STAT 992 (Fall 2015)

Applied Network Analysis

Prerequisite:	programming in R, linear algebra, probability	**Credits:**	1-3
Classes:	TTh 9:30-10:45	**Room:**	INGRAHAM 222
Professor:	Karl Rohe	**Office:**	MSC 1239
Phone:	263-8531	**E-mail:**	karlrohe@stat.wisc.edu
Office Hours:	Mon 2:30 - 3:30 or by appointment		

web: `http://pages.stat.wisc.edu/~karlrohe/netsci/main.html`
`goo.gl/8VGkll`

Description: We are connected by friendships. Webpages are connected by hyperlinks. Proteins are connected by chemical interactions. The language of graphs and networks allows statisticians to find common ground among these three domains; we seek to understand not merely the elements (the person, the webpage, the protein), but the broad structure of their relationships and how this structure relates to the processes that happen "on the network".

This course will provide an introduction to the increasingly popular field of network data analysis; here, the network is provided as data. Special attention will be paid to spectral clustering and network driven sampling. Technical leitmotifs include Markov chains and spectral graph theory (the necessary basics will be reviewed in class).

This course is not organized by certain methodologies or the various types of analysis. Instead, it is centered on projects inspired collaborators and their data sets. These collaborators will share their data and their ideas. They will point the way, or you will find your way, towards interesting questions. These will be the class projects.

As a graduate seminar, this course will seek to identify the limit of current knowledge and point towards the unexplored frontiers. A necessary part of this frontier is that collaborators will discuss work that is being formulated. You are encouraged to ask questions and provide feedback. Of course, this is no different from any other discussion pre-publication; if you provide valuable feedback, this does not make you an author. By participating in class discussions, you implicitly agree to this.

Because network analysis is a relatively unexplored area in Statistics (e.g. compared to linear regression), the content of the course will be accessible to first year graduate students and motivated undergraduates. While there are no formal prerequisites, you should have previously had classes which covered (1) eigenvectors, in some fashion and (2) expectation and variance, defined with pdf's and integrals. Undergraduate courses in linear algebra and probability are sufficient.

Objectives: To ask, discuss, and answer substantively interesting questions with network data from various disciplines. Along the way, we will learn the basic language of network analysis (graph theory notation). We will identify various types of questions that can be answered using network data. Students will become comfortable analyzing network data in R.

Topics: Physician referral networks, education labor markets, and perhaps others.

Graph definitions, empirical regularities (network pseudo-science), igraph, models of networks, random walks on graphs / Cheegar's bound / spectral clustering. Potential topics: contextualized network analysis, local algorithms, network driven sampling, others.

Computing: We will make extensive use of R. It is assumed that you are comfortable coding in R. For example, you should have comfort writing "if" statements and "for" loops. You should have previously struggled to import data into R. You might find these resources helpful `http://cran.r-project.org/doc/manuals/R-intro.pdf` and `http://adv-r.had.co.nz`.

Text: Not required. ED Kolaczyk, G Csrdi. *Statistical Analysis of Network Data with R*, Springer, 2014.

Exams: There are no exams.

Homeworks: There is no homework. However, projects may have multiple sequential assignments.

Projects: There will be one to three projects in the course. These will be discussed in more detail later in the course. We will decide as a class on the number and whether they are individual or group projects.

Academic Honesty: You are encouraged to talk with others about your project. However, you may not present other people's work as your own.

Disability accommodation:

Any student with a documented disability (e.g., physical, learning, psychiatric, vision, hearing, etc.) who needs to arrange reasonable accommodations must contact the instructor and McBurney Disability Resource Center at the beginning of the semester (i.e. within the first two weeks). The instructor needs to keep a copy of the documented disability.