Syllabus
Network Analysis
XXXXX XXXXXX

Instructor: Jared Edgerton
Office: 2032 Derby Hall
Email: edgerton.37@osu.edu
Office Hours: TBD
Meeting Place & Time: TBD
Course Web Site: `jaredfedgerton.net/classes`

Rationale and Scope

This course aims to present statistical models for network data in detail. The course will integrate theoretical discussions with practical examples and software code to perform analyses.

Just like any other area of statistics, network analytic procedures can be divided into two categories – descriptive and inferential. While we will spend some time at the beginning of the semester reviewing descriptive network analysis, this course assumes you are familiar with the basics of network analysis (e.g. measures of centrality, methods of visualization) and begins where an intro course covering such topics ends. Methods of descriptive network analysis are suitable for many worthwhile research pursuits, but are inadequate for research problems that demand precise hypothesis testing with network data, or stochastic simulation of network processes. Within the last 20 years, methodological research on network analysis has seen several groundbreaking innovations in model formulation/specification and computation. The focus of this course is to cover the most important of these innovations theoretically, and then get practical experience working with their implementations in open source software. We will cover four general classes of models for statistical inference with networks stochastic block models, inferential network analysis (i.e., latent space, multiplicative effects, and exponential random graph models, network diffusion models, and multiplex network analysis. The reading load will be comparatively light for a graduate course (on average, 1-2 chapters per week and maybe an article), but will be more demanding in terms of coding requirements for the homeworks in pursuit of a strong final paper.

Prerequisites

The course also assumes a working knowledge of non-network based statistics as well as concepts required for that (e.g. calculus and linear algebra).

Evaluation

Your final grade will be based on several problem sets (40%) throughout the semester (many of which will be designed to help you along with your final paper), a final paper in which you produce a high quality manuscript (e.g. one that could eventually be published) using the techniques we cover (40%), and the presentation of this paper to the class and a general audience (20%). You should complete the scheduled reading *before the class listed!*.

I subscribe to OSU's grading rubric: A 93-100, A- 90-92.9, B+ 87-89.9, B 83-86.9, B- 80-82.9, C+ 77-79.9, C 73-76.9, C- 70-72.9, D+ 67-69.9, D 60-66.9, E 0-59.

Academic Misconduct

It is the responsibility of the Committee on Academic Misconduct to investigate or establish procedures for the investigation of all reported cases of student academic misconduct. The term "academic misconduct" includes all forms of student academic misconduct wherever committed; illustrated by, but not limited to, cases of plagiarism and dishonest practices in connection with examinations. Instructors shall report all instances of alleged academic misconduct to the committee (Faculty Rule 3335-5-487). For additional information, see the Code of Student Conduct http://studentlife.osu.edu/csc/.

Students with Disabilities

The University strives to make all learning experiences as accessible as possible. If you anticipate or experience academic barriers based on your disability (including mental health, chronic or temporary medical conditions), please let me know immediately so that we can privately discuss options. To establish reasonable accommodations, I may request that you register with Student Life Disability Services. After registration, make arrangements with me as soon as possible to discuss your accommodations so that they may be implemented in a timely fashion. The Office of Student Life Disability Services is located in 098 Baker Hall, 113 W. 12th Avenue; telephone 614-292-3307, slds@osu.edu; http://slds.osu.edu/.

Course Norms

- Speak up when you have a question.
- Teamwork and collaboration is *highly encouraged* on every aspect of the course. Students may work together on assignments but must write out their own homework and list everyone they worked with. However, you are not allowed to divvy up the problems such that one person does one problem an another the next. You are allowed to collaborate on the final paper if you like (max 2 authors and both get the same grade regardless of real or perceived contributions).
- *All* homework assignments must be written in LaTeX. Assignments not written in LaTeX (or `sweave` if you want to be really fancy) will be returned without a grade.

Texts

Skyler J. Cranmer, Bruce A. Desmarais, and Jason Morgan "Inferential Network Analysis" at *Cambridge University Press* 2020.

Thomas W. Valente "Network Models of the Diffusion of Innovations" at *Hampton Press* 1975.

Tentative Schedule

Part 1. Basics of Networks

- Basic concepts and properties of graphs, motivation for network analysis;
- Historical context of network analysis (in statistics, sociology, math, physics, computer science etc.);

– Network visualization;

– Application areas of network data; and

– Community detection.

Week 1 (TBD 8) **Introduction and the basics of networks, when to use network models, network centrality, power laws**

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook. “Birds of a feather: Homophily in social networks.” Annual Review of Sociology 27.1 (2001): 415-444.

Watts, Duncan J., and Steven H. Strogatz. “Collective dynamics of small-world networks.” Nature 393.6684 (1998): 440-442.

Clauset, Aaron, Cosma Rohilla Shalizi, and Mark EJ Newman. “Power-law distributions in empirical data.” SIAM Review 51.4 (2009): 661-703.

Page, Lawrence, et al. The PageRank citation ranking: Bringing order to the web. Stanford InfoLab, 1999.

Bickel, Peter J., and Aiyou Chen. “A nonparametric view of network models and Newman–Girvan and other modularities.” Proceedings of the National Academy of Sciences 106.50 (2009): 21068-21073.

Week 2 (TBD 15) **Random graph models, community detection, stochastic block models, network attributes**

Daudin, J-J., Franck Picard, and Stephane Robin. “A mixture model for random graphs.” Statistics and Computing 18.2 (2008): 173-183.

Newman, Mark EJ, Duncan J. Watts, and Steven H. Strogatz. “Random graph models of social networks.” Proceedings of the National Academy of Sciences 99.suppl 1 (2002): 2566-2572.

Blondel, Vincent D., et al. “Fast unfolding of communities in large networks.” Journal of Statistical Mechanics: theory and experiment 2008.10 (2008): P10008.

Choi, David S., Patrick J. Wolfe, and Edoardo M. Airoldi. ”Stochastic blockmodels with a growing number of classes.” Biometrika 99.2 (2012): 273-284.

Part 2. Dependence and Interdependence

Week 3 (TBD 29) **Network diffusion** *This lecture focuses specifically on how social networks impact a variety of political phenomena* and *showing the reader how to detect the presence of complex dependencies (e.g. violations of independence assumptions).*

– Creating latent diffusion networks; and

– Using real social networks to measure diffusion.

Thomas W. Valente "Network Models of the Diffusion of Innovations" at *Hampton Press* 1975.

Valente, Thomas W. "Social network thresholds in the diffusion of innovations." Social networks 18.1 (1996): 69-89.

Cranmer, Skyler J., Bruce A. Desmarais, and Benjamin W. Campbell. "The contagion of democracy through international networks." Social Networks 61 (2020): 87-98.

Linder, Fridolin, et al. "Text as policy: Measuring policy similarity through bill text reuse." Policy Studies Journal 48.2 (2020): 546-574.

Week 4 (TBD 29) **Network contagion, computation models** *This lecture focuses on the spread of ideas using simulated networks (Tying in COVID-19 transmission examples).*

Allen, Linda JS. "Some discrete-time SI, SIR, and SIS epidemic models." Mathematical biosciences 124.1 (1994): 83-105.

Larson, Jennifer M. "Networks and interethnic cooperation." The Journal of Politics 79.2 (2017): 546-559.

Dodds, Peter Sheridan, and Duncan J. Watts. "Universal behavior in a generalized model of contagion." Physical review letters 92.21 (2004): 218701.

Woo, Jiyoung, Jaebong Son, and Hsinchun Chen. "An SIR model for violent topic diffusion in social media." Proceedings of 2011 IEEE International Conference on Intelligence and Security Informatics. IEEE, 2011.

Part 3. Inferential network analysis

Week 5 (TBD 5) **Conditioning out network dependencies.** *These lectures focus on how to address dependence in network data. This lays the groundwork for how to explicitly model network characteristics instead of treating them as a nuisance.*

Jenatton, Rodolphe, et al. "A latent factor model for highly multi-relational data." Advances in neural information processing systems. 2012.

Hoff, Peter D. "Multiplicative latent factor models for description and prediction of social networks." Computational and mathematical organization theory 15.4 (2009): 261.

Jochmans, Koen, and Martin Weidner. "Fixed-Effect Regressions on Network Data." Econometrica 87.5 (2019): 1543-1560.

Fosdick, Bailey K., and Peter D. Hoff. "Testing and modeling dependencies between a network and nodal attributes." Journal of the American Statistical Association 110.511 (2015): 1047-1056.

Hoff, Peter D. "Bilinear mixed-effects models for dyadic data." Journal of the american Statistical association 100.469 (2005): 286-295.

Week 6 (TBD 12) **ERGM** *This section lays the theoretical groundwork for the introduction of the ERGM. Local emergence, self-organization, and the role of network topology. This lecture focuses on the exposition of the (very) many endogenous dependence structures that may be included in an ERGM. All discussions will proceed theoretically, mathematically, and present simulation studies of the behavior of each of these statistics.*

Skyler J. Cranmer, Bruce A. Desmarais, and Jason Morgan "Inferential Network Analysis" at *Cambridge University Press* 2020.

Week 7 (TBD 26) **ERG Type Models for Longitudinally Observed Networks** *Many substantively interesting network are not observed only once, but recur and are observed longitudinally. This lecture focuses on explicating extensions to the ERGM that allows the researcher to model longitudinally observed networks.*

Skyler J. Cranmer, Bruce A. Desmarais, and Jason Morgan "Inferential Network Analysis" at *Cambridge University Press* 2020.

Week 8 (TBD 5) **Latent space models and dimension reduction** *Projecting network structures onto low dimensional spaces to model binary and contentious data.*

Hoff, Peter D., Adrian E. Raftery, and Mark S. Handcock. "Latent space approaches to social network analysis." Journal of the American Statistical Association 97.460 (2002): 1090-1098.

Skyler J. Cranmer, Bruce A. Desmarais, and Jason Morgan "Inferential Network Analysis" at *Cambridge University Press* 2020.

Handcock, Mark S., Adrian E. Raftery, and Jeremy M. Tantrum. "Model-based clustering for social networks." Journal of the Royal Statistical Society: Series A (Statistics in Society) 170.2 (2007): 301-354.

Part 4. Complex networks

Week 9 (TBD 19) **Dynamic stochastic block models and complex networks.** *This lecture focuses on measuring communities across time with binary and continuous edges using stochastic block models.*

Boccaletti, Stefano, et al. "The structure and dynamics of multilayer networks." Physics Reports 544.1 (2014): 1-122.

Kivela, Mikko, et al. "Multilayer networks." Journal of complex networks 2.3 (2014): 203-271.

Week 10 (TBD 26) **Dynamic stochastic block models and complex networks (continued).** *This lecture focuses on measuring communities across time with binary and continuous edges using stochastic block models.*

Matias, Catherine, and Vincent Miele. "Statistical clustering of temporal networks through a dynamic stochastic block model." arXiv preprint arXiv:1506.07464 (2015).

Week 11 (TBD 26) **Modeling temporal networks.** *How networks change over time.*

Skyler J. Cranmer, Bruce A. Desmarais, and Jason Morgan "Inferential Network Analysis" at *Cambridge University Press* 2020.

Matias, Catherine, and Vincent Miele. "Statistical clustering of temporal networks through a dynamic stochastic block model." arXiv preprint arXiv:1506.07464 (2015).

Gollini, Isabella, and Thomas Brendan Murphy. "Joint modeling of multiple network views." Journal of Computational and Graphical Statistics 25.1 (2016): 246-265.

Sewell, Daniel K., and Yuguo Chen. "Latent space models for dynamic networks." Journal of the American Statistical Association 110.512 (2015): 1646-1657.

Part 5. Network structures, hierarchy, clustering coefficients

Week 12 (TBD 26) **Network hierarchy** *How does the structure of a network contribute to the observed outcomes?*

Clauset, Aaron, Cristopher Moore, and Mark EJ Newman. "Hierarchical structure and the prediction of missing links in networks." Nature 453.7191 (2008): 98-101.

Clauset, Aaron, Samuel Arbesman, and Daniel B. Larremore. "Systematic inequality and hierarchy in faculty hiring networks." Science advances 1.1 (2015): e1400005.

Hobson, Elizabeth A., and Simon DeDeo. "Social feedback and the emergence of rank in animal society." PLoS Comput Biol 11.9 (2015): e1004411.

Week 13 (TBD 2) **Presentation of research papers**

Week 14 (TBD 9) **Presentation of research papers**

Week 15 (TBD 16) **Presentation of research papers**