

# Social Network Analysis

SUMMER 2014

Kinga Makovi  
Knox Hall, Room 512  
km2730@columbia.edu

## COURSE DESCRIPTION

*Social Network Analysis* is intended to introduce social networks as conceptual tools that help theory building of social structure and social action; and as a developing set of methods or analytical approaches to data that could be thought of in network terms. The course bridges long-standing traditions of network thinking in sociology with new ways to visualize and analyze network data. It will not stick to any substantive area of inquiry or to any specific methodology, rather, it will give a broad map of substantive questions and approaches to answer them. As a consequence, actors and relations will be defined in many different ways throughout the course depending on the puzzles at hand, for example, they could be families tied to one another through marriage, firms engaging in exchange, or organizations linked through career mobility of individuals. The key emphasis is going to be on the structure of relationality which will be understood through middle-range theories of tie formation and dissolution. Systematic thinking about networks (i.e., modeling) has the potential to link micro-level processes with macro-level outcomes – the aim is to start to see the problem of social structure and social action through this lens. The theoretical concepts will be implemented using the statistical software **R**. No previous programming experience is required: familiarity with **R** and the relevant packages will be built up through the course; however, an open mind, patience and willingness to make some intellectual effort will be key for this endeavour to be successful. The skills you will acquire familiarizing yourself with **R** will not be specific to network analysis, and will help you in statistics classes in the future.

We start with the theoretical foundations of structural analysis, including the parsimonious ways of describing social structures (i.e., density, degree distribution and centrality) and the principles to account for change in structures over time (i.e., balance, homophily and transitivity). We build on these basics to then turn to modularity: the “groupiness” of structures and talk about cohesion and the relationship between categories and networks. We will introduce more nuanced concepts, such as structural equivalence, brokerage, structural folds and multiplexity of ties. Finally, we will delve into methods to compare network structures (QAP regression), to model network structures through local processes ( $p^*$  models and ERGMs) and to model structures over time (TERGMs and agent based simulations).

## WHAT TO EXPECT IN TERMS OF CLASS ORGANIZATION?

Sessions will begin with a short lecture to introduce concepts and notations, but the bulk of the time will be spent on the discussion of readings which requires participation and will be prompted by memos of the students. The course’s main goal is to lay down the theoretical foundations of network analysis while introducing the simple ways to visualize and analyze networks using **R**.

Hence, the sessions will be complemented with labs as we go on.

#### LEARNING OBJECTIVES

At the end of this class you should be able to:

- Understand the basic concepts associated with social network analysis as a conceptual tool and a methodology.
- Think creatively and constructively about network structures of varied kind and form.
- Describe parsimoniously and analyze networks in order to reveal their key structural features using **R**.
- Collect network data individually to ask and answer questions of varied focus of interest.

#### REQUIREMENTS AND EVALUATION CRITERIA

There are a number of requirements for this class. The first is that students do the readings and write a brief (no more than 2 single-spaced pages) memo on them for each class. Memos are due the day before class at noon, and you are required to read other's memos to facilitate discussion. The memos are designed to provide you with an opportunity to summarize, evaluate, and extend the core idea of a reading or readings, the emphasis being on evaluation and extension. There is no need to incorporate explicitly all the readings for a class, unless the idea of the memo is to compare the readings from a specific viewpoint. Each memo is worth 2 points (24 points in total).

The third requirement is to complete four lab exercises that introduce the students to different ways to visualize and analyze social network data – each lab exercise is worth 6 points (24 points in total). Completing these assignments will enable you to learn basic network analysis skills which will be used to collect and analyze your own network data.

This project could be accomplished individually or in pairs and will be presented at the end of the course accompanied with a short lab report (5-8 single-spaced pages). The presentation is worth 10 points, the lab report 30 (40 points in total).

To aid the preparation of this assignment, projects must have the instructor's approval which will be given in two stages. First, students need to submit a brief summary on the type of data they seek to collect identifying actors and relations of their network(s) and the source(s) they intend to use – due by end of the second week. Second, students need to submit a brief progress report on how the data collection goes and a summary of the concepts and methods they intend to apply in their analysis – due by the end of the fourth week. These assignments will not be graded, but without their approval students will not be able to present at the end of the course.

Last but not least, students must participate in class discussion. Speaking less but saying more is preferable to speaking more but saying less. Contributions to discussion is worth 12 points – the assignment of which will in itself will be a collective endeavour.

This: 24+24+10+30+12 should amount to 100 points.

#### STATEMENT ON ACADEMIC INTEGRITY

The intellectual venture in which we are all engaged requires of faculty and students alike the highest level of personal and academic integrity. As members of an academic community, each one of us bears the responsibility to participate in scholarly discourse and research in a manner characterized by intellectual honesty and scholarly integrity.

Scholarship, by its very nature, is an iterative process, with ideas and insights building one upon the other. Collaborative scholarship requires the study of other scholars' work, the free discussion of such work, and the explicit acknowledgement of those ideas in any work that inform our own. This exchange of ideas relies upon a mutual trust that sources, opinions, facts, and insights will be properly noted and carefully credited.

In practical terms, this means that, as students, you must be responsible for the full citations of others' ideas in all of your research papers and projects; you must be scrupulously honest when taking your examinations; you must always submit your own work and not that of another student, scholar, or internet agent.

Any breach of this intellectual responsibility is a breach of faith with the rest of our academic community. It undermines our shared intellectual culture, and it cannot be tolerated. Students failing to meet these responsibilities should anticipate being asked to leave Columbia.

#### FOR STUDENTS WITH DISABILITIES

In order to receive disability-related academic accommodations, students must first be registered with Disability Services (DS). More information on the DS registration process is available online at [www.health.columbia.edu/ods](http://www.health.columbia.edu/ods). Faculty must be notified of registered students' accommodations before exam or other accommodations will be provided. Students who have, or think they may have, a disability are invited to contact Disability Services for a confidential discussion at (212) 854-2388 (Voice/TTY) or by email at [km2730@columbia.edu](mailto:km2730@columbia.edu).

## BY-WEEK MENU

**Class 1: Introduction – Basic concepts**

The first class will be lecture heavy to introduce some basic concepts and notations used throughout the class. The lab associated with this session will familiarize you with **R** and some basic functions of the network packages (**sna**, **network** and **igraph**) which you will continue to use throughout the course.

*Assigned Readings*

Wasserman, S. and K. Faust. 1994. Relations and Networks in the Social and Behavioral Sciences. Chapter 1 in *Social Network Analysis: Methods and Applications*, New York:Cambridge University Press, 3-27.

Borgatti, S.P., A. Mehra, D.J. Brass, and G. Labianca. 2009. Network Analysis in the Social Sciences. *Science*, 323(5916):892-895.

Butts, C.T. 2009. Revisiting the Foundations of Network Analysis. *Science*, 325(5939):414-416.

Gould, R.V. Uses of Network Tools in Comparative Historical Research. 2003. Chapter 7 in J. Mahoney and D. Rueshemeyer (eds.), *Comparative Historical Analysis in the Social Sciences*, New York:Cambridge University Press, 241-269.

Moody, J. 2006. Fighting a Hydra: A Note on the Network Embeddedness of the War on Terror. *Structure and Dynamics: eJournal of Anthropological and Related Sciences*, 1(2).

*Further Readings*

Tilly, C. 1998. From transactions to structures. Chapter 2 in *Durable Inequality*, Berkeley:University of California Press, 41-73.

Erikson, E., 2013. Formalist and Relationalist Theory in Social Network Analysis. *Sociological Theory*, 31(3):219–242.

**Class 2: Collecting network data**

The second class should be eye-opening and show you a variety of different ways to collect network data, and reveal the fact that data on networks is all around us and comes in a number of different forms. The lab associated with this session will concentrate on how to digitize and make usable the data you will collect. This is an important class to have at the beginning of the course, given that you will be working on the data you collect yourself.

*Assigned Readings*

Wasserman, S. and K. Faust. 1994. Social Network Data. Chapter 2 in *Social Network Analysis: Methods and Applications*, New York:Cambridge University Press, 28-66.

Kadushin, C . 2005. Who Benefits from Network Analysis: Ethics of Social Network Research. *Social Networks*, 27(2):139-153.

Eagle, N., A. Pentland, and D. Lazer. 2009. Inferring Social Network Structure using Mobile Phone Data. *Proceedings of the National Academy of Sciences (PNAS)*, 106(36):15274-15278.

Bearman, P., and P. Parigi. 2004. Cloning Headless Frogs and Other Important Matters: Conversation Topics and Network Structure. *Social Forces*, 83 (2): 535-557.

Vehovar, V., K.L. Manfreda, G. Koren, and V. Hlebec 2008. Measuring Ego-Centered Social Networks on the Web: Questionnaire Design Issues. *Social Networks*, 30(3):213-222.

#### *Further Readings*

Marsden, P.V. 1990. Network Data and Measurement. *Annual Review of Sociology*, 16:435-463.

Emirbayer, M. 1997. Manifesto for Relational Sociology. *American Journal of Sociology*, 103(2):281-317.

Breiger, R.L. 2005. Introduction to Special Issue: Ethical Dilemmas in Social Network Research. *Social Networks*, 27(2):89-93.

Kadushin, C. 2010. Introduction. Chapter 1 in *Making Connections: An introduction to social network theory, concepts and findings*, Oxford University Press:New York, 3-12.

### **Class 3: Describing and analyzing ego-networks I. – Centrality and power**

The third class will introduce centrality measures and explain the conceptual thinking behind them. The lab associated with this class will show how to calculate these measures. The first assignment is due on network visualization.

#### *Assigned Readings*

Bonacich, P. 1987. Power and Centrality: A family of Measures. *American Journal of Sociology*, 92(5):1170-1182.

Freeman, L.C. 1979. Centrality in Social Networks Conceptual Clarification. *Social Networks*, 1(3):215-239.

Rodan, S. 2011. Choosing the  $\beta$  Parameter When Using the Bonacich Power Measure. *Journal of Social Structure*, 11(4).

Martin, J.L. 2005. Is Power Sexy? *American Journal of Sociology*, 111(2):408-446.

#### *Further Readings*

Wasserman, S. and K. Faust. 1994. Centrality, Prestige, and Related Actor and Group Measures. Chapter 5 in *Social Network Analysis: Methods and Applications*, New York:Cambridge University Press, 169-248.

Marsden, P.V. 1983. Restricted Access in Networks and Models of Power. *American Journal of Sociology*, 88(4):686-717.

Bell, D.C., J.S. Atkinson, and J.W. Carlson. 1999. Centrality Measures for Disease Transmission Networks. *Social Networks* 21(1):1-21.

Bonacich, P., and P. Lloyd. 2001. Eigenvector-Like Measures of Centrality for Asymmetric Relations. *Social Networks*, 23(3):191-201.

Bothner, M.S., E.B. Smith, and H.C. White. 2010. A Model of Robust Positions in Social Networks. *American Journal of Sociology*, 116(3):943-992.

#### **Class 4: Describing and analyzing ego-networks II. – Structural holes, structural folds and tie strength**

The fourth class will recap the simple ways of describing ego-networks and bring in a couple of more nuanced concepts to have a handle on local structure. Instead of a lab, ideas for student projects will be discussed. The plan for data collection is due.

##### *Assigned Readings*

Burt, R.S. 2004. Structural Holes and Good Ideas. *American Journal of Sociology*, 110(2):349-399.

Granovetter, M. 1973. The Strength of Weak Ties. *American Journal of Sociology*, 78(6):1360-1380.

Vedres, B., and D. Stark. 2010. Structural Folds: Generative Disruption in Overlapping Groups. *American Journal of Sociology*, 115(4):1150-1190.

Backstrom, L. and J. Kleinberg. 2013. Romantic Partnerships and the Dispersion of Social Ties: A Network Analysis of Relationship Status on Facebook. *arXiv*:1310.6753.

##### *Further Readings*

Burt, R.S. 2005. Brokerage and Closure. Oxford University Press:New York.

Kalish, Y., and G. Robins. 2006. Psychological Predispositions and Network Structure: The Relationship Between Individual Predispositions, Structural Holes and Network Closure. *Social Networks*, 28(1):56-84.

Stark, D., and B. Vedres. 2006. Social Times of Network Spaces: Network Sequences and Foreign Investment in Hungary. *American Journal of Sociology*, 111(5):1367-1411.

#### **Class 5: Describing and analyzing complete networks I. – Small worlds and cohesion**

The fifth class will revolve around the concept of small worlds. The lab will discuss degree distributions, and different ways of generating networks with certain pre-given properties.

##### *Assigned Readings*

Travers, J., and S. Milgram. 1969. An Experimental Study of the Small World Problem. *Sociometry*, 32(4):425-443.

Watts, D.J. 1999. Networks, Dynamics, and the Small-World Phenomenon. *American Journal of Sociology*, 105(2): 493-527.

Uzzi, B., and J. Spiro. 2005. Collaboration and Creativity: The Small World Problem. *American Journal of Sociology*, 111(2): 447-504.

Eom, Y.H. and H.H. Jo. 2014. Generalized friendship paradox in complex networks: The case of scientific collaboration. *Scientific Reports*, 4, article number: 4603.

### *Further Readings*

Gould, R.V. 1993. Collective Action and Network Structure. *American Sociological Review*, 58(2):182-196.

Erikson, E., and P. Bearman. 2006. Malfeasance and the Foundations for Global Trade: The Structure of English trade in the East Indies, 1601-1833. *American Journal of Sociology*, 112(1):195-230.

Kim, H., and P. Bearman. 1997. The Structure and Dynamics of Movement Participation. *American Sociological Review*, 62(1):70-93.

## **Class 6: Describing and analyzing complete networks II. – Modularity and structural equivalence**

The sixth class will discuss the concept of modularity, structural equivalence and the ways these concepts might relate to social action. The lab will show a number of ways to identify cohesive groups in networks and actors that are structurally equivalent to one another and introduce blockmodeling. The second assignment is due on centrality measures.

### *Assigned Readings*

Breiger, R.L. 1974. The Duality of Persons and Groups. *Social Forces*, 53(2):181-190.

White, H.C., S.A. Boorman, and R.L. Breiger. 1976. Social Structure from Multiple Networks. I. Blockmodels of Roles and Positions. *American Journal of Sociology*, 81(4):730-780.

Gould, R.V. 1996. Patron-Client Ties, State Centralization, and the Whiskey Rebellion. *American Journal of Sociology*, 102(2):400-429.

### *Further Readings*

Bearman, P. 1997. Generalized Exchange. *American Journal of Sociology*, 102(5):1383-1415.

Newman, M.E.J. 2006. Modularity and Community Structure in Networks. *Proceedings of the National Academy of Sciences (PNAS)*, 103(23):8577-8582.

## **Class 7: Multiplexity**

The seventh class will be a recap of the first half of the course, opening up the territory for multiple networks, i.e., a set of networks defined on the same actors where relations have different content. The lab will introduce QAP regression useful for comparing networks.

### *Assigned Readings*

Padgett, J.F. and C.K. Ansell. 1993. Robust Action and the Rise of the Medici, 1400-1434. *American Journal of Sociology*, 98(6):1259-1319.

Kilduff, M. and D. Krackhardt. 1994. Bringing the Individual Back in: A Structural Analysis of the Internal Market for Reputation in Organizations. *Academy of Management Journal*, 37(1):87-108.

## **Class 8: Local processes – Homophily, transitivity, triadic closure, balance and hierarchy formation**

The eighth class walks you through the logic of a number of different local processes or mechanisms for tie formation and dissolution, therefore puts together the building blocks of dynamic models. There is no lab for this class. The plan for data analysis is due.

#### *Assigned Readings*

Simmel, G. 1950. The Triad. Part II, Chapter 4 in the *Sociology of Georg Simmel*. Glencoe, Illinois: The Free Press, 118-169.

Chase, Ivan. 1982. Dynamics of Hierarchy Formation: The Sequential Development of Dominance Relationships. *Behaviour*, 80(3/4):218-240.

McPherson, M., L. Smith-Lovin, and J.M. Cook. 2001. Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27:415-444.

#### *Further Readings*

Gould, R.V. 2002. The Origins of Status Hierarchies: A Formal Theory and Empirical Test, *American Journal of Sociology* 107: 1143-1178.

Chalizi, C.R., and A.C. Thomas. 2011. Homophily and Contagion Are Generically Confounded in Observational Social Network Studies. *Sociological Methods & Research*, 40(2):211-239.

### **Class 9: Peer influence and diffusion**

The ninth class will show how difficult it is empirically to parse out processes of peer influence from clustering of behavior/outcomes that aren't the consequence of diffusion (a topic which is close to issues of selection and causality more generally). The lab will allow you to ask questions related to your own data and analysis strategies for the final project. The third assignment is due on community detection in networks.

#### *Assigned Readings*

Christakis, N., and J.H. Fowler. 2007. The Spread of Obesity in a Large Social Network over 32 Years. *New England Journal of Medicine*, 357:370-379.

Centola, D., and M. Macy. 2007. Complex Contagions and the Weakness of Long Ties. *American Journal of Sociology*, 113(3):702-734.

Wang, D.J., and S.A. Soule. 2012. Social Movement Organizational Collaboration: Networks of Learning and the Diffusion of Protest Tactics, 1960-1995. *American Journal of Sociology*, 117(6):1674-1722.

#### *Further Readings*

Friedkin, N.E., and K.S. Cook. 1990. Peer Group Influence. *Sociological Methods & Research*. 19(1):122-143.

Aral, S., L. Muchnik, and A. Sundararajan 2009. Distinguishing Influence-Based Contagion from Homophily-Driven Diffusion in Dynamic Networks. *Proceedings of the National Academy of Sciences (PNAS)*, 106(51):21544-21549.



Liu, K., M. King, and P. Bearman. 2010. Social Influence and the Autism Epidemic. *American Journal of Sociology*, 115(5):1387-1434.

Liu, K. and P. Bearman. 2012. Focal Points, Endogenous Processes and Exogenous Shocks in the Autism Epidemic. *Sociological Methods and Research*

Centola, D. 2010. The Spread of Behavior in an Online Social Network Experiment. *Science*, 329(5996):1194-1197.

Fowler, James H. and Nicholas A. Christakis. 2010. Cooperative Behavior Cascades in Human Social Networks. *Proceedings of the National Academy of Sciences (PNAS)*, 107(12):5334-5338.

### **Class 10: Statistical modeling I. – $p^*$ models and ERGMs**

The tenth class introduces  $p^*$  models, and most importantly, it will highlight the kinds of questions these models are useful for. The lab will give an introduction to show how these models work in practice.

#### *Assigned Readings*

Snijders, T.A.B., P.E. Pattison, G.L. Robins, and M.S. Handcock. 2006. New Specifications for Exponential Random Graph Models. *Sociological Methodology*, 36(1):99-153.

Robins, G., P. Pattison, Y. Kalish, and D. Lusher. 2007. An Introduction to Exponential Random Graph ( $p^*$ ) Models for Social Networks. *Social Networks* 29(2):173-191.

Goodreau, S.M., J.A. Kitts, and M. 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.

#### *Further Readings*

Lusher, D., J. Koskinen, and G. Robins. 2013. Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications (Structural Analysis in the Social Sciences). Cambridge University Press:New York.

### **Class 11: Statistical modeling II. – Networks over time**

The eleventh class should bring together the second half of the course and show innovative ways for thinking about networks over time. These issues are on the boundary of social networks research, thus, the aim of the class is not to delve deep into the state of the art methods, rather, comprehend the directions toward which they are moving. The lab will not discuss new material, but serves you with the opportunity to ask clarification questions and seek help with finishing up you own project.

#### *Assigned Readings*

Bearman, P., J. Moody and K. Stovel. 2004. Chains of Affection: The Structure of Adolescent Romantic and Sexual Networks. *American Journal of Sociology*, 110(1):44-91.

Snijders, T.A.B., G.G. vande Bunt, and C.E.G. Steglichc. 2010. Introduction to Stochastic Actor-Based Models for Network Dynamics. *Social Networks*, 32:44-60.

Snijders, T.A.B. 2013. Network Dynamics. Chapter 7 in R. Wittek, T.A.B. Snijders, and V. Nee (eds.), *The Handbook of Rational Choice Social Research*. Stanford:Stanford University Press, 252-279.

## **Class 12: Left out from the canon – esoteric approaches to networks**

The final class will leave students with a discussion on where social network analysis is going, and will provide insights into some esoteric applications as well. The fourth assignment is due on ERGMs.

### *Assigned Readings*

Bearman, P., J. Moody, and R. Faris. 2002. Networks and History. *Complexity*, 8(1): 61-71.

Bearman, P. and K. Stovel. 2000. Becoming a Nazi: A Model for Narrative Networks. *Poetics*, 27(2):69-90.

### *Further Readings*

Bearman, P. 1991. Desertion as Localism: Army Unit Solidarity and Group Norms in the U.S. Civil War. *Social Forces*, 70(2):321-342.

Emirbayer M., and J. Goodwin. 1994. Network Analysis, Culture and the Problem of Agency. *American Journal of Sociology*, 99(6):1411-1454.

Bearman, P., R. Faris, and J. Moody. 1999. Blocking the Future: New Solutions for Old Problems in Historical Social Science. *Social Science History*, 23(4):501-533.

Butts, C.T. 2008. A Relational Event Framework for Social Action. *Sociological Methodology*, 38(1):155-200.

## **After-Class**

Some of the readings assigned for this course date back to the 70s, others just came out. This highlights the feature emphasized through this course: network thinking is inherent to the social sciences, but doing it right is hard, this is why we see a proliferation of new methods and approaches to network data. If you would like to go back to the classics, read Simmel Georg, Harrison White, and Roger Gould (and read a lot of history ☺). Read also Karen Barkey, Peter Bearman, David Stark, Balazs Vedres, and Karoly Takacs to strengthen your theoretical thinking. If you are interested in methodological advances, keep an eye on Carter Butts, Jim Moody, Martina Morris, Mark Newman, Garry Robins, Tom Snijders and their students.