

Sociology 186: Advanced Social Network Analysis

Instructor: Jonathan Haynes
Email: jhaynes@stanford.edu
Office Hours: T/Th 2:05-3pm (120-210C)

Time: T/Th 12:50-2:05pm
Location: 200-305
Quarter: Fall 2009/10

Course Description

This course addresses how social scientists think about networks. We will cover a little bit of information across many disciplines - Soc, CS, Stats, MS&E, ICME - and the goal is to connect ideas together. Specifically, this course provides practical experience and an interdisciplinary perspective on the collection, management, exploration, and analysis of social network data. The emphasis is on developing technical skills for studying large-scale social networks. Topics include theories of social order, small worlds, scientific computing, network sampling, and network dynamics.

The course webpage is available at <http://www.stanford.edu/~jhaynes/soc186/>

Course Goals

The primary goal is to encourage critical thinking about the role of social networks in our society. This will involve connecting theory from Sociology with other disciplines. A second goal is to learn technical skills for actually studying large social network datasets. This will include a gentle introduction to the Python programming language¹ and learning NetworkX². NX is an open-source Python package for network analysis developed by mathematicians at Los Alamos National Laboratories. A third goal will be learning some basic data management skills and covering major trends in the study of social networks.

The Good

This course covers the cutting edge of one of the most exciting frontiers of social science. This is also a really good course for people who like learning by doing. The course material and assignments focus on conceptual understandings and practical knowledge.

The Bad

This course will not attempt to provide an in-depth literature survey of the field or an abstract discussion of networks. There is simply too much material to cover in just one course. We will jump around many topics where a common theme – social networks – is what holds everything together.

The Ugly

Inter-disciplinary fields by their very nature are less well defined than *intra*-disciplinary fields, so we will need to connect together what may initially seem to be disparate topics.

¹ <http://www.python.org/>

² <http://networkx.lanl.gov/>

At the end of the quarter, you will be knowledgeable about:

- 1) How social scientists think about networks,
- 2) The tools and techniques for working with large network data sets, and
- 3) How to incorporate social network measures into social science research.

Course Prerequisites

There are no prerequisites. While a background in any related discipline would be helpful, a sincere interest in a conceptual understanding of the material is all you need.

Course Section

Depending on enrollment, there may be an optional weekly discussion section.

Course Texts

REQUIRED:

- (1) "Theories of Social Order: A Reader (Paperback)", 2nd Edition, eds. Hechter and Horne
- (2) "Six Degrees: The Science of a Connected Age (Paperback)", Duncan Watts
- (3) "Regional Advantage: Culture and Competition in Silicon Valley and Route 128 (Paperback)", Annalee Saxenian
- (4) "Learning Python, 3rd Edition (Paperback)", Mark Lutz

All required articles are available on the web. See citations below.

RECOMMENDED:

- (5) "Network Analysis: Methodological Foundations (Paperback)", eds: Ulrik Brandes, Thomas Erlebach

Note: Previews for all the texts can be viewed at <http://books.google.com/>

Course Requirements

This course provides hands on experience working with large, real-world social network data. We will work through examples in class/section and problem sets will give students an opportunity to apply these techniques.

Short Tutorial/Setup (5%)	A short tutorial for getting started with Python/NX.
Problem Sets (30%)	Three problem sets, 10% each.
Midterm (25%)	One in-class midterm exam; M/C, T/F, short answer, essays
Project (40%)	Project proposal (5%), Paper (30%) and Presentation (5%).

Software

You may use any computer platform you wish, as the software we will use (Python, NetworkX, Graphviz) are all open-source. Personally, I recommend the Ubuntu Linux distribution.

Course Schedule

Note: (R) = required reading, (L) = optional reading, will be covered in lecture

Week 1: Introduction

9/22 (Tues) Course Overview

(R) Hedstrom, Peter. "Dissecting the Social". (Reader pp 12-16)

(R) Brandes, Ulrik and Thomas Erlebach. "Introduction". (Network Analysis, pp 1-6)

9/24 (Thurs) Why study social networks?

(R) Watts, Duncan. Chapter 1. (Six Degrees, pp13-42)

(R) Brandes, Ulrik and Thomas Erlebach. "Fundamentals". (Network Analysis, pp 7-15)

Week 2: Getting Started

9/29 (Tues) Small Worlds

(R) Watts, Duncan. Chapters 2-3. (Six Degrees, pp 43-100)

(R) Travers, Jeffrey and Stanley Milgram. 1969. "An Experimental Study of the Small World Problem." *Sociometry*, Vol. 32, No. 4, pp 425-443.

(L) Kleinberg, Jon. "The small-world phenomenon: An algorithmic perspective." *Proc. 32nd ACM Symposium on Theory of Computing*, 2000.

[Hand out Tutorial/Setup assignment and project memo 9/29/09]

10/1 (Thurs) Introduction to Python; Search in Networks

(R) Watts, Duncan. Chapter 5. (Six Degrees, pp130-161)

(R) Lutz, Mark. Chapters 1 (Introduction), 4 (Python Object Types). (Python, pp 3-21, 65-92)

(R) Hagberg, Aric, Dan Schult, and Pieter Swart. "Tutorial". *NX Documentation*. (pp 5-15)

(L) Koschützki, D. *et al.* "Centrality Indices". (Network Analysis, pp 16-61)

[Section 1: Start IMDB project: basic python, basic network measures.]

Week 3: Microstructure

[Tutorial assignment DUE MONDAY 5PM October 5, 2009]

10/6 (Tues) Strength of Weak Ties

(R) Granovetter, Mark. "The Strength of Weak Ties", *American Journal of Sociology*, 78:6, 1973: 1360-1380. (in Reader, pp 299-309)

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

(L) Hong, Harrison, Jeffrey Kubik, and Jeremy Stein. "Thy Neighbor's Portfolio: Word-of-Mouth Effects in the Holdings and Trades of Money Managers." *Journal of Finance*. December 2005. pp 2801-2824.

[1-page project proposal DUE THURSDAY 12:50PM October 8, 2009]

10/8 (Thurs) Peer Influence

- (R) Moody, James. "Peer Influence Groups: Identifying Dense Clusters in Large Networks". *Social Networks*. (2001) 23:261-283.
 - (R) Lutz, Mark. Chapter 8 (Lists and Dictionaries). (Python, pp 152-171)
 - (L) Hill, Shawndra, Foster Provost, and Chris Volinsky. "Network-Based Marketing: Identifying Likely Adopters via Consumer Networks". *Statistical Science*. (2006) 22:2, pp 256–276.
 - (L) Thompson, Clive. "If You Liked This, Sure to Love That". *The New York Times*, November 21, 2008.
- [Section 2: Continue IMDB project, more network measures, methods in NX reference.
PROBLEM SET #1 OUT]

Week 4: Macrostructure

10/13 (Tues) Communication networks: email, mobile

- (R) Onnela, J.-P., J. Saramaki, et. al. "Structure and tie strengths in mobile communication networks." *Applied Physical Sciences*. 10.1073, April 2007.
- (R) Diesner, Jana, Terrill L. Frantz, and Kathleen M. Carley. "Communication Networks from the Enron Email Corpus 'It's Always about the People. Enron is no Different'". *Computational & Mathematical Organization Theory*. Vol. 11:3, October 2005.

10/15 (Thurs) Communication networks: instant messenger

- (R) Leskovec, Jure and Eric Horvitz. "Worldwide Buzz: Planetary-Scale Views on an Instant-Messaging Network", Microsoft Research, Technical Pub. MSR-TR-2006-186, June 2007.
 - (R) Lutz, Mark. Chapters 12 (if Tests), 13 (while and for Loops). (Python, pp 236-277)
- [Section 3: Start SiVNAP project, ERDs; Problem Set #2 OUT]

[Problem Set #1 DUE FRIDAY 5PM, OCTOBER 16 2009]

Week 5: Culture, Community, and Social Order

10/20 (Tues) Culture

- (R) Saxenian, Annalee. Chapters 1-2. (Regional Advantage, pp 1-58)
- (R) Carley, Kathleen. "A Theory of Group Stability". *American Sociological Review*. 1991. 56:3, 331-354.

10/22 (Thurs) Spontaneous Order

- (R) Saxenian, Annalee. Chapters 3-4. (Regional Advantage, pp 59-104)
 - (R) Hayek, Friedrich. "Cosmos and Taxis". (Reader, pp 221-236)
- [Section 4: Continue SiVNAP project]

Week 6: Review/Midterm Week

[Problem Set #2 DUE MONDAY 5PM, OCTOBER 26 2009]

10/27 (Tues) Path Dependence, Unintended Consequences

(R) Saxenian, Annalee. Chapters 5-7. (Regional Advantage, pp 105-168)

[IN CLASS REVIEW SESSION]

10/29 (Thurs) **MIDTERM EXAM**

[No Section]

Week 7: Network Sampling

11/3 (Tues) Sampling

(R) Granovetter, Mark. "Network Sampling: Some First Steps", *American Journal of Sociology*, 81:6. (May 1976), pp 1287-1303.

(R) Salganik, Matthew and Douglas Heckathorn. "Sampling and Estimation in Hidden Populations Using Respondent-Driven Sampling". *Sociological Methodology*. 34:1. (Nov 2004), pp 193-240.

11/5 (Thurs) Measure Robustness

(R) Borgatti, Steve, Kathleen Carley, and David Krackhardt. "Robustness of Centrality Measures under Conditions of Imperfect Data". *Social Networks*. (2006) 28:124-136.
[Section 5: Start THOMAS project; Problem Set #3 OUT]

Week 8: Network Dynamics; Scale-Free Networks

11/10 (Tues) Norms and Affiliations

(R) Horne, Christine. "Explaining the Emergence of Norms". (Reader, pp 129-139)

(R) Simmel, Georg. "The Web of Group Affiliations". (Reader, pp 291-298)

11/12 (Thurs) Network Visualizations; Scale-Free Networks

(R) North, Stephen. "Drawing graphs with NEATO". 2004.
<http://www.graphviz.org/pdf/neatoguide.pdf>

(R) Watts, Duncan. Chapter 4. (Six Degrees, pp 101-129)

(R) Barabási, Albert-László "Scale-Free Networks". *Scientific American*, 288:60-69, May 2003.

[Section 6: Continue THOMAS project; visualizations]

Week 9: Applications

[Problem Set #3 DUE MONDAY 5PM, NOVEMBER 16 2009]

11/17 (Tues) Finding central actors

(R) Borgatti, Steve. "Identifying sets of key players in a network". *Computational, Mathematical and Organizational Theory*. 12:1. 2006. pp 21-34.

11/19 (Thurs) Special Topics

(R) Haynes, Jonathan and Igor Perisic. "Mapping Search Relevance to Social Networks." *Proceedings of the 3rd Annual SNA-KDD Workshop '09 (SNA-KDD '09)*. June 28, 2009. Paris, France. ACM 978-1-59593-848-0.

[Section 7: Work on individual student papers and presentations]

Week 10: Conclusion

[STUDENT PRESENTATIONS TUES/THURS IN CLASS, DEC 1 AND DEC 3, 2009]

12/1 (Tues) Student presentations of projects

12/3 (Thurs) More presentations; Conclusion

[Final Paper DUE MONDAY 5PM, DECEMBER 7 2009]

End of Quarter

The Fine Print – Course Policies, Honor Code

Independent work: All graded assignments for this course must be your own work, with assistance only from the course staff. Please use your common sense and remember that the primary goal is to learn the course material. At all times, you must abide by the Stanford Honor Code (<http://honorcode.stanford.edu>).

Participation: Attendance and participation are neither required nor graded; however, you should plan on attending regularly. It may be difficult to do well in the course if you miss a lot of class.

Auditing: In order to provide the best experience for enrolled students, attendance by auditors and outside visitors must be approved in advanced by the instructor. In general, the answer is "no". If the course topic interests you, you should enroll.

Credit/No Credit: Students who opt for pass/fail must complete all course requirements to pass the course.

Late Policy: A problem set will be marked down five percentage points for each 24 hour period (part or full) that it is late. Extensions will only be considered by the instructor when requested at least 72 hours before the due date and only if deemed absolutely necessary. Weekends and University holidays do not count as late days. Finally, the midterm, paper, and presentation dates are fixed.