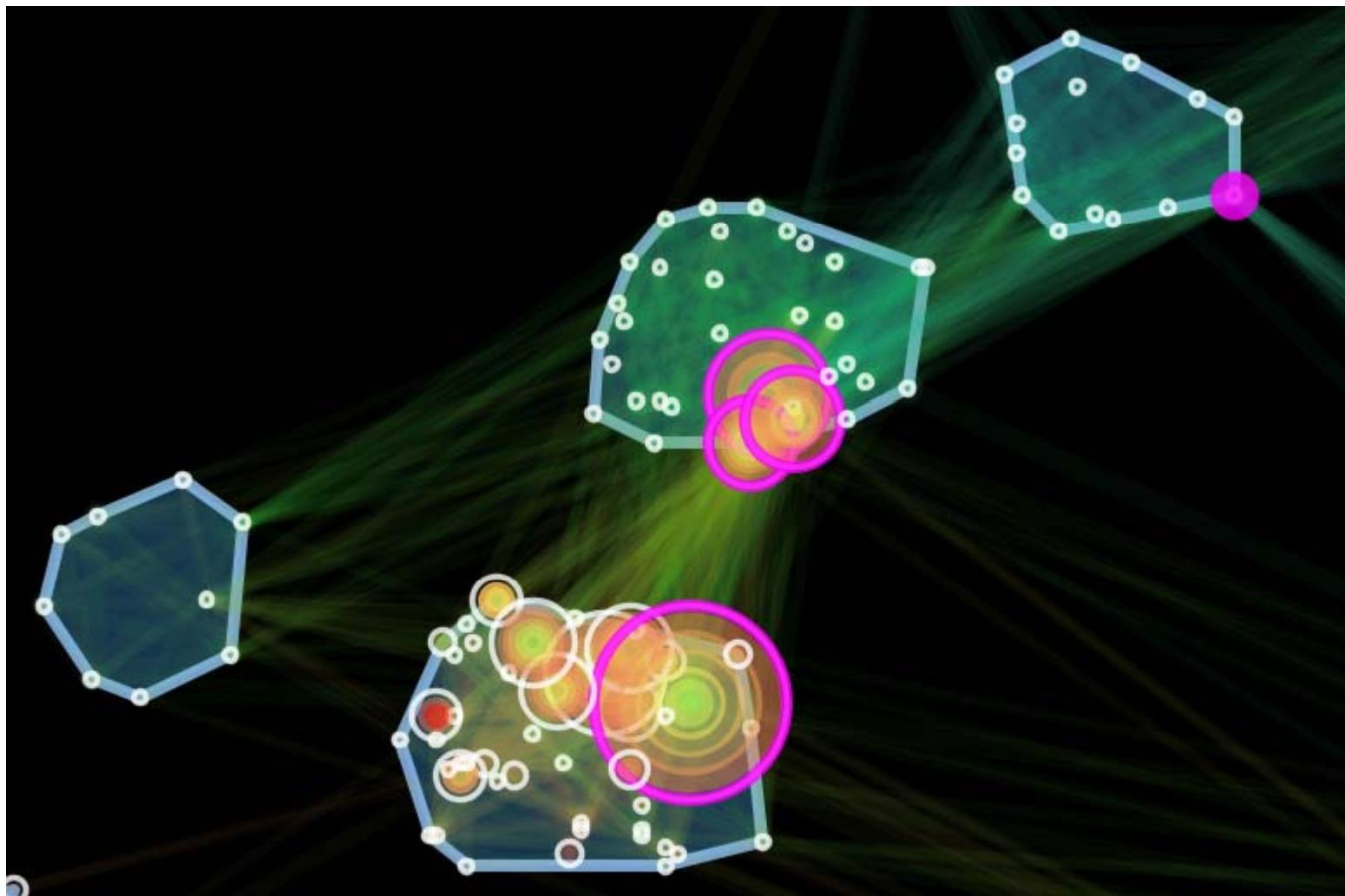
Influenza Pandemic



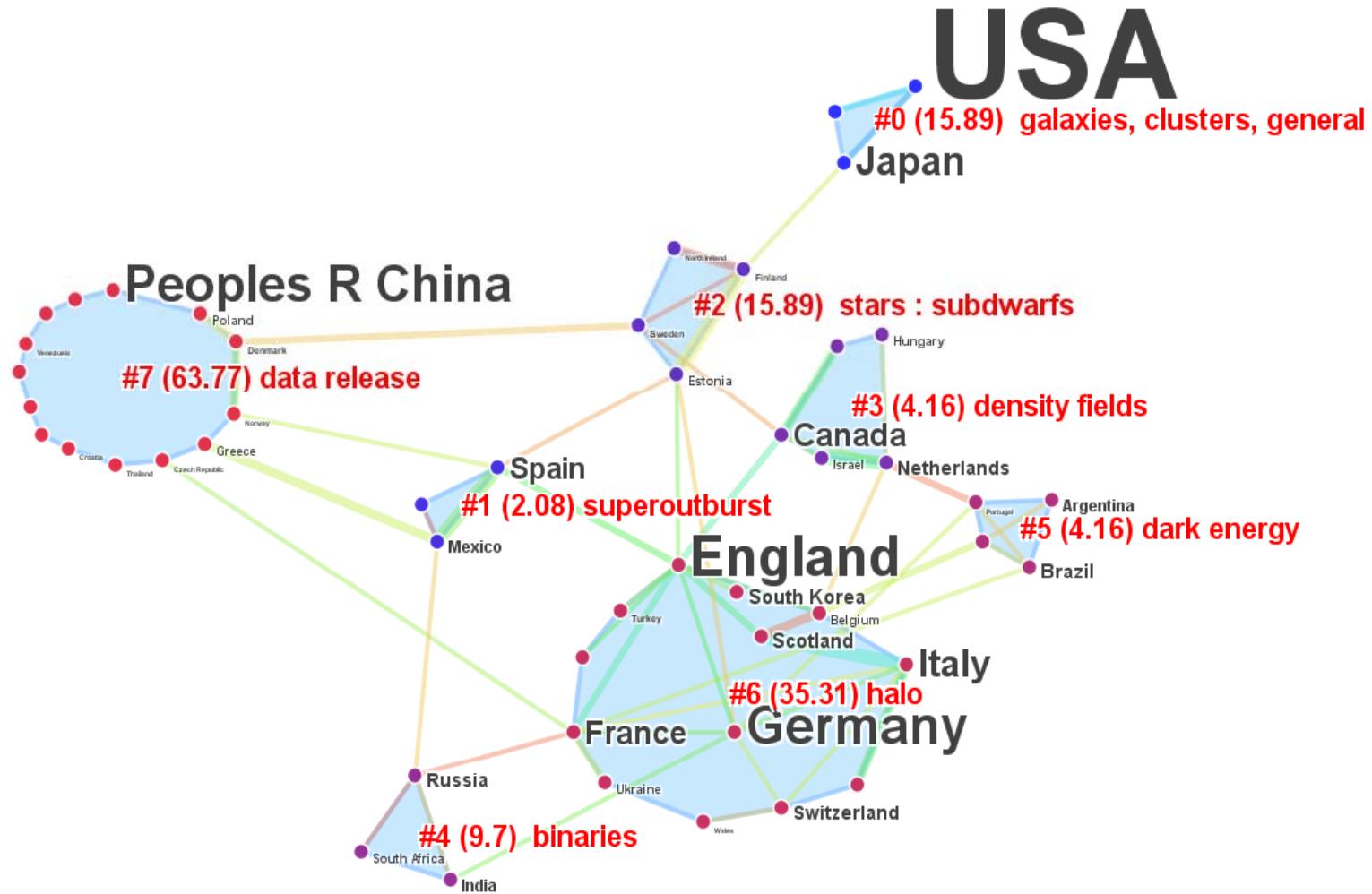
4. SDSS



Chen, C., Zhang, J. Vogeley, M. S. (2009) Visual analysis of scientific discoveries and knowledge diffusion. *Proceedings of the 12th International Conference on Scientometrics and Informetrics (ISSI 2009)*. July 14-17, 2009. Rio de Janeiro, Brazil.

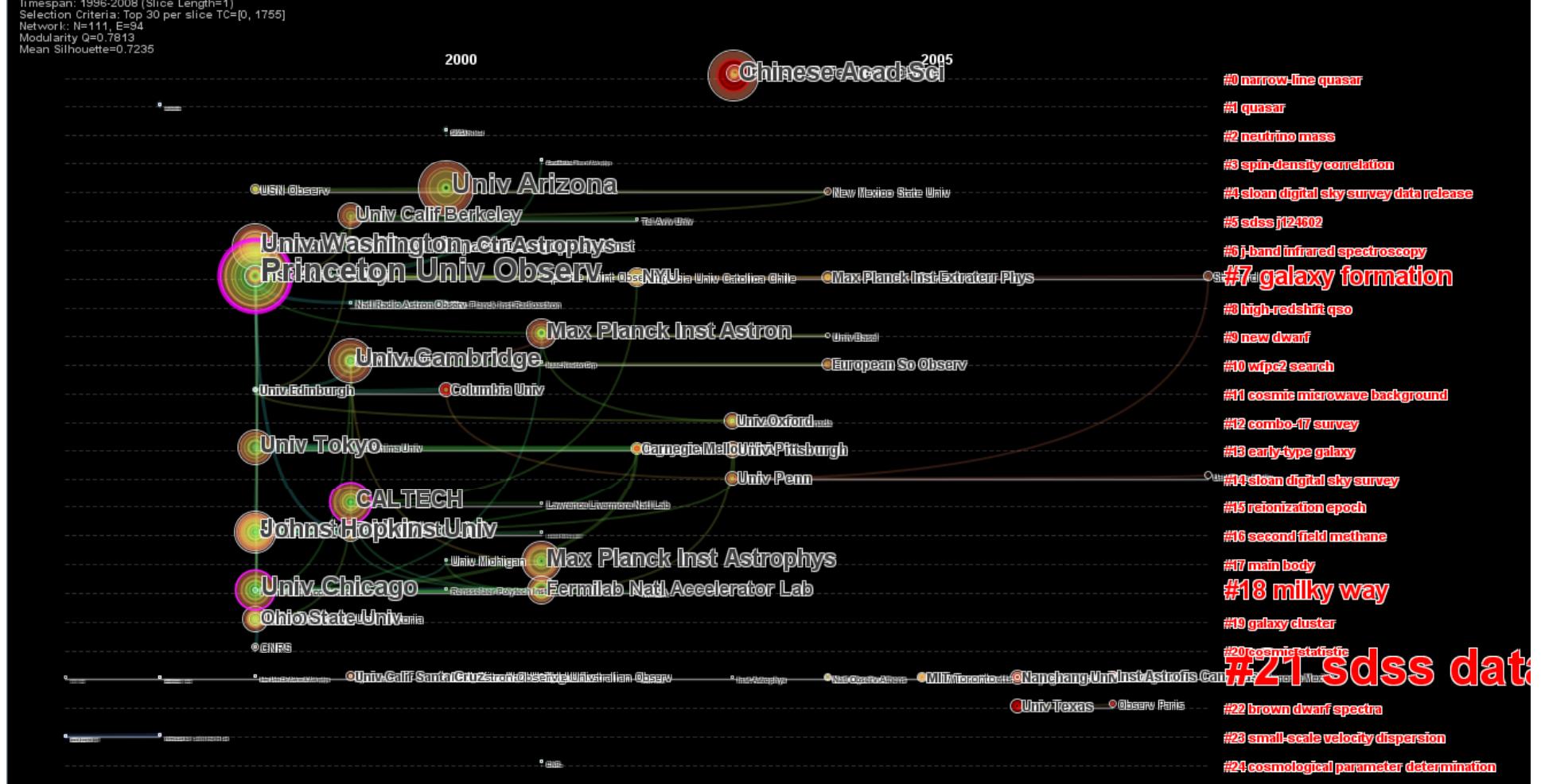


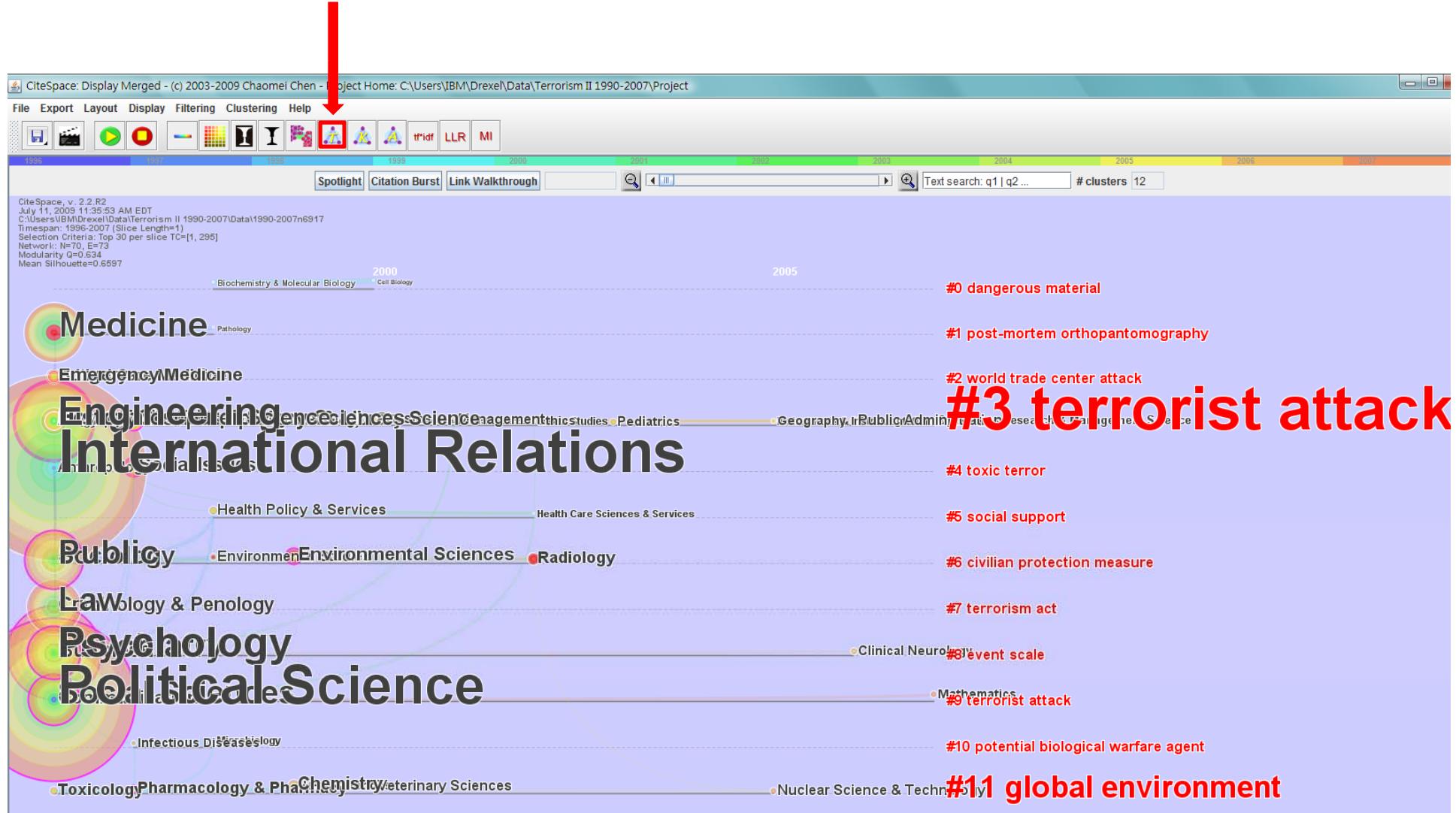
Chen, C., Zhang, J. Vogeley, M. S. (2009) Visual analysis of scientific discoveries and knowledge diffusion. *Proceedings of the 12th International Conference on Scientometrics and Informetrics (ISSI 2009)*. July 14-17, 2009. Rio de Janeiro, Brazil.



Chen, C., Zhang, J., Vogeley, M. S. (2009)
 Mapping the global impact of Sloan Digital Sky Survey. *IEEE Intelligent Systems*.

CiteSpace, v. 2.2.R2
July 11, 2009 11:22:53 AM EDT
C:\Users\IBM\Brexel\MyData\SDSS 1994-2008\Data
Timespan: 1994-2008 (Slice Length=1)
Selection Criteria: Top 30 per slice TC=[0, 1755]
Network: N=111, E=94
Modularity Q=0.7813
Mean Silhouette=0.7235





Questions and discussions