Political Network Analysis

POLISCI 791P-01
Location: Machmer E-23
Dates/Time: 1/22/2013–5/1/2013, M, 12:30pm-3pm
Instructor: Bruce A. Desmarais
desmarais@polsci.umass.edu
Office: 420 Thompson Hall

Course Overview: A network is a data structure composed of units and the relationships connecting them. The study of networks in political science, the social sciences and beyond has grown rapidly in recent years. This course is a comprehensive introduction to methods for analyzing network data. We will cover network data collection and management, the formulation of network theory and hypotheses, network visualization and description; and methods for the statistical analysis of networks. The course will make extensive use of real-world applications and students will gain a thorough background in the use of network analytic software. Most of the applications discussed will be drawn from political science, but this course will be relevant to anyone interested in the study of network data.

Course Objectives: In this course we will cover a sampling of important material in several topics. The objective is not to develop exhaustive coverage of any one topic area. Rather, the goal is that the material in each area will be mastered with sufficient depth that students will be able to independently broaden their expertise on a topic through independent study. The broad course objectives are that students will develop:

1. Fluency in the language of networks analysis; an in-depth understanding of the concepts that have proven most useful in the study of networks.
2. Awareness regarding how theory and hypotheses for networks are structured.
3. Command of network analysis software.
4. Understanding of how to explore and describe network data.
5. The ability to statistically model network data and formally test hypotheses about networks.
6. Practical experience in conducting research with network data.

Books: Students are not required to purchase any books for this course. Some of the readings come from the following books, but the required material is available on e-reserves. However, all three of these books make excellent references for network analysts, so students are encouraged to purchase them if possible.

1. Newman (2010)
2. Wasserman and Faust (1997)
3. Lusher, Koskinen and Robins (2012)

Prerequisites: This course will be accessible to students without prior training in quantitative research methods. However, students with background in basic descriptive and inferential statistics will likely get more out of the course than those who need to contemporaneously traverse these topics. Understanding of descriptive statistics, hypothesis testing, regression analysis, and some experience with a scripting-based statistical software will accelerate comprehension of the material.

Computing: All computing will be conducted in the `R` statistical software. We will use add-on packages, mostly from the `statnet` suite - `http://csde.washington.edu/statnet/`. It is strongly advisable that students download `R` onto a laptop and bring the laptop to class every week. The course will include an introduction to `R` for those unfamiliar with the software and we will regularly walk through applications during class. If students find they would like additional training in the use of `R` , the new Institute for Social Science Research at UMass is an excellent resource. Their consultants offer support in `R` - `http://www.umass.edu/issr/research/consultation/`.

Data: Many example datasets will be provided in order to illustrate application of the methods. Additionally, the course will provide ample opportunity for students to apply the methods to their own data. Appropriate data will include observations of tie existence/absence among actors (i.e., the network), covariates that correspond to dyads (e.g., other network variables, distances between actors in geographic space or some other metric), and data on actor attributes (i.e., node-level covariates). For homework, students may either use their own datasets or example datasets provided with the course materials.

Homework: One of the core objectives of this course is that students develop an ability to evaluate the applicability of the methods that we cover to their own research. There will be at least one homework for each of the top-level topics listed in the course schedule. For

each homework, students will be expected to develop a preliminary application of selected methods covered in the respective section. Homework is worth 30% of the final grade.

Online Quiz: Each week there will be a short multiple choice quiz posted on Moodle, designed to test students' comprehension of the reading. These should be completed individually before coming to class on Monday. Worth 15% of final grade.

Research Paper and CSS Poster Session: Students are required to complete a research paper. It should be a full paper, not a research design. The paper must contain an application of descriptive network analysis and an application of inferential network analysis. Students are required to present their research, in the form of a poster, at the Computational Social Science Seminar on 4/26, 12-2:30pm in LGRT 1634. If this presents a prohibitive scheduling conflict, please notify the instructor as soon as possible so that we can make alternative arrangements. The research paper and presentation is worth 40% of the final grade.

Collaboration: This course will include several exciting opportunities for collaboration on homework assignments and on the research paper. Prof. Krista Gile, assistant professor of statistics and core faculty affiliate of the computational social science initiative, is teaching a graduate course on statistical methods for the analysis of network data (STATISTC 697NS). Students in Prof. Gile's class and the current one will be asked to collaborate on selected homework assignments. Additionally, the research project requirement may be satisfied through a collaborative project with students in Prof. Gile's class, among students in the current class or both.

Project Labs: Two of the class periods will be devoted to 'Project Labs'. Students will be asked to deliver a 5-10 minute presentation describing the progress of their projects. We will discuss each project in detail. Participation in the project labs is worth 15% of the final grade.

Course Schedule: The schedule below gives the required reading. The readings listed for a particular day should be read before class time that day. The full citations for the readings can be found below in the references section.

1. **Section One:** Introduction to network data, network analysis and `R`

   1/28: Introduction to Network terminology and Network Data

   – Wasserman and Faust (1997) Chs. 1 – 2

- Lazer (2011)

2/4: Introduction to `R` and Network Analysis

- Free from UMass library website – Dalgaard (2008) Ch. 1
- Complete this tutorial – `http://www.cyclismo.org/tutorial/R/input.html`
- Butts (2008*a*)

2/11: Network theory and hypotheses

- Carpenter, Esterling and Lazer (2004)
- Mutz (2002)
- Ward, Stovel and Sacks (2011)

2. **Section Two:** Description and exploration of networks

2/19: Network Descriptive Statistics

- Newman (2010), Ch. 7
- Butts (2008*b*)
- Fowler (2006)

2/25: Network Visualization

- Freeman (2000)
- Read through the following documentation and run all of the examples at the end - `http://igraph.sourceforge.net/doc/R/plot.common.html`
- Bommarito, Katz and Zelner (2009)
- See `http://vimeo.com/12543669`

3/4: Community Detection

- Zhang, Friend, Traud, Porter, Fowler and Mucha (2008)
- Macon, Mucha and Porter (2012)
- Read through the following documentation and run the example at the end - `http://igraph.sourceforge.net/doc/R/communities.html`

3/11: Projects Lab

3. **Section Three:** Inferential network analysis: exponential random graph models.

3/25: Statistical Background and Intro to ERGM

- Lusher, Koskinen and Robins (2012), Chs. 2–4

- Goodreau, Kitts and Morris (2009)

4/1: ERGM Application and Interpretation

- Hunter, Handcock, Butts, Goodreau and Morris (2008)
- Cranmer and Desmarais (2011)

4/8: ERGM post-estimation

- Berardo and Scholz (2010)
- Desmarais and Cranmer (2012)

4/15: Projects Lab

References

Berardo, Ramiro and John T. Scholz. 2010. "Self-Organizing Policy Networks: Risk, Partner Selection, and Cooperation in Estuaries." *American Journal of Political Science* 54(3):632–649.

Bommarito, II, Michael J., Daniel Katz and Jon Zelner. 2009. Law as a seamless web?: comparison of various network representations of the United States Supreme Court corpus (1791-2005). In *Proceedings of the 12th International Conference on Artificial Intelligence and Law*. ICAIL '09 pp. 234–235.

Butts, Carter T. 2008*a*. "network: A Package for Managing Relational Data in R." *Journal of Statistical Software* 24(2):1–36.
URL: *http://www.jstatsoft.org/v24/i02*

Butts, Carter T. 2008*b*. "Social Network Analysis with sna." *Journal of Statistical Software* 24(6):1–51.

Carpenter, Daniel P., Kevin M. Esterling and David M. J. Lazer. 2004. "Friends, Brokers, and Transitivity: Who Informs Whom in Washington Politics?" *The Journal of Politics* 66:224–246.

Cranmer, Skyler J. and Bruce A. Desmarais. 2011. "Inferential Network Analysis with Exponential Random Graph Models." *Political Analysis* 19(1):66–86.

Dalgaard, Peter. 2008. *Introductory statistics with R*. New York, NY: Springer.

Desmarais, Bruce A. and Skyler J. Cranmer. 2012. "Micro-Level Interpretation of Exponential Random Graph Models with Application to Estuary Networks." *Policy Studies Journal* 40(3):402–434.

Fowler, James H. 2006. "Connecting the Congress: A Study of Cosponsorship Networks." *Political Analysis* 14(4):pp. 456–487.

Freeman, Linton. 2000. "Visualizing Social Networks." *Journal of Social Structure* 1(1).

Goodreau, Steven .M., James A. Kitts and Martina Morris. 2009. "Birds of a feather, or friend of a friend? Using exponential random graph models to investigate adolescent social networks." *Demography* 46(1):103–25.

Hunter, David R., Mark S. Handcock, Carter T. Butts, Steven M. Goodreau and Martina Morris. 2008. "ergm: A Package to Fit, Simulate and Diagnose Exponential-Family Models for Networks." *Journal of Statistical Software* 24(3):1–29.

Lazer, David. 2011. "Networks in Political Science: Back to the Future." *PS: Political Science & Politics* 44:61–68.

Lusher, Dean, Johan Koskinen and Garry Robins. 2012. *Exponential Random Graph Models for Social Networks*. New York, NY: Cambridge University Press.

Macon, Kevin T., Peter J. Mucha and Mason A. Porter. 2012. "Community structure in the United Nations General Assembly." *Physica A: Statistical Mechanics and its Applications* 391(1):343 – 361.

Mutz, Diana C. 2002. "Cross-Cutting Social Networks: Testing Democratic Theory in Practice." *The American Political Science Review* 96(1):pp. 111–126.

Newman, M.E.J. 2010. *Networks*. New York, NY: Oxford University Press.

Ward, Michael D., Katherine Stovel and Audrey Sacks. 2011. "Network Analysis and Political Science." *Annual Review of Political Science* 14(1):245–264.

Wasserman, Stanley and Katherine Faust. 1997. *Social Network Analysis*. New York, NY: Cambridge University Press.

Zhang, Yan, A.J. Friend, Amanda L. Traud, Mason A. Porter, James H. Fowler and Peter J. Mucha. 2008. "Community Structure in Congressional Cosponsorship Networks." *Physica A* 387(7):pp. 1705 – 1712.