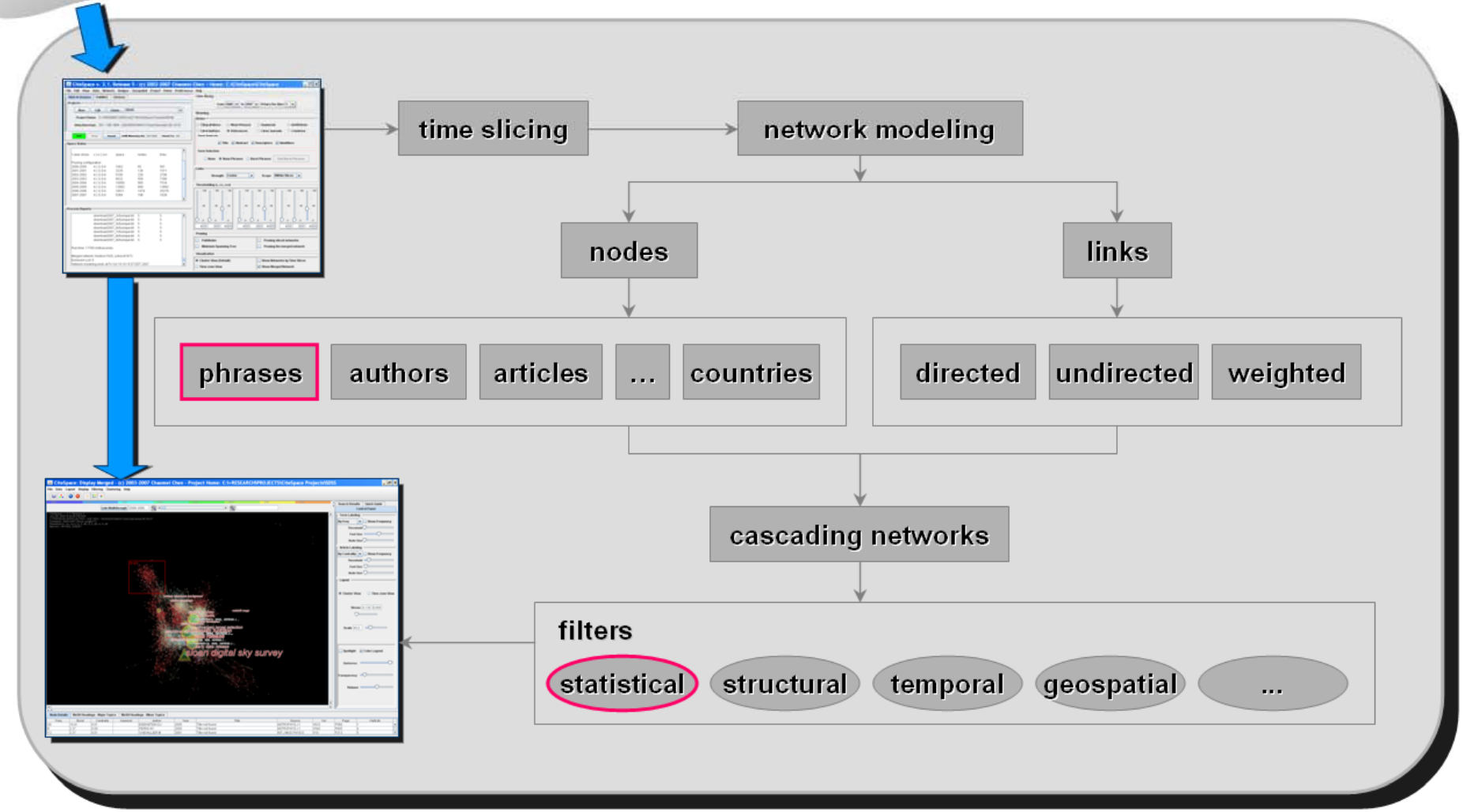
Workshop 2

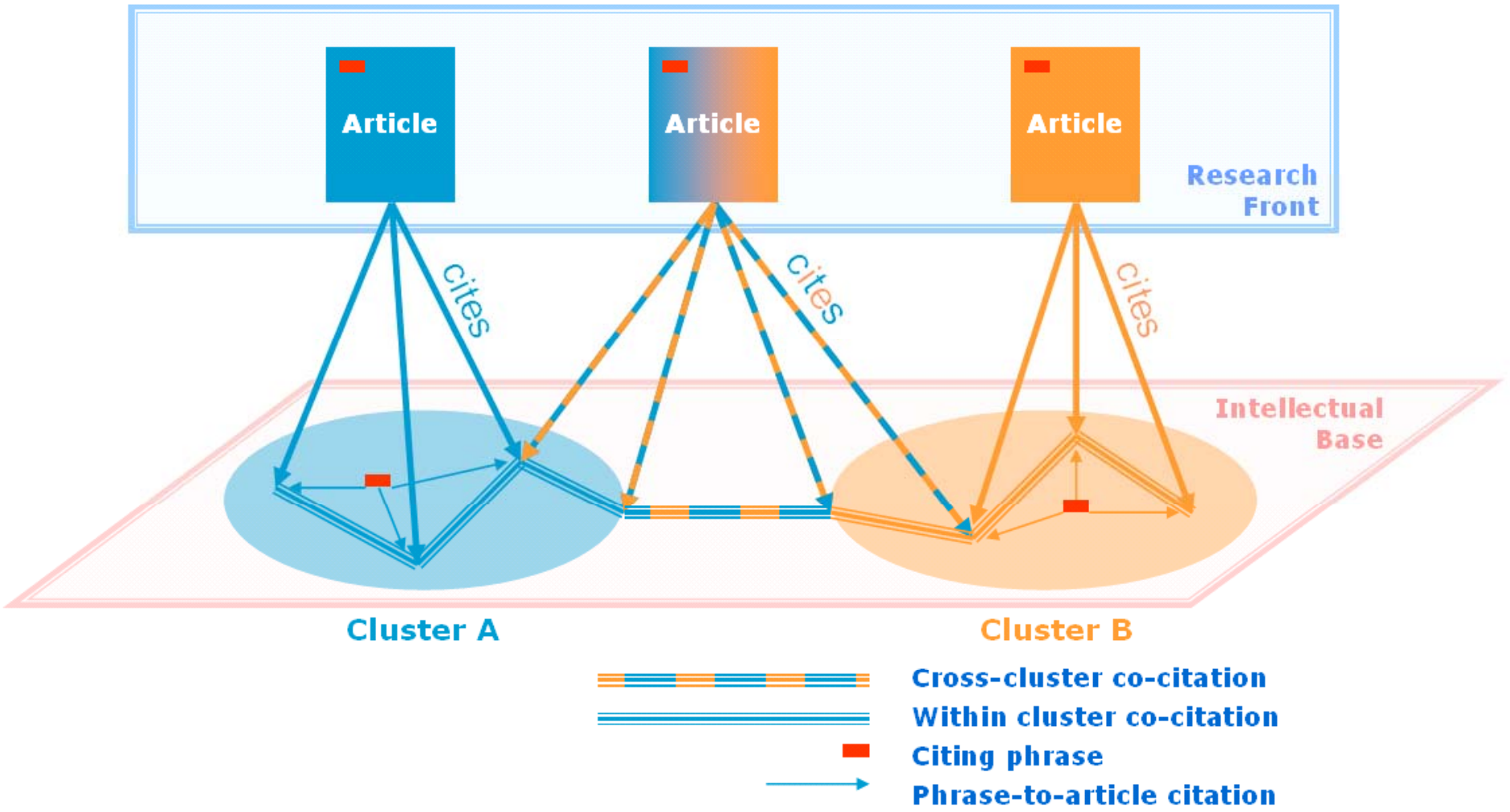
Visualizing and Analyzing Scientific Literature with *CiteSpace*

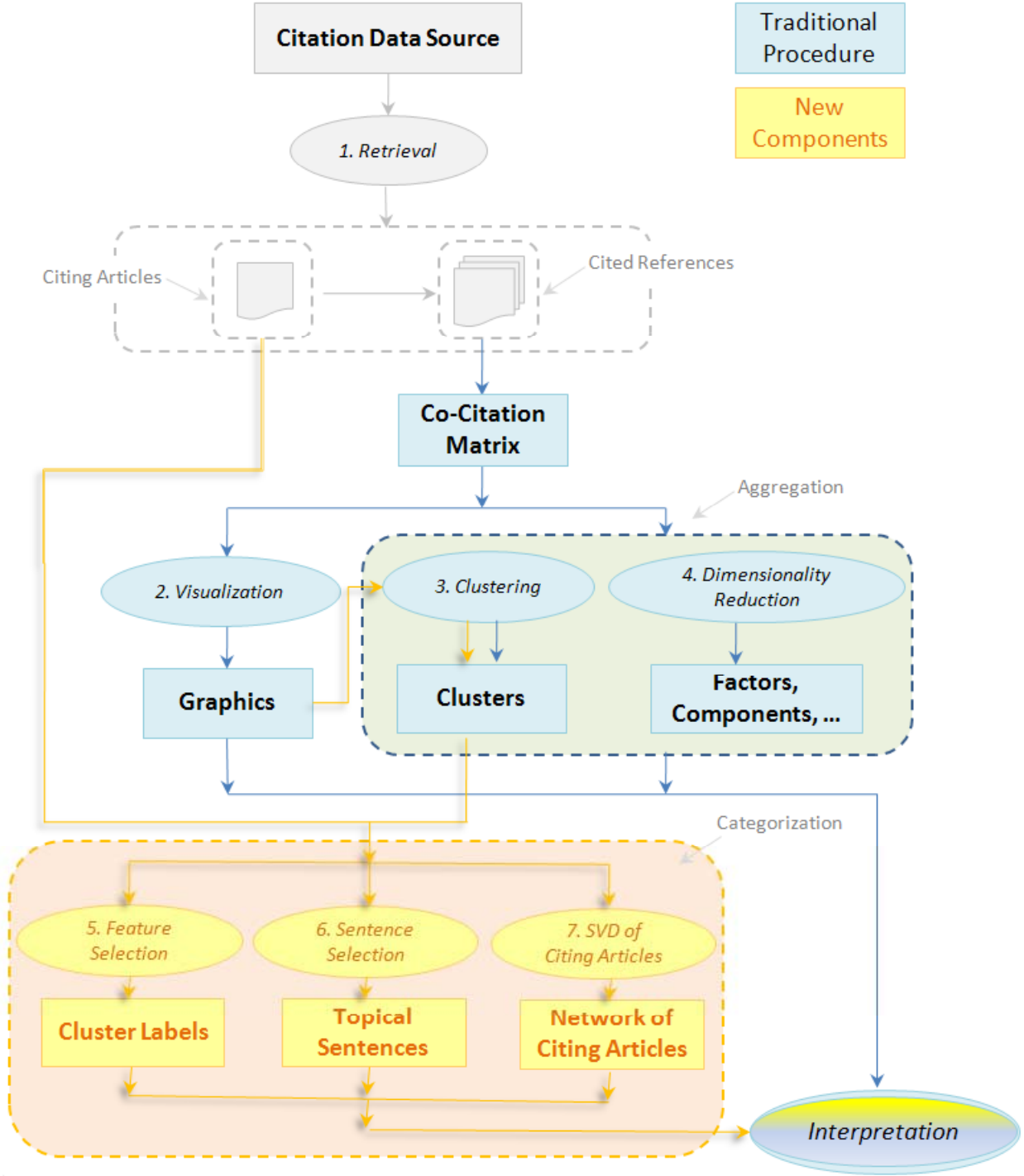
Chaomei Chen
College of Information Science and Technology
Drexel University
Email: chaomei.chen@cis.drexel.edu

OVERVIEW:
THE DESIGN OF CITESPACE

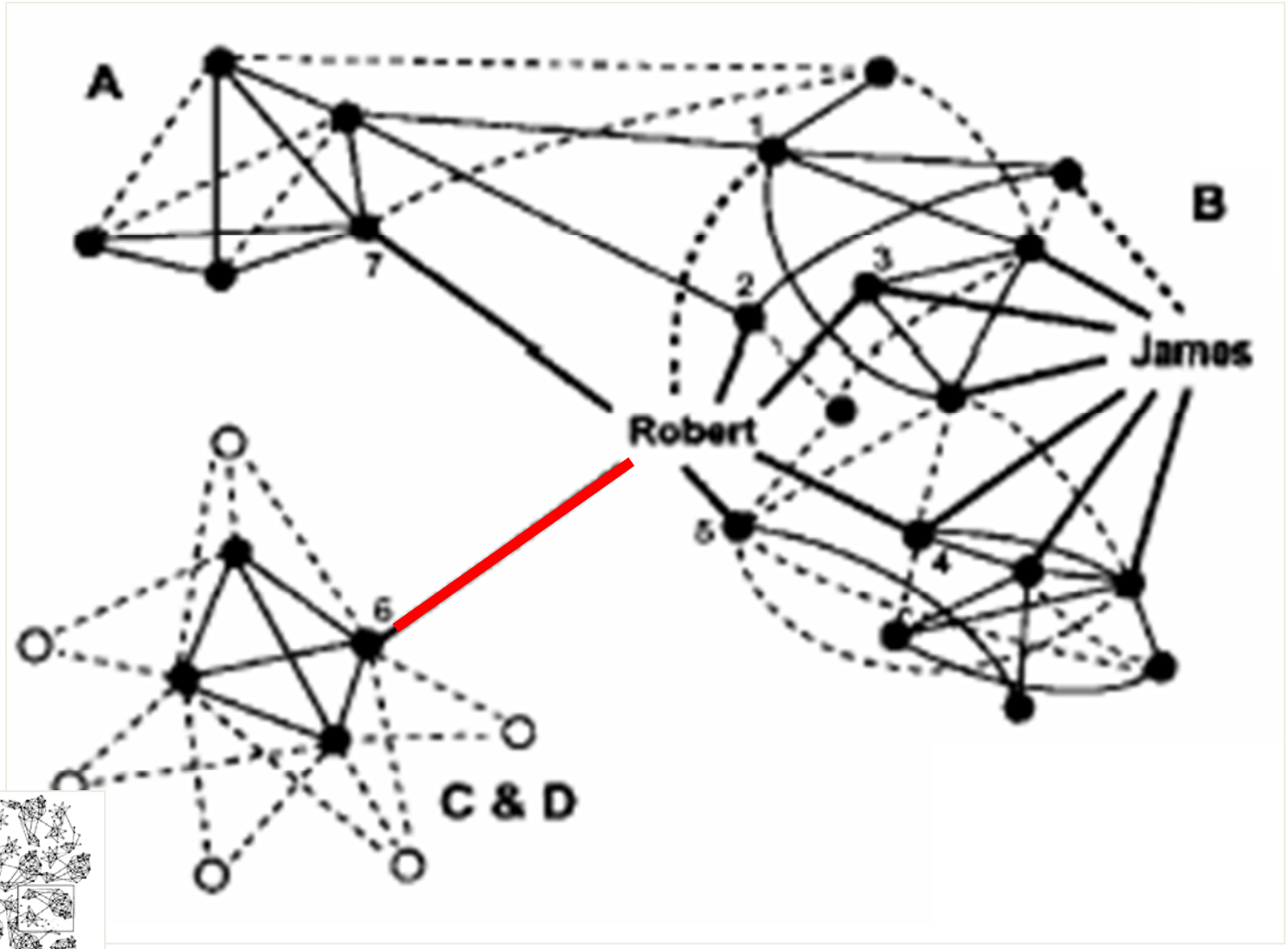
CiteSpace

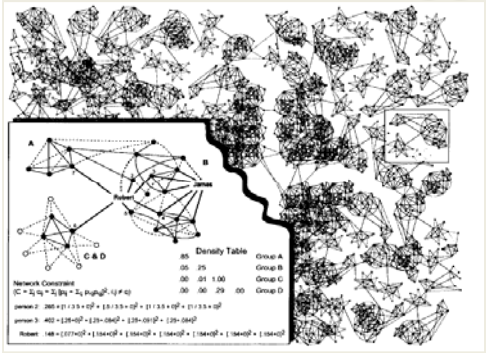

time slicing
network modeling
nodes
links
phrases
authors
articles
...
countries
directed
undirected
weighted
cascading networks
filters
statistical
structural
temporal
geospatial
...


Article
Article
Article
Research Front
cites
cites
cites
Intellectual Base
Cluster A
Cluster B
Cross-cluster co-citation
Within cluster co-citation
Citing phrase
Phrase-to-article citation

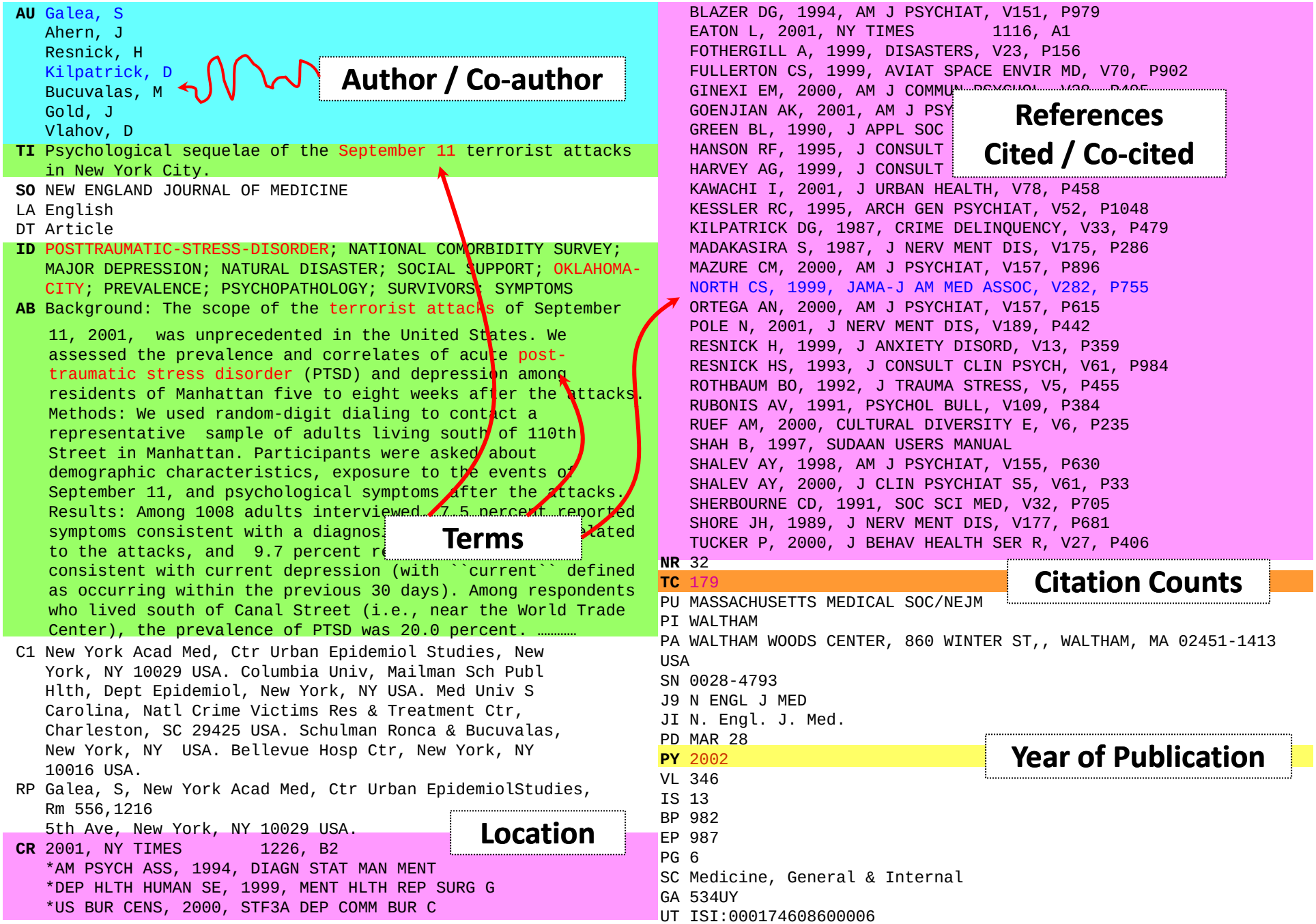

Citation Data Source
Traditional Procedure
New Components
1. Retrieval
Citing Articles
Cited References
Co-Citation Matrix
Aggregation
2. Visualization
3. Clustering
4. Dimensionality Reduction
Graphics
Clusters
Factors, Components, ...
Categorization
5. Feature Selection
6. Sentence Selection
7. SVD of Citing Articles
Cluster Labels
Topical Sentences
Network of Citing Articles
Interpretation

Structural Holes in Social Networks





http://www.wellcome.ac.uk/News/Media-office/Press-releases/2008/WTX050156.htm
An example of the areas to which we might allocate more memory when watching a relay race. Image courtesy of actionplus.co.uk.



AU Galea, S
Ahern, J
Resnick, H
Kilpatrick, D
Bucuvalas, M
Gold, J
Vlahov, D

Author / Co-author

TI Psychological sequelae of the September 11 terrorist attacks in New York City.

SO NEW ENGLAND JOURNAL OF MEDICINE
LA English
DT Article

ID POSTTRAUMATIC-STRESS-DISORDER; NATIONAL COMORBIDITY SURVEY; MAJOR DEPRESSION; NATURAL DISASTER; SOCIAL SUPPORT; OKLAHOMA-CITY; PREVALENCE; PSYCHOPATHOLOGY; SURVIVORS; SYMPTOMS

AB Background: The scope of the terrorist attacks of September 11, 2001, was unprecedented in the United States. We assessed the prevalence and correlates of acute post-traumatic stress disorder (PTSD) and depression among residents of Manhattan five to eight weeks after the attacks. Methods: We used random-digit dialing to contact a representative sample of adults living south of 110th Street in Manhattan. Participants were asked about demographic characteristics, exposure to the events of September 11, and psychological symptoms after the attacks. Results: Among 1008 adults interviewed, 7.5 percent reported symptoms consistent with a diagnos[illegible]lated to the attacks, and 9.7 percent re[illegible] consistent with current depression (with ``current`` defined as occurring within the previous 30 days). Among respondents who lived south of Canal Street (i.e., near the World Trade Center), the prevalence of PTSD was 20.0 percent. ……

Terms

C1 New York Acad Med, Ctr Urban Epidemiol Studies, New York, NY 10029 USA. Columbia Univ, Mailman Sch Publ Hlth, Dept Epidemiol, New York, NY USA. Med Univ S Carolina, Natl Crime Victims Res & Treatment Ctr, Charleston, SC 29425 USA. Schulman Ronca & Bucuvalas, New York, NY USA. Bellevue Hosp Ctr, New York, NY 10016 USA.
RP Galea, S, New York Acad Med, Ctr Urban EpidemiolStudies, Rm 556,1216
5th Ave, New York, NY 10029 USA.

Location

CR 2001, NY TIMES 1226, B2
*AM PSYCH ASS, 1994, DIAGN STAT MAN MENT
*DEP HLTH HUMAN SE, 1999, MENT HLTH REP SURG G
*US BUR CENS, 2000, STF3A DEP COMM BUR C
BLAZER DG, 1994, AM J PSYCHIAT, V151, P979
EATON L, 2001, NY TIMES 1116, A1
FOTHERGILL A, 1999, DISASTERS, V23, P156
FULLERTON CS, 1999, AVIAT SPACE ENVIR MD, V70, P902
GINEXI EM, 2000, AM J COMMUN PSYCHOL, V28, P405
GOENJIAN AK, 2001, AM J PSY[illegible]
GREEN BL, 1990, J APPL SOC [illegible]
HANSON RF, 1995, J CONSULT [illegible]
HARVEY AG, 1999, J CONSULT [illegible]
KAWACHI I, 2001, J URBAN HEALTH, V78, P458
KESSLER RC, 1995, ARCH GEN PSYCHIAT, V52, P1048
KILPATRICK DG, 1987, CRIME DELINQUENCY, V33, P479
MADAKASIRA S, 1987, J NERV MENT DIS, V175, P286
MAZURE CM, 2000, AM J PSYCHIAT, V157, P896
NORTH CS, 1999, JAMA-J AM MED ASSOC, V282, P755
ORTEGA AN, 2000, AM J PSYCHIAT, V157, P615
POLE N, 2001, J NERV MENT DIS, V189, P442
RESNICK H, 1999, J ANXIETY DISORD, V13, P359
RESNICK HS, 1993, J CONSULT CLIN PSYCH, V61, P984
ROTHBAUM BO, 1992, J TRAUMA STRESS, V5, P455
RUBONIS AV, 1991, PSYCHOL BULL, V109, P384
RUEF AM, 2000, CULTURAL DIVERSITY E, V6, P235
SHAH B, 1997, SUDAAN USERS MANUAL
SHALEV AY, 1998, AM J PSYCHIAT, V155, P630
SHALEV AY, 2000, J CLIN PSYCHIAT S5, V61, P33
SHERBOURNE CD, 1991, SOC SCI MED, V32, P705
SHORE JH, 1989, J NERV MENT DIS, V177, P681
TUCKER P, 2000, J BEHAV HEALTH SER R, V27, P406

References Cited / Co-cited

NR 32
TC 179

Citation Counts

PU MASSACHUSETTS MEDICAL SOC/NEJM
PI WALTHAM
PA WALTHAM WOODS CENTER, 860 WINTER ST,, WALTHAM, MA 02451-1413 USA
SN 0028-4793
J9 N ENGL J MED
JI N. Engl. J. Med.
PD MAR 28
PY 2002

Year of Publication

VL 346
IS 13
BP 982
EP 987
PG 6
SC Medicine, General & Internal
GA 534UY
UT ISI:000174608600006
ER

co-occurring burst terms

AU Galea, S
Ahern, J
Resnick, H
Kilpatrick, D
Bucuvalas, M
Gold, J
Vlahov, D

co-authorship

TI Psychological sequelae of the September 11 terrorist attacks in New York City.

SO NEW ENGLAND JOURNAL OF MEDICINE

LA English

DT Article

ID POSTTRAUMATIC-STRESS-DISORDER; NATIONAL COMORBIDITY SURVEY; MAJOR DEPRESSION; NATURAL DISASTER; SOCIAL SUPPORT; OKLAHOMA-CITY; PREVALENCE; PSYCHOPATHOLOGY; SURVIVORS; SYMPTOMS

AB Background: The scope of the terrorist attacks of September 11, 2001, was unprecedented in the United States. We assessed the prevalence and correlates of acute post-traumatic stress disorder (PTSD) and depression among residents of Manhattan five to eight weeks after the attacks. Methods: We used random-digit dialing to contact a representative sample of adults living south of 110th Street in Manhattan. Participants were asked about demographic characteristics, exposure to the events of September 11, and psychological symptoms after the attacks. Results: Among 1008 adults interviewed, 7.5 percent reported symptoms consistent with a diagnosis of current PTSD related to the attacks, and 9.7 percent reported symptoms consistent with current depression (with ``current`` defined as occurring within the previous 30 days). Among respondents who lived south of Canal Street (i.e., near the World Trade Center), the prevalence of PTSD was 20.0 percent. ……

```
CR 2001, NY TIMES         1226, B2
   *AM PSYCH ASS, 1994, DIAGN STAT MAN MENT
   *DEP HLTH HUMAN SE, 1999, MENT HLTH REP SURG G
   *US BUR CENS, 2000, STF3A DEP COMM BUR C
   BLAZER DG, 1994, AM J PSYCHIAT, V151, P979
   EATON L, 2001, NY TIMES         1116, A1
   FOTHERGILL A, 1999, DISASTERS, V23, P156
   FULLERTON CS, 1999, AVIAT SPACE ENVIR MD, V70, P902
   GINEXI EM, 2000, AM J COMMUN PSYCHOL, V28, P495
   GOENJIAN AK, 2001, AM J PSYCHIAT, V158, P788
   GREEN BL, 1990, J APPL SOC PSYCHOL, V20, P1033
   HANSON RF, 1995, J CONSULT CLIN PSYCH, V63, P987
   HARVEY AG, 1999, J CONSULT CLIN PSYCH, V67, P985
   KAWACHI I, 2001, J URBAN HEALTH, V78, P458
   KESSLER RC, 1995, ARCH GEN PSYCHIAT, V52, P1048
   KILPATRICK DG, 1987, CRIME DELINQUENCY, V33, P479
   MADAKASIRA S, 1987, J NERV MENT DIS, V175, P286
   MAZURE CM, 2000, AM J PSYCHIAT, V157, P896
   NORTH CS, 1999, JAMA-J AM MED ASSOC, V282, P755
   ORTEGA AN, 2000, AM J PSYCHIAT, V157, P615
   POLE N, 2001, J NERV MENT DIS, V189, P442
   RESNICK H, 1999, J ANXIETY DISORD, V13, P359
   RESNICK HS, 1993, J CONSULT CLIN PSYCH, V61, P984
   ROTHBAUM BO, 1992, J TRAUMA STRESS, V5, P455
   RUBONIS AV, 1991, PSYCHOL BULL, V109, P384
   RUEF AM, 2000, CULTURAL DIVERSITY E, V6, P235
   SHAH B, 1997, SUDAAN USERS MANUAL
   SHALEV AY, 1998, AM J PSYCHIAT, V155, P630
   SHALEV AY, 2000, J CLIN PSYCHIAT S5, V61, P33
   SHERBOURNE CD, 1991, SOC SCI MED, V32, P705
   SHORE JH, 1989, J NERV MENT DIS, V177, P681
   TUCKER P, 2000, J BEHAV HEALTH SER R, V27, P406
```

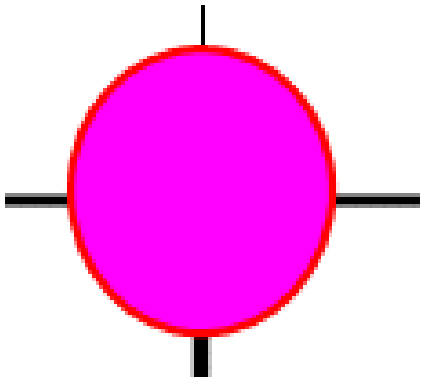
author co-citation

document co-citation

journal co-citation

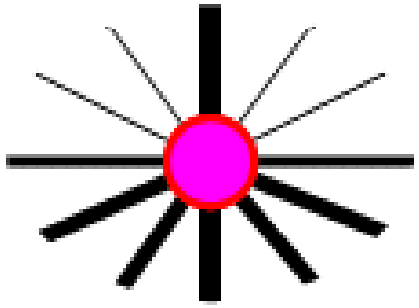
ACA/DCA/JCA

CiteSpace



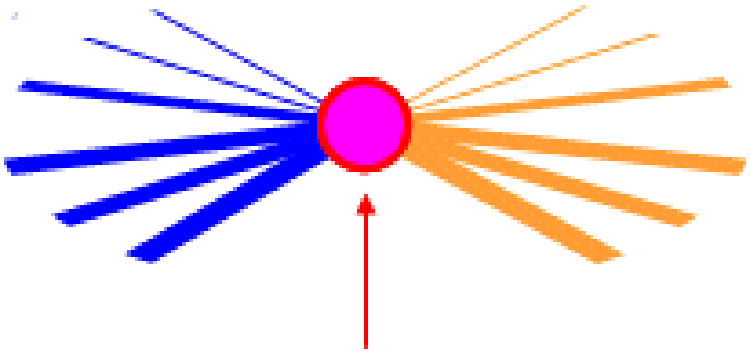
Landmark node

large radius



Hub node

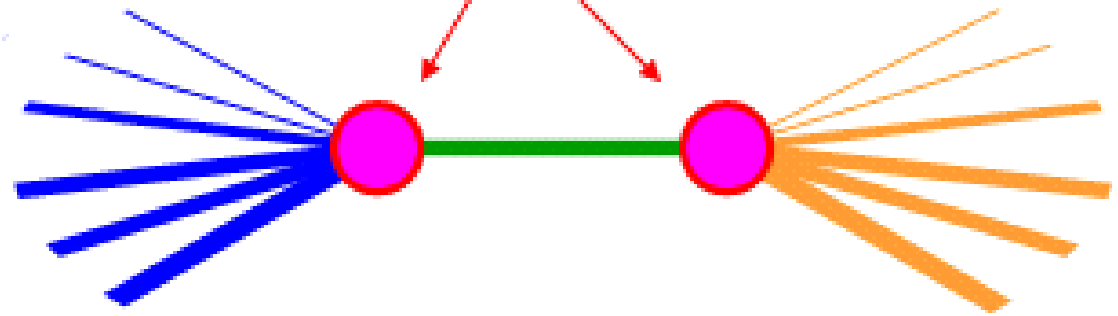
large degree



Pivot node

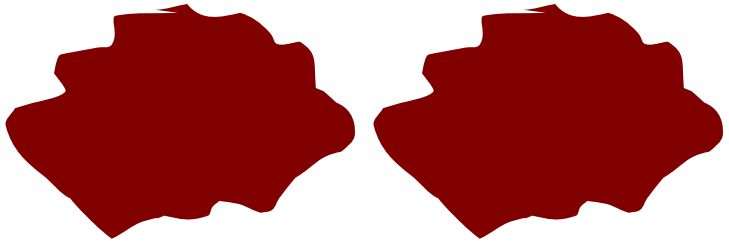
exclusive joints of clusters

or network patches

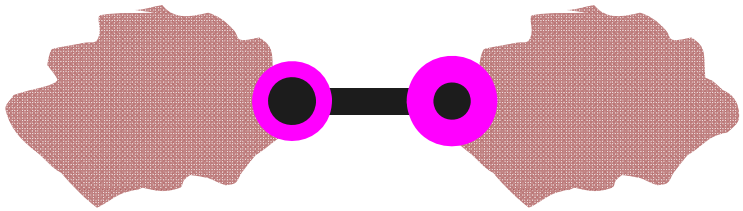


Expected Patterns

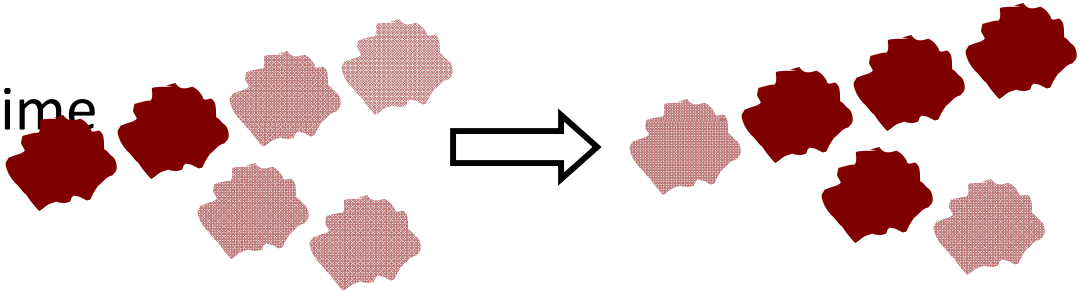
- Thematic grouping



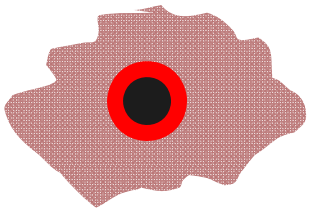
- Intellectual turning points

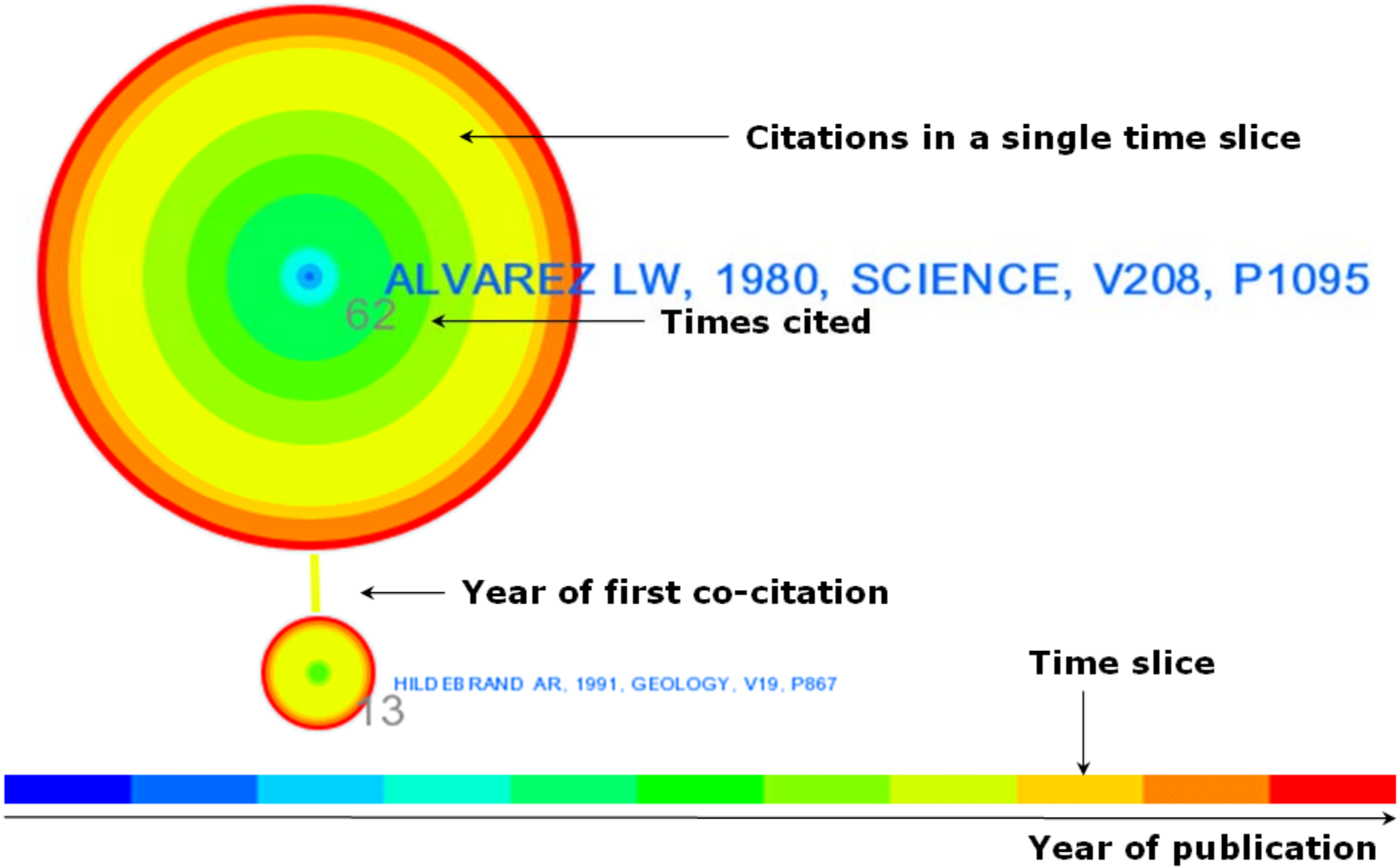


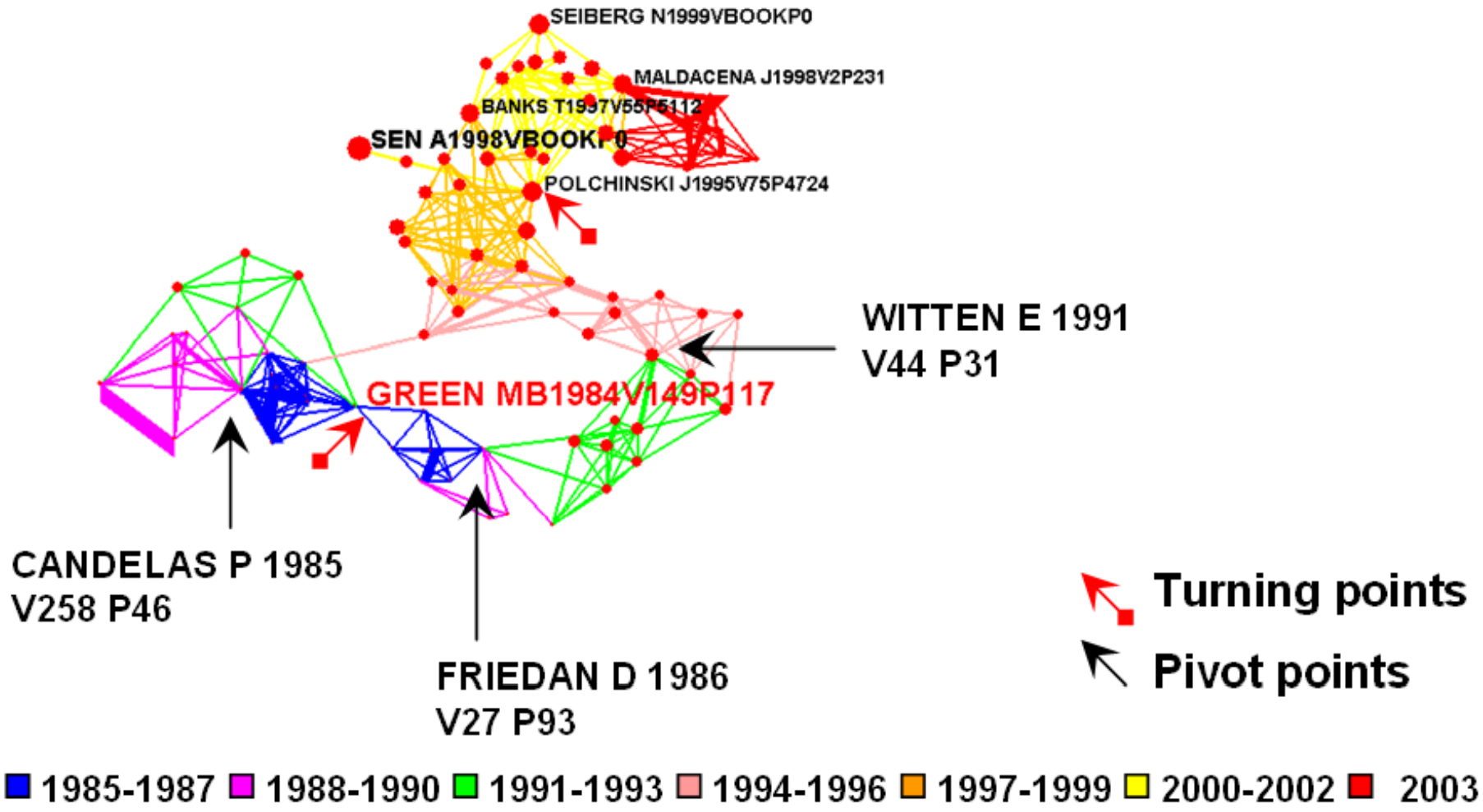
- Thematic change over time

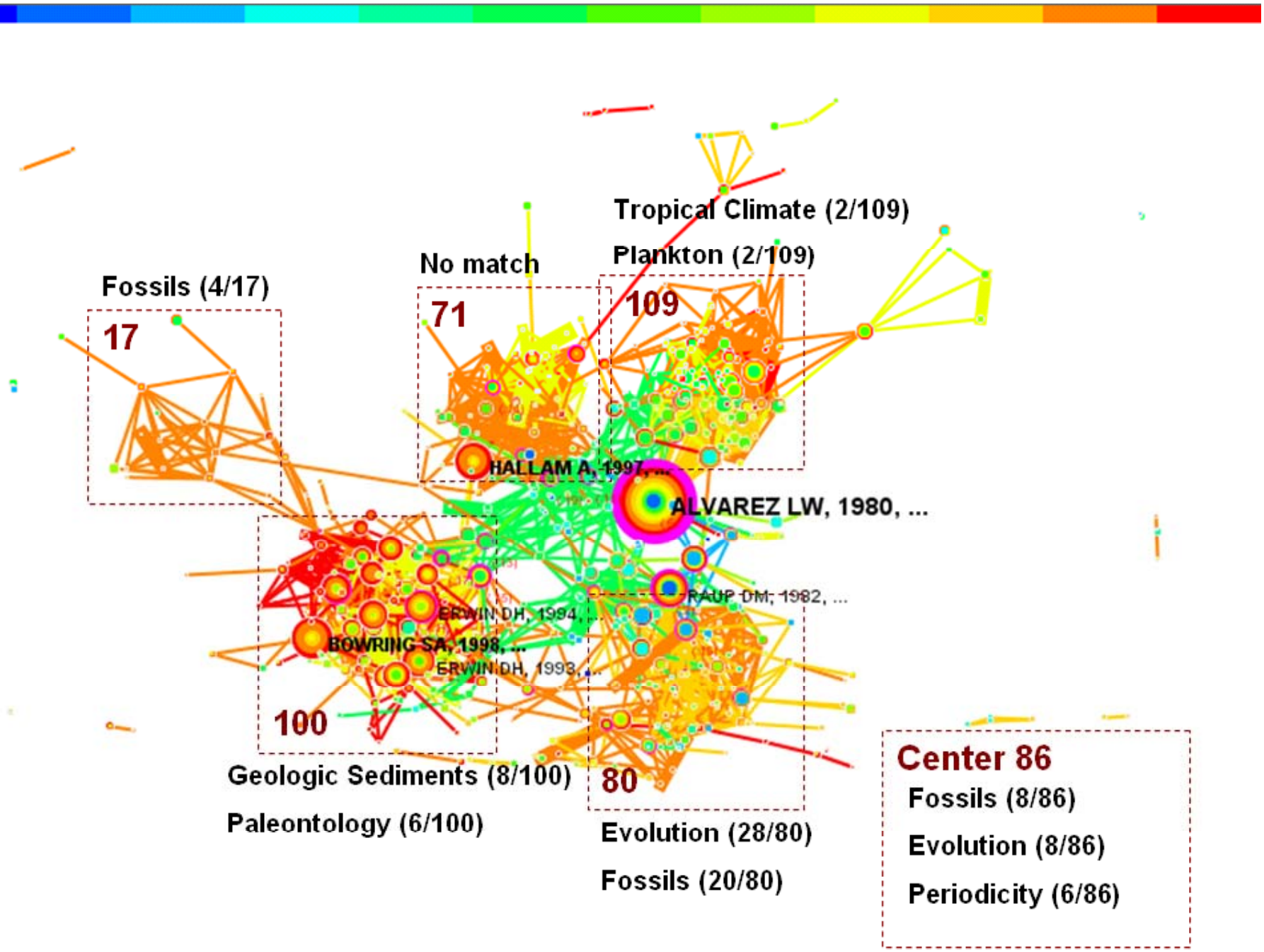


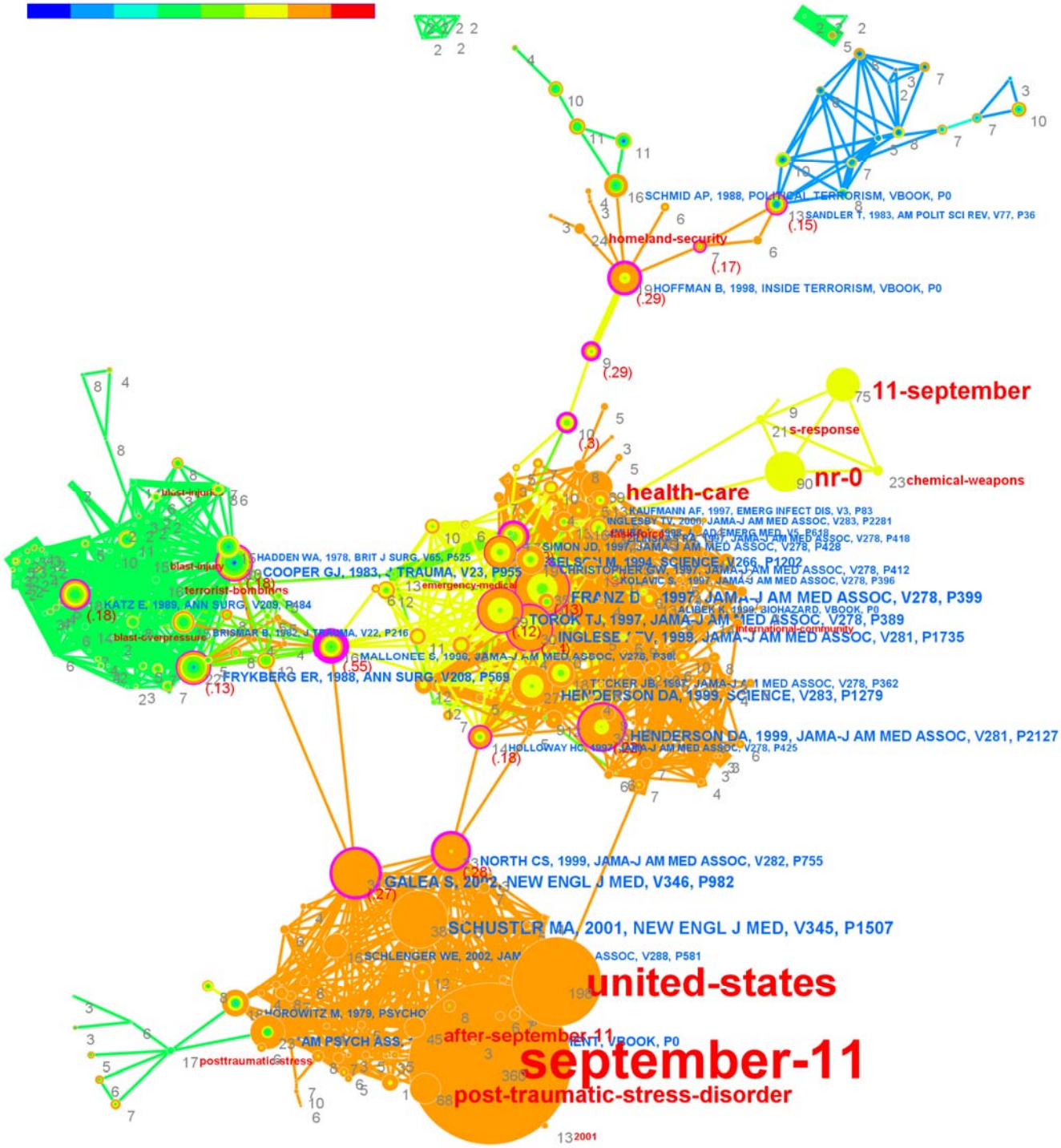
- Abrupt changes associated with triggers

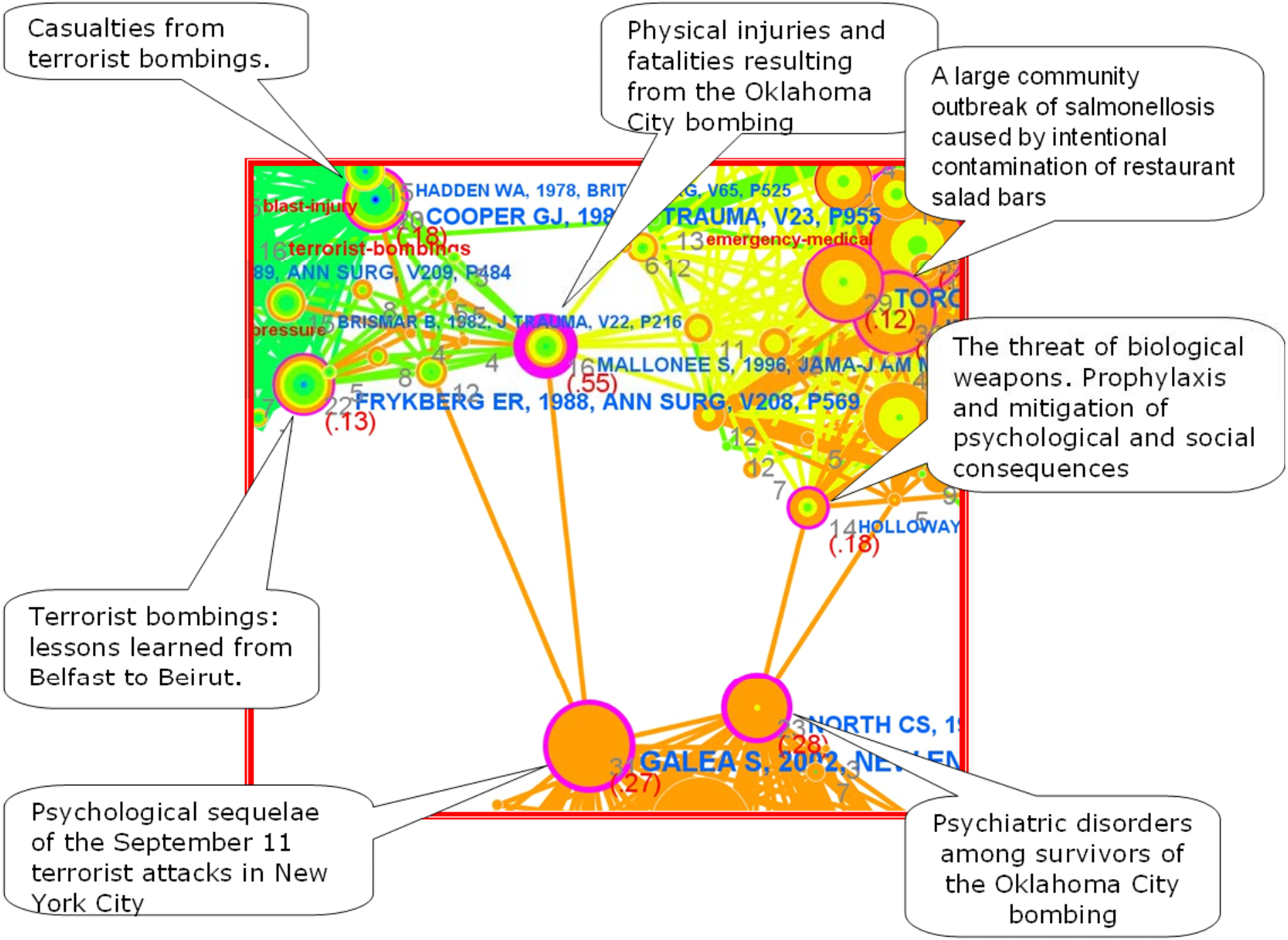


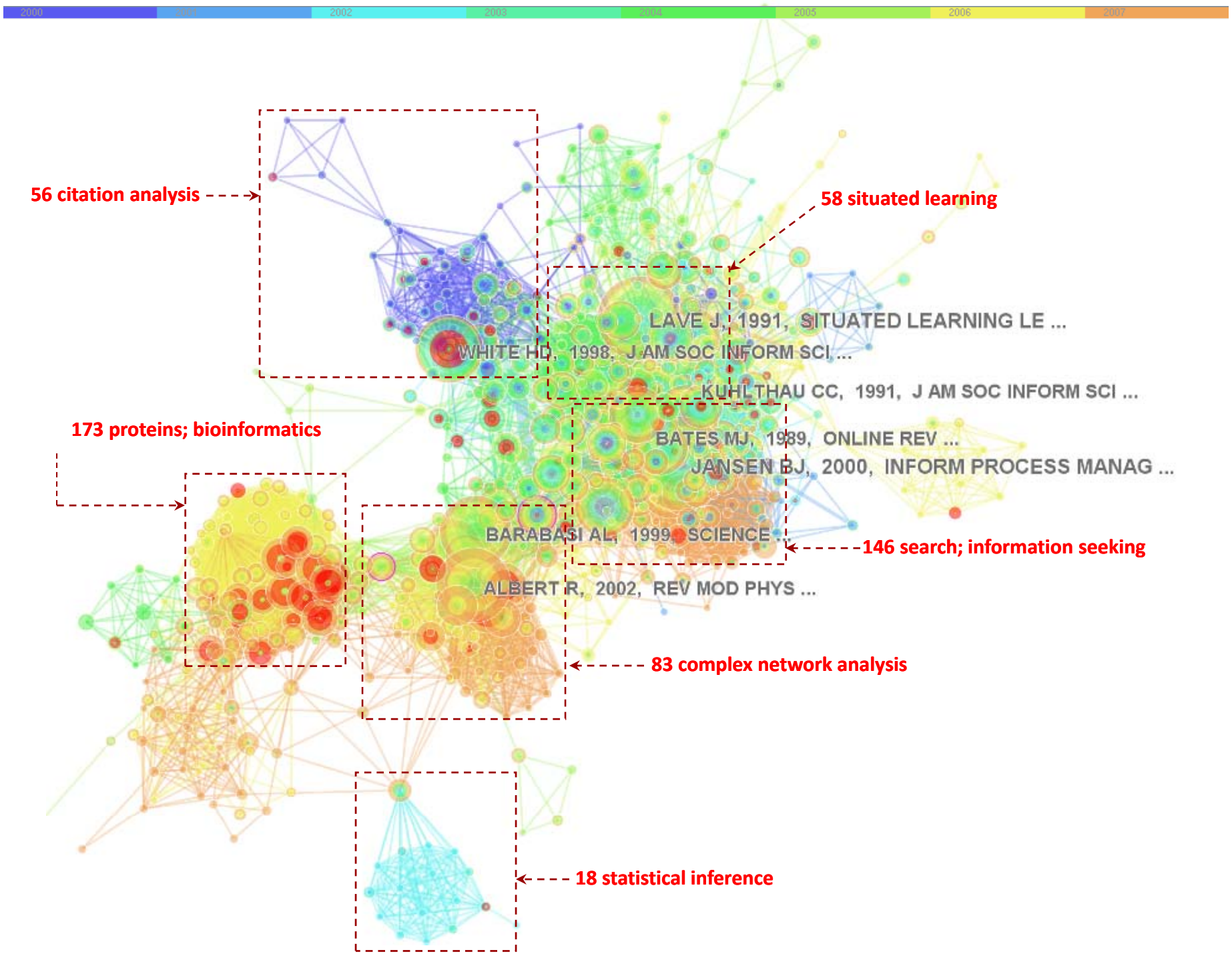

Citations in a single time slice
ALVAREZ LW, 1980, SCIENCE, V208, P1095
62
Times cited
Year of first co-citation
HILDEBRAND AR, 1991, GEOLOGY, V19, P867
13
Time slice
Year of publication

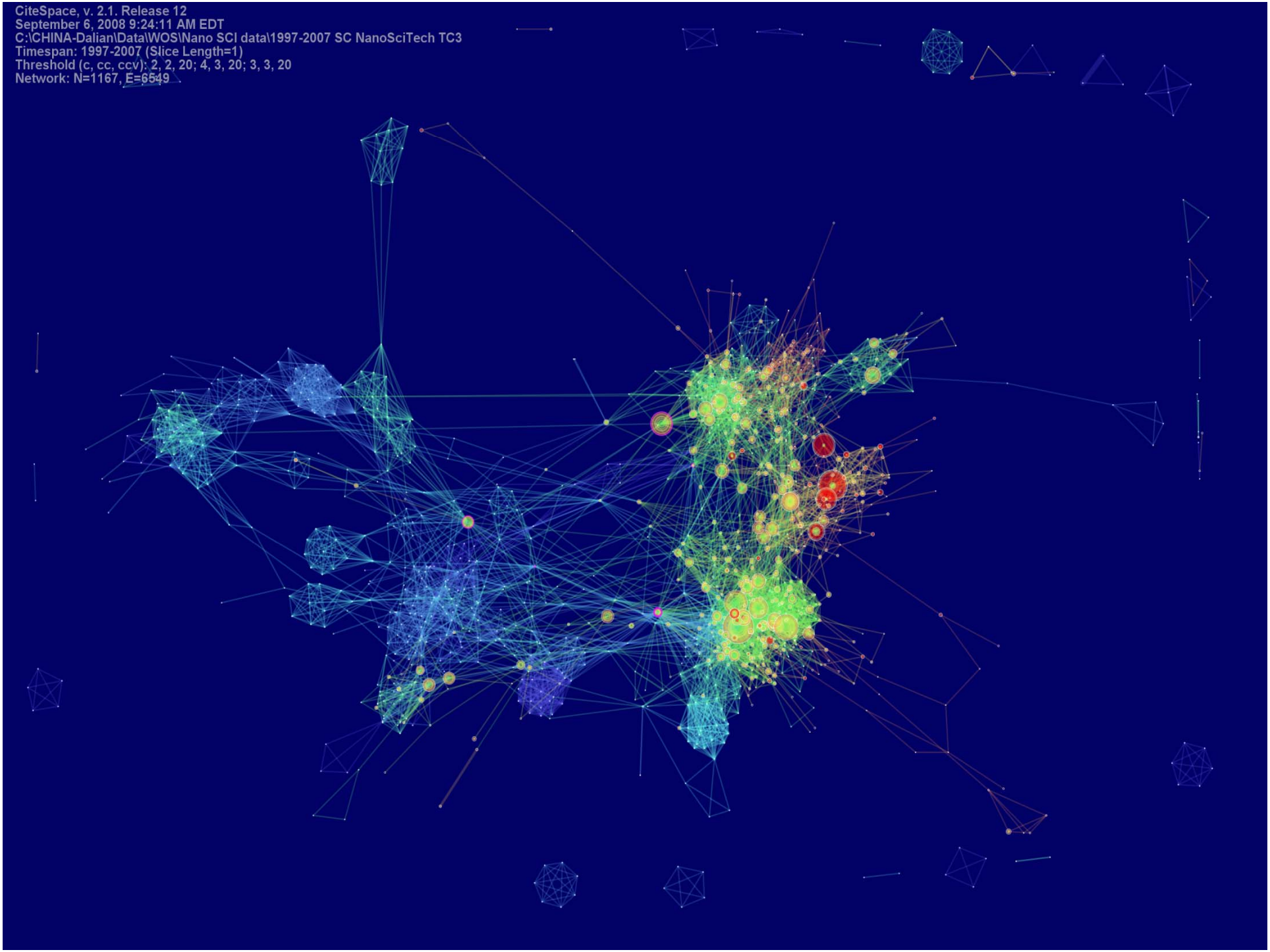

SEIBERG N1999VBOOKP0
MALDACENA J1998V2P231
BANKS T1997V55P5112
SEN A1998VBOOKP0
POLCHINSKI J1995V75P4724
WITTEN E 1991
V44 P31
GREEN MB1984V149P117
CANDELAS P 1985
V258 P46
FRIEDAN D 1986
V27 P93
Turning points
Pivot points
1985-1987
1988-1990
1991-1993
1994-1996
1997-1999
2000-2002
2003

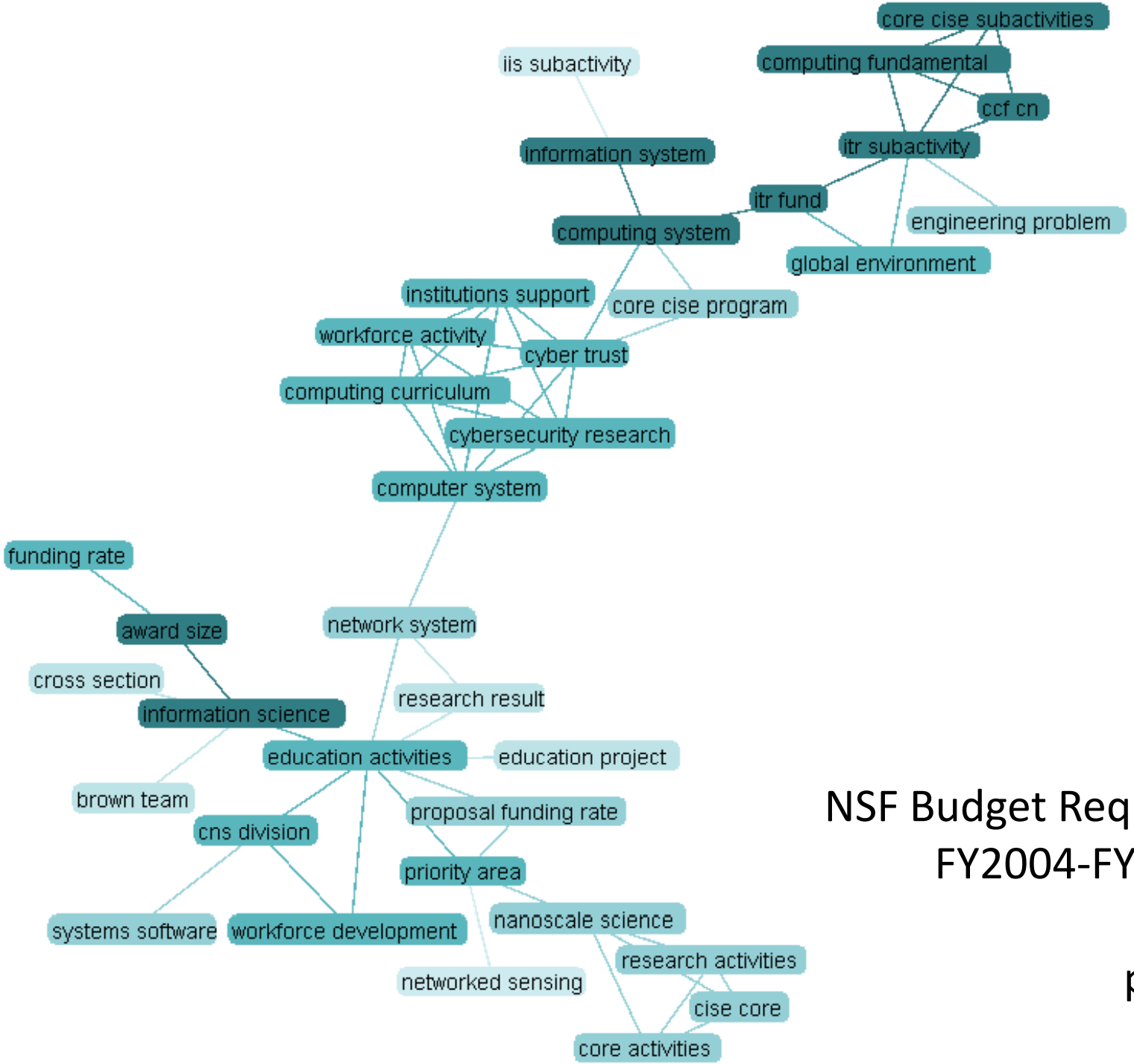

Fossils (4/17)
17
No match
71
Tropical Climate (2/109)
Plankton (2/109)
109
HALLAM A, 1997, ...
ALVAREZ LW, 1980, ...
RAUP DM, 1982, ...
ERWIN DH, 1994, ...
BOWRING SA, 1998, ...
ERWIN DH, 1993, ...
100
Geologic Sediments (8/100)
Paleontology (6/100)
80
Evolution (28/80)
Fossils (20/80)
Center 86
Fossils (8/86)
Evolution (8/86)
Periodicity (6/86)


SCHMID AP, 1988, POLITICAL TERRORISM, VBOOK, P0
SANDLER T, 1983, AM POLIT SCI REV, V77, P36
(.15)
homeland-security
(.17)
HOFFMAN B, 1998, INSIDE TERRORISM, VBOOK, P0
(.29)
(.29)
11-september
s-response
nr-0
chemical-weapons
health-care
(.3)
KAUFMANN AF, 1997, EMERG INFECT DIS, V3, P83
INGLESBY TV, 2000, JAMA-J AM MED ASSOC, V283, P2281
SIMON JD, 1997, JAMA-J AM MED ASSOC, V278, P428
SELSCH M, 1994, SCIENCE, V266, P1202
CHRISTOPHER GW, 1997, JAMA-J AM MED ASSOC, V278, P412
KOLAVIC S, 1997, JAMA-J AM MED ASSOC, V278, P396
FRANZ D, 1997, JAMA-J AM MED ASSOC, V278, P399
ALIBEK K, 1999, BIOHAZARD, VBOOK, P0
TOROK TJ, 1997, JAMA-J AM MED ASSOC, V278, P389
INGLESE J, 1999, JAMA-J AM MED ASSOC, V281, P1735
international-community
TUCKER JB, 1997, JAMA-J AM MED ASSOC, V278, P362
HENDERSON DA, 1999, SCIENCE, V283, P1279
HENDERSON DA, 1999, JAMA-J AM MED ASSOC, V281, P2127
HOLLOWAY HC, 1997, JAMA-J AM MED ASSOC, V278, P425
(.18)
HADDEN WA, 1978, BRIT J SURG, V65, P525
COOPER GJ, 1983, J TRAUMA, V23, P955
blast-injury
terrorist-bombings
KATZ E, 1989, ANN SURG, V209, P484
(.18)
BRISMAR B, 1982, J TRAUMA, V22, P216
blast-overpressure
MALLONEE S, 1996, JAMA-J AM MED ASSOC, V276, P38
(.55)
FRYKBERG ER, 1988, ANN SURG, V208, P569
(.13)
emergency-medical
(.12)
(.4)
NORTH CS, 1999, JAMA-J AM MED ASSOC, V282, P755
(.28)
GALEA S, 2002, NEW ENGL J MED, V346, P982
(.27)
SCHUSTER MA, 2001, NEW ENGL J MED, V345, P1507
SCHLENGER WE, 2002, JAM
ASSOC, V288, P581
united-states
GROWITZ M, 1979, PSYCHO
AM PSYCH ASS,
IENT, VBOOK, P0
after-september-11
september-11
posttraumatic-stress
post-traumatic-stress-disorder
2001


Casualties from terrorist bombings.
Physical injuries and fatalities resulting from the Oklahoma City bombing
A large community outbreak of salmonellosis caused by intentional contamination of restaurant salad bars
The threat of biological weapons. Prophylaxis and mitigation of psychological and social consequences
Terrorist bombings: lessons learned from Belfast to Beirut.
Psychological sequelae of the September 11 terrorist attacks in New York City
Psychiatric disorders among survivors of the Oklahoma City bombing
HADDEN WA, 1978, BRIT
G, V65, P525
blast-injury
COOPER GJ, 198
TRAUMA, V23, P955
(.18)
terrorist-bombings
emergency-medical
89, ANN SURG, V209, P484
TORC
(.12)
pressure
BRISMAR B, 1982, J TRAUMA, V22, P216
MALLONEE S, 1996, JAMA-J AM
(.55)
FRYKBERG ER, 1988, ANN SURG, V208, P569
(.13)
HOLLOWAY
(.18)
NORTH CS, 1
(.28)
GALEA S, 2002, NE
(.27)


2000
2001
2002
2003
2004
2005
2006
2007
56 citation analysis
58 situated learning
LAVE J, 1991, SITUATED LEARNING LE ...
WHITE HD, 1998, J AM SOC INFORM SCI ...
KUHLTHAU CC, 1991, J AM SOC INFORM SCI ...
173 proteins; bioinformatics
BATES MJ, 1989, ONLINE REV ...
JANSEN BJ, 2000, INFORM PROCESS MANAG ...
BARABASI AL, 1999, SCIENCE ...
146 search; information seeking
ALBERT R, 2002, REV MOD PHYS ...
83 complex network analysis
18 statistical inference


CiteSpace, v. 2.1. Release 12
September 6, 2008 9:24:11 AM EDT
C:\CHINA-Dalian\Data\WOS\Nano SCI data\1997-2007 SC NanoSciTech TC3
Timespan: 1997-2007 (Slice Length=1)
Threshold (c, cc, ccv): 2, 2, 20; 4, 3, 20; 3, 3, 20
Network: N=1167, E=6549



NSF Budget Requests
FY2004-FY2008
CISE
p=0.5

Contents lists available at ScienceDirect

Journal of Informetrics

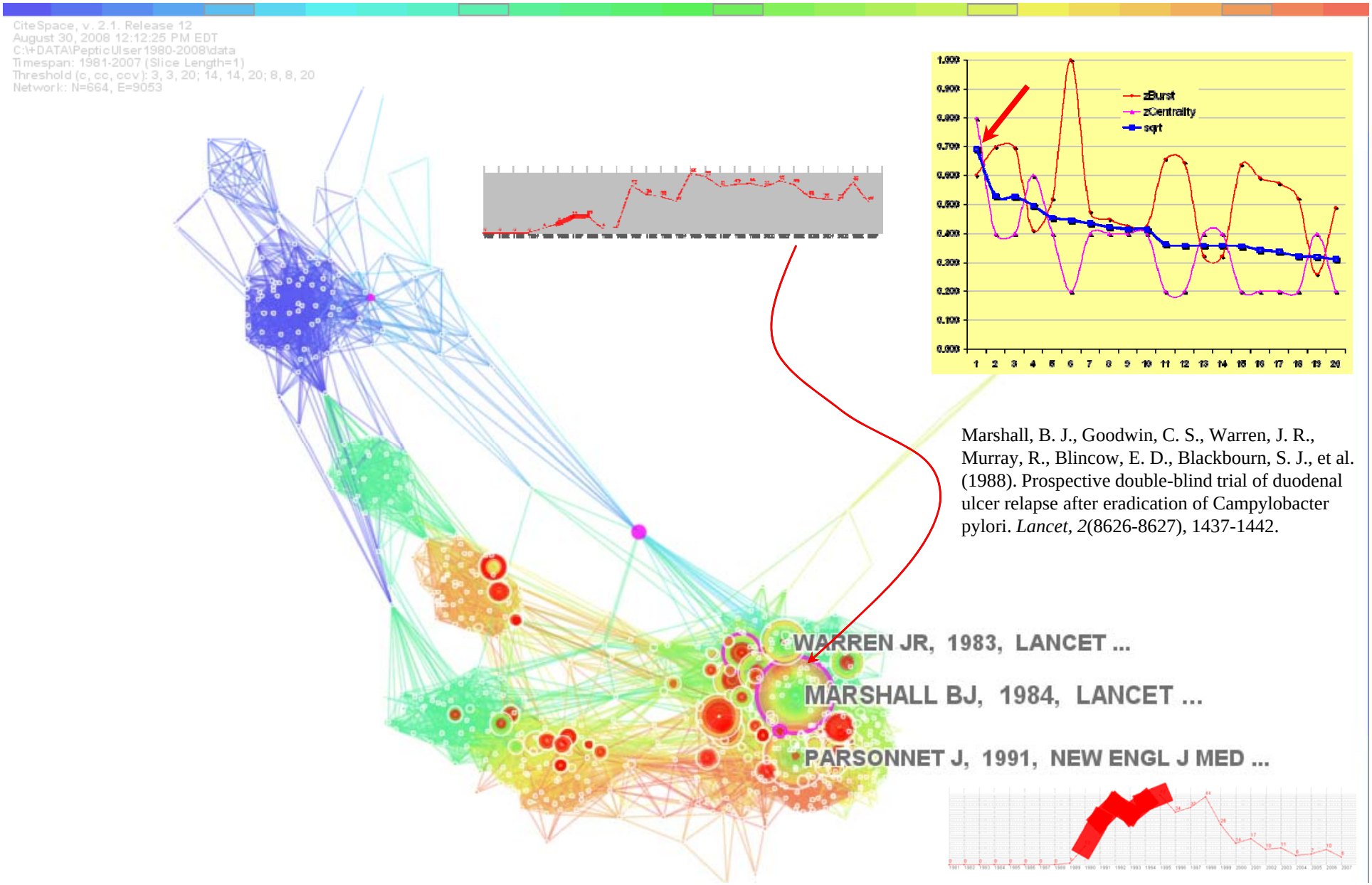
journal homepage: www.elsevier.com/locate/joi

Towards an explanatory and computational theory of scientific discovery☆

Chaomei Chen[a,b,*], Yue Chen[b], Mark Horowitz[a], Haiyan Hou[b], Zeyuan Liu[b], Donald Pellegrino[a]

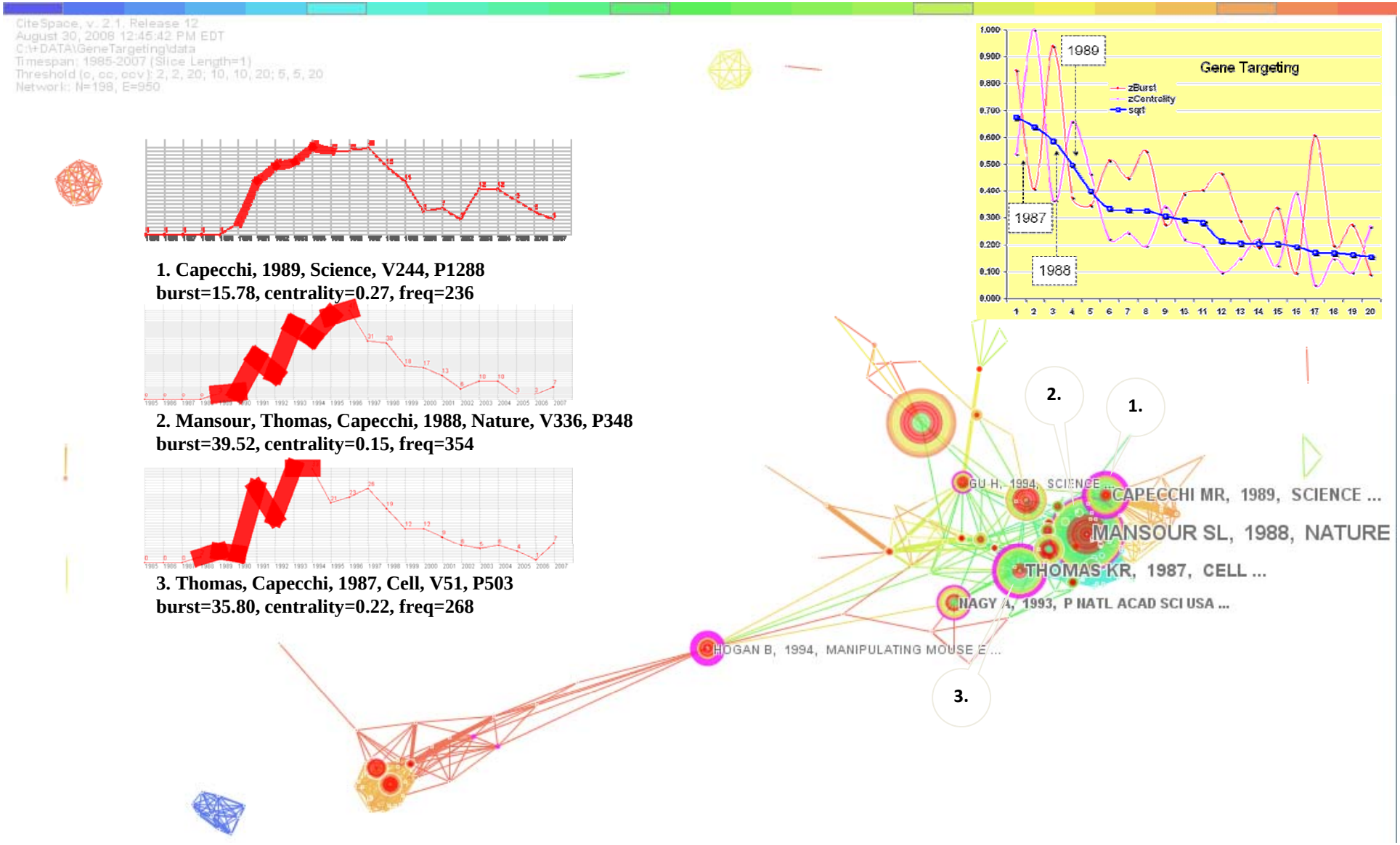
[a] College of Information Science and Technology, Drexel University, USA
[b] The WISE Lab, Dalian University of Technology, China

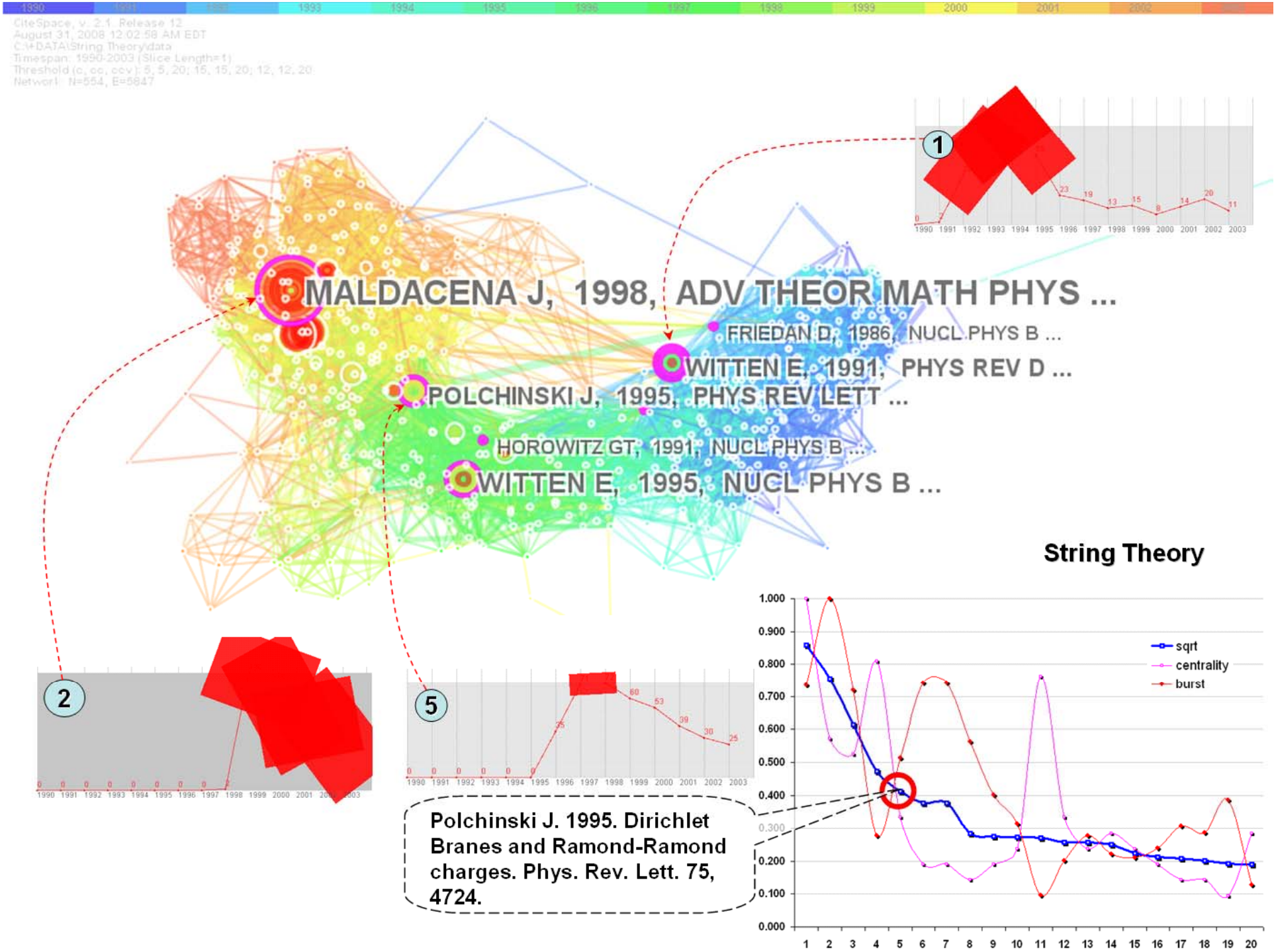
The 2005 Nobel Prize in Medicine to Barry J. Marshall and J. Robin Warren for their discovery of "the bacterium *Helicobacter pylori* and its role in gastritis and peptic ulcer disease."



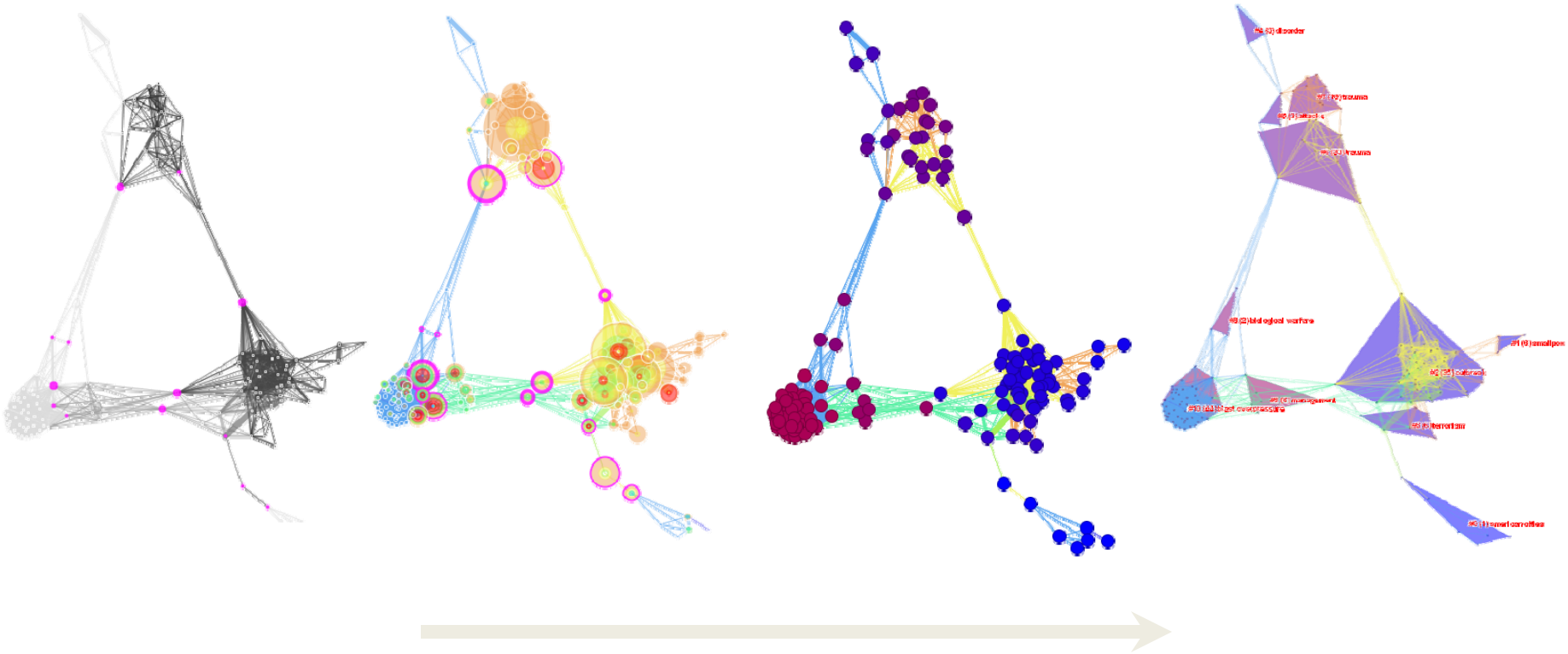
Marshall, B. J., Goodwin, C. S., Warren, J. R., Murray, R., Blincow, E. D., Blackbourn, S. J., et al. (1988). Prospective double-blind trial of duodenal ulcer relapse after eradication of Campylobacter pylori. *Lancet*, *2*(8626-8627), 1437-1442.

The 2007 Nobel Prize in Medicine to Mario R. Capecchi, Martin J. Evans and Oliver Smithies for their discoveries of "*principles for introducing specific gene modifications in mice by the use of embryonic stem cells.*"




1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003
CiteSpace, v. 2.1. Release 12
August 31, 2008 12:02:58 AM EDT
C:\#DATA\String Theory\data
Timespan: 1990-2003 (Slice Length=1)
Threshold (c, cc, ccv): 5, 5, 20; 15, 15, 20; 12, 12, 20
Network: N=554, E=5847
1
MALDACENA J, 1998, ADV THEOR MATH PHYS ...
FRIEDAN D, 1986, NUCL PHYS B ...
WITTEN E, 1991, PHYS REV D ...
POLCHINSKI J, 1995, PHYS REV LETT ...
HOROWITZ GT, 1991, NUCL PHYS B ...
WITTEN E, 1995, NUCL PHYS B ...
2
5
String Theory
sqrt
centrality
burst
Polchinski J. 1995. Dirichlet Branes and Ramond-Ramond charges. Phys. Rev. Lett. 75, 4724.

CiteSpace



Visualization Viewpoints

Editor:
Theresa-Marie Rhyne

An Information-Theoretic View of Visual Analytics

Chaomei Chen
Drexel University

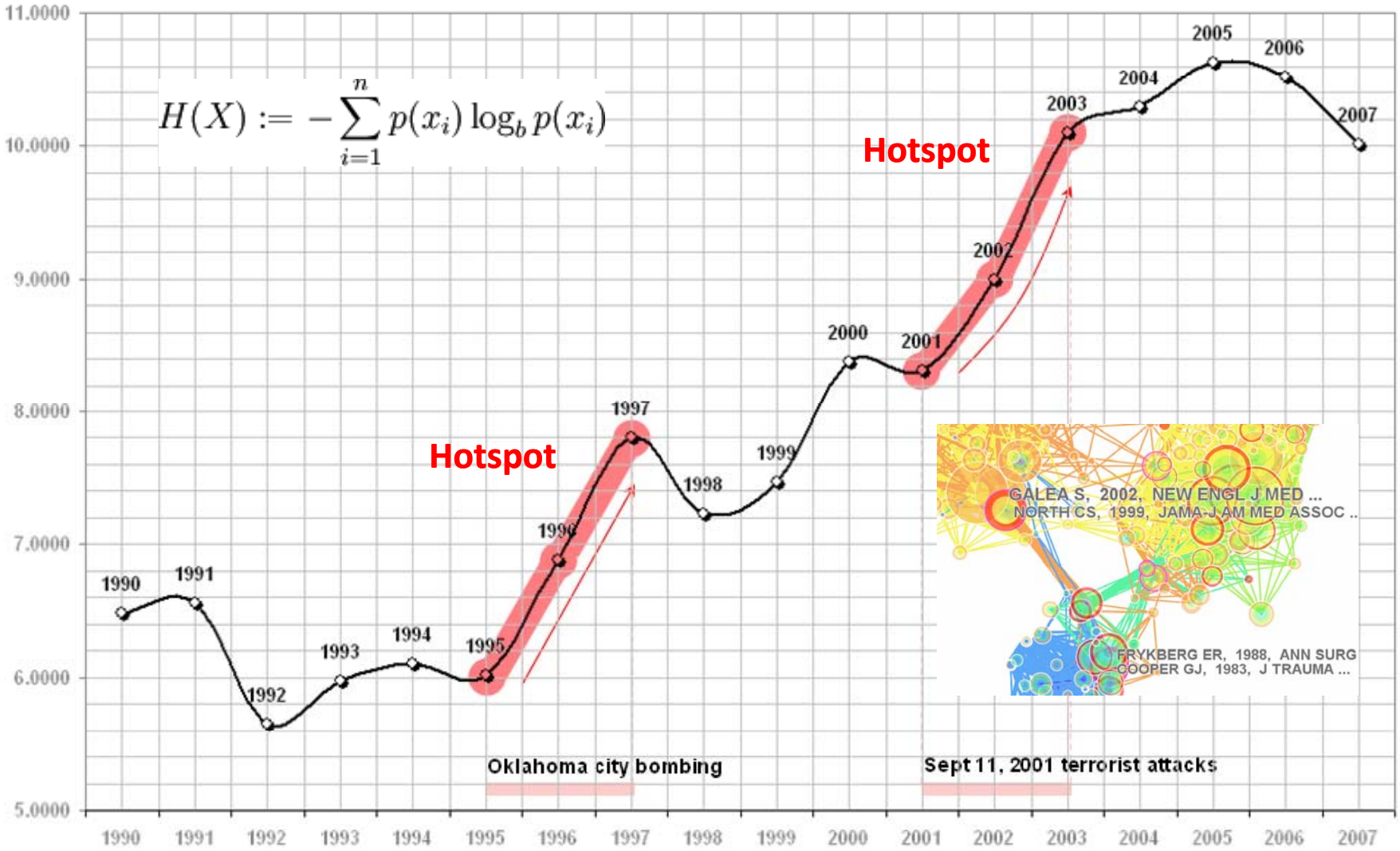
Prior to the 9/11 terrorist attacks, several foreign nationals enrolled in US civilian flying schools to learn how to fly large commercial aircraft. They wanted to learn how to navigate civilian airlines, but they were not interested in landings or takeoffs. They all paid cash for the lessons. So, the 9/11 investigations raised questions about whether intelligence agencies could have connected the dots and prevented the attacks.[1] But how do you connect these seemingly isolated dots and reveal the hidden story?

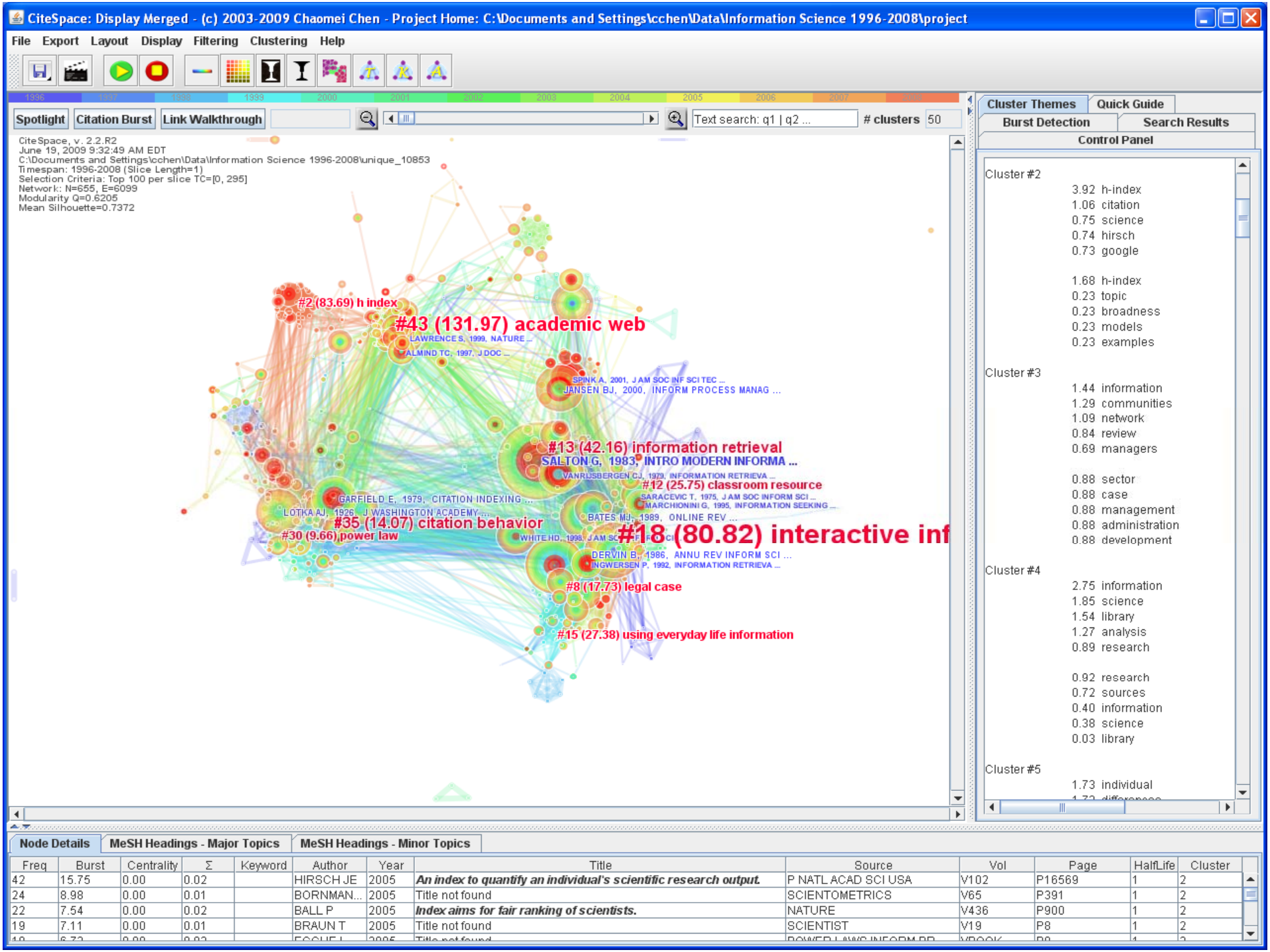
tion. We can view information entropy as the divergence from a uniform distribution. This is consistent with the common interpretation of information entropy as a measure of uncertainty, or lack of uniformity.

A useful alternative interpretation of the K-L divergence is the expected extra message length the communication entails if the message is transmitted without using a coding scheme based on the true distribution. In computer science, an object's *Kolmogorov complexity*, also known as algorithmic entropy or program-size com-

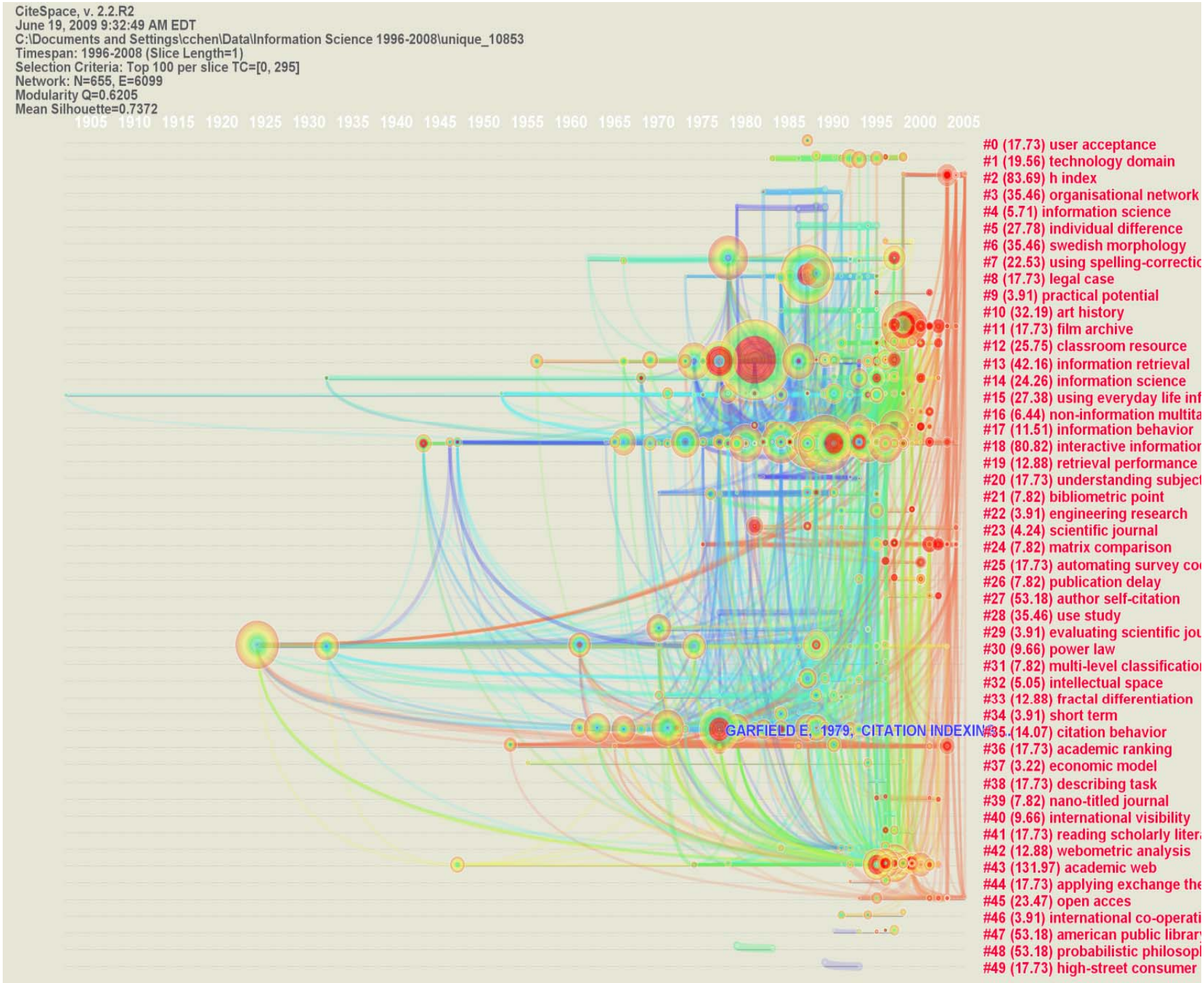
‘Profitability’ of Information Contents

Information Entropy (Vocabulary)





Freq	Burst	Centrality	Σ	Keyword	Author	Year	Title	Source	Vol	Page	HalfLife	Cluster
42	15.75	0.00	0.02		HIRSCH JE	2005	***An index to quantify an individual's scientific research output.***	P NATL ACAD SCI USA	V102	P16569	1	2
24	8.98	0.00	0.01		BORNMAN...	2005	Title not found	SCIENTOMETRICS	V65	P391	1	2
22	7.54	0.00	0.02		BALL P	2005	***Index aims for fair ranking of scientists.***	NATURE	V436	P900	1	2
19	7.11	0.00	0.01		BRAUN T	2005	Title not found	SCIENTIST	V19	P8	1	2


CiteSpace, v. 2.2.R2
June 19, 2009 9:32:49 AM EDT
C:\Documents and Settings\cchen\Data\Information Science 1996-2008\unique_10853
Timespan: 1996-2008 (Slice Length=1)
Selection Criteria: Top 100 per slice TC=[0, 295]
Network: N=655, E=6099
Modularity Q=0.6205
Mean Silhouette=0.7372
1905 1910 1915 1920 1925 1930 1935 1940 1945 1950 1955 1960 1965 1970 1975 1980 1985 1990 1995 2000 2005
#0 (17.73) user acceptance
#1 (19.56) technology domain
#2 (83.69) h index
#3 (35.46) organisational network
#4 (5.71) information science
#5 (27.78) individual difference
#6 (35.46) swedish morphology
#7 (22.53) using spelling-correctio
#8 (17.73) legal case
#9 (3.91) practical potential
#10 (32.19) art history
#11 (17.73) film archive
#12 (25.75) classroom resource
#13 (42.16) information retrieval
#14 (24.26) information science
#15 (27.38) using everyday life inf
#16 (6.44) non-information multita
#17 (11.51) information behavior
#18 (80.82) interactive information
#19 (12.88) retrieval performance
#20 (17.73) understanding subject
#21 (7.82) bibliometric point
#22 (3.91) engineering research
#23 (4.24) scientific journal
#24 (7.82) matrix comparison
#25 (17.73) automating survey co
#26 (7.82) publication delay
#27 (53.18) author self-citation
#28 (35.46) use study
#29 (3.91) evaluating scientific jou
#30 (9.66) power law
#31 (7.82) multi-level classificatio
#32 (5.05) intellectual space
#33 (12.88) fractal differentiation
#34 (3.91) short term
GARFIELD E., 1979, CITATION INDEXING
#35 (14.07) citation behavior
#36 (17.73) academic ranking
#37 (3.22) economic model
#38 (17.73) describing task
#39 (7.82) nano-titled journal
#40 (9.66) international visibility
#41 (17.73) reading scholarly litera
#42 (12.88) webometric analysis
#43 (131.97) academic web
#44 (17.73) applying exchange the
#45 (23.47) open acces
#46 (3.91) international co-operati
#47 (53.18) american public library
#48 (53.18) probabilistic philosophi
#49 (17.73) high-street consumer

QUESTIONS AND DISCUSSIONS