# Social Network Analysis: Theory and Methods (Sociology 8412)

Zack W. Almquist
Fall Semester, 2015

## Class Schedule

Lecture: W 2:30 – 5:00 pm Social Sciences Building 1114
**URL:** `http://moodle.umn.edu`
*Note:* Requires UMN login and registration in class to access.
**URL:** `http://www.github.com`
*Note:* Requires github login.

## Professor

**Name:** Zack W. Almquist
**Office:** 960 Social Science Building
**Office Hours:** W 5:00-6:00PM
**Email:** almquist@umn.edu
**Telephone:** 612-624-4300 (not recommended)

## Course Objectives

This course is an introduction to the theory and methods employed in modern social network analysis. Specific topics covered include data-collection, node and graph-level indices, conditional uniform graph tests, network regression, network autocorrelation models, network inference, elementary exponential random graph models, and missing data models. Both descriptive and inferential approaches will be covered, with an emphasis on the use of network methods for theory testing. The approach taken to the material will be a combination of lecture, labs, homework, and reading discussions. The class will include regular homework assignments that focus on the analysis and interpretation of a variety of network data sets. By the end of the class, each student should be familiar with the most commonly employed methods in social network analysis, should have a clear understanding of the capabilities, limitations, and indications for use of these methods within typical settings, and should be competent at both applying and interpreting these methods for both

exploratory analysis and theory testing in the social network domain. Students will also be exposed to the use of the `R` statistical computing system for network analysis, Rstudio, and github development platforms. The student will also be given the opportunity to acquire competency in basic data management and analysis tasks within the `R` environment.

## Prerequisites

Graduate level probability and statistics (including standard hypothesis testing and regression methods) is assumed, as is basic literacy in elementary scientific computing (ability to manipulate data, use command-line/scripting tools, etc.). Additional mathematical and computational background is not required, but students are expected to take initiative in learning to solve unfamiliar problems. Prior knowledge of network analysis is not assumed.

## Course Requirements

### Computers

It is not required that students bring their computers/laptops to lecture and lab (if one is owned), but it is *highly* recommended since lecture will make extensive use of the computer software `R`. Computer labs are available on campus, please consult with the office if you have trouble finding the various locations of campus computer labs.

### Readings

Weekly readings assignments can be found on the course syllabus. All readings are assumed to be completed before each lecture/seminar. You are expected to read over the class notes each week and make sure you are familiar with the material as the course progresses. Questions are encouraged.

### Homework

Homework assignments will normally be administered on a bi-weekly basis and will be due on every other Wednesday. Homework assignments are meant to achieve three results: (1) provide practice with the statistical concepts discussed in class and, and (2) provide practice with the computational and statistical programing language `R` and (3) provide a chance to demonstrate your mastery of material and highlight areas where more work is needed. You may work in a group, but all write-ups must be done independently. All collaborators should be appropriately cited in your write up and any detailed R code should also be provided.

## Exams

To assess mastery of course material, one take-home exam will be administered over the course of the semester. The examination will cover all material presented in lecture and assigned readings. In order to prepare for exams, students are advised to keep up-to-date on reading assignments and to attend lectures regularly.

## Participation

Individuals are expected to attend every course, to have completed every reading, and to participate with questions and discussion on each topic as presented. If you plan on missing any class period you are responsible for all material and for contacting the instructor in a timely manner.

## Grading

| | |
|---|---|
| Participation: | 20% |
| Homework: | 60% |
| Exam/Project: | 20% |

Lectures, readings, labs, and review sessions are provided for each student's benefit. It is the responsibility of the student to take advantage of these opportunities to acquire and demonstrate mastery of course material, so as to achieve his or her desired grade.

### Letter grade assignment

| | |
|---|---|
| A | 93%+ |
| A- | 90-92.99% |
| B+ | 87-89.99% |
| B | 83-86.99% |
| B- | 80-82.99% |
| C+ | 77-79.99% |
| C | 73-76.99% |
| C- | 70-72.99% |
| D | 60-69.99% |
| F | <59.99% |

## Required Texts

- Wasserman, S. and Faust, K. (1994). Social Networks: Methods and Applications. Cambridge: Cambridge University Press.
- Almquist, Zack (2015). Lecture Notes for Social Network Analysis: Theory and Methods. University of Minnesota.

### Readings

Be prepared to discuss all readings assigned at anytime in lecture/seminar.

## Required Software

We will be using the `R` statistical programming language. `R` can be downloaded at `http://www.r-project.org/`.

RStudio IDE (Integrated Development Environment) is a software application which facilitates interaction with the `R` statistical programming language. It is often preferred to the GUI (Graphic User Interface) made available through CRAN. You can download it at `http://www.rstudio.com/`.

Latex is a word processor and a document markup language. It can be downloaded and installed on Windows (`http://miktex.org/`), OSX (`https://tug.org/mactex/`) or Linux (use the package manager of your choice).

A github account will be required of all students. One can register for a github account at `https://github.com/`. You can find information about how github works with Rstudio at `http://z.umn.edu/rstudiogit`, and github maintains a quite good help-system at `https://help.github.com/`.

## Course Policies

### Missing Class, etc.

It is expected that each member of the class will attend every lecture/discussion. If there is an appropriate reason to miss class it is expected that the individual will email or discuss in person with the instructor at least one week in advance. For any medical issues please see the UMN website for university policies.

### Cheating, etc.

All work is assumed to be your own and all individuals are expected to follow the university policy on cheating and misconduct. If you have any questions please consult the UMN website for university policies.

## Class Structure

This class meets for two and half hours one day a week for approximately 15 weeks. Each week the course will be divided into three pieces: (1) lecture, (2) lab, and (3) discussion. Lecture will focus on introducing the key methods and theory of modern social network analysis. Lab will focus on introducing the student to network analysis in `R` in combination with github and Rstudio. Discussion will focus on going over the readings assigned each week. An email will be sent around each week assigning the following roles: (1) summary, (2) question, and (3) response.

## Assignments

### Week 1 (09/09/15): Introduction, Basic Concepts and Data Collection

- *Readings:*
  - Radcliffe-Brown, A. R. (1940). On Social Structure. The Journal of the Royal Anthropological Institute of Great Britain and Ireland, 70(1): 1–12.
  - Butts, C.T. (2008). Social Network Analysis: A Methodological Introduction. Asian Journal of Social Psychology, 11: 13–41; read journal pages 13–22.
  - Butts, C. T. (2009). Revisiting the foundations of network analysis. science, 325(5939): 414.
- *Lab:*
  - Load `networkMethods` package and run the code in lab 1.
- *Homework:*
  - No homework.

### Week 2 (09/16/15): Node-level Indices as Descriptives and Covariates

- *Readings:*
  - Friedkin, N. (1991). Theoretical Foundations for Centrality Measures. American Journal of Sociology, 96(6):1478–1504.
  - Freeman, L. (1979). Centrality in social networks: Conceptual clarification. Social Networks 1:215–239.
  - Good, P. (2000). A practical Guide to Resampling Methods, 2nd edition. Springer; read chapters 1 and 2.
- *Lab:*

  - Load `networkMethods` package and run the code in lab 2.

- *Homework:*
  - Go to github and edit the homework_1.Rnw file (make sure push your edits to github regularly).

### Week 3 (09/23/15): Network Autocorrelation Models

- *Readings:*
  - Butts, C.T. (2008). Social Network Analysis: A Methodological Introduction. Asian Journal of Social Psychology, 11: 13–41; read journal pages 34-36.
  - Leenders, R.T.A.J. (2002). Modeling Social Influence Through Network Autocorrelation: Constructing the Weight Matrix. Social Networks, 24: 21–47.
- *Lab:*
  - Load `networkMethods` package and run the code in lab 3.
- *Homework:*
  - Homework 1 due!

### Week 4 (09/30/15): Reciprocity, the Dyad Census

- *Readings:*
  - Wasserman, S. and Faust, K. (1994). Social Networks: Methods and Applications. Cambridge: Cambridge University Press. p.p. 503–555.
- *Lab:*
  - Load `networkMethods` package and run the code in lab 4.
- *Homework:*
  - Go to github and edit the homework_2.Rnw file (make sure push your edits to github regularly).

### Week 5 (10/07/15): Transitivity, the Triad Census

- *Readings:*
  - Butts, C.T. (2008). Social Network Analysis: A Methodological Introduction. Asian Journal of Social Psychology, 11: 13–41; read journal pages 22–29.
  - Faust, K. (2007). Very Local Structure in Social Networks. Sociological Methodology, 37:209–256.

  - Holland, P. and Leinhardt S. (1971). Transitivity in structural models of small groups. Comparative Group Studies 2:107–124.
  - Cartwright, D. and Harary F. (1979). Balance and clusterability: An overview. In Holland, P.W. and Leinhardt, S. (eds.), Perspectives on Social Network Research, pages 25-50. New York: Academic Press.
  - Davis, J. (1967). Clustering and structural balance in graphs. Human Relations. 20:181–187.
- *Lab:*
  - Load `networkMethods` package and run the code in lab 5.
- *Homework:*
  - Homework 2 due!

### Week 6 (10/14/15): Introduction to Random Graphs

- *Readings:*
  - Wasserman, S. and Faust, K. (1994). Social Networks: Methods and Applications. Cambridge: Cambridge University Press. Chapters 13.1-13.5.
  -
- *Lab:*
  - Load `networkMethods` package and run the code in lab 6.
- *Homework:*
  - Go to github and edit the homework_3.Rnw file (make sure push your edits to github regularly).

### Week 7 (10/21/15): Dyadic Mixing, Basic Blockmodels

- *Readings:*
  - McPherson, M.; Smith-Lovin, L.; and Cook, J.M. (2001). Birds of a Feather: Homophily in Social Networks. Annual Review of Sociology, 47:415–444.
  - Morris, M. (1991). A Log-linear Modeling Framework for Selective Mixing. Mathematical Biosciences, 107: 349–377.
- *Lab:*
  - Load `networkMethods` package and run the code in lab 7.
- *Homework:*
  - Homework 3 due!

### Week 8 (10/28/15): Positions, Roles, and Generalized Blockmodels

- *Readings:*
  - Doreian, P.; Batagelj, V.; and Ferligoj, A. (2005). Positional Analyses of Sociometric Data. In P.J. Carrington, J. Scott, and S. Wasserman (Eds.), Models and Methods in Social Network Analysis, pp77–97. Cambridge: Cambridge University Press.
- *Lab:*
  - Load `networkMethods` package and run the code in lab 8.
- *Homework:*
  - Go to github and edit the homework_4.Rnw file (make sure push your edits to github regularly).

### Week 9 (11/04/15): Cohesive Subgroups

- *Readings:*
  - Borgatti, S.P. and Everett, M.G. (1999). Models of Core/Periphery Structures. Social Networks, 21: 375–395.
  - Freeman, L. (1992). The sociological concept of group: An empirical test of two models. American Journal of Sociology. 98(1):152–166.
  - Breiger, R. (1974). The duality of persons and groups. Social Forces, 53:191-190.
  - Moody, J. and White, D. R. (2003). Structural Cohesion and Embeddedness: A hierarchical conception of Social Groups. American Sociological Review, 68:103–127
- *Lab:*
  - Load `networkMethods` package and run the code in lab 9.
- *Homework:*
  - Homework 4 due!

### Week 10 (11/11/15): Graph Correlation, QAP, and Network Regression

- *Readings:*
  - Krackhardt, D. (1987). QAP Partialling as a Test of Spuriousness. Social Networks, 9(2): 171–186.

  - Dekker, D., Krackhardt, D., and Snijders, T.A.B. (2008). Sensitivity of MRQAP Tests to Collinearity and Autocorrelation Conditions. Psychometrika, 72(4): 563-581.
  - Anderson, C. J., Wasserman, S., and Crouch, B. (1999), A p* Primer: Logit Models for Social Networks, Social Networks, 21: 37–66.
- *Lab:*
  - Load `networkMethods` package and run the code in lab 10.
- *Homework:*
  - Go to github and edit the homework_5.Rnw file (make sure push your edits to github regularly).

### Week 11 (11/18/15): Modeling Cohesive Groups in Social Space

- *Readings:*
  - Nowicki, K. and Snijders, T. (2001). Estimation and prediction for stochastic block models. Journal of the American Statistical Association, 96: 1077–1087.
- *Lab:*
  - Load `networkMethods` package and run the code in lab 11.
- *Homework:*
  - Homework 5 due!

### Week 12 (11/25/15): Multivariate Analysis of Graph Sets

[*Thanksgiving*]

- *Readings:*
  - Butts, C.T. (2007). Permutation Models for Relational Data. Sociological Methodology, 37: 257–281.
  - Butts, C.T. (2008). Social Network Analysis: A Methodological Introduction. Asian Journal of Social Psychology, 11: 13–41; read journal pages 31–34.
- *Lab:*
  - Load `networkMethods` package and run the code in lab 12.
- *Homework:*
  - Go to github and edit the homework_6.Rnw file (make sure push your edits to github regularly).

### Week 13 (12/02/15): Exponential Random Graph Models (ERGMs)

- *Readings:*
  - Hunter, David R.; Handcock, Mark S.;, Butts, Carter T.; Goodreau, Steven M.; and Morris, Martina. (2008). ergm: A Package to Fit, Simulate and Diagnose Exponential-Family Models for Networks. Journal of Statistical Software, 24(3).
  - Wasserman, S. and Robins, G. (2005). An Introduction to Random Graphs, Dependence Graphs, and p?. In P.J. Carrington, J. Scott, and S. Wasserman (Eds.), Models and Methods in Social Network Analysis, pp148–160. Cambridge: Cambridge University Press.
  - Morris, M.; Handcock, M.S.; and Hunter, D.R. (2008). Specification of Exponential-Family Random Graph Models: Terms and Computational Aspects. Journal of Statistical Software, 24(4).
- *Lab:*
  - Load `networkMethods` package and run the code in lab 13.
- *Homework:*
  - Homework 6 due!

### Week 14 (12/09/15): Adequacy and Degeneracy in ERG Models

- *Readings:*
  - Goodreau, S.M.; Handcock, M.S.; Hunter, D.R.; Butts, C.T.; and Morris, M. (2008). A statnet Tutorial. Journal of Statistical Software, 24(9).
  - Handcock, M.S. (2003). Statistical Models for Social Networks: Inference and Degeneracy. In R. Breiger, K.M. Carley, and P. Pattison (Eds.), Dynamic Social Network Modeling and Analysis, pp 229–240. Washington, D.C.: National Academies Press.
- *Lab:*
  - Load `networkMethods` package and run the code in lab 14.
- *Homework:*
  - No homework.

**Week 15 (12/16/15): Inference for partially observed networks**

- *Readings:*
  - Gile, K. J. and Handcock, M. S. (2010). Modeling Networks from Sampled Data. Annals of Applied Statistics 4: 5–25
  - Gile, K. J., and Handcock, M. S. (2015). Network model-assisted inference from respondent-driven sampling data. Journal of the Royal Statistical Society: Series A (Statistics in Society), 178(3), 619–639.
- *Lab:*
  - Load `networkMethods` package and run the code in lab 15.
- *Homework:*
  - Take home exam handed out.