

# G8325: Topics in Advanced Statistics Models and Methods for Relational Data

Department of Statistics, Columbia University

Fall 2010

Tentative as July 8, 2010

## Basic Course Information

- **Lecture Time:** Tuesdays 2:30-4:30 PM
- **Lecture Room:** TBD
- **Instructor:** Tian Zheng (Office hours: Mondays 1-3PM, Room 1007 SSW)
- **Email:** tzheng@stat.columbia.edu
- **Course website:** <http://courseworks.columbia.edu>

## Description

In this course we will explore current statistical models and methods for relational data, especially on network analysis. we will cover major models for relational data/networks, collection and estimation/inference of network data, visualization, major applications such as in social network analysis and biology.

## Textbook

No textbook required. We will use lectures notes and relevant journal articles, which will be posted on the class website.

## Reference texts

Topics covered in this course may be found in one of the following references. (This list will be updated as we progress in the course.)

1. Kolaczyk, E. D. (2009) Statistical analysis of network data, methods and models. Springer.
2. Newman, M., Barabasi, A.-L., Watts, D. J. (eds) (2006) The structure and dynamics of networks. Princeton University Press.

## Assignments and Grading

Each officially registered student is required to give an in-class presentation and submit a final report based on a reading/evaluation project. It can be a literature review on a collection of related papers, *or* a data analysis project based on one or several papers (or their extensions). The presentations will be scheduled in the second half of the semester. The officially registered students are encouraged to start discussing with the instructor about possible project ideas as early as possible (from week one, that is).

Grading is based on this *reading/evaluation project* and in-class participation. Students can choose any statistical computational package. Lecture examples will be given using R.