# A network analysis of committees in the U.S. House of Representatives

**Mason A. Porter*†‡, Peter J. Mucha*, M. E. J. Newman§, and Casey M. Warmbrand¶**

*School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160; †Center for Nonlinear Science, School of Physics, Georgia Institute of Technology, Atlanta, GA 30332-0430; §Department of Physics and Center for the Study of Complex Systems, University of Michigan, Ann Arbor, MI 48109-1040; and ¶Department of Mathematics, University of Arizona, Tucson, AZ 85721-0089

Edited by Simon A. Levin, Princeton University, Princeton, NJ, and approved April 5, 2005 (received for review January 9, 2005)

**Network theory provides a powerful tool for the representation and analysis of complex systems of interacting agents. Here, we investigate the U.S. House of Representatives network of committees and subcommittees, with committees connected according to "interlocks," or common membership. Analysis of this network reveals clearly the strong links between different committees, as well as the intrinsic hierarchical structure within the House as a whole. We show that network theory, combined with the analysis of roll-call votes using singular value decomposition, successfully uncovers political and organizational correlations between committees in the House without the need to incorporate other political information.**

Congress | politics | hierarchical clustering | singular value decomposition

Much of the detailed work in making U.S. law is performed by congressional committees and subcommittees, which is in contrast to parliamentary democracies such as those of Great Britain and Canada in which a larger part of the legislative process is directly in the hands of political parties or is conducted in sessions of the entire parliament. Although the legislation drafted by committees in the U.S. Congress is subject ultimately to roll-call votes by the full House of Representatives and Senate, the important role played by committees and subcommittees makes the study of their formation and composition vital to understanding the work of the American legislature.

Several contrasting theories of committee-assignment strategies have been developed in the political science literature (mostly through qualitative studies, although there have been some quantitative ones as well) (1–6), but there is no consensus explanation of how committee assignments are initially determined or how they are modified from one session of Congress to the next. A question of particular interest is whether political parties assign committee memberships essentially at random or if important congressional committees can be seen by using objective analysis to be "stacked" with partisan party members.

The work presented here approaches these issues by using a different set of analytical tools from those used previously. We use the tools of network theory, which have been applied successfully in recent years to characterize a wide variety of complex systems (7, 8). As we show, network theory is particularly effective at uncovering structure among committee and subcommittee assignments without the need to incorporate any specific knowledge about committee members or their political positions.

Although there has been only limited previous work on networks of congressional committees, there is a considerable body of literature on other collaboration networks such as the boards of directors of corporations (9–13), which occupy a position in the business world somewhat analogous to that occupied by committees in Congress. It has been shown that board memberships and the networks that they create play a major role in the spread of attitudes, ideas, and practices through the corporate world, affecting political donations (10), investment strategies (14), and even the stock market on which a company is listed (15). Studies of the structure of corporate networks have shed considerable light on the mechanisms and pathways of information diffusion (16–18), and it seems plausible that the structure of congressional committees will be similarly revealing.

## Networks of Committees

We study the U.S. House of Representatives and construct bipartite, or "two-mode," networks based on assignments of Representatives to committees and subcommittees (called just "committees" for simplicity hereafter) in the 101st–108th Houses (1989–2004). (Table 1 lists the House leadership during this period.) These networks have two types of nodes, Representatives and committees, with edges connecting each Representative to the committees on which they sit.

We project these two-mode committee-assignment networks onto one-mode networks with nodes that represent the committees and edges that represent common membership, or "interlocks," between committees. Fig. 1 shows a visualization of the network of committees for the 107th House (2001–2002), an example that we analyze in some depth.

The more common members that two committees have, the stronger their connection is in the network. We quantify the strength of connection by the "normalized interlock," defined as the number of common members divided by the expected number of such common members if committees of the same size were randomly and independently assigned from the members of the House. Committees with as many common members as would be expected by chance have a normalized interlock value of 1, those with twice as many have an interlock value of 2, those with none have an interlock value of 0, and so forth.

Some of the connections depicted in Fig. 1 are expected and unsurprising. For example, one finds that sets of subcommittees of the same larger committee share many of the same members, thereby forming a group or clique in the network. For example, the four subcommittees of the 107th Permanent Select Committee on Intelligence each include at least half of the full 20-member committee and at least one third of each of the other subcommittees. These tight connections result in normalized interlocks with values in the range of 14.4–26.6, which are substantially higher than average and cause these five nodes to be drawn close together in the graph visualization, forming the small pentagon in the middle right of Fig. 1.

We also find more surprising connections between committees. For example, the nine-member Select Committee on Homeland Security, formed in June 2002 during the 107th Congress in the aftermath of the terrorist attacks of September 11, 2001, is observed to have a strong connection to the

This paper was submitted directly (Track II) to the PNAS office.

Abbreviation: SVD, singular value decomposition.

‡To whom correspondence should be addressed. E-mail: mason@math.gatech.edu.

**Table 1. U.S. House of Representatives leadership for the 101st–108th Congresses**

| Congress | Speaker | Majority Leader | Minority Leader | Majority Whip | Minority Whip |
|---|---|---|---|---|---|
| 101st (1989–1990) | T. S. Foley | R. A. Gephardt | R. H. Michel | T. Coelho, W. H. Gray, III | D. Cheney, N. L. Gingrich |
| 102nd (1991–1992) | T. S. Foley | R. A. Gephardt | R. H. Michel | W. H. Gray III, D. E. Bonior | N. L. Gingrich |
| 103rd (1993–1994) | T. S. Foley | R. A. Gephardt | R. H. Michel | D. E. Bonior | N. L. Gingrich |
| 104th (1995–1996) | N. L. Gingrich | R. K. Armey | R. A. Gephardt | T. D. DeLay | D. E. Bonior |
| 105th (1997–1998) | N. L. Gingrich | R. K. Armey | R. A. Gephardt | T. D. DeLay | D. E. Bonior |
| 106th (1999–2000) | J. D. Hastert | R. K. Armey | R. A. Gephardt | T. D. DeLay | D. E. Bonior |
| 107th (2001–2002) | J. D. Hastert | R. K. Armey | R. A. Gephardt | T. D. DeLay | N. Pelosi |
| 108th (2003–2004) | J. D. Hastert | T. D. DeLay | N. Pelosi | R. Blunt | S. Hoyer |

The Democrats held the House majority in the 101st–103rd Congresses (1989–1994), and the Republicans held it in the 104th–108th Congresses (1995–2004).

13-member Rules Committee (with a normalized interlock value of 7.4 from two common members), which is the committee charged with deciding the rules and order of business under which legislation will be considered by other committees and the full House (see http://thomas.loc.gov). The Homeland Security Committee is also connected to the seven-member Legislative and Budget Process Subcommittee of Rules by the same two common members (with a normalized interlock value of 13.7). In the 108th Congress (data not shown), the Homeland Security Committee swelled to 50 members but maintained a close association with the Rules Committee (with a normalized interlock value of 4.1 from six common members).

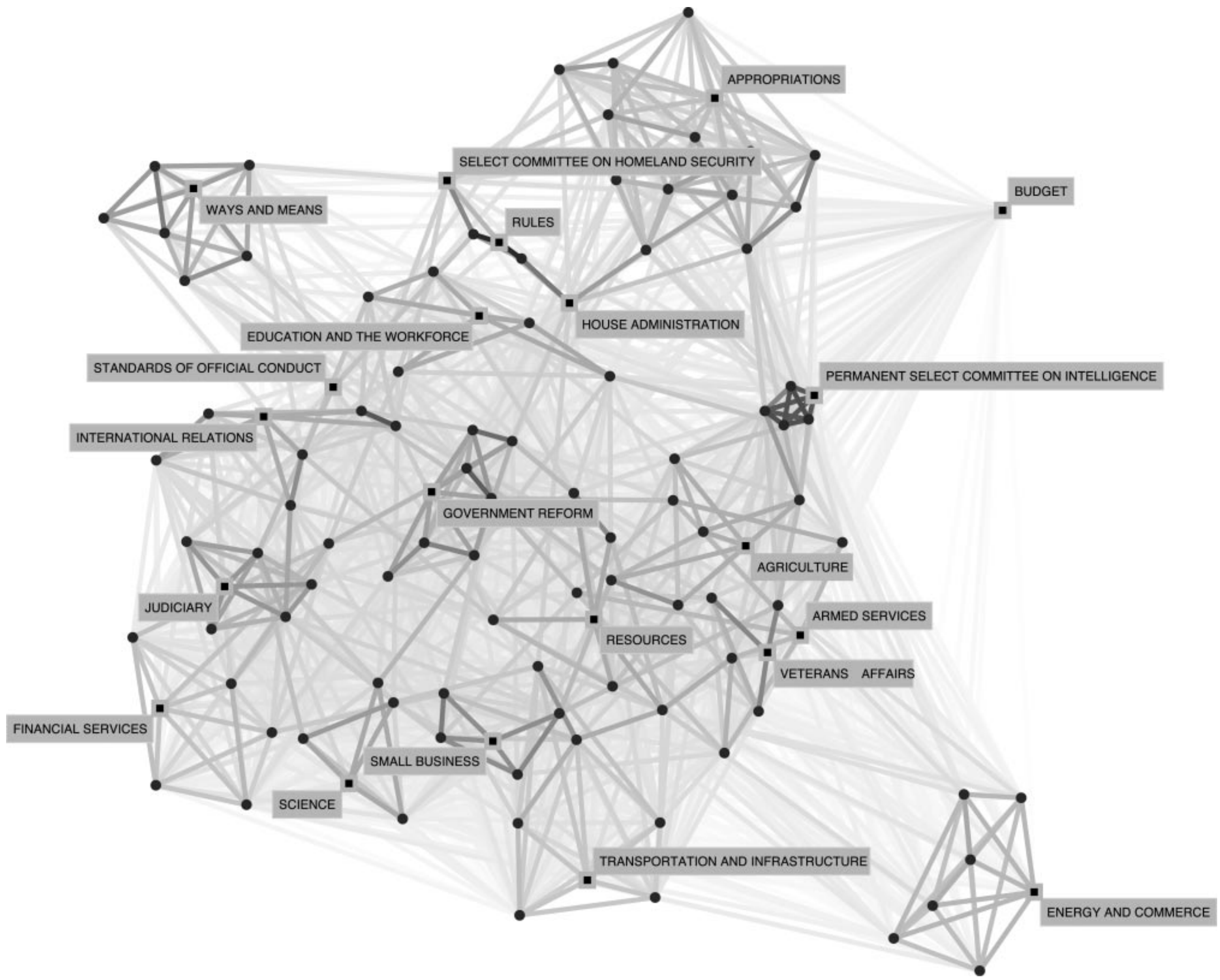


**Fig. 1.** Network of committees (■) and subcommittees (●) in the 107th U.S. House of Representatives, with standing and select committees labeled. (Subcommittees tend to be closely tied to their main committee and, therefore, are unlabeled.) Each link between two (sub)committees is assigned a strength that is equal to the normalized interlock. Thus, lines between pairs of circles or pairs of squares represent a normalized degree of joint membership between (sub)committees (lines between squares are typically very light because of this normalization), and lines between squares and circles represent the fraction of standing committee members on subcommittees. This figure was drawn by using a variant of the Kamada–Kawai spring-embedding visualization, which takes link strengths into account (25).

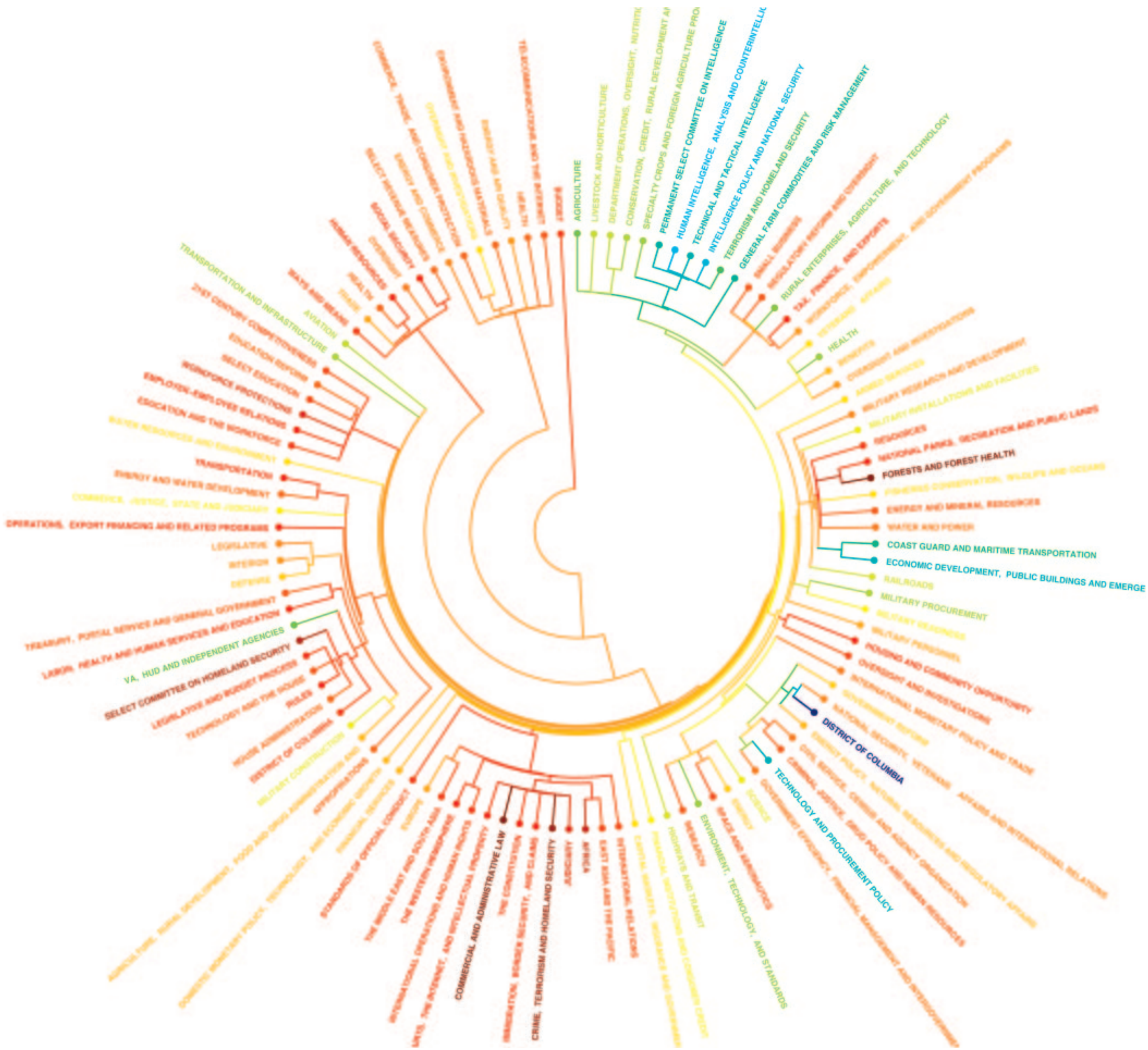


**Fig. 2.** Dendrogram representing the hierarchical clustering of the committees of the 107th U.S. House of Representatives, determined by single-linkage clustering on normalized committee interlocks. Each committee is color-coded according to the mean ''extremism'' of its members (defined in *Voting Patterns*), from less extreme (blue) to more extreme (red). The clusters at each level are color-coded according to the average of their constituent committee extremism scores.

## Structure of the House of Representatives

Connections between committees can be quantified in greater detail by applying the technique of single-linkage clustering (19). Starting with the complete set of committees for a given Congress, committees are joined together sequentially starting with the pair with the greatest normalized interlock value, followed by the next greatest, and so forth. This process generates ''clusters'' of committees, which can be represented by using a tree or dendrogram such as that shown in Fig. 2 for the 107th House. There appear to be essentially four hierarchical levels of clustering within this dendrogram: subcommittees, committees, groups of committees, and the entire House (20). (There is also some indication of a weaker fifth level of organization corresponding to groups of subcommittees inside larger standing committees.)

Here, we are interested primarily in the third hierarchical level: the connections between committees. For example, we see near the eight o'clock position in Fig. 2 a tightly grouped cluster that includes the House Rules Committee and the Select Committee on Homeland Security. Because assignments to select committees are ordinarily determined by drawing selectively from legislative bodies with overlapping jurisdiction, one might naively expect a close connection between the Select Committee on Homeland Security and, for example, the Terrorism and Homeland Security Subcommittee (of the Intelligence Select Committee), which was formed originally as a bipartisan ''working group'' and was designated on September 13, 2001, by House Speaker Dennis Hastert (Republican, Illinois) as the lead congressional entity assigned to investigate the September 11, 2001, terrorist attacks (see www.homelandsecurity.org/journal/

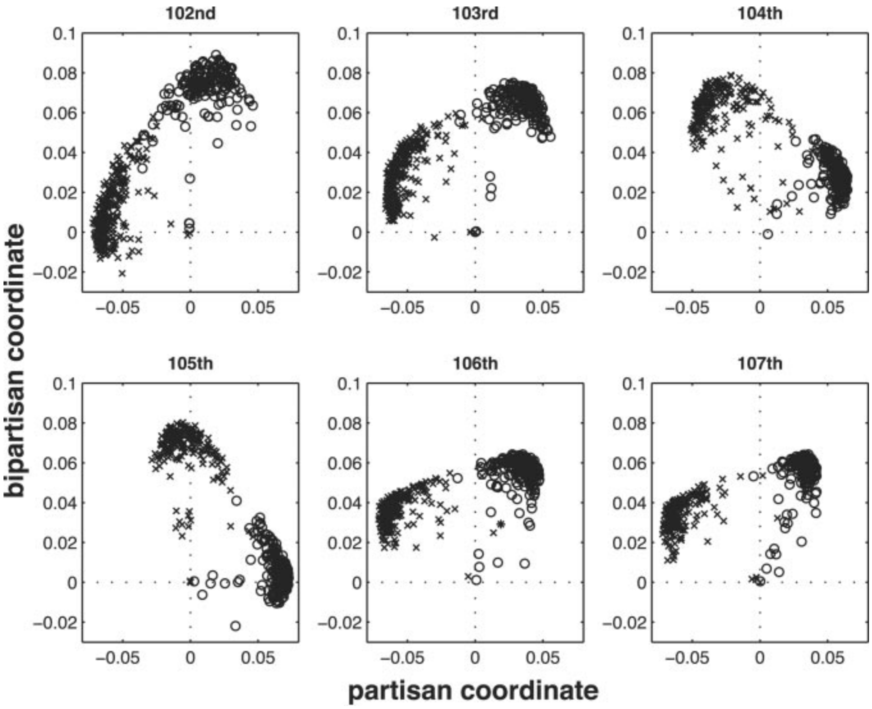


**Fig. 3.** SVD of the voting record of the House of Representatives in the 102nd–107th U.S. Congresses. Each point represents a projection of a Representative's votes onto eigenvectors corresponding to the leading two singular values. The two axes are denoted ''partisan'' and ''bipartisan,'' as described in *Voting Patterns.* Democrats (×) are shown on the left, and Republicans (○) are shown on the right. The few independents (*) are also shown.

articles/kaniewskilegislative.htm). However, the 107th Homeland Security Committee shares only one common member (normalized interlock value of 2.4) with the Intelligence Select Committee (located near the one o'clock position in Fig. 2) and has no interlock with any of the four Intelligence subcommittees.

## Voting Patterns

An additional twist can be introduced by considering how the network of interlocks between committees is related to the political positions of their constituent Representatives. One way to characterize political positions is to tabulate the voting records of individuals on key issues, but such a method is subjective by nature, and a method that involves less personal judgment on the part of the observer is preferable. Here, we use a singular value decomposition (SVD) (21) of voting records (22–24). Other data-mining methods can also be used (see www.ailab.si/aleks/politics).

We define an $n \times m$ voting matrix $\mathbf{B}$ with one row for each of the $n$ Representatives in the House and one column for each of the $m$ votes taken during the session. For example, the 107th House had $n = 444$ Representatives (including midterm replacements) and took $m = 990$ roll-call votes. The element $B_{ij}$ is $+1$ if Representative $i$ voted "yea" on vote $j$ and $-1$ if he or she voted "nay." If a Representative did not vote because of absence or abstention, the corresponding element is 0.

The SVD identifies groups of Representatives who voted in a similar fashion on many votes. The grouping that has the largest mean-square overlap with the actual groups voting for or against each measure is given by the leading (normalized) eigenvector $\mathbf{u}^{(1)}$ of the matrix $\mathbf{B}^T\mathbf{B}$, the next largest by the second eigenvector, and so on (21, 24). If we denote by $\sigma_k^2$ the corresponding eigenvalues (which are provably nonnegative) and by $\mathbf{v}^{(k)}$ the normalized eigenvectors of $\mathbf{B}\mathbf{B}^T$ (which have the same eigenvalues), then it can be shown that

$$B_{ij} = \sum_{k=1}^{n} \sigma_k u_i^{(k)} v_j^{(k)}, \qquad [1]$$

and that the matrix $\mathbf{B}^{(r)}$ with elements

$$B_{ij}^{(r)} = \sum_{k=1}^{r} \sigma_k u_i^{(k)} v_j^{(k)} \qquad [2]$$

approximates the full voting matrix $\mathbf{B}$, with the sum of the squares of the errors on the elements equal to $\Sigma_{k=r+1}^{n}\sigma_k^2$, which vanishes in the limit $r \rightarrow n$. Assuming that the quantities $\sigma_k$, which are called the "singular values," are ordered such that $\sigma_1 \geq \sigma_2 \geq \sigma_3$ …, then $\mathbf{B}^{(r)}$ is an excellent approximation to the original voting matrix if the singular values fall off sufficiently rapidly with increasing $k$.

Alternatively, one can say that the $k$th term in the SVD (Eq. **1**) accounts for a fraction $\sigma_k^2/\Sigma_{k=1}^{n}\sigma_k^2$ of the sum of the squares of the elements of the voting matrix. For the 107th House, we find that the leading eigenvector accounts for $\approx$45.3% of the voting matrix, the second eigenvector accounts for $\approx$29.6%, and no other eigenvector accounts for $>$1.6%. Thus, to an excellent approximation, a Representative's voting record can be characterized by just two coordinates, measuring the extent to which they align (or do not align) with the groups represented by the first two eigenvectors. That is,

$$B_{ij}^{(2)} = \sigma_1 u_i^{(1)} v_j^{(1)} + \sigma_2 u_i^{(2)} v_j^{(2)} \qquad [3]$$

is a good approximation to $B_{ij}$. Similar results are obtained for other sessions of Congress, with two eigenvectors giving a good approximation to the voting matrix in every case. It has been shown previously by using other methods that congressional voting positions are well approximated by just two coordinates

**Table 2. SVD rank ordering of the most and least partisan Representatives in the 107th U.S. House**

| Least partisan | Farthest left | Farthest right |
|---|---|---|
| K. Lucas (R) | J. D. Schakowsky (D) | T. G. Tancredo (R) |
| C. A. Morella (R) | J. P. McGovern (D) | J. B. Shadegg (R) |
| R. M. Hall (D) | H. L. Solis (D) | J. Ryun (R) |
| R. Shows (D) | L. C. Woolsey (D) | B. Schaffer (R) |
| G. Taylor (R) | J. F. Tierney (D) | P. Sessions (R) |
| C. W. Stenholm (D) | S. Farr (D) | S. Johnson (R) |
| R. E. Cramer (D) | N. Pelosi (D) | B. D. Kerns (R) |
| V. H. Goode (R) | E. J. Markey (D) | P. M. Crane (R) |
| C. John (D) | J. W. Olver (D) | W. T. Akin (R) |
| C. C. Peterson (D) | L. Roybal-Allard (D) | J. D. Hayworth (R) |

The first column gives the least-partisan Representatives as determined by an SVD of the roll-call votes. The second column gives the SVD rank ordering of the most partisan Representatives. They are all Democrats; thus, this also gives the rank of the Representatives farthest to the left. The third column gives the rank of the Representatives farthest to the right. The SVD rank ordering was determined for Representatives after midterm replacements (432 total congressmen) using all 990 roll calls; it classifies 92.7% of individual votes correctly. By contrast, in Poole and Rosenthal's Optimal Classification method (23), a rank ordering of the Representatives in the 107th House is determined by using 443 total Representatives and 749 of 990 roll calls (votes with <0.5% of the votes in the minority were removed from consideration). It classifies 92.8% of the individual Representatives' votes correctly. R, Republican; D, Democrat.

(see www.voteview.com), but the SVD does so in a particularly simple fashion directly from the roll-call data. In Fig. 3, we plot the two coordinates for every member of the House for each of the 102nd–107th Congresses.

We find that one of the two leading eigenvectors (the first in the 101st–105th Houses; the second in the 106th and 107th Houses) corresponds closely to the acknowledged political party affiliation of the Representatives, with Democrats (×) shown on the left and Republicans (○) shown on the right in the plots. Therefore, we call this the "partisan coordinate," and Representatives who score highly on it (either positively or negatively) tend often to vote with members of their own party. From the partisan coordinate, we also compute a measure of "extremism" for each Representative as the absolute value of their partisan coordinate relative to the mean partisan score of the full House. That is, we define the extremism $e_i$ of a Representative by $e_i = |p_i - \mu|$, where $p_i$ is the Representative's partisan coordinate, and $\mu$ is the mean of that coordinate for the entire House (which is usually skewed slightly toward the majority party). In Table 2, we list the most and least partisan Representatives from each party computed from the roll call of the 107th House. We also compare the vote reconstruction to those obtained by using an alternative method, the Optimal Classification technique of Poole and Rosenthal (23) (as applied only to the 107th House).

In contrast, the other leading eigenvector groups essentially all Representatives together regardless of party affiliation and thus appears to represent voting actions in which most members of the House either approve or disapprove of a motion simultaneously. We call this the "bipartisan coordinate," because Representatives who score highly on it tend often to vote with the majority of the House.

Using our SVD results, we also can calculate the positions of the votes (as opposed to the voters) along the same two leading dimensions to quantify the nature of the issues being decided. We show this for the 107th House in Fig. 4. One application of this analysis is a measurement of the reproducibility of individual



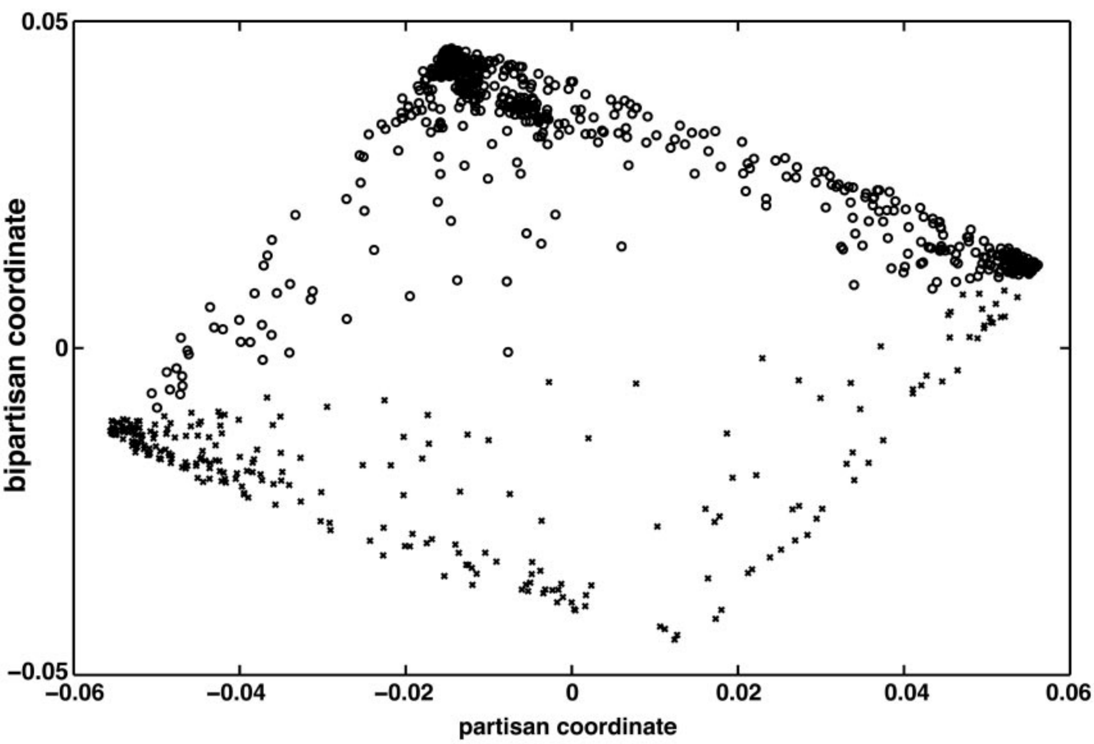


**Fig. 4.** SVD of the roll call of the 107th House projected onto the voting coordinates. Points represent projections of the votes cast on a measure onto eigenvectors corresponding to the leading two singular values. There is a clear separation between measures that passed (○) and those that did not (×). The four corners of the plot are interpreted as follows: measures with broad bipartisan support (north) all passed; those supported mostly by the right (east) passed because the Republicans constituted the majority party of the 107th House; measures supported by the left (west) failed because of the Democratic minority; and the (obviously) very few measures supported by almost nobody (south) also failed.

votes and outcomes. By reconstituting the voting matrix using only the information contained in the two leading singular values and the corresponding eigenvectors and subsequently summing the resulting approximated votes over all Representatives, we derive a single score for each vote. Making a simple assignment of "pass" to those votes that have a positive score and "fail" to all others successfully reconstructs the outcome of 984 of the 990 total votes ($\approx$99.4%). [Overall, 735 ($\approx$74.2%) of the votes passed; thus, simply guessing that every vote passed would be considerably less effective.] If we throw out known absences and abstentions, the analysis still identifies 975 of the 990 results correctly. Even with the most conservative measure of this computation's success rate, in which we throw out abstentions and absences first and then examine yeas/nays of the individual Representatives ($\approx$92.7% of which are correctly identified by the sign of the elements in the projection of the voting matrix), the two-dimensional reconstruction still identifies 939 ($\approx$94.8%) of the votes correctly. We repeated these calculations for the 101st–106th Houses and found similar results in each case. [The remarkable accuracy of SVDs in reconstructing votes was observed previously for the example of the U.S. Supreme Court (24).]

The SVD analysis gives a simple way of classifying the voting positions of Representatives without making subjective judgments. In Fig. 2, we have combined our clustering analysis of committees with the SVD results by color coding each committee according to the mean "extremism" of its members, so that committees populated by highly partisan members of either party appear in red and committees containing more moderate Representatives appear in blue. Taking again the examples of Intelligence and Homeland Security, the figures immediately identify the former as moderate and the latter as more partisan. Indeed, the Select Committee on Homeland Security has a larger mean extremism than any of the 19 standing committees and has the fourth largest mean extremism among the 113 committees of the 107th House, which perhaps is not so surprising when we see that its constituent Representatives included House Majority Leader Richard Armey (Republican, Texas), Majority Whip Tom DeLay (Republican, Texas), and Minority Whip Nancy Pelosi (Democrat, California). However, this characterization of the committee was made mathematically, using no knowledge beyond the roll-call votes of the 107th House. As another example, the 107th House Rules Committee is the second most extreme of the 19 standing committees (after the Judiciary Committee) and ranks 18th out of 113 committees overall. By contrast, the Permanent Select Committee on Intelligence of the 107th House has a smaller mean extremism than any of the 19 standing committees, and the Intelligence Committee and its four subcommittees all rank among the 10 least extreme of all 113 committees.

## Conclusions

To conclude, a network-theory approach coupled with an SVD analysis of roll-call votes is demonstrably useful in analyzing organizational structure in the committees of the U.S. House of Representatives. We found evidence of several levels of hierarchy within the network of committees and, without incorporating any knowledge of political events or positions, identified some close connections between committees, such as that between the House Rules Committee and the Select Committee on Homeland Security in the 107th Congress, as well as correlations between committee assignments and political positions of the Representatives. Our analysis of committee interlocks and voting patterns strongly suggests that committee assignments are not determined at random (i.e., that some committees are indeed stacked) and also indicates the degree of departure from randomness. We have discussed here a few observations in detail, but a rich variety of other results can be derived from similar analyses. We hope that additional studies using similar techniques will provide key insights into the structure of the House of Representatives and other political bodies.

We gratefully acknowledge Thomas Callaghan, Michael Cohen, Gordon Kingsley, and Scott Page for useful conversations during this research. We also thank Michael Abraham for developing some of the computer code used in this research and Thomas Bartsch, Steven Lansel, Tom Maccarone, Slaven Peleš, Steve Van Hooser, and Julia Wang for critical readings of the manuscript. We also thank the two anonymous referees, whose constructive comments improved the manuscript greatly. We obtained the roll calls for the 102nd–107th Congresses from the Voteview web site (http://voteview.com), the roll calls for the 101st Congress from the Inter-University Consortium for Political and Social Research (www.icpsr.umich.edu), and the committee assignments for the 101st–108th Congresses from the web site of the House of Representatives' Office of the Clerk (http://clerk.house.gov). This work was supported in part by the National Science Foundation, including the contribution from C.M.W. through a National Science Foundation Vertical Integration of Research and Education (VIGRE) Research Experiences for Undergraduates program and M.A.P.'s visiting faculty position at Georgia Institute of Technology.

1. Boyce, J. R. & Bischak, D. P. (2002) *J. Law. Econ. Organ.* **18,** 1–38.
2. Niskanen, W. A. (1971) *Bureaucracy and Representative Government* (Aldine–Atherton, Chicago).
3. Gilligan, T. W. & Krehbiel, K. (1987) *J. Law. Econ. Organ.* **3,** 287–335.
4. Krehbiel, K. (1990) *Am. Polit. Sci. Rev.* **84,** 149–163.
5. Cox, G. W. & McCubbins, M. D. (1993) *Legislative Leviathan: Party Government in the House* (Univ. of California Press, Berkeley).
6. Shepsle, K. A. & Weingast, B. R., eds. (1995) *Positive Theories of Congressional Institutions: A Comparison of Rational Choice Models of Congress* (Univ. of Michigan Press, Ann Arbor).
7. Strogatz, S. H. (2001) *Nature* **410,** 268–276.
8. Newman, M. E. J. (2003) *SIAM Rev.* **45,** 167–256.
9. Mariolis, P. (1975) *Soc. Sci. Q.* **56,** 425–439.
10. Useem, M. (1984) *The Inner Circle: Large Corporations and the Rise of Business Political Activity in the US and UK* (Oxford Univ. Press, Oxford).
11. Mintz, B. & Schwartz, M. (1985) *The Power Structure of American Business* (Univ. of Chicago Press, Chicago).
12. Robins, G. L. & Alexander, M. (2004) *J. Comput. Math. Organ. Theor.* **10,** 69–94.
13. Mizruchi, M. S. (1996) *Annu. Rev. Sociol.* **22,** 271–298.
14. Haunschild, P. R. (1993) *Adm. Sci. Q.* **38,** 564–592.
15. Rao, H., Davis, G. F. & Ward, A. (2000) *Adm. Sci. Q.* **45,** 268–292.
16. Davis, G. F., Yoo, M. & Baker, W. E. (2003) *Strateg. Organ.* **1,** 301–326.
17. Burt, R. S. (2005) *Brokerage and Closure: An Introduction to Social Capital* (Oxford Univ. Press, Oxford).
18. Burt, R. S. (2004) *Am. J. Sociol.* **110,** 349–399.
19. Johnson, S. C. (1967) *Psychometrika* **32,** 241–254.
20. Strahler, A. N. (1952) *Bull. Geol. Soc. Am.* **63,** 1117–1142.
21. Golub, G. H. & Van Loan, C. F. (1996) *Matrix Computations* (The Johns Hopkins Univ. Press, Baltimore), 3rd Ed.
22. Poole, K. T. & Rosenthal, H. (1997) *Congress: A Political-Economic History of Roll Call Voting* (Oxford Univ. Press, Oxford).
23. Poole, K. T. & Rosenthal, H. (2000) *Polit. Anal.* **8,** 211–237.
24. Sirovich, L. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 7432–7437.
25. Kamada, T. & Kawai, S. (1989) *Inf. Process. Lett.* **31,** 7–15.