Network Analysis: Advanced Topics

Inferential Network Analysis

Instructor: Bruce A. Desmarais is, Penn State University
desmarais@polsci.umass.edu
Office: 221
Office Hours: 9-10 & 12:30-1:30

TA: Matt Denny, Penn State University
matthewjdenny@gmail.com
Office: TBD
Office Hours: TBD

Course Overview: This is a course on inferential network analysis. The conventional categorization of data analytic methods into descriptive and inferential statistics can be fruitfully applied to network analysis. Descriptive methods of network analysis are important for illuminating structural features of a given network, but they cannot be used to build and/or test theories about the generation of networks. Inferential methods of network analysis can be used to test hypotheses about the generation and evolution of a network, derive measures of uncertainty for network indices, and find probabilistic models that accurately describe the overall features of a network. The first week will focus on exponential random graph models (ERGMs), which can be parametrized to represent complex dependence processes and the effects of exogenous covariates. The second week will cover latent space models and quadratic assignment procedure, which are both designed to account for network dependencies without formulating complex network statistics. In the third and fourth weeks we will cover statistical models for longitudinal networks. These will include a longitudinal extension of the ERGM – the Temporal ERGM, and the actor-oriented model of network dynamics (i.e., SIENA). We will present each model mathematically, discuss published social science applications of them, and utilize the models on example datasets.

Course Objectives: This course is designed as a series of weekly modules that build upon each other. Each module covers one or more state-of-the-art approaches to statistical analysis of network data. For each model covered, the objectives are that students will:

1. Develop a firm grasp of the assumptions and formulation of the model.
2. Understand how to interpret results.
3. Understand how to evaluate the fit and assumptions of the model.
4. Use software to apply the mode.
5. Read published substantive applications of the model.
6. Evaluate the potential for application of the model to their own research.

Two overarching objectives are that students will (1) develop an ability to compare the relative merits of the various models covered for a given empirical application and (2) develop a comprehensive sense of the state of the literature on statistical models for social networks.

Prerequisites: Students in this course are expected to have a background in descriptive network analysis. The *Network Analysis* course taught by Ann McCranie during the first four weeks of the summer program would provide excellent preparation for this course. Familiarity with statistical modeling will also be essential. If you know how to run and interpret a logistic regression model, you have a sufficient background in statistics. The course on *Maximum Likelihood Estimation for Generalized Linear Models*, taught by Dean Lacy during the first session of the summer program would provide excellent preparation in statistics.

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. The course will include an introduction to `R` for those unfamiliar with the software. However, for those who would like to develop an in-depth background, I strongly recommend taking the *R Statistical Computing Environment* offered by John Fox in the first session.

Homework: Students will have a weekly homework typically assigned on Wednesday and due on Monday. For each homework, students will be expected to develop a preliminary application of the methods covered that week to their own data or to a dataset made available by the instructor. Each student will hand in and present their application briefly on Monday, with time for each presentation dependent upon enrollment in the course. We will discuss each presentation as a class and provide feedback. The write-up should be 1–3 pages, single-spaced.

Grading: Grades will be assigned based on performance on the homework assignments. Each assignment will be graded (0 – no submission, 1 – does not demonstrate comprehension of the methodology, 2 – adequate comprehension demonstrated, 3 – excellent comprehension demonstrated).

Grade Scale (based on sum of assignment grades):

- A+ (12)
- A (10–11)
- A- (8–9)
- B+ (6–7)
- B- (< 6)

Application Sessions: During each application session (e.g., "ERGM Application", "LSM Application") we will run through applications `R` during. Students are strongly encouraged to bring their laptops to class. Students will be provided with data, but may also use their own datasets.

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). Moreover, datasets in which all variables are observed at multiple points in time will be particularly appropriate. If you have any questions about whether a given dataset would be appropriate, please contact me immediately.

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. **Week One:** Exponential Random Graph Models
   - **Monday:** No Meeting
   - **Tuesday:** Course Overview and Introduction to Network Analysis, (Wasserman and Faust, 1997) Ch.1–2
   - **Wednesday:** Introduction to R for Network Analysis,(Butts, 2008)
   - **Thursday:** ERGM Introduction (Lusher, Koskinen and Robins, 2012; Cranmer and Desmarais, 2011; Goodreau, Kitts and Morris, 2009, chs.2-4)
   - **Friday:** ERGM Application (Hunter, Handcock, Butts, Goodreau and Morris, 2008)
2. **Week Two:** Alternatives to ERGM and
   - **Monday:** ERGM for Bipartite Networks
   - **Tuesday:** ERGM Wrap-up and HW Discussion
   - **Wednesday:** LSM (Hoff, Raftery and Handcock, 2002; Krivitsky and Handcock, 2008)
   - **Thursday:** QAP (Krackhardt, 1988; Dreiling and Darves, 2011; Butts, 2008)
   - **Friday:** Network Autocorrelation (Butts, 2008)
3. **Week Three:** Longitudinal Network Analysis I
   - **Monday:** QAP/LSM/Network Autocorrelation Wrap-up and HW Discussion
   - **Tuesday:** TERGM Introduction (Robins and Pattison, 2001; Hanneke, Fu and Xing, 2010; Desmarais and Cranmer, 2012*b*)
   - **Wednesday:** TERGM Application (Cranmer, Desmarais and Kirkland, 2012)
   - **Thursday:** SIENA Introduction (Snijders, van de Bunt and Steglich, 2010; Desmarais and Cranmer, 2012*a*)
   - **Friday:** SIENA Application `http://www.stats.ox.ac.uk/~snijders/siena/RSiena_Manual.pdf`
4. **Week Four:** Longitudinal Network Analysis II
   - **Monday:** Network Inference
   - **Tuesday:** Longitudinal Network Analysis Lab
   - **Wednesday:** Longitudinal Network Analysis Wrap-up and HW Discussion
   - **Thursday:** Project Presentations

References

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

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

Cranmer, Skyler J., Bruce A. Desmarais and Justin H. Kirkland. 2012. "Toward a Network Theory of Alliance Formation." *International Interactions* 38(3):295–324.

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

Desmarais, Bruce A. and Skyler J. Cranmer. 2012*b*. "Statistical Mechanics of Networks: Estimation and Uncertainty." *Physica A* 391(4):1865–1876.

Dreiling, Michael and Derek Darves. 2011. "Corporate Unity in American Trade Policy: A Network Analysis of Corporate-Dyad Political Action." *American Journal of Sociology* 116(5):pp. 1514–63.

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.

Hanneke, Steve, Wenjie Fu and Eric P. Xing. 2010. "Discrete Temporal Models of Social Networks." *Electronic Journal of Statistics* 4:585–605.

Hoff, Peter D., Adrian E. Raftery and Mark S. Handcock. 2002. "Latent Space Approaches to Social Network Analysis." *Journal of the American Statistical Association* 97(460):pp. 1090–1098.

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.

Krackhardt, David. 1988. "Predicting with networks: Nonparametric multiple regression analysis of dyadic data." *Social Networks* 10(4):359 – 381.

Krivitsky, Pavel N. and Mark S. Handcock. 2008. "Fitting Latent Cluster Models for Networks with latentnet." *Journal of Statistical Software* 24(5):1–23.
URL: *http://www.jstatsoft.org/v24/i05*

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

Robins, Garry and Philippa Pattison. 2001. "Random Graph Models for Temporal Processes in Social Networks." *Journal of Mathematical Sociology* 25(1):5 – 41.

Snijders, Tom A.B., Gerhard G. van de Bunt and Christian E.G. Steglich. 2010. "Introduction to stochastic actor-based models for network dynamics." *Social Networks* 32(1):44 – 60. Dynamics of Social Networks.

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