

Complex Networks: From the U.S. Congress to U.S. College Football

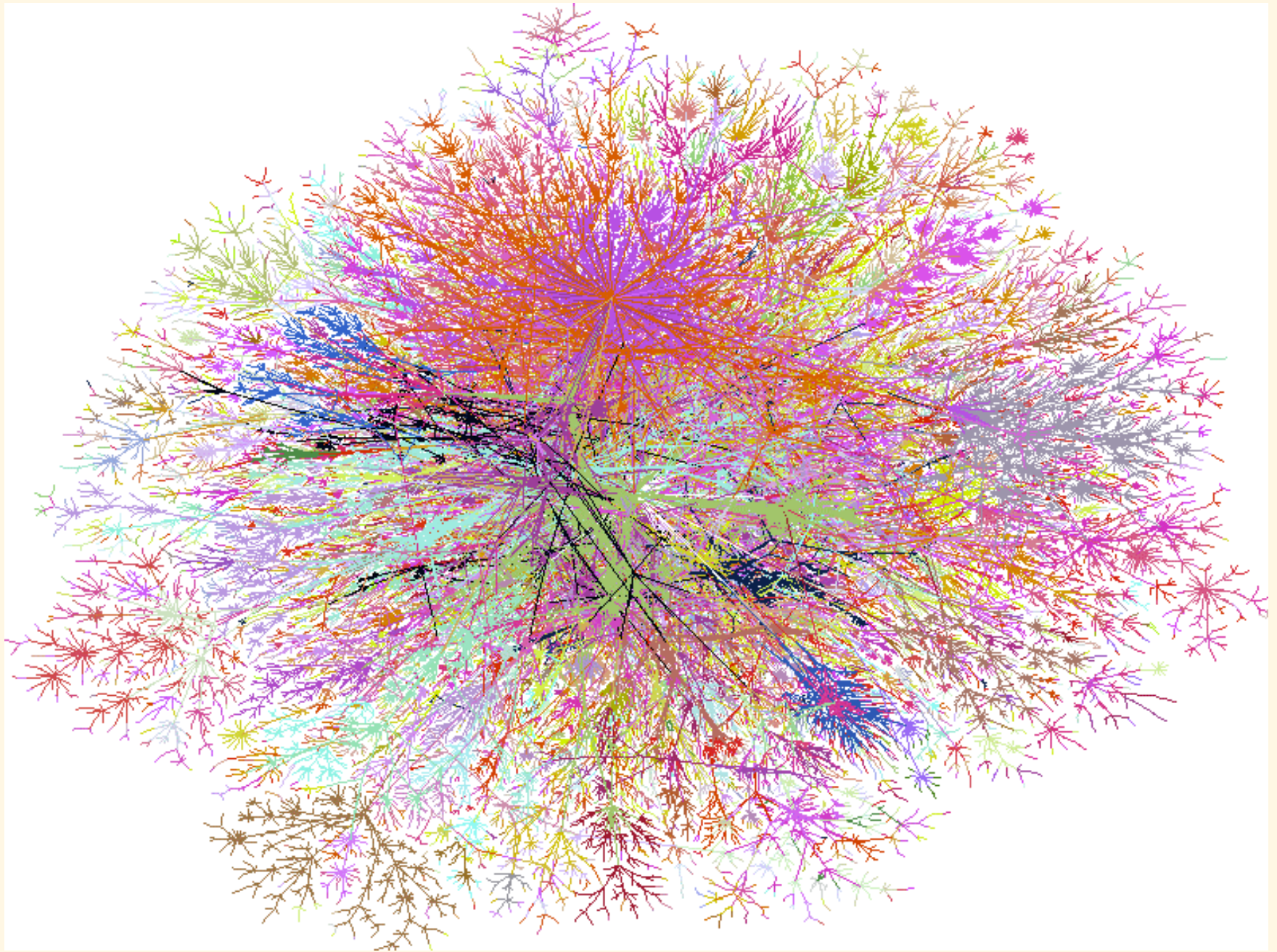
Mason A. Porter

Oxford Centre for Industrial and Applied Mathematics

Mathematical Institute

University of Oxford

COLLABORATORS: THOMAS CALLAGHAN, JAMES FOWLER, A. J. FRIEND, ERIC
KELSIC, OLGA MANDELSHTAM, PETER MUCHA, MARK NEWMAN, YE PEI, THOM



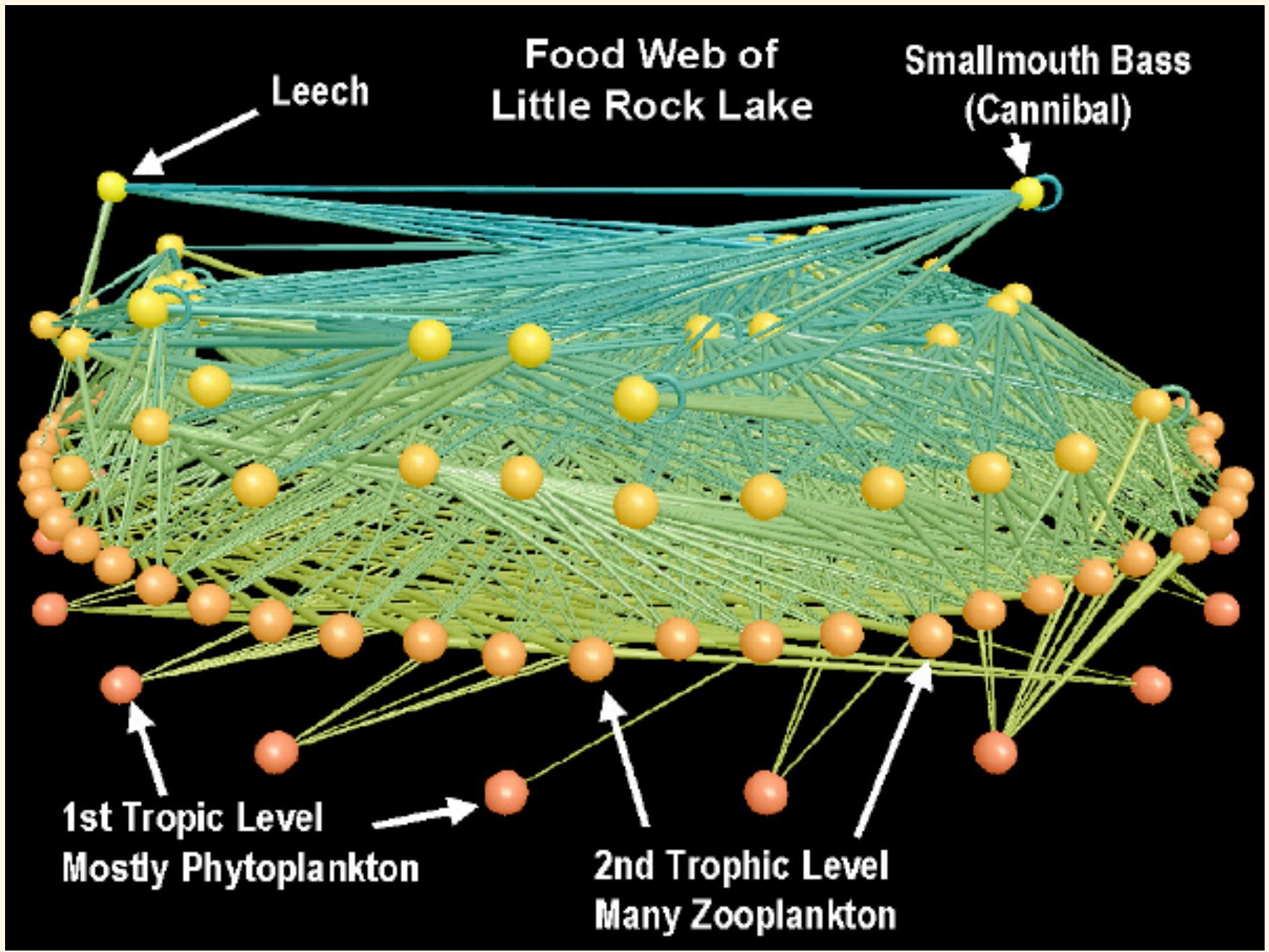
10/30/07. Oxford

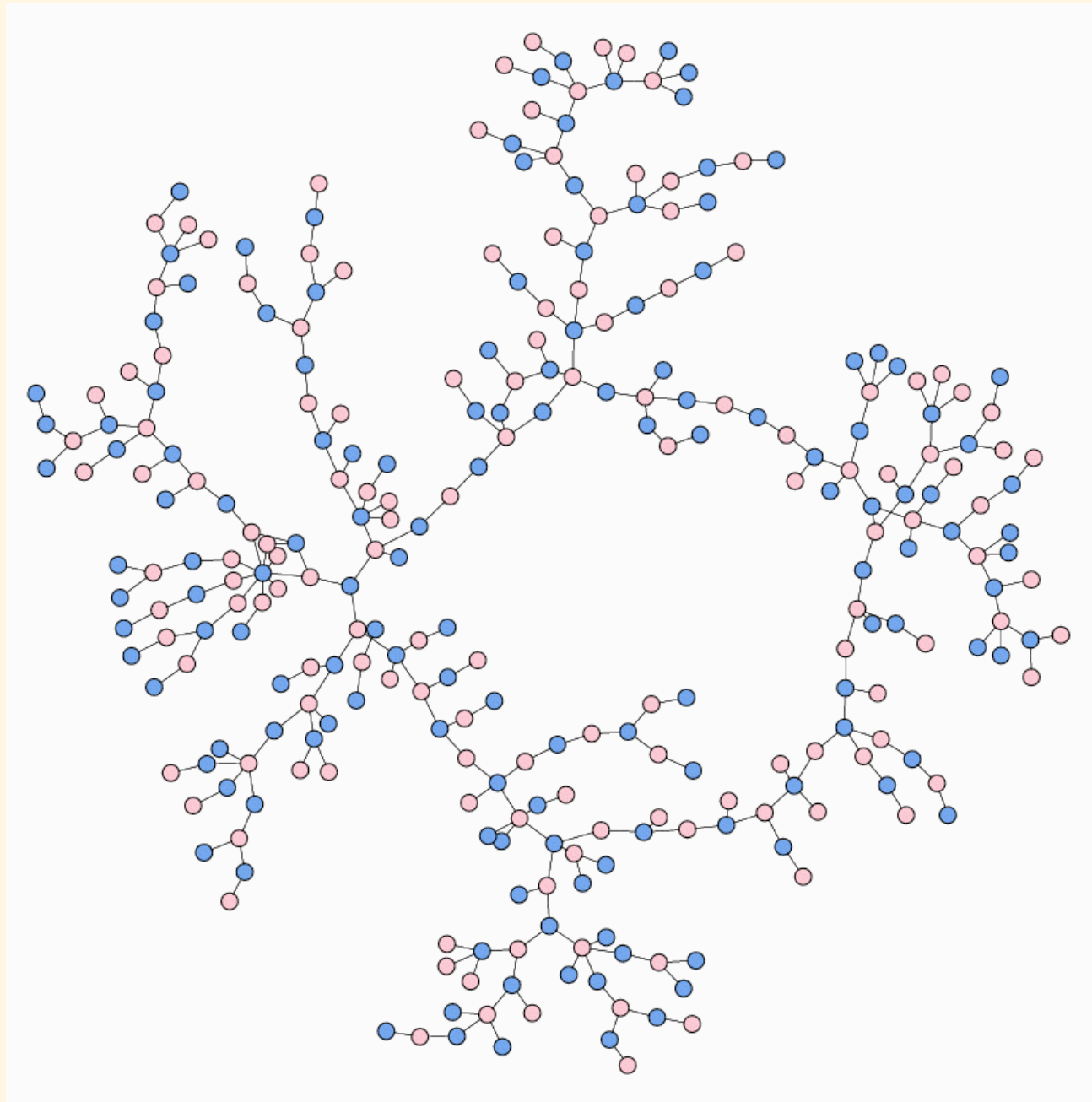
Outline

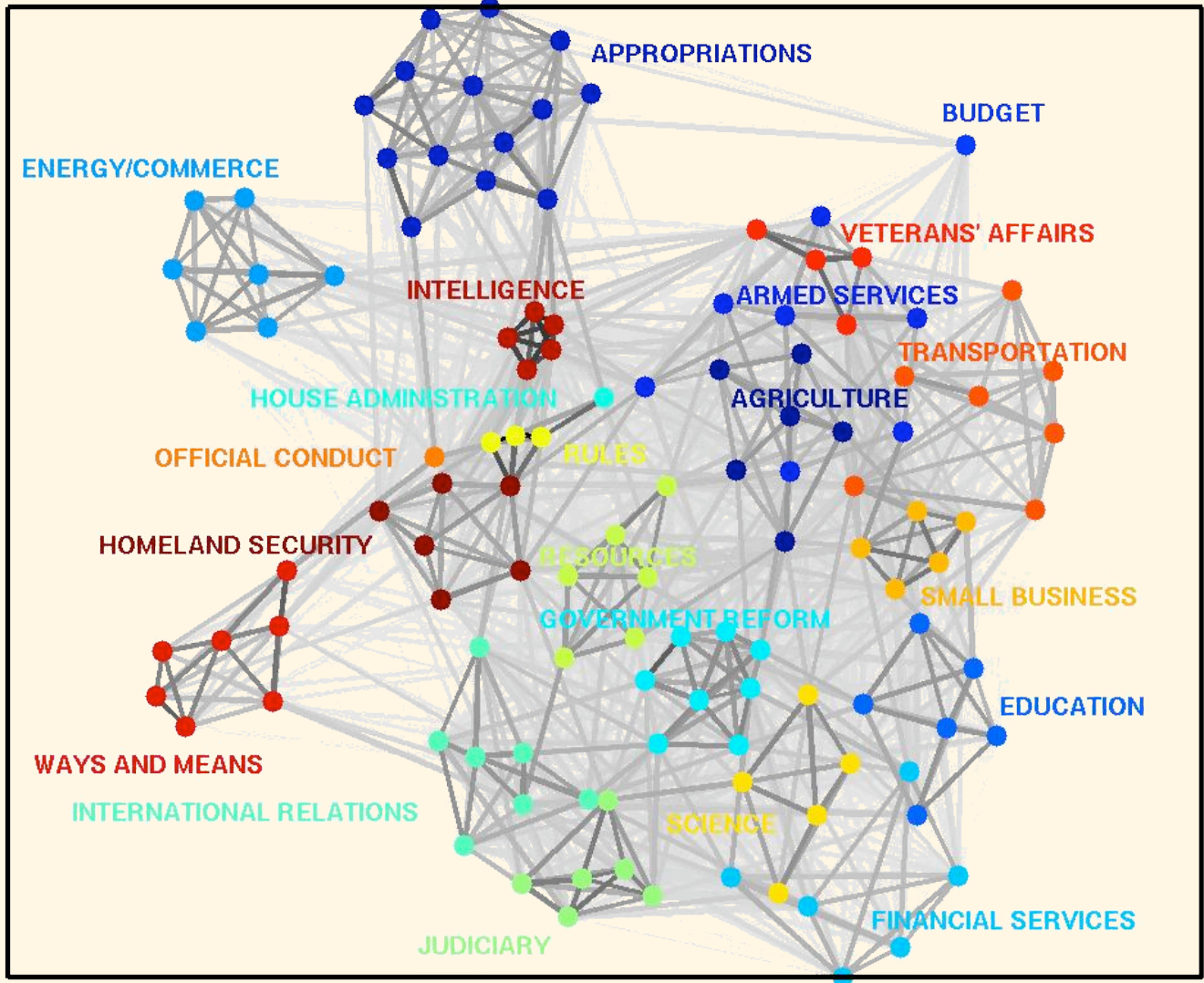
- "Complex networks"
- Communities in networks
- NCAA Division-IA Football
 - Rankings from biased random walks
- United States Congress
 - Committee assignment network
 - Quantifying the politics of Representatives and committees
 - Legislation cosponsorship and roll call voting networks
- Facebook networks and other current projects
- Summary

General References

- Survey/review articles
 - S. H. Strogatz [2001], "Exploring Complex Networks," *Nature* **410**, 268-276.
 - M. E. J. Newman [2003], "The Structure and Function of Complex Networks," *SIAM Review* **45**(2), 167-256.
- Netwiki: <http://netwiki.amath.unc.edu/>







From leaves to root...

- 1) Start without connections
- 2) Identify connection with strongest weight
- 3) Connect
- 4) Check to see if any components merged
- 5) Return to Step 2

Fewer options for unweighted networks, as it is unclear how to start this process...

From root to leaves...

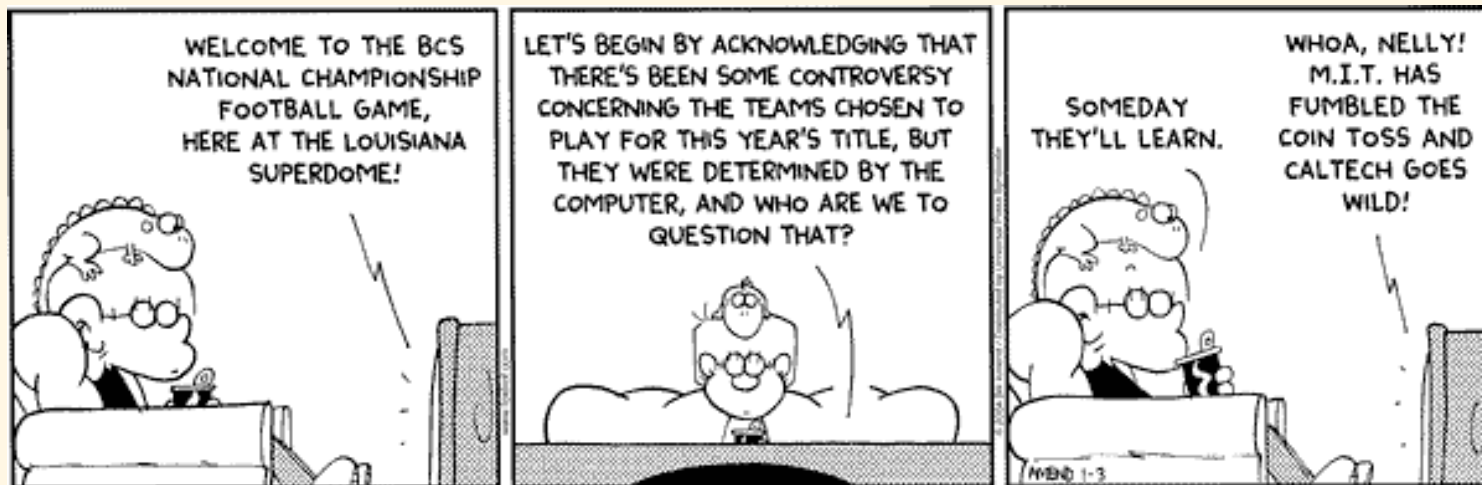
- 1) Identify weakest connection/edge (e.g., by weight or betweenness)
- 2) Remove
- 3) Check to see if component breaks
- 4) Return to Step 1

Different ways to identify “strength,” depending on size of network and whether it is weighted or unweighted

Recent theory: Eigenvector-based modularity maximization of M. E. J. Newman, *PNAS/PRE* 2006.

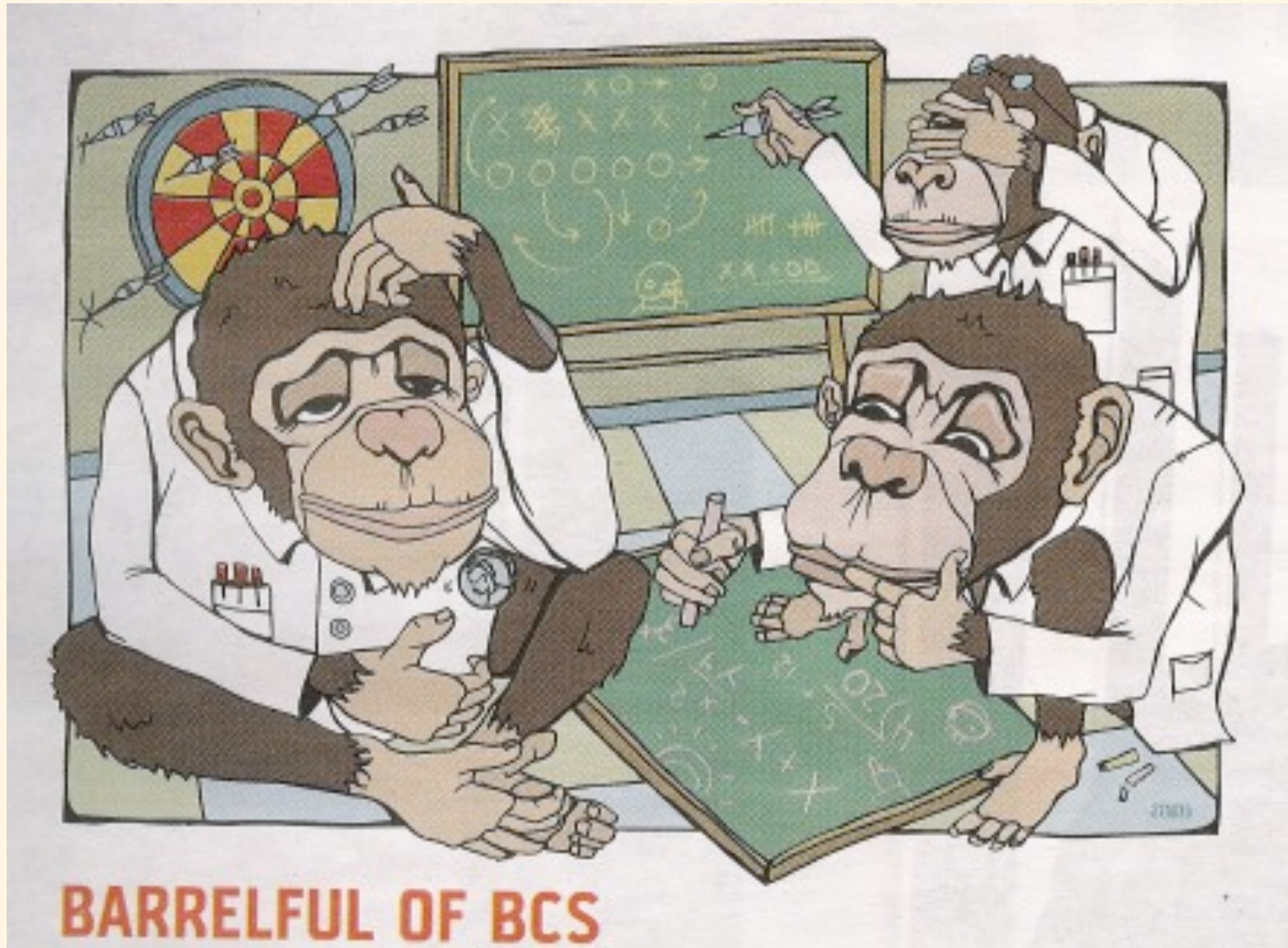
College Football

- T. Callaghan, P. J. Mucha, & MAP [2004], "The Bowl Championship Series: A mathematical review," *Notices of the AMS* **51**, 887-893.
- TC, PJM, & MAP [2007], "Random walker ranking for NCAA Division I-A football," *American Mathematical Monthly* **114**(9), 761-777.



<http://rankings.amath.unc.edu/>

Disclaimer



ESPN The Magazine

10/30/07. Oxford

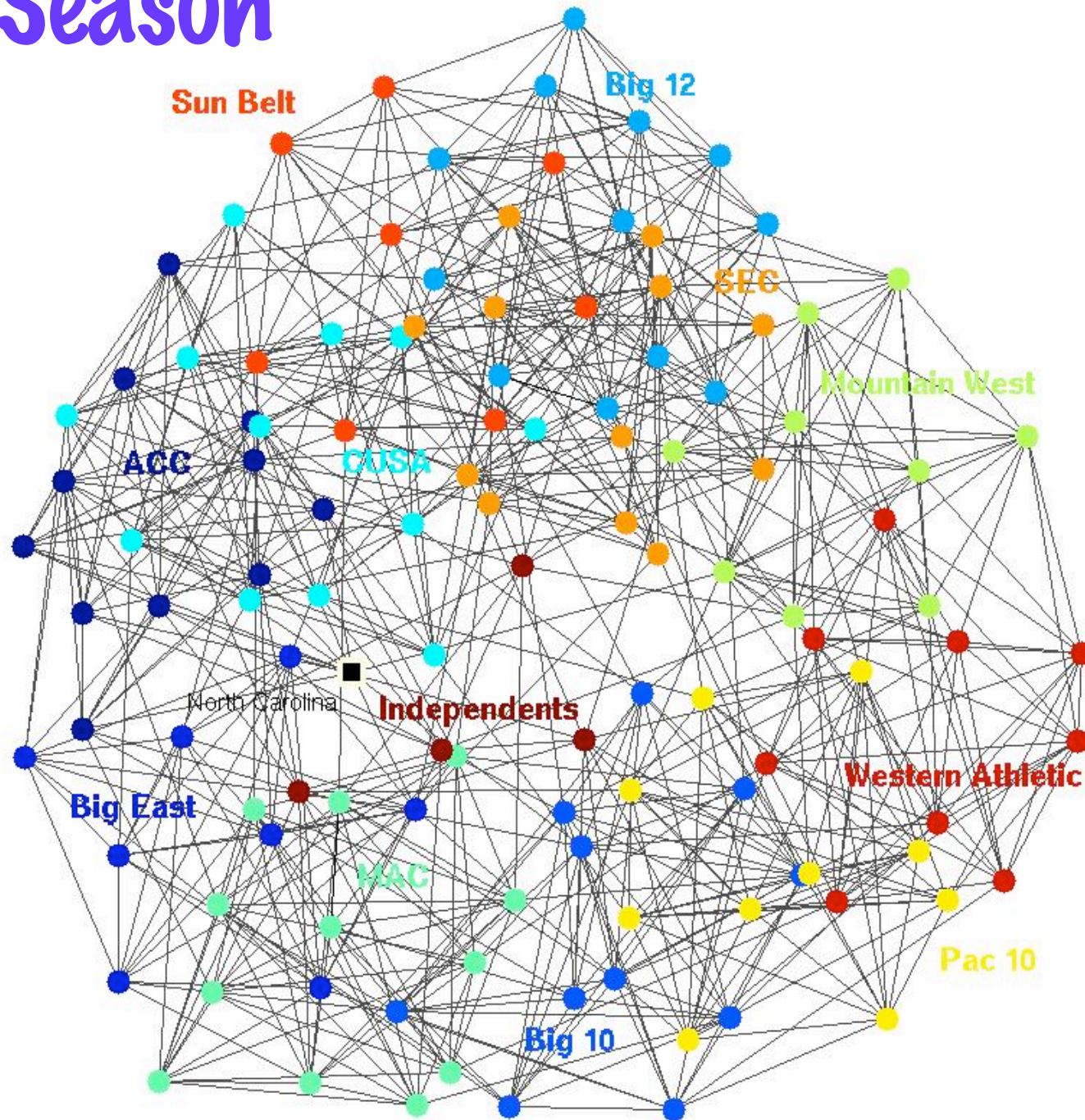


10/30/07. Oxford

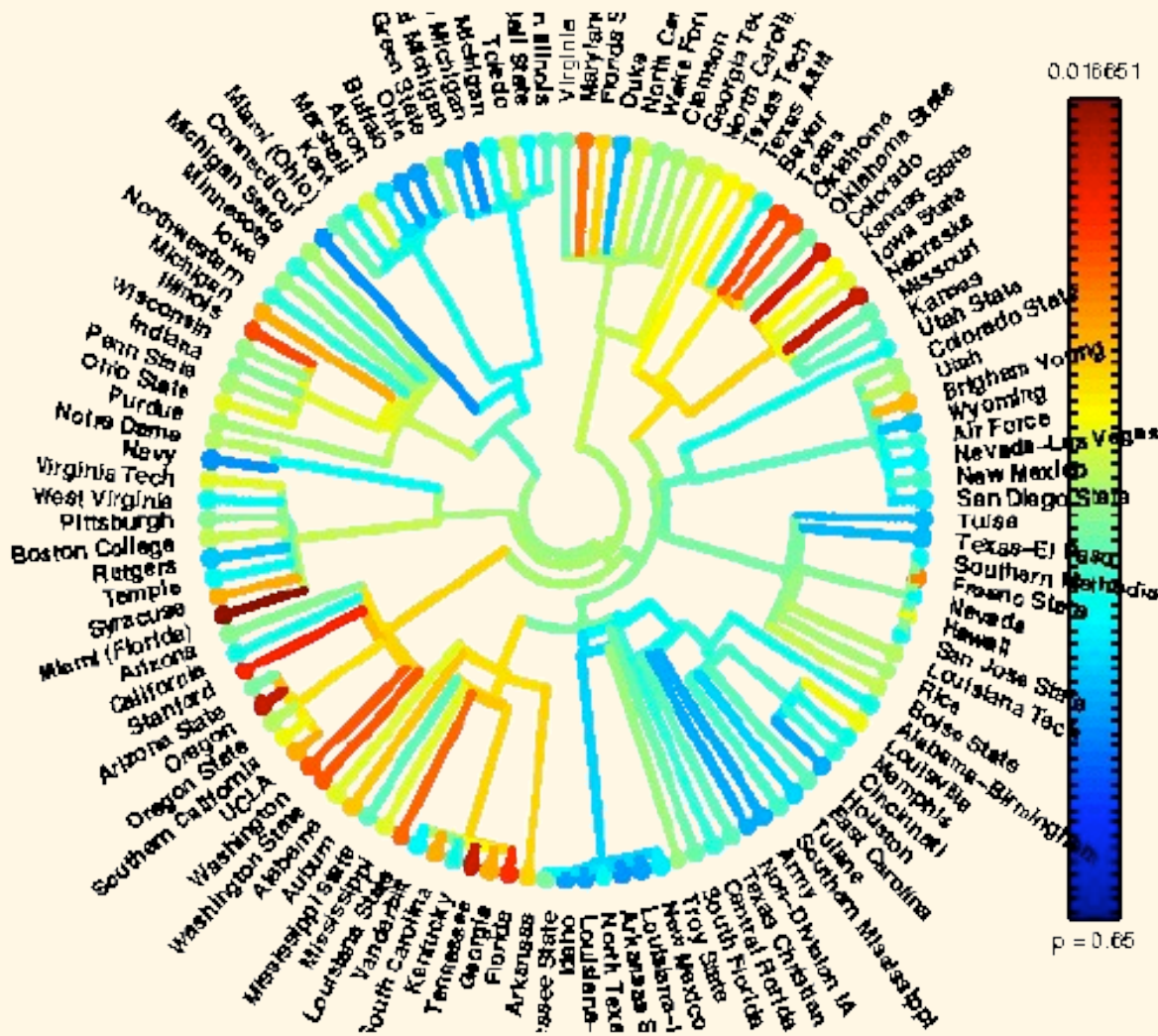
NCAA Division-IA Football

- Teams (nodes) connected to each other by games played (edges)
- In 2005, the 119 Division I-A teams played a total of 690 games prior to end-of-season bowl games
- Diameter = 4
- Single connected component in 3-4 weeks
- Most teams play majority of games inside their own conferences (ACC, SEC, etc.)
- One of the only sports at *any* level that doesn't determine champions in a playoff

2005 Season

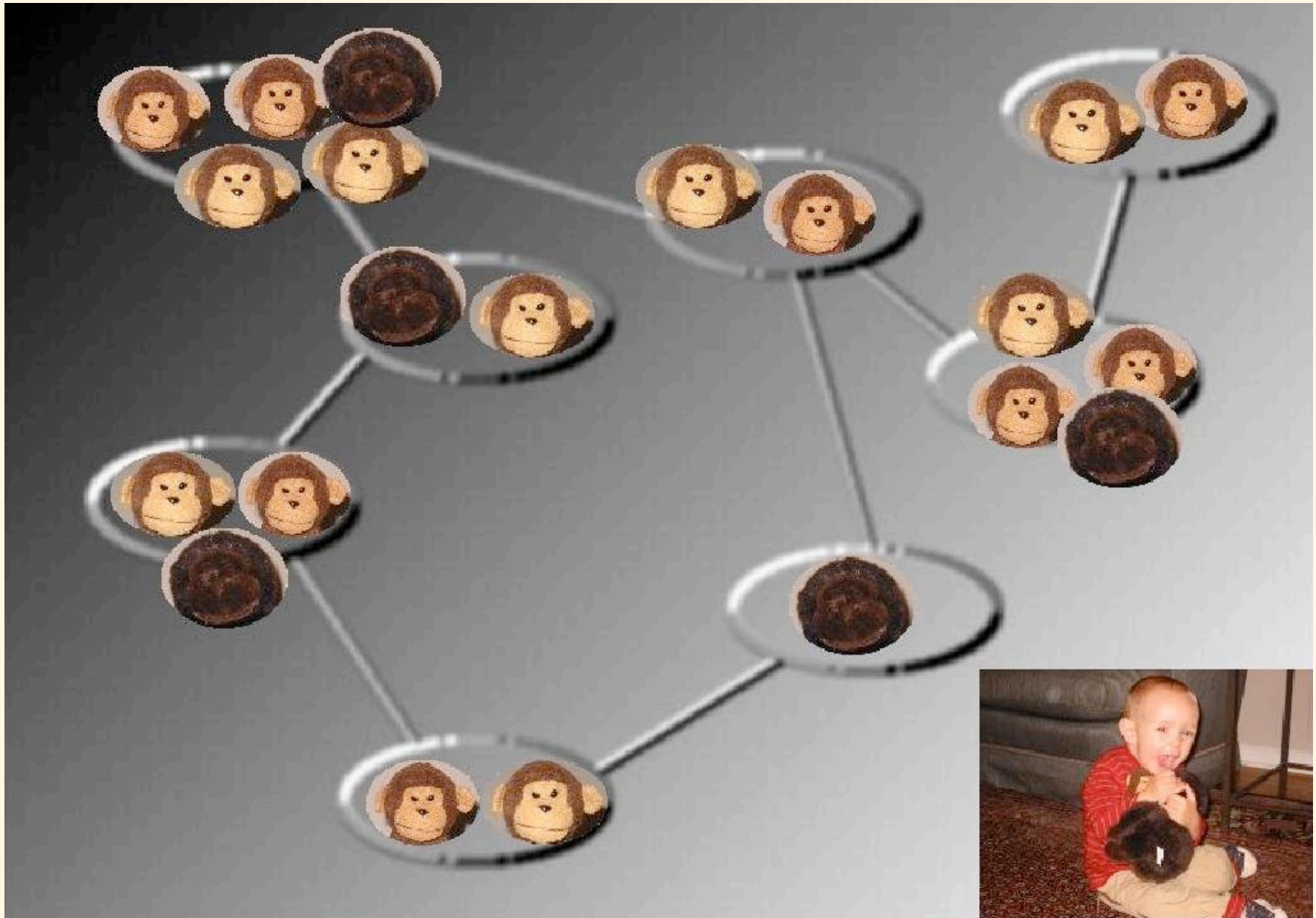


Community Structure



- Strong conference structure in Div-IA
- Girvan-Newman betweenness-based algorithm (PNAS, 2002), counting geodesics through each edge, clearly identifies different conferences

Biased Random Walk on Graph



Random-Walker Rankings

- 1) Randomly select a single game played by your "favorite" team
- 2) Flip weighted coin (heads with prob. p)
- 3) Heads: go with winner; tails: go with loser
- 4) Return to Step 1

An individual random walker will never settle down, but an ensemble has well-defined steady-state statistics
Interesting mathematics in the asymptotics for different values of p and in round-robin tournaments.

2007 Rankings (10/27/07)

Random walkers ($p = 0.75$)

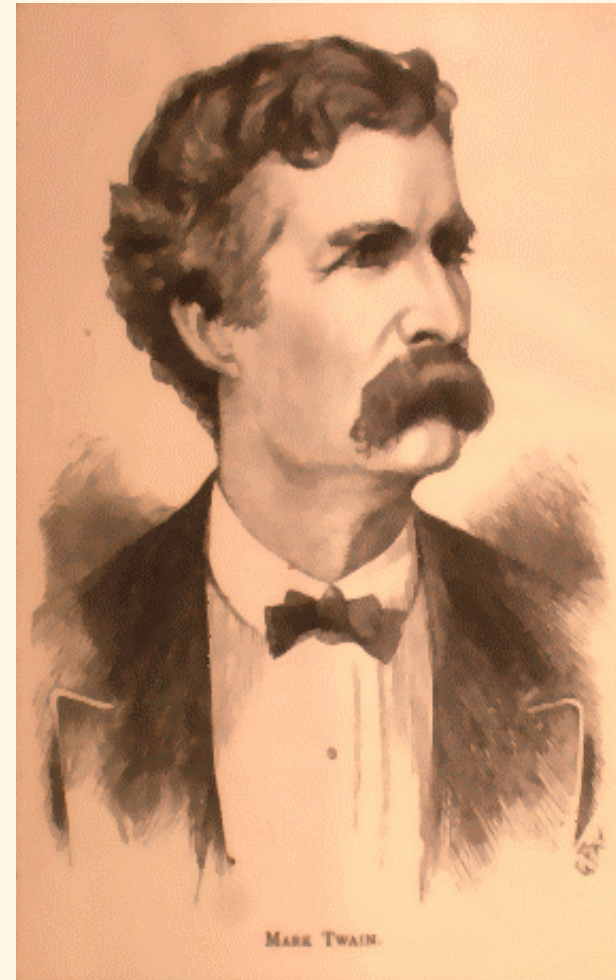
1. Arizona State
2. Boston College
3. LSU
4. Oregon
5. Kansas
6. Ohio State
7. Georgia (13th for BCS)
8. West Virginia
9. Oklahoma
10. Connecticut (15th for BCS)

BCS (now called FBS)

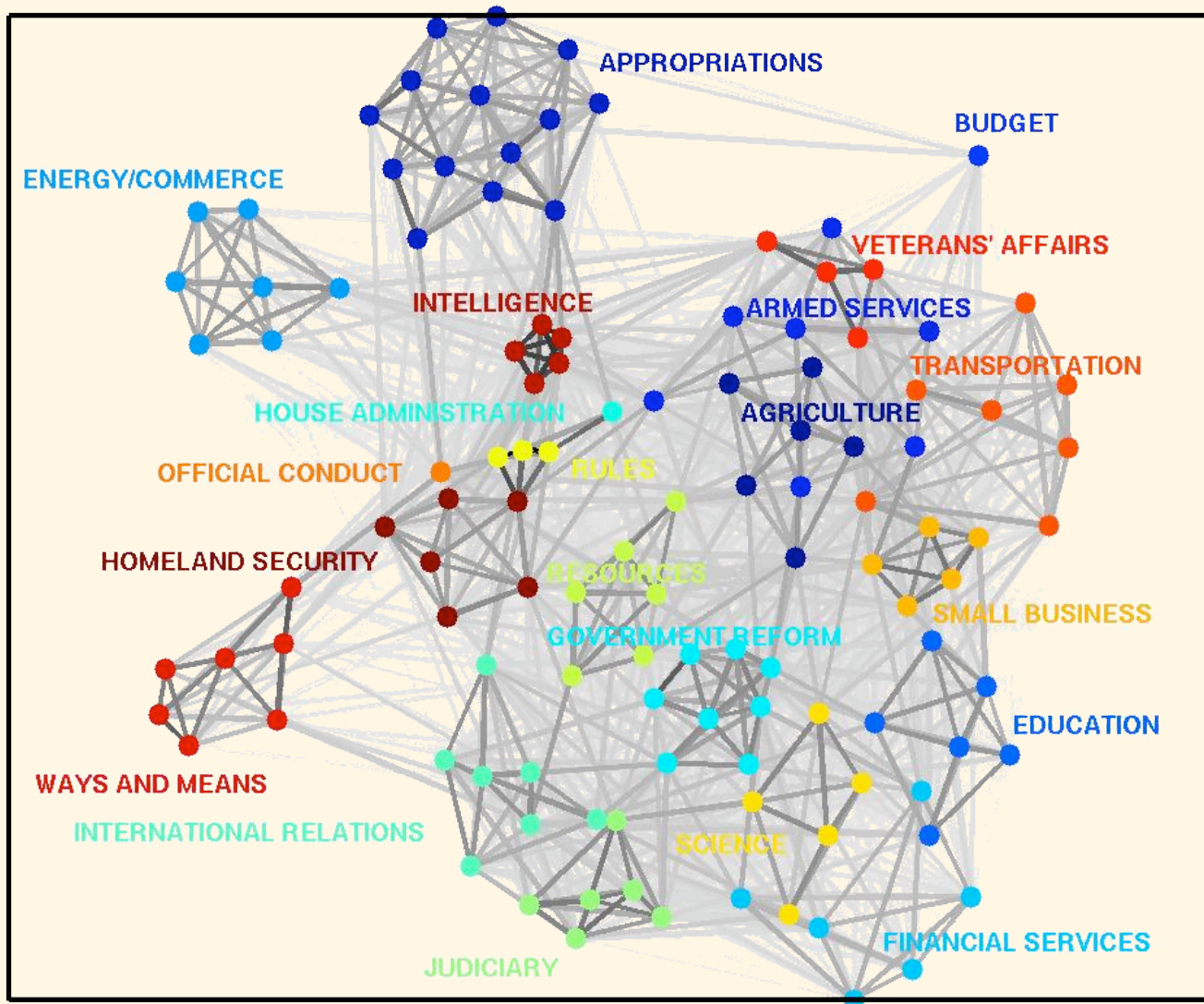
1. Ohio State
2. LSU
3. Arizona State
4. Oregon
5. Boston College
6. Kansas
7. West Virginia
8. Oklahoma
9. South Florida (11th for us)
10. Missouri (14th for us)

Congress: A Popular American Villain

- "It could be probably be shown by facts and figures that there is no distinctly American criminal class except Congress."
-- Mark Twain
- "Suppose you were an idiot and suppose you were a member of Congress. But I repeat myself."
-- Mark Twain



Congressional Committee Assignment Networks



Committees and the subcommittees connected by the Representatives through committee assignments.

Weights assigned to either (a) raw interlock of common members or (b) normalized interlock in terms of expected overlap.

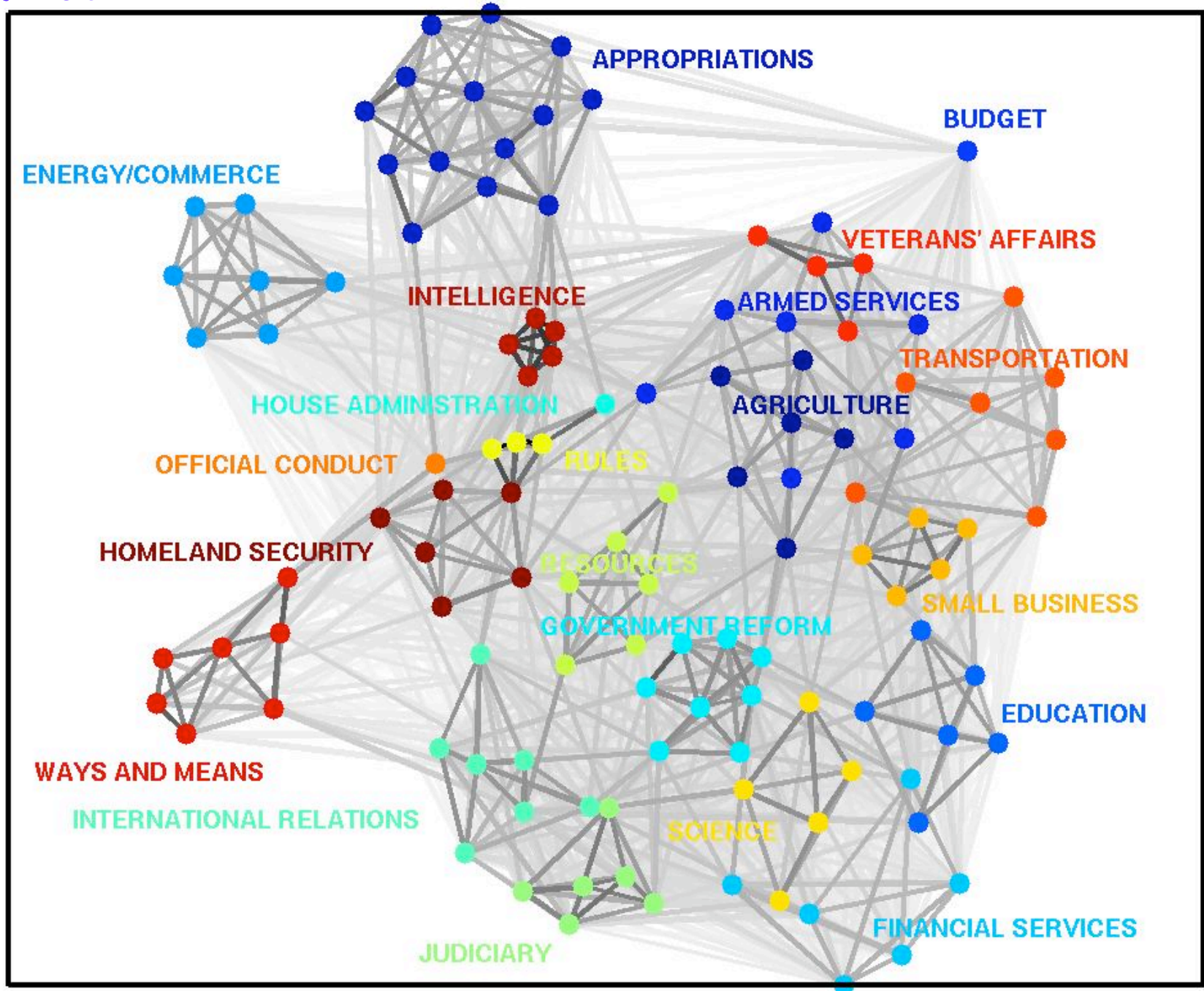
Congressional Committees Assignments

- AMS Mathematical Moment: "Unearthing Power Lines"
- MAP, P. J. Mucha, M. E. J. Newman, & C. M. Warmbrand [2005] "A network analysis of committees in the U.S. House of Representatives," *Proc. Nat. Acad. Sci.* **102**, 7057-62.
- MAP, A. J. Friend, PJM, & MEJN [2006], "Community structure in the U.S. House of Representatives," *Chaos*, **16**(4), 041106.
- MAP, PJM, MEJN, & AJF [2007], "Community structure in the United States House of Representatives," *Physica A* **386**(1), 414-438 .

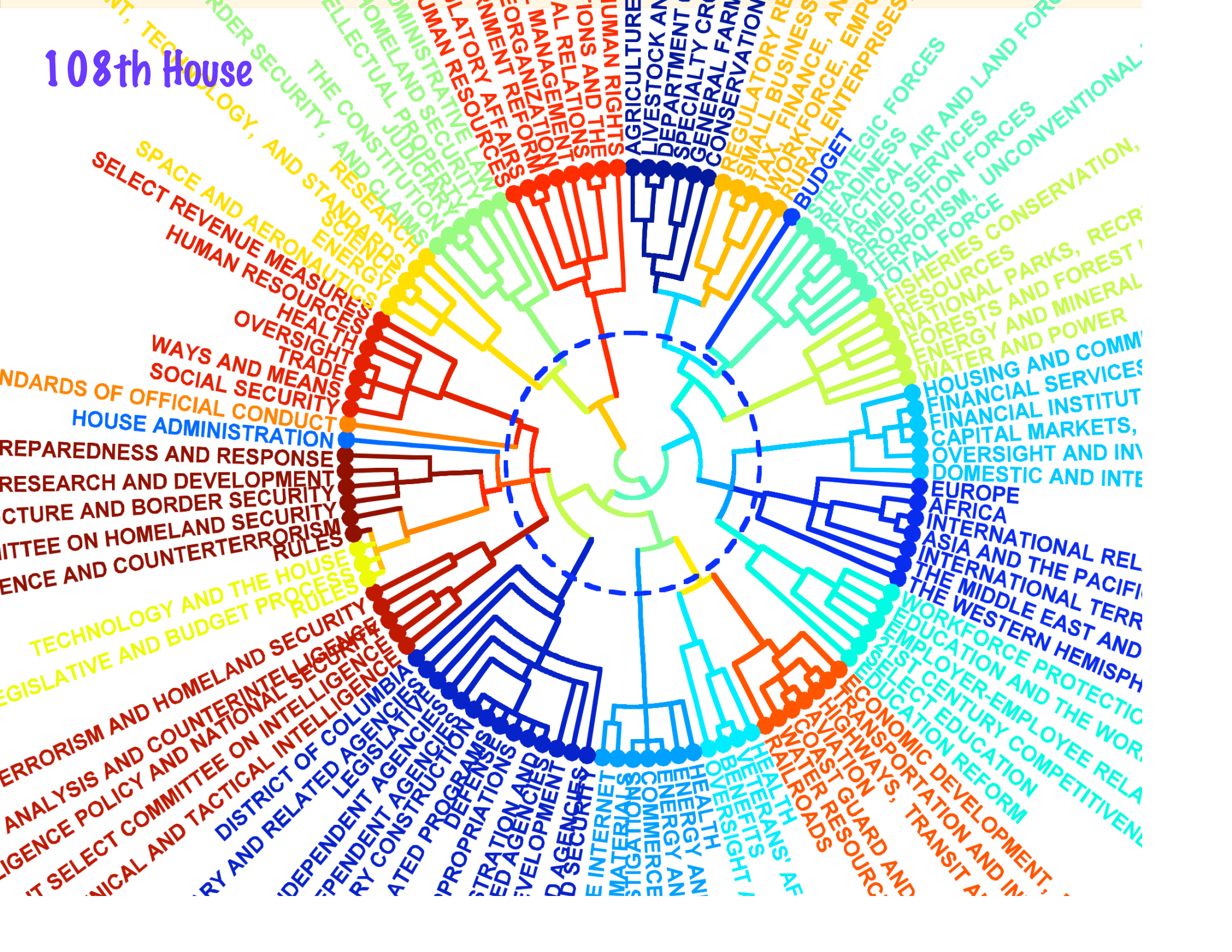
Committee Assignment Network

- Bipartite graph of 115-165 committees and about 440 Representatives and Delegates assigned to committees.
- Typical Representative sits on 2 Standing or Select committees, and about 2 subcommittees of each.
- Much of detailed work in making U.S. law occurs in committees and subcommittees.
- *Network is dense relative to many popular examples (good warmup for phylogenetics).*
- Major recent changes:
 - 1994 elections ("Republican Revolution")
 - 9/11 and Homeland Security

108th House

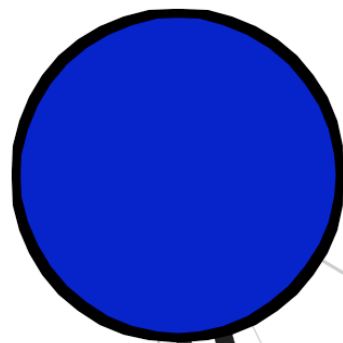


108th House

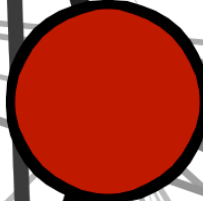
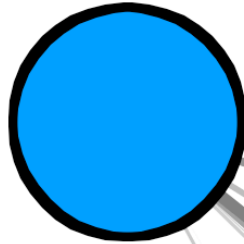


108th House

ENERGY/COMMERCE



APPROPRIATIONS



INTELLIGENCE

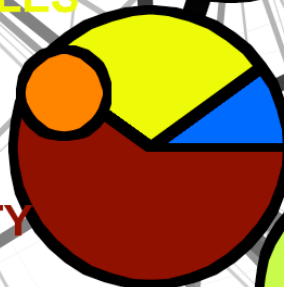
ARMED SERVICES



BUDGET

VETERANS' AFI

RULES



HOUSE ADMINISTRATION



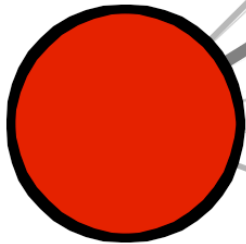
TRANSPORTA

OFFICIAL CONDUCT

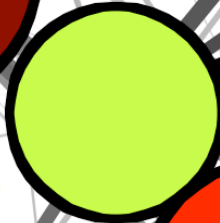
AGRICULTURE

HOMELAND SECURITY

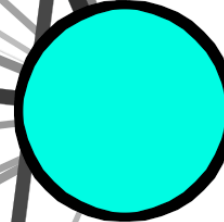
SMALL BUSINESS



RESOURCES

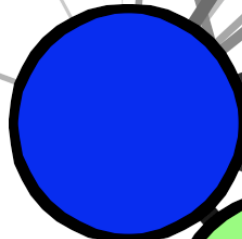


GOVERNMENT REFORM



EDUCATION

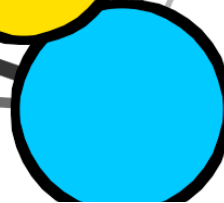
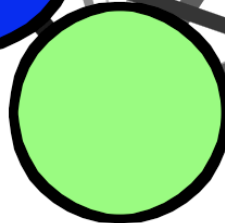
WAYS AND MEANS



INTERNATIONAL RELATIONS



SCIENCE



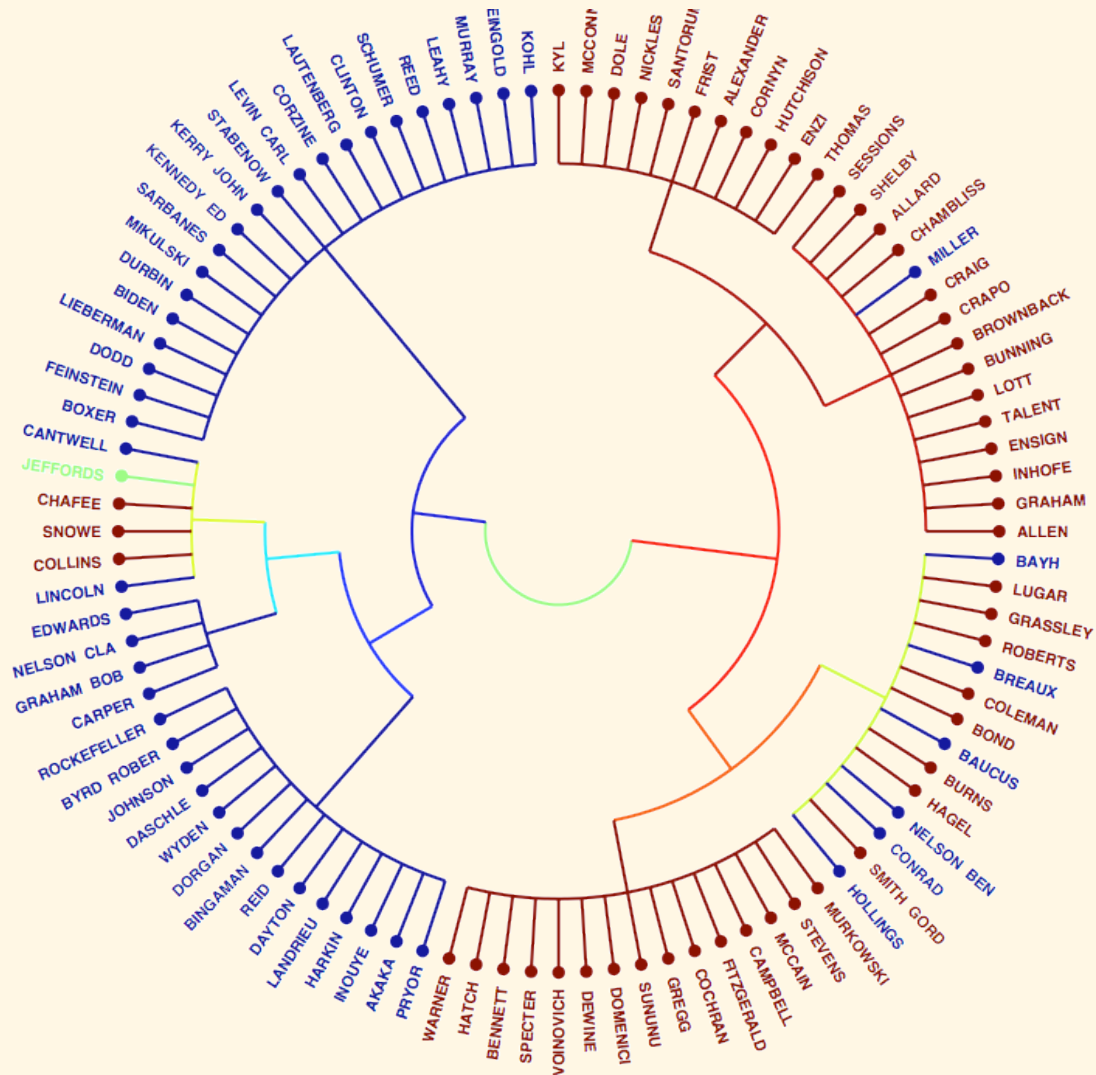
Quantifying Politics

- Voting matrix of roll call, +1/-1
(Representatives vs. measures)
- Singular value decomposition (SVD)
identifies that most of the variance of the
votes is in first two modes (eigenvectors)
[see Poole & Rosenthal, Sirovich]
- First mode ~ "Partisanship"
- Second mode ~ "Bipartisanship"

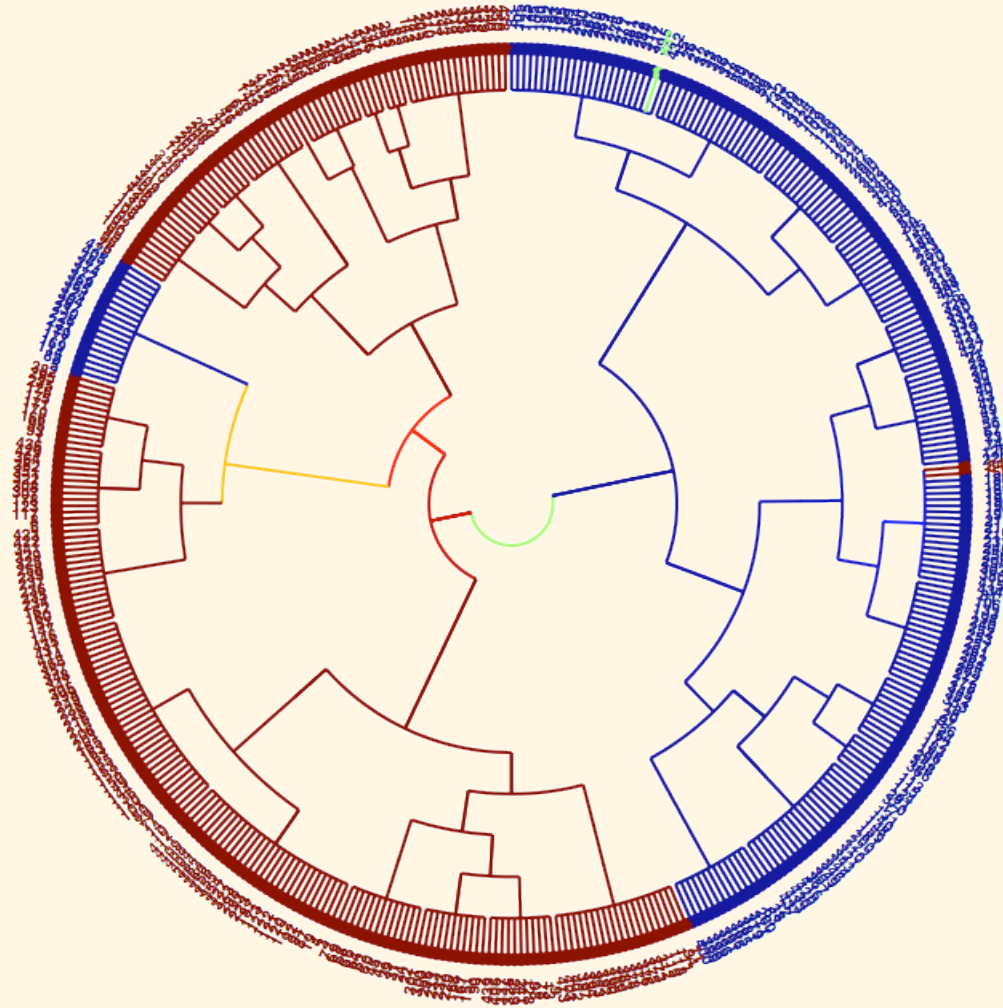
Legislation Cosponsorship Network

- Two Congressmen are connected if they sponsor/cosponsor legislation
- “Higher dimensional” data than committee assignments
 - Can be seen using modularity maximization
- Shows that polarization in Congress was gradual rather than abrupt
 - Can be quantified using modularity
- Y. Zhang, AJF, A. L. Traud, MAP, J. H. Fowler, PJM, submitted to *Physica A* (arXiv: 0708.1191)

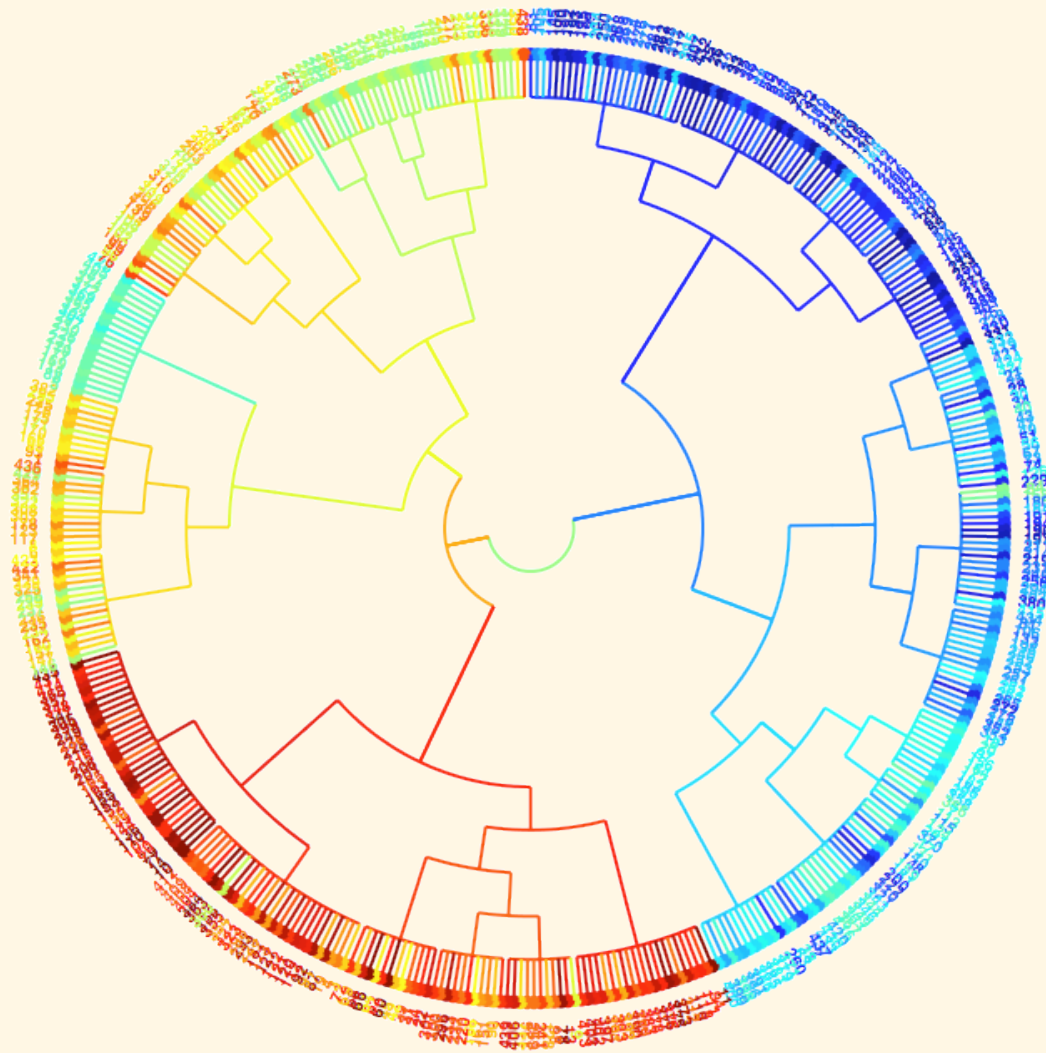
108th Senate (colored by party)



108th House (colored by party)

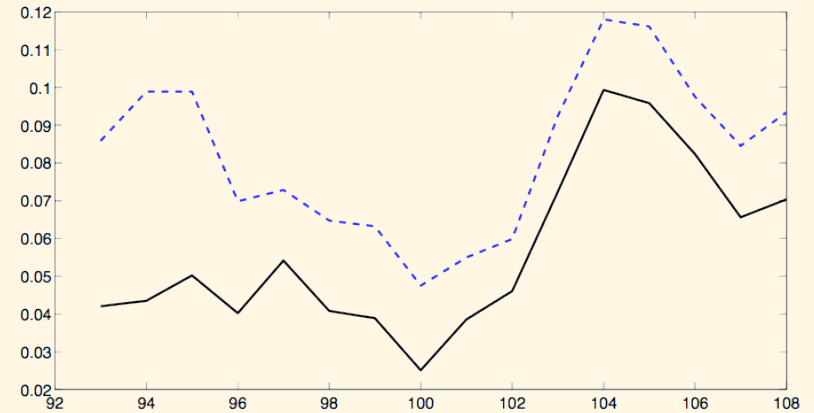
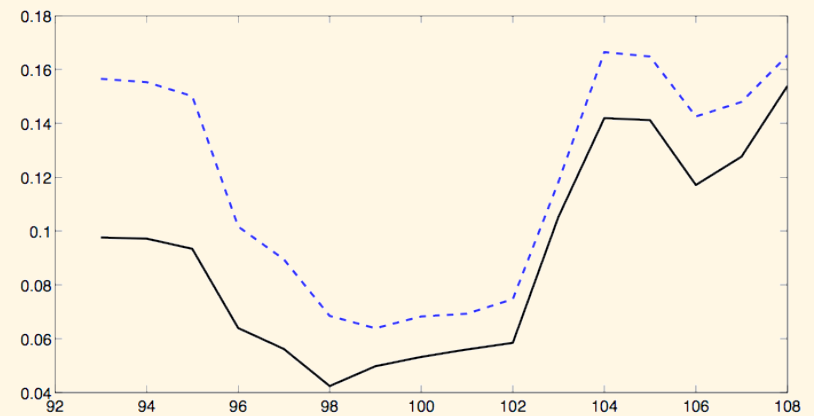
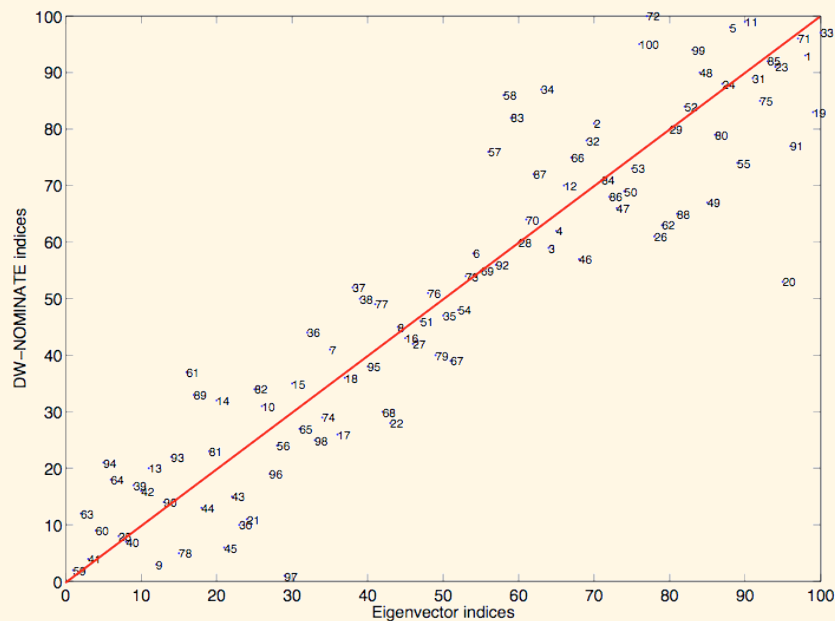


108th House (colored by DW-Nominate)



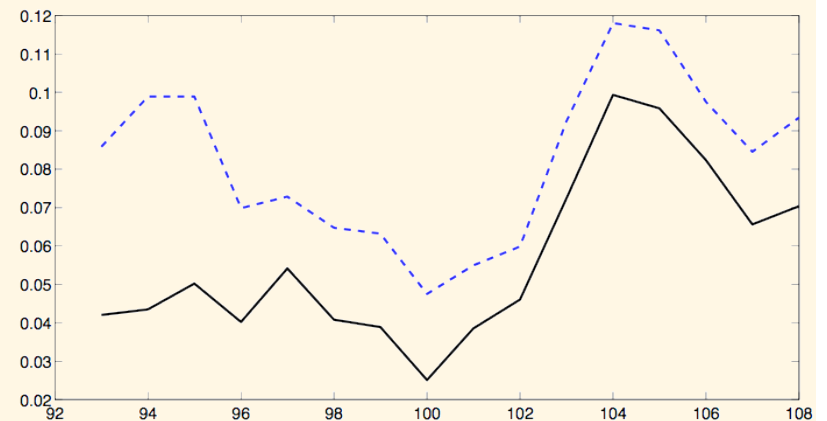
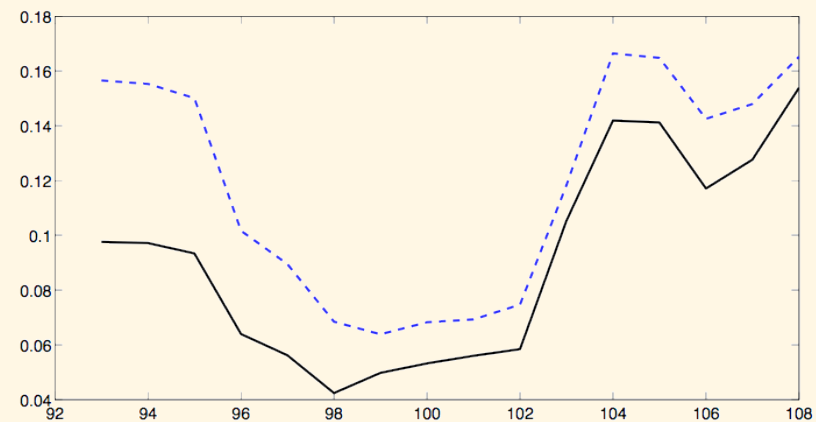
Partisanship via modularity

- Strong rank correlation: DW-Nominate versus components of leading modularity eigenvector

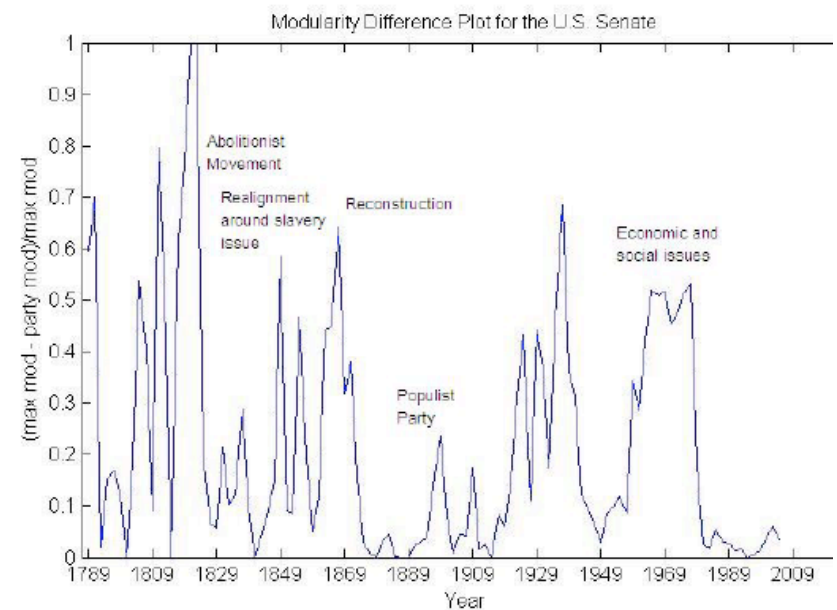
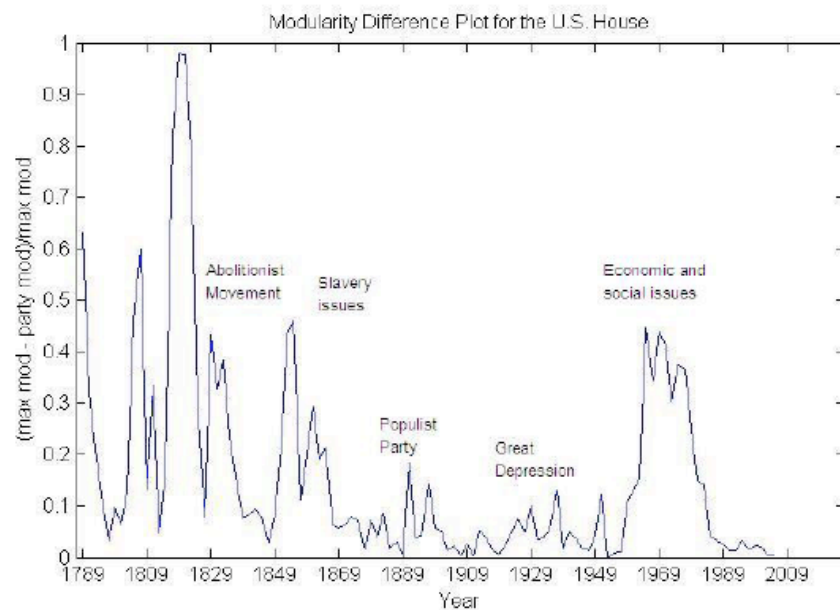


Partisanship via modularity

- Modularity at first leading-eigenvector split (good approximation of maximum) up sharply in early 1990s in both houses of Congress
- Modularity obtained when partitioning by party lines also up sharply and becomes closer to that given by eigenvector
- *Increased polarization in Congress appears in bill cosponsorship (and roll call)*



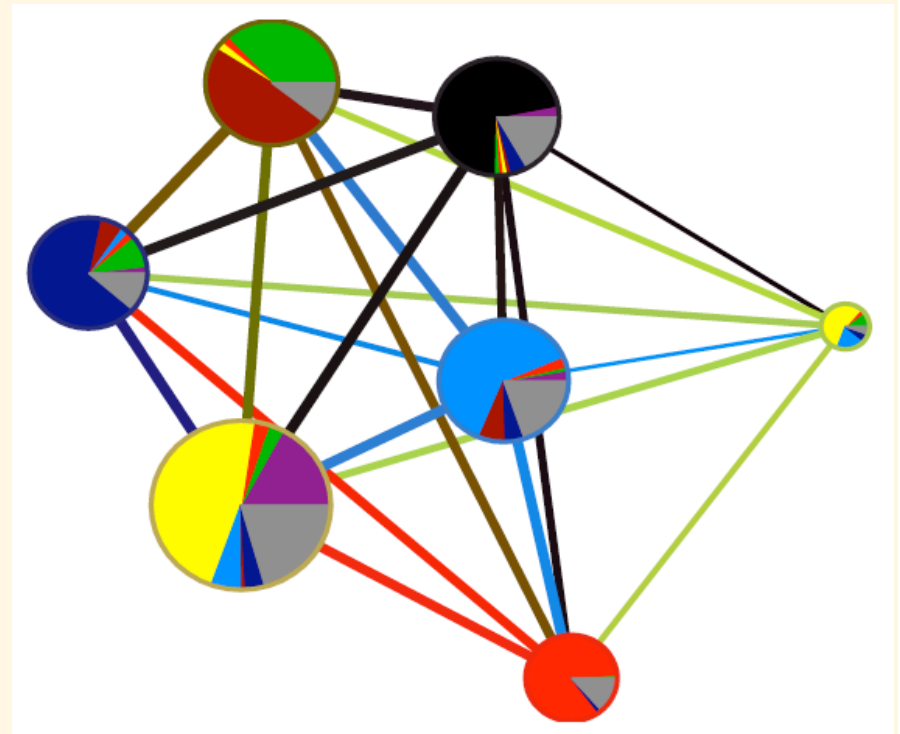
Political realignments via modularity



- A. Waugh, L. Pei, ALT, MAP, JFH, & PJM, in preparation.
 - Note: being sent to a political science journal...
- Uses roll call voting data
- Future work: voting in UK parliament (need students/postdoc)

Facebook

- Some community detection results (a tutorial with Facebook as working example)
 - ALT, E. Kelsic, PJM, & MAP, in preparation
- Friendship network among college students
- Data for 100 schools
- Different structures from different network growth mechanisms?
 - Olga Mandelshtam, Summer 2007
 - Need students/postdocs!



Caltech network

Current and Future Work

- Comparison of different Congressional networks
 - Committee/subcommittee assignments, legislation cosponsorship, roll call votes
 - Note: committee data available on request
- Some generalizations on eigenvector community detection for three-way splittings (UNC students)
- U.S. Supreme Court precedent network (anyone?)
- Baseball Hall of Fame rankings (anyone?)
- Baseball pitcher rankings (anyone?)
- Network growth mechanisms with Facebook and Supreme Court networks (anyone?)
- UK voting networks (anyone?)
- Always trying to acquire other interesting data...
- I'm actively trying to recruit students and postdocs...

Summary

- Investigate processes on networks (football team rankings, Congressional collaborations, collegiate social life, etc.) by studying hierarchical structure
- Development and focus: novel data
- Undergraduate students leading or involved
- Reprints and preprints:
<http://www.maths.ox.ac.uk/~porterm> or
porterm@maths.ox.ac.uk
- Netwiki: <http://www.netwiki.amath.unc.edu>
 - Wiki for network science