INFO 485 / SOCI 485
Computational Methods for Complex Networks

Cornell University, Spring 2008
206 Hollister Hall
Tuesday & Thursday, 10:10–11:25

Instructor: Gueorgi Kossinets, 368 Uris Hall, gk67@cornell.edu

Teaching Assistant: Douglas Forster, dpf8@cornell.edu

Office hours: TBA

Description: This is a survey of research methods and techniques used in the study of complex networks. Interdisciplinary in nature, the course is intended for anyone interested in computational techniques of network analysis e.g. students in Information Science, Communication, Sociology, Economics, and Bioinformatics. You may think of it as a practicum sequel for ECON204/ SOC209/ CS285/ INFO 204 and preparation for CS 685 and similar graduate courses. The course covers a range of modern methods and explores a variety of real-life datasets, from social to technological to biological networks. Appropriate tools will be learned (making use of the Python NetworkX toolkit and the R statistical environment).

Prerequisites: (1) ECON 204 / SOC 209 / CS 285 / INFO 204 or similar level of familiarity with network concepts and methods; (2) an introductory course in computer programming such as CS 211 or INFO 230 or equivalent. If you are unsure about your preparation, please contact the instructor.

Homework: There will be several homework assignments and a final group project. The objective is to prepare the students for writing research papers for conferences or technical reports. The exercises will correspond to the main topics of the course (data preparation and exploratory analysis; modeling network structure; simulating dynamic processes on networks), so that the students have the necessary skills and tools to tackle a question of interest to them for their final project. Final reports should be 5-10 pages long in the ACM format (see http://www.acm.org/sigs/publications/proceedings-templates).

Lectures will mostly follow: Ulrik Brandes and Thomas Erlebach (Eds). Network Analysis: Methodological Foundations. Springer, 2005. The book is accessible online at http://www.springerlink.com/content/nv20c2jfpf28/

The following textbooks may serve as useful references:
- Jon Kleinberg and Eva Tardos. *Algorithm Design.* Addison-Wesley, 2005.
- S. Bornholdt and H. G. Schuster (Eds). *Handbook of Graphs and Networks.* Wiley, 2003.
- J. Zelle. *Python Programming: An Introduction to Computer Science.* Franklin Beedle, 2003.

Tentative course outline
(depending on the students' background and interests, we may substitute or add some topics)

I. Network concepts and data
- basic graph-theoretical concepts (nodes, edges, paths, components)
- survey of datasets (social, economic, transportation, computer, biological networks)
- network types and representations (directed, weighted, bipartite)
- implications of missing data

II. Exploratory analysis
- network statistics and summary measures
- visualizing networks
- structural similarity; assortative mixing; homophily
- understanding community structure
- enumerating network motifs
- measuring and describing networks over time

III. Modeling network structure
- random graphs; networks with arbitrary degree distribution
- small world networks; scale-free networks
- network growth; densification; preferential attachment
- network identification (which generative model fits better)

IV. Measuring and simulating dynamic processes on networks
- information flows; social influence and diffusion
- disease epidemics and intervention strategies
- resource location and strategic search
- signaling in metabolic and neuronal networks
- games on networks; cooperation and coordination; selfish routers
- network segregation; opinion dynamics
- predicting with networks; regression and classification using network data